US 20250262297A1

(19) United States
(12) Patent Application Publication
Matsunaga et al.
(10) Pub. No.: US 2025/0262297 A1
(43) Pub. Date: Aug. 21, 2025

(54) **ALLERGY ANTIGEN AND EPITOPE THEREOF**

(71) Applicant: **HOYU CO., LTD.**, Aichi (JP)

(72) Inventors: **Kayoko Matsunaga**, Aichi (JP); **Akiko Yagami**, Aichi (JP); **Naoshi Shimojo**, Aichi (JP); **Fumiaki Ohno**, Aichi (JP)

(21) Appl. No.: **19/036,628**

(22) Filed: **Jan. 24, 2025**

Related U.S. Application Data

(62) Division of application No. 16/490,092, filed on Aug. 30, 2019, filed as application No. PCT/JP2017/032005 on Sep. 5, 2017.

(30) **Foreign Application Priority Data**

Mar. 3, 2017 (WO) PCT/JP2017/008593

Publication Classification

(51) **Int. Cl.**

A61K 39/35	(2006.01)
A23L 11/30	(2016.01)
A61K 38/00	(2006.01)
A61K 49/00	(2006.01)
A61P 37/08	(2006.01)
C07K 4/10	(2006.01)
C07K 14/415	(2006.01)
C07K 14/47	(2006.01)
G01N 33/53	(2006.01)

(52) **U.S. Cl.**

CPC ***A61K 39/35*** (2013.01); ***A23L 11/30*** (2016.08); ***A61K 49/00*** (2013.01); ***A61K 49/0006*** (2013.01); ***A61P 37/08*** (2018.01); ***C07K 4/10*** (2013.01); ***C07K 14/415*** (2013.01); ***C07K 14/47*** (2013.01); ***G01N 33/53*** (2013.01); *A61K 38/00* (2013.01); *G01N 2800/24* (2013.01)

(57) **ABSTRACT**

The present invention provides novel antigens of an allergy to soybean, methods and kits for diagnosing an allergy to soybean, pharmaceutical compositions comprising such an antigen, soybeans or processed products of soybean in which such an antigen is eliminated, and a tester for determining the presence or absence of a soybean antigen in an object of interest. The present invention also relates to polypeptides comprising an epitope of an allergen, kits, compositions and methods for diagnosing an allergy, comprising such a polypeptide, pharmaceutical compositions comprising such a polypeptide, and food raw materials or edible processed products in which an antigen comprising such a polypeptide is eliminated or reduced, or the polypeptide is cleaved or removed. The present invention further relates to a tester for determining the presence or absence of an antigen in an object of interest.

Specification includes a Sequence Listing.

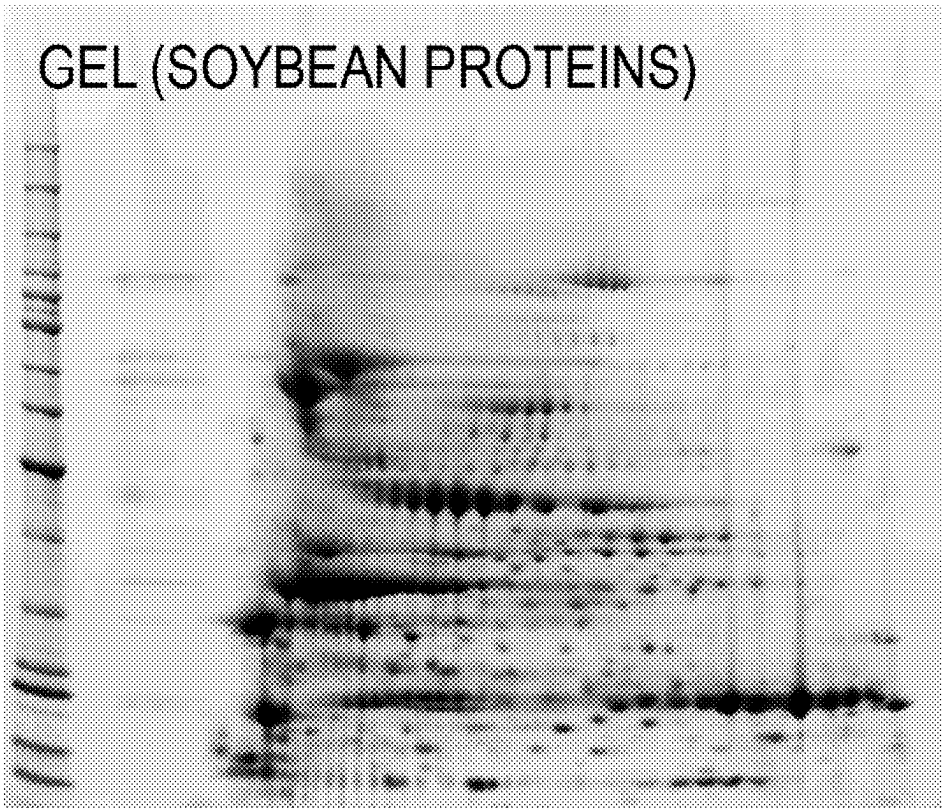


Fig. 1

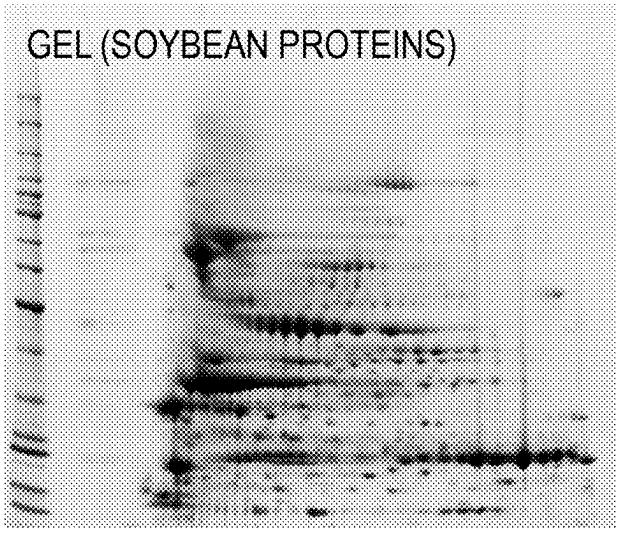


Fig. 2

HEALTHY SUBJECT

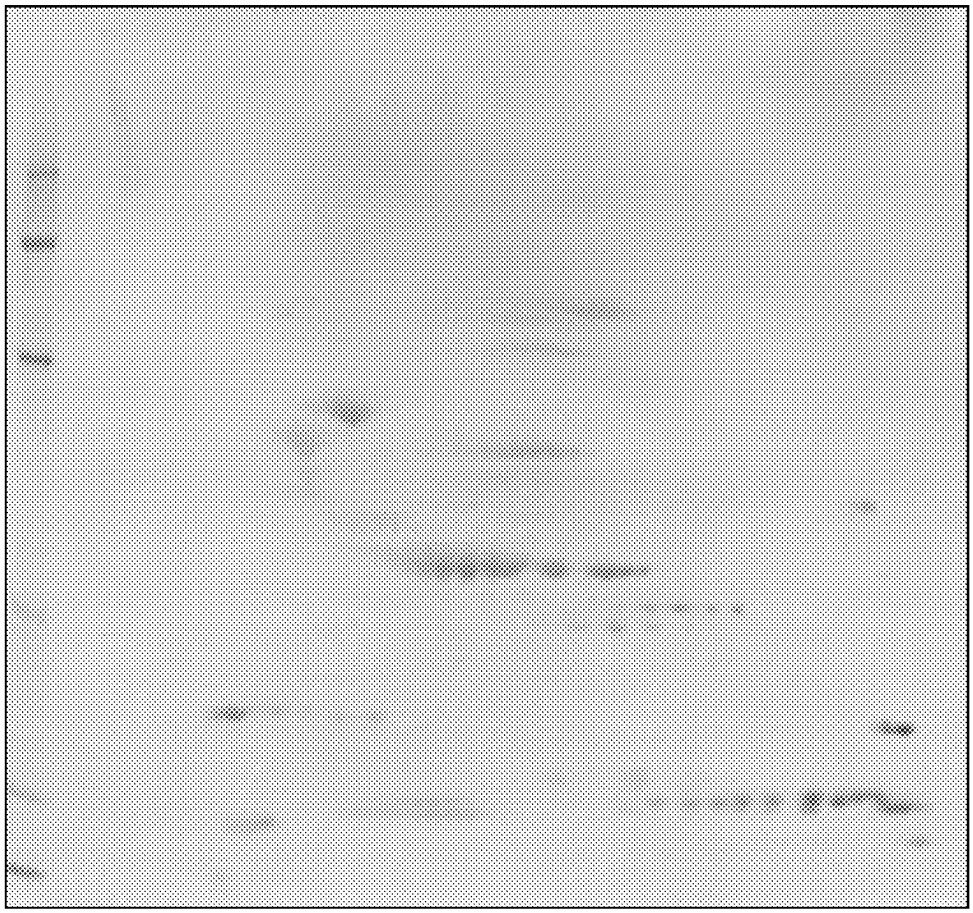


Fig. 3A

PATIENT
1α

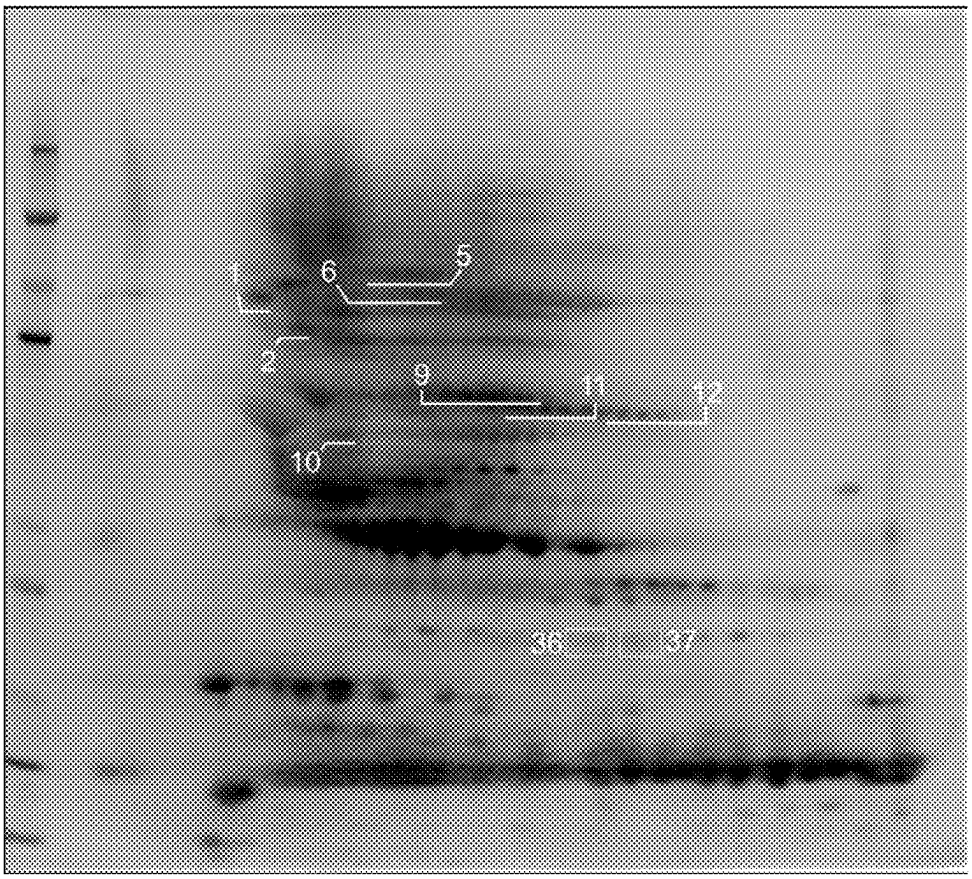


Fig. 3B

PATIENT
2α

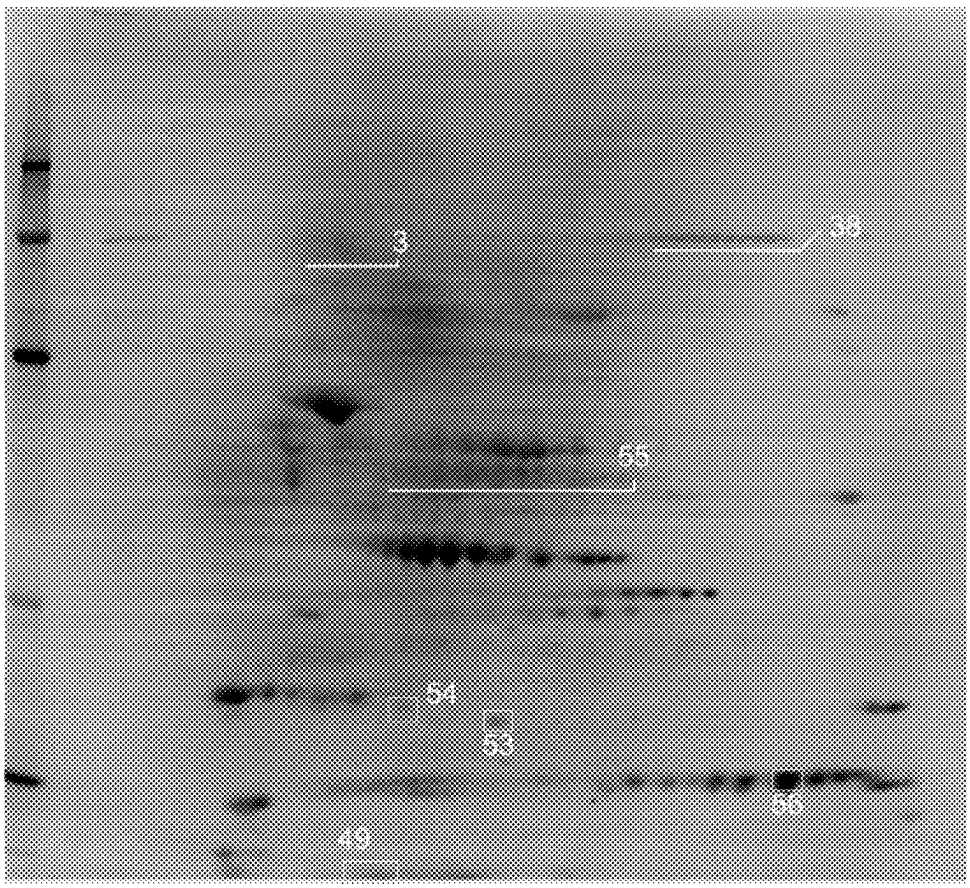


Fig. 3C

PATIENT
3α

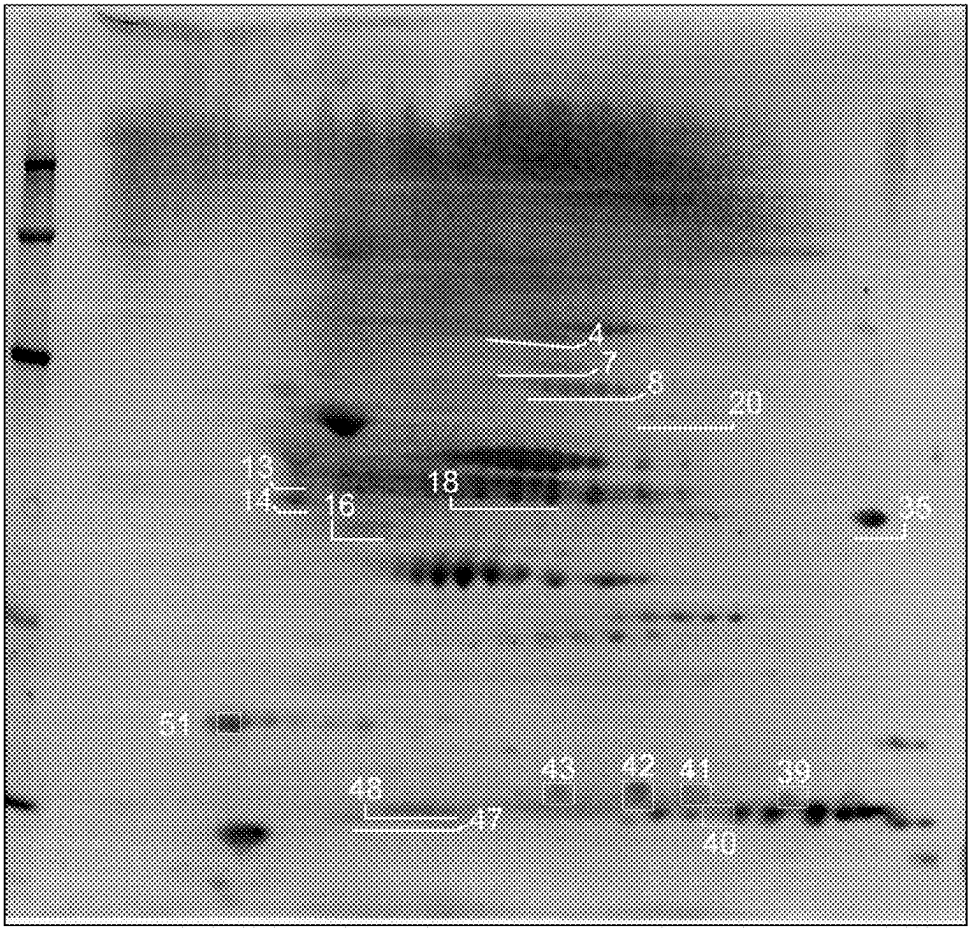


Fig. 3D

PATIENT
4α

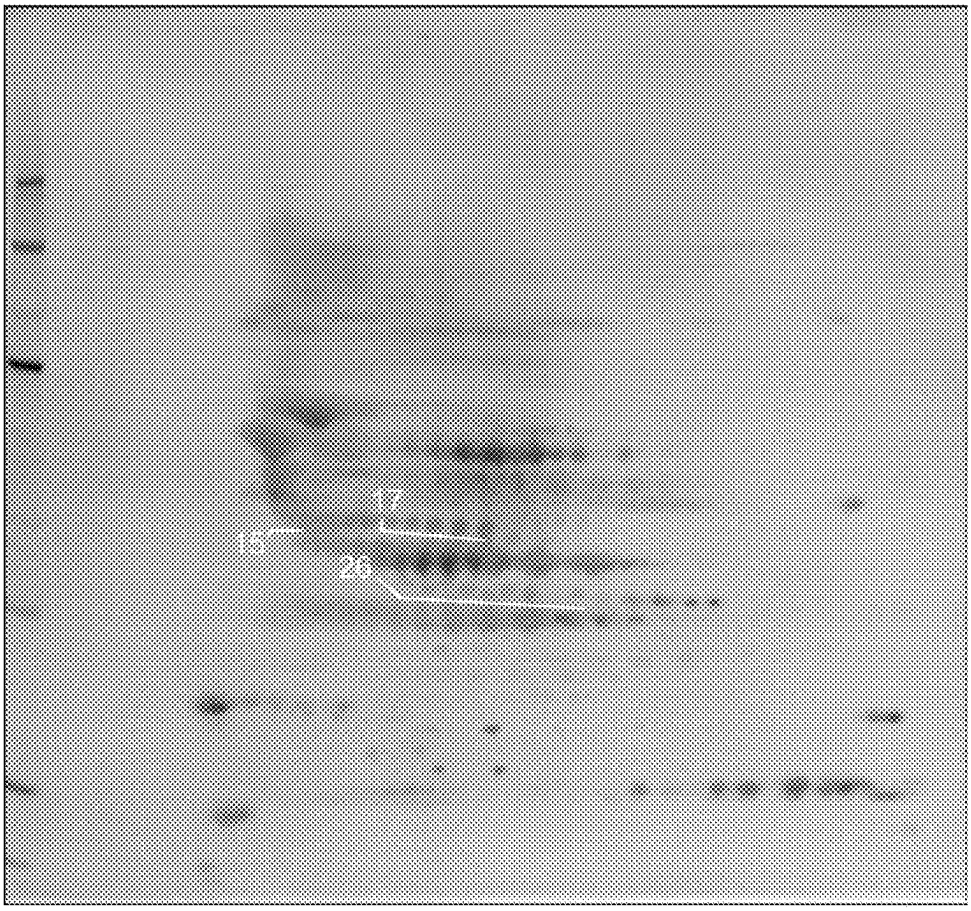


Fig. 3E

PATIENT
5α

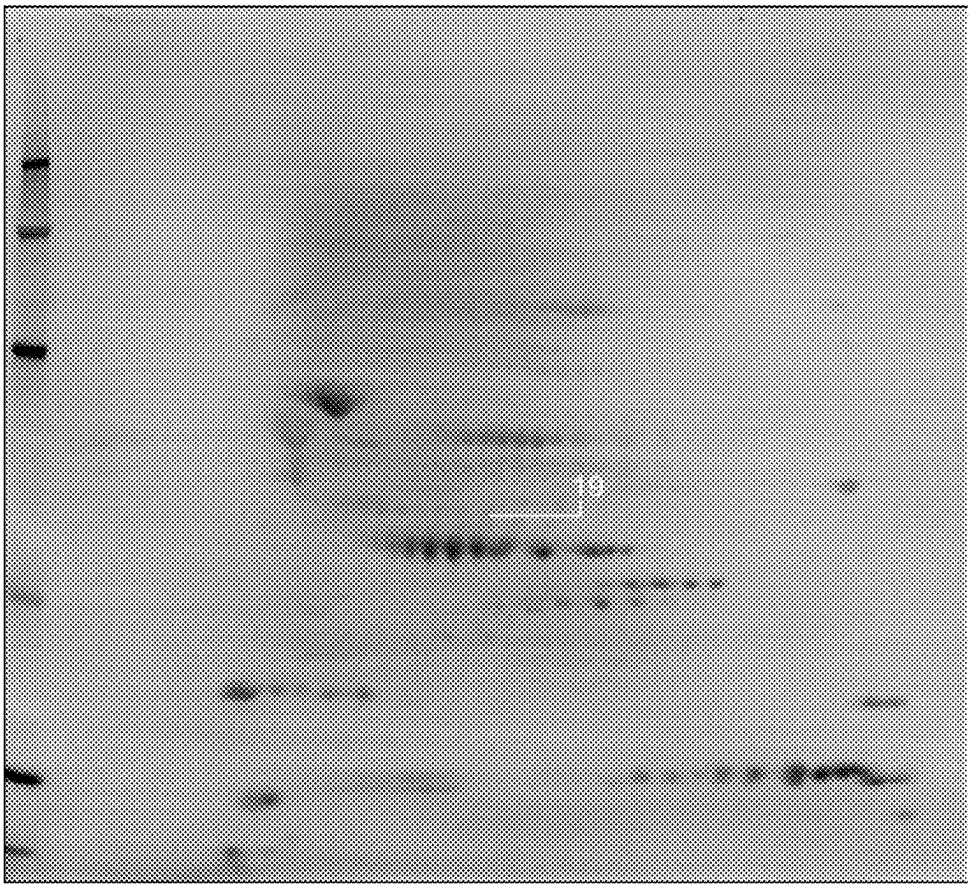


Fig. 3F

PATIENT
6α

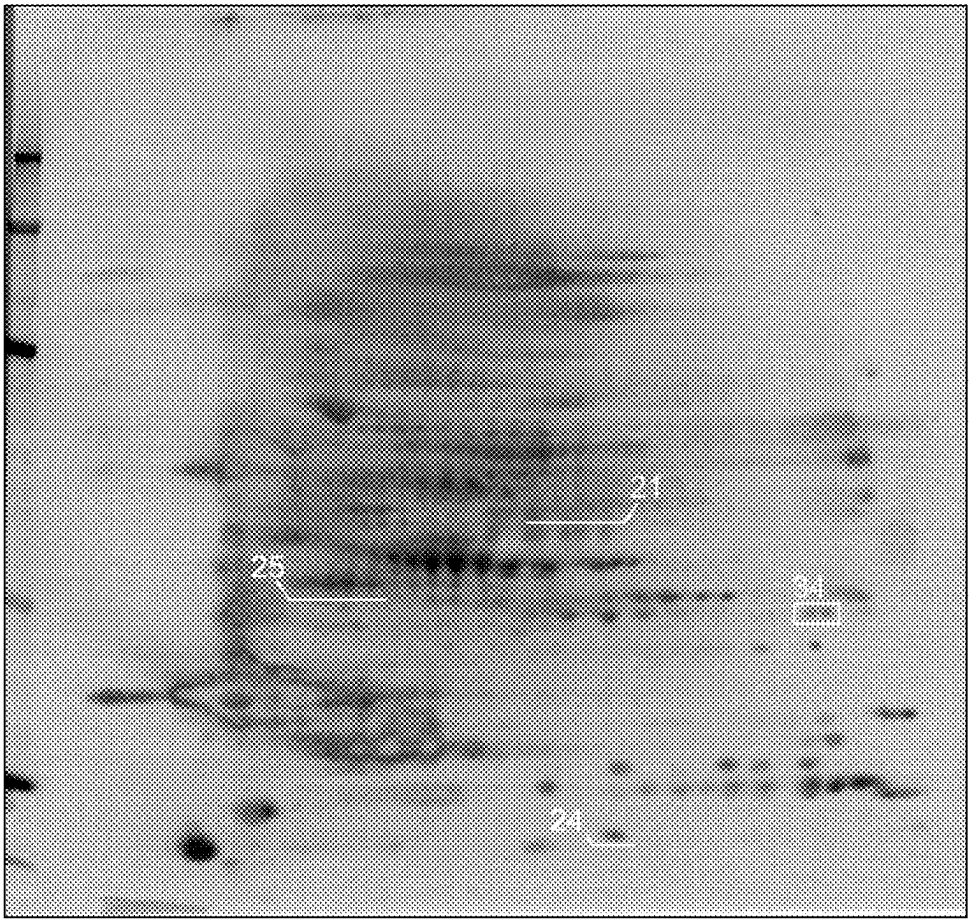


Fig. 3G

PATIENT
7α

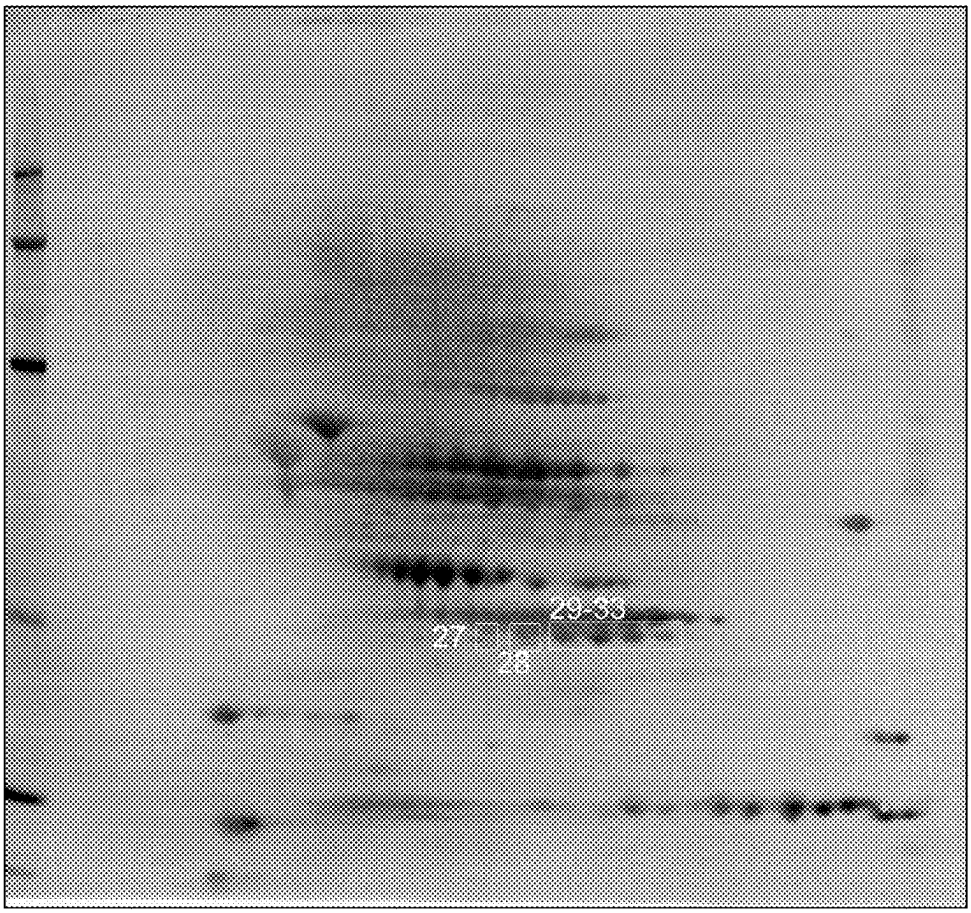


Fig. 3H

PATIENT
8α

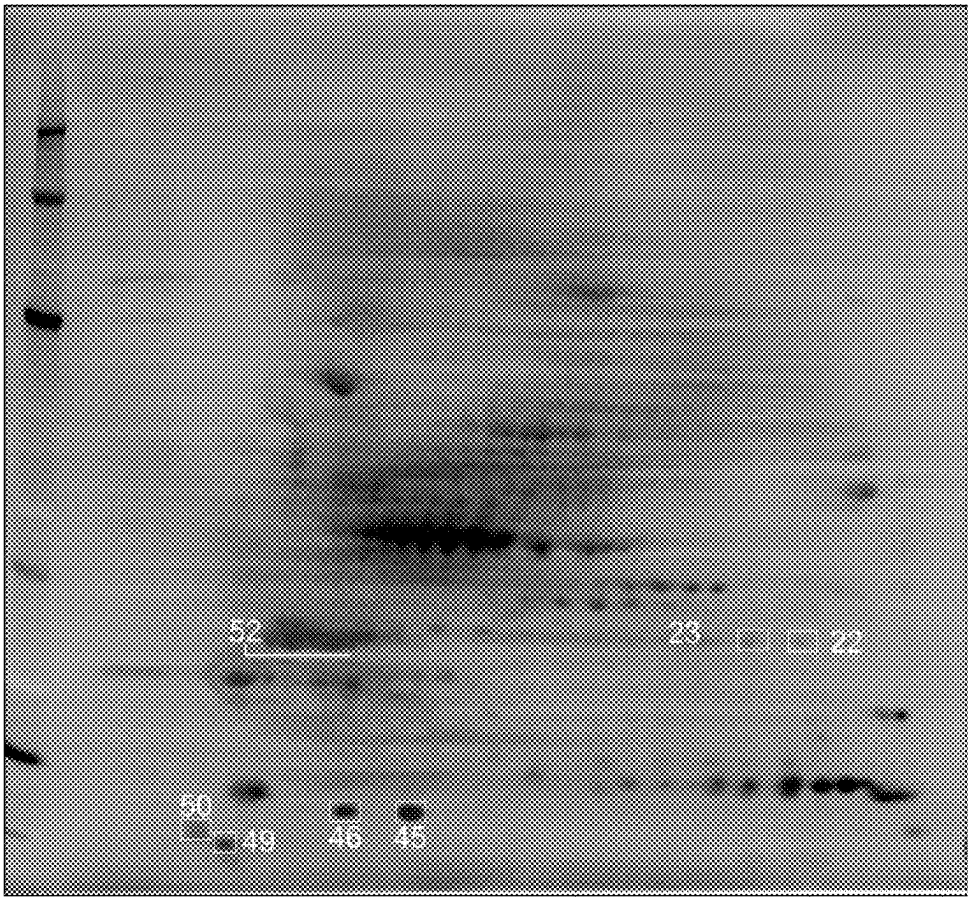


Fig. 3I

PATIENT
9α

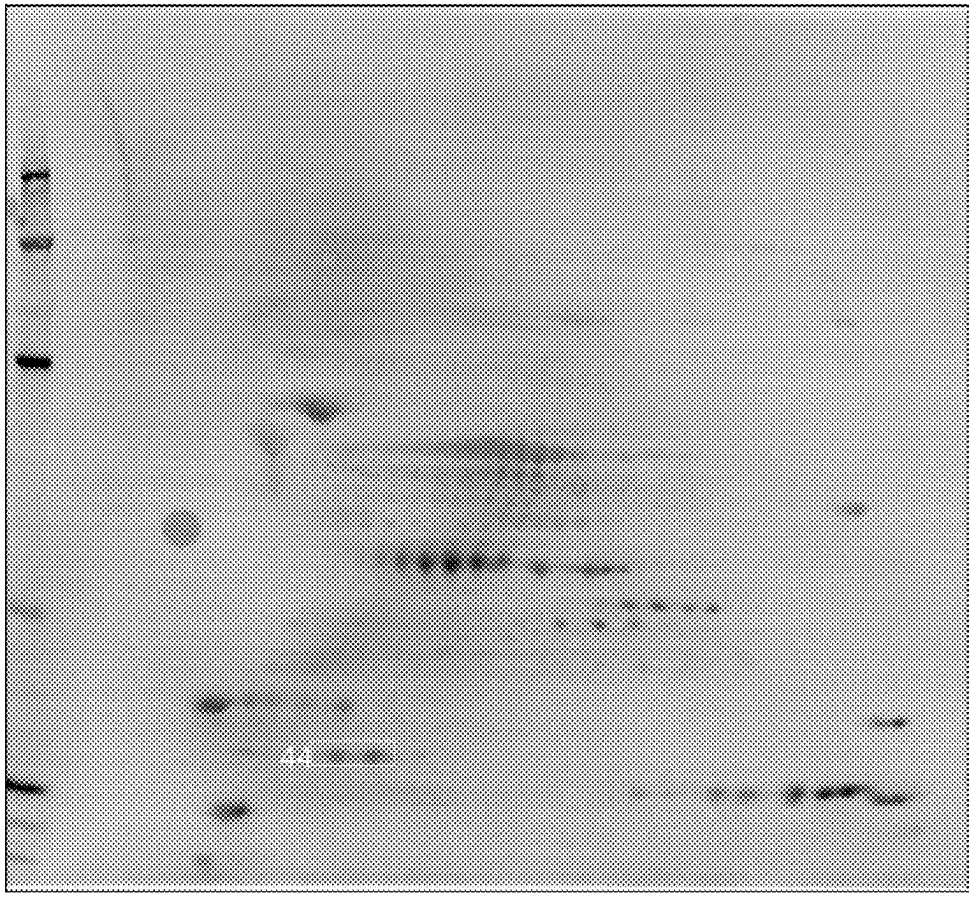


Fig. 4

SSPRGKVEAEKARPMEETKV (SEQ ID NO. 964)

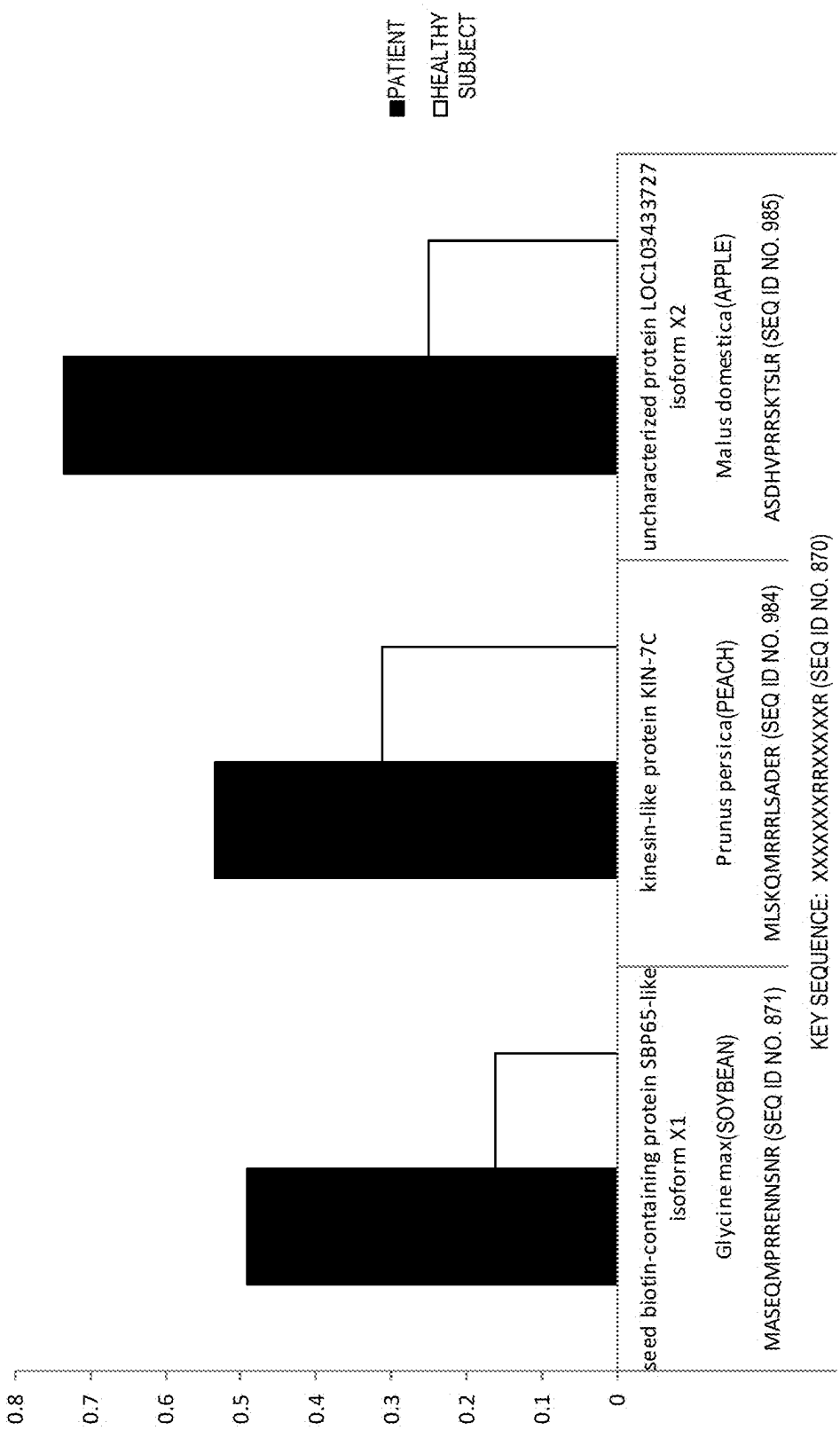
SSPRGKVEAE (SEQ ID NO. 965)
SPRGKVEAEK (SEQ ID NO. 966)
PRGKVEAEKA (SEQ ID NO. 967)
RGKVEAEKAR (SEQ ID NO. 968)
GKVEAEKARP (SEQ ID NO. 969)
KVEAEKARPM (SEQ ID NO. 970)
VEAEKARPME (SEQ ID NO. 971)
EAEKARPMEE (SEQ ID NO. 972)
AEKARPMEET (SEQ ID NO. 973)
EKARPMEETK (SEQ ID NO. 974)
KARPMEETKV (SEQ ID NO. 975)

Fig. 5

RGKVEAE (SEQ ID NO. 864) (RGKXXXX (SEQ ID NO. 862))

ARGKVEAE (SEQ ID NO. 976)
PAGKVEAE (SEQ ID NO. 977)
PRAKVEAE (SEQ ID NO. 978)
PRGAVEAE (SEQ ID NO. 979)
PRGKAEAE (SEQ ID NO. 980)
PRGKVAAE (SEQ ID NO. 981)
PRGKVEAE (SEQ ID NO. 982)
PRGKVEAA (SEQ ID NO. 983)

Fig. 6


PATIENT
HEALTHY SUBJECT
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
seed biotin-containing protein SBP65-like isoform X1
Glycine max(SOYBEAN)
MASEQMPRRENNSNR (SEQ ID NO. 871)
kinesin-like protein KIN-7C
Prunus persica(PEACH)
MLSKQMRRRLSADER (SEQ ID NO. 984)
uncharacterized protein LOC103433727 isoform X2
Malus domestica(APPLE)
ASDHVPRRSKTSLR (SEQ ID NO. 985)
KEY SEQUENCE: XXXXXXXRRXXXXXR (SEQ ID NO. 870)

ALLERGY ANTIGEN AND EPITOPE THEREOF

RELATED APPLICATIONS

[0001] This application is a divisional application of U.S. patent application Ser. No. 16/490,092, filed on Aug. 30, 2019, which is a U.S. National Stage Application of International Patent Application No. PCT/JP2017/032005, filed on Sep. 5, 2017, which claims priority to International Patent Application No. PCT/JP2017/008593, filed on Mar. 3, 2017. The contents of each of the aforementioned applications, including sequence listings, are incorporated herein by reference in their entirety.

REFERENCE TO SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing which has been submitted electronically in XML format and is hereby incorporated by reference in its entirety. Said XML copy, created on May 5, 2025, is named "129249_00603_SL.xml" and is 1,159,840 bytes in size.

TECHNICAL FIELD

[0003] The present invention relates to a novel antigen of an allergy to soybean. The present invention also relates to a kit, a composition, and a method for diagnosing allergy to soybean. The present invention also relates to a pharmaceutical composition comprising such an antigen and soybean or processed products of soybean in which such an antigen is eliminated or reduced. The present invention further relates to a tester for determining the presence or absence of a soybean antigen in an object of interest.

[0004] The present invention also relates to a polypeptide comprising an epitope of an allergen. The present invention also relates to a kit, a composition and a method for diagnosing an allergy, comprising such a polypeptide. The present invention also relates to a pharmaceutical composition comprising such a polypeptide, and a food raw material or an edible processed product in which an antigen comprising such a polypeptide is eliminated or reduced, or such a polypeptide is cleaved or removed. The present invention further relates to a method for producing an edible processed product in which such an antigen is eliminated or reduced. The present invention further relates to a tester for determining the presence or absence of an antigen comprising such a polypeptide in an object of interest.

BACKGROUND ART

[0005] In blood and tissues of allergic patients, IgE antibodies specific to particular antigens are produced. Physiological consequences caused by interaction between such IgE antibodies and such particular antigens elicit allergic reactions.

[0006] In the process of production of conventional allergy testing agents, antigen reagents are commonly prepared simply by grinding a candidate allergenic food, material or the like (Patent Literature 1). As seen above, conventional antigen reagents do not necessarily contain only a particular antigenic protein inducing an allergic reaction (allergen component), and rather contain different types of protein components. Thus, conventional antigen reagents contain varied amounts of allergen components. For this reason, the only case where conventional allergy tests have permitted detection of a positive allergic reaction is when in a conventional antigen reagent containing many types of proteins, a particular protein acting as an allergen component is present in an amount exceeding a threshold that allows determination of a positive reaction for binding to an IgE antibody. However, no determination of a positive reaction was possible and diagnosis efficiency was not sufficiently high when using a conventional allergy testing agent in patients possessing an IgE antibody binding to an allergen component present in small amounts in an allergen such as food.

[0007] The severity and symptoms of an allergic reaction do not necessarily correlate with the content of an allergen component. Even when a patient's IgE antibody reacts with an allergen component present in trace amounts in a candidate allergic food and material, the allergic reaction may develop allergic symptoms or may affect the severity of those symptoms.

[0008] An attempt to increase the diagnostic efficiency is being made by examining IgE antibodies to components composing a crude antigen to distinguish sensitization that directly contributes to a diagnosis from sensitization based on cross-antigenicity by a pan-allergen or the like. 8 soybean allergens are currently registered to the World Health Organization/International Union of Immunological Societies (WHO/IUIS).

[0009] However, while it is necessary to exhaustively identify allergen components in candidate allergic foods and materials in order to enhance the reliability of allergy tests, the patient detection rate by the measurement of such allergenic components is far from sufficient. Identification of novel allergens in soybean is very important not only for increasing the precision of diagnosis, but also for determining targets of therapeutic agents and low allergenic foods.

[0010] Meanwhile, in the field of protein separation and purification, various efforts have conventionally been made to develop methods for separating and purifying a protein or nucleic acid of interest from cell extracts or the like. Such methods may well be exemplified by dialysis based on salt concentration, and centrifugal separation.

[0011] Other efforts have been made to develop many purification methods based on electric charges of protein or nucleic acid residues or on the difference in molecular weight. Electric charge-based purification methods can be exemplified by column chromatography using ion exchange resins, and isoelectric focusing. Purifications based on molecular weight difference can be exemplified by centrifugal separation, molecular-sieve column chromatography, and SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis).

[0012] In recent years, a method for separating and purifying many different proteins from a small amount of sample has been used, which is more specifically a two-dimensional electrophoresis consisting of isoelectric focusing in the first dimension, followed by SDS-PAGE in the second dimension. The present applicant has conventionally developed some 2D electrophoresis methods with high separation ability (Patent Literature 2-5).

[0013] Allergen-specific IgE antibodies recognize and bind to epitopes that are particular amino acid sequences in allergen components. However, only a slight number of analyses have been made on epitopes as to the allergen components (Non Patent Literature 1), but such analyses are

still totally quite rare. Furthermore, any kit for diagnosing an allergy using a polypeptide comprising an epitope has not yet emerged in the market.

CITATION LIST

Patent Literature

[0014] PTL1: Japanese Patent Application Publication No. JP 2002-286716

[0015] PTL2: Japanese Patent Application Publication No. JP 2011-33544

[0016] PTL3: Japanese Patent Application Publication No. JP 2011-33546

[0017] PTL4: Japanese Patent Application Publication No. JP 2011-33547

[0018] PTL5: Japanese Patent Application Publication No. JP 2011-33548

Non Patent Literature

[0019] NPL 1: Matsuo, H., et al., J. Biol. Chem., (2004), Vol.279, No. 13, pp. 12135-12140

[0020] NPL 2: Zhao, L., et al., PLOS One. 2017 Aug. 11; 12 (8): c0182935. doi: 10.1371/journal.pone.0182935. eCollection 2017.

[0021] NPL 3: Cong, Y., et al., J Dairy Sci. 2013; 96 (11): 6870-6. doi: 10.3168/jds.2013-6880. Epub 2013 Sep. 12.

[0022] NPL 4: Beczhold, D. H., et al., J Allergy Clin Immunol. 2001 June; 107 (6): 1069-76.

SUMMARY OF INVENTION

Technical Problem

[0023] The present invention provides novel antigens of an allergy to soybean. The present invention also provides methods and kits for diagnosing allergy to soybean. The present invention also provides pharmaceutical compositions comprising such an antigen and soybean or processed products of soybean in which such an antigen is eliminated or reduced. The present invention further provides testers for determining the presence or absence of a soybean antigen in an object of interest.

[0024] The present invention also provides polypeptides comprising an epitope of an allergen. The present invention also provides kits, compositions and methods for diagnosing an allergy. The present invention also provides pharmaceutical compositions comprising such a polypeptide, and food raw materials or edible processed products in which an antigen comprising such a polypeptide is eliminated or reduced, or such a polypeptide is cleaved or removed. The present invention further relates to methods for producing an edible processed product in which such an antigen is eliminated or reduced. The present invention further relates to testers for determining the presence or absence of an antigen comprising such a polypeptide in an object of interest.

Solution to Problem

[0025] In order to solve the aforementioned problems, the present inventors had made intensive studies to identify causative antigens of an allergy. As a result, the inventors succeeded in identifying novel antigens to which an IgE antibody in the serum of an allergic patient specifically binds. The present invention has been completed based on this finding.

[0026] Thus, in one embodiment, the present invention can be as defined below.

[0027] [1] A kit for diagnosing an allergy to a soybean, the kit comprising, as an antigen, at least one of proteins defined below in any one of (1a) to (55a):

(1α) (1αA1) endoplasmin homolog isoform X2 or a variant thereof, which is defined below in any of (1αA1-a) to (1αA1-e):

[0028] (1αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 70;

[0029] (1αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 70;

[0030] (1αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 69;

[0031] (1αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 69; or

[0032] (1αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 69;

[0033] (1aB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 71-87,

[0034] (1αA2) endoplasmin homolog isoform X2 or a variant thereof, which is defined below in any of (1αA2-a) to (1αA2-e):

[0035] (1αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 89;

[0036] (1αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 89;

[0037] (1αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 88;

[0038] (1αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 88; or

[0039] (1αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 88; or

[0040] (1αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 90-104;

(2α) (2αA1) heat shock cognate protein 80 or a variant thereof, which is defined below in any of (2αA1-a) to (2αA1-e):

[0041] (2αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 106;

[0042] (2αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 106;

[0043] (2αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 105;

[0044] (2αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 105; or

[0045] (2αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 105; or

[0046] (2αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 107-120;

[0047] (2αA2) heat shock protein 90-1 or a variant thereof, which is defined below in any of (2αA2-a) to (2αA2-e):

[0048] (2αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 122;

[0049] (2A2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 122;

[0050] (2αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 121;

[0051] (2αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 121; or

[0052] (2αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 121; or

[0053] (2αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 123-137;

[0054] (2αA3) hypothetical protein GLYMA_02G302500 or a variant thereof, which is defined below in any of (2αA3-a) to (2αA3-e):

[0055] (2αA3-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 139;

[0056] (2αA3-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 139;

[0057] (2αA3-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 138;

[0058] (2αA3-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 138; or

[0059] (2αA3-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 138; or

[0060] (2αB3) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 140-153;

(4α) (4aA) embryo-specific urease isoform X1 or a variant thereof, which is defined below in any of (4αA-a) to (4αA-e):

[0061] (4αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 155;

[0062] (4αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 155;

[0063] (4αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 154;

[0064] (4αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 154; or

[0065] (4αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 154; or

[0066] (4αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 156 and 157;

(5α) (5αA) alpha-1,4 glucan phosphorylase L isozyme, chloroplastic/amyloplastic or a variant thereof, which is defined below in any of (5αA-a) to (5αA-e):

[0067] (5αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 159;

[0068] (5αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 159;

[0069] (5αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 158;

[0070] (5αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 158; or

[0071] (5αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 158; or

[0072] (5αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 160 and 161;

(6α) (6αA1) aconitate hydratase 1 or a variant thereof, which is defined below in any of (6αA1-a) to (6αA1-e):

[0073] (6αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 163;

[0074] (6αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 163;

[0075] (6αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with dele-

tion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 162;

[0076] (6αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 162; or

[0077] (6αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 162;

[0078] (6αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 164-167; or

[0079] (6αA2) aconitate hydratase 1 or a variant thereof, which is defined below in any of (6αA2-a) to (6αA2-e):

[0080] (6αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 169;

[0081] (6αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 169;

[0082] (6αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 168;

[0083] (6αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 168; or

[0084] (6αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 168; or

[0085] (6B2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 170-172;

(7α) (7αA) methionine synthase or a variant thereof, which is defined below in any of (7αA-a) to (7αA-e):

[0086] (7αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 174;

[0087] (7αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 174;

[0088] (7αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 173;

[0089] (7αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 173; or

[0090] (7αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 173; or

[0091] (7αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 175-187;

(8%) (8αA) transketolase, chloroplastic or a variant thereof, which is defined below in any of (8αA-a) to (8αA-e):

[0092] (8αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 189;

[0093] (8αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 189;

[0094] (8αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 188;

[0095] (8αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 188; or

[0096] (8αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 188; or

[0097] (8αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 190-199;

(9α) (9αA) lysine-tRNA ligase, cytoplasmic-like isoform X2 or a variant thereof, which is defined below in any of (9αA-a) to (9αA-e):

[0098] (9αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 201;

[0099] (9αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 201;

[0100] (9αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 200;

[0101] (9αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 200; or

[0102] (9αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 200; or

[0103] (9B) a protein comprising an amino acid sequence of SEQ ID NO: 202;

(10α) (10αA1) chaperonin CPN60-2, mitochondrial or a variant thereof, which is defined below in any of (10αA1-a) to (10αA1-e):

[0104] (10αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 204;

[0105] (10αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 204;

[0106] (10αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 203;

[0107] (10αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 203; or

[0108] (10αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 203;

[0109] (10αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 205-218;

[0110] (10αA2) chaperonin CPN60-2, mitochondrial-like isoform X1 or a variant thereof, which is defined below in any of (10αA2-a) to (10αA2-e):

[0111] (10αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 220;

[0112] (10αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 220;

[0113] (10αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 219;

[0114] (10αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 219; or

[0115] (10αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 219; or

[0116] (10αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 221-231;

(11α) (11αA) polyadenylate-binding protein 8 isoform X3 or a variant thereof, which is defined below in any of (11αA-a) to (11αA-e):

[0117] (11αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 233;

[0118] (11αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 233;

[0119] (11αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 232;

[0120] (11αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 232; or

[0121] (11αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 232; or

[0122] (11αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 234-237;

(12α) (12αA) pyrophosphate-fructose 6-phosphate 1-phosphotransferase subunit alpha-like or a variant thereof, which is defined below in any of (12αA-a) to (12αA-e):

[0123] (12αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 239;

[0124] (12αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 239;

[0125] (12αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 238;

[0126] (12αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 238; or

[0127] (12αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 238; or

[0128] (12αB) a protein comprising an amino acid sequence of SEQ ID NO: 240;

(13α) (13αA1) protein disulfide isomerase-like protein precursor or a variant thereof, which is defined below in any of (13αA1-a) to (13αA1-e):

[0129] (13αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 242;

[0130] (13αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 242;

[0131] (13αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 241;

[0132] (13αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 241; or

[0133] (13αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 241;

[0134] (13αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 243-252;

[0135] (13αA2) protein disulfide-isomerase or a variant thereof, which is defined below in any of (13αA2-a) to (13αA2-e):

[0136] (13αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 254;

[0137] (13αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 254;

[0138] (13αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 253;

[0139] (13αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 253; or

[0140] (13αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 253; or

[0141] (13αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 255-278;

(14α) (14αA1) V-type proton ATPase subunit B 2 or a variant thereof, which is defined below in any of (14αA1-a) to (14αA1-e):

[0142] (14αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 280;

[0143] (14αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 280;

[0144] (14αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 279;

[0145] (14αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 279; or

[0146] (14αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 279;

[0147] (14αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 281-284;

[0148] (14αA2) V-type proton ATPase subunit B 2 or a variant thereof, which is defined below in any of (14αA2-a) to (14αA2-e):

[0149] (14αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 286;

[0150] (14αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 286;

[0151] (14αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 285;

[0152] (14αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 285; or

[0153] (14αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 285; or

[0154] (14αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 287-289;

(15, 16) (15, 16αA1) UTP-glucose-1-phosphate uridylyltransferase or a variant thereof, which is defined below in any of (15, 16αA1-a) to (15, 16αA1-e):

[0155] (15, 16αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 291;

[0156] (15, 16αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 291;

[0157] (15, 16αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 290;

[0158] (15, 16αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 290; or

[0159] (15, 16αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 290;

[0160] (15, 16αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 292-304;

[0161] (15, 16αA2) hypothetical protein GLYMA_02G241100 or a variant thereof, which is defined below in any of (15, 16αA2-a) to (15, 16αA2-e):

[0162] (15, 16αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 306;

[0163] (15, 16αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 306;

[0164] (15, 16αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 305;

[0165] (15, 16αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 305; or

[0166] (15, 16αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 305; or

[0167] (15, 16αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 307-318;

(17α) (17αA1) guanosine nucleotide diphosphate dissociation inhibitor 2 or a variant thereof, which is defined below in any of (17αA1-a) to (17αA1-e):

[0168] (17αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 320;

[0169] (17αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 320;

[0170] (17αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with dele-

tion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 319;

[0171] (17αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 319; or

[0172] (17αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 319;

[0173] (17αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 321-325;

[0174] (17αA2) rab GDP dissociation inhibitor alpha-like or a variant thereof, which is defined below in any of (17αA2-a) to (17αA2-e):

[0175] (17αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 327;

[0176] (17αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 327;

[0177] (17αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 326;

[0178] (17αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 326; or

[0179] (17αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 326; or

[0180] (17αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 328-332;

(18α) (18αA) hypothetical protein GLYMA_10G028300 or a variant thereof, which is defined below in any of (18αA-a) to (18αA-e):

[0181] (18αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 334;

[0182] (18αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 334;

[0183] (18αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 333;

[0184] (18αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 333; or

[0185] (18αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 333; or

[0186] (18αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 335-339;

(19α) (19αA) sucrose binding protein homolog S-64 or a variant thereof, which is defined below in any of (19αA-a) to (19αA-e):

[0187] (19αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 341;

[0188] (19A-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 341;

[0189] (19αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 340;

[0190] (19αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 340; or

[0191] (19αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 340; or

[0192] (19αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 342-346;

(25α) (25αA1) succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial or a variant thereof, which is defined below in any of (25αA1-a) to (25αA1-e):

[0193] (25αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 348;

[0194] (25αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 348;

[0195] (25αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 347;

[0196] (25αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 347; or

[0197] (25αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 347;

[0198] (25αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 349-356;

[0199] (25αA2) succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial or a variant thereof, which is defined below in any of (25αA2-a) to (25αA2-e):

[0200] (25αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 358;

[0201] (25αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 358;

[0202] (25αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 357;

[0203] (25αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 357; or

[0204] (25αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 357; or

[0205] (25αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 359-365;

(26α) (26αA1) phosphoglycerate kinase, cytosolic or a variant thereof, which is defined below in any of (26αA1-a) to (26αA1-e):

[0206] (26αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 367;

[0207] (26αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 367;

[0208] (26αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 366;

[0209] (26αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 366; or

[0210] (26αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 366;

[0211] (26αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 368-382;

[0212] (26αA2) phosphoglycerate kinase, cytosolic or a variant thereof, which is defined below in any of (26αA2-a) to (26αA2-e):

[0213] (26αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 384;

[0214] (26αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 384;

[0215] (26αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 383;

[0216] (26αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 383; or

[0217] (26A2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 383;

[0218] (26αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 385-392;

[0219] (26αA3) phosphoglycerate kinase, cytosolic or a variant thereof, which is defined below in any of (26αA3-a) to (26αA3-e):

[0220] (26αA3-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 394;

[0221] (26αA3-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 394;

[0222] (26αA3-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 393;

[0223] (26αA3-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 393; or

[0224] (26αA3-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 393; or

[0225] (26αB3) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 395-409;

(27α) (27αA1) alcohol dehydrogenase 1 or a variant thereof, which is defined below in any of (27αA1-a) to (27αA1-e):

[0226] (27αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 411;

[0227] (27αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 411;

[0228] (27αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 410;

[0229] (27αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 410; or

[0230] (27αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 410;

[0231] (27αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 412-416;

[0232] (27αA2) alcohol dehydrogenase 1 or a variant thereof, which is defined below in any of (27αA2-a) to (27αA2-e):

[0233] (27αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 418;

[0234] (27αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 418;

[0235] (27αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 417;

[0236] (27αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 417; or

[0237] (27αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 417;

[0238] (27αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 419 and 420;

[0239] (27αA3) alcohol dehydrogenase 1-like or a variant thereof, which is defined below in any of (27αA3-a) to (27αA3-e):

[0240] (27αA3-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 422;

[0241] (27αA3-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 422;

[0242] (27αA3-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 421;

[0243] (27αA3-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 421; or

[0244] (27αA3-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 421; or

[0245] (27αB3) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 423-427;

(28α) (28αA) 35 kDa seed maturation protein isoform X1 or a variant thereof, which is defined below in any of (28αA-a) to (28αA-e):

[0246] (28αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 429;

[0247] (28αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 429;

[0248] (28αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 428;

[0249] (28αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 428; or

[0250] (28αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 428; or

[0251] (28αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 430-433;

(29-33) (29-33αA1) glyceraldehyde-3-phosphate dehydrogenase isoform X1 or a variant thereof, which is defined below in any of (29-33αA1-a) to (29-33αA1-e):

[0252] (29-33αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 435;

[0253] (29-33αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 435;

[0254] (29-33αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 434;

[0255] (29-33αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 434; or

[0256] (29-33αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 434;

[0257] (29-33αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 436-439;

[0258] (29-33αA2) glyceraldehyde-3-phosphate dehydrogenase, cytosolic or a variant thereof, which is defined below in any of (29-33αA2-a) to (29-33αA2-e):

[0259] (29-33αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 441;

[0260] (29-33αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 441;

[0261] (29-33αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 440;

[0262] (29-33αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 440; or

[0263] (29-33αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 440;

[0264] (29-33αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 442-445;

[0265] (29-33αA3) uncharacterized protein LOC100782924 or a variant thereof, which is defined below in any of (29-33αA3-a) to (29-33αA3-e):

[0266] (29-33αA3-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 447;

[0267] (29-33αA3-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 447;

[0268] (29-33αA3-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 446;

[0269] (29-33αA3-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 446; or

[0270] (29-33αA3-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 446; or

[0271] (29-33αB3) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 448-451,

(34α) (34αA1) bifunctional UDP-glucose 4-epimerase and UDP-xylose 4-epimerase 1-like or a variant thereof, which is defined below in any of (34αA1-a) to (34αA1-e):

[0272] (34αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 453;

[0273] (34αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 453;

[0274] (34αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 452;

[0275] (34αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 452; or

[0276] (34αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 452;

[0277] (34αB1) a protein comprising an amino acid sequence of SEQ ID NO: 454;

[0278] (34αA2) uncharacterized protein LOC100803483 or a variant thereof, which is defined below in any of (34αA2-a) to (34αA2-e):

[0279] (34αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 456;

[0280] (34αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 456;

[0281] (34αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 455;

[0282] (34αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 455; or

[0283] (34αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 455; or

[0284] (34αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 457 and 458;

(35α) (35αA1) elongation factor 1-alpha or a variant thereof, which is defined below in any of (35αA1-a) to (35αA1-e):

[0285] (35αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 460;

[0286] (35αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 460;

[0287] (35αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 459;

[0288] (35αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 459; or

[0289] (35αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 459;

[0290] (35αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 461-466;

[0291] (35αA2) elongation factor 1-alpha or a variant thereof, which is defined below in any of (35αA2-a) to (35αA2-e):

[0292] (35αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 468;

[0293] (35αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 468;

[0294] (35αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 467;

[0295] (35αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 467; or

[0296] (35αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 467;

[0297] (35αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 469-471;

[0298] (35αA3) elongation factor-1A isoform X1 or a variant thereof, which is defined below in any of (35αA3-a) to (35αA3-e):

[0299] (35αA3-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 473;

[0300] (35αA3-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 473;

[0301] (35αA3-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with dele-

tion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 472;

[0302] (35αA3-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 472; or

[0303] (35αA3-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 472; or

[0304] (35αB3) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 474-477;

(36, 37α) (36, 37αA1) seed maturation protein PM34 or a variant thereof, which is defined below in any of (36, 37αA1-a) to (36, 37αA1-e):

[0305] (36, 37αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 479;

[0306] (36, 37αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 479;

[0307] (36, 37αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 478;

[0308] (36, 37αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 478; or

[0309] (36, 37αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 478;

[0310] (36, 37αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 480-483;

[0311] (36, 37αA2) uncharacterized protein LOC100809384 or a variant thereof, which is defined below in any of (36, 37αA2-a) to (36, 37αA2-e):

[0312] (36, 37αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 485;

[0313] (36, 37αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 485;

[0314] (36, 37αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 484;

[0315] (36, 37αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 484; or

[0316] (36, 37αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 484;

[0317] (36, 37αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 486 and 487;

[0318] (36, 37αA3) hypothetical protein GLYMA_10G155300 or a variant thereof, which is defined below in any of (36, 37αA3-a) to (36, 37αA3-e):

[0319] (36, 37αA3-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 489;

[0320] (36, 37αA3-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 489;

[0321] (36, 37αA3-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 488;

[0322] (36, 37αA3-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 488; or

[0323] (36, 37αA3-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 488; or

[0324] (36, 37αB3) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 490-493;

(38α) (38αA) seed biotin-containing protein SBP65-like isoform X1 or a variant thereof, which is defined below in any of (38αA-a) to (38αA-e):

[0325] (38αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 495;

[0326] (38αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 495;

[0327] (38αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 494;

[0328] (38αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 494; or

[0329] (38αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 494; or

[0330] (38αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 496-513;

(39α) (39αA) ATP synthase subunit O, mitochondrial-like or a variant thereof, which is defined below in any of (39αA-a) to (39αA-e):

[0331] (39αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 515;

[0332] (39αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 515;

[0333] (39αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 514;

[0334] (39αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 514; or

[0335] (39αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 514; or

[0336] (39αB) a protein comprising an amino acid sequence of SEQ ID NO: 516;

(40, 56α) (40, 56αA) P24 oleosin isoform A or a variant thereof, which is defined below in any of (40, 56αA-a) to (40, 56αA-e):

[0337] (40, 56αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 518;

[0338] (40, 56αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 518;

[0339] (40, 56αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 517;

[0340] (40, 56αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 517; or

[0341] (40, 56αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 517; or

[0342] (40, 56αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 519 and 520;

(41α) (41αA) uncharacterized protein At3g49720-like isoform X2 or a variant thereof, which is defined below in any of (41αA-a) to (41αA-e):

[0343] (41αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 522;

[0344] (41αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 522;

[0345] (41αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 521;

[0346] (41αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 521; or

[0347] (41αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 521; or

[0348] (41αB) a protein comprising an amino acid sequence of SEQ ID NO: 523; (42, 43α) (42, 43αA1) superoxide dismutase [Mn], mitochondrial or a variant thereof, which is defined below in any of (42, 43αA1-a) to (42, 43αA1-e):

[0349] (42, 43αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 525;

[0350] (42, 43αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 525;

[0351] (42, 43αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 524;

[0352] (42, 43αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 524; or

[0353] (42, 43αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 524;

[0354] (42, 43αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 526-531;

[0355] (42, 43αA2) hypothetical protein GLYMA_04G221300 or a variant thereof, which is defined below in any of (42, 43αA2-a) to (42, 43αA2-e):

[0356] (42, 43αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 533;

[0357] (42, 43αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 533;

[0358] (42, 43αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 532;

[0359] (42, 43αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 532; or

[0360] (42, 43αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 532; or

[0361] (42, 43αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 534-539;

(45α) (45αA1) uncharacterized protein LOC100499771 or a variant thereof, which is defined below in any of (45αA1-a) to (45αA1-e):

[0362] (45αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 541;

[0363] (45αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 541;

[0364] (45αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with dele-

tion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 540;

[0365] (45αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 540; or

[0366] (45αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 540;

[0367] (45αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 542-545;

[0368] (45αA2) hypothetical protein GLYMA_09G192800 or a variant thereof, which is defined below in any of (45αA2-a) to (45αA2-e):

[0369] (45αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 547;

[0370] (45αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 547;

[0371] (45αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 546;

[0372] (45αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 546; or

[0373] (45αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 546; or

[0374] (45αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 548-551;

(46α) (46αA) seed maturation protein PM22 or a variant thereof, which is defined below in any of (46αA-a) to (46αA-e):

[0375] (46αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 553;

[0376] (46αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 553;

[0377] (46αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 552;

[0378] (46αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 552; or

[0379] (46αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 552; or

[0380] (46αB) a protein comprising an amino acid sequence of SEQ ID NO: 554;

(47α) (47αA1) eukaryotic translation initiation factor 5A-2 or a variant thereof, which is defined below in any of (47αA1-a) to (47αA1-e):

[0381] (47αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 556;

[0382] (47αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 556;

[0383] (47αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 555;

[0384] (47αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 555; or

[0385] (47αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 555;

[0386] (47αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 557 and 558;

[0387] (47αA2) uncharacterized protein LOC100305510 or a variant thereof, which is defined below in any of (47αA2-a) to (47αA2-e):

[0388] (47αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 560;

[0389] (47αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 560;

[0390] (47αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 559;

[0391] (47αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 559; or

[0392] (47αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 559; or

[0393] (47αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 561-563;

(48α) (48αA) uncharacterized protein LOC100306570 or a variant thereof, which is defined below in any of (48αA-a) to (48αA-e):

[0394] (48αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 565;

[0395] (48αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 565;

[0396] (48αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 564;

[0397] (48αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 564; or

[0398] (48αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 564; or

[0399] (48αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 566-570;

(50) (50αA) uncharacterized protein LOC100813859 or a variant thereof, which is defined below in any of (50αA-a) to (50αA-e):

[0400] (50αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 572;

[0401] (50αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 572;

[0402] (50αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 571;

[0403] (50αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 571; or

[0404] (50αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 571; or

[0405] (50αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 573-576;

(51α) (51αA) 14-3-3-like protein or a variant thereof, which is defined below in any of (51αA-a) to (51αA-e):

[0406] (51αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 578;

[0407] (51αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 578;

[0408] (51αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 577;

[0409] (51αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 577; or

[0410] (51αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 577; or

[0411] (51αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 579-582;

(52α) (52αA) probable fructokinase-4 or a variant thereof, which is defined below in any of (52αA-a) to (52αA-e):

[0412] (52αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 584;

[0413] (52αA-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 584;

[0414] (52αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 583;

[0415] (52αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 583; or

[0416] (52αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 583; or

[0417] (52αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 585-588;

(53α) (53αA1) triosephosphate isomerase isoform X1 or a variant thereof, which is defined below in any of (53αA1-a) to (53αA1-e):

[0418] (53αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 590;

[0419] (53αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 590;

[0420] (53αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 589;

[0421] (53αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 589; or

[0422] (53αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 589;

[0423] (53αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 591-597;

[0424] (53αA2) triosephosphate isomerase, cytosolic or a variant thereof, which is defined below in any of (53αA2-a) to (53αA2-e):

[0425] (53αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 599;

[0426] (53αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 599;

[0427] (53αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 598;

[0428] (53αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 598; or

[0429] (53αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 598; or

[0430] (53αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 600-604;

(54α) (54αA1) superoxide dismutase [Fe], chloroplastic isoform X1 or a variant thereof, which is defined below in any of (54αA1-a) to (54αA1-e):

[0431] (54αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 606;

[0432] (54αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 606;

[0433] (53αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 605;

[0434] (54αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 605; or

[0435] (54αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 605;

[0436] (54αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 607-611;

[0437] (54αA2) iron-superoxide dismutase or a variant thereof, which is defined below in any of (54αA2-a) to (54αA2-e):

[0438] (54αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 613;

[0439] (54αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 613;

[0440] (54αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 612;

[0441] (54αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 612; or

[0442] (54αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 612; or

[0443] (54αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 614-616; or

(55α) (55αA1) 51 kDa seed maturation protein precursor or a variant thereof, which is defined below in any of (55αA1-a) to (55αA1-e):

[0444] (55αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 618;

[0445] (55αA1-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 618;

[0446] (55αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 617;

[0447] (55αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 617; or

[0448] (55αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 617;

[0449] (55αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 619-635;

[0450] (55αA2) Lea protein precursor or a variant thereof, which is defined below in any of (55αA2-a) to (55αA2-e):

[0451] (55αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 637;

[0452] (55αA2-b) a protein comprising an amino acid sequence having at least 70% identity to the amino acid sequence of SEQ ID NO: 637;

[0453] (55αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 636;

[0454] (55αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70% identity to the nucleotide sequence of SEQ ID NO: 636; or

[0455] (55αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 636; or

[0456] (55αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 638-649.

[0457] [2] A composition for diagnosing an allergy to a soybean, comprising, as an antigen, at least one of proteins as defined above in any of (1α) to (55α) of [1].

[0458] [3] A method for providing an indicator for diagnosing an allergy to a soybean in a subject, the method comprising the steps of:

[0459] (i) contacting a sample obtained from the subject with an antigen, wherein the sample is a solution comprising an Ig antibody;

[0460] (ii) detecting binding between the IgE antibody present in the sample from the subject and the antigen; and

[0461] (iii) when the binding between the IgE antibody in the subject and the antigen is detected, an indicator of the fact that the subject is allergic to a soybean is provided;

[0462] wherein the antigen is at least one of proteins as defined above in any of (1α) to (55α) of [1].

[0463] [4] A pharmaceutical composition comprising at least one of proteins as defined above in any of (1α) to (55α) of [1].

[0464] [5] The pharmaceutical composition as set forth in [4], wherein the pharmaceutical composition is intended for the treatment of an allergy to a soybean.

[0465] [6] A soybean or processed products of soybean in which an antigen is eliminated or reduced, wherein the antigen is at least one of proteins as defined above in any of (1α) to (55α) of [1].

[0466] [7] A method for producing processed products of soybean in which an antigen is eliminated or reduced, the method comprising the step of confirming that the antigen is eliminated or reduced in a production process of the processed products, wherein the antigen is at least one of proteins as defined above in any of (1α) to (55α) of [1].

[0467] [8] A tester for determining the presence or absence of a soybean antigen in an object of interest, comprising an antibody that binds to at least one of proteins as defined above in any of (1α) to (55α) of [1].

[0468] [9] A soybean-derived antigen which is at least one of proteins as defined above in any of (1α) to (55α) of [1] and is causative of an allergy to soybean.

[0469] The present inventors also succeeded in finding epitopes as to the novel antigens.

[0470] As mentioned above, interaction between IgE antibodies from allergic patients and particular antigens elicit allergic reactions. These IgE antibodies recognize and bind to epitopes that are particular amino acid sequences in allergen components. Since the epitopes have a relatively short amino acid sequence, the IgE antibodies are capable of binding to different allergen components if the same amino acid sequence is present in the different allergen components. When different allergen components have a common epitope so that IgE antibodies from allergic patients bind to both of them, the allergens have cross-reactivity (or also called cross-antigenicity). Thus, the epitopes defined in the present invention enable diagnosis or treatment of an allergy including cross-reactivity, and detection of a plurality of allergen components comprising the epitopes, etc.

[0471] The present invention has been completed based on this finding. Thus, in another embodiment, the present invention can be as defined below.

[0472] [10] A polypeptide specifically binding to an IgE antibody from an allergic patient, the polypeptide being any one of the following:

[0473] (1β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 652-660 or the sequences XXKXEXXX and XTKXEXXX;

[0474] (2) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 662-666, 668-678 and 680-692 or the sequences DXXDKXXXX, XPDKX and XXXXDXXXXEXXXDX;

[0475] (4β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 718-731 or the sequence XXXXXXXKED;

[0476] (5β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 732-738;

[0477] (6) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 739 and 740;

[0478] (7β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 741-744;

[0479] (9) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 746-749 or the sequence XXEXXXXKDX;

[0480] (11β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 751-754;

[0481] (12) a polypeptide comprising the amino acid sequence of SEQ ID NO: 755;

[0482] (13β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 757-758 or the sequence XXGXXXXXXXXXXE;

[0483] (14β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 759-764;

[0484] (15β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NO: 766 or the sequence XGKXXXXXXXXXXXX;

[0485] (17β) a polypeptide comprising the amino acid sequence of SEQ ID NO: 767;

[0486] (18β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 750 and 768-775;

[0487] (19β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 776-778, 780-781, 783-784 and 786-787 or the sequences XXXXXXXXRXRRXXX, XXXXXXXXXXEEEXX and XXGKXXXXXXXXXXX;

[0488] (25β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 814 and 815;

[0489] (26β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 816 and 818-819 or the sequence FDKXXXX;

[0490] (27β) a polypeptide comprising the amino acid sequence of SEQ ID NO: 820;

[0491] (28) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 821-823 and 825-835 or the sequence XKXXXXEKXXXXXXX;

[0492] (29β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 837-839 or the sequence XXXXXXXDKP;

[0493] (34β) a polypeptide comprising at least one amino acid sequence selected from the group consisting

of SEQ ID NOs: 841-846 or the sequence XXXXXXXXXXXDDXE;

[0494] (35β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 847-853 and 855-856 or the sequence XXEXLDXXXX;

[0495] (36) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 857-861;

[0496] (38β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 864, 866-867, 869 and 871-892 or the sequences RGKXXXX, RGKXXXX, XRGKXXXXXX, RXKXXXXXXXXXXXE and XXXXXXXRRXXXXXR;

[0497] (39β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 893-895;

[0498] (40β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 896-899 and 901-902 or the sequence XXXXXVTAX;

[0499] (41β) a polypeptide comprising the amino acid sequence of SEQ ID NO: 903;

[0500] (42β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 904-906;

[0501] (45β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 910-912;

[0502] (46) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 914-915 or the sequence XDXDXXXDX;

[0503] (47β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 916-918;

[0504] (48) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 920-922 or the sequence XXXXXXFDK;

[0505] (50) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 924 and 927 or the sequences XXXXKQXXX and XXDXKQXXX;

[0506] (51) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 928-930;

[0507] (52β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 932-938 or the sequence XXXXXXSXXEXLXXX;

[0508] (53β) a polypeptide comprising the amino acid sequence of SEQ ID NO: 939;

[0509] (54β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 940 and 943 or the sequences KXXXXTX and KXXXXTQ; and

[0510] (55β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 944-949 and 951-952 or the sequence KDXXXXXXQ.

[0511] [11] A polypeptide specifically binding to an IgE antibody from an allergic patient, the polypeptide being any one of the following:

[0512] (3β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 694-695, 697 and 699-716 or the sequences XXXXXXXXXXRLXEX, XXXXXXXXXXXXXDE and XXXXGDR;

[0513] (20β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 788-792;

[0514] (21) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 793-802;

[0515] (22β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 803-808;

[0516] (23β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 809-812;

[0517] (24β) a polypeptide comprising the amino acid sequence of SEQ ID NO: 813;

[0518] (44β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 907-909;

[0519] (49) a polypeptide comprising the amino acid sequence of SEQ ID NO: 923;

[0520] (57) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 953-955;

[0521] (58β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 956-960;

[0522] (59β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 961 and 962; and

[0523] (60β) a polypeptide comprising the amino acid sequence of SEQ ID NO: 963.

[0524] [12] The polypeptide according to the above or [11], wherein the number of amino acid residues is 500 or less.

[0525] [13] A kit for diagnosing an allergy, comprising at least one of polypeptides according to any one of the above to [12].

[0526] [14] A composition for diagnosing an allergy, the composition comprising at least one of polypeptides according to any one of the above to as an antigen.

[0527] [15] A method for providing an indicator for diagnosing an allergy in a subject, the method comprising the steps of:

[0528] (i) contacting a sample obtained from the subject with an antigen, wherein the sample is a solution comprising an Ig antibody;

[0529] (ii) detecting binding between the IgE antibody present in the sample obtained from the subject and the antigen; and

[0530] (iii) when the binding between the IgE antibody in the subject and the antigen is detected, an indicator of the fact that the subject is allergic is provided;

[0531] wherein the antigen is at least one of polypeptides according to any one of the above to [12].

[0532] [16] A pharmaceutical composition comprising at least one of polypeptides according to any one of the above to [12].

[0533] [17] The pharmaceutical composition according to the above [16], wherein the pharmaceutical composition is intended for the treatment of an allergy.

[0534] [18] A tester for determining the presence or absence of an antigen in an object of interest, the tester comprising an antibody that binds to at least one of polypeptides according to any one of the above to [12].

[0535] [19] A food raw material or an edible processed product in which an antigen is eliminated or reduced, or a polypeptide is cleaved or removed, wherein the antigen is at least one of polypeptides according to any one of the above to [12].

[0536] [20] A method for producing an edible processed product in which an antigen is eliminated or reduced, or a polypeptide is cleaved or removed, the method comprising the step of confirming that the antigen is eliminated or reduced, or the polypeptide is cleaved or removed, in a production process of the processed product, wherein the antigen is at least one of polypeptides according to any one of the above to [12].

Advantageous Effects of Invention

[0537] The present invention can provide novel antigens of an allergy to soybean. Since the novel antigens (allergen components) that trigger a soybean allergy were identified according to this invention, this invention can provide highly sensitive methods and kits for diagnosing an allergy to soybean, pharmaceutical compositions comprising such an antigen, soybean or processed products of soybean in which such an antigen is eliminated or reduced.

[0538] The present invention can provide novel polypeptides comprising an epitope of an allergen. Use of the polypeptide of the present invention enables provision of highly sensitive methods and kits for diagnosing an allergy, comprising such a polypeptide, pharmaceutical compositions comprising such a polypeptide, food raw materials or edible processed products in which an antigen comprising such a polypeptide is eliminated or reduced, or such a polypeptide is cleaved or removed, and a method for producing the edible processed products.

BRIEF DESCRIPTION OF DRAWINGS

[0539] FIG. **1** is a photograph of a gel showing a protein electrophoretic pattern in two-dimensional electrophoresis of proteins contained in soybean. The bands at the left of the photograph are bands of molecular weight markers.

[0540] FIG. **2** is a photograph of an immunoblot of a two-dimensional electrophoretic pattern of proteins contained in soybean stained with serum of a healthy subject.

[0541] FIG. **3**A is a photograph of an immunoblot of a two-dimensional electrophoretic pattern of proteins contained in soybean stained with serum of Soybean-allergic patient 1α. The spots indicated with a white line or enclosed in a square are spot where the serum of the soybean-allergic patient specifically reacted. The numbers (n) indicated are the numbers of the spots and correspond to the spots indicated with "nα" herein.

[0542] FIG. **3**B is a photograph of an immunoblot of a two-dimensional electrophoretic pattern of proteins contained in soybean stained with serum of Soybean-allergic patient 2α. The spots indicated with a white line or enclosed in a square are spot where the serum of the soybean-allergic patient specifically reacted. The numbers (n) indicated are the numbers of the spots and correspond to the spots indicated with "nα" herein.

[0543] FIG. **3**C is a photograph of an immunoblot of a two-dimensional electrophoretic pattern of proteins contained in soybean stained with serum of Soybean-allergic patient 3α. The spots indicated with a white line or enclosed in a square are spot where the serum of the soybean-allergic patient specifically reacted. The numbers (n) indicated are the numbers of the spots and correspond to the spots indicated with "nα" herein.

[0544] FIG. **3**D is a photograph of an immunoblot of a two-dimensional electrophoretic pattern of proteins contained in soybean stained with serum of Soybean-allergic patient 4α. The spots indicated with a while line are spots where the serum of the soybean-allergic patient specifically reacted. The numbers (n) indicated are the numbers of the spots and correspond to the spots indicated with "nα" herein.

[0545] FIG. **3**E is a photograph of an immunoblot of a two-dimensional electrophoretic pattern of proteins contained in soybean stained with serum of Soybean-allergic patient 5α. The spots indicated with a while line are spots where the serum of the soybean-allergic patient specifically reacted. The numbers (n) indicated are the numbers of the spots and correspond to the spots indicated with "nα" herein.

[0546] FIG. **3**F is a photograph of an immunoblot of a two-dimensional electrophoretic pattern of proteins contained in soybean stained with serum of Soybean-allergic patient 6α. The spots indicated with a white line or enclosed in a square are spot where the serum of the soybean-allergic patient specifically reacted. The numbers (n) indicated are the numbers of the spots and correspond to the spots indicated with "nα" herein.

[0547] FIG. **3**G is a photograph of an immunoblot of a two-dimensional electrophoretic pattern of proteins contained in soybean stained with serum of Soybean-allergic patient 7α. The spots enclosed in a square are spots where the serum of the soybean-allergic patient specifically reacted. The numbers (n) indicated are the numbers of the spots and correspond to the spots indicated with "nα" herein.

[0548] FIG. **3**H is a photograph of an immunoblot of a two-dimensional electrophoretic pattern of proteins contained in soybean stained with serum of Soybean-allergic patient 8α. The spots indicated with a white line or enclosed in a square are spot where the serum of the soybean-allergic patient specifically reacted. The numbers (n) indicated are the numbers of the spots and correspond to the spots indicated with "nα" herein.

[0549] FIG. **31** is a photograph of an immunoblot of a two-dimensional electrophoretic pattern of proteins contained in soybean stained with serum of Soybean-allergic patient 9α. The spots indicated with a white line or enclosed in a square are spot where the serum of the soybean-allergic patient specifically reacted. The numbers (n) indicated are the numbers of the spots and correspond to the spots indicated with "nα" herein.

[0550] FIG. **4** is a diagram showing one example of results of determining peptides having the minimum number of amino acids functioning as an epitope by an overlapping technique.

[0551] FIG. **5** is a diagram showing one example of results of studying positions important for exertion of antigenicity as to the identified epitopes by an alanine scanning technique.

[0552] FIG. **6** is a graph showing results of studying cross-reactivity.

DESCRIPTION OF EMBODIMENTS

[0553] The present invention will be described in detail below, but the present invention is not limited to them.

[0554] Unless otherwise defined herein, all scientific and technical terms used in relation to the present invention shall have meanings commonly understood by those skilled in the art.

[0555] Amino acid sequences herein are represented by one-letter amino acid symbols well known to those skilled in the art and the left end corresponds to the amino terminal and the right end corresponds to the carboxy terminal.

[0556] As referred to herein, the "allergy" refers to the state in which, when a certain antigen enters the body of a living individual sensitized to said antigen, the living individual shows a hypersensitive reaction detrimental to him/her. In blood and tissues of individuals with many food-allergic diseases, IgE antibodies specific to antigens are produced. IgE antibodies bind to mast cells or basophils. When an antigen specific to such an IgE antibody enters again the body of a patient with an allergic disease, said antigen combines with the IgE antibody bound to mast cells or basophils, and the IgE antibody crosslinks said antigen on the cell surface, resulting in physiological effects of IgE antibody-antigen interaction. Examples of such physiological effects include release of histamine, serotonin, heparin, cosinophil chemotactic factors, leucotrienes, or the like. These released substances provoke an allergic reaction resulting from the combination of an IgE antibody with particular antigens. Such allergic reactions caused by particular antigens occur through the aforementioned pathway.

[0557] In the present invention, the allergy of interest is not particularly limited as long as it is an allergy to an allergen comprising an epitope to be used. Such an allergy may include allergies to plants of the family Malvaceae, the family Rubiaceae, the family Brassicaceae, the family Poaceae, the family Cucurbitaceae, the family Compositae, the family Asparagaceae, the family Juglandaceae, the family Rhamnaceae, the family Pedaliaceae, the family Umbelliferae, the family Euphorbiaceae, the family Solanaceae, the family Bromeliaceae, the family Caricaceae, the family Rosaceae, the family Amaranthaceae, the family Vitaceae, the family Leguminosae, the family Lythraceae and/or the family Betulaceae. Examples of the plant of the family Malvaceae include nalta jute (Jew's mallow, scientific name: Corchorus olitorius) and cacao (scientific name: *Theobroma cacao*). Examples of the plant of the family Rubiaceae include *Robusta* coffee (scientific name: *Coffea canephora*). Examples of the plant of the family Brassicaceae include Qing geng cai (scientific name: *Brassica rapa*), radish (scientific name: *Raphanus sativus*), and rapeseed (scientific name: *Brassica napus*). Examples of the plant of the family Poaceae include great millet (scientific name: *Sorghum bicolor*), corn (scientific name: *Zea mays*), and bread wheat (scientific name: *Triticum aestivum*). Examples of the plant of the family Cucurbitaceae include melon (scientific name: *Cucumis melo*), bitter melon (bitter gourd, scientific name: *Momordica charantia*), and cucumber (scientific name: *Cucumis sativus*). Examples of the plant of the family Compositae include sunflower (scientific name: *Helianthus annuus*). Examples of the plant of the family Asparagaceae include asparagus (scientific name: *Asparagus officinalis*). Examples of the plant of the family Juglandaceae include Persian walnut (scientific name: *Juglans regia*). Examples of the plant of the family Rhamnaceae include Chinese date (scientific name: Ziziphus jujuba). Examples of the plant of the family Pedaliaceae include sesame (scientific name: *Sesamum indicum*). Examples of the plant of the family Umbelliferae include carrot (scientific name: *Daucus carota* subsp. *Sativus*). Examples of the plant of the family Euphorbiaceae include cassava (scientific name: *Manihot esculenta*). Examples of the plant of the family Solanaceae include red pepper (scientific name: *Capsicum annuum*) and tomato (scientific name: *Solanum lycopersicum*). Examples of the plant of the family Bromeliaceae include pineapple (scientific name: *Ananas comosus*). Examples of the plant of the family Caricaceae include *papaya* (scientific name: *Carica papaya*). Examples of the plant of the family Rosaceae include apple (scientific name: *Malus domestica*), wild cherry (scientific name: *Prunus avium*), and peach (scientific name: *Prunus persica*). Examples of the plant of the family Amaranthaceae include quinua (scientific name: *Chenopodium quinoa*), spinach (scientific name: *Spinacia oleracea*), and beet (scientific name: *Beta vulgaris* subsp. vulgaris). Examples of the plant of the family Vitaceae include common grape (scientific name: *Vitis vinifera*). Examples of the plant of the family Leguminosae include pigeon pea (scientific name: Cajanus cajan), adzuki bean (scientific name: *Vigna angularis*), wild soybean (scientific name: *Glycine soja*), and soybean (scientific name: *Glycine max*). Examples of the plant of the family Lythraceae include pomegranate (scientific name: Punica granatum). Examples of the plant of the family Betulaceae include white birch (scientific name: *Betula platyphylla*). The allergy can produce an allergic reaction upon contact with an antigen or consumption of the antigen. Here, the contact refers to touch to an object and, particularly, as for the human body, refers to attachment to the skin, the mucosa (eyes, lips, etc.) or the like. The consumption refers to incorporation into the body and refers to incorporation by inhalation or through an oral route. In general, allergic reactions caused by consumption of foods are particularly referred to as food allergies. In a preferred mode, the allergy may be a food allergy.

[0558] As referred to herein, the "allergy to soybean" refers to the state in which an individual has an allergic reaction caused by proteins, etc. present in soybean which act as an antigen.

[0559] As referred to herein, the "antigen" refers to a substance that provokes an allergic reaction, and is also referred to as an "allergen component". The antigen is preferably a protein.

[0560] As referred to herein, the "protein" refers to a molecule having a structure in which naturally occurring amino acids are joined together by peptide bond. The number of amino acids present in a protein is not particularly limited. As referred to herein, the term "polypeptide" also means a molecule having a structure in which naturally occurring amino acids are joined together by peptide bond. The number of amino acids present in a polypeptide is not particularly limited. The "polypeptide" conceptually includes the "protein". Also, polypeptides having about 2 to 50 amino acids joined together by peptide bond are in some

cases called "peptides", especially. In the case where amino acids can form different enantiomers, the amino acids are understood to form an L-enantiomer, unless otherwise indicated. The amino acid sequences of proteins, polypeptides, or peptides as used herein are represented by one-letter symbols of amino acids in accordance with standard usage and the notational convention commonly used in the art. The leftward direction represents the amino-terminal direction, and the rightward direction represents the carboxy-terminal direction. In the one-letter symbols of amino acids, X can be any substance having an amino group and a carboxyl group that can bind to amino acids at both ends, and particularly represents that any of 20 types of naturally occurring amino acids are acceptable.

Identification of Antigens

[0561] Proteins contained in soybean were analyzed by the aforementioned technique to identify causative antigens of an allergy to soybean. To be specific, soybean proteins were subjected to two-dimensional electrophoresis under the conditions described below.

[0562] The electrophoresis in the first dimension was isoelectric focusing, which was performed using isoelectric focusing gels with a gel-strip length of 5 to 10 cm and a gel pH range of 3 to 10. The pH gradient of the gels in the direction of electrophoresis was as follows: when the total gel-strip length is taken as 1, the gel-strip length up to pH 5 is taken as "a", the gel-strip length from pH 5 to 7 is taken as "b", and the gel-strip length above pH 7 is taken as "c", "a" is in the range of 0.15 to 0.3, "b" is in the range of 0.4 to 0.7, and "c" is in the range of 0.15 to 0.3. More specifically, the isoelectric focusing was performed using the IPG gels, Immobiline Drystrip (pH3-10NL), produced by GE Healthcare Bio-Sciences Corporation (hereinafter abbreviated as "GE"). The electrophoresis system used was IPGphor produced by GE. The maximum current of the electrophoresis system was limited to 75 μA per gel strip. The voltage program adopted to perform the first-dimensional isoelectric focusing was as follows: (1) a constant voltage step was performed at a constant voltage of 300 V until the volt-hours reached 750 Vhr (the current variation width during electrophoresis for 30 minutes before the end of this step was 5 μA); (2) the voltage was increased gradually to 1000 V for 300 Vhr; (3) the voltage was further increased gradually to 5000 V for 4500 Vhr; and then (4) the voltage was held at a constant voltage of 5000 V until the total Vhr reached 12000.

[0563] The electrophoresis in the second dimension was SDS-PAGE, which was performed using polyacrylamide gels whose gel concentration at the distal end in the direction of electrophoresis was set to 3 to 6% and whose gel concentration at the proximal end was set to a higher value than that at the distal end. More specifically, the SDS-PAGE was performed using NuPAGE 4-12% Bris-Tris Gels (IPG well, Mini, 1 mm) produced by Life Technologies. The electrophoresis system used was XCell SureLock Mini-Cell produced by Life Technologies. The electrophoresis was run at a constant voltage of 200 V for about 45 minutes using an electrophoresis buffer composed of 50 mM MOPS, 50 mM Tris base, 0.1% (w/v) SDS and 1 mM EDTA.

[0564] As a result, the following spots in a two-dimensional electrophoresis gel run under the conditions described above for proteins in soybean have been revealed to exhibit specific binding to IgE antibodies from soybean-allergic patients.

[0565] Spot 1α: Molecular weight 70 to 120 kDa, pI 2.5 to 6.5

[0566] Spot 2α: Molecular weight 50 to 120 kDa, pI 2.5 to 6.5

[0567] Spot 4α: Molecular weight 70 to 120 kDa, pI 3.5 to 7.5

[0568] Spot 5α: Molecular weight 70 to 170 kDa, pI 3.5 to 7.5

[0569] Spot 6α: Molecular weight 70 to 120 kDa, pI 3.5 to 7.5

[0570] Spot 7α: Molecular weight 50 to 120 kDa, pI 3.5 to 7.5

[0571] Spot 8α: Molecular weight 50 to 120 kDa, pI 3.5 to 8.5

[0572] Spot 9α: Molecular weight 50 to 120 kDa, pI 3.5 to 8.5

[0573] Spot 10α: Molecular weight 30 to 90 kDa, pI 3.5 to 7.5

[0574] Spot 11α: Molecular weight 30 to 90 kDa, pI 3.5 to 8.5

[0575] Spot 12α: Molecular weight 30 to 90 kDa, pI 3.5 to 8.5

[0576] Spot 13α: Molecular weight 30 to 90 kDa, pI 3.5 to 7.5

[0577] Spot 14α: Molecular weight 30 to 90 kDa, pI 2.5 to 6.5

[0578] Spot 15α, 16α: Molecular weight 20 to 90 kDa, pI 3.5 to 7.5

[0579] Spot 17α: Molecular weight 20 to 90 kDa, pI 3.5 to 7.5

[0580] Spot 18α: Molecular weight 30 to 90 kDa, pI 3.5 to 8.5

[0581] Spot 19α: Molecular weight 30 to 90 kDa, pI 3.5 to 8.5

[0582] Spot 25α: Molecular weight 20 to 90 kDa, pI 3.5 to 7.5

[0583] Spot 26α: Molecular weight 20 to 90 kDa, pI 3.5 to 7.5

[0584] Spot 27α: Molecular weight 20 to 90 kDa, pI 3.5 to 7.5

[0585] Spot 28α: Molecular weight 10 to 70 kDa, pI 3.5 to 7.5

[0586] Spot 29α, 30α, 31α, 32α, 33α: Molecular weight 10 to 70 kDa, pI 3.5 to 8.5

[0587] Spot 34α: Molecular weight 10 to 70 kDa, pI 5.5 to 10.5

[0588] Spot 35α: Molecular weight 20 to 90 kDa, pI 5.5 to 10.5

[0589] Spot 36α, 37α: Molecular weight 10 to 70 kDa, pI 3.5 to 8.5

[0590] Spot 38α: Molecular weight 70 to 170 kDa, pI 5.5 to 10.5

[0591] Spot 39α: Molecular weight 5 to 60 kDa, pI 5.5 to 10.5

[0592] Spot 40α, 56α: Molecular weight 5 to 60 kDa, pI 5.5 to 10.5

[0593] Spot 41α: Molecular weight 5 to 60 kDa, pI 5.5 to 10.5

[0594] Spot 42α, 43α: Molecular weight 5 to 60 kDa, pI 5.5 to 10.5

[0595] Spot 45α: Molecular weight 2 to 50 kDa, pI 3.5 to 7.5

[0596] Spot 46α: Molecular weight 2 to 50 kDa, pI 3.5 to 7.5

[0597] Spot 47α: Molecular weight 5 to 50 kDa, pI 3.5 to 7.5

[0598] Spot 48α: Molecular weight 2 to 50 kDa, pI 3.5 to 8.5

[0599] Spot 50α: Molecular weight 2 to 50 kDa, pI 2.5 to 6.5

[0600] Spot 51α: Molecular weight 5 to 60 kDa, pI 2.5 to 6.5

[0601] Spot 52α: Molecular weight 10 to 70 kDa, pI 3.5 to 7.5

[0602] Spot 53α: Molecular weight 5 to 60 kDa, pI 3.5 to 8.5

[0603] Spot 54α: Molecular weight 5 to 60 kDa, pI 3.5 to 8.5 Spot 55α: Molecular weight 30 to 90 kDa, pI 5.5 to 10.5 Antigen

(1α) Antigen in Spot 1α

[0604] As the result of sequence identification of the antigen in spot 1α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 7-104 were detected.

[0605] Also, the mass spectroscopic data obtained for spot 1α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, endoplasmin homolog isoform X2 (amino acid sequence: SEQ ID NO: 70, encoding nucleotide sequence: SEQ ID NO: 69) was identified as a protein comprising any of SEQ ID NOs: 71 to 87 and endoplasmin homolog isoform X2 (amino acid sequence: SEQ ID NO: 89, encoding nucleotide sequence: SEQ ID NO: 88) was identified as a protein comprising any of SEQ ID NOs: 90 to 104.

[0606] Accordingly, the antigen in spot 1α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (1αA1-a) to (1αA1-c), (1αB1), (1αA2-a) to (1αA2-c), and (1αB2):

[0607] (1αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 70;

[0608] (1αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 70;

[0609] (1αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 69;

[0610] (1αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 69;

[0611] (1αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 69;

[0612] (1αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 71-87, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 71 to 87 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[0613] (1αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 89;

[0614] (1αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 89;

[0615] (1αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 88;

[0616] (1αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 88;

[0617] (1αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 88;

[0618] (1αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 90-104, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or 14 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 71 to 87 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0619] The protein selected from the group consisting of the proteins as defined above in (1αA1-a) to (1αA1-c), (1αB1), (1αA2-a) to (1αA2-c), and (1αB2) can be proteins that are found in a protein spot with a molecular weight of around 70 to 120 kDa, preferably around 80 to 110 kDa, and an isoelectric point of 2.5 to 6.5, preferably 3.0 to 6.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0620] Epitopes of the aforementioned antigens (1αA1-a) to (1αA1-e) and (1αB1) reside in amino acids 546-553 (SEQ ID NO: 652), amino acids 653-661 (SEQ ID NO: 655), amino acids 781-795 (SEQ ID NO: 657), amino acids 252-266 (SEQ ID NO: 658), amino acids 771-785 (SEQ ID NO: 659), and amino acids 493-502 (SEQ ID NO: 660) of SEQ ID NO: 70. When the antigen in spot 1α contains a variation in the amino acid sequence of SEQ ID NO: 70, amino acids at positions corresponding to the region of at least one of these epitopes may retain the sequence in SEQ ID NO: 70.

[0621] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 1α may be the following variant.

[0622] In the epitope having the sequence of SEQ ID NO: 652, sites of amino acids other than X in XXKXEXXX or XTKXEXXX are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 1α contains a variation in the amino acids of SEQ ID NO: 70, amino acids 546-553 of SEQ ID NO: 70 in this antigen may have the amino acid sequence of XXKXEXXX or XTKXEXXX.

[0623] In the epitope having the sequence of SEQ ID NO: 655, sites of amino acids other than X in SEQ ID NO: 653 or 654 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 1α contains a variation in the amino acid sequence of SEQ ID NO: 70, amino acids 653-661 of SEQ ID NO: 70 in this antigen may have the amino acid sequence of SEQ ID NO: 653 or 654.

[0624] In the epitope having the sequence of SEQ ID NO: 657, sites of amino acids other than X in SEQ ID NO: 656 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 1α contains a variation in the amino acid sequence of SEQ ID NO: 70, amino acids 781-795 of SEQ ID NO: 70 in this antigen may have the amino acid sequence of SEQ ID NO: 656.

[0625] Preferably, a protein that is an antigen in spot 1α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 70 or 89 (endoplasmin homolog isoform X2) or causative of an allergy to soybean.

(2α) Antigen in Spot 2a

[0626] As the result of sequence identification of the antigen in spot 2α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 107-153 were detected.

[0627] Also, the mass spectroscopic data obtained for spot 2a on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, heat shock cognate protein 80 (amino acid sequence: SEQ ID NO: 106, encoding nucleotide sequence: SEQ ID NO: 105) was identified as a protein comprising any of SEQ ID NOs: 107 to 120, heat shock protein 90-1 (amino acid sequence: SEQ ID NO: 122, encoding nucleotide sequence: SEQ ID NO: 121) was identified as a protein comprising any of SEQ ID NOs: 123 to 137, and hypothetical protein GLYMA_02G302500 was identified as a protein comprising any of SEQ ID NOs: 140 to 153.

[0628] Accordingly, the antigen in spot 1α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (2αA1-a) to (2αA1-e), (2αB1), (2αA2-a) to (2αA2-c), (2αB2), (2αA3-a) to (2αA3-c), and (2αB3):

[0629] (2αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 106;

[0630] (2αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 106;

[0631] (2αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 105;

[0632] (2αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 105;

[0633] (2αA1-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 105;

[0634] (2αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 107-120, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 107 to 120 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[0635] (2αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 122;

[0636] (2αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 122;

[0637] (2αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 121;

[0638] (2αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 121;

[0639] (2αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 121;

[0640] (2αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 123-137, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or 14 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 123 to 137 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0641] (2αA3-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 139;

[0642] (2αA3-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 139;

[0643] (2αA3-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 138;

[0644] (2αA3-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at

least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 138;

[0645] (2αA3-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 138;

[0646] (2αB3) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 140-153, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 140 to 153 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0647] The protein selected from the group consisting of the proteins as defined above in (2αA1-a) to (2αA1-c), (2αB1), (2αA2-a) to (2αA2-c), (2αB2), (2αA3-a) to (2αA3-c), and (2αB3) can be proteins that are found in a protein spot with a molecular weight of around 50 to 120 kDa, preferably around 60 to 110 kDa, and an isoelectric point of 2.5 to 6.5, preferably 3.0 to 6.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0648] Epitopes of the aforementioned antigens (2αA1-a) to (2αA1-c) and (2αB1) reside in amino acids 42-50 (SEQ ID NO: 663), amino acids 51-65 (SEQ ID NO: 666), amino acids 69-73 (SEQ ID NO: 669), amino acids 221-235 (SEQ ID NO: 670), amino acids 351-365 (SEQ ID NO: 671), amino acids 361-375 (SEQ ID NO: 672), amino acids 411-425 (SEQ ID NO: 675), amino acids 551-565 (SEQ ID NO: 676), amino acids 671-685 (SEQ ID NO: 677), amino acids 681-695 (SEQ ID NO: 678), amino acids 41-55 (SEQ ID NO: 681), amino acids 451-465 (SEQ ID NO: 684), amino acids 271-285 (SEQ ID NO: 689), amino acids 51-59 (SEQ ID NO: 690), amino acids 65-72 (SEQ ID NO: 691), and amino acids 453-460 (SEQ ID NO: 692) of SEQ ID NO: 106. When the antigen in spot 2α contains a variation in the amino acid sequence of SEQ ID NO: 106, amino acids at positions corresponding to the region of at least one of these epitopes may retain the sequence in SEQ ID NO: 106.

[0649] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 2α may be the following variant.

[0650] In the epitope having the sequence of SEQ ID NO: 663, sites of amino acids other than X in DXXDKXXXX or SEQ ID NO: 662 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 2α contains a variation in the amino acid sequence of SEQ ID NO: 106, amino acids 42-50 of SEQ ID NO: 106 in this antigen may have the amino acid sequence of DXXDKXXXX or SEQ ID NO: 662.

[0651] In the epitope having the sequence of SEQ ID NO: 666, sites of amino acids other than X in SEQ ID NO: 664 or 665 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 2α contains a variation in the amino acid sequence of SEQ ID NO: 106, amino acids 51-65 of SEQ ID NO: 106 in this antigen may have the amino acid sequence of SEQ ID NO: 664 or 665.

[0652] In the epitope having the sequence of SEQ ID NO: 669, sites of amino acids other than X in XPDKX or SEQ ID NO: 668 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 2α contains a variation in the amino acid sequence of SEQ ID NO: 106, amino acids 69-73 of SEQ ID NO: 106 in this antigen may have the amino acid sequence of XPDKX or SEQ ID NO: 668.

[0653] In the epitope having the sequence of SEQ ID NO: 675, sites of amino acids other than X in SEQ ID NO: 673 or 674 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 2α contains a variation in the amino acid sequence of SEQ ID NO: 106, amino acids 411-425 of SEQ ID NO: 106 in this antigen may have the amino acid sequence of SEQ ID NO: 673 or 674.

[0654] In the epitope having the sequence of SEQ ID NO: 681, sites of amino acids other than X in XXXXDXXXXEXXXDX or SEQ ID NO: 680 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 2α contains a variation in the amino acid sequence of SEQ ID NO: 106, amino acids 41-55 of SEQ ID NO: 106 in this antigen may have the amino acid sequence of XXXXDXXXXEXXXDX or SEQ ID NO: 680.

[0655] In the epitope having the sequence of SEQ ID NO: 684, sites of amino acids other than X in SEQ ID NO: 682 or 685 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 2α contains a variation in the amino acid sequence of SEQ ID NO: 106, amino acids 451-465 of SEQ ID NO: 106 in this antigen may have the amino acid sequence of SEQ ID NO: 682 or 685.

[0656] In the epitope having the sequence of SEQ ID NO: 689, sites of amino acids other than X in SEQ ID NO: 688 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 2α contains a variation in the amino acid sequence of SEQ ID NO: 106, amino acids 271-285 of SEQ ID NO: 106 in this antigen may have the amino acid sequence of SEQ ID NO: 688.

[0657] Preferably, a protein that is an antigen in spot 2α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 106 (heat shock cognate protein 80), the protein having the amino acid sequence of SEQ ID NO: 122 (heat shock protein 90-1), or the protein having the amino acid sequence of SEQ ID NO: 139 (hypothetical protein GLYMA_02G302500), or causative of an allergy to soybean.

(4α) Antigen in Spot 4α

[0658] As the result of sequence identification of the antigen in spot 4α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 156 and 157 were detected.

[0659] Also, the mass spectroscopic data obtained for spot 4α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, embryo-specific urease isoform X1 (amino acid sequence: SEQ ID NO: 155, encoding nucleotide sequence: SEQ ID NO: 154) was identified as a protein comprising SEQ ID NO: 156 or 157.

[0660] Accordingly, the antigens in spot 4α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (4αA-a) to (4αA-c) and (4αB):

[0661] (4αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 155;

[0662] (4αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 155;

[0663] (4αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 154;

[0664] (4αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 154;

[0665] (4αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 154;

[0666] (4αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 156 and 157, preferably a protein comprising both the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 156 and 157 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0667] The proteins selected from the group consisting of (4αA-a) to (4αA-c) and (4αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 70 to 120 kDa, preferably around 80 to 110 kDa and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0668] Epitopes of the antigen in spot 4α reside in amino acids 121-130 (SEQ ID NO: 719), amino acids 289-295 (SEQ ID NO: 722), amino acids 293-302 (SEQ ID NO: 725), amino acids 501-515 (SEQ ID NO: 726), amino acids 571-585 (SEQ ID NO: 727), amino acids 591-605 (SEQ ID NO: 728), and amino acids 771-785 (SEQ ID NO: 731) of SEQ ID NO: 155. The antigen in spot 4α may retain the sequence in SEQ ID NO: 155 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0669] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 4α may be the following variant.

[0670] In the epitope having the sequence of SEQ ID NO: 719, sites of amino acids other than X in XXXXXXXKED or SEQ ID NO: 718 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 121-130 of SEQ ID NO: 155 in the antigen in spot 4α may have the amino acid sequence of XXXXXXXKED or SEQ ID NO: 718.

[0671] In the epitope having the sequence of SEQ ID NO: 722, sites of amino acids other than X in SEQ ID NO: 720 or 721 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 289-295 of SEQ ID NO: 155 in the antigen in spot 4α may have the amino acid sequence of SEQ ID NO: 720 or 721.

[0672] In the epitope having the sequence of SEQ ID NO: 725, sites of amino acids other than X in SEQ ID NO: 723 or 724 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 293-302 of SEQ ID NO: 155 in the antigen in spot 4α may have the amino acid sequence of SEQ ID NO: 723 or 724.

[0673] In the epitope having the sequence of SEQ ID NO: 731, sites of amino acids other than X in SEQ ID NO: 729 or 730 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 771-785 of SEQ ID NO: 155 in the antigen in spot 4α may have the amino acid sequence of SEQ ID NO: 729 or 730.

[0674] Preferably, a protein that is an antigen in spot 4α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 155 (embryo-specific urease isoform X1) or causative of an allergy to soybean.

(5α) Antigen in Spot 5α

[0675] As the result of sequence identification of the antigen in spot 5α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 160 and 161 were detected.

[0676] Also, the mass spectroscopic data obtained for spot 5α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, alpha-1,4 glucan phosphorylase L isozyme, chloroplastic/amyloplastic (amino acid sequence: SEQ ID NO: 159, encoding nucleotide sequence: SEQ ID NO: 158) was identified as a protein comprising SEQ ID NO: 160 or 161.

[0677] Accordingly, the antigens in spot 5α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (5αA-a) to (5αA-c) and (5αB):

[0678] (5αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 159;

[0679] (5αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 159;

[0680] (5αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 158;

[0681] (5αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 158;

[0682] (5αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 158;

[0683] (5αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 160 and 161, preferably a protein comprising both the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 160 and 161 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0684] The proteins selected from the group consisting of (5αA-a) to (5αA-c) and (5αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 70 to 170 kDa, preferably around 80 to 160 kDa and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0685] Epitopes of the antigen in spot 5α reside in amino acids 861-875 (SEQ ID NO: 734), amino acids 871-885 (SEQ ID NO: 735), amino acids 891-905 (SEQ ID NO: 736), amino acids 901-915 (SEQ ID NO: 737), and amino acids 951-965 (SEQ ID NO: 738) of SEQ ID NO: 159. The antigen in spot 5α may retain the sequence in SEQ ID NO: 159 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0686] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 5α may be the following variant.

[0687] In the epitope having the sequence of SEQ ID NO: 734, sites of amino acids other than X in SEQ ID NO: 732 or 733 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 861-875 of SEQ ID NO: 159 in the antigen in spot 5a may have the amino acid sequence of SEQ ID NO: 732 or 733.

[0688] Preferably, a protein that is an antigen in spot 5α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 159 (alpha-1,4 glucan phosphorylase L isozyme, chloroplastic/amyloplastic) or causative of an allergy to soybean.

(6α) Antigen in Spot 6α

[0689] As the result of sequence identification of the antigen in spot 6α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 164-172 were detected.

[0690] Also, the mass spectroscopic data obtained for spot 6α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, aconitate hydratase 1 (amino acid sequence: SEQ ID NO: 163, encoding nucleotide sequence: SEQ ID NO: 162) was identified as a protein comprising any of SEQ ID NOs: 164 to 167 and aconitate hydratase 1 (amino acid sequence: SEQ ID NO: 169, encoding nucleotide sequence: SEQ ID NO: 168) was identified as a protein comprising any of SEQ ID NOs: 170 to 172.

[0691] Accordingly, the antigen in spot 6α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (6αA1-a) to (6αA1-e), (6αB1), (6αA2-a) to (6αA2-c), and (6αB2):

[0692] (6αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 163;

[0693] (6αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 163;

[0694] (6αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 162;

[0695] (6αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 162;

[0696] (6αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 162;

[0697] (6αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 164-167, preferably a protein comprising at least 2, 3, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 164 to 167 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[0698] (6αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 169;

[0699] (6αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 169;

[0700] (6αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 168;

[0701] (6αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 168;

[0702] (6αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 168;

[0703] (6αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 170-172, preferably a protein comprising at least 2, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 170 to 172 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0704] The protein selected from the group consisting of the proteins as defined above in (6αA1-a) to (6αA1-c), (6αB1), (6αA2-a) to (6αA2-c), and (6αB2) can be proteins that are found in a protein spot with a molecular weight of

around 70 to 120 kDa, preferably around 80 to 110 kDa, and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0705] Epitopes of the aforementioned antigens (6αA1-a) to (6αA1-e) and (6αB1) reside in amino acids 761-775 (SEQ ID NO: 739), amino acids 831-845 (SEQ ID NO: 740) of SEQ ID NO: 163. When the antigen in spot 6α contains a variation in the amino acid sequence of SEQ ID NO: 163, amino acids at positions corresponding to the region of at least one of these epitopes may retain the sequence in SEQ ID NO: 163.

[0706] Preferably, a protein that is an antigen in spot 6α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 163 or 169 (aconitate hydratase 1) or causative of an allergy to soybean.

(7α) Antigen in Spot 7α

[0707] As the result of sequence identification of the antigen in spot 7a by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 175-187 were detected.

[0708] Also, the mass spectroscopic data obtained for spot 7a on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, methionine synthase (amino acid sequence: SEQ ID NO: 174, encoding nucleotide sequence: SEQ ID NO: 173) was identified as a protein comprising any of SEQ ID NOs: 175 to 187.

[0709] Accordingly, the antigens in spot 7α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (7αA-a) to (7αA-c) and (7αB):

[0710] (7αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 174;

[0711] (7αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 174;

[0712] (7αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 173;

[0713] (7αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 173;

[0714] (7αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 173;

[0715] (7αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 175-187, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 175-187 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0716] The proteins selected from the group consisting of (7αA-a) to (7αA-c) and (7αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 50 to 120 kDa, preferably around 60 to 110 kDa and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0717] Epitopes of the antigen in spot 7α reside in amino acids 471-485 (SEQ ID NO: 741), amino acids 581-595 (SEQ ID NO: 744) of SEQ ID NO: 174. The antigen in spot 7α may retain the sequence in SEQ ID NO: 174 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0718] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 7a may be the following variant.

[0719] In the epitope having the sequence of SEQ ID NO: 744, sites of amino acids other than X in SEQ ID NO: 742 or 743 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 581-595 of SEQ ID NO: 174 in the antigen in spot 7a may have the amino acid sequence of SEQ ID NO: 742 or 743.

[0720] Preferably, a protein that is an antigen in spot 7α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 174 (methionine synthase) or causative of an allergy to soybean.

(8α) Antigen in Spot 8α

[0721] As the result of sequence identification of the antigen in spot 8α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 190-199 were detected.

[0722] Also, the mass spectroscopic data obtained for spot 8α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, transketolase, chloroplastic (amino acid sequence: SEQ ID NO: 189, encoding nucleotide sequence: SEQ ID NO: 188) was identified as a protein comprising any of SEQ ID NOs: 190 to 199.

[0723] Accordingly, the antigens in spot 8α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (8αA-a) to (8αA-c) and (8αB):

[0724] (8αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 189;

[0725] (8αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 189;

[0726] (8αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 188;

[0727] (8αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least

85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 188;

[0728] (8αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 188;

[0729] (8αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 190-199, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8 or 9 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 190-199 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0730] The proteins selected from the group consisting of (8αA-a) to (8αA-c) and (8αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 50 to 120 kDa, preferably around 60 to 110 kDa and an isoelectric point of 3.5 to 8.5, preferably 4.0 to 8.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0731] Preferably, a protein that is an antigen in spot 8α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 189 (transketolase, chloroplastic) or causative of an allergy to soybean.

(9α) Antigen in Spot 9α

[0732] As the result of sequence identification of the antigen in spot 9α by mass spectroscopy, the amino acid sequence of SEQ ID NO: 202 was detected.

[0733] Also, the mass spectroscopic data obtained for spot 9α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, lysine-tRNA ligase, cytoplasmic-like isoform X2 (amino acid sequence: SEQ ID NO: 201, encoding nucleotide sequence: SEQ ID NO: 200) was identified as a protein comprising SEQ ID NO: 202.

[0734] Accordingly, the antigens in spot 9α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (9αA-a) to (9αA-e) and (9αB):

[0735] (9αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 201;

[0736] (9αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 201;

[0737] (9αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 200;

[0738] (9αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 200;

[0739] (9αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 200;

[0740] (9αB) a protein comprising an amino acid sequence of SEQ ID NO: 202. As referred to above, the amino acid sequence(s) of SEQ ID NO: 202 may be an amino acid sequence(s) derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0741] The proteins selected from the group consisting of (9αA-a) to (9αA-c) and (9αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 50 to 120 kDa, preferably around 60 to 110 kDa and an isoelectric point of 3.5 to 8.5, preferably 4.0 to 8.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0742] Epitopes of the antigen in spot 9α reside in amino acids 236-245 (SEQ ID NO: 747), amino acids 283-290 (SEQ ID NO: 749) of SEQ ID NO: 201. The antigen in spot 9α may retain the sequence in SEQ ID NO: 201 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0743] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 9α may be the following variant.

[0744] In the epitope having the sequence of SEQ ID NO: 747, sites of amino acids other than X in XXEXXXXKDX or SEQ ID NO: 746 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 236-245 of SEQ ID NO: 201 in the antigen in spot 9α may have the amino acid sequence of XXEXXXXKDX or SEQ ID NO: 746.

[0745] In the epitope having the sequence of SEQ ID NO: 749, sites of amino acids other than X in SEQ ID NO: 748 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 283-290 of SEQ ID NO: 201 in the antigen in spot 9α may have the amino acid sequence of SEQ ID NO: 748.

[0746] Preferably, a protein that is an antigen in spot 9α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 201 (lysine-tRNA ligase, cytoplasmic-like isoform X2) or causative of an allergy to soybean.

(10α) Antigen in Spot 10α

[0747] As the result of sequence identification of the antigen in spot 10α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 205-231 were detected.

[0748] Also, the mass spectroscopic data obtained for spot 10α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, chaperonin CPN60-2, mitochondrial (amino acid sequence: SEQ ID NO: 204, encoding nucleotide sequence: SEQ ID NO: 203) was identified as a protein comprising any of SEQ ID NOs: 205 to 218 and chaperonin CPN60-2, mitochondrial-like isoform X1 (amino acid sequence: SEQ ID NO: 220, encoding nucleotide sequence: SEQ ID NO: 219) was identified as a protein comprising any of SEQ ID NOs: 221 to 231.

[0749] Accordingly, the antigen in spot 10α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (10αA1-a) to (10αA1-c), (10αB1), (10αA2-a) to (10αA2-c), and (10αB2):

[0750] (10αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 204;

[0751] (10αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 204;

[0752] (10αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 203;

[0753] (10αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 203;

[0754] (10αA1-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 203;

[0755] (10αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 205-218, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 205 to 218 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[0756] (10αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 220;

[0757] (10αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 220;

[0758] (10αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 219; (10A2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 219;

[0759] (10αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 219;

[0760] (10β2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 221-231, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, or 10, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 221 to 231 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0761] The protein selected from the group consisting of the proteins as defined above in (10αA1-a) to (10αA1-c), (10αB1), (10αA2-a) to (10αA2-e), and (10αB2) can be proteins that are found in a protein spot with a molecular weight of around 30 to 90 kDa, preferably around 40 to 80 kDa, and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0762] Preferably, a protein that is an antigen in spot 10α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 204 (haperonin CPN60-2, mitochondrial) or the protein having the amino acid sequence of SEQ ID NO: 220 (chaperonin CPN60-2, mitochondrial-like isoform X1) or causative of an allergy to soybean.

(11α) Antigen in Spot 11α

[0763] As the result of sequence identification of the antigen in spot 11a by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 234-237 were detected.

[0764] Also, the mass spectroscopic data obtained for spot 11a on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, polyadenylate-binding protein 8 isoform X3 (amino acid sequence: SEQ ID NO: 233, encoding nucleotide sequence: SEQ ID NO: 232) was identified as a protein comprising any of SEQ ID NOs: 234 to 237.

[0765] Accordingly, the antigens in spot 11α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (11αA-a) to (11αA-c) and (11αB):

[0766] (11αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 233;

[0767] (11αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 233;

[0768] (11αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 232;

[0769] (11αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 232;

[0770] (11αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 232;

[0771] (11αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 234-237, preferably a protein comprising at least 2 or 3 or all sequences of the amino acid sequences. As referred to above, the amino acid

sequence of any of SEQ ID NOs: 234-237 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0772] The proteins selected from the group consisting of (11αA-a) to (11αA-c) and (11αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 30 to 90 kDa, preferably around 40 to 80 kDa and an isoelectric point of 3.5 to 8.5, preferably 4.0 to 8.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0773] Epitopes of the antigen in spot 11a reside in amino acids 454-463 (SEQ ID NO: 753), amino acids 363-370 (SEQ ID NO: 754) of SEQ ID NO: 233. The antigen in spot 11α may retain the sequence in SEQ ID NO: 233 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0774] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 11a may be the following variant.

[0775] In the epitope having the sequence of SEQ ID NO: 753, sites of amino acids other than X in SEQ ID NO: 751 or 752 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 454-463 of SEQ ID NO: 223 in the antigen in spot 11α may have the amino acid sequence of SEQ ID NO: 751 or 752.

[0776] Preferably, a protein that is an antigen in spot 11α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 233 (polyadenylate-binding protein 8 isoform X3) or causative of an allergy to soybean.

(12α) Antigen in Spot 12α

[0777] As the result of sequence identification of the antigen in spot 12α by mass spectroscopy, the amino acid sequence of SEQ ID NO: 240 was detected.

[0778] Also, the mass spectroscopic data obtained for spot 12α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, pyrophosphate-fructose 6-phosphate 1-phosphotransferase subunit alpha-like (amino acid sequence: SEQ ID NO: 239, encoding nucleotide sequence: SEQ ID NO: 238) was identified as a protein comprising SEQ ID NO: 240.

[0779] Accordingly, the antigens in spot 12α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (12αA-a) to (12αA-c) and (12αB):

[0780] (12αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 239;

[0781] (12αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 239;

[0782] (12αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 238;

[0783] (12αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 238;

[0784] (12αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 238;

[0785] (12αB) a protein comprising an amino acid sequence of SEQ ID NO: 240. As referred to above, the amino acid sequence(s) of SEQ ID NO: 240 may be an amino acid sequence(s) derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0786] The proteins selected from the group consisting of (12αA-a) to (12αA-c) and (12αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 30 to 90 kDa, preferably around 40 to 80 kDa and an isoelectric point of 3.5 to 8.5, preferably 4.0 to 8.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0787] An epitope of the antigen in spot 12α resides in amino acids 551-565 (SEQ ID NO: 755) of SEQ ID NO: 239. The antigen in spot 11α may retain the sequence in SEQ ID NO: 239 as to amino acids at positions corresponding to the region of this epitope.

[0788] Preferably, a protein that is an antigen in spot 12α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 239 (pyrophosphate-fructose 6-phosphate 1-phosphotransferase subunit alpha-like) or causative of an allergy to soybean.

(13α) Antigen in Spot 13α

[0789] As the result of sequence identification of the antigen in spot 13α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 243-278 were detected.

[0790] Also, the mass spectroscopic data obtained for spot 13α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, protein disulfide isomerase-like protein precursor (amino acid sequence: SEQ ID NO: 242, encoding nucleotide sequence: SEQ ID NO: 241) was identified as a protein comprising any of SEQ ID NOs: 243 to 252 and protein disulfide-isomerase (amino acid sequence: SEQ ID NO: 254, encoding nucleotide sequence: SEQ ID NO: 253) was identified as a protein comprising any of SEQ ID NOs: 255 to 278.

[0791] Accordingly, the antigen in spot 13α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (13αA1-a) to (13αA1-e), (13αB1), (13αA2-a) to (13αA2-c), and (13αB2):

[0792] (13αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 242;

[0793] (13αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 242;

[0794] (13αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 241;

[0795] (13A1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 241;

[0796] (13αA1-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 241;

[0797] (13αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 243-252, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, or 9, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 205 to 218 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[0798] (13αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 254;

[0799] (13αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 254;

[0800] (13αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 253;

[0801] (13αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 253;

[0802] (13αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 253;

[0803] (13αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 255-278, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, or 23, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 221 to 231 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0804] The protein selected from the group consisting of the proteins as defined above in (13αA1-a) to (13αA1-c), (13αB1), (13αA2-a) to (13αA2-c), and (13αB2) can be proteins that are found in a protein spot with a molecular weight of around 30 to 90 kDa, preferably around 40 to 80 kDa, and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0805] Epitopes of the aforementioned antigens (13αA1-a) to (13αA1-c) and (13αB1) reside in amino acids 471-485 (SEQ ID NO: 758) of SEQ ID NO: 242. When the antigen in spot 13α contains a variation in the amino acid sequence of SEQ ID NO: 242, this antigen may retain the sequence in SEQ ID NO: 242 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0806] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 13α may be the following variant.

[0807] In the epitope having the sequence of SEQ ID NO: 758, sites of amino acids other than X in XXGXXXXXXXXXXXE or SEQ ID NO: 757 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 13α contains a variation in the amino acid sequence of SEQ ID NO: 242, amino acids 471-485 of SEQ ID NO: 242 in this antigen may have the amino acid sequence of XXGXXXXXXXXXXXE or SEQ ID NO: 757.

[0808] Preferably, a protein that is an antigen in spot 13α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 242 (protein disulfide isomerase-like protein precursor) or the protein having the amino acid sequence of SEQ ID NO: 254 (protein disulfide-isomerase) or causative of an allergy to soybean.

(14α) Antigen in Spot 14α

[0809] As the result of sequence identification of the antigen in spot 14α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 281-289 were detected.

[0810] Also, the mass spectroscopic data obtained for spot 14α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, V-type proton ATPase subunit B2 (amino acid sequence: SEQ ID NO: 280, encoding nucleotide sequence: SEQ ID NO: 279) was identified as a protein comprising any of SEQ ID NOs: 281 to 284 and V-type proton ATPase subunit B2 (amino acid sequence: SEQ ID NO: 286, encoding nucleotide sequence: SEQ ID NO: 285) was identified as a protein comprising any of SEQ ID NOs: 287 to 289.

[0811] Accordingly, the antigen in spot 14α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (14αA1-a) to (14αA1-c), (14αB1), (14αA2-a) to (14αA2-c), and (14«B2):

[0812] (14αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 280;

[0813] (14αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 280;

[0814] (14αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 279;

[0815] (14αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least

85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 279;

[0816] (14αA1-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 279;

[0817] (14αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 281-284, preferably a protein comprising at least 2, 3, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 281 to 284 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[0818] (14αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 286;

[0819] (14αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 286;

[0820] (14αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 285;

[0821] (14αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 285;

[0822] (14αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 285;

[0823] (14«B2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 287-289, preferably a protein comprising at least 2, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 287 to 289 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0824] The protein selected from the group consisting of the proteins as defined above in (14αA1-a) to (14αA1-c), (14αB1), (14αA2-a) to (14αA2-c), and (14αB2) can be proteins that are found in a protein spot with a molecular weight of around 30 to 90 kDa, preferably around 40 to 80 kDa, and an isoelectric point of 2.5 to 6.5, preferably 3.0 to 6.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0825] Epitopes of the aforementioned antigens (14αA1-a) to (14αA1-c) and (14αB1) reside in amino acids 191-205 (SEQ ID NO: 761) and amino acids 391-405 (SEQ ID NO: 764) of SEQ ID NO: 280. When the antigen in spot 14α contains a variation in the amino acid sequence of SEQ ID NO: 280, this antigen may retain the sequence in SEQ ID NO: 280 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0826] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 14α may be the following variant.

[0827] In the epitope having the sequence of SEQ ID NO: 761, sites of amino acids other than X in SEQ ID NO: 759 or 760 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 14α contains a variation in the amino acid sequence of SEQ ID NO: 280, amino acids 191-205 of SEQ ID NO: 280 in this antigen may each independently have the amino acid sequence of SEQ ID NO: 759 or 760.

[0828] In the epitope having the sequence of SEQ ID NO: 764, sites of amino acids other than X in SEQ ID NO: 762 or 763 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 14α contains a variation in the amino acid sequence of SEQ ID NO: 280, amino acids 391-405 of SEQ ID NO: 280 in this antigen may each independently have the amino acid sequence of SEQ ID NO: 762 or 763.

[0829] Preferably, a protein that is an antigen in spot 14α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 280 or 286 (V-type proton ATPase subunit B2) or causative of an allergy to soybean.

(15α, 16α) Antigens in Spots 15α, 16α

[0830] As the result of sequence identification of the antigens in spots 15α and 16α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 292-318 were detected.

[0831] Also, the mass spectroscopic data obtained for spots 15α, 16a on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, UTP-glucose-1-phosphate uridylyltransferase (amino acid sequence: SEQ ID NO: 291, encoding nucleotide sequence: SEQ ID NO: 290) was identified as a protein comprising any of SEQ ID NOs: 292 to 304 and the hypothetical protein GLYMA_02G241100 (amino acid sequence: SEQ ID NO: 306, encoding nucleotide sequence: SEQ ID NO: 305) was identified as a protein comprising any of SEQ ID NOs: 307 to 318.

[0832] Accordingly, the antigen in spot 15α, 16α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (15, 16αA1-a) to (15, 16αA1-c), (15, 16αB1), (15, 16αA2-a) to (15, 16αA2-c), and (15, 16αB2): (15, 16αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 291; (15, 16αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 291;

[0833] (15, 16αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 290;

[0834] (15, 16αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%,

at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 290;

[0835] (15, 16αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 290;

[0836] (15, 16αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 292-304, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 292 to 304 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[0837] (15, 16αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 306;

[0838] (15, 16αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 306;

[0839] (15, 16αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 305;

[0840] (15, 16αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 305;

[0841] (15, 16αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 305;

[0842] (15, 16αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 307-318, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 307 to 318 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0843] The protein selected from the group consisting of the proteins as defined above in (15, 16αA1-a) to (15, 16αA1-c), (15, 16αB1), (15, 16αA2-a) to (15, 16αA2-c), and (15, 16αB2) can be proteins that are found in a protein spot with a molecular weight of around 20 to 90 kDa, preferably around 30 to 80 kDa, and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0844] Epitopes of the aforementioned antigens (15, 16αA1-a) to (15, 16αA1-e) and (15, 16β1) reside in amino acids 211-225 (SEQ ID NO: 766) of SEQ ID NO: 291. When the antigen in spot 15α, 16α contains a variation in the amino acid sequence of SEQ ID NO: 291, this antigen may retain the sequence in SEQ ID NO: 291 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0845] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 15α, 16a may be the following variant.

[0846] In the epitope having the sequence of SEQ ID NO: 766, sites of amino acids other than X in XGKXXXXXXXXXXXX are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 15α, 16α contains a variation in the amino acid sequence of SEQ ID NO: 291, amino acids 211-225 of SEQ ID NO: 291 in this antigen may have the amino acid sequence of XGKXXXXXXXXXXXX.

[0847] Preferably, a protein that is an antigen in spot 15α or 16α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 291 (UTP-glucose-1-phosphate uridylyltransferase) or the protein having the amino acid sequence of SEQ ID NO: 306 (hypothetical protein GLYMA_02G241100) or causative of an allergy to soybean.

(17α) Antigen in Spot 17α

[0848] As the result of sequence identification of the antigen in spot 17 by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 321-332 were detected.

[0849] Also, the mass spectroscopic data obtained for spot 17a on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, guanosine nucleotide diphosphate dissociation inhibitor 2 (amino acid sequence: SEQ ID NO: 320, encoding nucleotide sequence: SEQ ID NO: 319) was identified as a protein comprising any of SEQ ID NOs: 321 to 325 and rab GDP dissociation inhibitor alpha-like (amino acid sequence: SEQ ID NO: 327, encoding nucleotide sequence: SEQ ID NO: 326) was identified as a protein comprising any of SEQ ID NOs: 328 to 332.

[0850] Accordingly, the antigen in spot 17α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (17αA1-a) to (17αA1-c), (17αB1), (17αA2-a) to (17αA2-c), and (17αB2):

[0851] (17αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 320;

[0852] (17αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 320;

[0853] (17αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 319;

[0854] (17αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 319;

[0855] (17αA1-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 319;

[0856] (17αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 321-325, preferably a protein comprising at least 2, 3, or 4 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 321 to 325 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0857] (17αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 327;

[0858] (17αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 327;

[0859] (17αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 326;

[0860] (17αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 326;

[0861] (17αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 326;

[0862] (17αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 328-332, preferably a protein comprising at least 2, 3, or 4 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 328 to 332 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[0863] The protein selected from the group consisting of the proteins as defined above in (17αA1-a) to (17αA1-c), (17αB1), (17αA2-a) to (17αA2-c), and (17αB2) can be proteins that are found in a protein spot with a molecular weight of around 20 to 90 kDa, preferably around 30 to 80 kDa, and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0864] Epitopes of the aforementioned antigens (17αA1-a) to (17αA1-c) and (17αB1) reside in amino acids 171-185 (SEQ ID NO: 767) of SEQ ID NO: 320. When the antigen in spot 17α contains a variation in the amino acid sequence of SEQ ID NO: 320, this antigen may retain the sequence in SEQ ID NO: 320 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0865] Preferably, a protein that is an antigen in spot 17a has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 320 (guanosine nucleotide diphosphate dissociation inhibitor 2) or the protein having the amino acid sequence of SEQ ID NO: 327 (rab GDP dissociation inhibitor alpha-like) or causative of an allergy to soybean.

(18α) Antigen in Spot 18α

[0866] As the result of sequence identification of the antigen in spot 18α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 335-339 were detected.

[0867] Also, the mass spectroscopic data obtained for spot 18α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, hypothetical protein GLYMA_10G028300 (amino acid sequence: SEQ ID NO: 334, encoding nucleotide sequence: SEQ ID NO: 333) was identified as a protein comprising any of SEQ ID NOs: 335 to 339.

[0868] Accordingly, the antigens in spot 18α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (18αA-a) to (18αA-c) and (18αB):

[0869] (18αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 334;

[0870] (18αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 334;

[0871] (18αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 333;

[0872] (18αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 333;

[0873] (18αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 333;

[0874] (18αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 335-339, preferably a protein comprising at least 2, 3 or 4 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 335-339 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0875] The proteins selected from the group consisting of (18αA-a) to (18αA-e) and (18αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 30 to 90 kDa, preferably around 40 to 80 kDa and an isoelectric point of 3.5 to 8.5, preferably 4.0 to 8.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0876] Epitopes of the antigen in spot 18α reside in amino acids 52-60 (SEQ ID NO: 750), amino acids 81-95 (SEQ ID NO: 768), amino acids 101-115 (SEQ ID NO: 771), amino

acids 281-295 (SEQ ID NO: 772), amino acids 491-500 (SEQ ID NO: 775) of SEQ ID NO: 334. The antigen in spot 18α may retain the sequence in SEQ ID NO: 334 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0877] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 18α may be the following variant.

[0878] In the epitope having the sequence of SEQ ID NO: 771, sites of amino acids other than X in SEQ ID NO: 769 or 770 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 101-115 of SEQ ID NO: 334 in the antigen in spot 18α may have the amino acid sequence of SEQ ID NO: 769 or 770.

[0879] In the epitope having the sequence of SEQ ID NO: 775, sites of amino acids other than X in SEQ ID NO: 773 or 774 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 491-500 of SEQ ID NO: 334 in the antigen in spot 18α may have the amino acid sequence of SEQ ID NO: 773 or 774.

[0880] Preferably, a protein that is an antigen in spot 18α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 334 (hypothetical protein GLYMA_10G028300) or causative of an allergy to soybean.

(19α) Antigen in Spot 19α

[0881] As the result of sequence identification of the antigen in spot 19α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 342-346 were detected.

[0882] Also, the mass spectroscopic data obtained for spot 19α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, sucrose binding protein homolog S-64 (amino acid sequence: SEQ ID NO: 341, encoding nucleotide sequence: SEQ ID NO: 340) was identified as a protein comprising any of SEQ ID NOs: 342 to 346.

[0883] Accordingly, the antigens in spot 19α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (19αA-a) to (19αA-c) and (19αB):

[0884] (19αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 341;

[0885] (19αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 341;

[0886] (19αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 340;

[0887] (19αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 340;

[0888] (19αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 340;

[0889] (19αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 342-346, preferably a protein comprising at least 2, 3 or 4 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 342-346 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0890] The proteins selected from the group consisting of (19αA-a) to (19αA-c) and (19αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 30 to 90 kDa, preferably around 40 to 80 kDa and an isoelectric point of 3.5 to 8.5, preferably 4.0 to 8.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0891] Epitopes of the antigen in spot 19α reside in amino acids 191-205 (SEQ ID NO: 778), amino acids 475-489 (SEQ ID NO: 781), amino acids 81-95 (SEQ ID NO: 784), amino acids 436-450 (SEQ ID NO: 786), amino acids 21-30 (SEQ ID NO: 787) of SEQ ID NO: 341. The antigen in spot 19α may retain the sequence in SEQ ID NO: 341 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0892] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 19α may be the following variant.

[0893] In the epitope having the sequence of SEQ ID NO: 778, sites of amino acids other than X in SEQ ID NO: 776 or 778 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 191-205 of SEQ ID NO: 341 in the antigen in spot 19α may have the amino acid sequence of SEQ ID NO: 776 or 778.

[0894] In the epitope having the sequence of SEQ ID NO: 781, sites of amino acids other than X in XXXXXXXXRXRRXXX or SEQ ID NO: 780 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 475-489 of SEQ ID NO: 341 in the antigen in spot 19α may have the amino acid sequence of XXXXXXXXRXRRXXX or SEQ ID NO: 780.

[0895] In the epitope having the sequence of SEQ ID NO: 784, sites of amino acids other than X in XXXXXXXXXXEEEXX or SEQ ID NO: 783 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 81-95 of SEQ ID NO: 341 in the antigen in spot 19α may have the amino acid sequence of XXXXXXXXXXEEEXX or SEQ ID NO: 783.

[0896] Preferably, a protein that is an antigen in spot 19α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 341 (sucrose binding protein homolog S-64) or causative of an allergy to soybean.

(25α) Antigen in Spot 25α

[0897] As the result of sequence identification of the antigen in spot 25α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 349-365 were detected.

[0898] Also, the mass spectroscopic data obtained for spot 25α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial (amino acid sequence: SEQ ID NO: 348, encoding nucleotide sequence: SEQ ID NO: 347) was identified as a protein comprising any of SEQ ID NOs: 349 to 356 and succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial (amino acid sequence: SEQ ID NO: 358, encoding nucleotide sequence: SEQ ID NO: 357) was identified as a protein comprising any of SEQ ID NOs: 359 to 365.

[0899] Accordingly, the antigen in spot 25α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (25αA1-a) to (25αA1-c), (25αB1), (25αA2-a) to (25αA2-c), and (25αB2):

[0900] (25αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 348;

[0901] (25αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 348;

[0902] (25αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 347;

[0903] (25αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 347;

[0904] (25αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 347;

[0905] (25αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 349-356, preferably a protein comprising at least 2, 3, 4, 5, 6, or 7 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 349 to 356 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[0906] (25αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 358;

[0907] (25αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 358;

[0908] (25αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 357;

[0909] (25αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 357;

[0910] (25αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 357;

[0911] (25% B2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 359-365, preferably a protein comprising at least 2, 3, 4, 5, or 6 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 359 to 365 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0912] The protein selected from the group consisting of the proteins as defined above in (25αA1-a) to (25αA1-c), (25αB1), (25αA2-a) to (25αA2-c), and (25αB2) can be proteins that are found in a protein spot with a molecular weight of around 20 to 90 kDa, preferably around 30 to 80 kDa, and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0913] Epitopes of the aforementioned antigens (25αA1-a) to (25αA1-c) and (25αB1) reside in amino acids 241-255 (SEQ ID NO: 814) and amino acids 321-335 (SEQ ID NO: 815) of SEQ ID NO: 348. When the antigen in spot 25α contains a variation in the amino acid sequence of SEQ ID NO: 348, this antigen may retain the sequence in SEQ ID NO: 348 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0914] Preferably, a protein that is an antigen in spot 25α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 348 or 358 (succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial) or causative of an allergy to soybean.

(26α) Antigen in Spot 26α

[0915] As the result of sequence identification of the antigen in spot 26α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 368-409 were detected.

[0916] Also, the mass spectroscopic data obtained for spot 26α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, phosphoglycerate kinase, cytosolic (amino acid sequence: SEQ ID NO: 367, encoding nucleotide sequence: SEQ ID NO: 366) was identified as a protein comprising any of SEQ ID NOs: 368 to 382, phosphoglycerate kinase, cytosolic (amino acid sequence: SEQ ID NO: 384, encoding nucleotide sequence: SEQ ID NO: 383) was identified as a protein comprising any of SEQ ID NOs: 385 to 392, and phosphoglycerate kinase, cytosolic (amino acid sequence: SEQ ID NO: 394, encoding nucleotide sequence: SEQ ID NO: 393) was identified as a protein comprising any of SEQ ID NOs: 395 to 409.

[0917] Accordingly, the antigen in spot 26α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (26αA1-a) to

(26αA1-c), (26αB1), (26αA2-a) to (26αA2-c), (26αB2), (26αA3-a) to (26αA3-c), and (26αB3):

[0918] (26αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 367;

[0919] (26αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 367;

[0920] (26αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 366;

[0921] (26αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 366;

[0922] (26αA1-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 366;

[0923] (26αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 368-382, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or 14 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 368 to 382 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[0924] (26αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 384;

[0925] (26αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 384;

[0926] (26αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 383;

[0927] (26αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 383;

[0928] (26αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 383;

[0929] (26αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 385-392, preferably a protein comprising at least 2, 3, 4, 5, 6, or 7 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 385 to 392 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0930] (26αA3-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 394;

[0931] (26αA3-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 394;

[0932] (26αA3-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 393;

[0933] (26αA3-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 393;

[0934] (26αA3-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 393;

[0935] (26αB3) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 395-409, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or 14 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 395 to 409 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0936] The protein selected from the group consisting of the proteins as defined above in (26αA1-a) to (26αA1-c), (26αB1), (26αA2-a) to (26αA2-c), (26αB2), (26αA3-a) to (26αA3-c), and (26αB3) can be proteins that are found in a protein spot with a molecular weight of around 20 to 90 kDa, preferably around 30 to 80 kDa, and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens". Epitopes of the aforementioned antigens (26αA1-a) to (26αA1-c) and (26αB1) reside in amino acids 121-135 (SEQ ID NO: 816) and amino acids 329-335 (SEQ ID NO: 819) of SEQ ID NO: 367. When the antigen in spot 26α contains a variation in the amino acid sequence of SEQ ID NO: 367, this antigen may retain the sequence in SEQ ID NO: 361 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0937] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 26α may be the following variant.

[0938] In the epitope having the sequence of SEQ ID NO: 819, sites of amino acids other than X in FDKXXXX or SEQ ID NO: 818 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 26α contains a variation in the amino acid sequence of SEQ ID NO: 367, amino acids 329-335 of SEQ

ID NO: 367 in this antigen may have the amino acid sequence of FDKXXXX or SEQ ID NO: 818.

[0939] Preferably, a protein that is an antigen in spot 25α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 367, 384, or 394 (phosphoglycerate kinase, cytosolic) or causative of an allergy to soybean.

(27α) Antigen in Spot 27α

[0940] As the result of sequence identification of the antigen in spot 27α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 412-427 were detected.

[0941] Also, the mass spectroscopic data obtained for spot 27a on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, alcohol dehydrogenase 1 (amino acid sequence: SEQ ID NO: 411, encoding nucleotide sequence: SEQ ID NO: 410) was identified as a protein comprising any of SEQ ID NOs: 412 to 416, alcohol dehydrogenase 1 (amino acid sequence: SEQ ID NO: 418, encoding nucleotide sequence: SEQ ID NO: 417) was identified as a protein comprising any of SEQ ID NOs: 419 and 420, and alcohol dehydrogenase 1-like (amino acid sequence: SEQ ID NO: 422, encoding nucleotide sequence: SEQ ID NO: 421) was identified as a protein comprising any of SEQ ID NOs: 423 to 427.

[0942] Accordingly, the antigen in spot 27 in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (27αA1-a) to (27αA1-e), (27αB1), (27αA2-a) to (27αA2-c), (27αB2), (27αA3-a) to (27αA3-c), and (27αB3): (27αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 411;

[0943] (27αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 411;

[0944] (27αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 410;

[0945] (27αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 410;

[0946] (27αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 410;

[0947] (27αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 412-416, preferably a protein comprising at least 2, 3, or 4 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 412 to 416 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[0948] (27αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 418;

[0949] (27αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 418;

[0950] (27αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 417;

[0951] (27αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 417;

[0952] (27αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 417;

[0953] (27αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 419 and 420, preferably a protein comprising both the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 419, 420 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0954] (27αA3-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 422;

[0955] (27αA3-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 422;

[0956] (27αA3-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 421; (27αA3-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 421;

[0957] (27αA3-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 421;

[0958] (27αB3) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 423-427, preferably a protein comprising at least 2, 3, or 4 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 423 to 427 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0959] The protein selected from the group consisting of the proteins as defined above in (27αA1-a) to (27αA1-c),

(27αB1), (27αA2-a) to (27αA2-c), (27αB2), (27αA3-a) to (27αA3-c), and (27αB3) can be proteins that are found in a protein spot with a molecular weight of around 20 to 90 kDa, preferably around 30 to 80 kDa, and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0960] Epitopes of the aforementioned antigens (27αA1-a) to (27αA1-e) and (27αB1) reside in amino acids 21-35 (SEQ ID NO: 820) of SEQ ID NO: 411. When the antigen in spot 27α contains a variation in the amino acid sequence of SEQ ID NO: 411, this antigen may retain the sequence in SEQ ID NO: 411 as to amino acids at positions corresponding to the region of this epitope.

[0961] Preferably, a protein that is an antigen in spot 27α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 411 or 418 (alcohol dehydrogenase 1) or the protein having the amino acid sequence of SEQ ID NO: 422 (alcohol dehydrogenase 1-like) or causative of an allergy to soybean.

(28α) Antigen in Spot 28α

[0962] As the result of sequence identification of the antigen in spot 28α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 430-433 were detected.

[0963] Also, the mass spectroscopic data obtained for spot 28α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, 35 kDa seed maturation protein isoform X1 (amino acid sequence: SEQ ID NO: 429, encoding nucleotide sequence: SEQ ID NO: 428) was identified as a protein comprising any of SEQ ID NO: 430 to 433.

[0964] Accordingly, the antigens in spot 28α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (28αA-a) to (28αA-e) and (28αB):

[0965] (28αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 429;

[0966] (28αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 429;

[0967] (28αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 428;

[0968] (28αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 428;

[0969] (28αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 428;

[0970] (28αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 430-433, preferably a protein comprising at least 2 or 3 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 430-433 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0971] The proteins selected from the group consisting of (28αA-a) to (28αA-c) and (28β) as defined above can be proteins that are found in a protein spot with a molecular weight of around 10 to 70 kDa, preferably around 20 to 60 kDa and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[0972] Epitopes of the antigen in spot 28α reside in amino acids 76-85 (SEQ ID NO: 820), amino acids 191-205 (SEQ ID NO: 826), amino acids 231-245 (SEQ ID NO: 829), amino acids 111-125 (SEQ ID NO: 830), amino acids 131-145 (SEQ ID NO: 831), amino acids 161-175 (SEQ ID NO: 832), amino acids 181-195 (SEQ ID NO: 833), amino acids 291-305 (SEQ ID NO: 834), amino acids 295-309 (SEQ ID NO: 835) of SEQ ID NO: 429. The antigen in spot 28a may retain the sequence in SEQ ID NO: 429 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[0973] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 28α may be the following variant.

[0974] In the epitope having the sequence of SEQ ID NO: 823, sites of amino acids other than X in SEQ ID NO: 821 or 822 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 76-85 of SEQ ID NO: 429 in the antigen in spot 28a may have the amino acid sequence of SEQ ID NO: 821 or 822.

[0975] In the epitope having the sequence of SEQ ID NO: 826, sites of amino acids other than X in XKXXXXEKXXXXXXX or SEQ ID NO: 825 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 191-205 of SEQ ID NO: 429 in the antigen in spot 28α may have the amino acid sequence of XKXXXXEKXXXXXXX or SEQ ID NO: 825.

[0976] In the epitope having the sequence of SEQ ID NO: 829, sites of amino acids other than X in SEQ ID NO: 827 or 828 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 231-2455 of SEQ ID NO: 429 in the antigen in spot 28α may have the amino acid sequence of SEQ ID NO: 827 or 828.

[0977] Preferably, a protein that is an antigen in spot 28α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 422 (35 kDa seed maturation protein isoform X1) or causative of an allergy to soybean.

(29α, 30α, 31α, 32α, 33α) Antigens in Spots 29α to 33α

[0978] As the result of sequence identification of the antigens in spots 29α to 33a by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 436-451 were detected.

[0979] Also, the mass spectroscopic data obtained for spots 29a to 33a on a mass spectrometer was analyzed by

comparing the data against the NCBI protein data, and as a result, glyceraldehyde-3-phosphate dehydrogenase isoform X1 (amino acid sequence: SEQ ID NO: 435, encoding nucleotide sequence: SEQ ID NO: 434) was identified as a protein comprising any of SEQ ID NO: 436 to 439, glyceraldehyde-3-phosphate dehydrogenase, cytosolic (amino acid sequence: SEQ ID NO: 441, encoding nucleotide sequence: SEQ ID NO: 440) was identified as a protein comprising any of SEQ ID NO: 442 to 445, and uncharacterized protein LOC100782924 (amino acid sequence: SEQ ID NO: 447, encoding nucleotide sequence: SEQ ID NO: 446) was identified as a protein comprising any of SEQ ID NO: 448 to 451.

[0980] Accordingly, the antigen in spot 29α to 33α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (29-33αA1-a) to (29-33αA1-c), (29-33αB1), (29-33αA2-a) to (29-33αA2-c), (29-33αB2), (29-33αA3-a) to (29-33αA3-c), and (29-33αB3):

[0981] (29-33αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 435;

[0982] (29-33αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 435;

[0983] (29-33αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 434;

[0984] (29-33αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 434;

[0985] (29-33αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 434;

[0986] (29-33αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 436-439, preferably a protein comprising at least 2, 3, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 436 to 439 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[0987] (29-33αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 441;

[0988] (29-33αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 441;

[0989] (29-33αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 440;

[0990] (29-33αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 440;

[0991] (29-33αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 440;

[0992] (29-33αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 442-445, preferably a protein comprising at least 2, 3, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 442 to 445 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0993] (29-33αA3-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 447;

[0994] (29-33αA3-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 447;

[0995] (29-33αA3-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 446;

[0996] (29-33αA3-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 446;

[0997] (29-33αA3-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 446;

[0998] (29-33αB3) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 448-451, preferably a protein comprising at least 2, 3, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 448 to 451 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[0999] The protein selected from the group consisting of the proteins as defined above in (29-33αA1-a) to (29-33αA1-c), (29-33αB1), (29-33αA2-a) to (29-33αA2-c), (29-33αB2), (29-33αA3-a) to (29-33αA3-c), and (29-33αB3) can be proteins that are found in a protein spot with a molecular weight of around 10 to 70 kDa, preferably around 20 to 60 kDa, and an isoelectric point of 3.5 to 8.5, preferably 4.0 to 8.0 on gels used in the two-dimensional

electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1000] Epitopes of the aforementioned antigens (29-33αA1-a) to (29-33αA1-c) and (29-33αB1) reside in amino acids 66-75 (SEQ ID NO: 838), amino acids 301-315 (SEQ ID NO: 839) of SEQ ID NO: 435. When the antigen in spot 29-33α contains a variation in the amino acid sequence of SEQ ID NO: 435, this antigen may retain the sequence in SEQ ID NO: 435 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[1001] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 29-33α may be the following variant.

[1002] In the epitope having the sequence of SEQ ID NO: 838, sites of amino acids other than X in XXXXXXXDKP or SEQ ID NO: 837 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 29-33α contains a variation in the amino acid sequence of SEQ ID NO: 435, amino acids 66-75 of SEQ ID NO: 435 in this antigen may have the amino acid sequence of XXXXXXXDKP or SEQ ID NO: 837.

[1003] Preferably, a protein that is an antigen in any of spots 29α to 33α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 435 (glyceraldehyde-3-phosphate dehydrogenase isoform X1), the protein having the amino acid sequence of SEQ ID NO: 441 (glyceraldehyde-3-phosphate dehydrogenase, cytosolic), or the protein having the amino acid sequence of SEQ ID NO: 447 (uncharacterized protein LOC100782924) or causative of an allergy to soybean.

(34α) Antigen in Spot 34α

[1004] As the result of sequence identification of the antigen in spot 34α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 454-458 were detected.

[1005] Also, the mass spectroscopic data obtained for spot 34α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, bifunctional UDP-glucose 4-epimerase and UDP-xylose 4-epimerase 1-like (amino acid sequence: SEQ ID NO: 453, encoding nucleotide sequence: SEQ ID NO: 452) was identified as a protein comprising SEQ ID NO: 454 and uncharacterized protein LOC100803483 (amino acid sequence: SEQ ID NO: 456, encoding nucleotide sequence: SEQ ID NO: 455) was identified as a protein comprising SEQ ID NO: 457 or 458.

[1006] Accordingly, the antigen in spot 34α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (34αA1-a) to (34αA1-c), (34αB1), (34αA2-a) to (34αA2-c), and (34αB2):

[1007] (34αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 453;

[1008] (34αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 453;

[1009] (34αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 452;

[1010] (34αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 452;

[1011] (34αA1-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 452;

[1012] (34αB1) a protein comprising an amino acid sequence of SEQ ID NOs: 454. As referred to above, the amino acid sequence of any of SEQ ID NOs: 464 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[1013] (34αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 456;

[1014] (34αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 456;

[1015] (34αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 455;

[1016] (34αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 455;

[1017] (34αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 455;

[1018] (34αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 457 and 458, preferably a protein comprising both the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 457, 458 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1019] The protein selected from the group consisting of the proteins as defined above in (34αA1-a) to (34αA1-c), (34αB1), (34αA2-a) to (34αA2-c), and (34αB2) can be proteins that are found in a protein spot with a molecular weight of around 10 to 70 kDa, preferably around 20 to 60 kDa, and an isoelectric point of 5.5 to 10.5, preferably 6.0 to 10.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1020] Epitopes of the aforementioned antigens (34αA1-a) to (34αA1-c) and (34αB1) reside in amino acids 61-75 (SEQ ID NO: 842), amino acids 31-45 (SEQ ID NO: 843), amino acids 141-155 (SEQ ID NO: 844), amino acids 221-235 (SEQ ID NO: 845), amino acids 231-245 (SEQ ID

NO: 846) of SEQ ID NO: 454. When the antigen in spot 34α contains a variation in the amino acid sequence of SEQ ID NO: 454, this antigen may retain the sequence in SEQ ID NO: 454 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[1021] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 34α may be the following variant.

[1022] In the epitope having the sequence of SEQ ID NO: 842, sites of amino acids other than X in XXXXXXXXXXXXDDXE or SEQ ID NO: 841 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 34α contains a variation in the amino acid sequence of SEQ ID NO: 454, amino acids 61-75 of SEQ ID NO: 454 in this antigen may have the amino acid sequence of XXXXXXXXXXXDDXE or SEQ ID NO: 841.

[1023] Preferably, a protein that is an antigen in spot 34α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 453 (bifunctional UDP-glucose 4-epimerase and UDP-xylose 4-epimerase 1-like) or the protein having the amino acid sequence of SEQ ID NO: 456 (uncharacterized protein LOC100803483) or causative of an allergy to soybean.

(35α) Antigen in Spot 35α

[1024] As the result of sequence identification of the antigen in spot 35α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 461-477 were detected.

[1025] Also, the mass spectroscopic data obtained for spot 35α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, elongation factor 1-alpha (amino acid sequence: SEQ ID NO: 460, encoding nucleotide sequence: SEQ ID NO: 459) was identified as a protein comprising any of SEQ ID NOs: 461 to 466, elongation factor 1-alpha (amino acid sequence: SEQ ID NO: 468, encoding nucleotide sequence: SEQ ID NO: 467) was identified as a protein comprising any of SEQ ID NOs: 469 to 471, and elongation factor-1A isoform X1 (amino acid sequence: SEQ ID NO: 473, encoding nucleotide sequence: SEQ ID NO: 472) was identified as a protein comprising any of SEQ ID NOs: 474 to 477.

[1026] Accordingly, the antigen in spot 35α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (35αA1-a) to (35QAl-c), (35αB1), (35αA2-a) to (35αA2-c), (35αB2), (35αA3-a) to (35αA3-c), and (35αB3): (35αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 460;

[1027] (35αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 460;

[1028] (35αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 459;

[1029] (35αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 459;

[1030] (35αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 459;

[1031] (35αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 461-466, preferably a protein comprising at least 2, 3, 4, or 5 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 412 to 416 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[1032] (35αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 468;

[1033] (35αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 468;

[1034] (35αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 467;

[1035] (35αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 467;

[1036] (35αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 467;

[1037] (35αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 469-471, preferably a protein comprising at least 2, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 469 to 471 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1038] (35αA3-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 473;

[1039] (35αA3-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 473;

[1040] (35αA3-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 472;

[1041] (35αA3-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least

85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 472;

[1042] (35αA3-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 472;

[1043] (35αB3) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 474-477, preferably a protein comprising at least 2, 3, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 474 to 477 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1044] The protein selected from the group consisting of the proteins as defined above in (35αA1-a) to (35αA1-c), (35αB1), (35αA2-a) to (35αA2-c), (35αB2), (35αA3-a) to (35αA3-c), and (35αB3) can be proteins that are found in a protein spot with a molecular weight of around 20 to 90 kDa, preferably around 30 to 80 kDa, and an isoelectric point of 5.5 to 10.5, preferably 6.0 to 10.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1045] Epitopes of the aforementioned antigens (35αA1-a) to (35αA1-e) and (35αB1) reside in amino acids 31-45 (SEQ ID NO: 849), amino acids 121-135 (SEQ ID NO: 850), amino acids 192-220 (SEQ ID NO: 853), amino acids 216-225 (SEQ ID NO: 856) of SEQ ID NO: 460. When the antigen in spot 35α contains a variation in the amino acid sequence of SEQ ID NO: 460, amino acids at positions corresponding to the region of at least one of these epitopes may retain the sequence in SEQ ID NO: 460.

[1046] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 35α may be the following variant.

[1047] In the epitope having the sequence of SEQ ID NO: 849, sites of amino acids other than X in SEQ ID NO: 847 or 848 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 35α contains a variation in the amino acid sequence of SEQ ID NO: 460, amino acids 31-45 of SEQ ID NO: 460 in this antigen may have the amino acid sequence of SEQ ID NO: 847 or 848.

[1048] In the epitope having the sequence of SEQ ID NO: 853, sites of amino acids other than X in SEQ ID NO: 851 or 852 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 35α contains a variation in the amino acid sequence of SEQ ID NO: 460, amino acids 192-220 of SEQ ID NO: 460 in this antigen may have the amino acid sequence of SEQ ID NO: 851 or 852.

[1049] In the epitope having the sequence of SEQ ID NO: 856, sites of amino acids other than X in XXEXLDXXXX or SEQ ID NO: 855 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 35α contains a variation in the amino acid sequence of SEQ ID NO: 460, amino acids 216-225 of SEQ ID NO: 460 in this antigen may have the amino acid sequence of XXEXLDXXXX or SEQ ID NO: 855.

[1050] Preferably, a protein that is an antigen in spot 35α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 460 or 468 (elongation factor 1-alpha) or the protein having the amino acid sequence of SEQ ID NO: 473 (elongation factor-1A isoform X1) or causative of an allergy to soybean.

(36α, 37α) Antigens in Spots 36α, 37α

[1051] As the result of sequence identification of the antigens in spots 36α and 37α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 480-493 were detected.

[1052] Also, the mass spectroscopic data obtained for spots 36α and 37a on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, seed maturation protein PM34 (amino acid sequence: SEQ ID NO: 479, encoding nucleotide sequence: SEQ ID NO: 478) was identified as a protein comprising any of SEQ ID NOs: 480 to 483, uncharacterized protein LOC100809384 (amino acid sequence: SEQ ID NO: 485, encoding nucleotide sequence: SEQ ID NO: 484) was identified as a protein comprising SEQ ID NO: 486 or 487, and hypothetical protein GLYMA_10G155300 (amino acid sequence: SEQ ID NO: 489, encoding nucleotide sequence: SEQ ID NO: 488) was identified as a protein comprising any of SEQ ID NOs: 4490 to 493.

[1053] Accordingly, the antigen in spot 36α, 37a in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (36, 37αA1-a) to (36, 37αA1-e), (36, 37αB1), (36, 37αA2-a) to (36, 37αA2-e), (36, 37αB2), (36, 37αA3-a) to (36, 37αA3-e), and (36, 37αB3):

[1054] (36, 37αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 479;

[1055] (36, 37αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 479;

[1056] (36, 37αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 478;

[1057] (36, 37αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 478;

[1058] (36, 37αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 478;

[1059] (36, 37αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 480-483, preferably a protein comprising at least 2, 3, or all sequences of the amino acid sequences. As referred to above, the amino acid

sequence of any of SEQ ID NOs: 480 to 483 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[1060] (36, 37αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 485;

[1061] (36, 37αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 485;

[1062] (36, 37αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 484;

[1063] (36, 37αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 484;

[1064] (36, 37αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 484;

[1065] (36, 37αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 486 and 487, preferably a protein comprising both the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 486, 487 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1066] (36, 37αA3-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 489;

[1067] (36, 37αA3-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 489;

[1068] (36, 37αA3-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 488;

[1069] (36, 37αA3-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 488;

[1070] (36, 37αA3-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 488;

[1071] (36, 37αB3) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 490-493, preferably a protein comprising at least 2, 3, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 490 to 493 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1072] The protein selected from the group consisting of the proteins as defined above in (36, 37αA1-a) to (36, 37αA1-c), (36, 37αB1), (36, 37αA2-a) to (36, 37αA2-c), (36, 37αB2), (36, 37αA3-a) to (36, 37αA3-c), and (36, 37αB3) can be proteins that are found in a protein spot with a molecular weight of around 10 to 70 kDa, preferably around 20 to 60 kDa, and an isoelectric point of 3.5 to 8.5, preferably 4.0 to 8.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1073] Epitopes of the aforementioned antigens (36, 37αA1-a) to (36, 37αA1-c) and (36, 37αB1) reside in amino acids 141-155 (SEQ ID NO: 857), amino acids 231-245 (SEQ ID NO: 858), amino acids 7-14 (SEQ ID NO: 861) of SEQ ID NO: 479. When the antigen in spot 36α, 37a contains a variation in the amino acid sequence of SEQ ID NO: 479, this antigen may retain the sequence in SEQ ID NO: 479 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[1074] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 36α,37a may be the following variant.

[1075] In the epitope having the sequence of SEQ ID NO: 861, sites of amino acids other than X in SEQ ID NO: 859 or 860 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 36α,37a contains a variation in the amino acid sequence of SEQ ID NO: 479, amino acids 7-14 of SEQ ID NO: 479 in this antigen may have the amino acid sequence of SEQ ID NO: 859 or 860.

[1076] Preferably, a protein that is an antigen in spot 36α or 37a has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 479 (seed maturation protein PM34), the protein having the amino acid sequence of SEQ ID NO: 485 (uncharacterized protein LOC100809384), or the protein having the amino acid sequence of SEQ ID NO: 489 (hypothetical protein GLYMA_10G155300) or causative of an allergy to soybean.

(38) Antigen in Spot 38α

[1077] As the result of sequence identification of the antigen in spot 38α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 496-513 were detected.

[1078] Also, the mass spectroscopic data obtained for spot 38α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, seed biotin-containing protein SBP65-like isoform X1 (amino acid sequence: SEQ ID NO: 495, encoding nucleotide sequence: SEQ ID NO: 494) was identified as a protein comprising any of SEQ ID NOs: 496 to 513.

[1079] Accordingly, the antigens in spot 38α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (38αA-a) to (38αA-c) and (38«B):

[1080] (38αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 495;

[1081] (38αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 495;

[1082] (38αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 494;

[1083] (38αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 494;

[1084] (38αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 494;

[1085] (38αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 496-513, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 or 17 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 496-513 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1086] The proteins selected from the group consisting of (38αA-a) to (38αA-c) and (38αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 70 to 170 kDa, preferably around 80 to 160 kDa and an isoelectric point of 5.5 to 10.5, preferably 6.0 to 10.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1087] Epitopes of the antigen in spot 38α reside in amino acids 440-446 or 574-580 (SEQ ID NO: 864), amino acids 71-80 (SEQ ID NO: 867), amino acids 521-535 (SEQ ID NO: 869), amino acids 1-15 (SEQ ID NO: 871), amino acids 31-45 (SEQ ID NO: 874), amino acids 221-235 (SEQ ID NO: 877), and amino acids 461-470 (SEQ ID NO: 882), and amino acids 481-490 (SEQ ID NO: 884), and amino acids 821-835 (SEQ ID NO: 887), and amino acids 861-870 (SEQ ID NO: 889), and amino acids 1031-1040 (SEQ ID NO: 892) of SEQ ID NO: 495. The antigen in spot 38α may retain the sequence in SEQ ID NO: 495 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[1088] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 38α may be the following variant.

[1089] In the epitope having the sequence of SEQ ID NO: 864, sites of amino acids other than X in RGKXXXX are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 440-446 and/or 574-580 of SEQ ID NO: 495 in the antigen in spot 38α may have the amino acid sequence of RGKXXXX.

[1090] In the epitope having the sequence of SEQ ID NO: 867, sites of amino acids other than X in XRGKXXXXXX or SEQ ID NO: 866 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 71-80 of SEQ ID NO: 495 in the antigen in spot 38α may have the amino acid sequence of XRGKXXXXXX or SEQ ID NO: 866.

[1091] In the epitope having the sequence of SEQ ID NO: 869, sites of amino acids other than X in RXKXXXXXXXXXXXE are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 521-535 of SEQ ID NO: 495 in the antigen in spot 38α may have the amino acid sequence of RXKXXXXXXXXXXXE.

[1092] In the epitope having the sequence of SEQ ID NO: 871, sites of amino acids other than X in XXXXXXXRRXXXXXR are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 1-15 of SEQ ID NO: 495 in the antigen in spot 38α may have the amino acid sequence of XXXXXXXRRXXXXXR.

[1093] In the epitope having the sequence of SEQ ID NO: 874, sites of amino acids other than X in SEQ ID NO: 872 or 873 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 31-45 of SEQ ID NO: 495 in the antigen in spot 38α may have the amino acid sequence of SEQ ID NO: 872 or 873.

[1094] In the epitope having the sequence of SEQ ID NO: 877, sites of amino acids other than X in SEQ ID NO: 875 or 876 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 221-235 of SEQ ID NO: 495 in the antigen in spot 38α may have the amino acid sequence of SEQ ID NO: 875 or 876.

[1095] In the epitope having the sequence of SEQ ID NO: 880, sites of amino acids other than X in SEQ ID NO: 878 or 879 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 311-325 of SEQ ID NO: 495 in the antigen in spot 38α may have the amino acid sequence of SEQ ID NO: 878 or 879.

[1096] In the epitope having the sequence of SEQ ID NO: 882, sites of amino acids other than X in SEQ ID NO: 881 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 461-470 of SEQ ID NO: 495 in the antigen in spot 38α may have the amino acid sequence of SEQ ID NO: 881.

[1097] In the epitope having the sequence of SEQ ID NO: 884, sites of amino acids other than X in SEQ ID NO: 883 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 481-490 of SEQ ID NO: 495 in the antigen in spot 38 may have the amino acid sequence of SEQ ID NO: 883.

[1098] In the epitope having the sequence of SEQ ID NO: 887, sites of amino acids other than X in SEQ ID NO: 885 or 886 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 821-835 of SEQ ID NO: 495 in the antigen in spot 38α may have the amino acid sequence of SEQ ID NO: 885 or 886.

[1099] In the epitope having the sequence of SEQ ID NO: 889, sites of amino acids other than X in SEQ ID NO: 888 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids

861-870 of SEQ ID NO: 495 in the antigen in spot 38α may have the amino acid sequence of SEQ ID NO: 888.

[1100] In the epitope having the sequence of SEQ ID NO: 892, sites of amino acids other than X in SEQ ID NO: 890 or 891 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 1031-1040 of SEQ ID NO: 495 in the antigen in spot 38α may have the amino acid sequence of SEQ ID NO: 890 or 891.

[1101] Preferably, a protein that is an antigen in spot 38α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 495 (seed biotin-containing protein SBP65-like isoform X1) or causative of an allergy to soybean.

(39α) Antigen in Spot 39α

[1102] As the result of sequence identification of the antigen in spot 39α by mass spectroscopy, the amino acid sequence of SEQ ID NO: 516 was detected.

[1103] Also, the mass spectroscopic data obtained for spot 39α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, ATP synthase subunit O, mitochondrial-like (amino acid sequence: SEQ ID NO: 515, encoding nucleotide sequence: SEQ ID NO: 514) was identified as a protein comprising SEQ ID NO: 516.

[1104] Accordingly, the antigens in spot 39α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (39αA-a) to (39αA-c) and (39β):

- **[1105]** (39αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 515;
- **[1106]** (39αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 515;
- **[1107]** (39αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 514;
- **[1108]** (39αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 514;
- **[1109]** (39αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 514;
- **[1110]** (39αB) a protein comprising an amino acid sequence of SEQ ID NO: 516. As referred to above, the amino acid sequence(s) of SEQ ID NO: 516 may be an amino acid sequence(s) derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1111] The proteins selected from the group consisting of (39αA-a) to (39αA-c) and (39β) as defined above can be proteins that are found in a protein spot with a molecular weight of around 5 to 60 kDa, preferably around 10 to 50 kDa and an isoelectric point of 5.5 to 10.5, preferably 6.0 to 10.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1112] An epitope of the antigen in spot 39α resides in amino acids 71-85 (SEQ ID NO: 895) of SEQ ID NO: 515. The antigen in spot 39α may retain the sequence in SEQ ID NO: 515 as to amino acids at positions corresponding to the region of this epitope.

[1113] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 39α may be the following variant.

[1114] In the epitope having the sequence of SEQ ID NO: 895, sites of amino acids other than X in SEQ ID NO: 893 or 894 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 71-85 of SEQ ID NO: 515 in the antigen in spot 39a may have the amino acid sequence of SEQ ID NO: 893 or 894.

[1115] Preferably, a protein that is an antigen in spot 39α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 515 (ATP synthase subunit O, mitochondrial-like) or causative of an allergy to soybean.

(40α, 56α) Antigens in Spots 40α, 56α

[1116] As the result of sequence identification of the antigens in spots 40α and 56α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 519 and 520 were detected.

[1117] Also, the mass spectroscopic data obtained for spots 40α and 56α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, P24 oleosin isoform A (amino acid sequence: SEQ ID NO: 518, encoding nucleotide sequence: SEQ ID NO: 517) was identified as a protein comprising SEQ ID NO: 519 or 520.

[1118] Accordingly, the antigens in spot 40α, 56α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (40, 56αA-a) to (40, 56αA-c) and (40, 56αB):

- **[1119]** (40, 56αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 518;
- **[1120]** (40, 56αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 518;
- **[1121]** (40, 56αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 517;
- **[1122]** (40, 56αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 517;
- **[1123]** (40, 56αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a

nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 517;

[1124] (40, 56αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 519 and 520, preferably a protein comprising both the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 519 and 520 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1125] The proteins selected from the group consisting of (40, 56αA-a) to (40, 56αA-c) and (40, 56αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 5 to 60 kDa, preferably around 10 to 50 kDa and an isoelectric point of 5.5 to 10.5, preferably 6.0 to 10.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1126] Epitopes of the antigens in spots 40α and 56α reside in amino acids 35-39 (SEQ ID NO: 898), amino acids 191-205 (SEQ ID NO: 899) and amino acids 214-222 (SEQ ID NO: 902) of SEQ ID NO: 518. The antigens in spots 40α and 56α may retain the sequence in SEQ ID NO: 518 as to amino acids at positions corresponding to the region of these epitopes.

[1127] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 40α, 56a may be the following variant.

[1128] In the epitope having the sequence of SEQ ID NO: 898, sites of amino acids other than X in SEQ ID NO: 896 or 897 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 35-39 of SEQ ID NO: 518 in the antigen in spot 40α, 56a may have the amino acid sequence of SEQ ID NO: 896 or 897.

[1129] In the epitope having the sequence of SEQ ID NO: 902, sites of amino acids other than X in XXXXXVTAX or SEQ ID NO: 901 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 214-2229 of SEQ ID NO: 518 in the antigen in spot 40α, 56α may have the amino acid sequence of XXXXXVTAX or SEQ ID NO: 901.

[1130] Preferably, a protein that is an antigen in spot 40α or 56α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 518 (P24 oleosin isoform A) or causative of an allergy to soybean.

(41α) Antigen in Spot 41α

[1131] As the result of sequence identification of the antigen in spot 41α by mass spectroscopy, the amino acid sequence of SEQ ID NO: 523 was detected.

[1132] Also, the mass spectroscopic data obtained for spot 41a on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, uncharacterized protein At3g49720-like isoform X2 (amino acid sequence: SEQ ID NO: 522, encoding nucleotide sequence: SEQ ID NO: 521) was identified as a protein comprising SEQ ID NO: 523.

[1133] Accordingly, the antigens in spot 41a in the present invention can be any protein selected from the group consisting of the proteins as defined below in (41αA-a) to (41αA-c) and (41αB):

[1134] (4*l* αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 522;

[1135] (41αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 522;

[1136] (41αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 521;

[1137] (41αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 521;

[1138] (41αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 521;

[1139] (41αB) a protein comprising an amino acid sequence of SEQ ID NO: 523. As referred to above, the amino acid sequence of SEQ ID NO: 523 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1140] The proteins selected from the group consisting of (41αA-a) to (41αA-c) and (41αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 5 to 60 kDa, preferably around 10 to 50 kDa and an isoelectric point of 5.5 to 10.5, preferably 6.0 to 10.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1141] An epitope of the antigen in spot 41α resides in amino acids 111-125 (SEQ ID NO: 903) of SEQ ID NO: 522. The antigen in spot 41α may retain the sequence in SEQ ID NO: 522 as to amino acids at positions corresponding to the region of this epitope.

[1142] Preferably, a protein that is an antigen in spot 41a has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 515 (ATP synthase subunit O, mitochondrial-like) or causative of an allergy to soybean.

(42α, 43α) Antigen in Spots 42α, 43α

[1143] As the result of sequence identification of the antigens in spots 42α and 43α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 526-539 were detected.

[1144] Also, the mass spectroscopic data obtained for spots 42α and 43α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, superoxide dismutase [Mn], mitochondrial (amino acid sequence: SEQ ID NO: 525, encoding nucleotide sequence: SEQ ID NO: 524) was identified as a protein comprising any of SEQ ID NOs: 526 to 531 and hypothetical protein GLYMA_04G221300 (amino acid sequence: SEQ ID NO: 533, encoding nucleotide sequence: SEQ ID NO: 532) was identified as a protein comprising any of SEQ ID NOs: 534 to 539.

[1145] Accordingly, the antigen in spot 42α, 43α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (42, 43αA1-a) to (42, 43αA1-c), (42, 43αB1), (42, 43αA2-a) to (42, 43A2-c), and (42, 43αB2):

[1146] (42, 43αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 525;

[1147] (42, 43αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 525;

[1148] (42, 43αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 524;

[1149] (42, 43αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 524;

[1150] (42, 43αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 524;

[1151] (42, 43β1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 526-531, preferably a protein comprising at least 2, 3, 4, or 5 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 526 to 531 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[1152] (42, 43αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 533;

[1153] (42, 43αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 533;

[1154] (42, 43αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 532;

[1155] (42, 43αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 532;

[1156] (42, 43αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 532;

[1157] (42, 43αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 534-539, preferably a protein comprising at least 2, 3, 4, or 5 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 534 to 539 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1158] The protein selected from the group consisting of the proteins as defined above in (42, 43αA1-a) to (42, 43αA1-c), (42, 43αB1), (42, 43αA2-a) to (42, 43αA2-c), and (42, 43αB2) can be proteins that are found in a protein spot with a molecular weight of around 5 to 60 kDa, preferably around 10 to 50 kDa, and an isoelectric point of 5.5 to 10.5, preferably 6.0 to 10.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1159] Epitopes of the aforementioned antigens (42, 43αA1-a) to (42, 43αA1-c) and (42, 43αB1) reside in amino acids 227-241 (SEQ ID NO: 906) of SEQ ID NO: 525. When the antigen in spot 42α, 43a contains a variation in the amino acid sequence of SEQ ID NO: 525, this antigen may retain the sequence in SEQ ID NO: 525 as to amino acids at positions corresponding to the region of at least one of these epitopes.

[1160] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 42α, 43 may be the following variant.

[1161] In the epitope having the sequence of SEQ ID NO: 906, sites of amino acids other than X in SEQ ID NO: 904 or 905 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 42α, 43a contains a variation in the amino acid sequence of SEQ ID NO: 525, amino acids 227-241 of SEQ ID NO: 525 in this antigen may have the amino acid sequence of SEQ ID NO: 904 or 905.

[1162] Preferably, a protein that is an antigen in spot 42α or 43α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 525 (superoxide dismutase [Mn], mitochondrial) or the protein having the amino acid sequence of SEQ ID NO: 533 (hypothetical protein GLYMA_04G221300) or causative of an allergy to soybean.

(45α) Antigen in Spot 45α

[1163] As the result of sequence identification of the antigen in spot 45α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 542-551 were detected.

[1164] Also, the mass spectroscopic data obtained for spot 45α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, uncharacterized protein LOC100499771 (amino acid sequence: SEQ ID NO: 541, encoding nucleotide sequence: SEQ ID NO: 540) was identified as a protein comprising any of SEQ ID NOs: 542 to 545 and hypothetical protein GLYMA_09G192800 (amino acid sequence: SEQ ID NO: 547, encoding nucleotide sequence: SEQ ID NO: 546) was identified as a protein comprising any of SEQ ID NOs: 548 to 551.

[1165] Accordingly, the antigen in spot 45α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (45αA1-a) to (45αA1-c), (45αB1), (45αA2-a) to (45αA2c), and (45αB2):

[1166] (45αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 541;

[1167] (45αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 541;

[1168] (45αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 540;

[1169] (45αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 540;

[1170] (45αA1-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 540;

[1171] (45αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 542-545, preferably a protein comprising at least 2, 3, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 542 to 545 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[1172] (45αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 547;

[1173] (45αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 547;

[1174] (45αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 546;

[1175] (45αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 546;

[1176] (45αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 546;

[1177] (45β2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 548-551, preferably a protein comprising at least 2, 3, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 548 to 551 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1178] The protein selected from the group consisting of the proteins as defined above in (45αA1-a) to (45αA1-c), (45αB1), (45αA2-a) to (45αA2-c), and (45αB2) can be proteins that are found in a protein spot with a molecular weight of around 2 to 50 kDa, preferably around 5 to 40 kDa, and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1179] Epitopes of the aforementioned antigens (45αA1-a) to (45αA1-c) and (45αB1) reside in amino acids 11-25 (SEQ ID NO: 910), amino acids 51-65 (SEQ ID NO: 911), amino acids 111-125 (SEQ ID NO: 912) of SEQ ID NO: 541. When the antigen in spot 45a contains a variation in the amino acid sequence of SEQ ID NO: 541, amino acids at positions corresponding to the region of at least one of these epitopes may retain the sequence in SEQ ID NO: 541.

[1180] Preferably, a protein that is an antigen in spot 45α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 541 (uncharacterized protein LOC100499771) or the protein having the amino acid sequence of SEQ ID NO: 547 (hypothetical protein GLYMA_09G192800) or causative of an allergy to soybean.

(46α) Antigen in Spot 46α

[1181] As the result of sequence identification of the antigen in spot 46α by mass spectroscopy, the amino acid sequence of SEQ ID NOs: 554 was detected.

[1182] Also, the mass spectroscopic data obtained for spot 46α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, seed maturation protein PM22 (amino acid sequence: SEQ ID NO: 553, encoding nucleotide sequence: SEQ ID NO: 552) was identified as a protein comprising SEQ ID NO: 554.

[1183] Accordingly, the antigens in spot 46α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (46αA-a) to (46αA-c) and (46αB):

[1184] (46αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 553;

[1185] (46αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 553;

[1186] (46αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 552;

[1187] (46αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 552;

[1188] (46αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 552;

[1189] (46αB) a protein comprising an amino acid sequence of SEQ ID NO: 554. As referred to.: above, the amino acid sequence(s) of SEQ ID NO: 554 may be

an amino acid sequence(s) derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1190] The proteins selected from the group consisting of (46αA-a) to (46αA-c) and (46αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 2 to 50 kDa, preferably around 5 to 40 kDa and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1191] An epitope of the antigen in spot 46α resides in amino acids 112-120 (SEQ ID NO: 915) of SEQ ID NO: 553. The antigen in spot 46α may retain the sequence in SEQ ID NO: 553 as to amino acids at positions corresponding to the region of this epitope.

[1192] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 46α may be the following variant.

[1193] In the epitope having the sequence of SEQ ID NO: 915, sites of amino acids other than X in XDXDXXXDX or SEQ ID NO: 914 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 112-120 of SEQ ID NO: 553 in the antigen in spot 46α may have the amino acid sequence of XDXDXXXDX or SEQ ID NO: 914.

[1194] Preferably, a protein that is an antigen in spot 46α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 553 (seed maturation protein PM22) or causative of an allergy to soybean.

(47α) Antigen in Spot 47α

[1195] As the result of sequence identification of the antigen in spot 47α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 557-563 were detected.

[1196] Also, the mass spectroscopic data obtained for spot 47a on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, eukaryotic translation initiation factor 5A-2 (amino acid sequence: SEQ ID NO: 556, encoding nucleotide sequence: SEQ ID NO: 555) was identified as a protein comprising SEQ ID NO: 557 or 558 and uncharacterized protein LOC100305510 (amino acid sequence: SEQ ID NO: 560, encoding nucleotide sequence: SEQ ID NO: 559) was identified as a protein comprising any of SEQ ID NOs: 561 to 563.

[1197] Accordingly, the antigen in spot 47α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (47αA1-a) to (47αA1-e), (47αB1), (47αA2-a) to (47αA2-c), and (47αB2):

[1198] (47αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 556;

[1199] (47αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 556;

[1200] (47αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 555;

[1201] (47αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 555;

[1202] (47αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 555;

[1203] (47αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 557 and 558, preferably a protein comprising both the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 557, 558 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[1204] (47αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 560;

[1205] (47αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 560;

[1206] (47αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 559;

[1207] (47αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 559;

[1208] (47αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 559;

[1209] (47αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 561-563, preferably a protein comprising at least 2, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 548 to 551 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1210] The protein selected from the group consisting of the proteins as defined above in (47αA1-a) to (47αA1-c), (47αB1), (47αA2-a) to (47αA2-c), and (47αB2) can be proteins that are found in a protein spot with a molecular weight of around 2 to 50 kDa, preferably around 5 to 40 kDa, and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1211] Epitopes of the aforementioned antigens (47αA1-a) to (47αA1-e) and (47αB1) reside in amino acids 11-25 (SEQ ID NO: 916), amino acids 81-95 (SEQ ID NO: 917),

amino acids 131-145 (SEQ ID NO: 918) of SEQ ID NO: 556. When the antigen in spot 47a contains a variation in the amino acid sequence of SEQ ID NO: 556, amino acids at positions corresponding to the region of at least one of these epitopes may retain the sequence in SEQ ID NO: 556.

[1212] Preferably, a protein that is an antigen in spot 47α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 556 (eukaryotic translation initiation factor 5A-2) or the protein having the amino acid sequence of SEQ ID NO: 560 (uncharacterized protein LOC100305510) or causative of an allergy to soybean.

(48α) Antigen in Spot 48α

[1213] As the result of sequence identification of the antigen in spot 48 by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 566-570 were detected.

[1214] Also, the mass spectroscopic data obtained for spot 48α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, uncharacterized protein LOC100306570 (amino acid sequence: SEQ ID NO: 565, encoding nucleotide sequence: SEQ ID NO: 564) was identified as a protein comprising any of SEQ ID NOs: 566 to 570.

[1215] Accordingly, the antigens in spot 48α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (48αA-a) to (48αA-c) and (48«B):

- **[1216]** (48αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 565;
- **[1217]** (48αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 565;
- **[1218]** (48αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 564;
- **[1219]** (48αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 564;
- **[1220]** (48αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 564;
- **[1221]** (48αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 566-570, preferably a protein comprising at least 2, 3 or 4 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 566-570 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1222] The proteins selected from the group consisting of (48αA-a) to (48αA-c) and (48&B) as defined above can be proteins that are found in a protein spot with a molecular weight of around 2 to 50 kDa, preferably around 5 to 40 kDa and an isoelectric point of 3.5 to 8.5, preferably 4.0 to 8.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1223] Epitopes of the antigen in spot 48α reside in amino acids 94-102 (SEQ ID NO: 921) and amino acids 131-145 (SEQ ID NO: 922) of SEQ ID NO: 565. The antigen in spot 48α may retain the sequence in SEQ ID NO: 565 as to amino acids at positions corresponding to the region of these epitopes.

[1224] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 48α may be the following variant.

[1225] In the epitope having the sequence of SEQ ID NO: 921, sites of amino acids other than X in XXXXXXFDK or SEQ ID NO: 918 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 94-102 of SEQ ID NO: 565 in the antigen in spot 48α may have the amino acid sequence of XXXXXXFDK or SEQ ID NO: 918.

[1226] Preferably, a protein that is an antigen in spot 48α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 565 (uncharacterized protein LOC100306570) or causative of an allergy to soybean.

(50α) Antigen in Spot 50α

[1227] As the result of sequence identification of the antigen in spot 50α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 573-576 were detected.

[1228] Also, the mass spectroscopic data obtained for spot 50α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, uncharacterized protein LOC100813859 (amino acid sequence: SEQ ID NO: 572, encoding nucleotide sequence: SEQ ID NO: 571) was identified as a protein comprising any of SEQ ID NOs: 573 to 576.

[1229] Accordingly, the antigens in spot 50α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (50αA-a) to (50αA-e) and (50«B):

- **[1230]** (50αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 572;
- **[1231]** (50αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 572;
- **[1232]** (50αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 571;
- **[1233]** (50αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 571;
- **[1234]** (50αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a

nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 571;

[1235] (50αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 573-576, preferably a protein comprising at least 2 or 3 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 573-576 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1236] The proteins selected from the group consisting of (50αA-a) to (50αA-c) and (50αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 2 to 50 kDa, preferably around 5 to 40 kDa and an isoelectric point of 2.5 to 6.5, preferably 3.0 to 6.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1237] Epitopes of the antigen in spot 50α reside in amino acids 91-105 (SEQ ID NO: 924) and amino acids 62-70 (SEQ ID NO: 927) of SEQ ID NO: 572. The antigen in spot 50α may retain the sequence in SEQ ID NO: 572 as to amino acids at positions corresponding to the region of these epitopes.

[1238] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 50α may be the following variant.

[1239] In the epitope having the sequence of SEQ ID NO: 927, sites of amino acids other than X in XXXXKQXXX or XXDXKQXXX are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 62-70 of SEQ ID NO: 572 in the antigen in spot 50α may have the amino acid sequence of XXXXKQXXX or XXDXKQXXX.

[1240] Preferably, a protein that is an antigen in spot 50α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 572 (uncharacterized protein LOC100813859) or causative of an allergy to soybean.

(51α) Antigen in Spot 51α

[1241] As the result of sequence identification of the antigen in spot 51 by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 579-582 were detected.

[1242] Also, the mass spectroscopic data obtained for spot 51a on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, 14-3-3-like protein (amino acid sequence: SEQ ID NO: 578, encoding nucleotide sequence: SEQ ID NO: 577) was identified as a protein comprising any of SEQ ID NOs: 579 to 582.

[1243] Accordingly, the antigens in spot 51α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (51αA-a) to (51αA-c) and (51αB):

[1244] (51αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 578;

[1245] (51αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 578;

[1246] (51αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 577;

[1247] (51αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 577;

[1248] (51αA-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 577;

[1249] (51αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 579-582, preferably a protein comprising at least 2 or 3 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 579-582 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1250] The proteins selected from the group consisting of (51αA-a) to (51αA-c) and (51αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 5 to 60 kDa, preferably around 10 to 50 kDa and an isoelectric point of 2.5 to 6.5, preferably 3.0 to 6.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1251] Epitopes of the antigen in spot 51α reside in amino acids 71-85 (SEQ ID NO: 928), amino acids 141-155 (SEQ ID NO: 929) and amino acids 211-225 (SEQ ID NO: 930) of SEQ ID NO: 578. The antigen in spot 51α may retain the sequence in SEQ ID NO: 578 as to amino acids at positions corresponding to the region of these epitopes.

[1252] Preferably, a protein that is an antigen in spot 51α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 578 (14-3-3-like protein) or causative of an allergy to soybean.

(52α) Antigen in Spot 52α

[1253] As the result of sequence identification of the antigen in spot 52α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 585-588 were detected.

[1254] Also, the mass spectroscopic data obtained for spot 52α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, probable fructokinase-4 (amino acid sequence: SEQ ID NO: 584, encoding nucleotide sequence: SEQ ID NO: 583) was identified as a protein comprising any of SEQ ID NOs: 585 to 588.

[1255] Accordingly, the antigens in spot 52α in the present invention can be any protein selected from the group consisting of the proteins as defined below in (52αA-a) to (52αA-c) and (52«B):

[1256] (52αA-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 584;

[1257] (52αA-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 584;

[1258] (52αA-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 583;

[1259] (52αA-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 583;

[1260] (52αA-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 583;

[1261] (52αB) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 585-588, preferably a protein comprising at least 2 or 3 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 585-588 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1262] The proteins selected from the group consisting of (52αA-a) to (52αA-c) and (52αB) as defined above can be proteins that are found in a protein spot with a molecular weight of around 5 to 60 kDa, preferably around 10 to 50 kDa and an isoelectric point of 3.5 to 7.5, preferably 4.0 to 7.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1263] Epitopes of the antigen in spot 52α reside in amino acids 11-25 (SEQ ID NO: 933), amino acids 71-85 (SEQ ID NO: 934), amino acids 121-135 (SEQ ID NO: 935), amino acids 201-215 (SEQ ID NO: 936), amino acids 211-225 (SEQ ID NO: 937), and amino acids 281-295 (SEQ ID NO: 938) of SEQ ID NO: 584. The antigen in spot 52α may retain the sequence in SEQ ID NO: 584 as to amino acids at positions corresponding to the region of these epitopes.

[1264] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 52α may be the following variant.

[1265] In the epitope having the sequence of SEQ ID NO: 933, sites of amino acids other than X in XXXXXXSXXEXLXXX or SEQ ID NO: 932 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, amino acids 11-25 of SEQ ID NO: 584 in the antigen in spot 52α may have the amino acid sequence of XXXXXXSXXEXLXXX or SEQ ID NO: 932.

[1266] Preferably, a protein that is an antigen in spot 52α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 584 (probable fructokinase-4) or causative of an allergy to soybean.

(53α) Antigen in Spot 53α

[1267] As the result of sequence identification of the antigen in spot 53α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 591-604 were detected.

[1268] Also, the mass spectroscopic data obtained for spot 53α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, triosephosphate isomerase isoform X1 (amino acid sequence: SEQ ID NO: 590, encoding nucleotide sequence: SEQ ID NO: 589) was identified as a protein comprising any of SEQ ID NOs: 591 to 597 and triosephosphate isomerase, cytosolic (amino acid sequence: SEQ ID NO: 599, encoding nucleotide sequence: SEQ ID NO: 598) was identified as a protein comprising any of SEQ ID NOs: 600 to 604.

[1269] Accordingly, the antigen in spot 53α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (53αA1-a) to (53αA1-c), (53αB1), (53αA2-a) to (53αA2-c), and (53αB2):

[1270] (53αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 590;

[1271] (53αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 590;

[1272] (53αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 589;

[1273] (53αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 589;

[1274] (53αA1-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 589;

[1275] (53αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 591-597, preferably a protein comprising at least 2, 3, 4, 5, or 6 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 591 to 597 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[1276] (53αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 599;

[1277] (53αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 599;

[1278] (53αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 598;

[1279] (53αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 598;

[1280] (53αA2-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 598;

[1281] (53αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 600-604, preferably a protein comprising at least 2, 3, or 4 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 600 to 604 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1282] The protein selected from the group consisting of the proteins as defined above in (53αA1-a) to (53αA1-c), (53αB1), (53αA2-a) to (53αA2-c), and (53αB2) can be proteins that are found in a protein spot with a molecular weight of around 5 to 60 kDa, preferably around 10 to 50 kDa, and an isoelectric point of 3.5 to 8.5, preferably 4.0 to 8.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1283] Epitopes of the aforementioned antigens (53αA1-a) to (53αA1-e) and (53«B1) reside in amino acids 71-85 (SEQ ID NO: 939) of SEQ ID NO: 590. When the antigen in spot 53α contains a variation in the amino acid sequence of SEQ ID NO: 590, amino acids at positions corresponding to the region of at least one of these epitopes may retain the sequence in SEQ ID NO: 590.

[1284] Preferably, a protein that is an antigen in spot 53α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 590 (triosephosphate isomerase isoform X1) or the protein having the amino acid sequence of SEQ ID NO: 599 (triosephosphate isomerase, cytosolic) or causative of an allergy to soybean.

(54α) Antigen in Spot 54α

[1285] As the result of sequence identification of the antigen in spot 54α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 607-616 were detected.

[1286] Also, the mass spectroscopic data obtained for spot 54α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, superoxide dismutase [Fe], chloroplastic isoform X1 (amino acid sequence: SEQ ID NO: 606, encoding nucleotide sequence: SEQ ID NO: 605) as a protein comprising any of SEQ ID NOs: 607 to 611 and iron-superoxide dismutase (amino acid sequence: SEQ ID NO: 613, encoding nucleotide sequence: SEQ ID NO: 612) was identified as a protein comprising any of SEQ ID NOs: 614 to 616.

[1287] Accordingly, the antigen in spot 54α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (54αA1-a) to (54αA1-e), (54αB1), (54αA2-a) to (54αA2-c), and (54αB2):

[1288] (54αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 606;

[1289] (54αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 606;

[1290] (54αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 605;

[1291] (54αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 605;

[1292] (54αA1-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 605;

[1293] (54αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 607-611, preferably a protein comprising at least 2, 3, or 4 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 591 to 597 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[1294] (54αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 613;

[1295] (54αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 613;

[1296] (54αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 612;

[1297] (54αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 612;

[1298] (54αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 612;

[1299] (54αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 614-616, preferably a protein comprising at least 2, or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 614 to 616 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1300] The protein selected from the group consisting of the proteins as defined above in (54αA1-a) to (54αA1-c), (54αB1), (54αA2-a) to (54αA2-c), and (54αB2) can be proteins that are found in a protein spot with a molecular weight of around 5 to 60 kDa, preferably around 10 to 50 kDa, and an isoelectric point of 3.5 to 8.5, preferably 4.0 to 8.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1301] Epitopes of the aforementioned antigens (54αA1-a) to (54αA1-c) and (54&B1) reside in amino acids 121-135 (SEQ ID NO: 940), amino acids 148-154 (SEQ ID NO: 943) of SEQ ID NO: 606. When the antigen in spot 54α contains a variation in the amino acid sequence of SEQ ID NO: 606, amino acids at positions corresponding to the region of at least one of these epitopes may retain the sequence in SEQ ID NO: 606.

[1302] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 54α may be the following variant.

[1303] In the epitope having the sequence of SEQ ID NO: 943, sites of amino acids other than X in KXXXXTX or KXXXXTQ are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 54α contains a variation in the amino acid sequence of SEQ ID NO: 606, amino acids 148-154 of SEQ ID NO: 606 in this antigen may have the amino acid sequence of KXXXXTX or KXXXXTQ.

[1304] Preferably, a protein that is an antigen in spot 54α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 606 (superoxide dismutase [Fe], chloroplastic isoform X1) or the protein having the amino acid sequence of SEQ ID NO: 613 (iron-superoxide dismutase) or causative of an allergy to soybean.

(55α) Antigen in Spot 55α

[1305] As the result of sequence identification of the antigen in spot 55α by mass spectroscopy, the amino acid sequences of SEQ ID NOs: 619-649 were detected.

[1306] Also, the mass spectroscopic data obtained for spot 55α on a mass spectrometer was analyzed by comparing the data against the NCBI protein data, and as a result, 51 kDa seed maturation protein precursor (amino acid sequence: SEQ ID NO: 618, encoding nucleotide sequence: SEQ ID NO: 617) was identified as a protein comprising any of SEQ ID NOs: 619 to 635 and Lea protein precursor (amino acid sequence: SEQ ID NO: 637, encoding nucleotide sequence: SEQ ID NO: 636) was identified as a protein comprising any of SEQ ID NOs: 638 to 649.

[1307] Accordingly, the antigen in spot 55α in the present invention can be any of proteins selected from the group consisting of the proteins as defined below in (55αA1-a) to (55αA1-e), (55αB1), (55αA2-a) to (55αA2-c), and (55αB2):

[1308] (55αA1-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 618;

[1309] (55αA1-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 618;

[1310] (55αA1-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 617;

[1311] (55αA1-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 617;

[1312] (55αA1-e) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 617;

[1313] (55αB1) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 619-635, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or 16 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 619 to 635 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids;

[1314] (55αA2-a) a protein comprising an amino acid sequence with deletion, substitution, insertion or addition of one or several amino acids in SEQ ID NO: 637;

[1315] (55αA2-b) a protein comprising an amino acid sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of SEQ ID NO: 637;

[1316] (55αA2-c) a protein comprising an amino acid sequence encoded by a nucleotide sequence with deletion, substitution, insertion or addition of one or several nucleotides in SEQ ID NO: 636;

[1317] (55αA2-d) a protein comprising an amino acid sequence encoded by a nucleotide sequence having at least 70%, preferably at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% identity to the nucleotide sequence of SEQ ID NO: 636;

[1318] (55αA2-c) a protein comprising an amino acid sequence encoded by a nucleic acid that hybridizes under stringent conditions with a nucleic acid having a nucleotide sequence complementary to the nucleotide sequence of SEQ ID NO: 636;

[1319] (55αB2) a protein comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 638-649, preferably a protein comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11 or all sequences of the amino acid sequences. As referred to above, the amino acid sequence of any of SEQ ID NOs: 638 to 649 may be an amino acid sequence derived therefrom by deletion, substitution, insertion or addition of one or several amino acids.

[1320] The protein selected from the group consisting of the proteins as defined above in (55αA1-a) to (55αA1-c), (55αB1), (55αA2-a) to (55αA2-c), and (55αB2) can be proteins that are found in a protein spot with a molecular weight of around 30 to 90 kDa, preferably around 40 to 80 kDa, and an isoelectric point of 5.5 to 10.5, preferably 6.0

to 10.0 on gels used in the two-dimensional electrophoresis performed under the conditions described above in the subsection titled "Identification of antigens".

[1321] Epitopes of the aforementioned antigens (55αA1-a) to (55αA1-c) and (55αB1) reside in amino acids 201-209 (SEQ ID NO: 946), amino acids 318-325 (SEQ ID NO: 949), amino acids 202-210 (SEQ ID NO: 952) of SEQ ID NO: 618. When the antigen in spot 55α contains a variation in the amino acid sequence of SEQ ID NO: 618, amino acids at positions corresponding to the region of at least one of these epitopes may retain the sequence in SEQ ID NO: 618.

[1322] From the viewpoint of improving sensitivity (a degree to which a patient can be diagnosed as being positive) or specificity (a degree to which a healthy subject is not diagnosed as being positive), the antigen in spot 55α may be the following variant.

[1323] In the epitope having the sequence of SEQ ID NO: 946, sites of amino acids other than X in SEQ ID NO: 944 or 945 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 55α contains a variation in the amino acid sequence of SEQ ID NO: 618, amino acids 201-209 of SEQ ID NO: 618 in this antigen may have the amino acid sequence of SEQ ID NO: 944 or 945.

[1324] In the epitope having the sequence of SEQ ID NO: 949, sites of amino acids other than X in SEQ ID NO: 947 or 948 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 55α contains a variation in the amino acid sequence of SEQ ID NO: 618, amino acids 318-325 of SEQ ID NO: 618 in this antigen may have the amino acid sequence of SEQ ID NO: 947 or 948.

[1325] In the epitope having the sequence of SEQ ID NO: 952, sites of amino acids other than X in KDXXXXXXQ or SEQ ID NO: 951 are important for binding to IgE antibodies from allergic patients. Thus, in another preferred mode, when the antigen in spot 55α contains a variation in the amino acid sequence of SEQ ID NO: 618, amino acids 202-210 of SEQ ID NO: 618 in this antigen may have the amino acid sequence of KDXXXXXXQ or SEQ ID NO: 951.

[1326] Preferably, a protein that is an antigen in spot 55α has activity equivalent to that of the protein having the amino acid sequence of SEQ ID NO: 618 (51 kDa seed maturation protein precursor) or the protein having the amino acid sequence of SEQ ID NO: 637 (Lea protein precursor) or causative of an allergy to soybean.

[1327] The proteins that are the aforementioned antigens (1α) to (55α) include those proteins whose amino acid residues are modified by phosphorylation, sugar chain modification, aminoacylation, ring-opening, deamination or the like.

[1328] By stating herein "deletion, substitution, insertion or addition of one or several amino acids" in relation to amino acid sequence, it is meant that in an amino acid sequence of interest, one or several amino acids (e.g., 30%, preferably 25%, 20%, 15%, 10%, 5%, 3%, 2% or 1%, of amino acids with respect to the total length of the amino acid sequence, or e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 amino acids) are deleted, one or several amino acids are substituted by any other amino acids, any other amino acids are inserted, and/or any other amino acids are added.

[1329] Among the aforementioned modifications, substitution is preferably conservative substitution. The "conservative substitution" refers to the substitution of a certain amino acid residue by a different amino acid residue having similar physicochemical characteristics, and can be any type of substitution as long as it does not substantially change the characteristics of the structure of the original sequence for example, any type of substitution is acceptable as long as any substituted amino acids do not disrupt the helical structure of the original sequence or other secondary structures that characterize the original sequence. The following gives examples of separate groups of amino acid residues that are conservatively substitutable with each other, but substitutable amino acid residues are not limited to the examples given below.

[1330] Group A: leucine, isoleucine, valine, alanine, methionine

[1331] Group B: aspartic acid, glutamic acid

[1332] Group C: asparagine, glutamine

[1333] Group D: lysine, arginine

[1334] Group E: serine, threonine

[1335] Group F: phenylalanine, tyrosine

[1336] In the case of non-conservative substitution, one member belonging to one of the aforementioned groups can be replaced with a member belong to any other group. For example, in order to eliminate the possibility of unwanted sugar-chain modification, amino acids of group B, D or E as listed above may be substituted by those of any other group. Also, cysteines may be deleted or substituted by any other amino acids to prevent them from being folded into a protein in its tertiary structure. Also, in order to maintain the balance between hydrophilicity and hydrophobicity or to increase hydrophilicity for the purpose of facilitating synthesis, any amino acids may be substituted in consideration of the hydropathy scales of amino acids, which are a measure of the hydrophilic and hydrophobic properties of amino acids (J. Kyte and R. Doolittle, J. Mol. Biol., Vol. 157, p. 105-132, 1982).

[1337] As referred to herein, the percent identity between two amino acid sequences can be determined by visual inspection and mathematical calculation. Alternatively, the percent identity can be determined using a computer program. Examples of such computer programs include BLAST, FASTA (Altschul, et al., J. Mol. Biol., 215:403-410 (1990)), and ClustalW. In particular, various conditions (parameters) for identity searches with the BLAST program are described in Altschul, et al. (Nucl. Acids. Res., 25, p.3389-3402, 1997), and are publicly available on the websites of the National Center for Biotechnology Information (NCBI) and DNA Data Bank of Japan (DDBJ) (Altschul, et al., BLAST Manual, Altschul, et al., NCB/NLM/NIH Bethesda, MD 20894). Also, the percent identity can be determined using a genetic information processing software program, such as GENETYX Ver.7 (Genetyx Corporation), DNASIS Pro (Hitachi Software Engineering Co., Ltd.), or Vector NTI (Infomax Inc.).

[1338] By stating herein "deletion, substitution, insertion or addition of one or several nucleotides" in relation to nucleotide sequence, it is meant that in a nucleotide sequence of interest, one or several nucleotides (e.g., 30%, preferably 25%, 20%, 15%, 10%, 5%, 3%, 2% or 1%, of nucleotides with respect to the total length of the nucleotide sequence, or e.g., 1, 5, 10, 15, 20, 25 or 30 nucleotides) are deleted, one or several nucleotides are substituted by any other nucleotides, any other nucleotides are inserted, and/or any other nucleotides are added. It is preferable that such a

nucleotide deletion, substitution, insertion or addition should not give rise to a frame shift in an amino acid coding sequence.

[1339] As referred to herein, the percent identity between two nucleotide sequences can be determined by visual inspection and mathematical calculation. Alternatively, the percent identity can be determined using a computer program. Examples of such sequence comparison computer programs include the BLASTN program, version 2.2.7, available on the website of the National Library of Medicine (http://www.ncbi.nlm.nih.gov/blast/b12seq/bls.html) (Altschul, et al. (1990) J. Mol. Biol., 215:403-10), or the WU-BLAST 2.0 algorithm. Standard default parameter settings for WU-BLAST 2.0 are found and available on the following website: http://blast.wustl.edu.

[1340] As referred to herein, "under stringent conditions" means that hybridization takes place under moderately or highly stringent conditions. To be specific, the moderately stringent conditions can be easily determined by those having ordinary skill in the art on the basis of, for example, the length of DNA. Basic conditions are described in Sambrook, et al., Molecular Cloning: A Laboratory Manual, 3rd ed., ch.6-7, Cold Spring Harbor Laboratory Press, 2001. The moderately stringent conditions include hybridization under the conditions of preferably 1×SSC to 6×SSC at 42° C. to 55° C., more preferably 1×SSC to 3×SSC at 45° C. to 50° C., most preferably 2×SSC at 50° C. In the case of using a hybridization solution containing, for example, about 50% formamide, a temperature around 5 to 15° C. lower than the foregoing should be adopted. Washing is also carried out under the conditions of 0.5×SSC to 6×SSC at 40° C. to 60° C. In the process of hybridization and washing, generally 0.05% to 0.2% SDS, preferably about 0.1% SDS, may be added. Likewise, the highly stringent conditions can be easily determined by those having ordinary skill in the art on the basis of, for example, the length of DNA. Generally, the highly stringent (high stringent) conditions include hybridization and/or washing at a higher temperature and/or a lower salt concentration than those adopted under the moderately stringent conditions. For example, hybridization is carried out under the conditions of 0.1×SSC to 2×SSC at 55° C. to 65° C., more preferably 0.1×SSC to 1×SSC at 60° C. to 65° C., most preferably 0.2×SSC at 63° C. Washing is carried out under the conditions of 0.2×SSC to 2×SSC at 50° C. to 68° C., more preferably 0.2×SSC at 60 to 65° C.

[1341] The residue of X as to the epitopes of the aforementioned antigens (1α) to (55) is an amino acid residue at a site where binding activity against IgE antibodies from allergic patients was maintained even after substitution by alanine in alanine scanning described in Example 4. It is well known to those skilled in the art that even when such a site is substituted by any other amino acids, it is highly probable that this binding activity against IgE antibodies is maintained.

[1342] Antigens may be obtained by separating and purifying them from soybean using a combination of protein purification methods well known to those skilled in the art. Also, antigens may be obtained by expressing them as recombinant proteins using a genetic recombination technique well known to those skilled in the art and by separating and purifying them using protein purification methods well known to those skilled in the art.

[1343] Exemplary protein purification methods include: solubility-based purification methods such as salt precipitation and solvent precipitation; purification methods based on molecular weight difference, such as dialysis, ultrafiltration, gel filtration and SDS-PAGE; charge-based purification methods such as ion exchange chromatography and hydroxylapatite chromatography; specific affinity-based purification methods such as affinity chromatography; purification methods based on hydrophobicity difference, such as reverse-phase high-performance liquid chromatography; and purification methods based on isoelectric point difference, such as isoelectric focusing.

[1344] Preparation of a protein by a genetic recombination technique is carried out by preparing an expression vector comprising an antigen-encoding nucleic acid, introducing the expression vector into appropriate host cells by gene transfer or genetic transformation, culturing the host cells under suitable conditions for expression of a recombinant protein, and recovering the recombinant protein expressed in the host cells.

[1345] The "vector" refers to a nucleic acid that can be used to introduce a nucleic acid attached thereto into host cells. The "expression vector" is a vector that can induce the expression of a protein encoded by a nucleic acid introduced therethrough. Exemplary vectors include plasmid vectors and viral vectors. Those skilled in the art can select an appropriate expression vector for the expression of a recombinant protein depending on the type of host cells to be used.

[1346] The "host cells" refers to cells that undergo gene transfer or genetic transformation by a vector. The host cells can be appropriately selected by those skilled in the art depending on the type of the vector to be used. The host cells can be derived from prokaryotes such as *E. coli*. When prokaryotic cells like *E coli*. are used as host cells, the antigen of the present invention may be designed to contain an N-terminal methionine residue in order to facilitate the expression of a recombinant protein in the prokaryotic cells. The N-terminal methionine can be cleaved from the recombinant protein after expression. Also, the host cells may be cells derived from eukaryotes, such as single-cell eukaryotes like yeast, plant cells and animal cells (e.g., human cells, monkey cells, hamster cells, rat cells, murine cells or insect cells) or silkworm.

[1347] Gene transfer or genetic transformation of an expression vector into host cells can be carried out as appropriate by following a technique known to those skilled in the art. Those skilled in the art can make possible the expression of a recombinant protein by selecting suitable conditions for the expression of the recombinant protein as appropriate depending on the type of host cells and culturing the host cells under the selected conditions. Then, those skilled in the art can homogenize the host cells having the expressed recombinant protein, and separate and purify an antigen expressed as the recombinant protein from the resulting homogenate by using an appropriate combination of such protein purification methods as mentioned above.

Diagnosis Kit and Method (1)

[1348] The present invention provides a method for providing an indicator for diagnosing an allergy to a soybean in a subject, the method comprising the steps of:

[1349] (i) contacting a sample obtained from the subject with an antigen, wherein the sample is a solution comprising an Ig antibody;

[1350] (ii) detecting binding between an IgE antibody present in the sample from the subject and the antigen; and

[1351] (iii) when the binding between the IgE antibody in the subject and the antigen is detected, an indicator of the fact that the subject is allergic to a soybean is provided;

[1352] wherein the antigen is at least one of proteins as defined above in any of (1α) to (55α).

[1353] The sample obtained from a subject is a solution containing an Ig antibody, preferably an IgE antibody, as collected from the subject. Examples of such solutions include blood, saliva, sputum, snivel, urine, sweat, and tear. The sample obtained from the subject may be subjected to pretreatment for increasing the concentration of an Ig antibody in the sample before being contacted with an antigen. The pretreatment of a sample may involve, for example, collection of the serum or the plasma from the blood. Furthermore, a Fab moiety that is an antigen-binding moiety may be purified. In a particularly preferred mode, the step (i) mentioned above is carried out by contacting an IgE antibody present in the serum obtained from a subject with an antigen.

[1354] The IgE antibody may be the IgE antibody itself or may be mast cells or the like bound to IgE antibodies.

[1355] Detection of contact and binding between the sample obtained from a subject and an antigen can be carried out by using a known method. Examples of such methods that can be used include detection by ELISA (Enzyme-Linked Immunosorbent Assay), sandwich immunoassay, immunoblotting, immunoprecipitation, or immunochromatography. These are all techniques for detecting binding between an antigen and an IgE antibody from a subject by contacting and binding the IgE antibody from a subject with the antigen, allowing an enzymatically labelled secondary antibody to act on the IgE antibody specifically bound to the antigen, and adding an enzyme substrate (generally, chromogenic or luminescent reagent) to detect an enzymatic reaction product. Also, a method for detecting a fluorescently labeled secondary antibody can be used. Alternatively, detection by a measurement method capable of evaluating binding between an antigen and an IgE antibody, such as surface plasmon resonance (SPR), can also be used. A plurality of antigen-specific IgE antibodies may be mixed.

[1356] The antigen may be provided as an isolated antigen in a state immobilized on a carrier. In this case, the steps (i) and (ii) mentioned above can be carried out using ELISA, sandwich immunoassay, immunochromatography, surface plasmon resonance, or the like. Also, the step (i) mentioned above can be carried out by contacting the sample obtained from a subject with an antigen-immobilized surface. The isolated antigen may be obtained by separating and purifying it from soybean using a combination of protein purification methods well known to those skilled in the art, or by preparing it using a genetic recombination technique. The IgE antibody from the subject may be used in a state immobilized on a carrier, and binding to the antigen may be detected by the aforementioned technique.

[1357] The antigen may be in a state unimmobilized on a carrier. In this case, flow cytometry or the like can be used in the aforementioned steps (i) and (ii), and the presence of the antigen bound to the antibody can be confirmed with laser beam. Examples include a basophil activation test (BAT). Another example includes a histamine release test (HRT) which examines whether histamine is released by further contacting an antigen with blood cells in a sample.

[1358] The antigen may be detected by immunoblotting in a state separated by two-dimensional electrophoresis. The two-dimensional electrophoresis is a technique for separating a protein sample by performing isoelectric focusing in the first dimension and performing SDS-PAGE (SDS-polyacrylamide gel electrophoresis) in the second dimension. The conditions for two-dimensional electrophoresis are not particularly limited as long as the conditions permit the separation of the antigen of the present invention. For example, the conditions for two-dimensional electrophoresis as described above in the subsection titled "Identification of antigens" can be adopted. Also, the electrophoresis conditions may be defined by reference to the descriptions in Patent Literatures 1 to 4 mentioned above. For example, two-dimensional electrophoresis can be carried out under the conditions that satisfy at least one selected from the group consisting of the following requirements:

[1359] (A) the isoelectric focusing gels used in the first dimension should have a gel-strip length of 5 to 10 cm and a gel pH range of 3 to 10, and the pH gradient of the gels in the direction of electrophoresis should be as follows: where the gel-strip length up to pH 5 is taken as "a", that length from pH 5 to 7 as "b", and that length above pH 7 as "c", the relations "a<b" and "b>c" are satisfied;

[1360] (B) in the case of (A), when the total gel-strip length is taken as 1, "a" should be in the range of 0.15 to 0.3, "b" should be in the range of 0.4 to 0.7, and "c" should be in the range of 0.15 to 0.3;

[1361] (C) in the first dimensional isoelectric focusing, a constant voltage step should be performed by applying a constant voltage ranging from 100 V to 600 V per gel strip containing a sample, and after the electrophoresis variation width during electrophoresis for 30 minutes falls within the range of 5 μA, a voltage-increasing step should be started at which the voltage is increased from the aforementioned constant voltage;

[1362] (D) in the case of (C), the final voltage at the voltage-increasing step should be in the range of 3000 V to 6000 V;

[1363] (E) the isoelectric focusing gels used in the first dimension should have a longitudinal gel-strip length of 5 to 10 cm, and the electrophoresis gels used in the second dimension should have a gel concentration at the distal end in the direction of electrophoresis, which is in the range of 3 to 6%; and

[1364] (F) in the case of (E), the electrophoresis gels used in the second dimension should have a gel concentration at the proximal end in the direction of electrophoresis, which is set to a higher value than that at the distal end in the direction of electrophoresis.

[1365] The aforementioned antigens (1α) to (55α) are antigens that specifically bind to IgE antibodies from patients with allergy to soybean. Therefore, when binding between an IgE antibody from a subject and the antigen is detected, an indicator of the fact that the subject is allergic to soybean is provided.

[1366] The present invention further provides a kit for diagnosing an allergy to a soybean, comprising at least one of the aforementioned antigens (1α) to (55α). The diagnosis kit of this invention may be used in the aforementioned method for providing an indicator for diagnosing an allergy

to a soybean or in a diagnosis method as described later. The diagnosis kit of this invention may comprise not only the at least one of the aforementioned antigens (1α) to (55α), but also an anti-IgE antibody labeled with an enzyme and a chromogenic or luminescent substrate serving as a substrate for the enzyme. Also, a fluorescent-labeled anti-IgE antibody may be used. In the diagnosis kit of this invention, the antigen may be provided in a state immobilized on a carrier. The diagnosis kit of this invention may also be provided together with instructions on the procedure for diagnosis or a package containing said instructions.

[1367] In another mode, the aforementioned diagnosis kit comprises a companion diagnostic agent for an allergy to a soybean. The companion diagnostic agent is used for identifying patients expected to respond to pharmaceutical products or identifying patients having the risk of severe adverse reactions to pharmaceutical products, or for studying the reactivity of pharmaceutical products in order to optimize treatment using the pharmaceutical products. Here, the optimization of treatment includes, for example, determination of dosage and administration, judgment regarding discontinuation of administration, and confirmation of an allergen component that is used to cause immunological tolerance.

[1368] The present invention further provides a composition for diagnosing an allergy to a soybean, comprising at least one of the aforementioned antigens (1α) to (55α). The diagnosis composition of this invention can be used in a diagnosis method as described below. The diagnosis composition of this invention may further comprise a pharmaceutically acceptable carrier and/or additives commonly used with the antigen of this invention depending on the need.

[1369] In one mode, the present invention provides a method for diagnosing an allergy to a soybean in a subject, the method comprising:

[1370] (i) contacting a sample obtained from the subject with an antigen;

[1371] (ii) detecting binding between an IgE antibody present in the sample from the subject and the antigen; and

[1372] (iii) when the binding between the IgE antibody in the subject and the antigen is detected, diagnosing the subject as being allergic to a soybean;

[1373] wherein the antigen is at least one of proteins as defined above in any of (1α) to (55α). In this method, the steps (i) and (ii) are performed as described above regarding the corresponding steps of the method for providing an indicator for diagnosing an allergy to a soybean.

[1374] In another mode, the present invention provides a method for diagnosing an allergy to a soybean in a subject, the method comprising administering to the subject at least one of the aforementioned antigens (1α) to (55α). This method may be performed in the form of a skin test characterized by applying the antigen onto the skin. Examples of the skin test include various forms of tests, such as: a prick test in which a diagnosis composition is applied onto the skin and then a tiny prick to such an extent as not to provoke bleeding is made in the skin to allow an antigen to penetrate the skin, thereby observing a skin reaction; a scratch test in which a diagnosis composition is applied onto the skin and then the skin is lightly scratched to observe a reaction; a patch test in which a diagnosis composition in the form of cream, ointment, etc. is applied onto the skin to observe a reaction; and an intracutaneous test in which an antigen is administered intracutaneously to observe a reaction. If a skin reaction such as swelling occurs in a skin portion to which the antigen has been applied, the subject of interest is diagnosed as having an allergy to a soybean. The amount of the antigen to be applied to the skin in such tests can be, for example, not more than 100 μg per dose.

[1375] In the process of allergy diagnosis, a load test aiming to identify an antigen is often adopted. At least one of the aforementioned antigens (1α) to (55α) can be used as an active ingredient for a load test to diagnose an allergy to a soybean. Here, the antigen protein to be used in the load test may be a protein that has been expressed and purified and may be a protein that has been expressed in food or cooking ingredients, such as rice-based vaccine expressing pollen allergens which are obtained by transforming rice with a gene of a cedar pollen antigen and expressing the antigen protein in the rice.

[1376] In yet another mode, the present invention provides at least one of the aforementioned antigens (1α) to (55α), intended for use in the diagnosis of an allergy to a soybean. This also includes the provision of at least one of the aforementioned antigens (1α) to (55α) mixed with a known antigen.

[1377] In still another mode, the present invention provides use of at least one of the aforementioned antigens (1α) to (55α) for the production of a diagnostic agent for an allergy to a soybean.

Pharmaceutical Composition and Treatment Method (1)

[1378] The present invention provides a pharmaceutical composition comprising at least one of the aforementioned antigens (1α) to (55α). In one mode, the aforementioned pharmaceutical composition is used for the treatment and diagnosis of an allergy to a soybean.

[1379] The present invention also provides a method for treating an allergy to a soybean, the method comprising administering at least one of the aforementioned antigens (1α) to (55α) to a patient in need of a treatment for an alle-rgy to a soybean.

[1380] In another mode, the present invention provides at least one of the aforementioned antigens (1α) to (55α), intended for use in the treatment for an allergy to a soybean. In yet another mode, the present invention provides use of at least one of the aforementioned antigens (1α) to (55α) for the production of a therapeutic agent for an allergy to a soybean.

[1381] In the process of allergy treatment, a hyposensitization therapy aiming to induce immunological tolerance by administering an antigen to a patient is often adopted. The at least one of the aforementioned antigens (1α) to (55α) can be used as an active ingredient for a hyposensitization therapy for an allergy to a soybean. Here, the antigen protein to be used in the hyposensitization therapy may be a protein that has been expressed and purified and may be a protein that has been expressed in food or cooking ingredients, such as rice-based vaccine expressing pollen allergens which are obtained by transforming rice with a gene of a cedar pollen antigen and expressing the antigen protein in the rice.

[1382] The pharmaceutical composition of the present invention can be administered by common administration routes. Examples of common administration routes include oral, sublingual, percutaneous, intracutaneous, subcutane-

ous, intravascular, intranasal, intramuscular, intraperitoneal, and intrarectal administrations.

[1383] The pharmaceutical composition of the present invention can be used as a pharmaceutical composition to which a commonly used pharmaceutically acceptable adjuvant or excipient or any other additives (e.g., stabilizer, solubilizer, emulsifier, buffer, preservative, colorant) are added by a conventional method together with the antigen of this invention depending on the need. The dosage form of the pharmaceutical composition can be selected by those skilled in the art as appropriate depending on the administration route. The pharmaceutical composition can be in the form of, for example, tablet, capsule, troche, sublingual tablet, parenteral injection, intranasal spray, poultice, solution, cream, or lotion. The administration dose, frequency and/or period of the pharmaceutical composition of this invention can be selected by a physician as appropriate depending on the administration route and the patient's condition and characteristics such as age and body weight. For example, the pharmaceutical composition may be administered to an adult patient at a dose of not more than 100 μg per dose. The administration interval can be, for example, once daily, once weekly, twice weekly, once per three months or so. The administration period can be, for example, several weeks to several years. The pharmaceutical composition may be administered in such a manner that the dose is increased in incremental steps over the administration period.

Tester (1)

[1384] The present invention provides a tester comprising an antibody for at least one of the aforementioned antigens (1α) to (55α).

[1385] The antibody can be prepared by a conventional method. For example, the antibody may be prepared by immunizing a mammal such as rabbit with one of the aforementioned antigens (1α) to (55α). The antibody may be an Ig antibody, a polyclonal antibody, a monoclonal antibody, or an antigen-binding fragment thereof (e.g., Fab, F(ab') 2, Fab').

[1386] Further, in the aforementioned tester, the antibody may be provided in a form bound to a carrier. The type of the carrier is not particularly limited as long as it is usable for detection of binding between an antibody and an antigen. Any given carrier known to those skilled in the art can be used.

[1387] Examples of a method for determining the presence or absence of an antigen include the following methods:

[1388] a method in which a prepared tester comprising an IgE antibody is contacted with a sample obtained from a food, a cooking ingredient, etc., ELISA or the like method is used to detect whether there is a binding between the IgE antibody and an antigen in the sample, and if the binding between the IgE antibody and the antigen is detected, it is determined that the antibody remains in the food, the cooking ingredient, etc. of interest; and

[1389] a method in which filter paper is impregnated with a food or a cooking ingredient and reacted with an antibody solution so as to detect an antigen contained therein.

[1390] Another mode of the present invention includes a tester for determining the presence or absence of an antigen of an allergy to a soybean in an object of interest, the tester comprising a primer having a nucleotide sequence complementary to a portion of at least one of the nucleotide sequences of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 69, 88, 105, 121, 138, 154, 158, 162, 168, 173, 188, 200, 203, 219, 232, 238, 241, 253, 279, 285, 290, 305, 319, 326, 333, 340, 347, 357, 366, 383, 393, 410, 417, 421, 428, 434, 440, 446, 452, 455, 459, 467, 472, 478, 484, 488, 494, 514, 517, 521, 524, 532, 540, 546, 552, 555, 559, 564, 571, 577, 583, 589, 598, 605, 612, 617 and 636. The primer has, for example, but not limited to, a nucleotide sequence complementary to, preferably, 12 residues, 15 bases, 20 bases, or 25 bases, in a 3'-terminal or central sequence in a partial sequence of at least one of the nucleotide sequences of SEQ ID NOs defined above in this paragraph. Particularly, when mRNA is of interest, the tester has a complementary primer of a poly-A tail. In a preferred mode, the tester comprising the primer mentioned above may further comprise a primer comprising a 5'-terminal nucleotide sequence, preferably a nucleotide sequence of 12 bases, 15 bases, 20 bases, or 25 bases, of at least one of the nucleotide sequences of SEQ ID NOs defined above in this paragraph.

[1391] For example, cDNA is amplified by PCR (Polymerase Chain Reaction) including RT-PCR using templated DNA or mRNA obtained from a soybean and the aforementioned complementary primer, and the sequence of the amplified cDNA is compared with the nucleotide sequences of SEQ ID NOs defined above in the previous paragraph to determine the presence or absence of the antigen. Amplification methods by PCR can be exemplified by RACE. In this respect, even if there exists a point mutation encoding the same amino acid in the comparison of the amplified cDNA with the nucleotide sequences of SEQ ID NOs defined above in the previous paragraph, or even if the nucleotide sequence of the amplified cDNA has insertion, deletion, substitution or addition of bases in the nucleotide sequences of SEQ ID NOS defined above in the previous paragraph, it is determined that the antigen is present when the amino acid sequence encoded by the cDNA has at least 70%, preferably at least 80, 90, 95, 98, or 99% identity to the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 70, 89, 106, 122, 139, 155, 159, 163, 169, 174, 189, 201, 204, 220, 233, 239, 242, 254, 280, 286, 291, 306, 320, 327, 334, 341, 348, 358, 367, 384, 394, 411, 418, 422, 429, 435, 441, 447, 453, 456, 460, 468, 473, 479, 485, 489, 495, 515, 518, 522, 525, 533, 541, 547, 553, 556, 560, 565, 572, 578, 584, 590, 599, 606, 613, 618 or 637.

[1392] In one mode, the aforementioned tester is used to determine the presence or absence of an antigen in cooking ingredients or in products of interest in a food production line. The tester may also be used for quality inspection of production lines and pre-shipment products by manufacturers, or may be used for self-checking of the presence or absence of an antigen in a food or cooking ingredients of interest by consumers.

Allergen-Free Food and the Like (1)

[1393] The present invention provides a soybean or processed product of soybean in which at least one of the aforementioned antigens (1α) to (55α) is eliminated or reduced.

[1394] The method for eliminating or reducing the antigen of the present invention in a soybean or processed products of soybean is not limited. The elimination or reduction of the

antigen may be conducted by any method, as long as the method permits the elimination or reduction of the antigen.

[1395] For example, the soybean of the present invention whose antigen is eliminated or reduced may be obtained by preparing a soybean in which the expression of the antigen of the present invention is knocked out, using a gene knock-out technique. Any technique known to those skilled in the art such as genetic modification can be used as the gene knock-out technique.

[1396] An antigen of the present invention may be the artefact that an antigen of the present invention assumed the removal or a reduced soybean raw material as for the removal or the reduced processed products of soybean. In the case of using an ordinary soybean as a source ingredient, a treatment for removing or reducing the antigen of this invention is performed before or after preparation of a processed product of soybean. The methods for removing or reducing the antigen of the present invention in the processed products of soybean which assumed an ordinary soybean raw material include a method to remove protein component in food such as a high pressure treatment and elution with the neutral salt solution or the high temperature steam, and a method to perform hydrolysis, denaturation, or amino acid alteration (chemical modification, elimination, or the like of a side chain) by heat treatment and acid treatment.

Method for Producing Antigen-Free Processed Product of Soybean

[1397] The present invention provides a method for producing a processed product of soybean in which an antigen is eliminated or reduced, the method comprising the step of confirming that the antigen is eliminated or reduced, in a production process of the processed product, wherein the antigen is at least one of the aforementioned antigens (1) to (55α).

[1398] The step of confirming that the antigen is eliminated or reduced, in a production process of the processed product of soybean in which an antigen is eliminated or reduced may be performed by confirming the presence or absence of an antigen by the method described above in the subsection titled "Tester (1)".

[1399] The production of the processed product of soybean in which an antigen is eliminated or reduced may be performed by the method described above in the subsection titled "Allergen-free food and the like (1)".

Epitope of Antigen

[1400] Epitopes and amino acids important for binding activity against IgE antibodies from allergic patients within the epitopes were identified as shown in Examples 4 and 5 as to antigens identified as shown in Example 3 and known antigens Gly m 4, Gly m 5, Gly m 6 and Gly m 8.

[1401] The present invention provides polypeptides of the following (1β) to (60β) as polypeptides comprising an amino acid sequence specifically binding to an IgE antibody from an allergic patient.

(1β) Epitope of Endoplasmin Homolog Isoform X2 (Spot 1α)

[1402] In the present invention, the polypeptide (1β) can be any polypeptide selected from the group consisting of (1β-1) to (1β-4) as defined below:

[1403] (1β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 652, 655 and 657-660;

[1404] (1β-2) a polypeptide comprising the amino acid sequence of XXKXEXXX, preferably a polypeptide comprising the amino acid sequence of XTKXEXXX, more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 1, 2, 4, and 6-8 of SEQ ID NO: 652 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 1, 4, and 6-8 of SEQ ID NO: 652 are substituted by other amino acids;

[1405] (1-3) a polypeptide comprising the amino acid sequence of XXXKDNVXX (SEQ ID NO: 653), preferably a polypeptide comprising the amino acid sequence of XLXKDNVXX (SEQ ID NO: 654), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 1-3, 8, and 9 of SEQ ID NO: 655 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 1, 3, 8, and 9 of SEQ ID NO: 655 are substituted by other amino acids; and

[1406] (1β-4) a polypeptide comprising the amino acid sequence of EEEDDXXXXXEXXXX (SEQ ID NO: 656), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 6-10 and 12-15 of SEQ ID NO: 657 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 6-10 and 12-15 of SEQ ID NO: 657 are substituted by other amino acids.

(2β) Epitope of Heat Shock Cognate Protein 80 (Spot 2α)

[1407] In the present invention, the polypeptide (2β) can be any polypeptide selected from the group consisting of (2β-1) to (2β-9) as defined below:

[1408] (2β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 663, 666, 669, 670-672, 675, 676-678, 681, 684, 687, 689 and 690-692;

[1409] (2-2) a polypeptide comprising the amino acid sequence of DXXDKXXXX, preferably a polypeptide comprising the amino acid sequence of DXXDKXRXX (SEQ ID NO: 662), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 2, 3, and 6-9 of SEQ ID NO: 663 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 2, 3, 6, 8, and 9 of SEQ ID NO: 663 are substituted by other amino acids;

[1410] (2β-3) a polypeptide comprising the amino acid sequence of XXXXKSKLXXXXXXX (SEQ ID NO: 664), preferably a polypeptide comprising the amino acid sequence of XXXXKSKLXXXPXXX (SEQ ID NO: 665), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 of the amino acids at positions 1-4, and 9-15

of SEQ ID NO: 666 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 of the amino acids at positions 1-4, 9-11, and 13-15 of SEQ ID NO: 666 are substituted by other amino acids;

[1411] (2β-4) a polypeptide comprising the amino acid sequence of XPDKX, preferably a polypeptide comprising the amino acid sequence of IPDKX (SEQ ID NO: 668), more preferably a polypeptide comprising an amino acid sequence in which any 1 or 2 of the amino acids at positions 1 and 5 of SEQ ID NO: 669 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1 or 2 of the amino acids at positions 1 and 5 of SEQ ID NO: 669 are substituted by other amino acids;

[1412] (2β-5) a polypeptide comprising the amino acid sequence of XXFSKXXXXXXXEDX (SEQ ID NO: 673), preferably a polypeptide comprising the amino acid sequence of XXFSKXXXXXXXEDS (SEQ ID NO: 674), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 of the amino acids at positions 1, 2, 6-12 and 15 of SEQ ID NO: 675 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, or 9 of the amino acids at positions 1, 2 and 6-12 of SEQ ID NO: 675 are substituted by other amino acids;

[1413] (2β-6) a polypeptide comprising the amino acid sequence of XXXXDXXXXEXXXDX, preferably a polypeptide comprising the amino acid sequence of XXXXDKIRXEXXXDK (SEQ ID NO: 680), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 of the amino acids at positions 1-4, 6-9, 11-13 and 15 of SEQ ID NO: 681 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7 or 8 of the amino acids at positions 1-4, 9, 11-13 of SEQ ID NO: 681 are substituted by other amino acids;

[1414] (2β-7) a polypeptide comprising the amino acid sequence of XYXXRXXXXQNXXXX (SEQ ID NO: 682), preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 of the amino acids at positions 1, 3, 4, 6-9, and 12-15 of SEQ ID NO: 684 are substituted by alanine, more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 of the amino acids at positions 1, 3, 4, 6-9, and 12-15 of SEQ ID NO: 684 are substituted by other amino acids;

[1415] (2β-8) a polypeptide comprising the amino acid sequence of DXXXRXXXXQNXXXX (SEQ ID NO: 685), preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 of the amino acids at positions 2-4, 6-9, and 12-15 of SEQ ID NO: 687 are substituted by alanine, more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 of the amino acids at positions 2-4, 6-9, and 12-15 of SEQ ID NO: 687 are substituted by other amino acids;

[1416] (2β-9) a polypeptide comprising the amino acid sequence of RKPXXXXXXXXAAXX (SEQ ID NO: 688), preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 of the amino acids at positions 4-11, 14 and 15 of SEQ ID NO: 689 are substituted by alanine, more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 of the amino acids at positions 4-11, 14 and 15 of SEQ ID NO: 689 are substituted by other amino acids.

(3β) Epitope of Cell Division Cycle Protein 48 Homolog (Spot 3α)

[1417] In the present invention, the polypeptide (3β) can be any polypeptide selected from the group consisting of (3β-1) to (3β-4) as defined below:

[1418] (3β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 695, 697, 699 and 700-716;

[1419] (3β-2) a polypeptide comprising the amino acid sequence of XXXXXXXXXXRLXEX, preferably a polypeptide comprising the amino acid sequence of XXXLKREDEERLDEV (SEQ ID NO: 694), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 of the amino acids at positions 1-10, 13 and 15 of SEQ ID NO: 695 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2 or 3 of the amino acids at positions 1-3 of SEQ ID NO: 695 are substituted by other amino acids;

[1420] (3β-3) a polypeptide comprising the amino acid sequence of XXXXXXXXXXXXXDE, preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or 13 of the amino acids at positions 1-13 of SEQ ID NO: 697 are substituted by alanine, more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or 13 of the amino acids at positions 1-13 of SEQ ID NO: 697 are substituted by other amino acids;

[1421] (3β-4) a polypeptide comprising the amino acid sequence of XXXXGDR, preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 1~4 of SEQ ID NO: 699 are substituted by alanine, more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 1-4 of SEQ ID NO: 699 are substituted by other amino acids.

[1422] In one mode, the polypeptide (3β) may not comprise the full-length amino acid sequence (SEQ ID NO: 16) of Cell division cycle protein 48 homolog or a variant or a homolog having 100% or at least 90%, at least 80% or at least 70% identity to the total length of this amino acid sequence.

[1423] SEQ ID NOs: 695, 697, 699 and 700-716 correspond to amino acids 191-205 (SEQ ID NO: 695), amino acids 21-35 (SEQ ID NO: 697), amino acids 779-785 (SEQ ID NO: 699), amino acids 8-14 (SEQ ID NO: 700), amino acids 49-56 (SEQ ID NO: 701), amino acids 101-107 (SEQ ID NO: 702), amino acids 161-175 (SEQ ID NO: 703), amino acids 291-305 (SEQ ID NO: 704), amino acids 316-323 (SEQ ID NO: 705), amino acids 327-333 (SEQ ID NO: 706), amino acids 341-355 (SEQ ID NO: 707), amino acids 361-369 (SEQ ID NO: 708), amino acids 431-445 (SEQ ID NO: 709), amino acids 451-465 (SEQ ID NO: 710), amino acids 501-515 (SEQ ID NO: 711), amino acids

565-572 (SEQ ID NO: 712), amino acids 631-645 (SEQ ID NO: 713), amino acids 703-712 (SEQ ID NO: 714), amino acids 711-715 (SEQ ID NO: 715), and amino acids 721-735 (SEQ ID NO: 716), respectively, of SEQ ID NO: 16.

(4β) Epitope of Embryo-Specific Urease Isoform X1 (Spot 4α)

[1424] In the present invention, the polypeptide (4β) can be any polypeptide selected from the group consisting of (4β-1) to (4β-5) as defined below:

[1425] (4β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 719, 722, 725, 726-728 and 731;

[1426] (4β-2) a polypeptide comprising the amino acid sequence of XXXXXXXKED, preferably a polypeptide comprising the amino acid sequence of XXXXXEXKED (SEQ ID NO: 718), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6 or 7 of the amino acids at positions 1-7 of SEQ ID NO: 719 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 1-5 and 7 of SEQ ID NO: 719 are substituted by other amino acids;

[1427] (4β-3) a polypeptide comprising the amino acid sequence of KXXLGDX (SEQ ID NO: 720), preferably a polypeptide comprising the amino acid sequence of KIXLGDX (SEQ ID NO: 721), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2 or 3 of the amino acids at positions 2, 3 and 7 of SEQ ID NO: 722 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2 or 3 of the amino acids at positions 3 and 7 of SEQ ID NO: 722 are substituted by other amino acids;

[1428] (4β-4) a polypeptide comprising the amino acid sequence of XXXDXFXKXE (SEQ ID NO: 723), preferably a polypeptide comprising the amino acid sequence of XXXDXFXKIE (SEQ ID NO: 724), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 1-3, 5, 7 and 9 of SEQ ID NO: 725 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 1-3, 5 and 7 of SEQ ID NO: 725 are substituted by other amino acids;

[1429] (4β-5) a polypeptide comprising the amino acid sequence of XXXHXLXXXNKXXEA (SEQ ID NO: 729), preferably a polypeptide comprising the amino acid sequence of XXVHXLXXXNKXXEA (SEQ ID NO: 730), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 1-3, 5, 7-9,12 and 13 of SEQ ID NO: 731 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7 or 8 of the amino acids at positions 1, 2, 5, 7-9, 12 and 13 of SEQ ID NO: 731 are substituted by other amino acids.

(53) Alpha-1,4 Glucan Phosphorylase L Isozyme, Chloroplastic/Amyloplastic (Spot 5α)

[1430] In the present invention, the polypeptide (5β) can be any polypeptide selected from the group consisting of (5β-1) to (5β-2) as defined below:

[1431] (5β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 734 and 735-738;

[1432] (5β-2) a polypeptide comprising the amino acid sequence of XXXXXXGLXKXRXEX (SEQ ID NO: 732), preferably a polypeptide comprising the amino acid sequence of XXXXXXGLXKXRXEG (SEQ ID NO: 733), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 8, 9 or 10 of the amino acids at positions 1-6, 9, 11, 13 and 15 of SEQ ID NO: 734 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 1-6, 9, 11 and 13 of SEQ ID NO: 734 are substituted by other amino acids.

(6β) Epitope of Aconitate Hydratase 1 (Spot 6α)

[1433] In the present invention, the polypeptide (6β) can be a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 739 and 740.

(7β) Epitope of Methionine Synthase (Spot 7α)

[1434] In the present invention, the polypeptide (7β) can be any polypeptide selected from the group consisting of (7β-1) to (7β-2) as defined below:

[1435] (7β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 741 and 744;

[1436] (7β-2) a polypeptide comprising the amino acid sequence of XXXXXXXKDXVEDXE (SEQ ID NO: 742), preferably a polypeptide comprising the amino acid sequence of XXXXXXXKDEVEDXE (SEQ ID NO: 743), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 1-7, 10 and 14 of SEQ ID NO: 744 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7 or 8 of the amino acids at positions 1-7 and 14 of SEQ ID NO: 744 are substituted by other amino acids.

(9β) Epitope of Lysine—tRNA Ligase, Cytoplasmic-Like Isoform X2 (Spot 9α)

[1437] In the present invention, the polypeptide (9β) can be any polypeptide selected from the group consisting of (9β-1) to (9β-3) as defined below:

[1438] (9β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 747 and 749;

[1439] (9β-2) a polypeptide comprising the amino acid sequence of XXEXXXXKDX, preferably a polypeptide comprising the amino acid sequence of NXEXXXXKDX (SEQ ID NO: 746), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6 or 7 of the amino acids at positions 1, 2, 4-7 and 10 of SEQ ID NO: 747 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 2, 4-7 and 10 of SEQ ID NO: 747 are substituted by other amino acids;

[1440] (9-3) a polypeptide comprising the amino acid sequence of LDFLEXET (SEQ ID NO: 748), prefer-

ably a polypeptide comprising an amino acid sequence in which the amino acid at position 6 of SEQ ID NO: 749 is substituted by alanine, more preferably a polypeptide comprising an amino acid sequence in which the amino acid at position 6 of SEQ ID NO: 749 is substituted by another amino acid.

(11β) Epitope of Polyadenylate-Binding Protein 8 Isoform X3 (Spot 11α)

[1441] In the present invention, the polypeptide (11β) can be any polypeptide selected from the group consisting of (11β-1) to (11β-2) as defined below:

[1442] (11β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 753 and 754;

[1443] (11β-2) a polypeptide comprising the amino acid sequence of PLXQQXXXXX (SEQ ID NO: 751), preferably a polypeptide comprising the amino acid sequence of PLXQQGXXXX (SEQ ID NO: 752), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 3 and 6-10 of SEQ ID NO: 753 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 3 and 7-10 of SEQ ID NO: 753 are substituted by other amino acids.

(12β) Epitope of Pyrophosphate—Fructose 6-Phosphate 1-Phosphotransferase Subunit Alpha-Like (Spot 12α)

[1444] In the present invention, the polypeptide (12) can be a polypeptide comprising the amino acid sequence of SEQ ID NO: 755.

(13β) Epitope of Protein Disulfide Isomerase-Like Protein Precursor (Spot 13α)

[1445] In the present invention, the polypeptide (13β) can be any polypeptide selected from the group consisting of (13β-1) to (13β-2) as defined below:

[1446] (13-1) a polypeptide comprising the amino acid sequence of SEQ ID NO: 758;

[1447] (13β-2) a polypeptide comprising the amino acid sequence of XXGXXXXXXXXXXE, preferably a polypeptide comprising the amino acid sequence of XXGXLXXXXGXRTXE (SEQ ID NO: 757), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or 13 of the amino acids at positions 1, 2 and 4-14 of SEQ ID NO: 758 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 1, 2, 4, 6-9, 11 and 14 of SEQ ID NO: 758 are substituted by other amino acids.

(14β) Epitope of V-Type Proton ATPase Subunit B 2 (Spot 14α)

[1448] In the present invention, the polypeptide (14β) can be any polypeptide selected from the group consisting of (14β-1) to (14β-3) as defined below:

[1449] (14β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 761 and 764;

[1450] (14β-2) a polypeptide comprising the amino acid sequence of XXXXXLLXXXXEEDN (SEQ ID NO: 759), preferably a polypeptide comprising the amino acid sequence of XXXXXLLEXXXEEDN (SEQ ID NO: 760), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 1-5 and 8-11 of SEQ ID NO: 761 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7 or 8 of the amino acids at positions 1-5 and 9-11 of SEQ ID NO: 761 are substituted by other amino acids;

[1451] (14β-3) a polypeptide comprising the amino acid sequence of RKXXXXXXXXXXYXNX (SEQ ID NO: 762), preferably a polypeptide comprising the amino acid sequence of RKDHSXXXXXLYANX (SEQ ID NO: 763), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 of the amino acids at positions 3-11, 13 and 15 of SEQ ID NO: 764 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 6-10 and 15 of SEQ ID NO: 764 are substituted by other amino acids.

(15β) Epitope of UTP-Glucose-1-Phosphate Uridylyltransferase (Spot 15α, 16α)

[1452] In the present invention, the polypeptide (15β) can be any polypeptide selected from the group consisting of (15β-1) to (15β-2) as defined below:

[1453] (15-1) a polypeptide comprising the amino acid sequence of SEQ ID NO: 766;

[1454] (15β-2) a polypeptide comprising the amino acid sequence of XGKXXXXXXXXXXXX, preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or 13 of the amino acids at positions 1 and 4-15 of SEQ ID NO: 766 are substituted by alanine, more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or 13 of the amino acids at positions 1 and 4-15 of SEQ ID NO: 766 are substituted by other amino acids.

(17) Epitope of Guanosine Nucleotide Diphosphate Dissociation Inhibitor 2 (Spot 17α)

[1455] In the present invention, the polypeptide (17β) can be a polypeptide comprising the amino acid sequence of SEQ ID NO: 767.

(18) Epitope of Hypothetical Protein GLYMA 10G028300 (Spot 18α)

[1456] In the present invention, the polypeptide (18β) can be any polypeptide selected from the group consisting of (18β-1) to (18β-3) as defined below:

[1457] (18β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 750, 768, 771, 772 and 775;

[1458] (18β-2) a polypeptide comprising the amino acid sequence of XXXXXXXNXXIFEED (SEQ ID NO: 769), preferably a polypeptide comprising the amino acid sequence of XXXXXXXNXYIFEED (SEQ ID NO: 770), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 1-7, 9 and 10 of SEQ ID NO: 771 are substituted by alanine, further

preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7 or 8 of the amino acids at positions 1-7 and 9 of SEQ ID NO: 771 are substituted by other amino acids;

[1459] (18β-3) a polypeptide comprising the amino acid sequence of XELPREXRXX (SEQ ID NO: 773), preferably a polypeptide comprising the amino acid sequence of XELPREERXX (SEQ ID NO: 774), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 1, 7, 9 and 10 of SEQ ID NO: 775 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2 or 3 of the amino acids at positions 1, 9 and 10 of SEQ ID NO: 775 are substituted by other amino acids.

(19β) Epitope of Sucrose Binding Protein Homolog S-64 (Spot 19α)

[1460] In the present invention, the polypeptide (19β) can be any polypeptide selected from the group consisting of (19β-1) to (19β-5) as defined below:

[1461] (19β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 778, 781, 784, 786 and 787;

[1462] (19-2) a polypeptide comprising the amino acid sequence of XXXXXPLYXXXXDEX (SEQ ID NO: 776), preferably a polypeptide comprising the amino acid sequence of XXXXXPLYIXXXDEX (SEQ ID NO: 777), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 of the amino acids at positions 1-5, 9-12 and 15 of SEQ ID NO: 778 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 1-5, 10-12 and 15 of SEQ ID NO: 778 are substituted by other amino acids;

[1463] (19β-3) a polypeptide comprising the amino acid sequence of XXXXXXXXRXRRXXX, preferably a polypeptide comprising the amino acid sequence of XXXXXSXXRXRRXXX (SEQ ID NO: 780), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 of the amino acids at positions 1-8, 10 and 13-15 of SEQ ID NO: 781 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 of the amino acids at positions 1-5, 7, 8, 10 and 13-15 of XXX XXXXEEEXX are substituted by other amino acids;

[1464] (19β-4) a polypeptide comprising the amino acid sequence of XXXXXXXXXXXEEEXX, preferably a polypeptide comprising the amino acid sequence of XKEEEHQEXXEEEQX (SEQ ID NO: 783), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 of the amino acids at positions 1-10, 14 and 15 of SEQ ID NO: 784 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 1, 9, 10 and 15 of SEQ ID NO: 784 are substituted by other amino acids;

[1465] (19β-5) a polypeptide comprising the amino acid sequence of XXGKXXXXXXXXXXX, preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or 13 of the amino acids at positions 1, 2 and 5-15 of SEQ ID NO: 786 are substituted by alanine, more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or 13 of the amino acids at positions 1, 2 and 5-15 of SEQ ID NO: 786 are substituted by other amino acids.

(20β) Epitope of Seed Biotin-Containing Protein SBP65 (Spot 20α)

[1466] In the present invention, the polypeptide (20β) can be a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 788-792.

[1467] In one mode, the polypeptide (20) may not comprise the full-length amino acid sequence (SEQ ID NO: 12) of Seed biotin-containing protein SBP65 or a variant or a homolog having 100% or at least 90%, at least 80% or at least 70% identity to the total length of this amino acid sequence.

[1468] SEQ ID NOs: 788-792 correspond to amino acids 131-145 (SEQ ID NO: 788), amino acids 251-258 (SEQ ID NO: 789), amino acids 281-295 (SEQ ID NO: 790), amino acids 441-448 (SEQ ID NO: 791), and amino acids 531-545 (SEQ ID NO: 792), respectively, of SEQ ID NO: 12.

(21β) Epitope of Ribulose Bisphosphate Carboxylase Large Chain (Spot 21α)

[1469] In the present invention, the polypeptide (21β) can be a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 793-802.

[1470] In one mode, the polypeptide (21β) may not comprise the full-length amino acid sequence (SEQ ID NO: 4) of Ribulose bisphosphate carboxylase large chain or a variant or a homolog having 100% or at least 90%, at least 80% or at least 70% identity to the total length of this amino acid sequence.

[1471] SEQ ID NOs: 793-802 correspond to amino acids 31-34 (SEQ ID NO: 793), amino acids 101-108 (SEQ ID NO: 794), amino acids 201-209 (SEQ ID NO: 795), amino acids 221-235 (SEQ ID NO: 796), and amino acids 237-245 (SEQ ID NO: 797), and amino acids 291-295 (SEQ ID NO: 798), amino acids 351-365 (SEQ ID NO: 799), amino acids 411-417 (SEQ ID NO: 800), amino acids 431-445 (SEQ ID NO: 801), and amino acids 461-475 (SEQ ID NO: 802), respectively, of SEQ ID NO: 4.

(22β) Epitope of Gamma-Glutamyl Hydrolase (Spot 22α)

[1472] In the present invention, the polypeptide (22β) can be a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 803-808.

[1473] In one mode, the polypeptide (22) may not comprise the full-length amino acid sequence (SEQ ID NO: 6) of Gamma-glutamyl hydrolase or a variant or a homolog having 100% or at least 90%, at least 80% or at least 70% identity to the total length of this amino acid sequence.

[1474] SEQ ID NOs: 803-808 correspond to amino acids 61-75 (SEQ ID NO: 803), amino acids 161-175 (SEQ ID NO: 804), amino acids 241-255 (SEQ ID NO: 805), amino acids 261-275 (SEQ ID NO: 806), amino acids 288-294

(SEQ ID NO: 807), and amino acids 301-315 (SEQ ID NO: 808), respectively, of SEQ ID NO: 6.

(23β) Epitope of Guanine Nucleotide-Binding Protein Subunit Beta-Like Protein (Spot 23α)

[1475] In the present invention, the polypeptide (23) can be a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 809-812.

[1476] In one mode, the polypeptide (23β) may not comprise the full-length amino acid sequence (SEQ ID NO: 8) of Guanine nucleotide-binding protein subunit beta-like protein or a variant or a homolog having 100% or at least 90%, at least 80% or at least 70% identity to the total length of this amino acid sequence.

[1477] SEQ ID NOs: 809-812 correspond to amino acids 41-55 (SEQ ID NO: 809), amino acids 111-125 (SEQ ID NO: 810), amino acids 227-235 (SEQ ID NO: 811), and amino acids 270-278 (SEQ ID NO: 812), respectively, of SEQ ID NO: 8.

(24β) Epitope of Protein SLE3 (Spot 24α)

[1478] In the present invention, the polypeptide (24β) can be a polypeptide comprising the amino acid sequence of SEQ ID NO: 813.

[1479] In one mode, the polypeptide (24β) may not comprise the full-length amino acid sequence (SEQ ID NO: 10) of Protein SLE3 or a variant or a homolog having 100% or at least 90%, at least 80% or at least 70% identity to the total length of this amino acid sequence.

[1480] SEQ ID NO: 813 corresponds to amino acids 71-85 of SEQ ID NO: 10.

(25β) Epitope of Succinyl-CoA Ligase [ADP-Forming] Subunit Beta, Mitochondrial (Spot 25α)

[1481] In the present invention, the polypeptide (25β) can be a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 814 and 815.

(26) Epitope of Phosphoglycerate Kinase, Cytosolic (Spot 26α)

[1482] In the present invention, the polypeptide (26β) can be any polypeptide selected from the group consisting of (26β-1) and (26β-2) as defined below:

[1483] (26β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 816 and 819; and

[1484] (26β-2) a polypeptide comprising the amino acid sequence of FDKXXXX, preferably a polypeptide comprising the amino acid sequence of FDKFXXX (SEQ ID NO: 818), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 4 to 7 of SEQ ID NO: 819 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2 or 3 of the amino acids at positions 5 to 7 of SEQ ID NO: 819 are substituted by other amino acids.

(27) Epitope of Alcohol Dehydrogenase 1 (Spot 27α)

[1485] In the present invention, the polypeptide (27β) can be a polypeptide comprising the amino acid sequence of SEQ ID NO: 820.

(28β) Epitope of 35 kDa Seed Maturation Protein Isoform X1 (Spot 28α)

[1486] In the present invention, the polypeptide (28β) can be any polypeptide selected from the group consisting of (28β-1) to (28β-4) as defined below:

[1487] (28β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 823, 826, 829 and 830-835;

[1488] (28β-2) a polypeptide comprising the amino acid sequence of XQXXXXENKP (SEQ ID NO: 821), preferably a polypeptide comprising the amino acid sequence of EQXQXXENKP (SEQ ID NO: 822), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 1 and 3-6 of SEQ ID NO: 823 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2 or 3 of the amino acids at positions 3, 5 and 6 of SEQ ID NO: 823 are substituted by other amino acids;

[1489] (28β-3) a polypeptide comprising the amino acid sequence of XKXXXXEKXXXXXXX, preferably a polypeptide comprising the amino acid sequence of XKDXXXEKXXXXXDX (SEQ ID NO: 825), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 of the amino acids at positions 1, 3-6 and 9-15 of SEQ ID NO: 826 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 of the amino acids at positions 1, 4-6, 9-13 and 15 of SEQ ID NO: 826 are substituted by other amino acids;

[1490] (28β-4) a polypeptide comprising the amino acid sequence of XXXDXXREEXXXXRX (SEQ ID NO: 827), preferably a polypeptide comprising the amino acid sequence of XXXDXXREEEEXXRX (SEQ ID NO: 828), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 of the amino acids at positions 1-3, 5, 6, 10-13 and 15 of SEQ ID NO: 829 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7 or 8 of the amino acids at positions 1-3, 5, 6, 12, 13 and 15 of SEQ ID NO: 829 are substituted by other amino acids.

(29β) Epitope of Glyceraldehyde-3-Phosphate Dehydrogenase Isoform X1 (Spot 29α, 30α, 31α, 32α, 33α)

[1491] In the present invention, the polypeptide (29β) can be any polypeptide selected from the group consisting of (29β-1) to (29β-2) as defined below:

[1492] (28β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 838 and 839;

[1493] (28β-2) a polypeptide comprising the amino acid sequence of XXXXXXXDKP, preferably a polypeptide comprising the amino acid sequence of XXXXXFXDKP (SEQ ID NO: 837), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6 or 7 of the amino acids at

positions 1-7 of SEQ ID NO: 838 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 1-5 and 7 of SEQ ID NO: 838 are substituted by other amino acids.

(34β) Epitope of Bifunctional UDP-Glucose 4-Epimerase and UDP-Xylose 4-Epimerase 1-Like (Spot 34α)

[1494] In the present invention, the polypeptide (34β) can be any polypeptide selected from the group consisting of (34β-1) to (34β-2) as defined below:

[1495] (34β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 842 and 843-846;

[1496] (34β-2) a polypeptide comprising the amino acid sequence of XXXXXXXXXXXDDXE, preferably a polypeptide comprising the amino acid sequence of LXXXXXXXXXXDDXE (SEQ ID NO: 841), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 of the amino acids at positions 1-11 and 14 of SEQ ID NO: 842 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 of the amino acids at positions 2-11 and 14 of SEQ ID NO: 842 are substituted by other amino acids.

(35β) Epitope of Elongation Factor 1-Alpha (Spot 35α)

[1497] In the present invention, the polypeptide (35β) can be any polypeptide selected from the group consisting of (35β-1) to (35β-4) as defined below:

[1498] (35β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 849, 850, 853 and 856;

[1499] (35β-2) a polypeptide comprising the amino acid sequence of XXXXDKXXXERXXXE (SEQ ID NO: 847), preferably a polypeptide comprising the amino acid sequence of XXXXDKRVXERXXXE (SEQ ID NO: 848), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 of the amino acids at positions 1-4, 7-9 and 12-14 of SEQ ID NO: 849 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7 or 8 of the amino acids at positions 1-4, 9 and 12-14 of SEQ ID NO: 849 are substituted by other amino acids;

[1500] (35β-3) a polypeptide comprising the amino acid sequence of PISXFEXXX (SEQ ID NO: 851), preferably a polypeptide comprising the amino acid sequence of PISGFEXXX (SEQ ID NO: 852), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 4 and 7-9 of SEQ ID NO: 853 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2 or 3 of the amino acids at positions 7-9 of SEQ ID NO: 853 are substituted by other amino acids;

[1501] (35β-4) a polypeptide comprising the amino acid sequence of XXEXLDXXXX, preferably a polypeptide comprising the amino acid sequence of LLEXLDXXXX (SEQ ID NO: 855), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6 or 7 of the amino acids at positions 1, 2, 4 and 7-10 of SEQ ID NO: 856 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 4 and 7-10 of SEQ ID NO: 856 are substituted by other amino acids.

(36β) Epitope of Seed Maturation Protein PM34 (Spot 36α, 37α)

[1502] In the present invention, the polypeptide (36β) can be any polypeptide selected from the group consisting of (36β-1) to (36β-2) as defined below:

[1503] (36β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 857, 858 and 861;

[1504] (36β-2) a polypeptide comprising the amino acid sequence of KXXXXQQQ (SEQ ID NO: 859), preferably a polypeptide comprising the amino acid sequence of KXXXQQQQ (SEQ ID NO: 860), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 2-5 of SEQ ID NO: 861 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2 or 3 of the amino acids at positions 2-4 of SEQ ID NO: 861 are substituted by other amino acids.

(38β) Epitope of Seed Biotin-Containing Protein SBP65-Like Isoform X1 (Spot 38α)

[1505] In the present invention, the polypeptide (38β) can be any polypeptide selected from the group consisting of (38β-1) to (38β-13) as defined below:

[1506] (38β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 864, 867, 869, 871, 874, 877, 880, 882, 884, 887, 889 and 892;

[1507] (38β-2) a polypeptide comprising the amino acid sequence of RGKXXXX, preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 4-7 of SEQ ID NO: 864 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 4-7 of SEQ ID NO: 864 are substituted by other amino acids;

[1508] (38β-3) a polypeptide comprising the amino acid sequence of XRGKXXXXXX, preferably a polypeptide comprising the amino acid sequence of RRGKXXXXXX (SEQ ID NO: 866), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6 or 7 of the amino acids at positions 1 and 5-10 of SEQ ID NO: 867 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 5-10 of SEQ ID NO: 867 are substituted by other amino acids;

[1509] (38β-4) a polypeptide comprising the amino acid sequence of RXKXXXXXXXXXXXE, preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 of the amino acids at positions 2 and 4-14 of SEQ ID NO: 869 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 of the amino

acids at positions 2 and 4-14 of SEQ ID NO: 869 are substituted by other amino acids;

[1510] (38β-5) a polypeptide comprising the amino acid sequence of XXXXXXXRRXXXXXR, preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 of the amino acids at positions 1-7 and 10-14 of SEQ ID NO: 871 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 of the amino acids at positions 1-7 and 10-14 of SEQ ID NO: 871 are substituted by other amino acids;

[1511] (38β-6) a polypeptide comprising the amino acid sequence of XXXKGRVVXEXXXXX (SEQ ID NO: 872), preferably a polypeptide comprising the amino acid sequence of LPDKGRVVXESEKKE (SEQ ID NO: 873), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 1-3, 9 and 11-15 of SEQ ID NO: 874 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 1-3, 9 and 11-15 of SEQ ID NO: 874 are substituted by other amino acids;

[1512] (38β-7) a polypeptide comprising the amino acid sequence of ATKAKDXXREXXXXX (SEQ ID NO: 875), preferably a polypeptide comprising the amino acid sequence of ATKAKDVTREXXXXX (SEQ ID NO: 876), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6 or 7 of the amino acids at positions 7, 8 and 11-15 of SEQ ID NO: 877 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 11-15 of SEQ ID NO: 877 are substituted by other amino acids;

[1513] (38β-8) a polypeptide comprising the amino acid sequence of XGXKGQXKKXXXXXX (SEQ ID NO: 878), preferably a polypeptide comprising the amino acid sequence of XGXKGQTKKXXGEEQ (SEQ ID NO: 879), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 1, 3, 7 and 10-15 of SEQ ID NO: 880 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 1, 3, 10 and 11 of SEQ ID NO: 880 are substituted by other amino acids;

[1514] (38β-9) a polypeptide comprising the amino acid sequence of RXXXETKEGX (SEQ ID NO: 881), preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 2-4 and 10 of SEQ ID NO: 882 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 2-4 and 10 of SEQ ID NO: 882 are substituted by other amino acids;

[1515] (38β-10) a polypeptide comprising the amino acid sequence of XXXXNKQSSL (SEQ ID NO: 883), preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 1~4 of SEQ ID NO: 884 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 1~4 of SEQ ID NO: 884 are substituted by other amino acids;

[1516] (38β-11) a polypeptide comprising the amino acid sequence of XKXVAXXXXXXXXEX (SEQ ID NO: 885), preferably a polypeptide comprising the amino acid sequence of XKXVAEXXXQXASEL (SEQ ID NO: 886), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 of the amino acids at positions 1, 3, 6-13 and 15 of SEQ ID NO: 887 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 1, 3, 7-9 and 11 of SEQ ID NO: 887 are substituted by other amino acids;

[1517] (38β-12) a polypeptide comprising the amino acid sequence of XXKVEAEXXX (SEQ ID NO: 888), preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 1, 2 and 8-10 of SEQ ID NO: 889 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 1, 2 and 8-10 of SEQ ID NO: 889 are substituted by other amino acids;

[1518] (38β-13) a polypeptide comprising the amino acid sequence of XXXXIAEIAE (SEQ ID NO: 890), preferably a polypeptide comprising the amino acid sequence of XXETIAEIAE (SEQ ID NO: 891), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 1~4 of SEQ ID NO: 892 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1 or 2 of the amino acids at positions 1 and 2 of SEQ ID NO: 892 are substituted by other amino acids.

(39β) Epitope of ATP Synthase Subunit O, Mitochondrial-Like (Spot 39α)

[1519] In the present invention, the polypeptide (39β) can be any polypeptide selected from the group consisting of (39β-1) to (39β-2) as defined below:

[1520] (39β-1) a polypeptide comprising the amino acid sequence of SEQ ID NOs: 895;

[1521] (39β-2) a polypeptide comprising the amino acid sequence of XYXXSXKXXXXEKVX (SEQ ID NO: 893), preferably a polypeptide comprising the amino acid sequence of XYIXSXKXXXXEKVX (SEQ ID NO: 894), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 1, 3, 4, 6, 8-11 and 15 of SEQ ID NO: 895 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7 or 8 of the amino acids at positions 1, 4, 6, 8-11 and 15 of SEQ ID NO: 895 are substituted by other amino acids.

(40β) Epitope of P24 Oleosin Isoform a (Spot 40α, 56α)

[1522] In the present invention, the polypeptide (40β) can be any polypeptide selected from the group consisting of (40β-1) to (40β-3) as defined below:

[1523] (40β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 898, 899 and 902;

[1524] (40β-2) a polypeptide comprising the amino acid sequence of YEXGV (SEQ ID NO: 896), preferably a polypeptide comprising the amino acid sequence of YEXGV (SEQ ID NO: 897), more preferably a polypeptide comprising an amino acid sequence in which any 1 or 2 of the amino acids at positions 3 and 6 of SEQ ID NO: 898 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1 or 2 of the amino acids at positions 3 and 6 of SEQ ID NO: 898 are substituted by other amino acids;

[1525] (40β-3) a polypeptide comprising the amino acid sequence of XXXXXVTAX, preferably a polypeptide comprising the amino acid sequence of VXXXXVTAX (SEQ ID NO: 901), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 1-5 and 9 of SEQ ID NO: 902 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 2-5 and 9 of SEQ ID NO: 902 are substituted by other amino acids.

(41β) Epitope of Uncharacterized Protein At3g49720-Like Isoform X2 (Spot 41α)

[1526] In the present invention, the polypeptide (41β) can be a polypeptide comprising the amino acid sequence of SEQ ID NO: 903.

(42β) Epitope of Superoxide Dismutase [Mn], Mitochondrial (Spot 42α, 43α)

[1527] In the present invention, the polypeptide (42) can be any polypeptide selected from the group consisting of (42β-1) to (42β-2) as defined below:

[1528] (42β-1) a polypeptide comprising the amino acid of SEQ ID NOs: 906;

[1529] (42β-2) a polypeptide comprising the amino acid sequence of XXWXXXXXXXEKEXX (SEQ ID NO: 904), preferably a polypeptide comprising the amino acid sequence of XXWXXXXEXXEKESX (SEQ ID NO: 905), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 of the amino acids at positions 1, 2, 4-10, 14 and 15 of SEQ ID NO: 906 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8 or 9 of the amino acids at positions 1, 2, 4-7, 9, 10 and 15 of SEQ ID NO: 906 are substituted by other amino acids.

(44) Epitope of Ferritin-2 (Spot 44α)

[1530] In the present invention, the polypeptide (44β) can be any polypeptide selected from the group consisting of (44β-1) to (44β-2) as defined below:

[1531] (44β-1) a polypeptide comprising the amino acid of SEQ ID NOs: 909;

[1532] (44β-2) a polypeptide comprising the amino acid sequence of XXXXEEEREHAXQLI (SEQ ID NO: 907), preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 1~4 and 12 of SEQ ID NO: 909 are substituted by alanine, more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 1~4 and 12 of SEQ ID NO: 909 are substituted by other amino acids.

[1533] In one mode, the polypeptide (44β) may not comprise the full-length amino acid sequence (SEQ ID NO: 14) of Ferritin-2 or a variant or a homolog having 100% or at least 90%, at least 80% or at least 70% identity to the total length of this amino acid sequence.

[1534] SEQ ID NO: 909 corresponds to amino acids 131-145 of SEQ ID NO: 14.

(45β) Epitope of Uncharacterized Protein LOC100499771 (Spot 45α)

[1535] In the present invention, the polypeptide (45β) can be a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 910-912.

(46β) Epitope of Seed Maturation Protein PM22 (Spot 46α)

[1536] In the present invention, the polypeptide (46β) can be any polypeptide selected from the group consisting of (46β-1) to (46β-2) as defined below:

[1537] (46β-1) a polypeptide comprising the amino acid of SEQ ID NOs: 915;

[1538] (46β-2) a polypeptide comprising the amino acid sequence of XDXDXXXDX, preferably a polypeptide comprising the amino acid sequence of XDIDXXXDX (SEQ ID NO: 914), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 1, 3, 5-7 and 9 of SEQ ID NO: 915 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 1 and 5-7 of SEQ ID NO: 915 are substituted by other amino acids.

(47β) Epitope of Eukaryotic Translation Initiation Factor 5A-2 (Spot 47α)

[1539] In the present invention, the polypeptide (47β) can be a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 916-918.

(48β) Epitope of Uncharacterized Protein LOC100306570 (Spot 48α)

[1540] In the present invention, the polypeptide (48β) can be any polypeptide selected from the group consisting of (48β-1) to (48β-2) as defined below:

[1541] (48β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 921 and 922;

[1542] (48β-2) a polypeptide comprising the amino acid sequence of XXXXXXFDK, preferably a polypeptide comprising the amino acid sequence of XXXXPXFDK (SEQ ID NO: 920), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 1-6 of SEQ ID NO: 921 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 1~4 and 6 of SEQ ID NO: 921 are substituted by other amino acids.

(49β) Epitope of Bowman-Birk Type Proteinase Inhibitor D-II (Spot 49α)

[1543] In the present invention, the polypeptide (49β) can be a polypeptide comprising the amino acid sequence of SEQ ID NO: 923.

[1544] In one mode, the polypeptide (49β) may not comprise the full-length amino acid sequence (SEQ ID NO: 2) of Bowman-Birk type proteinase inhibitor D-II or a variant or a homolog having 100% or at least 90%, at least 80% or at least 70% identity to the total length of this amino acid sequence.

[1545] SEQ ID NO: 923 corresponds to amino acids 79-83 of SEQ ID NO: 2.

(50β) Epitope of Uncharacterized Protein LOC100813859 (Spot 50)

[1546] In the present invention, the polypeptide (50β) can be any polypeptide selected from the group consisting of (50β-1) to (50β-2) as defined below:

[1547] (50β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 924 and 927;

[1548] (50β-2) a polypeptide comprising the amino acid sequence of XXXXKQXXX, preferably a polypeptide comprising the amino acid sequence of XXDXKQXXX, more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6 or 7 of the amino acids at positions 1~4 and 7-9 of SEQ ID NO: 927 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 1, 2, 4 and 7-9 of SEQ ID NO: 927 are substituted by other amino acids.

(51β) Epitope of 14-3-3-Like Protein (Spot 51α)

[1549] In the present invention, the polypeptide (51β) can be a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 928-930.

(52β) Epitope of Probable Fructokinase-4 (Spot 52α)

[1550] In the present invention, the polypeptide (52β) comprising an epitope can be any polypeptide selected from the group consisting of (52β-1) and (52β-2) as defined below:

[1551] (52β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 933 and 934-938; and

[1552] (52β-2) a polypeptide comprising the amino acid sequence of XXXXXXSXXEXLXXX, preferably a polypeptide comprising the amino acid sequence of XXGXXXSXXEXLXXX (SEQ ID NO: 932), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 of the amino acids at positions 1-6, 8, 9, 11, and 13-15 of SEQ ID NO: 933 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 of the amino acids at positions 1, 2, 4-6, 8, 9, 11, and 13-15 of SEQ ID NO: 933 are substituted by other amino acids.

(53β) Epitope of Triosephosphate Isomerase Isoform X1 (Spot 53α)

[1553] In the present invention, the polypeptide (53) can be a polypeptide comprising the amino acid sequence of SEQ ID NO: 939.

(54β) Epitope of Superoxide Dismutase [Fe], Chloroplastic Isoform X1 (Spot 54α)

[1554] In the present invention, the polypeptide (54β) can be any polypeptide selected from the group consisting of (54β-1) to (54β-2) as defined below:

[1555] (54β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 940 and 943;

[1556] (54β-2) a polypeptide comprising the amino acid sequence of KXXXXTX, preferably a polypeptide comprising the amino acid sequence of KXXXXTQ, more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 2-5 and 7 of SEQ ID NO: 943 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 2-5 of SEQ ID NO: 943 are substituted by other amino acids.

(55β) Epitope of 51 kDa Seed Maturation Protein Precursor (Spot 55α)

[1557] In the present invention, the polypeptide (55β) can be any polypeptide selected from the group consisting of (55β-1) to (55β-4) as defined below:

[1558] (55β-1) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 946, 949 and 952;

[1559] (55β-2) a polypeptide comprising the amino acid sequence of TKXXXSDXX (SEQ ID NO: 944), preferably a polypeptide comprising the amino acid sequence of TKDXXSDXX (SEQ ID NO: 945), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 3-5, 8 and 9 of SEQ ID NO: 946 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3 or 4 of the amino acids at positions 4, 5, 8 and 9 of SEQ ID NO: 946 are substituted by other amino acids;

[1560] (55β-3) a polypeptide comprising the amino acid sequence of VKDXXXDA (SEQ ID NO: 947), preferably a polypeptide comprising the amino acid sequence of VKDYXXDA (SEQ ID NO: 948), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2 or 3 of the amino acids at positions 4-6 of SEQ ID NO: 949 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1 or 2 of the amino acids at positions 5 and 6 of SEQ ID NO: 949 are substituted by other amino acids;

[1561] (55β-4) a polypeptide comprising the amino acid sequence of KDXXXXXXQ, preferably a polypeptide comprising the amino acid sequence of KDXXXXXAQ (SEQ ID NO: 951), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 3-8 of SEQ ID NO: 952 are substituted by alanine, further preferably a polypeptide comprising an

amino acid sequence in which any 1, 2, 3, 4 or 5 of the amino acids at positions 3-7 of SEQ ID NO: 952 are substituted by other amino acids.

(57β) Epitope of Gly m 4

[1562] In the present invention, the polypeptide (57β) can be any polypeptide selected from the group consisting of (56β-1) to (56β-2) as defined below:

[1563] (57β-1) a polypeptide comprising the amino acid of SEQ ID NOs: 955;

[1564] (57β-2) a polypeptide comprising the amino acid sequence of XXXXXXXEXYXLXHP (SEQ ID NO: 953), preferably a polypeptide comprising the amino acid sequence of XXXXXXIEAYLLAHP (SEQ ID NO: 954), more preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 of the amino acids at positions 1-7, 9, 11 and 13 of SEQ ID NO: 955 are substituted by alanine, further preferably a polypeptide comprising an amino acid sequence in which any 1, 2, 3, 4, 5 or 6 of the amino acids at positions 1-6 of SEQ ID NO: 955 are substituted by other amino acids.

[1565] In one mode, the polypeptide (57) may not comprise the full-length amino acid sequence (UniProt accession number: P26987.1) of Gly m 4 or a variant or a homolog having 100% or at least 90%, at least 80% or at least 70% identity to the total length of this amino acid sequence.

[1566] SEQ ID NO: 955 corresponds to amino acids 141-155 of UniProt accession number: P26987.1.

(58) Epitopes of Gly m 5α and Gly m 5α'

[1567] In the present invention, the polypeptide (58β) can be a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 956-960.

[1568] In one mode, the polypeptide (58β) comprising an epitope may not comprise the full-length amino acid sequence (NCBI accession number: NP_001236856.1) of Gly m 5a or the full-length amino acid sequence (SEQ ID NO: NCBI accession number: NP_001237316.1) of Gly m 5α', or a variant or a homolog having 100% or at least 90%, at least 80% or at least 70% identity to the total length of any of these amino acid sequences.

[1569] SEQ ID NOs: 956-960 correspond to amino acids 173-181 (SEQ ID NO: 956), amino acids 294-301 (SEQ ID NO: 957), amino acids 301-309 (SEQ ID NO: 958), amino acids 414-422 (SEQ ID NO: 959), and amino acids 294-301 (SEQ ID NO: 960), respectively, of NCBI accession number: NP_001236856.1.

(59β) Epitopes of Gly m 6.01 and Gly m 6.02

[1570] In the present invention, the polypeptide (59) can be a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 961 and 962.

[1571] In one mode, the polypeptide (59β) may not comprise the full-length amino acid sequence (UniProt accession number: CAA33215.1) of Gly m 6.01 or the full-length amino acid sequence (UniProt accession number: BAA00154.1) of Gly m 6.02, or a variant or a homolog having 100% or at least 90%, at least 80% or at least 70% identity to the total length of any of these amino acid sequences.

[1572] SEQ ID NOs: 961 and 962 correspond to amino acids 201-215 and amino acids 451-465, respectively, of UniProt accession number: CAA33215.1.

(60β) Epitope of Gly m 8

[1573] In the present invention, the polypeptide (60β) can be a polypeptide comprising the amino acid sequence of SEQ ID NO: 963.

[1574] In one mode, the polypeptide (60β) may not comprise the full-length amino acid sequence (UniProt accession number: AAB71140.1) of Gly m 8 or a variant or a homolog having 100% or at least 90%, at least 80% or at least 70% identity to the total length of this amino acid sequence.

[1575] SEQ ID NO: 963 corresponds to amino acids 71-85 of UniProt accession number: AAB71140.1.

[1576] The residue of X as to the aforementioned polypeptides (1β) to (60β) is an amino acid residue at a site where binding activity against IgE antibodies from allergic patients was maintained even after substitution by alanine in alanine scanning described in Example 4. It is well known to those skilled in the art that even when such a site is substituted by any other amino acids, it is highly probable that this binding activity against IgE antibodies is maintained.

[1577] The lengths of the aforementioned polypeptides (1β) to (60β) are not particularly limited. In a preferred mode, the lengths of the aforementioned polypeptides (1β) to (60β) can be at least 500 amino acids, at least 300 amino acids, at least 200 amino acids, at least 100 amino acids, at least 50 amino acids, at least 30 amino acids, at least 20 amino acids, at least 10 amino acids, or at least 5 amino acids.

[1578] The aforementioned polypeptides (1β) to (60β) may be prepared by a technique of chemical synthesis such as solid-phase peptide synthesis. Alternatively, polypeptides comprising an epitope may be obtained by expressing them as recombinant proteins using a genetic recombination technique well known to those skilled in the art and by separating and purifying them using protein purification methods well known to those skilled in the art.

Diagnosis Kit and Method (2)

[1579] The present invention provides a method for providing an indicator for diagnosing an allergy in a subject, the method comprising the steps of:

[1580] (i) contacting a sample obtained from the subject with an antigen, wherein the sample is a solution comprising an Ig antibody;

[1581] (ii) detecting binding between the IgE antibody present in the sample obtained from the subject and the antigen; and

[1582] (iii) when the binding between the IgE antibody in the subject and the antigen is detected, an indicator of the fact that the subject is allergic is provided; wherein the antigen is at least one of the aforementioned polypeptides (1β) to (60), or a polypeptide in which two or more of the aforementioned polypeptides (1β) to (60β) are joined together via or without a spacer.

[1583] Hereinafter, the at least one of the aforementioned polypeptides (1β) to (60β), or the polypeptide in which two or more of the aforementioned polypeptides (1β) to (60β) are joined together via or without a spacer is referred to as the "antigen including (1β) to (60β)" in the present speci-

fication. The type of the spacer is not particularly limited, and an ordinary spacer that is used by those skilled in the art for joining together two or more peptides can be used. The spacer may be, for example, a hydrocarbon chain such as Acp (6)-OH.

[1584] The sample obtained from a subject is as described above in the subsection titled "Diagnosis kit and method (1)".

[1585] Detection of contact and binding between the sample obtained from a subject and the antigen can be carried out by using a known method described above in the subsection titled "Diagnosis kit and method (1)", such as ELISA (Enzyme-Linked Immunosorbent Assay), sandwich immunoassay, immunoblotting, immunoprecipitation, or immunochromatography.

[1586] The antigen may be provided as an isolated antigen including (1β) to (60β) in a state immobilized on a carrier. In this case, the steps (i) and (ii) mentioned above can be carried out using ELISA or sandwich immunoassay, or the like. In this respect, the step (i) mentioned above can be carried out by contacting the sample obtained from a subject with an antigen-immobilized surface.

[1587] The antigen including (1β) to (60β) specifically binds to IgE antibodies from allergic patients. Therefore, when binding between an IgE antibody from a subject and the antigen is detected, an indicator of the fact that the subject is allergic to soybean is provided.

[1588] The present invention further provides a kit for diagnosing an allergy, comprising at least one antigen including (1β) to (60β). The diagnosis kit of this invention may be used in the aforementioned method for providing an indicator for diagnosing an allergy or in a diagnosis method as described later. The diagnosis kit of this invention may comprise not only the at least one antigen including (1β) to (60β), but also an anti-IgE antibody labeled with an enzyme and a chromogenic or luminescent substrate serving as a substrate for the enzyme. Also, a fluorescent-labeled anti-IgE antibody may be used. In the diagnosis kit of this invention, the antigen may be provided in a state immobilized on a carrier. The diagnosis kit of this invention may also be provided together with instructions on the procedure for diagnosis or a package containing said instructions.

[1589] The present invention further provides a composition for diagnosing an allergy, comprising at least one antigen including (1β) to (60β). The diagnosis composition of this invention can be used in a diagnosis method as described below. The diagnosis composition of this invention may further comprise a pharmaceutically acceptable carrier and/or additives commonly used with the antigen of this invention depending on the need.

[1590] In one mode, the present invention provides a method for diagnosing an allergy in a subject, the method comprising:

[1591] (i) contacting a sample obtained from the subject with an antigen;

[1592] (ii) detecting binding between an IgE antibody present in the sample obtained from the subject and the antigen; and

[1593] (iii) when the binding between the IgE antibody in the subject and the antigen is detected, diagnosing the subject as being allergic to a soybean;

[1594] wherein the antigen is at least one of the aforementioned polypeptides (1β) to (60β), or a polypeptide in which two or more of the aforementioned polypeptides (1β) to (60β) are joined together via or without a spacer (i.e., the antigen including (1β) to (60β)). In this method, the steps (i) and (ii) are performed as described above regarding the corresponding steps of the method for providing an indicator for diagnosing an allergy.

[1595] In another mode, the present invention provides a method for diagnosing an allergy in a subject, the method comprising administering to the subject at least one antigen including (1β) to (60β). This method may be performed in the form of a skin test characterized by applying the antigen onto the skin. The skin test includes forms of tests described above in the subsection titled "Diagnosis kit and method (1)". If a skin reaction such as swelling occurs in a skin portion to which the antigen has been applied, the subject of interest is diagnosed as having an allergy. The amount of the antigen to be applied to the skin in such tests can be, for example, not more than 100 μg per dose.

[1596] In yet another mode, the present invention provides at least one antigen including (1β) to (60), intended for use as an active ingredient for a load test for allergy diagnosis.

[1597] In still another mode, the present invention provides at least one antigen including (1β) to (60β), intended for use in the diagnosis of an allergy.

[1598] In still another mode, the present invention provides use of at least one antigen including (1β) to (60β) for the production of a diagnostic agent for an allergy.

[1599] In this subsection, the allergy to be diagnosed or detected can be allergies to allergens in which allergen components having the aforementioned polypeptides (1β) to (60β) are present. Thus, detection of the allergy can be detection of not only an allergy to a single allergen, but also allergies including cross-reactivity.

Pharmaceutical Composition and Treatment Method (2)

[1600] The present invention provides a pharmaceutical composition comprising at least one antigen including (1β) to (60β). In one mode, the aforementioned pharmaceutical composition is used for the treatment of an allergy.

[1601] The present invention also provides a method for treating an allergy, the method comprising administering at least one antigen including (1β) to (60β) to a patient in need of a treatment for an allergy.

[1602] In another mode, the present invention provides at least one antigen including (1β) to (60), intended for use in the treatment for an allergy. In yet another mode, the present invention provides use of at least one antigen including (1β) to (60β) for the production of a therapeutic agent for an allergy.

[1603] In still another mode, the at least one antigen including (1β) to (60β) can be used as an active ingredient for hyposensitization therapy for an allergy. Here, the antigen protein to be used in the hyposensitization therapy may be a polypeptide that has been expressed and purified and may be a polypeptide that has been expressed in food such as rice, such as rice-based vaccine expressing pollen allergens.

[1604] The administration route, administration dose, frequency and/or period of the pharmaceutical composition of this invention, and other ingredients to be contained in the pharmaceutical composition, and the dosage form can be as described above in the subsection titled "Pharmaceutical composition and treatment method (1)". In the case of using

the antigen including (1β) to (60β), for example, the dose to an adult patient may be a dose of not more than 100 μg per dose.

[1605] In this subsection, the allergy to be treated can be allergies to allergens in which allergen components having the aforementioned polypeptides (1β) to (60β) are present. Thus, detection of the allergy can be detection of not only an allergy to a single allergen, but also allergies including cross-reactivity.

Tester (2)

[1606] The present invention provides a tester comprising an antibody for at least one antigen including (1β) to (60β).

[1607] The antibody can be prepared by a conventional method. For example, the antibody may be prepared by immunizing a mammal such as rabbit with the aforementioned antigen including (1β) to (60β). The antibody may be an Ig antibody, a polyclonal antibody, a monoclonal antibody, or an antigen-binding fragment thereof (e.g., Fab, F(ab') 2, Fab').

[1608] Further, in the aforementioned tester, the antibody may be provided in a form bound to a carrier. The type of the carrier is not particularly limited as long as it is usable for detection of binding between an antibody and the antigen including (1β) to (60β). Any given carrier known to those skilled in the art can be used.

[1609] In one mode, the aforementioned tester can be used for the purposes described above in the subsection titled "Tester (1)".

[1610] A method for determining the presence or absence of an antigen is as described above in the subsection titled "Tester (1)".

[1611] Another mode of the present invention includes a tester for determining the presence or absence of an antigen of an allergy in an object of interest, the tester comprising a primer appropriate for an epitope. The primer is not limited and may be designed so as to comprise, for example, a portion of the nucleotide sequence of a nucleic acid encoding any of the amino acid sequences defined above in (1β) to (60β), or a complementary strand thereof. Alternatively, the primer may be designed so as to be the nucleotide sequence of a region upstream or downstream of a portion encoding an epitope that is any of the amino acid sequences defined above in (1β) to (60β), in a nucleic acid encoding a protein comprising the epitope, or a complementary strand thereof. Here, the position of the epitope in the full-length sequence of an antigen is as defined above in the subsections titled "Antigen" and "Epitope of antigen". Particularly, when mRNA is of interest, the tester has a complementary primer of a poly-A tail.

[1612] For example, DNA is amplified by PCR (Polymerase Chain Reaction) including RT-PCR using templated DNA or mRNA obtained from a sample and the aforementioned primer, and the presence or absence of a nucleic acid encoding the amino acid sequences defined above in (1β) to (60β) in the sequence of the amplified DNA is determined to determine the presence or absence of the antigen. Amplification methods by PCR can be exemplified by RACE.

Allergen-Free Food and the Like (2)

[1613] The present invention provides a food raw material or an edible processed product in which at least one antigen including (1β) to (60β) is eliminated or reduced.

[1614] The method for eliminating or reducing the antigen of the present invention in a food raw material or an edible processed product is not limited. The techniques described above in the subsection titled "Allergen-free food and the like (1)" may be used. Elimination or reduction of at least one antigen including (1β) to (60β) may be achieved by eliminating or reducing the whole antigen or may be achieved by cleaving or removing the polypeptides defined in (1) to (60β) from the antigen protein. The "removal" includes deletion and modification.

[1615] The edible processed product in which the antigen of the present invention is eliminated or reduced may be a processed product of the food raw material in which the antigen of the present invention is eliminated or reduced, such as powdered milk obtained with purified amino acids as a raw material. In the case of using an ordinary food raw material, a treatment for removing or reducing the polypeptide of this invention is performed before or after preparation of an edible processed product. The techniques described in the subsection titled "Allergen-free food and the like (1)" may be used as methods for removing or reducing the antigen of the present invention in an edible processed product obtained with an ordinary food raw material.

Method for Producing Allergen-Free Food (2)

[1616] The present invention provides a method for producing an edible processed product in which an antigen is eliminated or reduced, the method comprising the step of confirming that the antigen is eliminated or reduced, in a production process of the processed product, wherein the antigen is at least one of the aforementioned antigen including (1β) to (60β) and antigens (1α) to (55α).

[1617] In the production method, elimination or reduction of an antigen means that at least one antigen including (1β) to (60β) is eliminated or reduced, or the polypeptides defined in (1) to (60β) are cleaved or removed from the antigen protein.

[1618] A technique of confirming that the antigen is eliminated or reduced in the production process of the edible processed product is not particularly limited, and any technique capable of detecting at least one of the aforementioned antigen including (1) to (60β) and antigens (1α) to (55α) may be used. For example, the presence or absence of the polypeptide or the antigen in the processed product may be confirmed from the binding activity of an antibody for at least one of the aforementioned antigen including (1β) to (60β) and antigens (1α) to (55α) against a sample containing a material resulting from the production process of the processed product. Details of such a method are as described above in the subsection titled "Diagnosis kit and method (1)". Thus, in the production method, the "IgE antibody from a subject" described above in the subsection titled "Diagnosis kit and method (1)" is replaced with the "antibody for at least one of the aforementioned polypeptides (1β) to (60) and antigens (1α) to (55α)", and the "antigen" described above in the subsection titled "Diagnosis kit and method (1)" is replaced with the "sample containing a material resulting from the production process of the edible processed product". The techniques described above in the subsection titled "Diagnosis kit and method (1)" can be used to confirm that the antigen is eliminated or reduced in the production process of the edible processed product.

EXAMPLES

[1619] The following describes examples of the present invention. The technical scope of this invention is not limited by these examples.

Example 1: Confirmation of a Protein Pattern

[1620] Proteins contained in soybean were investigated using a two-dimensional electrophoresis method described below.

Protein Extraction

[1621] Commercially available dry soybeans (several beans) were powdered. 500 μl of water was added to 50 mg of the powder and shaking extraction was conducted using a vortex mixer at 25° C. for 10 minutes. After the shaking extraction, a liquid protein extract was collected.

[1622] Thereafter, the precipitation procedure was repeated twice using a 2D-CleanUP Kit (produced by GE). In the first round of precipitation, the collected liquid protein extract was precipitated by adding TCA (trichloroacetic acid) thereto and the precipitated product produced by this procedure (TCA-precipitated product) was collected. In the second round of precipitation, the TCA-precipitated product collected above was further precipitated by adding acetone thereto and the precipitated product (sample) produced by this procedure was collected.

Preparation of a Sample Solution

[1623] Part of the collected sample (50 μg on a protein weight basis) was dissolved in 150 μL of a DeStreak Rehydration Solution (produced by GE), which is a swelling buffer for first-dimensional isoelectric focusing gels, thereby obtaining a sample solution for first-dimensional isoelectric focusing (sample solution for swelling). The constituents of the DeStreak Rehydration Solution are as mentioned below.

[1624] 7M thiourea
[1625] 2M urea
[1626] 4% (w/v) of CHAPS (3-[(3-cholamidopropyl) dimethylammonio]propanesulfonate)
[1627] 0.5% (v/v) IPG buffer; produced by GE
[1628] Moderate amount of BPB (bromophenol blue)

Penetration of the Sample into First-Dimensional Isoelectric Focusing Gels

[1629] First-dimensional isoelectric focusing gel strips (Immobiline Drystrip IPG gels (pH3-10NL); produced by GE) were immersed in 140 μL of the foregoing sample solution for first-dimensional isoelectric focusing (sample solution for swelling) and impregnated with the solution at room temperature overnight.

[1630] In this example, an IPGphor electrophoresis system produced by GE was used.

[1631] An electrophoresis tray was filled with silicone oil. Filter paper moisten with water was positioned at both ends of the gel strips impregnated with the sample, and the gel strips were set in the electrophoresis tray such that the gel strips were covered with silicone oil. Electrodes were placed on the gel strips with the filter paper intervening therebetween.

[1632] The maximum current of the isoelectric focusing system was set to 75 μA per gel strip, and the first-dimensional isoelectric focusing was carried out according to the following voltage program: (1) a constant voltage step was performed at a constant voltage of 300 V until the volt-hours reached 750 Vhr (the current variation width during electrophoresis for 30 minutes before the end of this step was 5 μA); (2) the voltage was increased gradually to 1000 V for 300 Vhr; (3) the voltage was further increased gradually to 5000 V for 4500 Vhr; and then (4) the voltage was held at a constant voltage of 5000 V until the total Vhr reached 12000.

SDS Equilibration of Isoelectric Focusing Gels

[1633] After the aforementioned first-dimensional isoelectric focusing was done, the gel strips were taken out of the isoelectric focusing system, immersed in an equilibration buffer containing a reducing agent, and shaken at room temperature for 15 minutes. The constituents of the equilibration buffer containing the reducing agent are as mentioned below.

[1634] 100 mM Tris-HCl (pH 8.0)
[1635] 6M urea
[1636] 30% (v/v) glycerol
[1637] 2% (w/v) SDS
[1638] 1% (w/v) DTT

[1639] Next, the equilibration buffer containing the reducing agent was removed, and then the gel strips were immersed in an equilibration buffer containing an alkylating agent and shaken at room temperature for 15 minutes to obtain SDS-equilibrated gels. The constituents of the equilibration buffer containing the alkylating agent are as mentioned below.

[1640] 100 mM Tris-HCl (pH 8.0)
[1641] 6M urea
[1642] 30% (v/v) glycerol
[1643] 2% (w/v) SDS
[1644] 2.5% (w/v) iodoacetamide

Second-Dimensional SDS-PAGE

[1645] In this example, the XCell SureLock Mini-Cell electrophoresis system produced by Life Technologies was used. The second-dimensional electrophoresis gels used were NuPAGE 4-12% Bis-Tris Gels produced by Life Technologies. Also, an electrophoresis buffer composed of the following constituents was prepared and used.

[1646] 50 mM MOPS
[1647] 50 mM Tris base
[1648] 0.1% (w/v) SDS
[1649] 1 mM EDTA

[1650] Further, an agarose solution for gel adhesion was used in this example, which was prepared by dissolving 0.5% (w/v) Agarose S (produced by Nippon Gene Co., Ltd.) and a moderate amount of BPB (bromophenol blue) in the electrophoresis buffer.

[1651] SDS-PAGE wells were washed well with the electrophoresis buffer, and then the buffer used for the washing was removed. Next, the washed wells were charged with the fully dissolved agarose solution for gel adhesion. Next, the SDS-equilibrated gel strips were immersed in agarose and closely adhered to second-dimensional electrophoresis gels using tweezers. After it was confirmed that agarose was fully fixed with the gels being closely adhered to each other, electrophoresis was performed at a constant voltage of 200 V for about 45 minutes.

Fluorescent Staining of Gels

[1652] The gels were fluorescently stained with SYPRO Ruby (produced by Life Technologies).

[1653] First, an airtight container to be used was washed well in advance with 98% (v/v) ethanol. The electrophoresed second-dimensional electrophoresis gel strips were taken out of the SDS-PAGE system, placed onto the washed airtight container, and treated twice by immersion in 50% (v/v) methanol and 7% (v/v) aqueous solution containing acetic acid for 30 minutes. Then, a further immersion treatment was done for 10 minutes, with the solution being replaced by water. Next, the second-dimensional electrophoresis gel strips were immersed in 40 mL of SYPRO Ruby and shaken at room temperature overnight. Thereafter, the SYPRO Ruby was removed, and then the second-dimensional electrophoresis gel strips were washed with water and shaken in 10% (v/v) methanol and 7% (v/v) aqueous solution containing acetic acid for 30 minutes. Further shaking was done for at least 30 minutes, with the solution being replaced by water.

Analysis

[1654] The second-dimensional electrophoresis gels obtained through the foregoing series of treatments were subjected to fluorescent image scanning on Typhoon 9400 (produced by GE). The results of the two-dimensional electrophoresis are shown in FIG. **1**. Molecular weight marker bands are found at the left of the panels. The positions of the bands denote particular molecular weights (in KDa).

Example 2: Identification of Antigens by Immunoblotting

[1655] Identification of antigens by immunoblotting was carried out by taking all the steps up to the step of "Second-dimensional SDS-PAGE" as described above in Example 1, followed by the steps of "Transfer to membrane", "Immunoblotting" and "Analysis" as described below.

Transfer to Membrane

[1656] Transfer to membrane was done using the following transfer system and transfer buffer.

[1657] Transfer system: XCell SureLock Mini-Cell and XCell II Blot Module (produced by Life Technologies)

[1658] Transfer buffer: NuPAGE Transfer Buffer (X20) (produced by Life Technologies), used in a form diluted 20-fold with milliQ water.

[1659] To be specific, proteins in the two-dimensional electrophoresis gels were transferred to a membrane (PVDF membrane) according to the following procedure.

[1660] (1) The PVDF membrane was immersed in 100% methanol followed by milliQ water, and then moved into the transfer buffer to hydrophilize the PVDF membrane.

[1661] (2) After sponge, filter paper, the gels treated by second-dimensional SDS-PAGE, the hydrophilized PVDF membrane, filter paper, and sponge were put in place in this order, the transfer system was energized at a constant voltage of 30 V for one hour.

Immunoblotting

[1662] Immunoblotting of the membrane was carried out using, as a primary antibody, a serum sample from a patient with a soybean allergy or a serum sample from a healthy subject.

[1663] Immunoblotting of the membrane was carried out according to the following procedure.

[1664] (1) The transferred membrane was shaken in a 5% skim milk/PBST solution (a PBS buffer containing 0.3% Tween 20 nonionic surfactant) at room temperature for one hour.

[1665] (2) The membrane was left to stand in a solution of 3% primary antibody serum in 5% skim milk/PBST at room temperature for one hour.

[1666] (3) The membrane was washed with a PBST solution (5 min.×3 times).

[1667] (4) The membrane was left to stand in a 1:2500 dilution of the secondary antibody, anti-human IgE-HRP (horseradish peroxidase), with a 3% skim milk/PBST solution at room temperature for one hour.

[1668] (5) The membrane was washed with a PBST solution (5 min.×3 times).

[1669] (6) The membrane was left to stand in Pierce Western Blotting Substrate Plus (produced by Thermo Fisher Scientific) for 5 minutes.

Analysis

[1670] The membrane obtained through the foregoing series of treatments was subjected to fluorescent image scanning on Typhoon 9400 (produced by GE).

[1671] The immunoblots obtained with the serum from the soybean-allergic patient were compared with those obtained with the control serum from the healthy subject. By immunoblotting using the serum from soybean-allergic Patient 1α, Patient 2α, Patient 3α, Patient 4α, Patient 5α, Patient 6α, Patient 7α, Patient 8α, and Patient 9α, 56 spots were detected (FIGS. **3**A to I), which are different from the spots detected when the serum of the healthy subject was used (FIG. **2**) and different from those of the known soybean allergen proteins.

[1672] The molecular weights and isoelectric points of the 56 spots are as follows.

[1673] Spot 1α: Molecular weight 80 to 110 kDa, pI 3.0 to 6.0

[1674] Spot 2α: Molecular weight 60 to 110 kDa, pI 3.0 to 6.0

[1675] Spot 3α: Molecular weight 100 to 150 kDa, pI 4.0 to 6.0

[1676] Spot 4α: Molecular weight 80 to 110 kDa, pI 4.0 to 7.0

[1677] Spot 5α: Molecular weight 80 to 160 kDa, pI 4.0 to 7.0

[1678] Spot 6α: Molecular weight 80 to 110 kDa, pI 4.0 to 7.0

[1679] Spot 7α: Molecular weight 60 to 110 kDa, pI 4.0 to 7.0

[1680] Spot 8α: Molecular weight 60 to 110 kDa, pI 4.0 to 8.0

[1681] Spot 9α: Molecular weight 60 to 110 kDa, pI 4.0 to 8.0

[1682] Spot 10α: Molecular weight 40 to 80 kDa, pI 4.0 to 7.0

[1683] Spot 11α: Molecular weight 40 to 80 kDa, pI 4.0 to 8.0

[1684] Spot 12α: Molecular weight 40 to 80 kDa, pI 4.0 to 8.0

[1685] Spot 13α: Molecular weight 40 to 80 kDa, pI 4.0 to 7.0

[1686] Spot 14α: Molecular weight 40 to 80 kDa, pI 3.0 to 6.0

[1687] Spot 15α, 16α: Molecular weight 30 to 80 kDa, pI 4.0 to 7.0

[1688] Spot 17α: Molecular weight 30 to 80 kDa, pI 4.0 to 7.0

[1689] Spot 18α: Molecular weight 40 to 80 kDa, pI 4.0 to 8.0

[1690] Spot 19α: Molecular weight 40 to 80 kDa, pI 4.0 to 8.0

[1691] Spot 20α: Molecular weight 60 to 75 kDa, pI 6.0 to 8.0

[1692] Spot 21α: Molecular weight 40 to 60 kDa, pI 5.0 to 8.0

[1693] Spot 22α: Molecular weight 25 to 37 kDa, pI 8.0 to 10.0

[1694] Spot 23α: Molecular weight 25 to 37 kDa, pI 7.0 to 9.0

[1695] Spot 24α: Molecular weight 5 to 20 kDa, pI 5.0 to 7.0

[1696] Spot 25α: Molecular weight 30 to 80 kDa, pI 4.0 to 7.0

[1697] Spot 26α: Molecular weight 30 to 80 kDa, pI 4.0 to 7.0

[1698] Spot 27α: Molecular weight 30 to 80 kDa, pI 4.0 to 7.0

[1699] Spot 28α: Molecular weight 20 to 60 kDa, pI 4.0 to 7.0

[1700] Spot 29α, 30α, 31α, 32α, 33α: Molecular weight 20 to 60 kDa, pI 4.0 to 8.0

[1701] Spot 34α: Molecular weight 20 to 60 kDa, pI 6.0 to 10.0

[1702] Spot 35α: Molecular weight 30 to 80 kDa, pI 6.0 to 10.0

[1703] Spot 36, 37α: Molecular weight 20 to 60 kDa, pI 4.0 to 8.0

[1704] Spot 38α: Molecular weight 80 to 160 kDa, pI 6.0 to 10.0

[1705] Spot 39α: Molecular weight 10 to 50 kDa, pI 6.0 to 10.0

[1706] Spot 40α, 56α: Molecular weight 10 to 50 kDa, pI 6.0 to 10.0

[1707] Spot 41α: Molecular weight 10 to 50 kDa, pI 6.0 to 10.0

[1708] Spot 42α, 43α: Molecular weight 10 to 50 kDa, pI 6.0 to 10.0

[1709] Spot 44α: Molecular weight 10 to 50 kDa, pI 4.0 to 7.0

[1710] Spot 45α: Molecular weight 5 to 40 kDa, pI 4.0 to 7.0

[1711] Spot 46α: Molecular weight 5 to 40 kDa, pI 4.0 to 7.0

[1712] Spot 47α: Molecular weight 5 to 40 kDa, pI 4.0 to 7.0

[1713] Spot 48α: Molecular weight 5 to 40 kDa, pI 4.0 to 8.0

[1714] Spot 49α: Molecular weight 5 to 20 kDa, pI 4.0 to 7.0

[1715] Spot 50α: Molecular weight 5 to 40 kDa, pI 3.0 to 6.0

[1716] Spot 51α: Molecular weight 10 to 50 kDa, pI 3.0 to 6.0

[1717] Spot 52α: Molecular weight 20 to 60 kDa, pI 4.0 to 7.0

[1718] Spot 53α: Molecular weight 10 to 50 kDa, pI 4.0 to 8.0

[1719] Spot 54α: Molecular weight 10 to 50 kDa, pI 4.0 to 8.0

[1720] Spot 55α: Molecular weight 40 to 80 kDa, pI 6.0 to 10.0

Example 3: Mass Spectrometry and Identification of Antigens

[1721] The amino acid sequences of the antigens that form the 56 protein spots were identified by mass spectroscopy.

[1722] To be specific, protein extraction and mass spectroscopy were done by the following procedure.

[1723] (1) Soybeans were subjected to protein extraction, two-dimensional electrophoresis and transfer to membrane by following the procedures described in Example 1 and 2, and the resulting membrane was stained by shaking in a solution of 0.008% Direct blue in 40% ethanol and 10% acetic acid.

[1724] (2) Then, the membrane was decolorized by repeating a 5-minute treatment with 40% ethanol and 10% acetic acid three times, washed with water for 5 minutes, and then dried by air.

[1725] (3) A protein spot of interest was cut out with a clean cutter blade and put into a centrifugal tube. The cut membrane was subjected to hydrophilization with 50 μL of methanol, followed by washing with 100 μL of water twice and then centrifugal cleaning. Thereafter, 20 μL of 20 mM NH_4HCO_3 and 50% acetonitrile were added.

[1726] (4) 1 μL of 1 pmol/μL lysyl endopeptidase (produced by WAKO) was added, and the solution was left to stand at 37° C. for 60 minutes and then collected in a new centrifugal tube. After 20 μL of 20 mM NH_4HCO_3 and 70% acetonitrile were added to the membrane, the membrane was immersed therein at room temperature for 10 minutes, and the resulting solution was further collected. The solution was dissolved with 0.1% formic acid and 10 μL of 4% acetonitrile and transferred to a tube.

[1727] (5) The collected solution was dried under reduced pressure, dissolved with 15 μL of solution A (a 0.1% formic acid/4% acetonitrile solution), and analyzed by mass spectroscopy (ESI-TOF5600, produced by AB Sciex).

[1728] (6) Identification of proteins based on the MS data obtained with the mass spectrometer was done by searching the NCBI database.

Results

[1729] The mass spectrometry of the spots resulted in the detection of the following amino acid sequences. Furthermore, the spots were identified as the following proteins by the analysis of the mass data obtained from the mass spectrometer for the spots at NCBI.

Spot 1α:

[1730] The amino acid sequences set forth in SEQ ID NOs: 71 to 87 were detected by the mass spectrometry. The protein identified from this mass data was endoplasmin homolog isoform X2 (NCBI accession number XP_006596543.1, amino acid sequence: SEQ ID NO: 70, encoding nucleotide sequence: SEQ ID NO: 69).

[1731] The amino acid sequences set forth in SEQ ID NOs: 90 to 104 were detected by the mass spectrometry. The protein identified from this mass data was endoplasmin homolog isoform X2 (NCBI accession number XP_006601356.1, amino acid sequence: SEQ ID NO: 89, encoding nucleotide sequence: SEQ ID NO: 88).

Spot 2α

[1732] The amino acid sequences set forth in SEQ ID NOs: 107 to 120 were detected by the mass spectrometry. The protein identified from this mass data was heat shock cognate protein 80 (NCBI accession number XP_003544594.1, amino acid sequence: SEQ ID NO: 106, encoding nucleotide sequence: SEQ ID NO: 105).

[1733] The amino acid sequences set forth in SEQ ID NOs: 123 to 137 were detected by the mass spectrometry. The protein identified from this mass data was heat shock protein 90-1 (NCBI accession number NP_001236612.1, amino acid sequence: SEQ ID NO: 122, encoding nucleotide sequence: SEQ ID NO: 121).

[1734] The amino acid sequences set forth in SEQ ID NOs: 140 to 153 were detected by the mass spectrometry. The protein identified from this mass data was hypothetical protein GLYMA_02G302500 (NCBI accession number KRH73936.1, amino acid sequence: SEQ ID NO: 139, encoding nucleotide sequence: SEQ ID NO: 138).

Spot 3α:

[1735] The amino acid sequences set forth in SEQ ID NOs: 62 to 68 were detected by the mass spectrometry. The protein identified from this mass data was Cell division cycle protein 48 homolog (UniProt accession number P54774, amino acid sequence: SEQ ID NO: 16, encoding nucleotide sequence: SEQ ID NO: 15).

Spot 4α:

[1736] The amino acid sequences set forth in SEQ ID NOs: 156, 157 were detected by the mass spectrometry. The protein identified from this mass data was embryo-specific urease isoform X1 (NCBI accession number XP_014630847.1, amino acid sequence: SEQ ID NO: 155, encoding nucleotide sequence: SEQ ID NO: 154).

Spot 5α:

[1737] The amino acid sequences set forth in SEQ ID NOs: 160, 161 were detected by the mass spectrometry. The protein identified from this mass data was alpha-1,4 glucan phosphorylase L isozyme, chloroplastic/amyloplastic (NCBI accession number XP_006603904.1, amino acid sequence: SEQ ID NO: 159, encoding nucleotide sequence: SEQ ID NO: 158).

Spot 6α:

[1738] The amino acid sequences set forth in SEQ ID NOs: 164 to 167 were detected by the mass spectrometry. The protein identified from this mass data was aconitate hydratase 1 (NCBI accession number XP_003537655.1, amino acid sequence: SEQ ID NO: 163, encoding nucleotide sequence: SEQ ID NO: 162).

[1739] The amino acid sequences set forth in SEQ ID NOs: 170 to 172 were detected by the mass spectrometry. The protein identified from this mass data was aconitate hydratase 1 (NCBI accession number XP_003517155.1, amino acid sequence: SEQ ID NO: 169, encoding nucleotide sequence: SEQ ID NO: 168).

Spot 7α:

[1740] The amino acid sequences set forth in SEQ ID NOs: 175 to 187 were detected by the mass spectrometry. The protein identified from this mass data was methionine synthase (NCBI accession number NP_001235794.1, amino acid sequence: SEQ ID NO: 174, encoding nucleotide sequence: SEQ ID NO: 173).

Spot 8α:

[1741] The amino acid sequences set forth in SEQ ID NOs: 190 to 199 were detected by the mass spectrometry. The protein identified from this mass data was transketolase, chloroplastic (NCBI accession number XP_003521870.1, amino acid sequence: SEQ ID NO: 189, encoding nucleotide sequence: SEQ ID NO: 188).

Spot 9α:

[1742] The amino acid sequence set forth in SEQ ID NO: 202 was detected by the mass spectrometry. The protein identified from this mass data was lysine-tRNA ligase, cytoplasmic-like isoform X2 (NCBI accession number XP_014625509.1, amino acid sequence: SEQ ID NO: 201, encoding nucleotide sequence: SEQ ID NO: 200).

Spot 10α

[1743] The amino acid sequences set forth in SEQ ID NOs: 205 to 218 were detected by the mass spectrometry. The protein identified from this mass data was chaperonin CPN60-2, mitochondrial (NCBI accession number XP_003556325.1, amino acid sequence: SEQ ID NO: 204, encoding nucleotide sequence: SEQ ID NO: 203).

[1744] The amino acid sequences set forth in SEQ ID NOs: 221 to 231 were detected by the mass spectrometry. The protein identified from this mass data was chaperonin CPN60-2, mitochondrial-like isoform X1 (NCBI accession number XP_003535967.1, amino acid sequence: SEQ ID NO: 220, encoding nucleotide sequence: SEQ ID NO: 219).

Spot 11α:

[1745] The amino acid sequences set forth in SEQ ID NOs: 234 to 237 were detected by the mass spectrometry. The protein identified from this mass data was polyadenylate-binding protein 8 isoform X3 (NCBI accession number XP_006578034.1, amino acid sequence: SEQ ID NO: 233, encoding nucleotide sequence: SEQ ID NO: 232).

Spot 12α:

[1746] The amino acid sequence set forth in SEQ ID NO: 240 was detected by the mass spectrometry. The protein identified from this mass data was pyrophosphate-fructose 6-phosphate 1-phosphotransferase subunit alpha-like (NCBI

accession number XP_003555655.1, amino acid sequence: SEQ ID NO: 239, encoding nucleotide sequence: SEQ ID NO: 238).

Spot 13α:

[1747] The amino acid sequences set forth in SEQ ID NOs: 243 to 252 were detected by the mass spectrometry. The protein identified from this mass data was protein disulfide isomerase-like protein precursor (NCBI accession number NP_001238009.1, amino acid sequence: SEQ ID NO: 242, encoding nucleotide sequence: SEQ ID NO: 241).

[1748] The amino acid sequences set forth in SEQ ID NOs: 255 to 278 were detected by the mass spectrometry. The protein identified from this mass data was protein disulfide-isomerase (NCBI accession number XP_006581588.1, amino acid sequence: SEQ ID NO: 254, encoding nucleotide sequence: SEQ ID NO: 253).

Spot 14α:

[1749] The amino acid sequences set forth in SEQ ID NOs: 281 to 284 were detected by the mass spectrometry. The protein identified from this mass data was V-type proton ATPase subunit B 2 (NCBI accession number XP_003552473.1, amino acid sequence: SEQ ID NO: 280, encoding nucleotide sequence: SEQ ID NO: 279).

[1750] The amino acid sequences set forth in SEQ ID NOs: 287 to 289 were detected by the mass spectrometry. The protein identified from this mass data was V-type proton ATPase subunit B 2 (NCBI accession number XP_006605857.1, amino acid sequence: SEQ ID NO: 286, encoding nucleotide sequence: SEQ ID NO: 275).

Spots 15α, 16α:

[1751] The amino acid sequences set forth in SEQ ID NOs: 292 to 304 were detected by the mass spectrometry. The protein identified from this mass data was UTP-glucose-1-phosphate uridylyltransferase (NCBI accession number XP_003544964.1, amino acid sequence: SEQ ID NO: 291, encoding nucleotide sequence: SEQ ID NO: 290).

[1752] The amino acid sequences set forth in SEQ ID NOs: 307 to 318 were detected by the mass spectrometry. The protein identified from this mass data was hypothetical protein GLYMA_02G241100 (NCBI accession number KRH72929.1, amino acid sequence: SEQ ID NO: 306, encoding nucleotide sequence: SEQ ID NO: 305).

Spot 17α:

[1753] The amino acid sequences set forth in SEQ ID NOs: 321 to 325 were detected by the mass spectrometry. The protein identified from this mass data was guanosine nucleotide diphosphate dissociation inhibitor 2 (NCBI accession number XP_003531293.1, amino acid sequence: SEQ ID NO: 320, encoding nucleotide sequence: SEQ ID NO: 319).

[1754] The amino acid sequences set forth in SEQ ID NOs: 328 to 332 were detected by the mass spectrometry. The protein identified from this mass data was rab GDP dissociation inhibitor alpha-like (NCBI accession number NP_001276273.1, amino acid sequence: SEQ ID NO: 327, encoding nucleotide sequence: SEQ ID NO: 326).

Spot 18α:

[1755] The amino acid sequences set forth in SEQ ID NOs: 335 to 339 were detected by the mass spectrometry. The protein identified from this mass data was hypothetical protein GLYMA_10G028300 (NCBI accession number KRH32039.1, amino acid sequence: SEQ ID NO: 334, encoding nucleotide sequence: SEQ ID NO: 333).

Spot 19α:

[1756] The amino acid sequences set forth in SEQ ID NOs: 342 to 346 were detected by the mass spectrometry. The protein identified from this mass data was sucrose binding protein homolog S-64 (NCBI accession number AAF05723.1, amino acid sequence: SEQ ID NO: 341, encoding nucleotide sequence: SEQ ID NO: 340).

Spot 20α:

[1757] The amino acid sequences set forth in SEQ ID NOs: 35 to 53 were detected by the mass spectrometry. The protein identified from this mass data was Seed biotin-containing protein SBP65 (UniProt accession number Q39846, amino acid sequence: SEQ ID NO: 12, encoding nucleotide sequence: SEQ ID NO: 11).

Spot 21α:

[1758] The amino acid sequences set forth in SEQ ID NOs: 21 to 28 were detected by the mass spectrometry. The protein identified from this mass data was Ribulose bisphosphate carboxylase large chain (UniProt accession number P27066, amino acid sequence: SEQ ID NO: 4, encoding nucleotide sequence: SEQ ID NO: 3).

Spot 22α:

[1759] The amino acid sequences set forth in SEQ ID NOs: 29, 30 were detected by the mass spectrometry. The protein identified from this mass data was Gamma-glutamyl hydrolase (UniProt accession number P93164, amino acid sequence: SEQ ID NO: 6, encoding nucleotide sequence: SEQ ID NO: 5).

Spot 23α:

[1760] The amino acid sequences set forth in SEQ ID NOs: 31, 32 were detected by the mass spectrometry. The protein identified from this mass data was Guanine nucleotide-binding protein subunit beta-like protein (UniProt accession number Q39836, amino acid sequence: SEQ ID NO: 8, encoding nucleotide sequence: SEQ ID NO: 7).

Spot 24α:

[1761] The amino acid sequences set forth in SEQ ID NOs: 33, 34 were detected by the mass spectrometry. The protein identified from this mass data was Protein SLE3 (UniProt accession number C6TOL2, amino acid sequence: SEQ ID NO: 10, encoding nucleotide sequence: SEQ ID NO: 9).

Spot 25α:

[1762] The amino acid sequences set forth in SEQ ID NOs: 349 to 356 were detected by the mass spectrometry. The protein identified from this mass data was succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial (NCBI

accession number XP_003537078.1, amino acid sequence: SEQ ID NO: 348, encoding nucleotide sequence: SEQ ID NO: 347).

[1763] The amino acid sequences set forth in SEQ ID NOs: 359 to 365 were detected by the mass spectrometry. The protein identified from this mass data was succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial (NCBI accession number XP_003542431.1, amino acid sequence: SEQ ID NO: 358, encoding nucleotide sequence: SEQ ID NO: 357).

Spot 26α:

[1764] The amino acid sequences set forth in SEQ ID NOs: 368 to 382 were detected by the mass spectrometry. The protein identified from this mass data was phosphoglycerate kinase, cytosolic (NCBI accession number XP_003546821.1, amino acid sequence: SEQ ID NO: 367, encoding nucleotide sequence: SEQ ID NO: 366).

[1765] The amino acid sequences set forth in SEQ ID NOs: 385 to 392 were detected by the mass spectrometry. The protein identified from this mass data was phosphoglycerate kinase, cytosolic (NCBI accession number XP_003531482.1, amino acid sequence: SEQ ID NO: 384, encoding nucleotide sequence: SEQ ID NO: 383).

[1766] The amino acid sequences set forth in SEQ ID NOs: 395 to 409 were detected by the mass spectrometry. The protein identified from this mass data was phosphoglycerate kinase, cytosolic (NCBI accession number XP_003531483.1, amino acid sequence: SEQ ID NO: 394, encoding nucleotide sequence: SEQ ID NO: 393).

Spot 27α:

[1767] The amino acid sequences set forth in SEQ ID NOs: 412 to 416 were detected by the mass spectrometry. The protein identified from this mass data was alcohol dehydrogenase 1 (NCBI accession number XP_003526673.1, amino acid sequence: SEQ ID NO: 411, encoding nucleotide sequence: SEQ ID NO: 410).

[1768] The amino acid sequences set forth in SEQ ID NOs: 419, 420 were detected by the mass spectrometry. The protein identified from this mass data was alcohol dehydrogenase 1 (NCBI accession number XP_003545664.3, amino acid sequence: SEQ ID NO: 418, encoding nucleotide sequence: SEQ ID NO: 417).

[1769] The amino acid sequences set forth in SEQ ID NOs: 423 to 427 were detected by the mass spectrometry. The protein identified from this mass data was alcohol dehydrogenase 1-like (NCBI accession number XP_003544738.1, amino acid sequence: SEQ ID NO: 422, encoding nucleotide sequence: SEQ ID NO: 421).

Spot 28α:

[1770] The amino acid sequences set forth in SEQ ID NOs: 430 to 433 were detected by the mass spectrometry. The protein identified from this mass data was 35 kDa seed maturation protein isoform X1 (NCBI accession number XP_006576267.1, amino acid sequence: SEQ ID NO: 429, encoding nucleotide sequence: SEQ ID NO: 428).

Spots 29α, 30α, 31α, 32α, 33α:

[1771] The amino acid sequences set forth in SEQ ID NOs: 436 to 439 were detected by the mass spectrometry. The protein identified from this mass data was glyceraldehyde-3-phosphate dehydrogenase isoform X1 (NCBI accession number XP_014631598.1, amino acid sequence: SEQ ID NO: 435, encoding nucleotide sequence: SEQ ID NO: 434).

[1772] The amino acid sequences set forth in SEQ ID NOs: 442 to 445 were detected by the mass spectrometry. The protein identified from this mass data was glyceraldehyde-3-phosphate dehydrogenase, cytosolic (NCBI accession number XP_006578692.1, amino acid sequence: SEQ ID NO: 441, encoding nucleotide sequence: SEQ ID NO: 440).

[1773] The amino acid sequences set forth in SEQ ID NOs: 448 to 451 were detected by the mass spectrometry. The protein identified from this mass data was uncharacterized protein LOC100782924 (NCBI accession number NP_001240046.1, amino acid sequence: SEQ ID NO: 447, encoding nucleotide sequence: SEQ ID NO: 446).

Spot 34α:

[1774] The amino acid sequence set forth in SEQ ID NO: 454 was detected by the mass spectrometry. The protein identified from this mass data was bifunctional UDP-glucose 4-epimerase and UDP-xylose 4-epimerase 1-like (NCBI accession number XP_003532591.1, amino acid sequence: SEQ ID NO: 453, encoding nucleotide sequence: SEQ ID NO: 452).

[1775] The amino acid sequences set forth in SEQ ID NOs: 457, 458 were detected by the mass spectrometry. The protein identified from this mass data was uncharacterized protein LOC100803483 (NCBI accession number NP_001241075.1, amino acid sequence: SEQ ID NO: 456, encoding nucleotide sequence: SEQ ID NO: 455).

Spot 35α:

[1776] The amino acid sequences set forth in SEQ ID NOs: 461 to 466 were detected by the mass spectrometry. The protein identified from this mass data was elongation factor 1-alpha (NCBI accession number XP_003553292.1, amino acid sequence: SEQ ID NO: 460, encoding nucleotide sequence: SEQ ID NO: 459).

[1777] The amino acid sequences set forth in SEQ ID NOs: 469 to 471 were detected by the mass spectrometry. The protein identified from this mass data was elongation factor 1-alpha (NCBI accession number XP_003547695.1, amino acid sequence: SEQ ID NO: 468, encoding nucleotide sequence: SEQ ID NO: 467).

[1778] The amino acid sequences set forth in SEQ ID NOs: 474 to 477 were detected by the mass spectrometry. The protein identified from this mass data was elongation factor-1A isoform X1 (NCBI accession number XP_006600131.1, amino acid sequence: SEQ ID NO: 473, encoding nucleotide sequence: SEQ ID NO: 472).

Spots 36α, 37α:

[1779] The amino acid sequences set forth in SEQ ID NOs: 480 to 483 were detected by the mass spectrometry. The protein identified from this mass data was seed maturation protein PM34 (NCBI accession number NP_001238285.1, amino acid sequence: SEQ ID NO: 479, encoding nucleotide sequence: SEQ ID NO: 478).

[1780] The amino acid sequences set forth in SEQ ID NOs: 486, 487 were detected by the mass spectrometry. The protein identified from this mass data was uncharacterized

protein LOC100809384 (NCBI accession number NP_001241158.1, amino acid sequence: SEQ ID NO: 485, encoding nucleotide sequence: SEQ ID NO: 484).

[1781] The amino acid sequences set forth in SEQ ID NOs: 490 to 493 were detected by the mass spectrometry. The protein identified from this mass data was hypothetical protein GLYMA_10G155300 (NCBI accession number KRH33956.1, amino acid sequence: SEQ ID NO: 489, encoding nucleotide sequence: SEQ ID NO: 488).

Spot 38α

[1782] The amino acid sequences set forth in SEQ ID NOs: 496 to 513 were detected by the mass spectrometry. The protein identified from this mass data was seed biotin-containing protein SBP65-like isoform X1 (NCBI accession number XP_003526237.1, amino acid sequence: SEQ ID NO: 495, encoding nucleotide sequence: SEQ ID NO: 494).

Spot 39α:

[1783] The amino acid sequence set forth in SEQ ID NO: 516 was detected by the mass spectrometry. The protein identified from this mass data was ATP synthase subunit O, mitochondrial-like (NCBI accession number XP_003546884.1, amino acid sequence: SEQ ID NO: 515, encoding nucleotide sequence: SEQ ID NO: 514).

Spots 40α, 56α:

[1784] The amino acid sequences set forth in SEQ ID NOs: 519, 520 were detected by the mass spectrometry. The protein identified from this mass data was P24 oleosin isoform A (NCBI accession number XP_003553822.1, amino acid sequence: SEQ ID NO: 518, encoding nucleotide sequence: SEQ ID NO: 517).

Spot 41α:

[1785] The amino acid sequence set forth in SEQ ID NO: 523 was detected by the mass spectrometry. The protein identified from this mass data was uncharacterized protein At3g49720-like isoform X2 (NCBI accession number XP_006600908.1, amino acid sequence: SEQ ID NO: 522, encoding nucleotide sequence: SEQ ID NO: 521).

Spots 42α, 43α

[1786] The amino acid sequences set forth in SEQ ID NOs: 526 to 531 were detected by the mass spectrometry. The protein identified from this mass data was superoxide dismutase [Mn], mitochondrial (NCBI accession number XP_003526813.1, amino acid sequence: SEQ ID NO: 525, encoding nucleotide sequence: SEQ ID NO: 524).

[1787] The amino acid sequences set forth in SEQ ID NOs: 534 to 539 were detected by the mass spectrometry. The protein identified from this mass data was hypothetical protein GLYMA_04G221300 (NCBI accession number KRH64185.1, amino acid sequence: SEQ ID NO: 533, encoding nucleotide sequence: SEQ ID NO: 532).

Spot 44α:

[1788] The amino acid sequences set forth in SEQ ID NOs: 54 to 61 were detected by the mass spectrometry. The protein identified from this mass data was Ferritin-2 (UniProt accession number Q941C4, amino acid sequence: SEQ ID NO: 14, encoding nucleotide sequence: SEQ ID NO: 13).

Spot 45α:

[1789] The amino acid sequences set forth in SEQ ID NOs: 542 to 545 were detected by the mass spectrometry. The protein identified from this mass data was uncharacterized protein LOC100499771 (NCBI accession number NP_001235797.1, amino acid sequence: SEQ ID NO: 541, encoding nucleotide sequence: SEQ ID NO: 540).

[1790] The amino acid sequences set forth in SEQ ID NOs: 548 to 551 were detected by the mass spectrometry. The protein identified from this mass data was hypothetical protein GLYMA_09G192800 (NCBI accession number KRH39322.1, amino acid sequence: SEQ ID NO: 547, encoding nucleotide sequence: SEQ ID NO: 546).

Spot 46α:

[1791] The amino acid sequence set forth in SEQ ID NO: 554 was detected by the mass spectrometry. The protein identified from this mass data was seed maturation protein PM22 (NCBI accession number NP_001237935.1, amino acid sequence: SEQ ID NO: 553, encoding nucleotide sequence: SEQ ID NO: 552).

Spot 47α:

[1792] The amino acid sequences set forth in SEQ ID NOs: 557, 558 were detected by the mass spectrometry. The protein identified from this mass data was eukaryotic translation initiation factor 5A-2 (NCBI accession number XP_003549690.1, amino acid sequence: SEQ ID NO: 556, encoding nucleotide sequence: SEQ ID NO: 555).

[1793] The amino acid sequences set forth in SEQ ID NOs: 561 to 563 were detected by the mass spectrometry. The protein identified from this mass data was uncharacterized protein LOC100305510 (NCBI accession number NP_001236395.1, amino acid sequence: SEQ ID NO: 560, encoding nucleotide sequence: SEQ ID NO: 559).

Spot 48α:

[1794] The amino acid sequences set forth in SEQ ID NOs: 566 to 570 were detected by the mass spectrometry. The protein identified from this mass data was uncharacterized protein LOC100306570 (NCBI accession number NP_001236895.1, amino acid sequence: SEQ ID NO: 565, encoding nucleotide sequence: SEQ ID NO: 564).

Spot 49α:

[1795] The amino acid sequences set forth in SEQ ID NOs: 17 to 20 were detected by the mass spectrometry. The protein identified from this mass data was Bowman-Birk type proteinase inhibitor D-II (UniProt accession number P01064, amino acid sequence: SEQ ID NO: 2, encoding nucleotide sequence: SEQ ID NO: 1).

Spot 50α:

[1796] The amino acid sequences set forth in SEQ ID NOs: 573 to 576 were detected by the mass spectrometry. The protein identified from this mass data was uncharacterized protein LOC100813859 (NCBI accession number NP_001239816.1, amino acid sequence: SEQ ID NO: 572, encoding nucleotide sequence: SEQ ID NO: 571).

Spot 51α:

[1797] The amino acid sequences set forth in SEQ ID NOs: 579 to 582 were detected by the mass spectrometry. The protein identified from this mass data was 14-3-3-like protein (NCBI accession number XP_003526571.1, amino acid sequence: SEQ ID NO: 578, encoding nucleotide sequence: SEQ ID NO: 577).

Spot 52α:

[1798] The amino acid sequences set forth in SEQ ID NOs: 585 to 588 were detected by the mass spectrometry. The protein identified from this mass data was probable fructokinase-4 (NCBI accession number XP_003543564.1, amino acid sequence: SEQ ID NO: 584, encoding nucleotide sequence: SEQ ID NO: 583).

Spot 53α:

[1799] The amino acid sequences set forth in SEQ ID NOs: 591 to 597 were detected by the mass spectrometry. The protein identified from this mass data was triosephosphate isomerase isoform X1 (NCBI accession number XP_006593480.2, amino acid sequence: SEQ ID NO: 590, encoding nucleotide sequence: SEQ ID NO: 589).

[1800] The amino acid sequences set forth in SEQ ID NOs: 600 to 604 were detected by the mass spectrometry. The protein identified from this mass data was triosephosphate isomerase, cytosolic (NCBI accession number XP_003547334.1, amino acid sequence: SEQ ID NO: 599, encoding nucleotide sequence: SEQ ID NO: 598).

Spot 54α:

[1801] The amino acid sequences set forth in SEQ ID NOs: 607 to 611 were detected by the mass spectrometry. The protein identified from this mass data was superoxide dismutase [Fe], chloroplastic isoform X1 (NCBI accession number XP_014627751.1, amino acid sequence: SEQ ID NO: 606, encoding nucleotide sequence: SEQ ID NO: 605).

[1802] The amino acid sequences set forth in SEQ ID NOs: 614 to 616 were detected by the mass spectrometry. The protein identified from this mass data was iron-superoxide dismutase (NCBI accession number NP_001237901.1, amino acid sequence: SEQ ID NO: 613, encoding nucleotide sequence: SEQ ID NO: 612).

Spot 55α:

[1803] The amino acid sequences set forth in SEQ ID NOs: 619 to 635 were detected by the mass spectrometry. The protein identified from this mass data was 51 kDa seed maturation protein precursor (NCBI accession number NP_001238138.1, amino acid sequence: SEQ ID NO: 618, encoding nucleotide sequence: SEQ ID NO: 617).

[1804] The amino acid sequences set forth in SEQ ID NOs: 638 to 649 were detected by the mass spectrometry. The protein identified from this mass data was Lea protein precursor (NCBI accession number NP_001238201.1, amino acid sequence: SEQ ID NO: 637, encoding nucleotide sequence: SEQ ID NO: 636).

Example 4: Identification of Epitopes (1)

[1805] Identification of epitopes was carried out by the following procedure as to each of the antigen proteins identified in Example 3.

(A) Epitope Mapping

[1806] Epitope mapping was carried out using a library of overlapping peptide fragments (length: 15 amino acids) corresponding to the sequences of proteins bound patient-specifically to IgE antibodies in serum, which were obtained as a result of 2D-western of soybeans.

[1807] The peptides to be synthesized were shifted by 10 amino acids. Thus, each peptide had an overlap of 5 amino acids with each of the peptides previous and subsequent thereto.

[1808] For preparation of peptide arrays, peptide synthesis of interest was carried out in incremental steps on a modified cellulose membrane by the Intavis SPOT synthesis technique, which is a technique of synthesizing peptides of interest through 9-fluorenylmethoxycarbonyl (Fmoc) chemical reaction, using an automated synthesis apparatus (Intavis MultiPep RS). The aforementioned modified cellulose membrane (hereinafter also referred to as the membrane) bound to the obtained peptides of interest was used in epitope mapping.

[1809] The presence or absence of binding, to each peptide fragment, of an IgE antibody present in the serum of a patient allergic to a soybean was measured by immunoblotting using the membrane on which the peptide fragments were synthesized by SPOT. Immunoblotting of the membrane was carried out according to the following procedure using, as a primary antibody, a serum sample from an allergic patient, and anti-human IgE-HRP as a secondary antibody.

[1810] (1) The membrane was subjected to hydrophilization with methanol for 5 minutes.

[1811] (2) The hydrophilized membrane was shaken overnight in a 5% skim milk/PBST solution (a PBS buffer containing 0.1% Tween 20 nonionic surfactant) at 4° C.

[1812] (3) The membrane was washed with a PBST solution three times.

[1813] (4) The serum sample (1:40, a 5% skim milk/PBST solution) was added thereto as a primary antibody and shaken overnight at 4° C. In this operation, a protease inhibitor (Complete, produced by Roche) was added.

[1814] (5) The membrane was washed with a PBST solution three times.

[1815] (6) Anti-human IgE-HRP (1:10,000, a 5% skim milk/PBST solution) was added thereto as a secondary antibody and shaken at room temperature for one hour.

[1816] (7) The membrane was washed with a PBST solution three times.

[1817] (8) The membrane was left to stand in ECL plus (produced by Thermo Fisher Scientific) for 5 minutes.

[1818] (9) The membrane treated as described above in (1) to (8) was subjected to fluorescent image scanning using Amersham Imager 600.

[1819] Peptide fragments bound to IgE antibodies from allergic patients in immunoblotting were used as candidate peptides of epitopes.

(B) Overlapping

[1820] Thereafter, for the sequences of peptides bound patient-specifically to IgE antibodies in serum in the epitope mapping described above in (A), a library of overlapping peptide fragments (length: 10 amino acids) was obtained as

sequences having the sequences of these peptides as well as additional sequences subsequent to the peptides in the amino acid sequences of antigen proteins on which the epitopes were detected. Synthesis of this library was carried out by the Fmoc chemical synthesis method in the same manner as in (A) mentioned above. The peptides to be synthesized were shifted by one amino acid. Thus, each peptide had an overlap of 9 amino acids with the peptide subsequent thereto.

[1821] The presence or absence of binding of an IgE antibody from an allergic patient to each peptide fragment in this library of overlapping peptide fragments was measured by immunoblotting in the same manner as in (A) mentioned above. A common sequence of peptides bound to IgE antibodies from allergic patients was identified to determine peptides having the minimum number of amino acids functioning as an epitope. The peptides determined in (B) are referred to as the "original sequences".

[1822] As one example of the results, the results of determining the original sequences by the overlapping technique as to a partial sequence (a region of amino acids 571-590 of SEQ ID NO: 495/SEQ ID NO: 964 of seed biotin-containing protein SBP65-like isoform X1 (spot 38α) are shown in FIG. **4**. Strong binding of polypeptides having the amino acid sequences of SEQ ID NOs: 965-967 to the IgE antibody from the allergic patient was observed. Slightly weak binding of SEQ ID NO: 968 to the IgE antibody from the allergic patient was observed. Binding of SEQ ID NOs: 969-975 to the IgE antibody from the allergic patient was not observed. As a result, SEQ ID NO: 864, a common sequence of SEQ ID NOs: 965-968, was identified as an original sequence.

(C) Alanine Scanning

[1823] From the "original sequences" determined as described above in (B), a library of peptide fragments in which N-terminal amino acids were substituted one by one by alanine according to a technique called alanine scanning (Non Patent Literature 1) was prepared by the Fmoc chemical synthesis method in the same manner as in (A) mentioned above. The presence or absence of binding of an IgE antibody present in the serum of an allergic patient to each peptide fragment was measured by immunoblotting in the same manner as in (A) mentioned above. Amino acids at positions where the binding activity against IgE antibodies from allergic patients was lost by the substitution by alanine were confirmed as positions important for exertion of antigenicity. Also, Amino acids at positions where the binding activity against IgE antibodies from allergic patients was maintained even after the substitution by alanine were confirmed as substitutable positions that were not important for exertion of antigenicity. Common sequences (consensus sequences) important for exertion of antigenicity were found by this analysis.

[1824] As one example of the results, the results of performing alanine scanning as to the amino acid sequence of SEQ ID NO: 864 identified in (B) are shown in FIG. **5**. Peptides having SEQ ID NOs: 977-979 in which the amino acids at positions 1-3, respectively, of SEQ ID NO: 864 were substituted by alanine lost their binding activity against IgE antibodies from allergic patients. Thus, the amino acids at positions 1-3 of SEQ ID NO: 864 were identified as amino acids important for the binding to IgE antibodies from allergic patients.

(D) Results

(1β) Epitope of Endoplasmin Homolog Isoform X2 (Spot 1α)

[1825] SEQ ID NOs: 652, 655, 657, and 658-660 were identified as epitopes of endoplasmin homolog isoform X2 by the procedure of (B) mentioned above.

[1826] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 652, 655, 657, and 658-660.

[1827] It was revealed for SEQ ID NO: 652 that the amino acids at positions 3 and 5 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 2 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 4, and 6-8 of SEQ ID NO: 652 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1828] It was revealed for SEQ ID NO: 655 that the amino acids at positions 4-7 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 2 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 3, 8 and 9 of SEQ ID NO: 655 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1829] For SEQ ID NO: 657, the amino acids at positions 1-5 and 11 were amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 6-10 and 12-15 of SEQ ID NO: 657 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1830] For SEQ ID NOs: 658-660, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(2β) Epitope of Heat Shock Cognate Protein 80 (Spot 2α)

[1831] SEQ ID NOs: 663, 666, 669, 670-672, 675, 676-678, 681, 684, 687, 689, and 690-692 were identified as epitopes of heat shock cognate protein 80 by the procedure of (B) mentioned above.

[1832] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 663, 666, 669, 670-672, 675-678, 681, 684, 687 and 689-692.

[1833] It was revealed for SEQ ID NO: 663 that the amino acids at positions 1, 4 and 5 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 7 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 2, 3, 6, 8 and 9 of SEQ ID NO: 663 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1834] It was revealed for SEQ ID NO: 666 that the amino acids at positions 5-8 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 12 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1-4, 9-11, and 13-15 of SEQ ID NO:

666 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1835] It was revealed for SEQ ID NO: 669 that the amino acids at positions 2-4 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 1 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acid at position 5 of SEQ ID NO: 669 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1836] It was revealed for SEQ ID NO: 675 that the amino acids at positions 3-5, 13 and 14 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 15 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 2 and 6-12 of SEQ ID NO: 675 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alaninc.

[1837] It was revealed for SEQ ID NO: 681 that the amino acids at positions 5, 10 and 14 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 6-8 and 15 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 1-4, 9 and 11-13 of SEQ ID NO: 681 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1838] It was revealed for SEQ ID NO: 684 that the amino acids at positions 2, 5, 10, and 11 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 3, 4, 6-9, and 12-15 of SEQ ID NO: 684 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alaninc.

[1839] It was revealed for SEQ ID NO: 687 that the amino acids at positions 1, 5, 10, and 11 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 2, 3, 4, 6-9 and 12-15 of SEQ ID NO: 687 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1840] It was revealed for SEQ ID NO: 689 that the amino acids at positions 1-3, 12 and 13 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 4-11, 14 and 15 of SEQ ID NO: 689 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1841] For SEQ ID NOs: 670-672, 676-678, and 690-692, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(3β) Epitope of Cell Division Cycle Protein 48 Homolog (Spot 3α)

[1842] SEQ ID NOs: 695, 697, 699, and 700-716 were identified as epitopes of Cell division cycle protein 48 homolog by the procedure of (B) mentioned above.

[1843] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 695, 697, 699 and 700-716.

[1844] It was revealed for SEQ ID NO: 695 that the amino acids at positions 11, 12 and 14 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 4-10, 13 and 15 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 1-3 of SEQ ID NO: 695 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1845] It was revealed for SEQ ID NO: 697 that the amino acids at positions 14 and 15 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 1-13 of SEQ ID NO: 697 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1846] It was revealed for SEQ ID NO: 699 that the amino acids at positions 5-7 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 1~4 of SEQ ID NO: 699 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1847] For SEQ ID NOs: 700-716, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(4β) Epitope of Embryo-Specific Urease Isoform X1 (Spot 4α)

[1848] SEQ ID NOs: 719, 722, 725, 726-728, and 731 were identified as epitopes of embryo-specific urease isoform X1 by the procedure of (B) mentioned above.

[1849] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 719, 722, 725, 726-728, and 731.

[1850] It was revealed for SEQ ID NO: 719 that the amino acids at positions 8-10 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 6 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1-5 and 7 of SEQ ID NO: 719 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1851] It was revealed for SEQ ID NO: 722 that the amino acids at positions 1 and 4-6 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 2 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 3 and 7 of SEQ ID NO: 722 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1852] It was revealed for SEQ ID NO: 725 that the amino acids at positions 4, 6, 8 and 10 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 9 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1-3, 5 and 7 of SEQ ID NO: 725 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1853] It was revealed for SEQ ID NO: 731 that the amino acids at positions 4, 6, 10, 11, 14 and 15 are amino acids

particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 3 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 2, 5, 7-9, 12 and 13 of SEQ ID NO: 731 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1854] For SEQ ID NOs: 726-728, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(5β) Epitope of Alpha-1,4 Glucan Phosphorylase L Isozyme, Chloroplastic/Amyloplastic (Spot 5α)

[1855] SEQ ID NOs: 734 and 735-738 were identified as epitopes of alpha-1,4 glucan phosphorylase L isozyme, chloroplastic/amyloplastic by the procedure of (B) mentioned above.

[1856] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 734, 735-738.

[1857] It was revealed for SEQ ID NO: 734 that the amino acids at positions 7, 8, 10, 12 and 14 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 15 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1-6, 9, 11 and 13 of SEQ ID NO: 734 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1858] For SEQ ID NOs: 735-738, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(6β) Epitope of Aconitate Hydratase 1 (Spot 6α)

[1859] SEQ ID NOs: 739 and 740 were identified as epitopes of aconitate hydratase 1 by the procedure of (B) mentioned above.

[1860] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 739 and 740.

[1861] For SEQ ID NOs: 739 and 740, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(7β) Epitope of Methionine Synthase (Spot 7α)

[1862] SEQ ID NOs: 741 and 744 were identified as epitopes of methionine synthase by the procedure of (B) mentioned above.

[1863] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 741 and 744.

[1864] It was revealed for SEQ ID NO: 744 that the amino acids at positions 8, 9, 11-13 and 15 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 10 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1-7 and 14 of SEQ ID NO: 744 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1865] For SEQ ID NO: 741, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(9β) Epitope of Lysine—tRNA Ligase, Cytoplasmic-Like Isoform X2 (Spot 9α)

[1866] SEQ ID NOs: 747 and 749 were identified as epitopes of lysine—tRNA ligase, cytoplasmic-like isoform X2 by the procedure of (B) mentioned above.

[1867] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 747 and 749.

[1868] It was revealed for SEQ ID NO: 747 that the amino acids at positions 3, 8 and 9 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 1 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 2, 4-7 and 10 of SEQ ID NO: 747 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1869] It was revealed for SEQ ID NO: 749 that the amino acids at positions 1-5, 7, and 8 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acid at position 6 of SEQ ID NO: 749 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

(11β) Epitope of Polyadenylate-Binding Protein 8 Isoform X3 (Spot 11α)

[1870] SEQ ID NOs: 753 and 754 were identified as epitopes of polyadenylate-binding protein 8 isoform X3 by the procedure of (B) mentioned above.

[1871] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 753 and 754.

[1872] It was revealed for SEQ ID NO: 753 that the amino acids at positions 1, 2, 4 and 5 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 6 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 3 and 7-10 of SEQ ID NO: 753 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1873] For SEQ ID NO: 754, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(12) Epitope of Pyrophosphate—Fructose 6-Phosphate 1-Phosphotransferase Subunit Alpha-Like (Spot 12α)

[1874] SEQ ID NO: 755 was identified as an epitope of pyrophosphate—fructose 6-phosphate 1-phosphotransferase subunit alpha-like by the procedure of (B) mentioned above.

[1875] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NO: 755.

[1876] For SEQ ID NO: 755, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(13β) Epitope of Protein Disulfide Isomerase-Like Protein Precursor (Spot 13α)

[1877] SEQ ID NO: 758 was identified as an epitope of protein disulfide isomerase-like protein precursor by the procedure of (B) mentioned above.

[1878] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NO: 758.

[1879] It was revealed for SEQ ID NO: 758 that the amino acids at positions 3 and 15 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 5, 10, 12 and 13 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 2, 4, 6-9, 11 and 14 of SEQ ID NO: 758 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

(14β) Epitope of V-Type Proton ATPase Subunit B 2 (Spot 14α)

[1880] SEQ ID NOs: 761 and 764 were identified as epitopes of V-type proton ATPase subunit B 2 by the procedure of (B) mentioned above.

[1881] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 761 and 764.

[1882] It was revealed for SEQ ID NO: 761 that the amino acids at positions 6, 7 and 12-15 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 8 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1-5 and 9-11 of SEQ ID NO: 761 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1883] It was revealed for SEQ ID NO: 764 that the amino acids at positions 1, 2, 12 and 14 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 3-5, 11 and 13 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 6-10 and 15 of SEQ ID NO: 746 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

(15β) Epitope of UTP-Glucose-1-Phosphate Uridylyltransferase (Spots 15α and 16α)

[1884] SEQ ID NO: 766 was identified as an epitope of UTP-glucose-1-phosphate uridylyltransferase by the procedure of (B) mentioned above.

[1885] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NOs: 766.

[1886] It was revealed for SEQ ID NO: 766 that the amino acids at positions 2 and 3 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 1 and 4-15 of SEQ ID NO: 766 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

(17β) Epitope of Guanosine Nucleotide Diphosphate Dissociation Inhibitor 2 (Spot 17α)

[1887] SEQ ID NO: 767 was identified as an epitope of guanosine nucleotide diphosphate dissociation inhibitor 2 by the procedure of (B) mentioned above.

[1888] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NO: 767.

[1889] For SEQ ID NO: 767, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(18β) Epitope of Hypothetical Protein GLYMA 10G028300 (Spot 18α)

[1890] SEQ ID NOs: 750, 768, 771, 772, and 775 were identified as epitopes of hypothetical protein GLYMA_10G028300 by the procedure of (B) mentioned above.

[1891] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 750, 768, 771, 772 and 775.

[1892] It was revealed for SEQ ID NO: 771 that the amino acids at positions 8 and 11-15 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 10 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1-7 and 9 of SEQ ID NO: 771 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1893] It was revealed for SEQ ID NO: 775 that the amino acids at positions 2-6 and 8 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 7 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 9 and 10 of SEQ ID NO: 775 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1894] For SEQ ID NOs: 750, 768 and 772, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(19β) Epitope of Sucrose Binding Protein Homolog S-64 (Spot 19α)

[1895] SEQ ID NOs: 778, 781, 784, 786, and 787 were identified as epitopes of sucrose binding protein homolog S-64 by the procedure of (B) mentioned above.

[1896] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 778, 781, 784, 786 and 787.

[1897] It was revealed for SEQ ID NO: 778 that the amino acids at positions 6-8, 13 and 14 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 9 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1-5, 10-12 and 15 of SEQ ID NO: 778 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1898] It was revealed for SEQ ID NO: 781 that the amino acids at positions 9, 11 and 12 are amino acids particularly

important for the binding to IgE antibodies from allergic patients, and the amino acid at position 6 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1-5, 7, 8, 10 and 13-15 of SEQ ID NO: 781 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1899] It was revealed for SEQ ID NO: 784 that the amino acids at positions 11-13 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 2-8 and 14 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 9, 10 and 15 of SEQ ID NO: 784 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alaninc.

[1900] It was revealed for SEQ ID NO: 786 that the amino acids at positions 2 and 3 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 1-2 and 5-15 of SEQ ID NO: 786 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1901] For SEQ ID NO: 787, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(20β) Epitope of Seed Biotin-Containing Protein SBP65 (Spot 20α)

[1902] SEQ ID NOs: 788-792 were identified as epitopes of Seed biotin-containing protein SBP65 by the procedure of (B) mentioned above.

[1903] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 788-792.

[1904] For SEQ ID NOs: 788-792, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(21β) Epitope of Ribulose Bisphosphate Carboxylase Large Chain (Spot 21α)

[1905] SEQ ID NOs: 793-802 were identified as epitopes of Ribulose bisphosphate carboxylase large chain by the procedure of (B) mentioned above.

[1906] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 793-802.

[1907] For SEQ ID NOs: 793-802, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(22β) Epitope of Gamma-Glutamyl Hydrolase (Spot 22α)

[1908] SEQ ID NOs: 803-808 were identified as epitopes of Gamma-glutamyl hydrolase by the procedure of (B) mentioned above.

[1909] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 803-808.

[1910] For SEQ ID NOs: 803-808, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(23β) Epitope of Guanine Nucleotide-Binding Protein Subunit Beta-Like Protein (Spot 23α)

[1911] SEQ ID NOs: 809-812 were identified as epitopes of Guanine nucleotide-binding protein subunit beta-like protein by the procedure of (B) mentioned above.

[1912] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 809-812.

[1913] For SEQ ID NOs: 809-812, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(24β) Epitope of Protein SLE3 (Spot 24α)

[1914] SEQ ID NO: 813 was identified as an epitope of Protein SLE3 by the procedure of (B) mentioned above.

[1915] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NO: 813.

[1916] For SEQ ID NO: 813, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(25) Epitope of Succinyl-CoA Ligase [ADP-Forming] Subunit Beta, Mitochondrial (Spot 25α)

[1917] SEQ ID NOs: 814 and 815 were identified as epitopes of succinyl-CoA ligase [ADP-forming] subunit beta, mitochondrial by the procedure of (B) mentioned above.

[1918] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 814 and 815.

[1919] For SEQ ID NOs: 814 and 815, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(26β) Epitope of Phosphoglycerate Kinase, Cytosolic (Spot 26α)

[1920] SEQ ID NOs: 816 and 819 were identified as epitopes of phosphoglycerate kinase, cytosolic by the procedure of (B) mentioned above.

[1921] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 816 and 819.

[1922] It was revealed for SEQ ID NO: 819 that the amino acids at positions 1-3 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 4 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 5-7 of SEQ ID NO: 819 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1923] For SEQ ID NO: 816, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(27) Epitope of Alcohol Dehydrogenase 1 (Spot 27α)

[1924] SEQ ID NO: 820 was identified as an epitope of alcohol dehydrogenase 1 by the procedure of (B) mentioned above.

[1925] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NO: 820.

[1926] For SEQ ID NO: 820, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(28β) Epitope of 35 kDa Seed Maturation Protein Isoform X1 (Spot 28α)

[1927] SEQ ID NOs: 823, 826, 829, and 830-835 were identified as epitopes of 35 kDa seed maturation protein isoform X1 by the procedure of (B) mentioned above.

[1928] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 823, 826, 829 and 830-835.

[1929] It was revealed for SEQ ID NO: 823 that the amino acids at positions 2 and 7-10 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 1, 4 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 3, 5 and 6 of SEQ ID NO: 823 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1930] It was revealed for SEQ ID NO: 826 that the amino acids at positions 2, 7 and 8 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 3 and 14 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 4-6, 9-13 and 15 of SEQ ID NO: 826 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1931] It was revealed for SEQ ID NO: 829 that the amino acids at positions 4, 7-9 and 14 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 10 and 11 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 1-3, 5, 6, 12, 13 and 15 of SEQ ID NO: 829 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1932] For SEQ ID NOs: 830-835, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(29β) Epitope of Glyceraldehyde-3-Phosphate Dehydrogenase Isoform X1 (Spots 29α, 30α, 31α, 32α, and 33α)

[1933] SEQ ID NOs: 838 and 839 were identified as epitopes of glyceraldehyde-3-phosphate dehydrogenase isoform X1 by the procedure of (B) mentioned above.

[1934] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 838 and 839.

[1935] It was revealed for SEQ ID NO: 838 that the amino acids at positions 8-10 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 6 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1-5 and 7 of SEQ ID NO: 838 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1936] For SEQ ID NO: 839, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(34β) Epitope of Bifunctional UDP-Glucose 4-Epimerase and UDP-Xylose 4-Epimerase 1-Like (Spot 34α)

[1937] SEQ ID NOs: 842 and 843-846 were identified as epitopes of bifunctional UDP-glucose 4-epimerase and UDP-xylose 4-epimerase 1-like by the procedure of (B) mentioned above.

[1938] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 842 and 843-846.

[1939] It was revealed for SEQ ID NO: 842 that the amino acids at positions 12, 13 and 15 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 1 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 2-11 and 14 of SEQ ID NO: 842 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1940] For SEQ ID NOs: 843-846, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(35β) Epitope of Elongation Factor 1-Alpha (Spot 35α)

[1941] SEQ ID NOs: 849, 850, 853, and 856 were identified as epitopes of elongation factor 1-alpha by the procedure of (B) mentioned above.

[1942] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 849, 850, 853 and 856.

[1943] It was revealed for SEQ ID NO: 849 that the amino acids at positions 5, 6, 10, 11 and 15 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 7 and 8 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 1-4, 9 and 12-14 of SEQ ID NO: 849 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1944] It was revealed for SEQ ID NO: 853 that the amino acids at positions 1-3, 5 and 6 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 4 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 7-9 of SEQ ID NO: 853 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1945] It was revealed for SEQ ID NO: 856 that the amino acids at positions 3, 5 and 6 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 1 and 2 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 4 and 7-10 of

SEQ ID NO: 856 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1946] For SEQ ID NO: 850, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(36β) Epitope of Seed Maturation Protein PM34 (Spots 36α and 37α)

[1947] SEQ ID NOs: 857, 858, and 861 were identified as epitopes of seed maturation protein PM34 by the procedure of (B) mentioned above.

[1948] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 857, 858, and 861.

[1949] It was revealed for SEQ ID NO: 861 that the amino acids at positions 1 and 6-8 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 5 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 2-4 of SEQ ID NO: 861 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1950] For SEQ ID NOs: 857 and 858, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(38β) Epitope of Seed Biotin-Containing Protein SBP65-Like Isoform X1 (Spot 38α)

[1951] SEQ ID NOs: 864, 867, 869, 871, 874, 877, 880, 882, 884, 887, 889, and 892 were identified as epitopes of seed biotin-containing protein SBP65-like isoform X1 by the procedure of (B) mentioned above.

[1952] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 864, 867, 869, 871, 874, 877, 880, 882, 884, 887, 889, and 892.

[1953] It was revealed for SEQ ID NO: 864 that the amino acids at positions 1-3 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 4-7 of SEQ ID NO: 864 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1954] It was revealed for SEQ ID NO: 867 that the amino acids at positions 2-4 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 1 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 5-10 of SEQ ID NO: 867 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1955] It was revealed for SEQ ID NO: 869 that the amino acids at positions 1, 3 and 15 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 2 and 4-14 of SEQ ID NO: 869 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1956] It was revealed for SEQ ID NO: 871 that the amino acids at positions 8, 9 and 15 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 1-7 and 10-14 of SEQ ID NO: 871 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1957] It was revealed for SEQ ID NO: 874 that the amino acids at positions 4-8 and 10 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 1-3 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acid at position 9 of SEQ ID NO: 874 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1958] It was revealed for SEQ ID NO: 877 that the amino acids at positions 1-6, 9 and 10 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 7 and 8 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 11-15 of SEQ ID NO: 877 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1959] It was revealed for SEQ ID NO: 880 that the amino acids at positions 2, 4-6, 8 and 9 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 7 and 12-15 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 3, 10 and 11 of SEQ ID NO: 880 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alaninc.

[1960] It was revealed for SEQ ID NO: 882 that the amino acids at positions 1 and 5-9 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 2-4 and 10 of SEQ ID NO: 882 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1961] It was revealed for SEQ ID NO: 884 that the amino acids at positions 5-10 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 1~4 of SEQ ID NO: 884 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alaninc.

[1962] It was revealed for SEQ ID NO: 887 that the amino acids at positions 2, 4, 5 and 14 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 6, 10, 12, 13 and 15 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 3, 7-9 and 11 of SEQ ID NO: 887 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alaninc.

[1963] It was revealed for SEQ ID NO: 889 that the amino acids at positions 3-7 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 2 and 8-10 of SEQ ID NO: 889 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1964] It was revealed for SEQ ID NO: 892 that the amino acids at positions 5-10 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 3 and 4 are amino acids that influence the binding to IgE antibodies from allergic

patients. The amino acids at positions 1 and 2 of SEQ ID NO: 892 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

(39β) Epitope of ATP Synthase Subunit O, Mitochondrial-Like (Spot 39α)

[1965] SEQ ID NO: 895 was identified as an epitope of ATP synthase subunit O, mitochondrial-like by the procedure of (B) mentioned above.

[1966] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NO: 895.

[1967] It was revealed for SEQ ID NO: 895 that the amino acids at positions 2, 5, 7, and 12-14 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 3 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 4, 6, 8-11, and 15 of SEQ ID NO: 895 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

(40β) Epitope of P24 Oleosin Isoform a (Spots 40α and 56α)

[1968] SEQ ID NOs: 898, 899, and 902 were identified as epitopes of P24 oleosin isoform A by the procedure of (B) mentioned above.

[1969] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 898, 899 and 902.

[1970] It was revealed for SEQ ID NO: 898 that the amino acids at positions 1, 2, 4 and 5 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 6 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acid at position 3 of SEQ ID NO: 898 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1971] It was revealed for SEQ ID NO: 902 that the amino acids at positions 6-8 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 1 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 2-5 and 9 of SEQ ID NO: 902 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1972] For SEQ ID NO: 899, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(41β) Epitope of Uncharacterized Protein At3g49720-Like Isoform X2 (Spot 41α)

[1973] SEQ ID NO: 903 was identified as an epitope of uncharacterized protein At3g49720-like isoform X2 by the procedure of (B) mentioned above.

[1974] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NO: 903.

[1975] For SEQ ID NO: 903, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(42) Epitope of Superoxide Dismutase [Mn], Mitochondrial (Spots 42α and 43α)

[1976] SEQ ID NO: 906 was identified as an epitope of superoxide dismutase [Mn], mitochondrial by the procedure of (B) mentioned above.

[1977] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NO: 906.

[1978] It was revealed for SEQ ID NO: 906 that the amino acids at positions 3 and 11-13 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 8 and 14 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 2, 4-7, 9, 10, and 15 of SEQ ID NO: 906 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

(44β) Epitope of Ferritin-2 (Spot 44α)

[1979] SEQ ID NO: 909 was identified as an epitope of Ferritin-2 by the procedure of (B) mentioned above.

[1980] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NO: 909.

[1981] It was revealed for SEQ ID NO: 909 that the amino acids at positions 5-11 and 13-15 are amino acids particularly important for the binding to IgE antibodies from allergic patients. The amino acids at positions 1~4 and 12 of SEQ ID NO: 909 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

(45β) Epitope of Uncharacterized Protein LOC100499771 (Spot 45α)

[1982] SEQ ID NOs: 910-912 were identified as epitopes of uncharacterized protein LOC100499771 by the procedure of (B) mentioned above.

[1983] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 910-912.

[1984] For SEQ ID NOs: 910-912, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(46β) Epitope of Seed Maturation Protein PM22 (Spot 46α)

[1985] SEQ ID NO: 915 was identified as an epitope of seed maturation protein PM22 by the procedure of (B) mentioned above.

[1986] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NO: 915.

[1987] It was revealed for SEQ ID NO: 915 that the amino acids at positions 2, 4, and 8 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 3 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 5-7, and 9 of SEQ ID NO: 915 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

(47β) Epitope of Eukaryotic Translation Initiation Factor 5A-2 (Spot 47α)

[1988] SEQ ID NOs: 916-918 were identified as epitopes of eukaryotic translation initiation factor 5A-2 by the procedure of (B) mentioned above.

[1989] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 916-918.

[1990] For SEQ ID NOs: 916-918, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(48β) Epitope of Uncharacterized Protein LOC100306570 (Spot 48α)

[1991] SEQ ID NOs: 921 and 922 were identified as epitopes of uncharacterized protein LOC100306570 by the procedure of (B) mentioned above.

[1992] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 921 and 922.

[1993] It was revealed for SEQ ID NO: 921 that the amino acids at positions 7-9 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 5 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1~4 and 6 of SEQ ID NO: 921 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[1994] For SEQ ID NO: 922, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(49β) Epitope of Bowman-Birk Type Proteinase Inhibitor D-II (Spot 49α)

[1995] SEQ ID NO: 923 was identified as an epitope of Bowman-Birk type proteinase inhibitor D-II by the procedure of (B) mentioned above.

[1996] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NO: 923.

[1997] For SEQ ID NO: 923, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(50β) Epitope of Uncharacterized Protein LOC100813859 (Spot 50α)

[1998] SEQ ID NOs: 924 and 927 were identified as epitopes of uncharacterized protein LOC100813859 by the procedure of (B) mentioned above.

[1999] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 924 and 927.

[2000] It was revealed for SEQ ID NO: 927 that the amino acids at positions 5 and 6 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 3 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 2, 4, and 7-9 of SEQ ID NO: 927 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[2001] For SEQ ID NO: 924, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(51) Epitope of 14-3-3-Like Protein (Spot 51α)

[2002] SEQ ID NOs: 928-930 were identified as epitopes of 14-3-3-like protein by the procedure of (B) mentioned above.

[2003] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 928-930.

[2004] For SEQ ID NOs: 928-930, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(52β) Epitope of Probable Fructokinase-4 (Spot 52α)

[2005] SEQ ID NOs: 933 and 934-938 were identified as epitopes of probable fructokinase-4 by the procedure of (B) mentioned above.

[2006] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 933 and 934-938.

[2007] It was revealed for SEQ ID NO: 933 that the amino acids at positions 7, 10, and 12 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 3 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 1, 2, 4-6, 8, 9, 11, and 13-15 of SEQ ID NO: 933 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[2008] For SEQ ID NOs: 934-938, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(53β) Epitope of Triosephosphate Isomerase Isoform X1 (Spot 53α)

[2009] SEQ ID NO: 939 was identified as an epitope of triosephosphate isomerase isoform X1 by the procedure of (B) mentioned above.

[2010] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to SEQ ID NO: 939.

[2011] For SEQ ID NO: 939, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(54β) Epitope of Superoxide Dismutase [Fe], Chloroplastic Isoform X1 (Spot 54α)

[2012] SEQ ID NOs: 940 and 943 were identified as epitopes of superoxide dismutase [Fe], chloroplastic isoform X1 by the procedure of (B) mentioned above.

[2013] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 940 and 943.

[2014] It was revealed for SEQ ID NO: 943 that the amino acids at positions 1 and 6 are amino acids particularly important for the binding to IgE antibodies from allergic

patients, and the amino acid at position 7 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 2-5 of SEQ ID NO: 943 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[2015] For SEQ ID NO: 940, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(55β) Epitope of 51 kDa Seed Maturation Protein Precursor (Spot 55α)

[2016] SEQ ID NOs: 946, 949, and 952 were identified as epitopes of 51 kDa seed maturation protein precursor by the procedure of (B) mentioned above.

[2017] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) mentioned above as to each of SEQ ID NOs: 946, 949, and 952.

[2018] It was revealed for SEQ ID NO: 946 that the amino acids at positions 1, 2, 6, and 7 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 3 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 4, 5, 8, and 9 of SEQ ID NO: 946 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[2019] It was revealed for SEQ ID NO: 949 that the amino acids at positions 1-3, 7, and 8 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 4 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 5 and 6 of SEQ ID NO: 949 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

[2020] It was revealed for SEQ ID NO: 952 that the amino acids at positions 1, 2, and 9 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acid at position 8 is an amino acid that influences the binding to IgE antibodies from allergic patients. The amino acids at positions 3-7 of SEQ ID NO: 952 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

Example 5: Identification of Epitopes (2)

[2021] Identification of epitopes was carried out by the same procedure as in Example 4 as to each of the following antigen proteins.

(57) Epitope of Gly m 4 (UniProt Accession Number: P26987.1)

[2022] SEQ ID NO: 955 was identified as an epitope of Gly m 4 by the procedure of (B) of Example 4.

[2023] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) of Example 4 as to SEQ ID NO: 955.

[2024] It was revealed for SEQ ID NO: 955 that the amino acids at positions 8, 10, 12, 14, and 15 are amino acids particularly important for the binding to IgE antibodies from allergic patients, and the amino acids at positions 7, 9, 11, and 13 are amino acids that influence the binding to IgE antibodies from allergic patients. The amino acids at positions 1-6 of SEQ ID NO: 955 had no influence on the binding activity against IgE antibodies from allergic patients when substituted by alanine.

(58β) Epitopes of Gly m 5a (NCBI Accession Number: NP_001236856.1) and Gly m 5α' (NCBI Accession Number: NP 001237316.1)

[2025] SEQ ID NOs: 956-960 were identified as epitopes of Gly m 5α or Gly m 5α' by the procedure of (B) of Example 4.

[2026] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) of Example 4 as to each of SEQ ID NOs: 956-960.

[2027] For SEQ ID NOs: 956-960, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(59β) Epitopes of Gly m 6.01 (UniProt Accession Number: CAA33215.1) and Gly m 6.02 (UniProt Accession Number: BAA00154.1)

[2028] SEQ ID NOs: 961 and 962 were identified as epitopes of Gly m 6.01 or Gly m 6.02 by the procedure of (B) of Example 4.

[2029] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) of Example 4 as to each of SEQ ID NOs: 961 and 962.

[2030] For SEQ ID NOs: 961 and 962, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

(60β) Epitope of Gly m 8 (UniProt Accession Number: AAB71140.1)

[2031] SEQ ID NO: 963 was identified as an epitope of Gly m 8 by the procedure of (B) of Example 4.

[2032] Amino acids important for binding to IgE antibodies from allergic patients were identified by the procedure of (C) of Example 4 as to SEQ ID NO: 963.

[2033] For SEQ ID NO: 963, all the amino acids were amino acids particularly important for the binding to IgE antibodies from allergic patients.

Example 6: Study on Cross-Reactivity

[2034] Amino acids important for binding activity against IgE antibodies from allergic patients were identified as to SEQ ID NO: 871, one of the polypeptides in spot 38β, to obtain XXXXXXXRRXXXXXR (hereinafter also referred to as a key sequence). A protein having this key sequence was compared against NCBI protein data and analyzed by PHI-BLAST. As a result, peach-derived kinesin-like protein KIN-7C comprising SEQ ID NO: 984 was identified as an amino acid sequence corresponding to the key sequence, and apple-derived uncharacterized protein LOC103433727 isoform X2 comprising SEQ ID NO: 985 was identified as an amino acid sequence corresponding to the key sequence.

[2035] The binding activity of peptides comprising the amino acid sequences of SEQ ID NOS: 871, 984 and 985 against an IgE antibody from one patient having allergies to a soybean, a peach and an apple was confirmed by ELISA. The peptides were synthesized such that the peptides were N-terminally biotinylated by the Fmoc method.

[2036] To be specific, ELISA was carried out according to the following procedure.

[2037] (1) The concentrations of the biotinylated peptides were adjusted to 1 μg/mL with PBST.

[2038] (2) Each peptide solution was added at 40 μL to each well of a 384-well plate coated with streptavidin, and shaken at room temperature for one hour. After collection of the solution, the wells were washed with PBS five times.

[2039] (3) A 1:5 dilution of Blocking one (diluted with MQ) was added at 40 μL thereto and shaken at room temperature for 15 minutes. After removal of the solution, the wells were washed with PBS five times.

[2040] (4) A diluted serum solution (1:10, diluting solution: Blocking one produced by Nacalai Tesque, Inc.) was added at 40 μL thereto and shaken at room temperature for one hour. After removal of the solution, the wells were washed with PBS five times.

[2041] (5) A diluted secondary antibody solution (1:5000, diluting solution: Blocking one produced by Nacalai Tesque, Inc.) was added at 40 μL thereto and shaken at room temperature for one hour. After removal of the solution, the wells were washed with PBS five times.

[2042] (6) 1-Step Ultra TMB-ELISA (Thermo Fisher Scientific) was added at 40 μL thereto and shaken at room temperature for 15 minutes.

[2043] (7) 2 M H_2SO_4 was added at 40 μL thereto. Absorbance at 450 nm was measured.

[2044] The results are shown in FIG. **6**. As shown in FIG. **6**, the peptides comprising the amino acid sequences of SEQ ID NOs: 871, 984 and 985 exhibited binding activity against the IgE antibody from the patient having allergies to a soybean, a peach and an apple. Also, these peptides bound more strongly to the IgE antibody from the allergic patient than to an IgE antibody from a healthy subject who was not an allergic patient. This indicates that this key sequence can be used for detecting the cross-reactivity of an allergen.

[2045] This result also indicates that the polypeptides (1β) to (60β) of the present invention can also be used for the detection of the cross-reactivity of an allergen.

SEQUENCE LISTING

```
Sequence total quantity: 985
SEQ ID NO: 1            moltype = DNA  length = 249
FEATURE                 Location/Qualifiers
source                  1..249
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 1
atgtgtattc tgagcttcct caaaagtgat cagtcatcaa gttatgatga tgatgagtat  60
tcaaaaccat gctgtgatct ctgcatgtgc acacgctcaa tgcctcctca atgcagctgt  120
gaagatatta ggctgaattc atgccactca gattgtaaga gctgtatgtg cacacgctca  180
cagccaggac agtgtcgttg tcttgacacc aacgacttct gctacaaacc ttgcaagtcc  240
agagatgac                                                          249

SEQ ID NO: 2            moltype = AA  length = 83
FEATURE                 Location/Qualifiers
source                  1..83
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 2
MCILSFLKSD QSSSYDDDEY SKPCCDLCMC TRSMPPQCSC EDIRLNSCHS DCKSCMCTRS  60
QPGQCRCLDT NDFCYKPCKS RDD                                          83

SEQ ID NO: 3            moltype = DNA  length = 1425
FEATURE                 Location/Qualifiers
source                  1..1425
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 3
atgtcaccac aaacagagac taaagcaagt gttgggttca aagctggtgt taaagattat  60
aaattgactt attactcc tgactatgaa accaaagata ctgatatctt ggcagcattc  120
cgagtaactc ctcaaccagg agttccgcct gaagaagcag gtgccgcggt agccgccgaa  180
tcttctactg gtacatggac aactgtgtgg accgatgggc ttaccagtct tgatcgttac  240
aaagggcgat gctacggcct tgaacctgtt gctggggaag aaaatcaata tattgcttat  300
gtagcttatc ccttagacct ttttgaagaa ggttctgtta ctaacatgtt tacttccatt  360
gtcggtaatg tatttgggtt caaggccctg cgtgctctac gtctggagga tttgcgaatc  420
cctactgctt atattaaaac tttccaaggt ccgcctcatg gcatccaagt tgagagagat  480
aaattgaaca agtatggtcg tcccctatta ggatgtacta ttaaacctaa attggggtta  540
tccgctaaga attatggtag agctgtttat gaatgtcttc gtgggggact tgattttacc  600
aaagatgatg aaaatgtgaa ttcccaacca tttatgcgtt ggagagaccg tttcttattt  660
tgtgccgaag ccatttttaa atcacaggct gaaacaggtg aaatcaaagg gcattacttg  720
aatgcaactg cgggtacatg cgaagaaatg atgaaaagag ctgtatttgc cagagaatta  780
ggcgttccta tcgtaatgca tgattattta acgggggat tcactgcaaa tactagcttg  840
gctcattatt gccgagataa tggtctactt cttcatatac accgtgcaat gcatgcagtt  900
atcgacagac aaaagaatca tggtatgcac tttcgtgtac tagctaaagc attacgtttg  960
tctggtggag atcatgttca cgccggtacc gtagtaggta aacttgaagg ggaaagagaa  1020
atcactttag gttttgttga tttactacgt gatgattttg ttgaaaaaga tcgaagtcgc  1080
ggtatttatt tcactcagga ttgggtttct ctaccaggtg ttttgcctgt agcttcggga  1140
ggtattcacg tttggcatat gcctgctctg accgagatct ttggggatga ttctgtactc  1200
```

-continued

```
caatttggcg gaggaacttt aggacaccct tggggaaatg caccaggtgc tgtagctaat  1260
cgagtagctc ttgaagcatg tgtgcaggct cgaaatgaag gacgcgatct tgctcgtgaa  1320
ggtaatgaaa ttatccgtga ggctagcaaa tggagtcctg aattagctgc tgcttgtgaa  1380
gtatggaagg agatcaaatt tgaattcgaa gcaatggata ctttg                  1425

SEQ ID NO: 4            moltype = AA  length = 475
FEATURE                 Location/Qualifiers
source                  1..475
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 4
MSPQTETKAS VGFKAGVKDY KLTYYTPDYE TKDTDILAAF RVTPQPGVPP EEAGAAVAAE  60
SSTGTWTTVW TDGLTSLDRY KGRCYGLEPV AGEENQYIAY VAYPLDLFEE GSVTNMFTSI  120
VGNVFGFKAL RALRLEDLRI PTAYIKTFQG PPHGIQVERD KLNKYGRPLL GCTIKPKLGL  180
SAKNYGRAVY ECLRGGLDFT KDDENVNSQP FMRWRDRFLF CAEAIFKSQA ETGEIKGHYL  240
NATAGTCEEM MKRAVFAREL GVPIVMHDYL TGGFTANTSL AHYCRDNGLL LHIHRAMHAV  300
IDRQKNHGMH FRVLAKALRL SGGDHVHAGT VVGKLEGERE ITLGFVDLLR DDFVEKDRSR  360
GIYFTQDWVS LPGVLPVASG GIHVWHMPAL TEIFGDDSVL QFGGGTLGHP WGNAPGAVAN  420
RVALEACVQA RNEGRDLARE GNEIIREASK WSPELAAACE VWKEIKFEFE AMDTL       475

SEQ ID NO: 5            moltype = DNA  length = 1026
FEATURE                 Location/Qualifiers
source                  1..1026
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 5
atgccaaacg acagcgttct gagtcttttc ttcttcgtta ccctcttcac atgccttctc  60
tctgctactt cccacgatga ccatattttc ctaccgagtc aactccatga cgatgactca  120
gtgagttgca ctgcaaccga cccaagcctt aactacaagc cggtgatcgg aatcctgacc  180
caccccggcg acggtgcctc tggccggctc agcaacgcca ccggggtttc ttacatcgcc  240
gcctcctacg ttaaattcgt ggagtccggt ggtgctagag ttattcccct tatctacaat  300
gagtcacctg agaaccttaa caagaagctc gatttggtca atggagtgct cttcactggt  360
gggtgggctg tttctggtcc atatttggat actcttggaa acatttttaa gaaagcttta  420
gagagaaatg atgctggaga ccatttccca gtaattgcct tcaacttagg tggaaacctc  480
gtaataagga tcgtaagcga gcaaactgat attcttgaac catttactgc gtcgagtttg  540
ccaagttctc ttgtattgtg gaacgaagca aatgctaaag gaagtctatt tcaaagattt  600
ccttcagatt tgttgacaca gctgaagaca gattgtctcg tcctgcacaa tcatcgttat  660
gccatatccc cacgcaagct tcaatataat accaagttat ctgacttttt tgagattctg  720
gcaacatctg gggatagaga tggaaagact ttcgtttcca cagcacgtgg taggaaatat  780
cccgtgactg tcaacctatg gcaaccagag aaaaatgcgt ttgagtgggc tacatctttg  840
aaagcccctc acacagagga tgcaattcgg gttactcagt ccacagcaaa ctttttcata  900
agtgaagcta ggaaatccac aaacacacct gatgctcaaa aagtgcgaga tagtctcatc  960
tacaattaca aacctacttt tggtggaacg gctgggaaag gatatgatca agtgtatctc  1020
tttgaa                                                             1026

SEQ ID NO: 6            moltype = AA  length = 342
FEATURE                 Location/Qualifiers
source                  1..342
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 6
MPNDSVLSLF FFVTLFTCLL SATSHDDHIF LPSQLHDDDS VSCTATDPSL NYKPVIGILT  60
HPGDGASGRL SNATGVSYIA ASYVKFVESG GARVIPLIYN ESPENLNKKL DLVNGVLFTG  120
GWAVSGPYLD TLGNIFKKAL ERNDAGDHFP VIAFNLGGNL VIRIVSEQTD ILEPFTASSL  180
PSSLVLWNEA NAKGSLFQRF PSDLLTQLKT DCLVLHNHRY AISPRKLQYN TKLSDFFEIL  240
ATSGDRDGKT FVSTARGRKY PVTVNLWQPE KNAFEWATSL KAPHTEDAIR VTQSTANFFI  300
SEARKSTNTP DAQKVRDSLI YNYKPTFGGT AGKGYDQVYL FE                     342

SEQ ID NO: 7            moltype = DNA  length = 975
FEATURE                 Location/Qualifiers
source                  1..975
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 7
atggcggaag gactcgtcct gaagggaacc atgcgcgccc acaccgacgt cgtgacggcg  60
attgccactc ccatcgacaa ctccgacatg atcgtgacgg cgtcgcgcga cagatccatc  120
attctctggc acctcaccaa ggaggacaag acctacggtg tcccccgccg ccgcctcacc  180
ggacattctc acttcgtcca ggacgtcgtc ctctcatccg acggtcagtt cgccctctcc  240
ggctcctggg acggcgagct ccgcctctgg gacctcgcgg ctggcacctc tgcccgccgc  300
ttcgttggcc acaccaagga cgtgctctcc gtggcgttct ccatcgacaa ccgtcagatc  360
gtgtcggcct ctcgtgaccg cacgatcaag ctgtggaaca ccctgggtga gtgcaagtac  420
acaatccaag atggcgatgc gcattcggat tgggtaagtt gcgtccgttt cagccctagc  480
actcttcagc caaccattgt ttctgcttca tgggacagga ccgttaaggt ttggaacctg  540
accaactgca agctgaggaa cacccttgct ggacacaatg ggtatgtgaa tactgttgct  600
gtttcccctg atggctctct ctgtgccagt ggtggcaaag atggggttat tcttctgtgg  660
gatttggctg agggtaagcg tctttactct ctcgatgctg gctcaatcat ccatgcactc  720
```

-continued

```
tgcttcagcc ccagcaggta ctggctctgc gccgccaccg agcagagcat caagatctgg  780
gatttggaga gcaagagtat cgttgaggat ttgaaggttg acctcaagac tgaggctgat  840
gccacctccg gtggtggtaa cgccaacaag aagaaggtta tttattgcac aagtttgaac  900
tggagtgcgg atggaagcac tttgtttagt ggctataccg atggtgtggc cagagtttgg  960
gctattggac gttat                                                   975

SEQ ID NO: 8            moltype = AA  length = 325
FEATURE                 Location/Qualifiers
source                  1..325
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 8
MAEGLVLKGT MRAHTDVVTA IATPIDNSDM IVTASRDRSI ILWHLTKEDK TYGVPRRRLT  60
GHSHFVQDVV LSSDGQFALS GSWDGELRLW DLAAGTSARR FVGHTKDVLS VAFSIDNRQI  120
VSASRDRTIK LWNTLGECKY TIQDGDAHSD WVSCVRFSPS TLQPTIVSAS WDRTVKVWNL  180
TNCKLRNTLA GHNGYVNTVA VSPDGSLCAS GGKDGVILLW DLAEGKRLYS LDAGSIIHAL  240
CFSPSRYWLC AATEQSIKIW DLESKSIVED LKVDLKTEAD ATSGGGNANK KKVIYCTSLN  300
WSADGSTLFS GYTDGVARVW AIGRY                                        325

SEQ ID NO: 9            moltype = DNA  length = 303
FEATURE                 Location/Qualifiers
source                  1..303
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 9
atggcatctc atcaacaaaa caagcaagag cttgatgaaa gggcaaggca aggagagact  60
gttgttccgg gtggcactgg tggcaagagc cttgaggctc agcaacatct tgctgaagga  120
aggagcaggg gagggaaaac aaggaaggag cagctgggga cagaagggta ccatgaaatg  180
ggacgcaaag gtgggttgag cactatggac aagtcaggag aagaacgtgc tcaagaggaa  240
gccattgaca tcgatgagtc caagttcagg actggtaata accaagacaa gaaccagaac  300
aag                                                                303

SEQ ID NO: 10           moltype = AA  length = 101
FEATURE                 Location/Qualifiers
source                  1..101
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 10
MASHQQNKQE LDERARQGET VVPGGTGGKS LEAQQHLAEG RSRGGKTRKE QLGTEGYHEM  60
GRKGGLSTMD KSGEERAQEE AIDIDESKFR TGNNQDKNQN K                      101

SEQ ID NO: 11           moltype = DNA  length = 1929
FEATURE                 Location/Qualifiers
source                  1..1929
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 11
atggcgtctg aacaattggc tcgcagagag aacaccacca ccgagaagga gattcacgtt  60
gagaaacaca gggttcccaa gatggccacc cacttcgagc accttgcgga gcaagctaaa  120
gaatctgaca tcaccgctgg gaaggatact ccgcagggta gcatagaagc cttgcaagct  180
ggagagaggg tcaaagatca tgcaggaaaa gccatgggag atattggtgg acgtgggaaa  240
gccagagaga ctcacgagct gggtgcacac tttgaatccc tggctgacaa agtaactgac  300
catgcagcag ccaatgtagt tggaaacaaa gagagtcaaa gagaagctcg tggtggtgtt  360
cgtgatgttg ggaagttcga aatgagaact gaagggggag aaaaaggaaa caaggataga  420
ccagaactga aaacacgaac cagagaggta ataggaagaa cggagaaaga aagaggaaga  480
gaaagtggag gacaagtagt ggcagagaaa ggaagagaaa cagaaactgc aaggggaagg  540
gttggtgctg aaaatgaagg tgcaagaaca accgctgtta taacttgcac gttggaaaaa  600
ggtggtggta ctcagaaacc aataagagaa gaagagagag agagtgaaag tgaaagatca  660
gcgtgggaac agatttccaa ctacagtgat caagccacac aaggtgtgaa ggaaaagtac  720
gagagagcaa agcaagcagc aagcgagaca ctgaacacca caacacaaac cgcacaagag  780
aaaagcgcac aggcaaagaa tcttgcagca caagcaaagg acgcaacgtt agagaagggt  840
cagcaaggct atgctgtgac aaaggacacc atttcaagtg cagcgaaaac cgcttcggag  900
aaaactgcgc cagtggcaga gaaggccaaa gactacactc ttcaggctgc ggagaaggcc  960
aagagtgcgg gtggcacaac ggcgagttac gtcggagaga aagcggtgca ggccaaggat  1020
gtcgccgtgg agagtgggaa gagtgcggct gggtatgctg ccaaggtggc tgctgatttg  1080
agggacaagg ccaccgcggt ggggtgggcc gcggcccatt tttcggcgga aaagacggtg  1140
gaagggacca aggccgccgc ccatgttgtc gagggagcgg ccgggtatgc cggacataag  1200
gccgcggagc ttgcttcgat gtcggccggt gctgttaaag ggctggctgc ttctgctggt  1260
gagactgcta aggaatacac cgcaaagaag aaagaggagg cacagcggga attggaggcc  1320
aaaaaaccat ctcaacctca ggaagctgaa gagaggccat cagaaggaat tggcgaaact  1380
gtgaggcagt atgcacaaaa accaaagcca agtgaaggaa accctcagaa ggagggaacc  1440
ggaagcatcg tgtttacagc aataggcgaa acagtaagca gtgccggtga gaaggtaaag  1500
aaaccgttta aaaacactat gggaggagaa agtgaaggag gaggagggaa agaagaggga  1560
aagagtgtca ttggtaaaag tttgaccagc attggtgaga agctgggtga tgcaaaacag  1620
agagaggagc tgttagacaa cgtcactggc aacatcaccg agggaggagg cgaggtgttg  1680
ggagcagtag gtgaaaccgt tgctgagatt ggacagaaca tgatgaagcc agcggagata  1740
```

-continued

```
gttcaggaac gcgctcacgt tcgacaagcg ggtggggtct tggatgccat tggtgaaacc  1800
attgctgaga ttgctgaaac gactcgagtc atggtttctg gagaggacga gagagtgctg  1860
aggcagagcg tggtgttgga gactcgcgta actggtcgtg ccaagcatga agaagggtcg  1920
cacggtgct                                                          1929

SEQ ID NO: 12          moltype = AA  length = 643
FEATURE                Location/Qualifiers
source                 1..643
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 12
MASEQLARRE NTTTEKEIHV EKHRVPKMAT HFEHLAEQAK ESDITAGKDT PQGSIEALQA  60
GERVKDHAGK AMGDIGGRGK ARETHELGAH FESLADKVTD HAAANVVGNK ESQREARGGV  120
RDVGKFEMRT EGGEKGNKDR PELKTRTREV IGRTEKERGR ESGGQVVAEK GRETETARGR  180
VGAENEGART TAVITCTLEK GGGTQKPIRE EERESESERS AWEQISNYSD QATQGVKEKY  240
ERAKQAASET LNTTTQTAQE KSAQAKNLAA QAKDATLEKG QQGYAVTKDT ISSAAKTASE  300
KTAPVAEKAK DYTLQAAEKA KSAGGTTASY VGEKAVQAKD VAVESGKSAA GYAAKVAADL  360
RDKATAVGWA AAHFSAEKTV EGTKAAAHVV EGAAGYAGHK AAELASMSAG AVKGLAASAG  420
ETAKEYTAKK KEEAQRELEA KKPSQPQEAE ERPSEGIGET VRQYAQKPKP SEGNPQKEGT  480
GSIVFTAIGE TVSSAGEKVK KPFKNTMGGE SEGGGGKEEG KSVIGKSLTS IGEKLGDAKQ  540
REELLDNVTG NITEGGGEVL GAVGETVAEI GQNMMKPAEI VQERAHVRQA GGVLDAIGET  600
IAEIAETTRV MVSGEDERVL RQSVVLETRV TGRAKHEEGS HGA                    643

SEQ ID NO: 13          moltype = DNA  length = 771
FEATURE                Location/Qualifiers
variation              242
                       note = a, c, g, or t
source                 1..771
                       mol_type = other DNA
                       organism = Glycine max
SEQUENCE: 13
atggcccttt cttgctccaa agttttaagc ttttatctct cgcccgttgt gggtggtggt  60
gatgttccga aaaagctcac tttttcttct tttttgggtc tttctaaggg tgttggtgga  120
agcaggagct cgagggtgtg tgctgcttca aatgcacccg caccactcgc tggggtcata  180
tttgaaccat ttcaagagct caagaaggat tatcttgctg ttccgattgc acacaatgtt  240
tncttggctc gtcagaacta tgcagatgac tctgaatctg caatcaacga gcagatcaat  300
gtggaataca acgtttccta tgtgtaccac gccttgtttg cgtacttcga cagagacaac  360
atagctctca aaggacttgc aaagttcttc aaggaatcaa gtgaagaaga aagagagcat  420
gccgagcagc ttataaagta tcagaacatt cgtggtggac gagtggtgct gcacccaatc  480
acaagtcctc cctcggaatt tgagcattcg gaaaaagggg atgccttgta tgccatggaa  540
ctagctttgt ctttggagaa gttgacgaat gagaaacttc tgcacgttca cagtgtggca  600
gaacgtaaca atgatcctca aytdgcagac ttcattgaga gcgagttctt gtatgagcag  660
gttaaatcga ttaaaaagat tgcagagtat gtggctcaac tgagattagt tggaaagggt  720
catggtgtgt ggcactttga tcaaaagctt cttcacgatg aagatcatgt a           771

SEQ ID NO: 14          moltype = AA  length = 257
FEATURE                Location/Qualifiers
VARIANT                81
                       note = any amino acid
VARIANT                208
                       note = any amino acid
source                 1..257
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 14
MALSCSKVLS FYLSPVVGGG DVPKKLTFSS FLGLSKGVGG SRSSRVCAAS NAPAPLAGVI  60
FEPFQELKKD YLAVPIAHNV XLARQNYADD SESAINEQIN VEYNVSYVYH ALFAYFDRDN  120
IALKGLAKFF KESSEEEREH AEQLIKYQNI RGGRVVLHPI TSPPSEFEHS EKGDALYAME  180
LALSLEKLTN EKLLHVHSVA ERNNDPQXAD FIESEFLYEQ VKSIKKIAEY VAQLRLVGKG  240
HGVWHFDQKL LHDEDHV                                                 257

SEQ ID NO: 15          moltype = DNA  length = 2421
FEATURE                Location/Qualifiers
source                 1..2421
                       mol_type = other DNA
                       organism = Glycine max
SEQUENCE: 15
atgtctcagc aaggcgaatc atcggacccg aaatcgggga agaaggactt ctcaaccgcg  60
attctggaac gcaaaaagtc gccgaaccgg cttgtcgtgg atgaggctgt caacgacgac  120
aactccgtgg tcaccatgca tccccaaacc atggaaaagc ttcagctttt ccgaggggac  180
actatcctca tcaagggtaa gaaacggaag gataccattt gcatagctct tgctgatgag  240
aactgtgagg agccgaagat caggatgaat aaagttgtga ggtcaaattt gcgggttcgg  300
cttggagatg ttgtgtctgt gcaccagtgc ccggatgtta aatatgggaa acgtgtgcac  360
attctgccta ttgatgatac cattgaaggt gtcactggca atctgtttga tgctttcttg  420
aaaccctatt tcttggaagc ttaccgtcct gtcagaaaag gagatctatt tcttgtccgg  480
ggagggatga gaagtgtaga gttcaaggtg gttgaaaccg atcctggaga atattgtgtt  540
```

-continued

```
gttgctccag acactgaaat tttctgtgag ggggagcctt tgaaaagaga agatgaagag  600
agactggatg aagtcgggta tgatgatgtg ggtggtgtca ggaaacaaat ggctcagatt  660
cgtgagttgg ttgagcttcc attgaggcat ccacagcttt ttaaatcaat tggtgtgaaa  720
ccacccaaag gaattttact ttatggaccc ccaggttctg ggaagacatt aatagcaaga  780
gctgtagcta atgaaactgg agctttttc ttttgtatta acggaccaga gattatgtcc  840
aaacttgctg gagagagtga aagcaatctc aggaaagcgt ttgaagaagc agaaaagaat  900
gcaccatcca ttatcttcat tgatgaaatt gattccattg cccccaagcg ggagaagaca  960
catggtgaag ttgaaaggag gattgtttca caacttttga ctcttatgga tggattaaaa  1020
tctcgtgctc atgttattgt tattggagct actaatcgtc caaatagcat tgacccagca  1080
ttgagaaggt ttggtagatt tgatagggaa attgatattg gtgttcccga tgaagttggg  1140
agacttgaag tccttcgcat acatactaaa aacatgaagc tatctgatga tgttgattta  1200
gagagaattg caaaagatac tcatgggtat gttggtgctg accttgctgc cctttgcact  1260
gaagctgctt tgcaatgcat tagggagaag atggatgtaa ttgacttgga ggatgagact  1320
attgatgcag aagtactgaa ttccatggca gttacaaatg agcatttcca gactgccctt  1380
ggaacaagca acccatcagc tttacgtgaa actgttgttg aagtgcctaa tgtcagctgg  1440
gaagacattg gaggccttga aaatgtcaag cgtgaactgc aagagactgt acaatatcca  1500
gttgagcacc cagaaaagtt tgagaagttt ggtatgtctc catctaaagg agtacttttc  1560
tatggtcctc caggatgtgg gaaaactttg ttagccaaag caattgctaa tgaatgtcaa  1620
gctaacttca tcagtgtgaa aggtccagaa ttgcttacaa tgtggtttgg tgaaagtgaa  1680
gctaatgtta gagaaatttt tgacaaggcc agacaatcgg ctccatgtgt cctcttcttt  1740
gatgagcttg actctattgc cactcagaga ggaagcagtg ttggggatgc tggtggtgcg  1800
gctgatagag ttctaaatca acttttgact gaaatggatg gcatgtctgc caaaaagact  1860
gtgttcataa ttggggcaac taacagacct gacatcattg acccagcact tcttcgacca  1920
ggccgtctgg atcaattgat ctatattcct ctccctgatg aggattcccg gcatcaaatc  1980
tttaaagcct gcttgaggaa gtcacctatc gcaaaaaatg ttgatttgag agcattggct  2040
aggcacactc aagggttcag tggtgctgat attacagaaa tatgccagcg agcatgcaaa  2100
tatgccataa gagaaaacat agagaaggac attgagcgag agaggaagag tagggaaaac  2160
cctgaggcca tggatgaaga taccgtcgac gacgaagttg cggagatcaa ggctgctcat  2220
tttgaggagt caatgaagtt tgcacggagg agtgttagtg atgctgatat tcgcaagtat  2280
caggcatttg ctcagacctt acagcaatca agagggtttg gaagtgaatt taggtttcca  2340
gagtctggag ataggactac tactggatct gaccccttg cagcttctgc tggtggggct  2400
gacgaagatg atctatatag t                                            2421

SEQ ID NO: 16          moltype = AA  length = 807
FEATURE                Location/Qualifiers
source                 1..807
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 16
MSQQGESSDP KSGKKDFSTA ILERKKSPNR LVVDEAVNDD NSVVTMHPQT MEKLQLFRGD  60
TILIKGKKRK DTICIALADE NCEEPKIRMN KVVRSNLRVR LGDVVSVHQC PDVKYGKRVH  120
ILPIDDTIEG VTGNLFDAFL KPYFLEAYRP VRKGDLFLVR GGMRSVEFKV VETDPGEYCV  180
VAPDTEIFCE GEPLKREDEE RLDEVGYDDV GGVRKQMAQI RELVELPLRH PQLFKSIGVK  240
PPKGILLYGP PGSGKTLIAR AVANETGAFF FCINGPEIMS KLAGESESNL RKAFEEAEKN  300
APSIIFIDEI DSIAPKREKT HGEVERRIVS QLLTLMDGLK SRAHVIVIGA TNRPNSIDPA  360
LRRFGRFDRE IDIGVPDEVG RLEVLRIHTK NMKLSDDVDL ERIAKDTHGY VGADLAALCT  420
EAALQCIREK MDVIDLEDET IDAEVLNSMA VTNEHFQTAL GTSNPSALRE TVVEVPNVSW  480
EDIGGLENVK RELQETVQYP VEHPEKFEKF GMSPSKGVLF YGPPGCGKTL LAKAIANECQ  540
ANFISVKGPE LLTMWFGESE ANVREIFDKA RQSAPCVLFF DELDSIATQR GSSVGDAGGA  600
ADRVLNQLLT EMDGMSAKKT VFIIGATNRP DIIDPALLRP GRLDQLIYIP LPDEDSRHQI  660
FKACLRKSPI AKNVDLRALA RHTQGFSGAD ITEICQRACK YAIRENIEKD IERERKSREN  720
PEAMDEDTVD DEVAEIKAAH FEESMKFARR SVSDADIRKY QAFAQTLQQS RGFGSEFRFP  780
ESGDRTTTGS DPFAASAGGA DEDDLYS                                      807

SEQ ID NO: 17          moltype = AA  length = 13
FEATURE                Location/Qualifiers
source                 1..13
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 17
DQSSSYDDDE YSK                                                     13

SEQ ID NO: 18          moltype = AA  length = 15
FEATURE                Location/Qualifiers
source                 1..15
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 18
KSDQSSSYDD DEYSK                                                   15

SEQ ID NO: 19          moltype = AA  length = 14
FEATURE                Location/Qualifiers
source                 1..14
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 19
```

-continued

```
SDQSSSYDDD EYSK                                                    14

SEQ ID NO: 20          moltype = AA  length = 9
FEATURE                Location/Qualifiers
source                 1..9
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 20
SYDDDEYSK                                                          9

SEQ ID NO: 21          moltype = AA  length = 18
FEATURE                Location/Qualifiers
source                 1..18
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 21
ALRALRLEDL RIPTAYIK                                                18

SEQ ID NO: 22          moltype = AA  length = 18
FEATURE                Location/Qualifiers
source                 1..18
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 22
ALRLSGGDHV HAGTVVGK                                                18

SEQ ID NO: 23          moltype = AA  length = 22
FEATURE                Location/Qualifiers
source                 1..22
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 23
LEGEREITLG FVDLLRDDFV EK                                           22

SEQ ID NO: 24          moltype = AA  length = 14
FEATURE                Location/Qualifiers
source                 1..14
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 24
LRLEDLRIPT AYIK                                                    14

SEQ ID NO: 25          moltype = AA  length = 11
FEATURE                Location/Qualifiers
source                 1..11
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 25
LTYYTPDYET K                                                       11

SEQ ID NO: 26          moltype = AA  length = 11
FEATURE                Location/Qualifiers
source                 1..11
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 26
TDGLTSLDRY K                                                       11

SEQ ID NO: 27          moltype = AA  length = 14
FEATURE                Location/Qualifiers
source                 1..14
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 27
TFQGPPHGIQ VERD                                                    14

SEQ ID NO: 28          moltype = AA  length = 15
FEATURE                Location/Qualifiers
source                 1..15
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 28
TFQGPPHGIQ VERDK                                                   15

SEQ ID NO: 29          moltype = AA  length = 24
FEATURE                Location/Qualifiers
```

-continued

```
source                  1..24
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 29
APHTEDAIRV TQSTANFFIS EARK                                          24

SEQ ID NO: 30           moltype = AA  length = 23
FEATURE                 Location/Qualifiers
source                  1..23
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 30
FVESGGARVI PLIYNESPEN LNK                                           23

SEQ ID NO: 31           moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 31
IWDLESK                                                             7

SEQ ID NO: 32           moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 32
SIVEDLK                                                             7

SEQ ID NO: 33           moltype = AA  length = 21
FEATURE                 Location/Qualifiers
source                  1..21
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 33
QELDERARQG ETVVPGGTGG K                                             21

SEQ ID NO: 34           moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 34
SGEERAQEEA IDIDESK                                                  17

SEQ ID NO: 35           moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 35
AAAHVVEGAA GYAGHK                                                   16

SEQ ID NO: 36           moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 36
AKDYTLQAAE K                                                        11

SEQ ID NO: 37           moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 37
ARETHELGAH FESLADK                                                  17

SEQ ID NO: 38           moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 38
```

-continued

```
ATAVGWAAAH FSAEK                                                    15

SEQ ID NO: 39          moltype = AA  length = 17
FEATURE                Location/Qualifiers
source                 1..17
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 39
DTPQGSIEAL QAGERVK                                                  17

SEQ ID NO: 40          moltype = AA  length = 9
FEATURE                Location/Qualifiers
source                 1..9
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 40
DYTLQAAEK                                                           9

SEQ ID NO: 41          moltype = AA  length = 10
FEATURE                Location/Qualifiers
source                 1..10
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 41
EEAQRELEAK                                                          10

SEQ ID NO: 42          moltype = AA  length = 21
FEATURE                Location/Qualifiers
source                 1..21
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 42
EGTGSIVFTA IGETVSSAGE K                                             21

SEQ ID NO: 43          moltype = AA  length = 14
FEATURE                Location/Qualifiers
source                 1..14
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 43
ERGRESGGQV VAEK                                                     14

SEQ ID NO: 44          moltype = AA  length = 11
FEATURE                Location/Qualifiers
source                 1..11
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 44
GLAASAGETA K                                                        11

SEQ ID NO: 45          moltype = AA  length = 9
FEATURE                Location/Qualifiers
source                 1..9
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 45
GQQGYAVTK                                                           9

SEQ ID NO: 46          moltype = AA  length = 26
FEATURE                Location/Qualifiers
source                 1..26
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 46
KPSQPQEAEE RPSEGIGETV RQYAQK                                        26

SEQ ID NO: 47          moltype = AA  length = 8
FEATURE                Location/Qualifiers
source                 1..8
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 47
PAEIVQER                                                            8

SEQ ID NO: 48          moltype = AA  length = 15
FEATURE                Location/Qualifiers
```

-continued

```
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 48
PQGSIEALQA GERVK                                                15

SEQ ID NO: 49           moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 49
QAASETLNTT TQTAQEK                                              17

SEQ ID NO: 50           moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 50
SAGGTTASYV GEK                                                  13

SEQ ID NO: 51           moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 51
SLTSIGEK                                                        8

SEQ ID NO: 52           moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 52
TPQGSIEALQ AGERVK                                               16

SEQ ID NO: 53           moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 53
VTDHAAANVV GNK                                                  13

SEQ ID NO: 54           moltype = AA  length = 22
FEATURE                 Location/Qualifiers
source                  1..22
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 54
AASNAPAPLA GVIFEPFQEL KK                                        22

SEQ ID NO: 55           moltype = AA  length = 21
FEATURE                 Location/Qualifiers
source                  1..21
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 55
ASNAPAPLAG VIFEPFQELK K                                         21

SEQ ID NO: 56           moltype = AA  length = 18
FEATURE                 Location/Qualifiers
source                  1..18
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 56
FFKESSEEER EHAEQLIK                                             18

SEQ ID NO: 57           moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 57
```

-continued

```
GHGVWHFDQK                                                          10

SEQ ID NO: 58           moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 58
IAEYVAQLRL VGK                                                      13

SEQ ID NO: 59           moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 59
KIAEYVAQLR LVGK                                                     14

SEQ ID NO: 60           moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 60
LLHDEDHV                                                            8

SEQ ID NO: 61           moltype = AA  length = 20
FEATURE                 Location/Qualifiers
source                  1..20
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 61
SNAPAPLAGV IFEPFQELKK                                               20

SEQ ID NO: 62           moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 62
AGESESNLRK                                                          10

SEQ ID NO: 63           moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 63
GESEANVREI FDK                                                      13

SEQ ID NO: 64           moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 64
GILLYGPPGS GK                                                       12

SEQ ID NO: 65           moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 65
LAGESESNLR K                                                        11

SEQ ID NO: 66           moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 66
LSDDVDLERI AK                                                       12

SEQ ID NO: 67           moltype = AA  length = 16
FEATURE                 Location/Qualifiers
```

-continued

```
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 67
RELQETVQYP VEHPEK                                                   16

SEQ ID NO: 68           moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 68
TVFIIGATNR PDIID                                                    15

SEQ ID NO: 69           moltype = DNA  length = 2448
FEATURE                 Location/Qualifiers
source                  1..2448
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 69
atgaggaagt ggacggttgc ttccgctctc cttcttctct ctctcctctt tctctttgca  60
gatcaaggtc gaaagtttca ggcgaatgcc gagggagatt ccgacgagct tgtagatcca  120
cccaaggtgg aggacaagat cggtgctgtg cctcatggcc tttcaaccga ttcagatgtc  180
gtcaagaggg aggcggagtc gatctcgaag agatctctcc ggagcaacgc ggagaagttt  240
gagttccaag cggaagtgtc gcggcttatg gatatcatca tcaattctct ctacagtaac  300
aaggatatct tcctcagaga gttgatttcc aatgcttctg atgctttgga caagattagg  360
ttcctctccc tcacggataa ggaggtgttg ggagaaggtg ataacaccaa gcttgatatt  420
cagattaagt tggacaagga aaagaaaatt ctctcaattc gagacagagg tatcggtatg  480
acgaaggagg atttgataaa gaatttgggg acaatagcga aatctggaac ttctgcattt  540
gttgagaaga tgcaaacaag tggagatctc aatctgattg ggcagtttgg agtcggcttc  600
tactctgttt atcttgtggc cgactatgtc gaagttatta gcaagaataa cgatgataag  660
cagtatgtat gggaatcaaa agctgatgga gcattcgcta tttcagaaga tacatggaat  720
gaaccacttg gacgaggaac tgagattaga ttacacctca aggaagaggc tggggaatac  780
ttggaggagt ctaaactgaa agaattggtg aagagatact ctgaatttat taacttccct  840
atctatattt gggcaagcaa agaagttgat gtggaggttc ccgctgatga agatgattcc  900
agtgatgaag aggattcatc tgaaagcagc tccaaggagg aaagcgaaga tgaagatgct  960
gataaaagtg aagatgaaga aaaaaagcct aagacaaaga cagtgaagga aacaacttat  1020
gagtgggaac ttttaaatga tgtgaaagct atatggctga ggaatccaaa agaggttaca  1080
gaagaagaat acaccaaatt ctaccactct cttgccaagg atttcagtga tgagaaacct  1140
ttagcatgga gtcatttcac tgctgagggt gatgttgagt tcaaggcagt cctgtttgta  1200
ccacctaagg cacctcaaga tttatatgag agctattata acgcaaacaa atcgaacctg  1260
aagttgtatg ttagacgtgt gttcatttca gatgaattta atgaactgtt gcccaagtac  1320
ctaaactttt tgctgggtct tgttgattct gacactttgc cacttaatgt atcacgagaa  1380
atgcttcaac aacatagtag tctgaagaca atcaagaaga aacttattag gaaagccctt  1440
gatatgatcc gcaggattgc tgatgaggat cctgacgagt caactgacaa agaaaagaaa  1500
gatgcatcct ctgacaatga tgagaagaaa ggtcaatatt ctaagttctg gaatgagttt  1560
ggtaaatcca ttaaacttgg tatcatcgag gatgccacca acagaaatcg tttggcaaaa  1620
ctgctcagat ttgagagcac caagtctgag ggtaaattga catctctgga tcagtacata  1680
tctagaatga aaactgggca gaaggatatc ttttacataa ctggaaccag caaagaacaa  1740
ctggaaaact ctcccttcct tgagcggctt aagaagaaaa attttgaggt tattttcttc  1800
acggatccag ttgatgagta cttgatgcaa tacctgatgg attatgaaga caagaagttc  1860
caaaatgtgt ccaaggaggg tttgaaactt gggaaagatt caaaagacaa ggaactgaag  1920
gaatctttta aggatcttac caagtggtgg aaaactgctc tttccaaaga caatgttgat  1980
gatgtgaaga tttccaaccg attagataac actccttgtg tagtggtgac atcaaaattt  2040
ggttggagtg ctaacatgga gagaatcatg cagtctcaaa ctttgtcaga tgctagcaag  2100
caagcataca tgcgcggcaa gagggttctt gagatcaatc ctaggcaccc tattatcaag  2160
gagctccgag agagagtagt aaagaaccct gaggatgagg gtgtgaagca tacagcacag  2220
ctgatgtacc agactgcact ctttgaaagc ggcttcctgc ttgatgatcc caaggatttt  2280
gcttcccgga tttatgattc agtgaagact agcctggaca tcagtcctga agcaacagtt  2340
gaggaggaag atgacactga agtagaggct gaaagtgacg cgaaagaaga ttcaaagcct  2400
gaaactgatg ctgttaacga tgataatgat gtcaaggacg agttgtag               2448

SEQ ID NO: 70           moltype = AA  length = 815
FEATURE                 Location/Qualifiers
source                  1..815
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 70
MRKWTVASAL LLLSLLFLFA DQGRKFQANA EGDSDELVDP PKVEDKIGAV PHGLSTDSDV  60
VKREAESISK RSLRSNAEKF EFQAEVSRLM DIIINSLYSN KDIFLRELIS NASDALDKIR  120
FLSLTDKEVL GEGDNTKLDI QIKLDKEKKI LSIRDRGIGM TKEDLIKNLG TIAKSGTSAF  180
VEKMQTSGDL NLIGQFGVGF YSVYLVADYV EVISKNNDDK QYVWESKADG AFAISEDTWN  240
EPLGRGTEIR LHLKEEAGEY LEESKLKELV KRYSEFINFP IYIWASKEVD VEVPADEDDS  300
SDEEDSSESS SKEESEDEDA DKSEDEEKKP KTKTVKETTY EWELLNDVKA IWLRNPKEVT  360
EEEYTKFYHS LAKDFSDEKP LAWSHFTAEG DVEFKAVLFV PPKAPQDLYE SYYNANKSNL  420
KLYVRRVFIS DEFNELLPKY LNFLLGLVDS DTLPLNVSRE MLQQHSSLKT IKKKLIRKAL  480
```

-continued

```
DMIRRIADED PDESTDKEKK DASSDNDEKK GQYSKFWNEF GKSIKLGIIE DATNRNRLAK  540
LLRFESTKSE GKLTSLDQYI SRMKTGQKDI FYITGTSKEQ LENSPFLERL KKKNFEVIFF  600
TDPVDEYLMQ YLMDYEDKKF QNVSKEGLKL GKDSKDKELK ESFKDLTKWW KTALSKDNVD  660
DVKISNRLDN TPCVVVTSKF GWSANMERIM QSQTLSDASK QAYMRGKRVL EINPRHPIIK  720
ELRERVVKNP EDEGVKHTAQ LMYQTALFES GFLLDDPKDF ASRIYDSVKT SLDISPEATV  780
EEEDDTEVEA ESDAKEDSKP ETDAVNDDND VKDEL                             815

SEQ ID NO: 71          moltype = AA  length = 14
FEATURE                Location/Qualifiers
source                 1..14
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 71
APQDLYESYY NANK                                                    14

SEQ ID NO: 72          moltype = AA  length = 8
FEATURE                Location/Qualifiers
source                 1..8
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 72
AVLFVPPK                                                           8

SEQ ID NO: 73          moltype = AA  length = 11
FEATURE                Location/Qualifiers
source                 1..11
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 73
DFASRIYDSV K                                                       11

SEQ ID NO: 74          moltype = AA  length = 17
FEATURE                Location/Qualifiers
source                 1..17
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 74
DIFLRELISN ASDALDK                                                 17

SEQ ID NO: 75          moltype = AA  length = 10
FEATURE                Location/Qualifiers
source                 1..10
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 75
DIFYITGTSK                                                         10

SEQ ID NO: 76          moltype = AA  length = 7
FEATURE                Location/Qualifiers
source                 1..7
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 76
DNVDDVK                                                            7

SEQ ID NO: 77          moltype = AA  length = 17
FEATURE                Location/Qualifiers
source                 1..17
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 77
EDSKPETDAV NDDNDVK                                                 17

SEQ ID NO: 78          moltype = AA  length = 11
FEATURE                Location/Qualifiers
source                 1..11
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 78
EEAGEYLEES K                                                       11

SEQ ID NO: 79          moltype = AA  length = 13
FEATURE                Location/Qualifiers
source                 1..13
                       mol_type = protein
                       organism = Glycine max
```

-continued

```
SEQUENCE: 79
EQLENSPFLE RLK                                                  13

SEQ ID NO: 80          moltype = AA  length = 10
FEATURE                Location/Qualifiers
source                 1..10
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 80
EVLGEGDNTK                                                      10

SEQ ID NO: 81          moltype = AA  length = 9
FEATURE                Location/Qualifiers
source                 1..9
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 81
EVTEEEYTK                                                       9

SEQ ID NO: 82          moltype = AA  length = 7
FEATURE                Location/Qualifiers
source                 1..7
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 82
FYHSLAK                                                         7

SEQ ID NO: 83          moltype = AA  length = 16
FEATURE                Location/Qualifiers
source                 1..16
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 83
IGAVPHGLST DSDVVK                                               16

SEQ ID NO: 84          moltype = AA  length = 15
FEATURE                Location/Qualifiers
source                 1..15
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 84
LGIIEDATNR NRLAK                                                15

SEQ ID NO: 85          moltype = AA  length = 8
FEATURE                Location/Qualifiers
source                 1..8
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 85
NPEDEGVK                                                        8

SEQ ID NO: 86          moltype = AA  length = 9
FEATURE                Location/Qualifiers
source                 1..9
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 86
SGTSAFVEK                                                       9

SEQ ID NO: 87          moltype = AA  length = 26
FEATURE                Location/Qualifiers
source                 1..26
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 87
TSLDISPEAT VEEEDDTEVE AESDAK                                    26

SEQ ID NO: 88          moltype = DNA  length = 2442
FEATURE                Location/Qualifiers
source                 1..2442
                       mol_type = other DNA
                       organism = Glycine max
SEQUENCE: 88
atgaggaagt ggacggttgc ttccgctctc cttcttctct ctctcctctt tctctttgca  60
gatcaaggtc gaaagtttca ggcgaatgcc gagggagatt ccgacgagct tgtagatcct  120
ccgaaggtgg aggacaagat cggtgctgtg cctcatggcc tttccaccga ttcagatgtc  180
```

-continued

```
gtcaagaggg agtcggagtc gatctcgaag agatctctcc ggagcaatgc ggagaagttt  240
gagttccaag cggaagtgtc acggcttatg gatattatca tcaattctct ctacagtaac  300
aaggatatct tcctcagaga gttgatttcc aatgcttctg atgctttgga caagattagg  360
ttcctttccc tcacggataa ggacgttttg ggagaaggtg ataacaccaa gcttgatatt  420
cagattaagt tggacaagga aaagaaaatt ctctcaattc gagacagagg tattggtatg  480
acgaaggagg atttgataaa gaatttgggg acaatagcaa aatctggaac ttctgcattt  540
gttgagaaga tgcaaacaag tggagatctc aatctgattg ggcagtttgg agtcggcttc  600
tactctgttt atctcgtggc tgactatgtc gaagtcatta gcaagaataa tgatgataaa  660
cagtatgtat gggaatcaaa agctgatgga gcatttgcaa tttcagaaga tacatggaat  720
gaacctctag ggcgtggaac tgagattaga ttacacctca aggaagaggc tggggagtac  780
ttgcaggagt ctaaactgaa agaattggtg aagagatact ctgaatttat taacttccct  840
atctatattt gggcaagcaa agaggttgat gtggaggttc ccgctgatga agatgattcc  900
agtgatgaag aggattcatc tgaaagcagc tccaaggagg aaagcgaaga tgaagatgct  960
gataaaagtg aagacgaaga aaaaaagccc aagacaaaga cggtgaagga aacaacttat  1020
gaatgggaac ttttaaatga tgtgaaagct atatggctga ggaatccaaa ggaggttaca  1080
gaagaagaat acaccaaatt ctaccactct cttgccaagg atttcagtga tgagaaacct  1140
ttagcatgga gtcatttcac tgctgagggc gatgtcgagt tcaaggcagt cctgtttgta  1200
ccacctaagg cacctcaaga tttatacgag agctattata acgcaaacaa atccaacctg  1260
aagttgtatg ttagacgtgt gttcatttca gatgaattta atgaactgtt gcctaagtac  1320
ctaaactttt tgctgggtct tgttgattct gacacattgc cacttaatgt atcgcgagaa  1380
atgcttcagc aacatagtag tctaaagaca atcaagaaga aacttattag gaaagccctt  1440
gatatgatcc gtaggattgc tgatgaggat cctgatgagt ccactgacaa agaaaagaaa  1500
gatacgtcat ctgacaataa tgagaagaaa ggtcaatatt ccaagttctg gaatgagttt  1560
ggtaaatcca ttaaacttgg tatcatcgag gatgccacca acagaaatcg tttggcaaaa  1620
ctgctcagat ttgagagcac caagtctgag ggtaaattga catctctgga tcagtacata  1680
tctagaatga aagctgggca gaaggatatc ttttacataa ctggaaccag caaagaacaa  1740
ctggaaaact ctccattcct tgagcggctt aagaagaaaa attttgaggt tattttcttc  1800
acggatccag ttgatgaata cttgatgcaa tacctgatgg attatgaaga caagaagttc  1860
caaaatgtgt ccaaggaggg tttgaaactc gggaaagatt caaaagacaa ggaactgaag  1920
gaatccttca aggaccttac caagtggtgg aaaactgctc tttccaaaga caatgttgat  1980
gatgtgaaga tttccaaccg attggacaac actccttgtg tagtggtaac atcaaaattt  2040
ggttggagtg ctaacatgga gagaatcatg cagtctcaaa ctttgtctga tgctagcaag  2100
caagcataca tgcgcggcaa gagggttctt gaggtcaatc ctaggcaccc tattatcaag  2160
gagctccgag agagagtagt aaagaaccct gaggatgagg gtgtgaaaca tacagcacag  2220
ttgatgtacc agaccgcact cttcgaaagt ggcttcctgc ttgatgatcc caaggatttt  2280
gcttcccgta tttatgattc agttaagacc agcttggaca tcagtcctga agcaacagtt  2340
gaggaggaag atgacaccga agtagaggct gaaagtgatt caaagcctga agctgatgcg  2400
gttcatgatg gtgatgttga tgatgccaag gacgagttgt ag                     2442

SEQ ID NO: 89          moltype = AA  length = 813
FEATURE                Location/Qualifiers
source                 1..813
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 89
MRKWTVASAL LLLSLLFLFA DQGRKFQANA EGDSDELVDP PKVEDKIGAV PHGLSTDSDV  60
VKRESESISK RSLRSNAEKF EFQAEVSRLM DIIINSLYSN KDIFLRELIS NASDALDKIR  120
FLSLTDKDVL GEGDNTKLDI QIKLDKEKKI LSIRDRGIGM TKEDLIKNLG TIAKSGTSAF  180
VEKMQTSGDL NLIGQFGVGF YSVYLVADYV EVISKNNDDK QYVWESKADG AFAISEDTWN  240
EPLGRGTEIR LHLKEEAGEY LQESKLKELV KRYSEFINFP IYIWASKEVD VEVPADEDDS  300
SDEEDSSESS SKEESEDEDA DKSEDEEKKP KTKTVKETTY EWELLNDVKA IWLRNPKEVT  360
EEEYTKFYHS LAKDFSDEKP LAWSHFTAEG DVEFKAVLFV PPKAPQDLYE SYYNANKSNL  420
KLYVRRVFIS DEFNELLPKY LNFLLGLVDS DTLPLNVSRE MLQQHSSLKT IKKKLIRKAL  480
DMIRRIADED PDESTDKEKK DTSSDNNEKK GQYSKFWNEF GKSIKLGIIE DATNRNRLAK  540
LLRFESTKSE GKLTSLDQYI SRMKAGQKDI FYITGTSKEQ LENSPFLERL KKKNFEVIFF  600
TDPVDEYLMQ YLMDYEDKKF QNVSKEGLKL GKDSKDKELK ESFKDLTKWW KTALSKDNVD  660
DVKISNRLDN TPCVVVTSKF GWSANMERIM QSQTLSDASK QAYMRGKRVL EVNPRHPIIK  720
ELRERVVKNP EDEGVKHTAQ LMYQTALFES GFLLDDPKDF ASRIYDSVKT SLDISPEATV  780
EEEDDTEVEA ESDSKPEADA VHDGDVDDAK DEL                               813

SEQ ID NO: 90          moltype = AA  length = 14
FEATURE                Location/Qualifiers
source                 1..14
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 90
APQDLYESYY NANK                                                    14

SEQ ID NO: 91          moltype = AA  length = 8
FEATURE                Location/Qualifiers
source                 1..8
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 91
AVLFVPPK                                                           8
```

-continued

```
SEQ ID NO: 92          moltype = AA  length = 11
FEATURE                Location/Qualifiers
source                 1..11
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 92
DFASRIYDSV K                                                    11

SEQ ID NO: 93          moltype = AA  length = 17
FEATURE                Location/Qualifiers
source                 1..17
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 93
DIFLRELISN ASDALDK                                              17

SEQ ID NO: 94          moltype = AA  length = 10
FEATURE                Location/Qualifiers
source                 1..10
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 94
DIFYITGTSK                                                      10

SEQ ID NO: 95          moltype = AA  length = 7
FEATURE                Location/Qualifiers
source                 1..7
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 95
DNVDDVK                                                         7

SEQ ID NO: 96          moltype = AA  length = 10
FEATURE                Location/Qualifiers
source                 1..10
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 96
DVLGEGDNTK                                                      10

SEQ ID NO: 97          moltype = AA  length = 11
FEATURE                Location/Qualifiers
source                 1..11
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 97
EEAGEYLQES K                                                    11

SEQ ID NO: 98          moltype = AA  length = 13
FEATURE                Location/Qualifiers
source                 1..13
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 98
EQLENSPFLE RLK                                                  13

SEQ ID NO: 99          moltype = AA  length = 9
FEATURE                Location/Qualifiers
source                 1..9
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 99
EVTEEEYTK                                                       9

SEQ ID NO: 100         moltype = AA  length = 7
FEATURE                Location/Qualifiers
source                 1..7
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 100
FYHSLAK                                                         7

SEQ ID NO: 101         moltype = AA  length = 16
FEATURE                Location/Qualifiers
source                 1..16
                       mol_type = protein
```

-continued

```
                        organism = Glycine max
SEQUENCE: 101
IGAVPHGLST DSDVVK                                                   16

SEQ ID NO: 102          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 102
LGIIEDATNR NRLAK                                                    15

SEQ ID NO: 103          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 103
NPEDEGVK                                                            8

SEQ ID NO: 104          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 104
SGTSAFVEK                                                           9

SEQ ID NO: 105          moltype = DNA  length = 2103
FEATURE                 Location/Qualifiers
source                  1..2103
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 105
atggcttcgg agactgagac gttcgctttt caggccgaga tcaaccagct cctcagtcta  60
ataatcaaca ccttctacag caacaaggag atctttcttc gtgagctcat cagcaatgcc  120
tctgacgctt tggacaagat tcgatttgag agtcttacag acaagagcaa gcttgatgct  180
cagccagagt tgttcattca cattattcct gacaagacca acaacacatt gtctatcatt  240
gacagtggta ttggcatgac taaggctgat ttggtgaata acctcggtac tattgcaagg  300
tctgggacca aggaattcat ggaagccctg gctgctggtg ctgacgtgag catgattggg  360
cagttcggtg ttggtttcta ctcggcttat cttgtagctg ataaggttat tgttaccaca  420
aagcacaatg atgatgaaca gtatgtctgg gagtcccatg ccggtggctc atttactgtg  480
acaagagata catctggtga gaaccttggt aggggaacaa agataacttt gtttcttaag  540
gaagatcaac ttgagtatct tgaggaacgc cgtctgaagg atttgataaa gaagcattct  600
gaattcatta gctatccaat ttctctctgg attgagaaga ctactgaaaa ggagatttct  660
gatgatgagg atgaggaaga gaagaaggat gaggagggta aagttgaaga tgtggatgaa  720
gataaggaga aggaggagaa gaaaaagaag accattaagg aggtatctca cgagtggtca  780
ttggtgaaca agcagaagcc tatttggatg aggaaacctg aagagattac gaaagaagaa  840
tatgctgctt tctacaagag tcttaccaat gattgggaag agcatttggc tgttaagcac  900
ttttctgttg agggtcagct ggagtttaag gctgtcctct ttatccccaa gagggcacct  960
tttgatctct ttgacaccag gaagaagcct aacaacatca aactgtatgt ccgccgtgtc  1020
ttcatcatgg acaactgtga ggagttgatg cctgaatatc ttagctttgt taagggtatt  1080
gtggattctg aggatcttcc tctcaacatt tctagagaaa tgctgcagca gaacaagatc  1140
ctgaaagtca tcaggaagaa cttggtcaag aagtgcattg agatgttctt tgaaattgct  1200
gaaaacaagg aagactataa caagttttat gaagccttct ctaagaacct gaaacttggt  1260
attcatgagg attctcagaa caagacaaag ctagctgaat tgctcaggta tcactccact  1320
aagagtggtg atgagatgac cagcctcaag gactatgtta ccaggatgaa ggaaggacag  1380
aatgacatct actacattac tggtgaaagc aagaaagctg tcgagaattc ccccttcctt  1440
gaaaagctta agaagaaggg atacgaggtt ctctacatgg ttgatgctat tgatgagtat  1500
gctgttggcc agcttaagga atttgaggga aagaagttgg tctctgctac caaggaaggc  1560
ctaaaacttg atgagagcga agacgagaag aaaaagaagg aagaactaaa ggataaattt  1620
gagggtcttt gccatgtgat caaggatgtg ttgggtgaca aggtggagaa agttgtggtg  1680
tctgatcgtg ttgtggattc tccttgctgt ctggtgacag gtgaatatgg gtggacagcc  1740
aacatggaaa ggattatgaa agcacaggca ttgagggaca gcagcatggc tggatacatg  1800
tcaagcaaga agacgatgga gattaaccct gagaacccca tcatggagga gctgaggaag  1860
cgagcagatg ctgataagaa tgacaagtct gtgaaggatc tcgtgctctt gctctttgag  1920
actgctcttc taacttctgg gttcagcctt gatgatccca acactttcgg taacaggatt  1980
cacaggatgc tgaagcttgg actgagcatt gatgaagatg ctggtgaagc agatgctgac  2040
atgcctcctc ttgaggacgc tgatgcagat gctgagggta gcaagatgga agaagttgat  2100
taa                                                                 2103

SEQ ID NO: 106          moltype = AA  length = 700
FEATURE                 Location/Qualifiers
source                  1..700
                        mol_type = protein
                        organism = Glycine max
```

-continued

```
SEQUENCE: 106
MASETETFAF QAEINQLLSL IINTFYSNKE IFLRELISNA SDALDKIRFE SLTDKSKLDA  60
QPELFIHIIP DKTNNTLSII DSGIGMTKAD LVNNLGTIAR SGTKEFMEAL AAGADVSMIG  120
QFGVGFYSAY LVADKVIVTT KHNDDEQYVW ESHAGGSFTV TRDTSGENLG RGTKITLFLK  180
EDQLEYLEER RLKDLIKKHS EFISYPISLW IEKTTEKEIS DDEDEEEKKD EEGKVEDVDE  240
DKEKEEKKKK TIKEVSHEWS LVNKQKPIWM RKPEEITKEE YAAFYKSLTN DWEEHLAVKH  300
FSVEGQLEFK AVLFIPKRAP FDLFDTRKKP NNIKLYVRRV FIMDNCEELM PEYLSFVKGI  360
VDSEDLPLNI SREMLQQNKI LKVIRKNLVK KCIEMFFEIA ENKEDYNKFY EAFSKNLKLG  420
IHEDSQNKTK LAELLRYHST KSGDEMTSLK DYVTRMKEGQ NDIYYITGES KKAVENSPFL  480
EKLKKKGYEV LYMVDAIDEY AVGQLKEFEG KKLVSATKEG LKLDESEDEK KKKEELKDKF  540
EGLCHVIKDV LGDKVEKVVV SDRVVDSPCC LVTGEYGWTA NMERIMKAQA LRDSSMAGYM  600
SSKKTMEINP ENPIMEELRK RADADKNDKS VKDLVLLLFE TALLTSGFSL DDPNTFGNRI  660
HRMLKLGLSI DEDAGEADAD MPPLEDADAD AEGSKMEEVD                        700

SEQ ID NO: 107          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 107
ADLVNNLGTI ARSGTK                                                  16

SEQ ID NO: 108          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 108
AVENSPFLEK                                                         10

SEQ ID NO: 109          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 109
EDQLEYLEER RLK                                                     13

SEQ ID NO: 110          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 110
EEYAAFYK                                                           8

SEQ ID NO: 111          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 111
EGQNDIYYIT GESK                                                    14

SEQ ID NO: 112          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 112
EIFLRELISN ASDALDK                                                 17

SEQ ID NO: 113          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 113
FYEAFSK                                                            7

SEQ ID NO: 114          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 114
```

-continued

```
HFSVEGQLEF K                                                        11

SEQ ID NO: 115          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 115
IRFESLTDK                                                           9

SEQ ID NO: 116          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 116
LAELLRYHST K                                                        11

SEQ ID NO: 117          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 117
LDESEDEK                                                            8

SEQ ID NO: 118          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 118
LGIHEDSQNK                                                          10

SEQ ID NO: 119          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 119
RAPFDLFDTR K                                                        11

SEQ ID NO: 120          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 120
VEDVDEDK                                                            8

SEQ ID NO: 121          moltype = DNA  length = 2109
FEATURE                 Location/Qualifiers
source                  1..2109
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 121
atggcggaga cagagacgtt cgcattccaa gcagagatca accagctcct cagtctcatc  60
atcaacacct tctactccaa caaggagatt ttccttcgtg aactcatcag caacgcctcc  120
gatgctttgg acaagattcg atttgagagt ttgacagata agagcaagct cgatgctcag  180
ccagagttgt tcattcatat tattccagac aagactaaca atacactgac cattgttgac  240
agtggtattg gcatgaccaa agctgatttg gtgaacaact tgggtacaat tgctagatct  300
ggtaccaagg aattcatgga agccttggct gctggtgctg atgttagtat gattggacag  360
tttggtgtgg ggttctactc tgcatatctt gttgctgaga aggtcattgt tacaaccaaa  420
cacaatgatg atgaacagta tgtttgggag tctcaagctg gaggttcttt cactgtcacc  480
agggataaca ctggtgaggt tcttggcagg ggtactaaga tcactctcta cctcaaggaa  540
gatcagcttg agtaccttga ggagcgtcgc ttgaaggatc tgatcaagaa gcattctgag  600
ttcattagct acccaatttc tctttggatt gagaagacta ctgagaagga gatctctgac  660
gatgaggatg aggaagagaa gaaggatgag gagggtaagg tggaggatgt tgacgaagag  720
aaggagaagg aagaaaaaaa gaaaaagaag atcaaggaag tgtcacatga atggtccttg  780
gtgaacaagc agaagcccat ttggatgagg aagcctgagg agatcacaaa agaggagtat  840
tctgctttct acaagagtct taccaatgac tgggaagagc atttggctgt caagcacttc  900
tctgttgagg gtcagttgga gtttaaggct atactctttg ttcccaagag ggcacctttt  960
gacctctttg acacaaggaa gaagcctaac aacattaagc tctatgtccg ccgtgtcttt  1020
attatggaca actgtgagga gttgattcct gagtatttgg gctttgtcaa gggtattgtt  1080
gattctgaag atctcccact caacatttca agagaaatgt tgcaacagaa caagatcttg  1140
aaggtcatcc gtaagaactt ggtcaagaag tgccttgagc tcttctttga aattgcagag  1200
```

-continued

```
aacaaggagg attacaacaa gttctatgag gctttctcta agaacttaaa gcttggtatc  1260
catgaggatt cccagaacaa gggaaagatt gctgaattgc tcaggtacca ctccaccaag  1320
agtggtgatg agctgacaag cctcaaggat tatgtgacca ggatgaagga aggtcagagt  1380
gacatctact acatcaccgg tgaaagcaag aaggctgttg agaactcccc attcttggag  1440
aagttgaaga agaaagggta tgaggttctt ttcatggttg atgccattga tgaatatgct  1500
gttggccagc tcaaggaatt tgagggcaag aagcttgtct ctgccaccaa ggagggtctc  1560
aaacttgacg agagtgaaga tgagaagaag aagcaagagg agttgaagga gaagtttgac  1620
aacctatgca aggtgatcaa ggatgttttg ggtgacaagg ttgagaaggt tgtagtctct  1680
gaccgtgttg ttgattcacc atgctgtctg gtgactggag aatacggttg gactgcaaac  1740
atggaaagga tcatgaaagc tcaagcccta agggacaaca gcatggcagg atacatgtca  1800
agcaagaaga caatggagat caaccctgag aacccaatca tggaggagct caggaagcgt  1860
gcggatgctg acaagaacga caaatccgtg aaggaccttg ttctgttgct gttcgagact  1920
gctcttctca cttctggttt cagccttgac gaacccaaca cttttggaaa caggatccac  1980
agaatgttga agcttggact gagcatcgac gaggatgcag ctgaagctga tgctgatgct  2040
gatatgcccc cactggagga ggaggctgaa gctgatgctg agggcagcaa aatggaggag  2100
gttgattaa                                                          2109

SEQ ID NO: 122          moltype = AA  length = 702
FEATURE                 Location/Qualifiers
source                  1..702
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 122
MAETETFAFQ AEINQLLSLI INTFYSNKEI FLRELISNAS DALDKIRFES LTDKSKLDAQ  60
PELFIHIIPD KTNNTLTIVD SGIGMTKADL VNNLGTIARS GTKEFMEALA AGADVSMIGQ  120
FGVGFYSAYL VAEKVIVTTK HNDDEQYVWE SQAGGSFTVT RDNTGEVLGR GTKITLYLKE  180
DQLEYLEERR LKDLIKKHSE FISYPISLWI EKTTEKEISD DEDEEEKKDE EGKVEDVDEE  240
KEKEEKKKKK IKEVSHEWSL VNKQKPIWMR KPEEITKEEY SAFYKSLTND WEEHLAVKHF  300
SVEGQLEFKA ILFVPKRAPF DLFDTRKKPN NIKLYVRRVF IMDNCEELIP EYLGFVKGIV  360
DSEDLPLNIS REMLQQNKIL KVIRKNLVKK CLELFFEIAE NKEDYNKFYE AFSKNLKLGI  420
HEDSQNKGKI AELLRYHSTK SGDELTSLKD YVTRMKEGQS DIYYITGESK KAVENSPFLE  480
KLKKKGYEVL FMVDAIDEYA VGQLKEFEGK KLVSATKEGL KLDESEDEKK KQEELKEKFD  540
NLCKVIKDVL GDKVEKVVVS DRVVDSPCCL VTGEYGWTAN MERIMKAQAL RDNSMAGYMS  600
SKKTMEINPE NPIMEELRKR ADADKNDKSV KDLVLLLFET ALLTSGFSLD EPNTFGNRIH  660
RMLKLGLSID EDAAEADADA DMPPLEEEAE ADAEGSKMEE VD                     702

SEQ ID NO: 123          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 123
ADLVNNLGTI ARSGTK                                                  16

SEQ ID NO: 124          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 124
AVENSPFLEK                                                         10

SEQ ID NO: 125          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 125
EDQLEYLEER RLK                                                     13

SEQ ID NO: 126          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 126
EEYSAFYK                                                           8

SEQ ID NO: 127          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 127
EGQSDIYYIT GESK                                                    14
```

-continued

```
SEQ ID NO: 128          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 128
EIFLRELISN ASDALDK                                                  17

SEQ ID NO: 129          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 129
FYEAFSK                                                             7

SEQ ID NO: 130          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 130
HFSVEGQLEF K                                                        11

SEQ ID NO: 131          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 131
IAELLRYHST K                                                        11

SEQ ID NO: 132          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 132
IRFESLTDK                                                           9

SEQ ID NO: 133          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 133
LDESEDEK                                                            8

SEQ ID NO: 134          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 134
LGIHEDSQNK                                                          10

SEQ ID NO: 135          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 135
RAPFDLFDTR K                                                        11

SEQ ID NO: 136          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 136
SGDELTSLK                                                           9

SEQ ID NO: 137          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
```

-continued

```
                        organism = Glycine max
SEQUENCE: 137
VEDVDEEK                                                            8

SEQ ID NO: 138          moltype = DNA  length = 2109
FEATURE                 Location/Qualifiers
source                  1..2109
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 138
atggcttcgg agactgagac gtttgctttt caggccgaga tcaaccagct cctcagtcta  60
atcatcaaca ccttctacag caacaaggag atctttcttc gtgagctcat cagcaatgcc  120
tctgatgctt tggacaagat tcgatttgag agtcttacag acaagagcaa gcttgatgct  180
cagccagagt tgttcattca cattattcct gacaagacca acaacacatt gtctatcatt  240
gacagtggta ttggcatgac taaggccgat ttggtgaata acctcggtac tattgcaagg  300
tctgggacca aggaattcat ggaagccctg gctgctggtg ctgatgtgag catgattgga  360
cagtttggtg ttggtttcta ctcggcttat cttgtagctg acaaggttat tgttaccaca  420
aagcacaatg atgatgagca gtatgtctgg gagtcccatg ctggtggctc atttactgtg  480
acaagggata catctggtga gaaccttggt aggggaacaa agataacttt gtttctcaag  540
gaagatcaac ttgagtatct tgaggaacgc cgtctgaagg atttgataaa gaagcattct  600
gaattcatta gctatccaat ttctctttgg attgagaaga ccactgaaaa ggagatttct  660
gatgatgagg acgaggaaga gaagaaggat gaggagggta aagttgaaga tgtggatgaa  720
gataaggaga aggaggagaa gaaaaagaag accattaagg aggtatctca tgagtggtca  780
ttggtgaaca agcagaagcc tatttggatg aggaaacctg aagagatcac aaaagaagaa  840
tatgctgctt tctacaagag tcttaccaat gattgggaag agcatttggc tgttaagcac  900
ttttctgttg agggtcagct ggagtttaag gctgtcctct ttatccccaa gagggcacct  960
tttgatcttt ttgacactaa gaagaagcct aacaacatca aactgtatgt ccgccgtgtt  1020
ttcatcatgg acaactgtga ggagttgatg ccagaatatc ttagctttgt taagggtatt  1080
gtggattctg aggatcttcc tctcaacatt tctagagaaa tgctgcagca gaacaagatc  1140
ctgaaagtca tcaggaagaa cttggtcaag aagtgcattg agatgttctt tgaaattgct  1200
gagaacaagg aagactataa caagttttat gaagccttct ctaagaatct gaaacttggt  1260
attcatgagg attctcagaa caagacaaaa cttgctgaat tgctcaggta tcactccacc  1320
aagagtggtg atgagatgac cagcctcaag gactatgtta ccaggatgaa ggaagggcag  1380
aatgacatct actacattac tggtgaaagc aagaaagctg ttgagaattc ccctttcctt  1440
gaaaagctta agaagaaggg gttcgaggtt ctctacatgg ttgatgctat tgacgagtat  1500
gctgttggtc agcttaagga atttgaggga aagaagttgg tctctgctac taaggaaggc  1560
ctgaaacttg acgagagcga agacgaaaag aaaaagaagg aagaattgaa ggagaaattt  1620
gagggtctct gccatgtgat caaggatgtg ttgggtgaca aggtggagaa agttgtggtg  1680
tctgatcgtg ttgtggattc tccttgctgt ctcgtgactg gtgaatatgg gtggacagcc  1740
aacatggaaa ggattatgaa agcacaggca ttgagggaca gcagcatggc tggatacatg  1800
tcgagcaaga agaccatgga gattaaccct gagaacccca tcatggatga actgaggaag  1860
cgagctgatg ctgacaagaa tgacaaatct gtgaaggatc tcgttctcct actctttgag  1920
actgctcttc taacttctgg gttcagcctt gatgatccca acactttcgg taacaggatt  1980
cacaggatgc tgaagcttgg actgagcatt gatgaagatg ctggtgaagc agatgctgac  2040
atgcctcctc tcgaggacgc cgatgctgat gccgatgctg agggtagcaa gatggaagaa  2100
gttgattaa                                                           2109

SEQ ID NO: 139          moltype = AA  length = 702
FEATURE                 Location/Qualifiers
source                  1..702
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 139
MASETETFAF QAEINQLLSL IINTFYSNKE IFLRELISNA SDALDKIRFE SLTDKSKLDA  60
QPELFIHIIP DKTNNTLSII DSGIGMTKAD LVNNLGTIAR SGTKEFMEAL AAGADVSMIG  120
QFGVGFYSAY LVADKVIVTT KHNDDEQYVW ESHAGGSFTV TRDTSGENLG RGTKITLFLK  180
EDQLEYLEER RLKDLIKKHS EFISYPISLW IEKTTEKEIS DDEDEEEKKD EEGKVEDVDE  240
DKEKEEKKKK TIKEVSHEWS LVNKQKPIWM RKPEEITKEE YAAFYKSLTN DWEEHLAVKH  300
FSVEGQLEFK AVLFIPKRAP FDLFDTKKKP NNIKLYVRRV FIMDNCEELM PEYLSFVKGI  360
VDSEDLPLNI SREMLQQNKI LKVIRKNLVK KCIEMFFEIA ENKEDYNKFY EAFSKNLKLG  420
IHEDSQNKTK LAELLRYHST KSGDEMTSLK DYVTRMKEGQ NDIYYITGES KKAVENSPFL  480
EKLKKKGFEV LYMVDAIDEY AVGQLKEFEG KKLVSATKEG LKLDESEDEK KKKEELKEKF  540
EGLCHVIKDV LGDKVEKVVV SDRVVDSPCC LVTGEYGWTA NMERIMKAQA LRDSSMAGYM  600
SSKKTMEINP ENPIMDELRK RADADKNDKS VKDLVLLLFE TALLTSGFSL DDPNTFGNRI  660
HRMLKLGLSI DEDAGEADAD MPPLEDADAD ADAEGSKMEE VD                     702

SEQ ID NO: 140          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 140
ADLVNNLGTI ARSGTK                                                   16

SEQ ID NO: 141          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
```

-continued

```
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 141
AVENSPFLEK                                                          10

SEQ ID NO: 142          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 142
EDQLEYLEER RLK                                                      13

SEQ ID NO: 143          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 143
EEYAAFYK                                                            8

SEQ ID NO: 144          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 144
EGQNDIYYIT GESK                                                     14

SEQ ID NO: 145          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 145
EIFLRELISN ASDALDK                                                  17

SEQ ID NO: 146          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 146
FYEAFSK                                                             7

SEQ ID NO: 147          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 147
HFSVEGQLEF K                                                        11

SEQ ID NO: 148          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 148
IRFESLTDK                                                           9

SEQ ID NO: 149          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 149
LAELLRYHST K                                                        11

SEQ ID NO: 150          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 150
```

-continued

```
LDESEDEK                                                            8

SEQ ID NO: 151          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 151
LGIHEDSQNK                                                          10

SEQ ID NO: 152          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 152
RAPFDLFDTK                                                          10

SEQ ID NO: 153          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 153
VEDVDEDK                                                            8

SEQ ID NO: 154          moltype = DNA  length = 2109
FEATURE                 Location/Qualifiers
source                  1..2109
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 154
atggcttcgg agactgagac gtttgctttt caggccgaga tcaaccagct cctcagtcta  60
atcatcaaca ccttctacag caacaaggag atctttcttc gtgagctcat cagcaatgcc  120
tctgatgctt tggacaagat tcgatttgag agtcttacag acaagagcaa gcttgatgct  180
cagccagagt tgttcattca cattattcct gacaagacca acaacacatt gtctatcatt  240
gacagtggta ttggcatgac taaggccgat ttggtgaata acctcggtac tattgcaagg  300
tctgggacca aggaattcat ggaagccctg gctgctggtg ctgatgtgag catgattgga  360
cagtttggtg ttggtttcta ctcggcttat cttgtagctg acaaggttat tgttaccaca  420
aagcacaatg atgatgagca gtatgtctgg gagtcccatg ctggtggctc atttactgtg  480
acaagggata catctggtga gaaccttggt aggggaacaa agataacttt gtttctcaag  540
gaagatcaac ttgagtatct tgaggaacgc cgtctgaagg atttgataaa gaagcattct  600
gaattcatta gctatccaat ttctctttgg attgagaaga ccactgaaaa ggagatttct  660
gatgatgagg acgaggaaga gaagaaggat gaggagggta aagttgaaga tgtggatgaa  720
gataaggaga aggaggagaa gaaaaagaag accattaagg aggtatctca tgagtggtca  780
ttggtgaaca agcagaagcc tatttggatg aggaaacctg aagagatcac aaaagaagaa  840
tatgctgctt tctacaagag tcttaccaat gattgggaag agcatttggc tgttaagcac  900
ttttctgttg agggtcagct ggagtttaag gctgtcctct ttatccccaa gagggcacct  960
tttgatcttt ttgacactaa gaagaagcct aacaacatca aactgtatgt ccgccgtgtt  1020
ttcatcatgg acaactgtga ggagttgatg ccagaatatc ttagctttgt taagggtatt  1080
gtggattctg aggatcttcc tctcaacatt tctagagaaa tgctgcagca gaacaagatc  1140
ctgaaagtca tcaggaagaa cttggtcaag aagtgcattg agatgttctt tgaaattgct  1200
gagaacaagg aagactataa caagttttat gaagccttct ctaagaatct gaaacttggt  1260
attcatgagg attctcagaa caagacaaaa cttgctgaat tgctcaggta tcactccacc  1320
aagagtggtg atgagatgac cagcctcaag gactatgtta ccaggatgaa ggaagggcag  1380
aatgacatct actacattac tggtgaaagc aagaaagctg ttgagaattc ccctttcctt  1440
gaaaagctta agaagaaggg gttcgaggtt ctctacatgg ttgatgctat tgacgagtat  1500
gctgttggtc agcttaagga atttgaggga aagaagttgg tctctgctac taaggaaggc  1560
ctgaaacttg acgagagcga agacgaaaag aaaaagaagg aagaattgaa ggagaaattt  1620
gagggtctct gccatgtgat caaggatgtg ttgggtgaca aggtggagaa agttgtggtg  1680
tctgatcgtg ttgtggattc tccttgctgt ctcgtgactg gtgaatatgg gtggacagcc  1740
aacatggaaa ggattatgaa agcacaggca ttgagggaca gcagcatggc tggatacatg  1800
tcgagcaaga agaccatgga gattaaccct gagaacccca tcatggatga actgaggaag  1860
cgagctgatg ctgacaagaa tgacaaatct gtgaaggatc tcgttctcct actctttgag  1920
actgctcttc taacttctgg gttcagcctt gatgatccca acactttcgg taacaggatt  1980
cacaggatgc tgaagcttgg actgagcatt gatgaagatg ctggtgaagc agatgctgac  2040
atgcctcctc tcgaggacgc cgatgctgat gccgatgctg agggtagcaa gatggaagaa  2100
gttgattaa                                                           2109

SEQ ID NO: 155          moltype = AA  length = 839
FEATURE                 Location/Qualifiers
source                  1..839
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 155
MKLSPREVEK LGLHNAGYLA QKRLARGLRL NYTEAVALIA TQIMEFARDG EKTVAQLMCI  60
```

-continued

```
GKHLLGRRQV LPEVQHLLNA VQVEATFPDG TKLVTVHDPI SCEHGDLGQA LFGSFLPVPS  120
LDKFAENKED NRIPGEIIYG DGSLVLNPGK NAVILKVVSN GDRPIQVGSH YHFIEVNPYL  180
TFDRRKAYGM RLNIAAGTAV RFEPGDSKSV KLVRIGGNKV IRGGNGIADG QVNETNLREA  240
MEAVCKRGFG HKEEEDASEG ITGDPDSPFT TIIPREEYAN KYGPTTGDKI RLGDTDLFAK  300
IEKDFALYGD ECVFGGGKVL RDGMGQSCGD PPAISLDTVI TNAVIIDYSG IIKADIGIKD  360
GLIVSIGKAG NPDIMDDVFF NMIIGANTEV IAGEGLIVTA GAIDCHVHYI CPQLVDEAIS  420
SGITTLVGGG TGPTAGTRAT TCTPAPSQMK LMLQSTDDLP LNFGFTGKGS SSKPDELHDI  480
IKAGAMGLKL HEDWGSTPAA IDSCLTVADQ YDIQINIHTD TLNEAGFVEH SIAAFKGRTI  540
HTYHSEGAGG GHAPDIIKVC GMKNVLPSST NPTRPLTLNT IDEHLDMLMV CHHLNREIPE  600
DLAFACSRIR EGTIAAEDIL HDIGAISIIS SDSQAMGRVG EVISRTWQTA NKMKVQRGPL  660
QPGESDNDNF RIKRYIAKYT INPAIANGFS QYVGSVEVGK LADLVMWKPS FFGAKPEMVI  720
KGGVVAWADM GDPNASIPTP EPVKMRPMFG TLGKAGGALS IAFVSKAAVD QRVHALYGLN  780
KRVEAVGNVR KLTKLDMKLN DSLPQITVDP DNYTVTADGE VLTSFATTFV PLSRNYFLF   839

SEQ ID NO: 156          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 156
AGGALSIAFV SK                                                      12

SEQ ID NO: 157          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 157
DGLIVSIGK                                                          9

SEQ ID NO: 158          moltype = DNA  length = 2946
FEATURE                 Location/Qualifiers
source                  1..2946
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 158
atggcttctt cttccacgac aatgcggttc tccgcggcct ccaccggcgc cgcggcggcg  60
cttcctcgcc ggagctccgt cgccggattc atcggcgtcg cggccagatc ctccgccaaa  120
tcgaggctcc gcttcatcgg aaggaacgcg aacttgagcc tccggatgag gaggatgagc  180
tctttctccg tcgtgaagtg cgtctccgga agcgaggcca aggtgcagga taccgtggcc  240
aaacaacaag aagctacaac ctctctcagc tccttcacgc cagatgcttc atctattgcc  300
tcaagtatta aataccatgc agagttcact ccattgtttt ctccagagaa ttttgatcta  360
ccacaggctt ttcttgcaac tgcacaaagt gttcgagatt ctctcattat aaactggaat  420
gctacctatg attactatga gaaactgaat gttaagcagg catattacct ttcgatggaa  480
tttttgcagg gtagagcatt attaaatgca ataggcaatc tggagctaac tggtccctat  540
gctgaggctt tgagcaagct cgggcataaa ttagaaaatg tggcatacca ggagccagat  600
gctgcacttg gaaatggtgg tcttgggcgt cttgcttcat gtttcttgga ctctttggca  660
accttgaact atccagcttg gggttatgga ttgagatata agtatggttt atttaaacag  720
cgaataacca aagatgggca ggaggaagtt gctgaagatt ggcttgagat gggcaatccc  780
tgggagatta ttagaaatga tgtttcatat cctgtaaagt tctatggcaa ggttgtttca  840
gggtcagatg gtaaaaaaca ttggattgga ggagaagata taaaagctgt tgcacatgat  900
gttcccatac caggatataa aactaagact acaatcaact tgagactttg gtctacaaaa  960
gctgcatcag aagaatttga tttatctgct tttaatgctg gaaggcacac tgaagcatct  1020
gaggctctag caaatgctga aaagatatgc tatatactct accctgggga tgaaccaata  1080
gagggcaaga tccttcgcct caagcaacaa tatactttat gttctgcttc tcttcaagac  1140
atcattgctc gttttgagag aagatcagga gcaaatgtta attgggaaga gtttcccgag  1200
aaggttgcag tgcagatgaa tgatactcac ccaacgttgt gcatcccaga gctcatgaga  1260
attttaatag atgtaaaggg cttaaattgg aaggatgctt ggaatattac acaaaggact  1320
gtagcatata caaaccacac tgttcttcca gaggcattag agaaatggag cttggatctt  1380
atgcagaaac tgcttccacg ccatattgag ataatagaaa tgattgacga ggagctggtt  1440
cggaccataa ttgcagaata tggcacagaa aattcagact tgctggagaa aaaattgaag  1500
gaaatgagaa tactagaaaa tgtcgaattg actgcagaat ttgcagatat actggttaag  1560
tccaaagagg ccattgatat tcctagtgag gagctgcaaa gttctgaaca agcagaagca  1620
gaagacgaaa aagatgatga tgaagtggaa gcagttgcta aaaagaatgg aacagatgaa  1680
agttccattg aggatgagaa agaagagttg cctgagccag taccagaacc accaaaattg  1740
gttcgtatgg caaatttgtg tgttgtgggt ggtcatgcag tgaatggggt tgcagagata  1800
catagtgaaa ttgtaaaaga tgatgttttc aatgcatttt ataagttgtg gcccgagaag  1860
ttccaaaaca aaacaaatgg tgtgacacca aggagatgga tcaggttctg caatccagat  1920
ctgagtaaaa taataaccga gtggataggt actgaggatt gggtcctaaa tactggaaaa  1980
ttggccgagc tacggaagtt tgtagacaat gaggatctac aagtgcagtg gagggaagca  2040
aagaggagta acaaagttaa agttgcagca ttcatcagag aaaaaacagg gtattctgtc  2100
agtcctgatg caatgtttga catccaggtg aagagaatcc atgaatataa gcgacaatta  2160
atgaacatat ttggaattgt ttaccgatac aagaagatga aagaaatgag tgcggcagaa  2220
agggaagcaa attttgttcc tagagtttgt atatttgggg gaaaagcatt tgctacatat  2280
gtgcaagcca agagaattgt gaaatttatt acagatgttg gggctacagt aaatcatgac  2340
cctgaaatag gagatctact taaggttgtt tttgtccctg attataatgt cagtgttgcc  2400
```

-continued

```
gagatgctta ttccagcaag tgaattgtca caacatatca gtactgctgg aatggaggca  2460
agtggaacta gcaatatgaa atttgcaatg aatggctgca tattgattgg aactttagat  2520
ggggccaatg ttgaaataag ggaagaggtt ggtgcagaca acttcttcct ctttggtgcc  2580
aaggctcatg aaattgctgg gctgaggaaa gaaagagcag agggaaagtt tgtgcctgat  2640
ccacgatttg aagaagttaa agaatttgtc agaagtggca tttttggatc ttacaactat  2700
gacgaattga tgggatcctt agaaggaaat gaaggttttg gtcgagcaga ttatttcctt  2760
gtagggaaag actttcccag ctacatagaa tgccaagaga aggttgatga agcataccgt  2820
gaccaaacta aatggacaag aatgtcaata ttgaacacag caggttcata caagtttagc  2880
agtgaccgca cgattcatga atatgcccga gagatatgga acattgaacc agtccaattg  2940
ccatag                                                             2946

SEQ ID NO: 159         moltype = AA  length = 981
FEATURE                Location/Qualifiers
source                 1..981
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 159
MASSSTTMRF SAASTGAAAA LPRRSSVAGF IGVAARSSAK SRLRFIGRNA NLSLRMRRMS  60
SFSVVKCVSG SEAKVQDTVA KQQEATTSLS SFTPDASSIA SSIKYHAEFT PLFSPENFDL  120
PQAFLATAQS VRDSLIINWN ATYDYYEKLN VKQAYYLSME FLQGRALLNA IGNLELTGPY  180
AEALSKLGHK LENVAYQEPD AALGNGGLGR LASCFLDSLA TLNYPAWGYG LRYKYGLFKQ  240
RITKDGQEEV AEDWLEMGNP WEIIRNDVSY PVKFYGKVVS GSDGKKHWIG GEDIKAVAHD  300
VPIPGYKTKT TINLRLWSTK AASEEFDLSA FNAGRHTEAS EALANAEKIC YILYPGDEPI  360
EGKILRLKQQ YTLCSASLQD IIARFERRSG ANVNWEEFPE KVAVQMNDTH PTLCIPELMR  420
ILIDVKGLNW KDAWNITQRT VAYTNHTVLP EALEKWSLDL MQKLLPRHIE IIEMIDEELV  480
RTIIAEYGTE NSDLLEKKLK EMRILENVEL TAEFADILVK SKEAIDIPSE ELQSSEQAEA  540
EDEKDDDEVE AVAKKNGTDE SSIEDEKEEL PEPVPEPPKL VRMANLCVVG GHAVNGVAEI  600
HSEIVKDDVF NAFYKLWPEK FQNKTNGVTP RRWIRFCNPD LSKIITEWIG TEDWVLNTGK  660
LAELRKFVDN EDLQVQWREA KRSNKVKVAA FIREKTGYSV SPDAMFDIQV KRIHEYKRQL  720
MNIFGIVYRY KKMKEMSAAE REANFVPRVC IFGGKAFATY VQAKRIVKFI TDVGATVNHD  780
PEIGDLLKVV FVPDYNVSVA EMLIPASELS QHISTAGMEA SGTSNMKFAM NGCILIGTLD  840
GANVEIREEV GADNFFLFGA KAHEIAGLRK ERAEGKFVPD PRFEEVKEFV RSGIFGSYNY  900
DELMGSLEGN EGFGRADYFL VGKDFPSYIE CQEKVDEAYR DQTKWTRMSI LNTAGSYKFS  960
SDRTIHEYAR EIWNIEPVQL P                                            981

SEQ ID NO: 160         moltype = AA  length = 9
FEATURE                Location/Qualifiers
source                 1..9
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 160
AFATYVQAK                                                          9

SEQ ID NO: 161         moltype = AA  length = 23
FEATURE                Location/Qualifiers
source                 1..23
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 161
QQEATTSLSS FTPDASSIAS SIK                                          23

SEQ ID NO: 162         moltype = DNA  length = 2706
FEATURE                Location/Qualifiers
source                 1..2706
                       mol_type = other DNA
                       organism = Glycine max
SEQUENCE: 162
atggcgactg agaatccttt caatagcatt ttgacgacgc tcgaaaagcc cggtggtgct  60
ggcgagtttg gaaagtactt tagcttgcct gctctcaacg atcgtagaat cgatagactt  120
ccttactctg tgaggatact tttggaatct gcaatccgga actgtgatga gtttcaagtt  180
aagagtaacg acgttgaaaa aattattgat tgggagaata cctctccgaa gctggtggag  240
attccattca agccagctag agtgcttctc caggatttta ctggcgtgcc tgctgttgtc  300
gatcttgctt gcatgcgaga tgccatgaac aaacttggtg gtgattctaa taaaattaat  360
cccttggtac cagttgatct tgtcattgat cactctgttc aagttgatgt ggcaagatca  420
gaaaatgctg tccaggcaaa tatggaactt gaattccaaa gaaacaagga aagatttggt  480
ttcctcaaat ggggttctaa tgcattcaat aatatgcttg tcgttcctcc tggatcaggg  540
atagttcacc aggttaattt agaatacctt ggtagagttg tgttcaacac aaatggtgtg  600
ctttatcctg acagcgttgt tggaactgat tcacacacaa ctatgattga tggcctgggt  660
gttgctggat ggggagttgg tggaatagaa gcagaagctg caatgcttgg ccaacccatg  720
agcatggtcc taccgggtgt cgttgggttc aaattattag gcaaactacg agatggtgtc  780
acagctactg acttggtgtt gactgttaca caaatgctaa gaaagcatgg ggttgttggc  840
aagtttgtag agttttatgg ggaaggcatg agtgaactgt ctttggcaga tcgtgccacc  900
atagcaaaca tgtctcctga gtatggtgca accatgggct tctttcctgt ggatcatgtc  960
actttgcaat atttgagact gactggcaga agtgatgaga ctgtttctat gatagaatcg  1020
tacttacgag caaataagat gtttgtggat tacagcgagc cccaagtgga gagagtgtac  1080
tcatcgtatt tggaactcaa tcttgaggat gtagaaccct gtgtctcagg tccaaaaagg  1140
```

-continued

```
cctcatgacc gtgtcccttt gagagaaatg aaggtagact ggcatgcctg cctcaacaat  1200
aaagttggat tcaagggttt tgctgtacca aaagaatctc agaataaggt tgcagagttc  1260
acattccaag ggacaccagc acatcttagg cacggtgatg ttgttatagc tgctatcacc  1320
agttgtacaa atacttcaaa tcctagcgta atgcttggag ctgcattggt tgcaaagaaa  1380
gcatgtgaat tgggattgca ggtgaagcca tggattaaaa caagccttgc cccaggttct  1440
ggtgttgtga ctaagtattt gcagaggagt ggcttgcaga agtatttgaa tgagctaggg  1500
tttaatatag ttgggtatgg atgcaccaca tgcattggaa attcagggga tattaatgaa  1560
gctgttgcat ctgcaattac tgaaaatgat atagtagctg cagctgtact gtctggaaat  1620
aggaactttg agggccgagt tcacccatta acaagagcta attatcttgc ttctcctcct  1680
cttgttgttg cctatgccct tgctggcaca gtggacattg actttgacac tgaacccatt  1740
gggataggga aggatggcac aaagatcttc ttcagggata tttggccatc cagtgaagaa  1800
atagcaaatg ttgtacaatc aagtgtgctg cctgctatgt ttagggacac atataatgcg  1860
atcacccaag ggaatcccat gtggaataat ttatctgttc caactggcac tctttatgct  1920
tgggacccta catcaactta tattcatgag ccaccttatt ttagagatat gagcatgtct  1980
cccccaggct ctcatggagt gaaagatgct tactgtttgc tcaactttgg agacagtatc  2040
acaactgacc acatctctcc agctggcagc atacataagg acagtcctgc agccagatac  2100
cttatagaac gtggggttga tagacgggac ttcaactcct atggaagtcg ccgtggtaat  2160
gatgaggtga tggcccgagg aacttttgcc aatattcgca ttgtcaacaa atttttgaat  2220
ggagaagttg gacctaaaac tattcatatt ccttctgggg agaagttatc agtctttgat  2280
gctgcagaga aatacaagag tgaggggcat gatatgatta tcttggctgg tgctgagtat  2340
gggagtggaa gctctcgtga ttgggctgca aaggggccaa tgctactggg cgtgaaagct  2400
gtcatagcaa aaagttttga aaggattcac cgaagtaatt tggtgggaat gggtatcatt  2460
ccattgtgtt ttaagcccgg ggatgatgct gattctcttg gattaactgg ccatgaacgt  2520
tacaccattg atctaccaag caatgtgaat gaaatacgac ccggccaaga tgtcacagtg  2580
gtgacagatg ccgggaaatc atttgtatct acactgagat ttgatacaga ggttgaacta  2640
gcttacttca accatggagg catcttacaa tatgtcatca gaaacatggt caatgctaaa  2700
cattga                                                             2706

SEQ ID NO: 163          moltype = AA  length = 901
FEATURE                 Location/Qualifiers
source                  1..901
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 163
MATENPFNSI LTTLEKPGGA GEFGKYFSLP ALNDRRIDRL PYSVRILLES AIRNCDEFQV  60
KSNDVEKIID WENTSPKLVE IPFKPARVLL QDFTGVPAVV DLACMRDAMN KLGGDSNKIN  120
PLVPVDLVID HSVQVDVARS ENAVQANMEL EFQRNKERFG FLKWGSNAFN NMLVVPPGSG  180
IVHQVNLEYL GRVVFNTNGV LYPDSVVGTD SHTTMIDGLG VAGWGVGGIE AEAAMLGQPM  240
SMVLPGVVGF KLLGKLRDGV TATDLVLTVT QMLRKHGVVG KFVEFYGEGM SELSLADRAT  300
IANMSPEYGA TMGFFPVDHV TLQYLRLTGR SDETVSMIES YLRANKMFVD YSEPQVERVY  360
SSYLELNLED VEPCVSGPKR PHDRVPLREM KVDWHACLNN KVGFKGFAVP KESQNKVAEF  420
TFQGTPAHLR HGDVVIAAIT SCTNTSNPSV MLGAALVAKK ACELGLQVKP WIKTSLAPGS  480
GVVTKYLQRS GLQKYLNELG FNIVGYGCTT CIGNSGDINE AVASAITEND IVAAAVLSGN  540
RNFEGRVHPL TRANYLASPP LVVAYALAGT VDIDFDTEPI GIGKDGTKIF FRDIWPSSEE  600
IANVVQSSVL PAMFRDTYNA ITQGNPMWNN LSVPTGTLYA WDPTSTYIHE PPYFRDMSMS  660
PPGSHGVKDA YCLLNFGDSI TTDHISPAGS IHKDSPAARY LIERGVDRRD FNSYGSRRGN  720
DEVMARGTFA NIRIVNKFLN GEVGPKTIHI PSGEKLSVFD AAEKYKSEGH DMIILAGAEY  780
GSGSSRDWAA KGPMLLGVKA VIAKSFERIH RSNLVGMGII PLCFKPGDDA DSLGLTGHER  840
YTIDLPSNVN EIRPGQDVTV VTDAGKSFVS TLRFDTEVEL AYFNHGGILQ YVIRNMVNAK  900
H                                                                  901

SEQ ID NO: 164          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 164
FLNGEVGPK                                                          9

SEQ ID NO: 165          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 165
LSVFDAAEK                                                          9

SEQ ID NO: 166          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 166
LSVFDVAEK                                                          9

SEQ ID NO: 167          moltype = AA  length = 12
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 167
TSLAPGSGVV TK                                                        12

SEQ ID NO: 168          moltype = DNA  length = 2706
FEATURE                 Location/Qualifiers
source                  1..2706
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 168
atggcgactg agaacccttt caatagcatt ttgaggacgc tcgagaagcc cggcggcgct  60
ggcgagtttg gaaagtactt tagcttgccc gctctcaacg atcctagaat cgatagactt  120
ccttactctg tgaggatact tttggaatct gcaatccgta actgtgatga gtttcaagtt  180
aagagtaacg acattgaaaa aattattgat tgggagaata cctctcccaa gctggtggag  240
attccattca agccagctag agtgcttctg caggacttta ctggggtgcc tgctgtggtt  300
gatcttgctt gcatgcgaga tgcaatgaac aaacttggtg gtgattctaa taaaattaac  360
cccttggtac cagttgatct tgtcattgat cactctgttc aagttgatgt ggcaagatca  420
gaaaatgctg tccaggcaaa tatggaactt gaattccaaa gaaacaagga aagatttggt  480
ttcctcaaat ggggttctaa tgcattcaat aatatgcttg tcgttcctcc tggatcaggg  540
atagttcatc aggttaattt agaatacctt ggtagagttg tgttcaacac aaatggtgtg  600
ctttatccag acagcgttgt cggaactgat tcacacacta ctatgataga tggcctgggt  660
gttgctggat ggggagttgg tggaatagaa gcagaagctg caatgcttgg ccagcccatg  720
agcatggtcc taccgggtgt tgttgggttc aaattattag gcaaactacg agatggtgtc  780
acagctactg acttggtgtt gactgttact caaatgctaa gaaagcatgg ggttgttggc  840
aagtttgtag agttttatgg ggaaggcatg agtgaactgt ctttggcaga tcgtgccact  900
atagcaaaca tgtctcctga gtatggtgca accatgggct tctttcctgt ggatcatgtc  960
actttgcaat atttgagact gactggcaga agtgatgaga ctgtttctat gatagaatcc  1020
tatttacgag caaataagat gtttgtggat tatagcgagc cccaagtgga gagagtgtac  1080
tcatcctatt tggaactcaa tcttgaggat gtagagccct gtgtctcagg tccaaaaagg  1140
cctcacgacc gtgtcccttt gagagaaatg aaggtagact ggcatgcctg cctcaacaat  1200
aaagttggat ttaagggttt tgctgtatca aaagaatctc agaataaggt tgcagagttc  1260
acattccaag ggacaccagc acatcttagg catggtgatg ttgttatagc tgctatcacc  1320
agttgtacaa atacttcaaa tcctagcgta atgcttggag ctgcattggt tgcaaagaaa  1380
gcatgtgaat tgggattgca ggtgaagcca tggattaaaa caagtcttgc cccaggttct  1440
ggtgttgtaa ctaagtattt gcagaggagt ggcttgcaga agtatttgaa tgagctaggg  1500
tttaatatag ttgggtatgg atgcaccaca tgcattggaa attcagggga tattaatgaa  1560
gctgttgcat ctgcgattac cgaaaatgat atagtagctg cagctgtatt gtctggaaat  1620
aggaactttg agggccgagt tcacccatta acaagagcta attatcttgc ttctcctcct  1680
cttgttgttg cctatgccct tgctggcaca gtggacattg actttgacac tgaacccatt  1740
gggataggga aggatggaac agagatcttc ttcaaggata tttggccatc cagtgaagaa  1800
atagcaaatg ttgtacaatc aagtgtgctg cctgatatgt ttagggacac atataatgca  1860
atcacccaag ggaatcccat gtggaataat ttatctgttc caacgggcac tctttatgct  1920
tgggacccta catcgactta tattcatgag ccaccttatt ttagagatat gagcatgtct  1980
cccccaggct ctcatggagt gaaggatgct tactgtttgc tcaactttgg agacagtatc  2040
acaactgacc acatctctcc agctggcagc atacataagg acagtcctgc ggccagatac  2100
cttatagaac gtggggttga tagacgggac ttcaactcct atggaagtcg ccgtggtaat  2160
gatgaggtga tggcccgagg aacgtttgcc aatattcgca ttgtcaacaa atttttgaat  2220
ggagaagttg gacctaaaac tattcatatt ccttctgggg agaagttgtc agtctttgac  2280
gttgcagaga aatacaagag tgaagggcat gatatgatta tcttggctgg tgctgagtat  2340
gggagtggaa gctctcgtga ttgggctgca aaggggccaa tgctactggg cgtgaaagct  2400
gtcatagcaa aaagttttga aaggattcac cgaagtaatt tggtgggaat gggtatcatt  2460
cctttgtgtt ttaagcccgg ggaggatgct gattctcttg gattaactgg tcaagaacgt  2520
tacaccattg atctaccaag caatgtgaat gaaatacgac ccggtcaaga tgtcacagtg  2580
gtgacagaca ccgggaaatc atttgtatct acattgagat ttgatactga ggttgaacta  2640
gcttacttca accatggagg tatcttacaa tatgtcatca gaaacttgat caatgctaaa  2700
cattga                                                              2706

SEQ ID NO: 169          moltype = AA  length = 901
FEATURE                 Location/Qualifiers
source                  1..901
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 169
MATENPFNSI LRTLEKPGGA GEFGKYFSLP ALNDPRIDRL PYSVRILLES AIRNCDEFQV  60
KSNDIEKIID WENTSPKLVE IPFKPARVLL QDFTGVPAVV DLACMRDAMN KLGGDSNKIN  120
PLVPVDLVID HSVQVDVARS ENAVQANMEL EFQRNKERFG FLKWGSNAFN NMLVVPPGSG  180
IVHQVNLEYL GRVVFNTNGV LYPDSVVGTD SHTTMIDGLG VAGWGVGGIE AEAAMLGQPM  240
SMVLPGVVGF KLLGKLRDGV TATDLVLTVT QMLRKHGVVG KFVEFYGEGM SELSLADRAT  300
IANMSPEYGA TMGFFPVDHV TLQYLRLTGR SDETVSMIES YLRANKMFVD YSEPQVERVY  360
SSYLELNLED VEPCVSGPKR PHDRVPLREM KVDWHACLNN KVGFKGFAVS KESQNKVAEF  420
TFQGTPAHLR HGDVVIAAIT SCTNTSNPSV MLGAALVAKK ACELGLQVKP WIKTSLAPGS  480
GVVTKYLQRS GLQKYLNELG FNIVGYGCTT CIGNSGDINE AVASAITEND IVAAAVLSGN  540
RNFEGRVHPL TRANYLASPP LVVAYALAGT VDIDFDTEPI GIGKDGTEIF FKDIWPSSEE  600
```

-continued

```
IANVVQSSVL PDMFRDTYNA ITQGNPMWNN LSVPTGTLYA WDPTSTYIHE PPYFRDMSMS  660
PPGSHGVKDA YCLLNFGDSI TTDHISPAGS IHKDSPAARY LIERGVDRRD FNSYGSRRGN  720
DEVMARGTFA NIRIVNKFLN GEVGPKTIHI PSGEKLSVFD VAEKYKSEGH DMIILAGAEY  780
GSGSSRDWAA KGPMLLGVKA VIAKSFERIH RSNLVGMGII PLCFKPGEDA DSLGLTGQER  840
YTIDLPSNVN EIRPGQDVTV VTDTGKSFVS TLRFDTEVEL AYFNHGGILQ YVIRNLINAK  900
H                                                                  901

SEQ ID NO: 170          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 170
FLNGEVGPK                                                          9

SEQ ID NO: 171          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 171
LSVFDVAEK                                                          9

SEQ ID NO: 172          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 172
TSLAPGSGVV TK                                                      12

SEQ ID NO: 173          moltype = DNA  length = 2292
FEATURE                 Location/Qualifiers
source                  1..2292
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 173
atggcatctc acatcgttgg ataccccgt atgggtccca agagagagct caagttcgct  60
ctcgagtctt tctgggatgg caagagcagc gccgaggatt tgcagaaggt gtcttctgat  120
ctcagggcat ccatctggaa gcagatggct gatgctggga tcaagtacat ccccagcaac  180
actttctctc actatgacca ggttctcgac gccaccgcca ccctcggtgc cgttccacca  240
aggtacggct ggaccggcgg cgagattggg tttgatacct acttctccat ggccagaggt  300
aatgctaccg tgccagctat ggagatgacc aagtggttcg acaccaacta ccactttatt  360
gtccctgaat tgggccctga tgtgaacttc acctatgctt ctcacaaggc tgttgatgaa  420
tacaaggagg ccaaggcgct tggagtggat accgttccgg tcctcgttgg ccctgttaca  480
tacctgttgc tctccaagcc tgccaaggga gttgagaaat ccttttctct cctctctctc  540
cttcccaagg ttcttgctgt ctacaaggaa gttattgctg accttaaggc agctggtgct  600
tcatggattc agtttgatga gcctaccctt gtcttggacc ttgagtctca caagttgcaa  660
gcattcactg acgcatatgc agaacttgcg cctgctttgt ctggtttgaa tgttcttgtt  720
gagacctact ttgctgacat ccctgctgag gcatacaaga ccctcacatc tctgaatggc  780
gtcactgcat atggatttga tttggtccgt ggaaccaaca ctcttgattt gatcaagggt  840
ggatttccca gcggaaaata cctctttgct ggagtggttg atggaaggaa catctgggcc  900
aatgaccttg ctgcttctct cactaccttg cagggtcttg agggcattgt gggcaaagat  960
aagcttgttg tgtccacctc ctcctccctt cttcacactg ctgttgatct agttaacgag  1020
accaagttgg atgatgagat caagtcatgg ctagcttttg ctgcccaaaa aattgttgaa  1080
gttaacgcat tggctaaagc attgtctggc cacaaggatg aggccttctt ctctggtaat  1140
gctgctgctc tggcttcaag gaagtcttct ccaagagtga ccaacgaggc tgttcagaag  1200
gctgctgctg cattgaaggg ttcagatcat cgccgtgcaa caaatgtcag tgccagactg  1260
gattctcaac aaaagaagct caaccttcca atcctgccaa ccaccactat tggatccttc  1320
cctcagactg tagaactgag gagggtacgc cgtgaattca aggctaacaa gatctccgag  1380
gaagagtatg taaagtcaat taaggaggaa attcgcaaag ttgttgagct tcaagaagag  1440
cttgatattg atgttcttgt tcatggagaa ccagagagaa atgatatggt tgagtacttc  1500
ggtgaacaat tgtcaggctt tgccttcacc gttaatgggt gggtgcaatc ctatggttcc  1560
cgttgcgtga agccaccgat catctatggt gatgtgagcc gcccaaagcc aatgaccgtc  1620
ttctggtcat ctctggctca gagctttacc aagcgcccaa tgaagggaat gcttaccggt  1680
cctgttacca ttctcaactg gtcctttgtt agaaatgacc aacctagatc tgagaccacc  1740
taccagattg ctttgtctat caaggacgaa gtggaagacc ttgaaaaggc tggcatcact  1800
gttatccaaa ttgatgaagc tgctttgaga gagggtcttc cactgaggaa atcagagcaa  1860
gctcactact tggactgggc tgtccatgcc ttcagaatca ccaatgttgg tgtccaggat  1920
accacccaga tccacactca catgtgctac tcgaatttca acgacatcat ccactccatc  1980
atcgacatgg acgccgatgt tatcaccatt gagaactctc gctccgacga gaagcttctg  2040
tcagtcttcc gcgaaggtgt gaagtatggt gctggaattg gccctggtgt ctatgacatc  2100
cactccccaa gaataccacc aactgaagaa attgctgaca gaatcaacaa gatgctggca  2160
gtgctcgaga agaacatctt gtgggtgaac cctgactgtg ggctcaagac ccgtaagtac  2220
actgaggtga agccagccct cacaaacatg gttgccgcag caaaactcat ccgcaacgaa  2280
cttgccaagt ga                                                      2292
```

-continued

```
SEQ ID NO: 174          moltype = AA  length = 763
FEATURE                 Location/Qualifiers
source                  1..763
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 174
MASHIVGYPR MGPKRELKFA LESFWDGKSS AEDLQKVSSD LRASIWKQMA DAGIKYIPSN  60
TFSHYDQVLD ATATLGAVPP RYGWTGGEIG FDTYFSMARG NATVPAMEMT KWFDTNYHFI  120
VPELGPDVNF TYASHKAVDE YKEAKALGVD TVPVLVGPVT YLLLSKPAKG VEKSFSLLSL  180
LPKVLAVYKE VIADLKAAGA SWIQFDEPTL VLDLESHKLQ AFTDAYAELA PALSGLNVLV  240
ETYFADIPAE AYKTLTSLNG VTAYGFDLVR GTNTLDLIKG GFPSGKYLFA GVVDGRNIWA  300
NDLAASLTTL QGLEGIVGKD KLVVSTSSSL LHTAVDLVNE TKLDDEIKSW LAFAAQKIVE  360
VNALAKALSG HKDEAFFSGN AAALASRKSS PRVTNEAVQK AAAALKGSDH RRATNVSARL  420
DSQQKKLNLP ILPTTTIGSF PQTVELRRVR REFKANKISE EEYVKSIKEE IRKVVELQEE  480
LDIDVLVHGE PERNDMVEYF GEQLSGFAFT VNGWVQSYGS RCVKPPIIYG DVSRPKPMTV  540
FWSSLAQSFT KRPMKGMLTG PVTILNWSFV RNDQPRSETT YQIALSIKDE VEDLEKAGIT  600
VIQIDEAALR EGLPLRKSEQ AHYLDWAVHA FRITNVGVQD TTQIHTHMCY SNFNDIIHSI  660
IDMDADVITI ENSRSDEKLL SVFREGVKYG AGIGPGVYDI HSPRIPPTEE IADRINKMLA  720
VLEKNILWVN PDCGLKTRKY TEVKPALTNM VAAAKLIRNE LAK                    763

SEQ ID NO: 175          moltype = AA  length = 21
FEATURE                 Location/Qualifiers
source                  1..21
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 175
AGITVIQIDE AALREGLPLR K                                            21

SEQ ID NO: 176          moltype = AA  length = 24
FEATURE                 Location/Qualifiers
source                  1..24
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 176
ALGVDTVPVL VGPVTYLLLS KPAK                                         24

SEQ ID NO: 177          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 177
DEAFFSGNAA ALASRK                                                  16

SEQ ID NO: 178          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 178
DEVEDLEK                                                           8

SEQ ID NO: 179          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 179
ISEEEYVK                                                           8

SEQ ID NO: 180          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 180
IVEVNALAK                                                          9

SEQ ID NO: 181          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 181
LIRNELAK                                                           8
```

-continued

```
SEQ ID NO: 182          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 182
LLSVFREGVK                                                          10

SEQ ID NO: 183          moltype = AA  length = 21
FEATURE                 Location/Qualifiers
source                  1..21
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 183
LVVSTSSSLL HTAVDLVNET K                                             21

SEQ ID NO: 184          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 184
SFSLLSLLPK                                                          10

SEQ ID NO: 185          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 185
SSPRVTNEAV QK                                                       12

SEQ ID NO: 186          moltype = AA  length = 26
FEATURE                 Location/Qualifiers
source                  1..26
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 186
TLTSLNGVTA YGFDLVRGTN TLDLIK                                        26

SEQ ID NO: 187          moltype = AA  length = 29
FEATURE                 Location/Qualifiers
source                  1..29
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 187
YGAGIGPGVY DIHSPRIPPT EEIADRINK                                     29

SEQ ID NO: 188          moltype = DNA  length = 2196
FEATURE                 Location/Qualifiers
source                  1..2196
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 188
atggcttctt ccacttcctc tatctctctc tcccaagccc tcctccgccc caccaccacc  60
tctctcccct cctcctcctc ctccctccgc ctcaccccca ctctccgccc taccccaccc  120
gccctctccc tccgccgccg cctcaccacc tccatccgcg ccgccgcctc cgtcaaaacc  180
gtcgagaagg caaccgcgga cgccgcgctg gtcgaaaaat ccgtcaacac gatccgcttc  240
ctcgccatcg acgccgtcga gaaggcgaat tcgggtcacc cggggctccc catgggatgc  300
gctccgatgg gtcacgtgct atatgatgaa accatgaagt ataaccctaa gaacccttc  360
tggttcaacc gcgatcgctt cgttctctcc gctgggcacg ggtgcatgct tcagtacgcg  420
cttctgcacc tcgctgggtt cgatagcgtc aaggaagagg atttgaggga atttcgccaa  480
tgggggagca gaactccggg acaccctgag aattttgaga ctccgggaat tgaagttaca  540
acaggtcctc ttggtcaggg gattgccaat gctgttggtt tggcccttgc agagaagcac  600
ttggctgcaa gatacaacaa gcctgacaat gagattgttg accattacac gtatgctatt  660
ctgggtgatg gttgtcaaat ggagggaatt gcaaatgaag catgctcact tgctggccac  720
tggggactgg ggaagctgat agcattttat gatgacaatc atatttctat tgatggtaac  780
acagagattg cgttcaccga gagtgttgat agtcgttttg agggacttgg gtggcatgtt  840
atttgggtaa agaatggaaa taatggctat gatgatattc gtgctgccat taaggaagca  900
aaagctgtca aagacaaacc cacattaatc aaggtcacaa caaccattgg ttatggttct  960
cctaacaagg ctaactccta cagtgtgcat ggaagtgcac tgggtgccaa agaagttgat  1020
gccaccaggc agaaccttgg atggtcacat gagccattcc acgtgcctga ggatgtcaaa  1080
aagcattgga gtcgccacac ccctgagggt gctgcacttg aagctgagtg gaatgctaag  1140
tttgctgagt atgaaaagaa atacaaggag gaggctgcag aattgaaatc tattatcaat  1200
ggtgaattcc ctgctggttg ggagaaagca cttccgacat acactccaga gagcccagcg  1260
```

-continued

```
gatgccacca gaaacctgtc tcaaacaaac cttaatgccc ttgcaaaggt tcttcccggt  1320
ctgcttggtg gcagtgcaga tcttgcttct tccaacatga ccttgctcaa aatgttcggg  1380
gacttccaaa aggatactcc agcagagcgt aatgttagat tcggtgttag agaacacgga  1440
atgggagcta tctgcaacgg cattgctctt cacagccctg gactgattcc atattgtgca  1500
accttctttg tattcactga ctacatgaga ggtgccataa ggctttctgc gctgtctgag  1560
gctggggtta tttatgtcat gacccatgat tcaataggac ttggagaaga tgggccaacc  1620
caccagccaa ttgagcacct agcaagcttc cgggcaatgc caaacatttt gatgcttcgt  1680
cccgccgacg gtaacgaaac agccggagca tacaaagtgg ccgtgctcaa caggaagaga  1740
ccctccattc ttgccctatc caggcaaaaa ctgccccagc ttcccggaac ttccattgaa  1800
ggagttgaaa agggtggtta caccatttcg gacaactcca ctggcaacaa gcctgatgtc  1860
attttgatcg gaactggttc ggaattggaa atcgctgcca aagctgctga tgacctaagg  1920
aaggaaggga aggctgttag agttgtttcc cttgtttctt gggaactttt tgatgagcaa  1980
tcagaagcct acaaggagag tgttttccct gctgctgttt cagccagagt tagcattgag  2040
gcaggatcaa catttgggtg ggagaaaatt gttggagcaa agggaaaagc aataggcatt  2100
gatcgttttg gagctagtgc tccagctgga agaatataca aagaatttgg tatcactaag  2160
gaagctgttg ttgctgcagc taaagagctt atctag                            2196

SEQ ID NO: 189          moltype = AA  length = 731
FEATURE                 Location/Qualifiers
source                  1..731
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 189
MASSTSSISL SQALLRPTTT SLPSSSSSLR LTPTLRPTPP ALSLRRRLTT SIRAAASVKT  60
VEKATADAAL VEKSVNTIRF LAIDAVEKAN SGHPGLPMGC APMGHVLYDE TMKYNPKNPF  120
WFNRDRFVLS AGHGCMLQYA LLHLAGFDSV KEEDLREFRQ WGSRTPGHPE NFETPGIEVT  180
TGPLGQGIAN AVGLALAEKH LAARYNKPDN EIVDHYTYAI LGDGCQMEGI ANEACSLAGH  240
WGLGKLIAFY DDNHISIDGN TEIAFTESVD SRFEGLGWHV IWVKNGNNGY DDIRAAIKEA  300
KAVKDKPTLI KVTTTIGYGS PNKANSYSVH GSALGAKEVD ATRQNLGWSH EPFHVPEDVK  360
KHWSRHTPEG AALEAEWNAK FAEYEKKYKE EAAELKSIIN GEFPAGWEKA LPTYTPESPA  420
DATRNLSQTN LNALAKVLPG LLGGSADLAS SNMTLLKMFG DFQKDTPAER NVRFGVREHG  480
MGAICNGIAL HSPGLIPYCA TFFVFTDYMR GAIRLSALSE AGVIYVMTHD SIGLGEDGPT  540
HQPIEHLASF RAMPNILMLR PADGNETAGA YKVAVLNRKR PSILALSRQK LPQLPGTSIE  600
GVEKGGYTIS DNSTGNKPDV ILIGTGSELE IAAKAADDLR KEGKAVRVVS LVSWELFDEQ  660
SEAYKESVFP AAVSARVSIE AGSTFGWEKI VGAKGKAIGI DRFGASAPAG RIYKEFGITK  720
EAVVAAAKEL I                                                       731

SEQ ID NO: 190          moltype = AA  length = 18
FEATURE                 Location/Qualifiers
source                  1..18
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 190
AIGIDRFGAS APAGRIYK                                                18

SEQ ID NO: 191          moltype = AA  length = 27
FEATURE                 Location/Qualifiers
source                  1..27
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 191
ALPTYTPESP ADATRNLSQT NLNALAK                                      27

SEQ ID NO: 192          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 192
ANSYSVHGSA LGAK                                                    14

SEQ ID NO: 193          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 193
ATADAALVEK                                                         10

SEQ ID NO: 194          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 194
DKPTLIK                                                            7
```

-continued

```
SEQ ID NO: 195          moltype = AA  length = 30
FEATURE                 Location/Qualifiers
source                  1..30
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 195
GGYTISDNST GNKPDVILIG TGSELEIAAK                                  30

SEQ ID NO: 196          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 196
LPQLPGTSIE GVEK                                                   14

SEQ ID NO: 197          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 197
NGNNGYDDIR AAIK                                                   14

SEQ ID NO: 198          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 198
SVNTIRFLAI DAVEK                                                  15

SEQ ID NO: 199          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 199
VTTTIGYGSP NK                                                     12

SEQ ID NO: 200          moltype = DNA  length = 1797
FEATURE                 Location/Qualifiers
source                  1..1797
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 200
atggaagttc cttcggaggt accatctact gacactggca ccagcaccgg caccggcgaa  60
accatgagca agaatgcgtt gaagcgggag ctgaagaaca aacagagaga agaagaaagg  120
aagcgcaagg atgaggagaa ggccaaaaag gcagctgaag cacaacaggc aaaggataac  180
aaatctgcac cagctgatga tgaggatatg gatccaactc aataccttga gaacaggcta  240
aagtatcttg cagttcaaaa gacagaaggg aagaacccgt atcctcataa attctttgtc  300
actatgtcta ttgatcaata catcaaggaa tatggaggtt taagcaacgg gcaacacatt  360
gaggatgtct ctgtgtcttt ggctggccgg atcatgcaca aacgcacctc tggttctaag  420
cttgtctttt atgacctgca cggtggtggt tgcaaggtcc aggttatggc tgatgctagt  480
aaatcagact tggatgaggc tggattttcc aaattccatt ctaatgtgaa gcgcggtgac  540
atagttggtg tcactgggtt cccaggcaaa agtaagaagg gcgaacttag tatttttccc  600
aagaattttg tgtcgctgtc tcattgtttg cacatgatgc caaggcaaaa gtctgctgct  660
gctgctattg ataatagaaa tccatggata ccgggaagta ctaggaatcc tgaaacatat  720
attttgaagg atcaggaaac tagatatcgg ttacgccatt tggatttgat gcttaatcca  780
gaggttcgag acatattcaa gaccaggtct aaaatcattt cttacattag gaggttcctt  840
gatgaccttg atttcttgga ggttgaaaca cctatgatga atatgattgc tggtggagct  900
gcagcccgtc catttgtaac acatcacaat gatcttaaca tgaggttatt catgaggatt  960
gctccagaac tgtatcttaa ggagttggtt gttggtggac tggatcgtgt ttatgaaatt  1020
ggtaaacaat ttaggaatga gggcatagat ttgactcata atcctgagtt tactacctgt  1080
gagttctata tggcttataa ggactacaat gacttaatgg atctaacaga gcaaatgttg  1140
agtggtatgg ttaaggagct tacgaaaggc agctataaaa tcaagtatca tgcagatggg  1200
attgataagg aacctattga aattgacttt actacacctt ttagaaggat tgacatgatt  1260
gaagagttag agaagatggc aggtctaagt attcccaaag acttgtcaag tgaggaagcc  1320
aatcaatatt tgaaggatgc atgcgtgaag tatgagatca aatgtcctcc ccctgagaca  1380
actgctcgtt tgttggataa acttgtgggt cactttttgg aagaaacgtg tgtaaatcct  1440
acattcatca aaaaccatcc tgagatcatg agtcctctag caaagtggca cagatcaaaa  1500
ccaggcctga ctgaacgttt tgaattgttt gttaataaac gtgaactttg caatgcatat  1560
actgaattga atgaccctgt agtacaacga cagagatttg ctgaacaact caaggatcgg  1620
caatctggcg atgatgaagc aatggcctta gatgaaacat tttgtacggc tctagaatat  1680
ggtttaccac ctactggtgg ttggggtttg ggcattgatc ggttgaccat gttactgact  1740
```

-continued

```
gattcacaga atattaagga ggttcttctc ttccctgcca tgaaacctca agactga     1797

SEQ ID NO: 201          moltype = AA  length = 598
FEATURE                 Location/Qualifiers
source                  1..598
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 201
MEVPSEVPST DTGTSTGTGE TMSKNALKRE LKNKQREEER KRKDEEKAKK AAEAQQAKDN  60
KSAPADDEDM DPTQYLENRL KYLAVQKTEG KNPYPHKFFV TMSIDQYIKE YGGLSNGQHI  120
EDVSVSLAGR IMHKRTSGSK LVFYDLHGGG CKVQVMADAS KSDLDEAGFS KFHSNVKRGD  180
IVGVTGFPGK SKKGELSIFP KNFVSLSHCL HMMPRQKSAA AAIDNRNPWI PGSTRNPETY  240
ILKDQETRYR LRHLDLMLNP EVRDIFKTRS KIISYIRRFL DDLDFLEVET PMMNMIAGGA  300
AARPFVTHHN DLNMRLFMRI APELYLKELV VGGLDRVYEI GKQFRNEGID LTHNPEFTTC  360
EFYMAYKDYN DLMDLTEQML SGMVKELTKG SYKIKYHADG IDKEPIEIDF TTPFRRIDMI  420
EELEKMAGLS IPKDLSSEEA NQYLKDACVK YEIKCPPPET TARLLDKLVG HFLEETCVNP  480
TFIKNHPEIM SPLAKWHRSK PGLTERFELF VNKRELCNAY TELNDPVVQR QRFAEQLKDR  540
QSGDDEAMAL DETFCTALEY GLPPTGGWGL GIDRLTMLLT DSQNIKEVLL FPAMKPQD    598

SEQ ID NO: 202          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 202
DLSSEEANQY LK                                                      12

SEQ ID NO: 203          moltype = DNA  length = 1728
FEATURE                 Location/Qualifiers
source                  1..1728
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 203
atgtaccgct ttgcaacctc cctcgcctcc aaagctagaa ttgccaggag cagcactcaa  60
cagattggaa gtagggtcac ttggaacagg aattatgcgg ccaaagacat taagtttggt  120
gtggaggctc gggctctgat gcttaagggt gtcgaagagc tcgctgatgc cgttaaagta  180
acaatgggcc ctaaggggcg taatgtggtt attgagcaaa gttttggagc acctaaagtg  240
acgaaagatg gagtaaccgt tgctaagagc attgaattca aggacaaagt caagaatatc  300
ggtgccagtc ttgtaaagca ggttgctaat gctactaatg acgtggctgg tgatggaacc  360
acttgtgcta caattcttac taaagcaata tttacggaag gatgtaaatc agttgcagct  420
ggaatgaatg cgatggacct gagacgaggt ataaatatgg ctgttgatgc tgtggtgaca  480
aatctgaaaa gcagagcacg tatgattagc acttctgaag aaattgccca ggttggaaca  540
atatctgcca atggggagag ggaaattggt gagttaattg caaaagctat ggagaaagtt  600
ggcaaagagg gtgtcatcac aatctcagat ggcaagacat tgtacaatga gttggaagtc  660
gttgaaggaa tgaagctcga caggggatac atttctccat atttcataac aaaccagaag  720
aaccagaaat gtgaacttga ggatcctctc attataatcc atgaaaagaa aatctcaagc  780
ataaatgcta tcgttaaagt attagagtta gctttgaaga gacaaagatc tttattgatt  840
gtcgccgagg atgtggaaag tgatgccctt gcaactctta ttctaaataa acttcgtgct  900
ggaatcaagg tatgtgccat caaagctcct ggttttggtg aaaaccgtaa atctggtctc  960
caggatcttg ctgttcttac aggaggtcaa cttatcaccg aagagcttgg tttgaatctt  1020
gaaaaagtag acctggactt atttggctct tgcaagaaga taacaatttc taaagatgac  1080
actgtcatcc ttgatggagc tggtgataag aaagcaattg aggaaagatg tgagcagatt  1140
agatcagcaa ttgaaaatag cacttcagat tatgacaagg aaaagttaca ggaaagatta  1200
gccaagcttt ctgggggtgt tgcagtactg aagattggag gagccagtga ggctgaagtt  1260
ggtgagaaga aggatagagt gacagatgcc ttaaatgcaa ccaaggcagc tgtagaggag  1320
ggcatagtac ctggtggtgg tgttgctctt ctttatgcat cgagtgagtt ggataagctt  1380
caaactgcca actttgatca aaagataggt gtccagatca tccaaaatgc tttgaagact  1440
cctgtgcaca caattgcttc aaatgcagga gttgagggtg ctgttgttgt tggcaaacta  1500
ttggaacaga atgatcctga tcttggatat gatgcagcca aaggtgaata tgttgatatg  1560
gttaaaactg gaattattga tccattgaag gttattagga ccgccttggt tgatgcagcc  1620
agtgtgtcat ctttgatgac aacaactgaa gctgttgtgt ctgaactccc gaaggacgat  1680
aaagatgttc ctgccatggg tggtggcatg ggtggcatgg attattaa               1728

SEQ ID NO: 204          moltype = AA  length = 575
FEATURE                 Location/Qualifiers
source                  1..575
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 204
MYRFATSLAS KARIARSSTQ QIGSRVTWNR NYAAKDIKFG VEARALMLKG VEELADAVKV  60
TMGPKGRNVV IEQSFGAPKV TKDGVTVAKS IEFKDKVKNI GASLVKQVAN ATNDVAGDGT  120
TCATILTKAI FTEGCKSVAA GMNAMDLRRG INMAVDAVVT NLKSRARMIS TSEEIAQVGT  180
ISANGEREIG ELIAKAMEKV GKEGVITISD GKTLYNELEV VEGMKLDRGY ISPYFITNQK  240
NQKCELEDPL IIIHEKKISS INAIVKVLEL ALKRQRSLLI VAEDVESDAL ATLILNKLRA  300
GIKVCAIKAP GFGENRKSGL QDLAVLTGGQ LITEELGLNL EKVDLDLFGS CKKITISKDD  360
TVILDGAGDK KAIEERCEQI RSAIENSTSD YDKEKLQERL AKLSGGVAVL KIGGASEAEV  420
```

-continued

```
GEKKDRVTDA LNATKAAVEE GIVPGGGVAL LYASSELDKL QTANFDQKIG VQIIQNALKT  480
PVHTIASNAG VEGAVVVGKL LEQNDPDLGY DAAKGEYVDM VKTGIIDPLK VIRTALVDAA  540
SVSSLMTTTE AVVSELPKDD KDVPAMGGGM GGMDY                             575

SEQ ID NO: 205          moltype = AA  length = 24
FEATURE                 Location/Qualifiers
source                  1..24
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 205
AAVEEGIVPG GGVALLYASS ELDK                                          24

SEQ ID NO: 206          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 206
DRVTDALNAT K                                                        11

SEQ ID NO: 207          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 207
EGVITISDGK                                                          10

SEQ ID NO: 208          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 208
GRNVVIEQSF GAPK                                                     14

SEQ ID NO: 209          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 209
GVEELADAVK                                                          10

SEQ ID NO: 210          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 210
IGGASEAEVG EK                                                       12

SEQ ID NO: 211          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 211
IGVQIIQNAL K                                                        11

SEQ ID NO: 212          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 212
ISSINAIVK                                                           9

SEQ ID NO: 213          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 213
LLEQNDPDLG YDAAK                                                    15
```

-continued

```
SEQ ID NO: 214          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 214
LQTANFDQK                                                           9

SEQ ID NO: 215          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 215
LSGGVAVLK                                                           9

SEQ ID NO: 216          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 216
NIGASLVK                                                            8

SEQ ID NO: 217          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 217
TGIIDPLK                                                            8

SEQ ID NO: 218          moltype = AA  length = 20
FEATURE                 Location/Qualifiers
source                  1..20
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 218
TPVHTIASNA GVEGAVVVGK                                               20

SEQ ID NO: 219          moltype = DNA  length = 1725
FEATURE                 Location/Qualifiers
source                  1..1725
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 219
atgtatcgct ttgcttctaa cctcgcttcc aaagcaagga ttgctaggag tagcagccac  60
cagattggaa gtaggctgag ttcgagcaga aactatgcgg ctaaagacat tagatttggt  120
gtggaggctc gggcactgat gcttaagggt gttgaagagc tagctgatgc tgtcaaagta  180
accatgggtc ccaaggggcg taatgttgtg attgagcaaa gttttggtgc ccctaaagta  240
accaaagatg gagtaactgt tgcaaagagc attgaattca aggataaagt taagaatgtt  300
ggtgccagtc ttgtaaagca ggttgcaaat gctactaatg atgtggctgg tgatggaacc  360
acgtgcgcta cagtccttac acgagcaata tttactgaag gctgcaaatc agttgcggct  420
ggaatgaatg caatggacct gaggcgtggt ataagtatgg ctgtcgatgc tgtggtgaca  480
aacttgaaaa gcagagcaag gatgattagc acatctgaag aaatagcaca ggtagggaca  540
atatctgcta acggagagag agaaattggt gagttgattg caaaagctat ggagaaagtt  600
ggcaaagagg gagtaatcac aatttcagat ggcaaaacct tatataatga gttggaagtt  660
gttgaaggaa tgaagctcga caggggctac atatctcctt atttcattac aaatgataag  720
aaccagaaat gtgaacttga agatcccctc attctaattc atgagaagaa aatctcgagt  780
ataaatgcta tagttaaagt cttagagttg gctttgaaga gacaaagatc tttattgatt  840
attgccgagg atgtggaaag tgatgcactt gcaactctta ttctaaataa gcttcgggct  900
ggaattaagg tatgtgctat taaagcccct ggttttggcg aaaataggaa agccaatctt  960
caagatcttg cagttctcac aggaggagca cttattactg aagagcttgg cttgaagctt  1020
gaaaaagtgg atttggatat gcttggcacc tgtaaaaaaa taacagtttc taaggatgac  1080
actgtcattc ttgacggagc aggtgacaag aaagcacttg aggaaagatg cgagcagatt  1140
aggtctgcaa ttgaaaacag cacctcagat tatgacaagg aaaagttaca ggaaagatta  1200
gccaagcttt ctggtggtgt tgccgtgctt aagattggag gagccagtga agctgaagtt  1260
ggtgagaaga aggatagagt gacagatgcc ttaaatgcaa ctaaggctgc tgtagaggaa  1320
ggcattgtat caggtggtgg tgttgctctt ctatatgcgt caaaagagtt ggataaactt  1380
caaactgcta actttgatca gaagattggt gtccagatta tccaaaatgc tttaaagaca  1440
cctgtgctca caattgcatc aaatgcgggt gttgagggtg cagttgttgt tggcaaacta  1500
ttggaacagg agaatcatga tctaggatat gatgcagcca aaggtgaata tgttgatatg  1560
gttaaagctg gaatcataga ccctttgaag gtgattagga cagccttggt tgatgcagcc  1620
agtgtatcat ctttgatgac aacaactgaa gctgttgtgt ctgaacttcc aaacgatgat  1680
aaagatactc cagctatggc tggtggcatg ggtggcatgg gttat                  1725
```

-continued

```
SEQ ID NO: 220          moltype = AA  length = 575
FEATURE                 Location/Qualifiers
source                  1..575
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 220
MYRFASNLAS KARIARSSSH QIGSRLSSSR NYAAKDIRFG VEARALMLKG VEELADAVKV  60
TMGPKGRNVV IEQSFGAPKV TKDGVTVAKS IEFKDKVKNV GASLVKQVAN ATNDVAGDGT  120
TCATVLTRAI FTEGCKSVAA GMNAMDLRRG ISMAVDAVVT NLKSRARMIS TSEEIAQVGT  180
ISANGEREIG ELIAKAMEKV GKEGVITISD GKTLYNELEV VEGMKLDRGY ISPYFITNDK  240
NQKCELEDPL ILIHEKKISS INAIVKVLEL ALKRQRSLLI IAEDVESDAL ATLILNKLRA  300
GIKVCAIKAP GFGENRKANL QDLAVLTGGA LITEELGLKL EKVDLDMLGT CKKITVSKDD  360
TVILDGAGDK KALEERCEQI RSAIENSTSD YDKEKLQERL AKLSGGVAVL KIGGASEAEV  420
GEKKDRVTDA LNATKAAVEE GIVSGGGVAL LYASKELDKL QTANFDQKIG VQIIQNALKT  480
PVLTIASNAG VEGAVVVGKL LEQENHDLGY DAAKGEYVDM VKAGIIDPLK VIRTALVDAA  540
SVSSLMTTTE AVVSELPNDD KDTPAMAGGM GGMGY                             575

SEQ ID NO: 221          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 221
AGIIDPLK                                                           8

SEQ ID NO: 222          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 222
DRVTDALNAT K                                                       11

SEQ ID NO: 223          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 223
EGVITISDGK                                                         10

SEQ ID NO: 224          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 224
GRNVVIEQSF GAPK                                                    14

SEQ ID NO: 225          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 225
GVEELADAVK                                                         10

SEQ ID NO: 226          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 226
IGGASEAEVG EK                                                      12

SEQ ID NO: 227          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 227
IGVQIIQNAL K                                                       11

SEQ ID NO: 228          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
```

-continued

```
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 228
ISSINAIVK                                                           9

SEQ ID NO: 229          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 229
LLEQENHDLG YDAAK                                                    15

SEQ ID NO: 230          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 230
LQTANFDQK                                                           9

SEQ ID NO: 231          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 231
LSGGVAVLK                                                           9

SEQ ID NO: 232          moltype = DNA  length = 1869
FEATURE                 Location/Qualifiers
source                  1..1869
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 232
atggctcaag ttcttgagaa ccccaccatt gacgcggcgg cttccggcgc gaacccctcg  60
ctgacgacca tctcgttgta cgtcggggat cttcaccacg acgtcaacga tccgcagctg  120
tatgatctct tcaaccaggt tgcccaagtc gtttccgttc gaatttgcag ggacgtggcc  180
acccaacaat cccttggtta cggctacgtc aatttcagca acgctcacga tgccgcaaaa  240
gcaattgatg tgctgaattt cactccactg aacggcaaaa tcattaggat aatgtactct  300
attcgggacc ctagtgctcg gaagagtggt gcagctaatg ttttcataaa gaatttggac  360
aaggcgattg atcacaaggc tttgtacgat acattttcgg cttttgggaa tattctatct  420
tgcaaggttg cgacagatgc ttctggccag tctaaaggcc atggttttgt gcagtttgag  480
agtgaagaat ctgcacaaaa tgcaattgac aagttaaacg gcatgctgat caatgataag  540
caggtctttg tgggtccttt cctgcggaag caggatagag agagtgctct cagtggcaca  600
aagtttaata acgttttcgt gaaaaacctg ttggattcaa tgactgaggc tgacttggag  660
agaatttttg gagaatatgg agctattact agcgctgtag tgatgagaga tgtagatggt  720
aaatcaaaag gttttggttt tgtcaatttt gcaaatgtag acgatgctgc taaagctgtg  780
gaggctctta atggaaagaa ttttgatggt aaggagtggt atgttggaaa agcccagaaa  840
aagtctgaaa gggagcttga actgaaaggg caacatgaac agattacaaa ggagactgtt  900
gataaatatc atggtacaaa cttgtatatc aagaacttgg atgatagtgt tggtgatgaa  960
gaactcatgg aactgttctc tgagtttggt acgataacct catgcaaggt tatgcgagat  1020
ccaaatggaa ttagtagagg atctggattt gtttcatttt caattgctga gggagcaact  1080
cgggctcttg gtgagatgaa tggtaaaatg gttgctggca aacctctcta tgttgccctt  1140
gcacagagaa aagaagatag aagagcaagg ttacaggcac aattttcaca atcgaggcct  1200
gctgcaataa cccctaatgt ttcacctcga atgcctctat accctctagg ggctcctgcg  1260
ataggacaac aatttttgta tgggcaagca gctccagcca cgatacctca acattttgtt  1320
cctggaatga ggccaggtgg tgctccaaac ttctatgttc cactggttca gcagggtcaa  1380
cagggacagc gcccaggagg acaccgagga tcaggtccta tgcaacagct ccagcagcca  1440
ttgccattga tgcaccaaca gatgcttcca aggggacatg tctatcgata tccttctggc  1500
tccaatatgc aaaatattca actggcagga gtacctggag gaatggtcac ttatggcatt  1560
ggtcaacaaa tgcccacgca ggccctggcc tcggcccttg ccaatgctac tcctgaacag  1620
cagaggacta tgcttggtga agctttatac ccgttagtag ataagctgga gcatgaagca  1680
gcagcaaagg ttactggtat gctcttagag atggatcagc ctgaggtatt acatctgatt  1740
gagtcaccgg atgctctcaa ggctaaagtt gttgaggcca tggatgtgtt gaagaaagtt  1800
acacagcagc aatcaaacag cccagccgat caactggctt cactctctct gaatgataat  1860
gtcgagtct                                                          1869

SEQ ID NO: 233          moltype = AA  length = 623
FEATURE                 Location/Qualifiers
source                  1..623
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 233
MAQVLENPTI DAAASGANPS LTTISLYVGD LHHDVNDPQL YDLFNQVAQV VSVRICRDVA  60
TQQSLGYGYV NFSNAHDAAK AIDVLNFTPL NGKIIRIMYS IRDPSARKSG AANVFIKNLD  120
```

-continued

```
KAIDHKALYD TFSAFGNILS CKVATDASGQ SKGHGFVQFE SEESAQNAID KLNGMLINDK  180
QVFVGPFLRK QDRESALSGT KFNNVFVKNL LDSMTEADLE RIFGEYGAIT SAVVMRDVDG  240
KSKGFGFVNF ANVDDAAKAV EALNGKNFDG KEWYVGKAQK KSERELELKG QHEQITKETV  300
DKYHGTNLYI KNLDDSVGDE ELMELFSEFG TITSCKVMRD PNGISRGSGF VSFSIAEGAT  360
RALGEMNGKM VAGKPLYVAL AQRKEDRRAR LQAQFSQSRP AAITPNVSPR MPLYPLGAPA  420
IGQQFLYGQA APATIPQHFV PGMRPGGAPN FYVPLVQQGQ QGQRPGGHRG SGPMQQLQQP  480
LPLMHQQMLP RGHVYRYPSG SNMQNIQLAG VPGGMVTYGI GQQMPTQALA SALANATPEQ  540
QRTMLGEALY PLVDKLEHEA AAKVTGMLLE MDQPEVLHLI ESPDALKAKV VEAMDVLKKV  600
TQQQSNSPAD QLASLSLNDN VES                                          623

SEQ ID NO: 234          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 234
AIDVLNFTPL NGK                                                      13

SEQ ID NO: 235          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 235
FNNVFVK                                                             7

SEQ ID NO: 236          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 236
GFGFVNFANV DDAAK                                                    15

SEQ ID NO: 237          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 237
SGAANVFIK                                                           9

SEQ ID NO: 238          moltype = DNA  length = 1851
FEATURE                 Location/Qualifiers
source                  1..1851
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 238
atggattctg attatggtat tcccagggaa ctctcagatc ttcaaaagat tcggtctttg  60
taccagccag agctccctcc ttgtctccag ggaaccactg tgagggttga atttggtgac  120
gcaaccacca ctgctgaccc cactgatgca ctcaccgtct gcagggcttt tcctaacact  180
tatggacacc ctttggctca ctttctcagg gccactgcca aagtccctga tgctcagatc  240
ataactgagt tccctcctat tagggtgggg attgtctttt gtgggagaca atctcctgga  300
gggcataatg tcatttgggg tatccacaat gctctcaaaa ttcacaatcc caacagtgtt  360
ttgcttggat ttcttggtgg ttcagaaggt ctatttgccc agaaaactct ggagataact  420
gacgatattc tttcgacgta caaaaatcaa ggtggttatg atttgcttgg gcggacgaag  480
gatcaaatta ggaccacaga acaagttaat gctgcacttg ctgcatgcaa caatttgaaa  540
cttgatggcc ttgtcataat cggaggggtg acatcaaaca cagatgctgc acagcttgct  600
gaaacatttg ttgtagcaaa atgccccaca aaggtggttg gtgttcctgt cactctgaat  660
ggagatctta aaaaccagtt tgtagaaact aatgttggtt ttgatacaat ttgtaaggtt  720
aattctcagc tcatcagcaa tgtctgcact gatgctctct ctgcagagaa gtattattat  780
ttcatccgtc ttatgggacg caaggcatca cacgttgctt tggaatgcac cctccagtcc  840
catccaaata tggtaattct tggtgaggag gttgctgcat caaagcttac cctttctgaa  900
attgcaaaac agattactga tgcaattcag gctagagcag agcaagacaa ataccatgga  960
gtcattctcc ttccagaagg gctgatagag agtattcctg aagtgtatgc cctcttgaag  1020
gaaattcatg gtttactcag acaaggtgtt gaagttgaca agatatcttc tcagctaaca  1080
ccatgggcat ctgctttatt tgaatttctg ccacccttta ttaggagaca gttgctcctt  1140
taccctgaat ccgacgactc agcacagtta tcacagattg aaactgagaa attactagca  1200
tatcttgtgg aagtcgagat taacaagcgt ctgaaagaag gcacatacaa ggggaagaaa  1260
ttcaatgcca tttgccactt ttttggttat caagctcgtg gatctctgcc atcaaaattt  1320
gactgtgatt acgcatatgt ccttgggcac atctgctacc atatactggc cgctggtcta  1380
aatggataca tggcaacttt aactaacctg aagaatccag taaacaagtg gaaatgtggg  1440
gctgctccaa ttgcatctat gatgactgta aagcgctggt ctccaaatcc aggagctaca  1500
tcaattggaa aacctgctat tcatccagct actgtggact tgagaggcaa agcatatgag  1560
ctgttgagac aaaaggcaca aagctttctt atggatgaca tctacagaaa tccaggtcca  1620
cttcagtttg atggtcccgg cgcagatgcc aaggttataa ctctttctgt tgaagaccag  1680
```

-continued

```
gactacatgg gccgcattaa gaaacttcag gagtaccttg aacaggttcg aaccattgtc  1740
aagccaggat gccctcagga ggttttgaaa gctgctctaa gcgtcatggc gtcggttacc  1800
gaagttctag ctgcaatgtc aacatcttat accaacaccc tggcatccct g           1851

SEQ ID NO: 239          moltype = AA  length = 617
FEATURE                 Location/Qualifiers
source                  1..617
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 239
MDSDYGIPRE LSDLQKIRSL YQPELPPCLQ GTTVRVEFGD ATTTADPTDA LTVCRAFPNT  60
YGHPLAHFLR ATAKVPDAQI ITEFPPIRVG IVFCGRQSPG GHNVIWGIHN ALKIHNPNSV  120
LLGFLGGSEG LFAQKTLEIT DDILSTYKNQ GGYDLLGRTK DQIRTTEQVN AALAACNNLK  180
LDGLVIIGGV TSNTDAAQLA ETFVVAKCPT KVVGVPVTLN GDLKNQFVET NVGFDTICKV  240
NSQLISNVCT DALSAEKYYY FIRLMGRKAS HVALECTLQS HPNMVILGEE VAASKLTLSE  300
IAKQITDAIQ ARAEQDKYHG VILLPEGLIE SIPEVYALLK EIHGLLRQGV EVDKISSQLT  360
PWASALFEFL PPFIRRQLLL YPESDDSAQL SQIETEKLLA YLVEVEINKR LKEGTYKGKK  420
FNAICHFFGY QARGSLPSKF DCDYAYVLGH ICYHILAAGL NGYMATLTNL KNPVNKWKCG  480
AAPIASMMTV KRWSPNPGAT SIGKPAIHPA TVDLRGKAYE LLRQKAQSFL MDDIYRNPGP  540
LQFDGPGADA KVITLSVEDQ DYMGRIKKLQ EYLEQVRTIV KPGCPQEVLK AALSVMASVT  600
EVLAAMSTSY TNTLASL                                                 617

SEQ ID NO: 240          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 240
VVGVPVTLNG DLK                                                     13

SEQ ID NO: 241          moltype = DNA  length = 1575
FEATURE                 Location/Qualifiers
source                  1..1575
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 241
atgcccaaat tttttcatag cattttcaga ggaacaatgg cgggtagggt ttcgacgtgc  60
ttcttcttcg tgtttgctct ttctctcctt cttccttttc agatctccgc tgaggaatca  120
tcggagaaag agttcgtact cacattggat cactctaact ttcacgacac tgtcagcaag  180
cacgatttca tcgtcgtcga gttctacgct ccatggtgtg gccactgtaa gaagcttgct  240
cccgagtatg agaaagctgc ttctatcttg agtagtcatg atcctccggt tgttttggct  300
aaaattgatg ccaatgagga gaagaacaaa gaccttgcat ctcaatatga tgttagggga  360
tacccaacaa ttaagattct gagaaatggg ggaaagaatg ttcaggaata caaaggtccc  420
cgtgaagcag atggcatcgt tgactatttg aagaaacaaa gtggtcctgc ttcaactgaa  480
attaaatctg ctgatgaagc tactgctttc attggtgaaa ataaagttgc tattgttgga  540
gttttcccga aattttctgg agaggagttt gacaactttt ctgccttagc agagaagttg  600
cgttctgact atgacttcgg tcacactctg aatgccaaac accttccaag aggagagtca  660
tcagtttctg gtcccgtagt taggttgttc aagccattcg atgaactttt tgttgacttc  720
caggatttca atgtggaagc tctagagaaa tttgttgagg aatctagtac ccctgttgtt  780
actgtcttta acaatgatcc gagcaatcac cctttcgttg ccaaattctt caacagtccc  840
aatgcaaagg caatgttgtt catcaacttc actgccgaag gtgctgaatc tttcaaatca  900
aaataccgtg aagctgctga gcaacataaa caacagggtg taagcttttt ggtgggagat  960
gttgagtcta gtcaaggtgc tttccagtat tttggactga aagaagaaca agtacctcta  1020
attatcattc aacacaatga cgggaaaaag ttttttaaac ccaatttgga agctgatcac  1080
attccaactt ggttgaaggc gtacaaggat ggtaatgttg caccatttgt gaagtctgaa  1140
cccattcctg aagctaacga tgaacctgtt aaagtggtag ttgggaacag tcttgaggac  1200
attgttttca aatctgggaa gaatgttctg ctggagtttt atgctccctg gtgtggtcat  1260
tgcaaacagt tggctccaat attggacgaa gttgctatct catatcaaag tgatgctgat  1320
gttgttattg caaaactgga tgcaactgcc aacgatatcc caagtgaaac ctttgatgtc  1380
caaggttatc caaccgtgta cttcaggtca gcaagtggaa agttatcaca atacgaaggg  1440
ggtaggacca aggaagacat catagaattc atcgaaaaga accgggacaa acctgctcag  1500
caagaacaag gacaagataa acctgctcag caagaacaag gacaagatga gcaagaaaaa  1560
ggaaaagatg agctt                                                   1575

SEQ ID NO: 242          moltype = AA  length = 525
FEATURE                 Location/Qualifiers
source                  1..525
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 242
MPKFFHSIFR GTMAGRVSTC FFFVFALSLL LPFQISAEES SEKEFVLTLD HSNFHDTVSK  60
HDFIVVEFYA PWCGHCKKLA PEYEKAASIL SSHDPPVVLA KIDANEEKNK DLASQYDVRG  120
YPTIKILRNG GKNVQEYKGP READGIVDYL KKQSGPASTE IKSADEATAF IGENKVAIVG  180
VFPKFSGEEF DNFSALAEKL RSDYDFGHTL NAKHLPRGES SVSGPVVRLF KPFDELFVDF  240
QDFNVEALEK FVEESSTPVV TVFNNDPSNH PFVAKFFNSP NAKAMLFINF TAEGAESFKS  300
KYREAAEQHK QQGVSFLVGD VESSQGAFQY FGLKEEQVPL IIIQHNDGKK FFKPNLEADH  360
```

-continued

```
IPTWLKAYKD GNVAPFVKSE PIPEANDEPV KVVVGNSLED IVFKSGKNVL LEFYAPWCGH  420
CKQLAPILDE VAISYQSDAD VVIAKLDATA NDIPSETFDV QGYPTVYFRS ASGKLSQYEG  480
GRTKEDIIEF IEKNRDKPAQ QEQGQDKPAQ QEQGQDEQEK GKDEL                  525

SEQ ID NO: 243          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 243
AASILSSHDP PVVLAK                                                  16

SEQ ID NO: 244          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 244
DGNVAPFVK                                                          9

SEQ ID NO: 245          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 245
DLASQYDVRG YPTIK                                                   15

SEQ ID NO: 246          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 246
EDIIEFIEK                                                          9

SEQ ID NO: 247          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 247
EEQVPLIIIQ HNDGK                                                   15

SEQ ID NO: 248          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 248
EFVLTLDHSN FHDTVSK                                                 17

SEQ ID NO: 249          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 249
FFKPNLEADH IPTWLK                                                  16

SEQ ID NO: 250          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 250
FFNSPNAK                                                           8

SEQ ID NO: 251          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 251
FSGEEFDNFS ALAEK                                                   15
```

-continued

```
SEQ ID NO: 252          moltype = AA  length = 25
FEATURE                 Location/Qualifiers
source                  1..25
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 252
FVEESSTPVV TVFNNDPSNH PFVAK                                          25

SEQ ID NO: 253          moltype = DNA  length = 1509
FEATURE                 Location/Qualifiers
source                  1..1509
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 253
atggcgggta cggtttcgac ttgtttcttc ttcgtgtttc ttctttctct ccttcttcct  60
tttcagatct ccgctgagga atcctcggag aaggagttcg tactcacatt ggatcactct  120
aactttcacg acactgtcag caagcacgat ttcatcgtcg tcgagttcta cgccccatgg  180
tgtggtcact gtaagaagct tgctccggag tatgagaaag ctgcttctat cttgagtagt  240
catgatcctc ctattgtttt ggctaaagta gatgccaatg aggagaagaa caaagacctt  300
gcatcacaat atgatgttaa gggattccca acaattaata ttctgagaaa tgggggaaag  360
aatgttcagg aatacaaagg tccccgtgaa gcagatggca tcgttgacta tttgaagaaa  420
caaagtggtc ctgcttcaac tgaaattaaa tctgctgatg aagctactgc ttttattggt  480
gaaaataaag ttgctattgt tggagttttc ccgaaatttt ctggagagga gtttgacaac  540
ttttctgcct tagcagagaa gttgcgttct gattatgact ttggtcacac tctgaatgcc  600
aaactccttc caagaggaga gtcatcagtt tctgggcccg tagttaggct gttcaagcca  660
ttcgatgaac tttttgttga cttccaggat ttcaatgtgg aagctctaga gaagtttgtt  720
gaggaatcta gtacccctgt tgttactgtc tttaacaatg aaccgagcaa tcacccttc  780
gttgtcaaat tcttcaacag tcccaacgca aaggcaatgt tgttcatcaa cttcactgcc  840
gaaggtgctg aagctatcaa atcaaaatac cgtgaagctg ctgagcaata taaacaacag  900
ggtgtcagct ttttggtggg agatgttgag tctagtcaag gtgctttcca gtattttgga  960
ctgaaagaag agcaagtacc tctaattatc attcaacata atgatgggaa aaagtttttt  1020
aaacccaatt tggaagctga tcacattcca acctggttga aggcatacaa ggatggtcat  1080
gttgcaccat ttgttaagtc tgaacccatt cccgagacta acgatgaacc tgttaaagtg  1140
gtagttgggg ccagtcttga ggacattgtt ttcaaatctg ggaagaatgt tctgctggag  1200
ttttatgctc cctggtgtgg tcattgcaaa cagttggctc caatattgga tgaagttgct  1260
atctcatatc aaaatgaagc tgatgttgtt attgcaaaac tggatgcaac tgccaacgat  1320
atcccaagtg aaacctttga tgtccaaggt tatcctaccg tgtacttcag gtcagcaagt  1380
ggaaagttat cacaatacga cggcggtagg accaaggaag acatcataga attcatcgaa  1440
aagaaccggg ataaacctgc tcagcaagaa caaggaaaag atgagcaaga acaaggaaaa  1500
gatgagctt                                                          1509

SEQ ID NO: 254          moltype = AA  length = 503
FEATURE                 Location/Qualifiers
source                  1..503
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 254
MAGTVSTCFF FVFLLSLLLP FQISAEESSE KEFVLTLDHS NFHDTVSKHD FIVVEFYAPW  60
CGHCKKLAPE YEKAASILSS HDPPIVLAKV DANEEKNKDL ASQYDVKGFP TINILRNGGK  120
NVQEYKGPRE ADGIVDYLKK QSGPASTEIK SADEATAFIG ENKVAIVGVF PKFSGEEFDN  180
FSALAEKLRS DYDFGHTLNA KLLPRGESSV SGPVVRLFKP FDELFVDFQD FNVEALEKFV  240
EESSTPVVTV FNNEPSNHPF VVKFFNSPNA KAMLFINFTA EGAEAIKSKY REAAEQYKQQ  300
GVSFLVGDVE SSQGAFQYFG LKEEQVPLII IQHNDGKKFF KPNLEADHIP TWLKAYKDGH  360
VAPFVKSEPI PETNDEPVKV VVGASLEDIV FKSGKNVLLE FYAPWCGHCK QLAPILDEVA  420
ISYQNEADVV IAKLDATAND IPSETFDVQG YPTVYFRSAS GKLSQYDGGR TKEDIIEFIE  480
KNRDKPAQQE QGKDEQEQGK DEL                                          503

SEQ ID NO: 255          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 255
AASILSSHDP PIVLAK                                                  16

SEQ ID NO: 256          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 256
DGHVAPFVK                                                          9

SEQ ID NO: 257          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
```

-continued

```
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 257
DLASQYDVK                                                          9

SEQ ID NO: 258          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 258
EDIIEFIEK                                                          9

SEQ ID NO: 259          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 259
EEQVPLIIIQ HNDGK                                                   15

SEQ ID NO: 260          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 260
EFVLTLDHSN FHDTVSK                                                 17

SEQ ID NO: 261          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 261
FFKPNLEADH IPTWLK                                                  16

SEQ ID NO: 262          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 262
FFNSPNAK                                                           8

SEQ ID NO: 263          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 263
FSGEEFDNFS ALAEK                                                   15

SEQ ID NO: 264          moltype = AA  length = 25
FEATURE                 Location/Qualifiers
source                  1..25
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 264
FVEESSTPVV TVFNNEPSNH PFVVK                                        25

SEQ ID NO: 265          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 265
GFPTINILRN GGK                                                     13

SEQ ID NO: 266          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 266
GPREADGIVD YLK                                                     13
```

-continued

```
SEQ ID NO: 267          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 267
LAPEYEK                                                             7

SEQ ID NO: 268          moltype = AA  length = 29
FEATURE                 Location/Qualifiers
source                  1..29
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 268
LDATANDIPS ETFDVQGYPT VYFRSASGK                                     29

SEQ ID NO: 269          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 269
LRSDYDFGHT LNAK                                                     14

SEQ ID NO: 270          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 270
LSQYDGGRTK                                                          10

SEQ ID NO: 271          moltype = AA  length = 23
FEATURE                 Location/Qualifiers
source                  1..23
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 271
QLAPILDEVA ISYQNEADVV IAK                                           23

SEQ ID NO: 272          moltype = AA  length = 24
FEATURE                 Location/Qualifiers
source                  1..24
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 272
QQGVSFLVGD VESSQGAFQY FGLK                                          24

SEQ ID NO: 273          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 273
QSGPASTEIK                                                          10

SEQ ID NO: 274          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 274
SADEATAFIG ENK                                                      13

SEQ ID NO: 275          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 275
SEPIPETNDE PVK                                                      13

SEQ ID NO: 276          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
```

-continued

```
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 276
VAIVGVFPK                                                           9

SEQ ID NO: 277          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 277
VVVGASLEDI VFK                                                      13

SEQ ID NO: 278          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 278
YREAAEQYK                                                           9

SEQ ID NO: 279          moltype = DNA  length = 1467
FEATURE                 Location/Qualifiers
source                  1..1467
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 279
atgggtgtgg ctcaaaacat tcccgatgcg gaggagggaa ccttggagat tggaatggaa  60
tacagaactg tttctggagt tgctgggcca ttggtcattc ttgataaagt taagggaccc  120
aagtttcaag agattgttaa cattcgctta ggagatggta ctactcggcg tggacaagtg  180
ctagaagttg atggtgaaaa ggctgttgtt caggtctttg agggtacatc tgggattgac  240
aataaattta caactgtgca atttactgga gaggtgttaa aaactccagt atcattggac  300
atgcttggcc gcatattcaa tggttctgga aaaccaattg ataatggtcc acccatttta  360
cctgaggctt acttggacat atcaggaagt tccatcaacc ctagtgagag aacctatcca  420
gaggagatga tacagacagg aatatctacc attgatgtca tgaattctat tgctagaggt  480
caaaagatcc ctctattctc agctgctggt ctcccccata atgaaattgc tgcacagata  540
tgccgtcagg ctggcttagt caagcgacta gaaaagtctg ataatcttct cgagggtgga  600
ggagaagagg acaattttgc cattgtgttt gcagctatgg gagttaacat ggagactgca  660
cagtttttca aacgtgactt tgaggaaaat ggctcgatgg agagagtgac ccttttcctt  720
aatctggcaa atgatcctac gattgagcgt atcatcactc ctcgtattgc tcttactact  780
gctgaatatt tggcatatga atgtggcaag catgttcttg tgatactaac tgatatgagt  840
tcttatgctg atgctcttcg tgaggtatct gctgcccgag aagaagtacc tggtagacgt  900
gggtatcccg gctacatgta tacagatctt gcaacaatat acgaacgtgc tggaagaatt  960
gaggggcgaa aaggctctat tacacaaatt ccaattctaa ccatgcctaa tgatgatatt  1020
acgcatccaa ctcctgatct tacaggatat attactgagg gacagatata tatcgacaga  1080
cagttgtata acaggcagat atatcctcca atcaatgtcc ttccatcatt gtctcgtttg  1140
atgaagagtg ccattggtga gggaatgaca cgaaaggatc attctgatgt gtccaaccag  1200
ctctatgcaa attatgccat tggaaaggat gtccaggcaa tgaaagccgt ggtcggagaa  1260
gaagcacttt catctgagga cctgctgtat ttggaatttt tggagaagtt tgagaggaag  1320
tttgttgccc aaggagcata cgatacacgc aatatcttcc agtcactaga tcttgcatgg  1380
actttgcttc ggattttccc tcgcgagctt ctccaccgta ttcctgcaaa gactcttgat  1440
cagttctaca gccgagatgc cagtaac                                      1467

SEQ ID NO: 280          moltype = AA  length = 489
FEATURE                 Location/Qualifiers
source                  1..489
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 280
MGVAQNIPDA EEGTLEIGME YRTVSGVAGP LVILDKVKGP KFQEIVNIRL GDGTTRRGQV  60
LEVDGEKAVV QVFEGTSGID NKFTTVQFTG EVLKTPVSLD MLGRIFNGSG KPIDNGPPIL  120
PEAYLDISGS SINPSERTYP EEMIQTGIST IDVMNSIARG QKIPLFSAAG LPHNEIAAQI  180
CRQAGLVKRL EKSDNLLEGG GEEDNFAIVF AAMGVNMETA QFFKRDFEEN GSMERVTLFL  240
NLANDPTIER IITPRIALTT AEYLAYECGK HVLVILTDMS SYADALREVS AAREEVPGRR  300
GYPGYMYTDL ATIYERAGRI EGRKGSITQI PILTMPNDDI THPTPDLTGY ITEGQIYIDR  360
QLYNRQIYPP INVLPSLSRL MKSAIGEGMT RKDHSDVSNQ LYANYAIGKD VQAMKAVVGE  420
EALSSEDLLY LEFLEKFERK FVAQGAYDTR NIFQSLDLAW TLLRIFPREL LHRIPAKTLD  480
QFYSRDASN                                                          489

SEQ ID NO: 281          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 281
DHSDVSNQLY ANYAIGK                                                 17
```

-continued

```
SEQ ID NO: 282          moltype = AA  length = 26
FEATURE                 Location/Qualifiers
source                  1..26
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 282
FQEIVNIRLG DGTTRRGQVL EVDGEK                                        26

SEQ ID NO: 283          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 283
FTTVQFTGEV LK                                                       12

SEQ ID NO: 284          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 284
TLDQFYSRDA SN                                                       12

SEQ ID NO: 285          moltype = DNA  length = 1464
FEATURE                 Location/Qualifiers
source                  1..1464
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 285
atgggtgtgg ctgagaacat tcccgatatg gaggagggta ccttggagat tggaatggaa  60
tacagaactg tatctggagt tgctggacca ttggtcattc ttgacaaagt taagggaccc  120
aaatttcaag agattgttaa tattcgttta ggagatggaa ctactcggcg tggacaagtt  180
ctagaggttg atggcgaaaa ggctgttgtt caggtatttg aaggcacatc tgggatcgac  240
aataagttta caactgtgca atttactgga gaggtgttaa aaactcctgt ctcactggac  300
atgcttggac gcatatttaa tggttctgga aaaccaattg ataatggtcc accaatttta  360
cctgaggctt acctggatat ttctggaagt tccatcaatc ctagtgagag aacctatcct  420
gaggagatga tacagacagg aatatctaca attgatgtca tgaattctat tgctagaggt  480
caaaagatcc ctctattctc agcagctggt ctcccccata atgaaattgc tgcacagatc  540
tgtcgtcagg ctggcttagt caagcgacta gagaaatctg ataatcttct tgagggtgga  600
gaagaggaca attttgctat tgtgtttgca gctatgggag ttaacatgga gactgcacag  660
tttttcaagc gtgactttga ggagaatgga tcaatggaga gagtgactct tttccttaat  720
ttggcaaatg atcctacaat tgagcgtatt atcactcctc gtattgctct tactactgct  780
gaatatttgg cttatgaatg tggcaagcac gttcttgtga tactcacaga catgagttct  840
tatgctgatg ctcttcgtga ggtatctgct gcccgagaag aagtgccggg tagacgtggt  900
tatcctggat atatgtatac agatcttgca acaatttatg aacgagctgg gagaattgag  960
gggcgtaaag gctctattac acaaattcca attctaacca tgcccaatga tgatattacg  1020
cacccaactc ctgatctaac aggatatatc actgaggggc agatatacat cgacaggcag  1080
ctgtataata gacagatata cccaccaatc aacgtcctcc catcactatc tcggttgatg  1140
aagagtgcta ttggtgaggg aatgactaga agggatcatt ctgatgtgtc caaccagctc  1200
tatgcaaatt atgctattgg aaaggatgtc caggcaatga aagcagtggt tggagaagaa  1260
gcactttcat ctgaggacct gctgtatctg gaatttttgg agaagtttga gaggaagttt  1320
gttgcccagg gagcctatga cacccgcaat atattccagt cacttgatct tgcatggact  1380
ttactccgga ttttccctcg tgaacttctt catcgtatac cagccaagac tcttgaccag  1440
ttttacagca gagatgctgg taat                                          1464

SEQ ID NO: 286          moltype = AA  length = 488
FEATURE                 Location/Qualifiers
source                  1..488
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 286
MGVAENIPDM EEGTLEIGME YRTVSGVAGP LVILDKVKGP KFQEIVNIRL GDGTTRRGQV  60
LEVDGEKAVV QVFEGTSGID NKFTTVQFTG EVLKTPVSLD MLGRIFNGSG KPIDNGPPIL  120
PEAYLDISGS SINPSERTYP EEMIQTGIST IDVMNSIARG QKIPLFSAAG LPHNEIAAQI  180
CRQAGLVKRL EKSDNLLEGG EEDNFAIVFA AMGVNMETAQ FFKRDFEENG SMERVTLFLN  240
LANDPTIERI ITPRIALTTA EYLAYECGKH VLVILTDMSS YADALREVSA AREEVPGRRG  300
YPGYMYTDLA TIYERAGRIE GRKGSITQIP ILTMPNDDIT HPTPDLTGYI TEGQIYIDRQ  360
LYNRQIYPPI NVLPSLSRLM KSAIGEGMTR RDHSDVSNQL YANYAIGKDV QAMKAVVGEE  420
ALSSEDLLYL EFLEKFERKF VAQGAYDTRN IFQSLDLAWT LLRIFPRELL HRIPAKTLDQ  480
FYSRDAGN                                                            488

SEQ ID NO: 287          moltype = AA  length = 26
FEATURE                 Location/Qualifiers
source                  1..26
```

-continued

```
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 287
FQEIVNIRLG DGTTRRGQVL EVDGEK                                          26

SEQ ID NO: 288          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 288
FTTVQFTGEV LK                                                         12

SEQ ID NO: 289          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 289
TLDQFYSRDA GN                                                         12

SEQ ID NO: 290          moltype = DNA  length = 1407
FEATURE                 Location/Qualifiers
source                  1..1407
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 290
atggccaccg ctgccgagaa actctccgct ctcaaatccg ccgtcgccgg attgaacgaa  60
atcagtgaga atgagaagaa cggattcatc agcctcgtcg gccgctatct cagtggcgaa  120
gcgcagcatg tggaatggag caagatccag acgcctacgg acgaagtggt tgtgccttac  180
gacactttgg cgccaactcc tgaaggttct tcggaggtga agaatctatt ggacaagctt  240
gtggtgttga agctaaatgg aggcttggga acaactatgg gttgcactgg tcctaaatct  300
gtaattgaag ttcgtgatgg gttgacattt ctagatttaa ttgtcatcca aattgagaat  360
ctcaattcca aatatggaag caatgttcct ttgcttttga tgaattcatt caacactcat  420
gatgacactc aaaagattgt tgaaaaatac caaaactcaa atattgagat tcatactttt  480
aaccagagcc agtatcctcg attggttgtt gaggactttt tgccattgcc ttccaaaggg  540
catactgaca aggatggatg gtaccctcct ggccatggtg atgtcttccc atcattattg  600
aacagtggca aacttgatgc actattgtca cagggtaaag agtatgtgtt tgttgccaat  660
tcggataact tgggagctat agttgacttg aaaatcttga atcatttgat ccagaacaag  720
aatgaatact gtatggaggt gactcccaaa acattggctg atgtaaaggg tggcactttg  780
atttcttacg aaggaagggt tcagctcctg gaaattgcac aagtccccga tgaacatgtc  840
aatgagttca agtcaataga gaagttcaaa attttcaaca caaataattt gtgggtgaac  900
ttaaacgcag ttaaaaggct tgttgaagct gatgctctta agatggaaat tattcccaat  960
ccaaaggaag ttgacggaat aaaagttctt cagctggaaa ctgcagctgg tgctgcaata  1020
aggttctttg acaaggctat tgggattaat gttcctcgat cacgattcct tcctgtgaag  1080
gcaacttcag atttgcttct tgtccagtct gacctctaca ctttggaaga cggatttgtc  1140
attcggaaca aagctaggga aaatcctgaa aacccttcta ttgaactggg accagaattt  1200
aagaaggtta gcaacttctt gggccgcttc aagtcaattc ctagtatcgt tgagcttgac  1260
agtctaaaag tggctggtga tgtatggttt ggagctggtg ttatcctcaa ggggaaagtc  1320
agcattgtat caaagcccgg tgtaaagctg gaagttcccg acggtgttgc cattgtagac  1380
aaggaaatta atggcccaga ggacctg                                      1407

SEQ ID NO: 291          moltype = AA  length = 469
FEATURE                 Location/Qualifiers
source                  1..469
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 291
MATAAEKLSA LKSAVAGLNE ISENEKNGFI SLVGRYLSGE AQHVEWSKIQ TPTDEVVVPY  60
DTLAPTPEGS SEVKNLLDKL VVLKLNGGLG TTMGCTGPKS VIEVRDGLTF LDLIVIQIEN  120
LNSKYGSNVP LLLMNSFNTH DDTQKIVEKY QNSNIEIHTF NQSQYPRLVV EDFLPLPSKG  180
HTDKDGWYPP GHGDVFPSLL NSGKLDALLS QGKEYVFVAN SDNLGAIVDL KILNHLIQNK  240
NEYCMEVTPK TLADVKGGTL ISYEGRVQLL EIAQVPDEHV NEFKSIEKFK IFNTNNLWVN  300
LNAVKRLVEA DALKMEIIPN PKEVDGIKVL QLETAAGAAI RFFDKAIGIN VPRSRFLPVK  360
ATSDLLLVQS DLYTLEDGFV IRNKARENPE NPSIELGPEF KKVSNFLGRF KSIPSIVELD  420
SLKVAGDVWF GAGVILKGKV SIVSKPGVKL EVPDGVAIVD KEINGPEDL              469

SEQ ID NO: 292          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 292
AIGINVPRSR FLPVK                                                      15

SEQ ID NO: 293          moltype = AA  length = 17
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 293
ARENPENPSI ELGPEFK                                                  17

SEQ ID NO: 294          moltype = AA  length = 24
FEATURE                 Location/Qualifiers
source                  1..24
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 294
ATSDLLLVQS DLYTLEDGFV IRNK                                          24

SEQ ID NO: 295          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 295
EINGPEDL                                                            8

SEQ ID NO: 296          moltype = AA  length = 18
FEATURE                 Location/Qualifiers
source                  1..18
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 296
EYVFVANSDN LGAIVDLK                                                 18

SEQ ID NO: 297          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 297
ILNHLIQNK                                                           9

SEQ ID NO: 298          moltype = AA  length = 26
FEATURE                 Location/Qualifiers
source                  1..26
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 298
IQTPTDEVVV PYDTLAPTPE GSSEVK                                        26

SEQ ID NO: 299          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 299
LDALLSQGK                                                           9

SEQ ID NO: 300          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 300
RLVEADALK                                                           9

SEQ ID NO: 301          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 301
SAVAGLNEIS ENEK                                                     14

SEQ ID NO: 302          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
```

-continued

```
SEQUENCE: 302
SIPSIVELDS LK                                                        12

SEQ ID NO: 303          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 303
VLQLETAAGA AIRFFDK                                                   17

SEQ ID NO: 304          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 304
VSNFLGRFK                                                            9

SEQ ID NO: 305          moltype = DNA  length = 1353
FEATURE                 Location/Qualifiers
source                  1..1353
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 305
atgaccaccg ccaccgagaa gctctccgct ctcaaatccg ccgtcgccgg attgaacgaa  60
atcagtgaga gtgagaagaa cggattcatc agcctcgtca gccgctatct cagtggcgaa  120
gcgcagcatg tggaatggag caagatccag acgcctacgg acgaagtggt tgtgccttac  180
gacactttgg cgccaactcc tgatggttct tcggacgtga agaatctatt ggacaagctt  240
gtggtgttga agctaaatgg aggcttgggc acaactatgg gttgcactgg tcctaaatct  300
gtaattgaag ttcgtgatgg gttgacattt ctagatttaa ttgtgatcca gattgagaat  360
ctcaattcca aatatggaag caatgttcct ttgcttttga tgaattcatt caacactcat  420
gatgacactc aaaagattgt tgaaaaatac caaaactcca atattgagat tcatactttt  480
aaccagggca tactgacaag gatggatggt accctcctgg cccatggaga tgtctttcca  540
tcattattga acagtggcaa acttgatgca ctattgtcac agggtaaaga gtatgtattt  600
gttgccaatt cagataactt gggagctata gttgacttga aaatcttaaa tcatttgatc  660
cagaacaaga atgaatactg tatggaggtg actcccaaaa cattggctga tgtaaagggt  720
ggcactttga tttcttacga aggaaggctt ttggaaattg cacaagtccc agatgaacat  780
gtcaatgagt tcaagtcaat agagaagttc aaaattttca acacaaataa tttgtgggtg  840
aacttaaatg cagttaaaag gcttgttgaa gctgatgctc ttaagatgga aattattccc  900
aatccaaagg aagttgatgg aataaaagtt cttcagctgg aaactgcagc tggtgctgca  960
ataaggttct ttgacaaggc tattgggatt aatgttcctc gatcacgatt ccttcctgtg  1020
aaggcaactt cagatttgct tcttgtccag tctgacctct acactttgga agatggattt  1080
gtcattcgga acaaagctag ggcaaatcct gaaaaccctt ctattgaact tgggccagaa  1140
tttaagaagg ttagcaactt cttgggccgc ttcaagtcaa ttcccagtat tgttgagctt  1200
gacagtctaa aagtggctgg caatgtatgg tttggagctg gtgttatcct caagggaaaa  1260
atcagtatcg tggccaatcc tggtgttaag ctggaagttc ccgatggtgc tgtcatttcg  1320
gataaggaaa ttaatggccc agaggacctc ctg                                1353

SEQ ID NO: 306          moltype = AA  length = 451
FEATURE                 Location/Qualifiers
source                  1..451
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 306
MTTATEKLSA LKSAVAGLNE ISESEKNGFI SLVSRYLSGE AQHVEWSKIQ TPTDEVVVPY  60
DTLAPTPDGS SDVKNLLDKL VVLKLNGGLG TTMGCTGPKS VIEVRDGLTF LDLIVIQIEN  120
LNSKYGSNVP LLLMNSFNTH DDTQKIVEKY QNSNIEIHTF NQGILTRMDG TLLAHGDVFP  180
SLLNSGKLDA LLSQGKEYVF VANSDNLGAI VDLKILNHLI QNKNEYCMEV TPKTLADVKG  240
GTLISYEGRL LEIAQVPDEH VNEFKSIEKF KIFNTNNLWV NLNAVKRLVE ADALKMEIIP  300
NPKEVDGIKV LQLETAAGAA IRFFDKAIGI NVPRSRFLPV KATSDLLLVQ SDLYTLEDGF  360
VIRNKARANP ENPSIELGPE FKKVSNFLGR FKSIPSIVEL DSLKVAGNVW FGAGVILKGK  420
ISIVANPGVK LEVPDGAVIS DKEINGPEDL L                                  451

SEQ ID NO: 307          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 307
AIGINVPRSR FLPVK                                                     15

SEQ ID NO: 308          moltype = AA  length = 24
FEATURE                 Location/Qualifiers
source                  1..24
                        mol_type = protein
```

-continued

```
                        organism = Glycine max
SEQUENCE: 308
ATSDLLLVQS DLYTLEDGFV IRNK                                          24

SEQ ID NO: 309          moltype = AA  length = 18
FEATURE                 Location/Qualifiers
source                  1..18
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 309
EYVFVANSDN LGAIVDLK                                                 18

SEQ ID NO: 310          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 310
ILNHLIQNK                                                           9

SEQ ID NO: 311          moltype = AA  length = 26
FEATURE                 Location/Qualifiers
source                  1..26
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 311
IQTPTDEVVV PYDTLAPTPD GSSDVK                                        26

SEQ ID NO: 312          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 312
LDALLSQGK                                                           9

SEQ ID NO: 313          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 313
LEVPDGAVIS DK                                                       12

SEQ ID NO: 314          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 314
RLVEADALK                                                           9

SEQ ID NO: 315          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 315
SAVAGLNEIS ESEK                                                     14

SEQ ID NO: 316          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 316
SIPSIVELDS LK                                                       12

SEQ ID NO: 317          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 317
VLQLETAAGA AIRFFDK                                                  17
```

-continued

```
SEQ ID NO: 318          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 318
VSNFLGRFK                                                           9

SEQ ID NO: 319          moltype = DNA  length = 1332
FEATURE                 Location/Qualifiers
source                  1..1332
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 319
atggatgagg agtacgatgt gatcgtgttg ggcacgggtc tcaaggagtg cattctcagc  60
ggtcttctct ccgtcgatgg cctcaaggtt ctgcacatgg acaggaatga ctattacgga  120
ggagaatcca cttcgctcaa tcttattcag ctttggaaga gattcagggg agatgacaag  180
cctcctccac atttgggagc tagcagggac tataacatcg atatgaaccc taagtttatt  240
atggctaatg gcaatttggt gcgggttctc atacatactg atgttactaa gtatctctct  300
ttcaaagctg tggatgggag ctttgtgttt aataaaggaa aggttcataa ggtgcccgca  360
aatgacatgg aggccttgaa gtctccactt atggggctat ttgaaaaacg ccgagcacgc  420
aagtttttca tttatgtgca aaattacaat gagagtgatc ccaagaccca tgaagggatg  480
gacttgacaa gagtgactac taaagagttg atagcaaaat acggtcttga tgataacact  540
atagacttca ttggtcatgc tttggcgctt catagtgatg atcgctactt ggctgagcca  600
gcactggaca ctgtgaagag aatgaagtta tatgcagagt ctcttgcacg ctttcaagga  660
gggtcacctt acatatatcc tttgtatggt ttaggagagc ttccccaggc atttgcacgg  720
cttagtgctg tttatggtgg gacatatatg ttgaataagc ctgagtgcaa ggtagaattt  780
gatgaggaag gaaaggttgt tggtgtcaca tctgaagggg aaactgcaaa gtgcaaaaaa  840
gttgtttgtg atccatcata tttgcccggc aaggtgagga aggttggtag ggttgcaaga  900
gcaatagcca taatgagtca cccaatcccc agcactgatg attcccattc ggtgcaaatt  960
attttaccac agaaacagtt gggtcggaag tcagacatgt acctattttg ttgttcatat  1020
actcacaatg ttgctccaaa gggaaaattt attgcatttg tatcaaccga ggcagagaca  1080
gatcagccag caattgaact aaaacctgga attgaccttc taggacctgt tgatgagata  1140
ttttttgata tctatgacag atatgaaccg gtcaatgaac ccactttgga caactgcttt  1200
gtatcaacaa gttatgatgc tactactcac tttgagtcga ccgtggatga tgttctcaac  1260
atgtatacat tgataactgg aaaggttatt gacctgagtg tggatcttag tgctgctagt  1320
gctgcagaag ag                                                       1332

SEQ ID NO: 320          moltype = AA  length = 444
FEATURE                 Location/Qualifiers
source                  1..444
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 320
MDEEYDVIVL GTGLKECILS GLLSVDGLKV LHMDRNDYYG GESTSLNLIQ LWKRFRGDDK  60
PPPHLGASRD YNIDMNPKFI MANGNLVRVL IHTDVTKYLS FKAVDGSFVF NKGKVHKVPA  120
NDMEALKSPL MGLFEKRRAR KFFIYVQNYN ESDPKTHEGM DLTRVTTKEL IAKYGLDDNT  180
IDFIGHALAL HSDDRYLAEP ALDTVKRMKL YAESLARFQG GSPYIYPLYG LGELPQAFAR  240
LSAVYGGTYM LNKPECKVEF DEEGKVVGVT SEGETAKCKK VVCDPSYLPG KVRKVGRVAR  300
AIAIMSHPIP STDDSHSVQI ILPQKQLGRK SDMYLFCCSY THNVAPKGKF IAFVSTEAET  360
DQPAIELKPG IDLLGPVDEI FFDIYDRYEP VNEPTLDNCF VSTSYDATTH FESTVDDVLN  420
MYTLITGKVI DLSVDLSAAS AAEE                                          444

SEQ ID NO: 321          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 321
AVDGSFVFNK                                                          10

SEQ ID NO: 322          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 322
FFIYVQNYNE SDPK                                                     14

SEQ ID NO: 323          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 323
VEFDEEGK                                                            8
```

-continued

```
SEQ ID NO: 324          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 324
VIDLSVDLSA ASAAEE                                                   16

SEQ ID NO: 325          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 325
VVGVTSEGET AK                                                       12

SEQ ID NO: 326          moltype = DNA  length = 1332
FEATURE                 Location/Qualifiers
source                  1..1332
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 326
atggatgagg agtacgatgt gatcgtgttg ggcacgggtc tcaaggagtg cattctcagc  60
ggtcttctct ccgtcgatgg cctcaaggtt ctgcatatgg ataggaacga ctattacgga  120
ggagaatcca cttcgctcaa tcttattcag ctttggaaga gattcagggg agatgacaag  180
cctcctccac atttgggcgc tagcagggac tacaacatcg atatgaaccc caagtttatt  240
atggctaatg gcaatttggt gcgggttctc atacataccg atgttacaaa gtatctctat  300
ttcaaagctg tggatgggag ctatgtgttt aataaaggaa aggttcataa ggtgcctgca  360
aacgacatgg aggccttgaa gtccccactt atggggctat ttgaaaaacg ccgggcacgc  420
aagtttttca tttatgtgca aaactacaac gagagcgatc ccaagaccca cgaggggatg  480
gacttgacaa gagtgactac taaagagttg atagcaaaat acggccttga tgacaacact  540
atagacttca ttggtcatgc tttggcgctt catagcgatg atcgctactt ggctgagcca  600
gcactggaca ccgtgaagag aatgaagtta tatgcagagt ctcttgcacg ctttcaagga  660
gggtcacctt acatatatcc tttgtatggt ttaggagagc ttccccaggc atttgcacgg  720
cttagtgctg tttatggtgg aacatatatg ttgaataagc ctgagtgcaa ggtagaattt  780
gatgaggaag gaaaggttgt tggtgttaca tctgaagggg aaactgcaaa gtgtaaaaaa  840
gttgtttgtg atccatcata tttgcctggc aaggtgagga aggttggtag ggttgcaagg  900
gcgatagcca tcatgagcca cccaatcccc agcactgatg attcccattc agtgcaaatt  960
attttaccac agaaacagtt gggtcggaag tcagacatgt acctattttg ttgttcgtat  1020
actcacaatg ttgctccaaa gggaaaattt attgcatttg tatcaactga ggcagagaca  1080
gatcagccgg caattggact gaaacctgga attgaccttc taggacctgt tgatgagata  1140
ttttttgata tctatgacag atatgaacca gtcaacgaac ccactttgga caactgcttt  1200
atatcaacaa gttatgatgc tactactcac tttgagtcga ctgtggatga cgttctcaac  1260
atgtatacat tgataactgg aaaggttatt gacctcagtg tggatcttag tgctgctagt  1320
gctgcagaag ag                                                      1332

SEQ ID NO: 327          moltype = AA  length = 444
FEATURE                 Location/Qualifiers
source                  1..444
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 327
MDEEYDVIVL GTGLKECILS GLLSVDGLKV LHMDRNDYYG GESTSLNLIQ LWKRFRGDDK  60
PPPHLGASRD YNIDMNPKFI MANGNLVRVL IHTDVTKYLY FKAVDGSYVF NKGKVHKVPA  120
NDMEALKSPL MGLFEKRRAR KFFIYVQNYN ESDPKTHEGM DLTRVTTKEL IAKYGLDDNT  180
IDFIGHALAL HSDDRYLAEP ALDTVKRMKL YAESLARFQG GSPYIYPLYG LGELPQAFAR  240
LSAVYGGTYM LNKPECKVEF DEEGKVVGVT SEGETAKCKK VVCDPSYLPG KVRKVGRVAR  300
AIAIMSHPIP STDDSHSVQI ILPQKQLGRK SDMYLFCCSY THNVAPKGKF IAFVSTEAET  360
DQPAIGLKPG IDLLGPVDEI FFDIYDRYEP VNEPTLDNCF ISTSYDATTH FESTVDDVLN  420
MYTLITGKVI DLSVDLSAAS AAEE                                         444

SEQ ID NO: 328          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 328
AVDGSYVFNK                                                          10

SEQ ID NO: 329          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 329
```

-continued

```
FFIYVQNYNE SDPK                                                    14

SEQ ID NO: 330          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 330
VEFDEEGK                                                           8

SEQ ID NO: 331          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 331
VIDLSVDLSA ASAAEE                                                  16

SEQ ID NO: 332          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 332
VVGVTSEGET AK                                                      12

SEQ ID NO: 333          moltype = DNA  length = 1512
FEATURE                 Location/Qualifiers
source                  1..1512
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 333
atggcgatga gaaccaaact ttctttagct atctttttct tttttctttt agccttattt  60
tcaaacctag cctttggcaa atgtaaagaa accgaggtag aagaagaaga ccctgagctc  120
gtaacatgca aacaccaatg ccaacagcca cagcaataca ctgagggtga caagcgggta  180
tgcttgcaaa gatgtgacag gtatcatcgt atgaagcaag agcgagagaa acaaatccaa  240
gaggaaactc gtgagaagaa agaggaggaa agccgtgaaa gggaagaaga acaacaagag  300
cagcatgagg aacaagatga aaatccctac atttttgagg aagataagga ttttgagacc  360
agagtcgaaa ctgaaggtgg cagaattcgg gttctcaaga agttcactga aaaatccaag  420
cttcttcagg gcattgagaa cttccgtttg gccattttgg aagctagagc acacacgttc  480
gtgtctcctc gccactttga ttccgaggtt gtttttttca acatcaaggg gcgagccgta  540
cttgggttgg tgagtgaaag tgaaacagaa aaaatcaccc ttgaacctgg ggacatgata  600
cacataccag ctggcacccc actgtacatc gttaacagag atgagaatga taagctcttc  660
cttgccatgc tccatatacc tgtctctgtc tctactcctg gaaaatttga ggaatttttc  720
gggcctggag gacgagaccc agaatcggtc ctctcagcat tcagctggaa tgtgctgcaa  780
gctgcgctcc aaaccccaaa aggaaagtta gaaaagcttt ttgatcaaca gaacgaggga  840
agtattttcg caataagcag agaacaggtg cgtgcgttgg cccctaccaa gaaaagctct  900
tggtggccat tcggtggcga atccaagcct caattcaata ttttcagcaa gcgtcccact  960
atctccaacg gatatggtcg tttaactgaa gttggtcctg acgatgatga aaagagttgg  1020
cttcaaagac tcaatctcat gcttaccttt accaacatca cccagagatc tatgagtact  1080
attcactaca actcacatgc aacgaagata gcactggtga tcgatggcag agggcatctt  1140
caaatatcat gtccacacat gtcatcaagg tcgtcacact cgaagcatga taagagtagc  1200
ccttcatacc atagaatcag ttccgatttg aagcctggaa tggtgtttgt tgtccctccc  1260
ggtcatccct tcgtcaccat agcttccaat aaagagaatc tcctcatgat ttgcttcgag  1320
gttaacgctc gagacaacaa gaagtttaca tttgcaggga aggacaacat tgtgagttct  1380
ctggacaacg ttgctaagga gcttgccttt aactatcctt ctgagatggt gaacggagta  1440
ttcgaccgaa aggagagttt ctttttccct tttgagttgc cgagagagga gcgtggtcgt  1500
cgcgctgatg cg                                                      1512

SEQ ID NO: 334          moltype = AA  length = 504
FEATURE                 Location/Qualifiers
source                  1..504
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 334
MAMRTKLSLA IFFFFLLALF SNLAFGKCKE TEVEEEDPEL VTCKHQCQQP QQYTEGDKRV  60
CLQRCDRYHR MKQEREKQIQ EETREKKEEE SREREEEQQE QHEEQDENPY IFEEDKDFET  120
RVETEGGRIR VLKKFTEKSK LLQGIENFRL AILEARAHTF VSPRHFDSEV VFFNIKGRAV  180
LGLVSESETE KITLEPGDMI HIPAGTPLYI VNRDENDKLF LAMLHIPVSV STPGKFEEFF  240
GPGGRDPESV LSAFSWNVLQ AALQTPKGKL EKLFDQQNEG SIFAISREQV RALAPTKKSS  300
WWPFGGESKP QFNIFSKRPT ISNGYGRLTE VGPDDDEKSW LQRLNLMLTF TNITQRSMST  360
IHYNSHATKI ALVIDGRGHL QISCPHMSSR SSHSKHDKSS PSYHRISSDL KPGMVFVVPP  420
GHPFVTIASN KENLLMICFE VNARDNKKFT FAGKDNIVSS LDNVAKELAF NYPSEMVNGV  480
FDRKESFFFP FELPREERGR RADA                                         504

SEQ ID NO: 335          moltype = AA  length = 17
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 335
DFETRVETEG GRIRVLK                                                  17

SEQ ID NO: 336          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 336
DNIVSSLDNV AK                                                       12

SEQ ID NO: 337          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 337
GRAVLGLVSE SETEK                                                    15

SEQ ID NO: 338          moltype = AA  length = 25
FEATURE                 Location/Qualifiers
source                  1..25
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 338
LFDQQNEGSI FAISREQVRA LAPTK                                         25

SEQ ID NO: 339          moltype = AA  length = 21
FEATURE                 Location/Qualifiers
source                  1..21
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 339
RPTISNGYGR LTEVGPDDDE K                                             21

SEQ ID NO: 340          moltype = DNA  length = 1467
FEATURE                 Location/Qualifiers
source                  1..1467
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 340
atggcgacca gagccaagct ttctttagct atcttccttt tctttctttt agccttgatt  60
tcaaacctag ccttgggcaa acttaaagaa accgaggtcg aagaagatcc cgagctcgta  120
acatgcaaac accagtgcca acagcaacgg caatacactg agagtgacaa gcgaacatgc  180
ttgcaacaat gtgacagtat gaagcaagag cgagagaaac aagtcgaaga ggaaactcgc  240
gagaaggaag aagaacatca agagcagcat gaggaggagc aagaccaaaa tccctacgtt  300
tttgaagaag ataaggattt ttcgaccaga gtcgaaacag aaggtggcag cattcgggtt  360
ctcaagaagt tcactgagaa atccaagctt cttcaaggca atgagaattt ccgtttggcc  420
atcttagaag ctagagcaca cacgttcgtg tccccacgcc actttgattc cgaggttgtc  480
ttgttcaaca ttaaggggag agccgtactt gggttggtga gggaaagtga aacagaaaaa  540
atcaccctag aacctggaga catgatacac ataccagcag gcacaccact gtacatcgtt  600
aacagagatg agaatgagaa gctcctcctt gccatgctcc atatacctgt ctctactcct  660
ggaaaatttg aggaattttt cgggcctgga ggacgagacc cagaatcggt cctctcagca  720
ttcagctgga atgtgctgca agctgcgctc caaaccccaa aaggaaagtt agaaaggctt  780
tttaatcaac agaacgaggg aagtattttc aaaataagca gagaacgggt gcgtgcgttg  840
gcccccacca agaaaagctc ttggtggcca ttcggcggcg aatccaaggc tcaattcaat  900
attttcagca agcgtcccac tttctccaac ggatatggcc gtttaactga agttggtcct  960
gatgatgaaa agagttggct tcaaagactc aacctcatgc ttacctttac caacatcacc  1020
cagagatcta tgagtactat tcactacaac tcacatgcaa cgaagatagc actggtgatg  1080
gatggtagag ggcatcttca aatatcatgt ccacacatgt catcaaggtc agactcaaag  1140
catgataaga gtagcccctc ataccataga atcagtgcgg acttgaagcc tggaatggtg  1200
tttgttgtcc ctcctggtca tcccttcgtc actatagctt ccaataaaga gaatctcctc  1260
ataatttgct tcgaggttaa cgttcgagac aacaagaagt ttacgtttgc agggaaggac  1320
aacattgtga gctctctgga caacgtagct aaggagctgg cctttaacta tccttctgag  1380
atggtgaacg gagtcttcga aagaaaggag agtctctttt tccccttcga gttgccgagc  1440
gaggagcgtg gtcgtcgcgc tgttgcg                                      1467

SEQ ID NO: 341          moltype = AA  length = 489
FEATURE                 Location/Qualifiers
source                  1..489
                        mol_type = protein
                        organism = Glycine max
```

-continued

```
SEQUENCE: 341
MATRAKLSLA IFLFFLLALI SNLALGKLKE TEVEEDPELV TCKHQCQQQR QYTESDKRTC  60
LQQCDSMKQE REKQVEEETR EKEEEHQEQH EEEQDQNPYV FEEDKDFSTR VETEGGSIRV  120
LKKFTEKSKL LQGNENFRLA ILEARAHTFV SPRHFDSEVV LFNIKGRAVL GLVRESETEK  180
ITLEPGDMIH IPAGTPLYIV NRDENEKLLL AMLHIPVSTP GKFEEFFGPG GRDPESVLSA  240
FSWNVLQAAL QTPKGKLERL FNQQNEGSIF KISRERVRAL APTKKSSWWP FGGESKAQFN  300
IFSKRPTFSN GYGRLTEVGP DDEKSWLQRL NLMLTFTNIT QRSMSTIHYN SHATKIALVM  360
DGRGHLQISC PHMSSRSDSK HDKSSPSYHR ISADLKPGMV FVVPPGHPFV TIASNKENLL  420
IICFEVNVRD NKKFTFAGKD NIVSSLDNVA KELAFNYPSE MVNGVFERKE SLFFPFELPS  480
EERGRRAVA                                                          489

SEQ ID NO: 342          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 342
AQFNIFSK                                                           8

SEQ ID NO: 343          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 343
DFSTRVETEG GSIRVLK                                                 17

SEQ ID NO: 344          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 344
DNIVSSLDNV AK                                                      12

SEQ ID NO: 345          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 345
GRAVLGLVRE SETEK                                                   15

SEQ ID NO: 346          moltype = AA  length = 20
FEATURE                 Location/Qualifiers
source                  1..20
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 346
RPTFSNGYGR LTEVGPDDEK                                              20

SEQ ID NO: 347          moltype = DNA  length = 1266
FEATURE                 Location/Qualifiers
source                  1..1266
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 347
atggtgagag gattgctcaa caagctcgtc tctcgctccc tctcggtcgc cggaaaatgg  60
cagcacaatc aactccgccg cctcaacatc cacgagtacc agggtgctga attgatgagc  120
aaacacggag tcaacgttcc gagaggcgtg gctgtttcct ctgttgaaga ggccagaaag  180
gtcatcaagg accttttccc taatgaaaat gagttggtgg ttaagagtca aattttagcc  240
ggcggacgag gtttgggaac tttcaaaagc ggccttaagg gaggagtgca cattgttaag  300
actgaccagg ttgaagacat agctgggaag atgcttggac agatactagt taccaaacag  360
actggtcctc agggaaaaat agttagcaag gtttacttgt gtgaaaagct gtcacttgta  420
aatgagatgt actttgctat aacactggat cgtacgagtg ctggtccaat tataattgct  480
tgtagtaagg gaggaaccag cattgaagac cttgcagaga aatttccgga catgattata  540
aaggtacctg ttgatgtttt tgaaggtatt actgatgaag gtgctgcaaa ggttgttgat  600
ggtttggctc tcaaagtggc tgatagaaat aagtcgattg aacaagtgaa gaatttgtat  660
aaactttttg ttgattgtga ctgcactctt ttggaaatca atcccatagc ggagacagct  720
gataaccaat tggtggctgc tgatgctaag ttgaattttg atgataatgc tgcatatcgc  780
cagaaagaga tattttctct ccgtgataca acacaagagg accctcgaga ggtgactgct  840
gcaaaggcag atttaaatta tattgggtta gatggagaaa ttggctgcat ggtgaatgga  900
gcagggttag caatggccac aatggatata attaaattgc atgggggtac tcctgccaat  960
tttctggatg taggtggcaa tgcctctgaa aaccaggttg ttgaggcatt taaaatattg  1020
actgctgatg acaaagtgaa ggcaattttg gttaacattt ttggtggtat aatgaagtgt  1080
gatgttatag caagtggaat agtgaatgct gcaaaagagg ttgcactgaa agtaccagtt  1140
```

-continued

```
gtagttcgtc tggaaggcac caatgttgat caaggaaaaa gaatattgaa ggaaagtggt  1200
atggcattaa taacggctga agatttagat gatgcagctc agaaagcagt gaaagctgct  1260
tacaaa                                                             1266

SEQ ID NO: 348         moltype = AA  length = 422
FEATURE                Location/Qualifiers
source                 1..422
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 348
MVRGLLNKLV SRSLSVAGKW QHNQLRRLNI HEYQGAELMS KHGVNVPRGV AVSSVEEARK  60
VIKDLFPNEN ELVVKSQILA GGRGLGTFKS GLKGGVHIVK TDQVEDIAGK MLGQILVTKQ  120
TGPQGKIVSK VYLCEKLSLV NEMYFAITLD RTSAGPIIIA CSKGGTSIED LAEKFPDMII  180
KVPVDVFEGI TDEGAAKVVD GLALKVADRN KSIEQVKNLY KLFVDCDCTL LEINPIAETA  240
DNQLVAADAK LNFDDNAAYR QKEIFSLRDT TQEDPREVTA AKADLNYIGL DGEIGCMVNG  300
AGLAMATMDI IKLHGGTPAN FLDVGGNASE NQVVEAFKIL TADDKVKAIL VNIFGGIMKC  360
DVIASGIVNA AKEVALKVPV VVRLEGTNVD QGKRILKESG MALITAEDLD DAAQKAVKAA  420
YK                                                                 422

SEQ ID NO: 349         moltype = AA  length = 12
FEATURE                Location/Qualifiers
source                 1..12
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 349
DLFPNENELV VK                                                      12

SEQ ID NO: 350         moltype = AA  length = 20
FEATURE                Location/Qualifiers
source                 1..20
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 350
EIFSLRDTTQ EDPREVTAAK                                              20

SEQ ID NO: 351         moltype = AA  length = 26
FEATURE                Location/Qualifiers
source                 1..26
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 351
LHGGTPANFL DVGGNASENQ VVEAFK                                       26

SEQ ID NO: 352         moltype = AA  length = 12
FEATURE                Location/Qualifiers
source                 1..12
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 352
LNFDDNAAYR QK                                                      12

SEQ ID NO: 353         moltype = AA  length = 14
FEATURE                Location/Qualifiers
source                 1..14
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 353
SQILAGGRGL GTFK                                                    14

SEQ ID NO: 354         moltype = AA  length = 10
FEATURE                Location/Qualifiers
source                 1..10
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 354
TDQVEDIAGK                                                         10

SEQ ID NO: 355         moltype = AA  length = 16
FEATURE                Location/Qualifiers
source                 1..16
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 355
VPVVVRLEGT NVDQGK                                                  16

SEQ ID NO: 356         moltype = AA  length = 8
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 356
VVDGLALK                                                            8

SEQ ID NO: 357          moltype = DNA  length = 1266
FEATURE                 Location/Qualifiers
source                  1..1266
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 357
atggtgaggg gattgctcaa caagctcgtc tctcgctccc tctccgtcgc cggaaaatgg  60
cagcacaacc aactccgccg cctcaacatc cacgagtacc agggcgctga attgatgagc  120
aaacacggag ttaatgttcc gagaggcgtg gctgtttcct ctgttgaaga ggtcagaaag  180
gccatcaagg accttttccc taatgaaaat gagttggtgg ttaagagtca aattttagcc  240
ggtggaagag gtttgggaac tttcaaaagc ggccttaagg gaggagtgca cattgttaag  300
actgaccagg ttgaagacat agctgggaag atgcttgggc agatactagt taccaaacag  360
actggtcctc agggaaaaat agttagcaag gtttacttgt gtgaaaagct gtcactagta  420
aatgagatgt actttgctat aacactggat cgtacgagtg ctggtccaat tataattgct  480
tgtagtaagg gaggaaccag cattgaagac ctagcagaga aatttccgga catgattata  540
aaggtacctg ttgatgtttt tgaaggaatt actgatgaag atgctgcaaa ggttgttgat  600
ggtttggctc tcaaagtggc tgatagaaat aagtcgattg aacaagtgaa gaatttgtat  660
aaactttttg ttgattgtga ctgcaccctt ttggaaatca atcccattgc ggagacagct  720
gataaccaat tggtggctgc tgatgctaag ttgaattttg acgataatgc tgcatatcgc  780
cagaaagaga tatttgctct ccgtgataca acacaagagg accctcgaga ggtgactgct  840
gcaaaggcag acttaaatta tattgggtta gatggagaaa taggctgcat ggtgaatgga  900
gcagggttag caatggccac aatggatata attaaattgc atgggggtac tcctgccaat  960
tttttggatg taggtggcaa tgcctctgaa ggccaggtgg ttgaggcatt taaaatattg  1020
actgctgatg acaaagtgaa ggcaattttg gttaacattt ttgggggtat aatgaagtgt  1080
gatgttatag caagtggaat agtgaatgct gcaaaagagg ttgcactgaa agtacccgta  1140
gtagttcgtc tggaaggcac caatgttgat catggaaaaa gaattttgaa ggaaagtggt  1200
atggcgttga taacggctga agatttagat gatgctgctc agaaagcagt gaaagctgct  1260
tacaaa                                                              1266

SEQ ID NO: 358          moltype = AA  length = 422
FEATURE                 Location/Qualifiers
source                  1..422
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 358
MVRGLLNKLV SRSLSVAGKW QHNQLRRLNI HEYQGAELMS KHGVNVPRGV AVSSVEEVRK  60
AIKDLFPNEN ELVVKSQILA GGRGLGTFKS GLKGGVHIVK TDQVEDIAGK MLGQILVTKQ  120
TGPQGKIVSK VYLCEKLSLV NEMYFAITLD RTSAGPIIIA CSKGGTSIED LAEKFPDMII  180
KVPVDVFEGI TDEDAAKVVD GLALKVADRN KSIEQVKNLY KLFVDCDCTL LEINPIAETA  240
DNQLVAADAK LNFDDNAAYR QKEIFALRDT TQEDPREVTA AKADLNYIGL DGEIGCMVNG  300
AGLAMATMDI IKLHGGTPAN FLDVGGNASE GQVVEAFKIL TADDKVKAIL VNIFGGIMKC  360
DVIASGIVNA AKEVALKVPV VVRLEGTNVD HGKRILKESG MALITAEDLD DAAQKAVKAA  420
YK                                                                  422

SEQ ID NO: 359          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 359
DLFPNENELV VK                                                       12

SEQ ID NO: 360          moltype = AA  length = 20
FEATURE                 Location/Qualifiers
source                  1..20
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 360
EIFALRDTTQ EDPREVTAAK                                               20

SEQ ID NO: 361          moltype = AA  length = 26
FEATURE                 Location/Qualifiers
source                  1..26
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 361
LHGGTPANFL DVGGNASEGQ VVEAFK                                        26

SEQ ID NO: 362          moltype = AA  length = 12
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 362
LNFDDNAAYR QK                                                       12

SEQ ID NO: 363          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 363
SQILAGGRGL GTFK                                                     14

SEQ ID NO: 364          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 364
TDQVEDIAGK                                                          10

SEQ ID NO: 365          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 365
VVDGLALK                                                            8

SEQ ID NO: 366          moltype = DNA  length = 1203
FEATURE                 Location/Qualifiers
source                  1..1203
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 366
atggcgacta agaggagcgt gggaacgttg aaggagggtg atttgaaggg gaagagggta  60
ttcgtgaggg tcgatctgaa cgtgcctttg gatgacaacc ttaacatcac cgatgacact  120
agagtccgtg ctgctgttcc caccatcaag tacttgactg gttatggtgc taaagtgatc  180
ctttctagcc atttgggacg tccaaaaggt gtaacaccta aatacagttt gaagcctctt  240
gtgccaaggc tgtctcaact tcttggaatt gaggttacga tggccaatga ctctattggg  300
gaggaagttg aaaagttggt tacacaactt ccagaaggtg gtgttttgct tctagagaat  360
gtgaggttct acaaggagga agagaagaat gatcctgaat ttgcaaagaa gttggcttct  420
ctagctgatc tctatgtgaa tgatgcattt ggcactgccc acagagctca tgcttctaca  480
gaaggagttg ccaaatactt gaagccctct gttgcaggat tcctaatgca gaaggagctt  540
gattatctag ttggggctgt gtcaaacccc aagagaccat ttgctgctat tgtcggtgga  600
tcgaaggtgt cttccaagat tggagttatt gaatccttgt tggagaaagt taatgttctc  660
ttgcttggtg gaggaatgat ctttaccttt tacaaggctc agggttattc agttgggtca  720
tctcttgtgg aggaagacaa gcttagcctt gcaactacac ttcttgaaaa ggccaaggct  780
aaaggggttt ctttgttgct tccaactgat gttgtcatag cagacaagtt tgctgctgat  840
gctaacagca agactgtgcc ggcatcaagc atcccagatg ggtggatggg gttggatatt  900
ggtcctgatt ccatcaagac atttggtgaa gcattggata cgactcagac cattatttgg  960
aatggaccaa tgggtgtttt tgagtttgat aagtttgcca caggaacaga ggctatagcc  1020
aagaaactgg cagagctgag tgggaaggga gtgacaacca tcattggagg aggtgactcc  1080
gttgccgctg tggagaaggt tggacttgca gacaagatga gccacatctc cactggtggt  1140
ggtgccagct tagagcttct tgagggaaag caactccctg gtgtccttgc tcttgatgat  1200
gct                                                                 1203

SEQ ID NO: 367          moltype = AA  length = 401
FEATURE                 Location/Qualifiers
source                  1..401
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 367
MATKRSVGTL KEGDLKGKRV FVRVDLNVPL DDNLNITDDT RVRAAVPTIK YLTGYGAKVI  60
LSSHLGRPKG VTPKYSLKPL VPRLSQLLGI EVTMANDSIG EEVEKLVTQL PEGGVLLLEN  120
VRFYKEEEKN DPEFAKKLAS LADLYVNDAF GTAHRAHAST EGVAKYLKPS VAGFLMQKEL  180
DYLVGAVSNP KRPFAAIVGG SKVSSKIGVI ESLLEKVNVL LLGGGMIFTF YKAQGYSVGS  240
SLVEEDKLSL ATTLLEKAKA KGVSLLLPTD VVIADKFAAD ANSKTVPASS IPDGWMGLDI  300
GPDSIKTFGE ALDTTQTIIW NGPMGVFEFD KFATGTEAIA KKLAELSGKG VTTIIGGGDS  360
VAAVEKVGLA DKMSHISTGG GASLELLEGK QLPGVLALDD A                      401

SEQ ID NO: 368          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
```

-continued

```
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 368
AQGYSVGSSL VEEDK                                                    15

SEQ ID NO: 369          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 369
ELDYLVGAVS NPK                                                      13

SEQ ID NO: 370          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 370
FAADANSK                                                            8

SEQ ID NO: 371          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 371
FATGTEAIAK                                                          10

SEQ ID NO: 372          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 372
GVSLLLPTDV VIADK                                                    15

SEQ ID NO: 373          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 373
GVTTIIGGGD SVAAVEK                                                  17

SEQ ID NO: 374          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 374
IGVIESLLEK                                                          10

SEQ ID NO: 375          moltype = AA  length = 28
FEATURE                 Location/Qualifiers
source                  1..28
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 375
LASLADLYVN DAFGTAHRAH ASTEGVAK                                      28

SEQ ID NO: 376          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 376
LSLATTLLEK                                                          10

SEQ ID NO: 377          moltype = AA  length = 20
FEATURE                 Location/Qualifiers
source                  1..20
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 377
LVTQLPEGGV LLLENVRFYK                                               20
```

-continued

```
SEQ ID NO: 378          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 378
NDPEFAK                                                          7

SEQ ID NO: 379          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 379
QLPGVLALDD A                                                     11

SEQ ID NO: 380          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 380
RPFAAIVGGS K                                                     11

SEQ ID NO: 381          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 381
VILSSHLGRP K                                                     11

SEQ ID NO: 382          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 382
YLTGYGAK                                                         8

SEQ ID NO: 383          moltype = DNA  length = 1455
FEATURE                 Location/Qualifiers
source                  1..1455
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 383
atggcctcag ccgcatcccc caccactttc tctctcctcc actccgccac cccttcctcc  60
tcctccaccg ccacctccgc ctccccctcc ccctcgcgcg tgcagctcac ctcctccctc  120
cgcgcccgcc cgctgcgccg cctgggcttc gccgcggcgg acccgctgct cgcggcgagt  180
gtagcggcgc gcgtggcggc ggtgagagga ggaggaaagg gcgcgagggg ggtggtgtcg  240
atggcgaaga agagcgtggg ggacctgagc gcggcggacc tgaaggggaa gaaggtcttc  300
gtcagggccg acctcaacgt cccgctcgac gacaaccaga acatcaccga cgacaccaga  360
atccgcgccg ccattcccac catcaaacat ctcatccaaa acggcgccaa agtcattctc  420
tccagccact tggggcggcc gaagggtgtc actcccaaat acagcttggc gcctcttgta  480
ccccggcttt ctgagctcat tggtatccag attgtcaagg ccgaggacag tattggtcca  540
gaagtagaaa agttggtggc ttctcttcca gatggaggtg ttctccttct agaaaatgtg  600
aggttttaca aggaggaaga aaagaatgac cccgagcatg caaaaaagct tgcctctctg  660
gcagatctgt atgtgaatga tgcatttggt actgctcaca gggcacatgc atcaacagag  720
ggtgttacta aatacctgaa gccatctgtt gctggttttc ttttgcaaaa ggaacttgat  780
taccttgttg gggcagtatc aaaccccaaa aggccatttg ctgccattgt tggtggttcc  840
aaagtctcat ctaaaattgg agtgattgag tcacttttag aaaaagttga catccttctt  900
cttggtggag gaatgatctt cacattttac aaggcacaag gtctttcagt gggttcatcc  960
cttgtagaag aagataagtt ggatcttgct acatcactac ttgcaaaagc caaggaaaag  1020
ggggtgtctc tcttgttacc aagtgatgtg gtgattgcag acaaatttgc tcctgatgct  1080
aatagcaagg tcgtgccagc atctgccatc cctgatggct ggatgggatt ggatattggt  1140
ccagattctg ttaaatcatt cagtgaagca ttggatacta cccaaaccat catatggaat  1200
ggaccaatgg gagtgtttga gtttgacaag tttgctgttg gtacagaggc tattgcaaag  1260
aagctggctg acctcagtgg aaagggggtg accacaatta ttggaggagg agattctgtt  1320
gcagctgtgg agaaagtagg agttgctagt gtcatgagcc acatatctac tggtggtggt  1380
gccagtttgg agttattgga aggcaaagag cttccaggag tccttgctct tgatgaagcc  1440
gtcccagtcg ctgtg                                                  1455

SEQ ID NO: 384          moltype = AA  length = 485
FEATURE                 Location/Qualifiers
source                  1..485
```

-continued

```
                      mol_type = protein
                      organism = Glycine max
SEQUENCE: 384
MASAASPTTF SLLHSATPSS SSTATSASPS PSRVQLTSSL RARPLRRLGF AAADPLLAAS  60
VAARVAAVRG GGKGARGVVS MAKKSVGDLS AADLKGKKVF VRADLNVPLD DNQNITDDTR  120
IRAAIPTIKH LIQNGAKVIL SSHLGRPKGV TPKYSLAPLV PRLSELIGIQ IVKAEDSIGP  180
EVEKLVASLP DGGVLLLENV RFYKEEEKND PEHAKKLASL ADLYVNDAFG TAHRAHASTE  240
GVTKYLKPSV AGFLLQKELD YLVGAVSNPK RPFAAIVGGS KVSSKIGVIE SLLEKVDILL  300
LGGGMIFTFY KAQGLSVGSS LVEEDKLDLA TSLLAKAKEK GVSLLLPSDV VIADKFAPDA  360
NSKVVPASAI PDGWMGLDIG PDSVKSFSEA LDTTQTIIWN GPMGVFEFDK FAVGTEAIAK  420
KLADLSGKGV TTIIGGGDSV AAVEKVGVAS VMSHISTGGG ASLELLEGKE LPGVLALDEA  480
VPVAV                                                              485

SEQ ID NO: 385         moltype = AA  length = 13
FEATURE                Location/Qualifiers
source                 1..13
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 385
ELDYLVGAVS NPK                                                     13

SEQ ID NO: 386         moltype = AA  length = 16
FEATURE                Location/Qualifiers
source                 1..16
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 386
ELPGVLALDE AVPVAV                                                  16

SEQ ID NO: 387         moltype = AA  length = 10
FEATURE                Location/Qualifiers
source                 1..10
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 387
FAVGTEAIAK                                                         10

SEQ ID NO: 388         moltype = AA  length = 15
FEATURE                Location/Qualifiers
source                 1..15
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 388
GVSLLLPSDV VIADK                                                   15

SEQ ID NO: 389         moltype = AA  length = 17
FEATURE                Location/Qualifiers
source                 1..17
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 389
GVTTIIGGGD SVAAVEK                                                 17

SEQ ID NO: 390         moltype = AA  length = 10
FEATURE                Location/Qualifiers
source                 1..10
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 390
IGVIESLLEK                                                         10

SEQ ID NO: 391         moltype = AA  length = 11
FEATURE                Location/Qualifiers
source                 1..11
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 391
RPFAAIVGGS K                                                       11

SEQ ID NO: 392         moltype = AA  length = 11
FEATURE                Location/Qualifiers
source                 1..11
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 392
VILSSHLGRP K                                                       11
```

-continued

```
SEQ ID NO: 393          moltype = DNA  length = 1203
FEATURE                 Location/Qualifiers
source                  1..1203
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 393
atggcgacta agaggagcgt gggaacgttg aaggagggtg atttgaaagg gaagagggtg  60
ttcgtgaggg tcgatctgaa cgtgcctttg gatgacaacc ttaacatcac cgatgacact  120
agagtccgtg ctgctgttcc caccatcaag tacttaactg gtcatggagc caaagtgatc  180
ctttctagcc acttgggacg tcctaaaggt gtaacaccta aatacagttt gaagcctctt  240
gtgccaaggc tgtctcaact tctcggaatt gaggttacga tggcaaatga ctctattggg  300
gaggaagttg agaagttggt tgcacaactt ccagaaggtg gtgttttgct tctagagaat  360
gtgaggttct ataaggagga agagaagaat gaccctgagt ttgcaaagaa gttggcttct  420
cttgctgatc tctacgtgaa tgatgctttc ggcactgccc acagagctca tgcttctaca  480
gaaggagtgg ccaaatactt gaagccctct gttgcaggat tccttatgca gaaggaactt  540
gactatctag ttggagctgt gtcaaacccc aagagaccat ttgctgctat cgttggtggg  600
tcaaaggtgt cttccaagat tggagttatt gagtccttgt tggagaaagt taacgttctc  660
ttgcttggtg gaggaatgat ctttaccttt tacaaggctc agggttattc agttggttca  720
tctcttgtgg aggaagacaa gcttagcctt gcaactacac ttctcgaaaa ggccaaggct  780
aaaggggttt ctttgttgct tccaactgat gtggtcatag cagacaagtt tgctgctgat  840
gctaacagca agactgtgcc agcatcaagc atcccagacg ggtggatggg attagatatt  900
ggtcccgatt ccatcaagac atttggtgaa gcattggata agacccagac cgtcatttgg  960
aatggaccaa tgggtgtttt tgagtttgat aagtttgcca caggaacaga ggccatagcc  1020
aagaaactgg cagagctgag tgggaaggga gtaacgacca tcattggagg aggtgactct  1080
gttgccgctg tggagaaggt tggtcttgca gacaagatga gccacatctc cactggtggt  1140
ggtgccagct tagagcttct tgaggggaag caactccctg gtgtccttgc ccttgatgat  1200
gct                                                                1203

SEQ ID NO: 394          moltype = AA  length = 401
FEATURE                 Location/Qualifiers
source                  1..401
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 394
MATKRSVGTL KEGDLKGKRV FVRVDLNVPL DDNLNITDDT RVRAAVPTIK YLTGHGAKVI  60
LSSHLGRPKG VTPKYSLKPL VPRLSQLLGI EVTMANDSIG EEVEKLVAQL PEGGVLLLEN  120
VRFYKEEEKN DPEFAKKLAS LADLYVNDAF GTAHRAHAST EGVAKYLKPS VAGFLMQKEL  180
DYLVGAVSNP KRPFAAIVGG SKVSSKIGVI ESLLEKVNVL LLGGGMIFTF YKAQGYSVGS  240
SLVEEDKLSL ATTLLEKAKA KGVSLLLPTD VVIADKFAAD ANSKTVPASS IPDGWMGLDI  300
GPDSIKTFGE ALDKTQTVIW NGPMGVFEFD KFATGTEAIA KKLAELSGKG VTTIIGGGDS  360
VAAVEKVGLA DKMSHISTGG GASLELLEGK QLPGVLALDD A                      401

SEQ ID NO: 395          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 395
AQGYSVGSSL VEEDK                                                   15

SEQ ID NO: 396          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 396
ELDYLVGAVS NPK                                                     13

SEQ ID NO: 397          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 397
FAADANSK                                                           8

SEQ ID NO: 398          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 398
FATGTEAIAK                                                         10

SEQ ID NO: 399          moltype = AA  length = 15
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 399
GVSLLLPTDV VIADK                                                    15

SEQ ID NO: 400          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 400
GVTTIIGGGD SVAAVEK                                                  17

SEQ ID NO: 401          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 401
IGVIESLLEK                                                          10

SEQ ID NO: 402          moltype = AA  length = 28
FEATURE                 Location/Qualifiers
source                  1..28
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 402
LASLADLYVN DAFGTAHRAH ASTEGVAK                                      28

SEQ ID NO: 403          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 403
LSLATTLLEK                                                          10

SEQ ID NO: 404          moltype = AA  length = 20
FEATURE                 Location/Qualifiers
source                  1..20
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 404
LVAQLPEGGV LLLENVRFYK                                               20

SEQ ID NO: 405          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 405
NDPEFAK                                                             7

SEQ ID NO: 406          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 406
QLPGVLALDD A                                                        11

SEQ ID NO: 407          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 407
RPFAAIVGGS K                                                        11

SEQ ID NO: 408          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
```

-continued

```
SEQUENCE: 408
TFGEALDK                                                            8

SEQ ID NO: 409          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 409
VILSSHLGRP K                                                        11

SEQ ID NO: 410          moltype = DNA  length = 1143
FEATURE                 Location/Qualifiers
source                  1..1143
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 410
atgtcgtcga gcacagttgg tcaagtcatc aaatgcaaag ctgcggtttc atgggaggca  60
gggaagccac ttgtgattga agaagtagag gtggcaccac cacaagctgg tgaagtgcgt  120
ttgaagatcc tctacacctc tctttgccac actgatgttt acttctggga agccaagggc  180
caaactccat tgtttcctcg tatatttggt catgaagctg gagggattgt ggagagcgta  240
ggtgagggtg tgactcatct caaaccaggg gaccatgccc tcccagtgtt taccggggag  300
tgtggagaat gccctcattg caagtcagag gagagcaaca tgtgtgatct tctcaggatc  360
aacactgata ggggtgtcat gatccatgac agccaaacaa gattttctat taagggacaa  420
cctatttacc attttgttgg tacctccaca ttcagtgaat acactgttgt tcatgccgga  480
tgtgttgcca aagtcaaccc tgctgcccca cttgacaaaa tttgtgttct cagttgtgga  540
atttgcacag gtcttggggc tactatcaat gttgcaaaac caaaacctgg ttcttctgtt  600
gccatttttg gacttggagc tgttggcctt gctgctgctg aaggggcaag gatttctggt  660
gcatcaagaa ttattggggt tgatttagtt tctagcagat ttgaagaagc taagaagttt  720
ggggtcaatg aatttgtgaa cccaaaagac catgacaagc ctgtacaaga ggtgattgct  780
gcaatgacga atgggggtgt ggatcgtgct gttgaatgta ctggcagcat ccaagccatg  840
atctcggcat tcgaatgcgt ccacgatggt tggggtgttg ctgtacttgt tggagtgcca  900
aacaaagatg atgcattcaa aactcatcca gtgaatttct tgaacgagag aactctaaag  960
ggtaccttct acggaaacta caaaccccgc actgatcttc cttctgttgt ggagaagtac  1020
atgaatgggg aattggaact ggagaaattc atcactcaca cagttccatt ctcagagatt  1080
aacaaggctt ttgattacat gctgaaaggg gagtccatca gatgcatcat tcgaatgggg  1140
gag                                                                 1143

SEQ ID NO: 411          moltype = AA  length = 381
FEATURE                 Location/Qualifiers
source                  1..381
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 411
MSSSTVGQVI KCKAAVSWEA GKPLVIEEVE VAPPQAGEVR LKILYTSLCH TDVYFWEAKG  60
QTPLFPRIFG HEAGGIVESV GEGVTHLKPG DHALPVFTGE CGECPHCKSE ESNMCDLLRI  120
NTDRGVMIHD SQTRFSIKGQ PIYHFVGTST FSEYTVVHAG CVAKVNPAAP LDKICVLSCG  180
ICTGLGATIN VAKPKPGSSV AIFGLGAVGL AAAEGARISG ASRIIGVDLV SSRFEEAKKF  240
GVNEFVNPKD HDKPVQEVIA AMTNGGVDRA VECTGSIQAM ISAFECVHDG WGVAVLVGVP  300
NKDDAFKTHP VNFLNERTLK GTFYGNYKPR TDLPSVVEKY MNGELELEKF ITHTVPFSEI  360
NKAFDYMLKG ESIRCIIRMG E                                             381

SEQ ID NO: 412          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 412
FGVNEFVNPK                                                          10

SEQ ID NO: 413          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 413
FITHTVPFSE INK                                                      13

SEQ ID NO: 414          moltype = AA  length = 19
FEATURE                 Location/Qualifiers
source                  1..19
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 414
GTFYGNYKPR TDLPSVVEK                                                19
```

-continued

```
SEQ ID NO: 415          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 415
THPVNFLNER TLK                                                  13

SEQ ID NO: 416          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 416
VNPAAPLDK                                                       9

SEQ ID NO: 417          moltype = DNA  length = 1113
FEATURE                 Location/Qualifiers
source                  1..1113
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 417
atgcgtgcag ctgcggtggc atgggaagca gggaagccac tgtcaataga aaccatagaa  60
gtggcaccac cccagaaagg tgaagtgcgt ttgaggatcc tcttcaactc cctttgccgt  120
agtgatgttt attggtggga tgctaaggac cagactcctc tgtttcctcg tatcttaggc  180
cacgaagctt cagggattgt ggagagcgtg ggcgagggtg tgacacatct gaaaccaggt  240
gaccatgcac tcccaatatt cactggggag tgtggtgaat gcacttattg taagtctgaa  300
gagagcaacc tgtgtgagct actcaggatc aacactgaca ggggtgttat gctgagtgat  360
ggaaaaacaa ggttctccaa aaatggccaa cccatttacc acttcgttgg aacctccact  420
ttcagtgaat acactgttct ccatgaagga tgtgtagcca agattaaccc taatgcccca  480
cttgacaaag ttgctattgt cagctgtgga ttctgcacag gttttggggc tactgtgaat  540
gttgcaaaac caaagcctaa taatactgtt gctgtctttg gcttaggagc tgttggtctt  600
gctgcctgtg aaggggcaag ggtttctggt gcatctagaa tcatcggggt tgatttactt  660
cccaatcgat ttgaacaagc caaaaaattt ggggtcactg attttgtaaa cccgaaagat  720
cacaacaaac cagtccaaga ggtcattgct gaaatgacta atggaggagt agatcgtgct  780
attgaatgca ctggaagcat ccaagcttca atctcagcat ttgaatgcac tcatgatggt  840
tggggtactg ctgtccttgt tggtgtgcca aagaaggatg ttgaattcaa aactaatcct  900
atgaagttca tggaagggag aactcttaag ggtaccttct atggtcacta cagaccccgc  960
actgacattc ctggtgttgt ggagaagtac ctcaacaagg agttggaact agacaaattt  1020
atcactcact ctgtgccatt ctcaaagatc aacactgcat ttgatcttat gctgaaaggg  1080
gagggcataa ggtgcctcat ctgcatggaa gag                               1113

SEQ ID NO: 418          moltype = AA  length = 371
FEATURE                 Location/Qualifiers
source                  1..371
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 418
MRAAAVAWEA GKPLSIETIE VAPPQKGEVR LRILFNSLCR SDVYWWDAKD QTPLFPRILG  60
HEASGIVESV GEGVTHLKPG DHALPIFTGE CGECTYCKSE ESNLCELLRI NTDRGVMLSD  120
GKTRFSKNGQ PIYHFVGTST FSEYTVLHEG CVAKINPNAP LDKVAIVSCG FCTGFGATVN  180
VAKPKPNNTV AVFGLGAVGL AACEGARVSG ASRIIGVDLL PNRFEQAKKF GVTDFVNPKD  240
HNKPVQEVIA EMTNGGVDRA IECTGSIQAS ISAFECTHDG WGTAVLVGVP KKDVEFKTNP  300
MKFMEGRTLK GTFYGHYRPR TDIPGVVEKY LNKELELDKF ITHSVPFSKI NTAFDLMLKG  360
EGIRCLICME E                                                       371

SEQ ID NO: 419          moltype = AA  length = 19
FEATURE                 Location/Qualifiers
source                  1..19
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 419
GTFYGHYRPR TDIPGVVEK                                            19

SEQ ID NO: 420          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 420
INPNAPLDK                                                       9

SEQ ID NO: 421          moltype = DNA  length = 1140
FEATURE                 Location/Qualifiers
source                  1..1140
                        mol_type = other DNA
```

-continued

```
                        organism = Glycine max
SEQUENCE: 421
atgtctagca ccgttggcca gaccatcaag tgcaaagctg cgattgcatg ggaggctggg  60
aagccgctgg tgattgaaga agtggaggtt gcgccaccgc aagccggtga agtccgtttg  120
aagatcctct acacctctct ttgccacact gatgtttact tctgggatgc gaagggccag  180
actccattgt ttcctcgcat ttttggccat gaagcttcgg ggattgtgga gagcgtaggc  240
gagggtgtga ctcatctgaa accaggtgac catgccctcc ctgtgttcac aggagagtgt  300
ggagattgcg cccattgcaa gtcagaggag agcaacatgt gtgagctact caggatcaac  360
accgataggg gtgtcatgat ccatgatggc caatcaaggt tctctaagaa tggacaaccc  420
atacaccatt tcttgggaac ctctacattc agtgaataca ctgttgtcca tgctggatgt  480
gttgcaaaga tcaaccctgc tgctccactt gacaaagttt gtgttctcag ttgtggaatt  540
tgcacaggtt ttggtgctac tgtaaatgta gcaaaaccaa aacctggttc ctctgttgcc  600
atatttggac ttggagctgt tggccttgcg gctgctgaag gtgcaagagt ttcaggtgct  660
tcaagaatca ttggagttga tttagtttct gcccgatttg aagaagctaa gaagtttggg  720
gttaatgagt ttgtgaaccc aaaggatcat gataaacctg tgcaacaggt aattgctgaa  780
atgaccaatg gaggtgtgga tcgggctgtt gaatgtactg gcagcatcca agccatggtc  840
tcagcattcg aatgcgtcca cgatggttgg ggtcttgctg tacttgttgg tgtgcctagt  900
aaagatgatg cattcaaaac tgctcctatt aatttcctga acgagaggac tcttaagggc  960
accttttatg gcaactacaa accacgcacc gatcttccat ctgttgtcga gaagtacatg  1020
agtggggagc tagaagtgga caaattcatc actcacacag ttccattctc agagatcaac  1080
aaagcttttg atttaatgct gaagggacag tccattaggt gtatcatccg catgcaagag  1140

SEQ ID NO: 422          moltype = AA  length = 380
FEATURE                 Location/Qualifiers
source                  1..380
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 422
MSSTVGQTIK CKAAIAWEAG KPLVIEEVEV APPQAGEVRL KILYTSLCHT DVYFWDAKGQ  60
TPLFPRIFGH EASGIVESVG EGVTHLKPGD HALPVFTGEC GDCAHCKSEE SNMCELLRIN  120
TDRGVMIHDG QSRFSKNGQP IHHFLGTSTF SEYTVVHAGC VAKINPAAPL DKVCVLSCGI  180
CTGFGATVNV AKPKPGSSVA IFGLGAVGLA AAEGARVSGA SRIIGVDLVS ARFEEAKKFG  240
VNEFVNPKDH DKPVQQVIAE MTNGGVDRAV ECTGSIQAMV SAFECVHDGW GLAVLVGVPS  300
KDDAFKTAPI NFLNERTLKG TFYGNYKPRT DLPSVVEKYM SGELEVDKFI THTVPFSEIN  360
KAFDLMLKGQ SIRCIIRMQE                                              380

SEQ ID NO: 423          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 423
FGVNEFVNPK                                                         10

SEQ ID NO: 424          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 424
FITHTVPFSE INK                                                     13

SEQ ID NO: 425          moltype = AA  length = 19
FEATURE                 Location/Qualifiers
source                  1..19
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 425
GTFYGNYKPR TDLPSVVEK                                               19

SEQ ID NO: 426          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 426
INPAAPLDK                                                          9

SEQ ID NO: 427          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 427
TAPINFLNER TLK                                                     13
```

-continued

```
SEQ ID NO: 428          moltype = DNA  length = 927
FEATURE                 Location/Qualifiers
source                  1..927
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 428
atggcgtcaa gacagcaatt caaggaagac agagccgagg cagctgcaaa actcgcagcg  60
aaagacatcg gcgacatcaa cagaacaaac gagcgcgacg aggggtacga cttgcacgac  120
gaagccaact tcaagcaagc acgagctgaa gcagctgcaa agctggcggc caaggatctg  180
gaggacgcta acagattaag agacaacact ttcaacacat ggaaagagca acatcagcat  240
gatgagaaca aacccggtgt gatagggtcc atgtttagag ccgccaagga agccgtggtt  300
gggaagcctc accacgagga agtttatgca aaggaaacaa agatgggtga atattcagat  360
tatgcgactc agaaggcgag ggaaacgaag cagaaggctg gggagtatac agactatgcg  420
gctcagaaag caaaagaaac caaagactat gctgctgaga aggctaagaa tgctaaagat  480
accactgtgc agaaagctgg agagtacaag gattataccg ctgagaaggc taagaatgct  540
aaagatacca ctgtgcagaa agctggagaa tacaaggatt ataccgctga gaaggcaatt  600
gaagggaagg acagtgcagt ggggaagctt ggggagctaa aggaatctgc tgctgatgct  660
gctaaaaggg ctatgggttt cttcaccggt ggcaacaaag accaaactag agaggaggag  720
gaggaaacga ggcggagtat gcaggaggtg agggtgcaag ataaggagta tggaacggga  780
cgtggtactg gagagaagtt ggtgataaaa atggaagagt cgcggccagg agcggtggcg  840
gatgcgctca aagcagcgaa tagggccatg gacggcgata tggaggagga aggcgtgctt  900
catgtggagc gtcgtaggga gaaaatg                                      927

SEQ ID NO: 429          moltype = AA  length = 309
FEATURE                 Location/Qualifiers
source                  1..309
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 429
MASRQQFKED RAEAAAKLAA KDIGDINRTN ERDEGYDLHD EANFKQARAE AAAKLAAKDL  60
EDANRLRDNT FNTWKEQHQH DENKPGVIGS MFRAAKEAVV GKPHHEEVYA KETKMGEYSD  120
YATQKARETK QKAGEYTDYA AQKAKETKDY AAEKAKNAKD TTVQKAGEYK DYTAEKAKNA  180
KDTTVQKAGE YKDYTAEKAI EGKDSAVGKL GELKESAADA AKRAMGFFTG GNKDQTREEE  240
EETRRSMQEV RVQDKEYGTG RGTGEKLVIK MEESRPGAVA DALKAANRAM DGDMEEEGVL  300
HVERRREKM                                                          309

SEQ ID NO: 430          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 430
AGEYTDYAAQ K                                                       11

SEQ ID NO: 431          moltype = AA  length = 24
FEATURE                 Location/Qualifiers
source                  1..24
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 431
DIGDINRTNE RDEGYDLHDE ANFK                                         24

SEQ ID NO: 432          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 432
DLEDANRLRD NTFNTWK                                                 17

SEQ ID NO: 433          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 433
EAVVGKPHHE EVYAK                                                   15

SEQ ID NO: 434          moltype = DNA  length = 1104
FEATURE                 Location/Qualifiers
source                  1..1104
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 434
atgggcaagg tcaagatcgg aatcaacgga tttggaagaa ttggccgttt ggtagccaga  60
gtggctctgc agagagacga tgttgaactc gttgccgtta acgacccttt catcaccacc  120
```

-continued

```
gattacatga catacatgtt taaatacgac agtgttcatg gacactggaa gcatcacgat  180
gtcaccgtta aggacgagaa gacccttctc ttcggtgaca agccagtcac tatttttgga  240
cacaggtttt tatttttcac tttaattttt cttactgttt atgattgttg tgtcctgcat  300
gtaaacgttg agtgttgttt tctctgtgaa ctcagaaacc ctgaagagat cccatggggg  360
tcaactggag ctgacatcat tgttgagtcc accggagttt tcaccgataa ggacaaggcc  420
gccgcacatt tgaagggtgg tgcaaagaag gttattattt ctgcccccag taaggatgcc  480
cccatgtttg ttgttggtgt caacgagcac gagtacaagc cagagcttga tattatttcc  540
aatgctagct gcacaaccaa ctgccttgcc ccacttgcca aggttatcaa tgacaggttt  600
ggcattgttg agggtttgat gaccactgtt cattccatca ccgctaccca gaagactgtt  660
gatggaccat cagccaagga ctggagaggt ggaagagctg cttcatttaa catcattcct  720
agcagcactg gagctgccaa ggctgttggg aaagtcctcc ctgctttgaa tggaaaattg  780
actggtatgg cattccgtgt tcccaccgtg gatgtctctg ttgttgacct cacagtgagg  840
ctggagaaag aagcttccta cgatgaaatt aaaaatgcta tcaaggagga atcagagggc  900
aagttgaagg gaattcttgg ttacactgaa gatgatgtgg tctccactga ctttatcggc  960
gatagcagat caagtatttt tgatgcaaag gctggaattg cattgaataa gaactttgtg  1020
aagcttgttt cttggtacga caacgagtgg ggatacagct cacgtgtcat tgatcttctt  1080
gtattcgttg ccaagaagtc tctt                                         1104

SEQ ID NO: 435          moltype = AA  length = 368
FEATURE                 Location/Qualifiers
source                  1..368
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 435
MGKVKIGING FGRIGRLVAR VALQRDDVEL VAVNDPFITT DYMTYMFKYD SVHGHWKHHD  60
VTVKDEKTLL FGDKPVTIFG HRFLFFTLIF LTVYDCCVLH VNVECCFLCE LRNPEEIPWG  120
STGADIIVES TGVFTDKDKA AAHLKGGAKK VIISAPSKDA PMFVVGVNEH EYKPELDIIS  180
NASCTTNCLA PLAKVINDRF GIVEGLMTTV HSITATQKTV DGPSAKDWRG GRAASFNIIP  240
SSTGAAKAVG KVLPALNGKL TGMAFRVPTV DVSVVDLTVR LEKEASYDEI KNAIKEESEG  300
KLKGILGYTE DDVVSTDFIG DSRSSIFDAK AGIALNKNFV KLVSWYDNEW GYSSRVIDLL  360
VFVAKKSL                                                           368

SEQ ID NO: 436          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 436
EASYDEIK                                                           8

SEQ ID NO: 437          moltype = AA  length = 27
FEATURE                 Location/Qualifiers
source                  1..27
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 437
GILGYTEDDV VSTDFIGDSR SSIFDAK                                      27

SEQ ID NO: 438          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 438
VIISAPSK                                                           8

SEQ ID NO: 439          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 439
VLPALNGK                                                           8

SEQ ID NO: 440          moltype = DNA  length = 1014
FEATURE                 Location/Qualifiers
source                  1..1014
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 440
atgggcaagg tcaagatcgg aatcaacgga tttggaagaa ttggccgttt ggttgccaga  60
gtcgctctgc aaagagacga tgtggagctc gttgccgtta acgacccttt cattaccacc  120
gattacatga catacatgtt taaatacgac agtgttcacg gacactggaa gcatcacgat  180
gtcaccgtta aggacgagaa gacccttctc ttcggtgaca agccagtcac tgtttttgga  240
cacaggaacc ctgaagagat cccatggggg tcaactggag ctgatatcat tgttgagtcc  300
accggagttt tcaccgataa ggacaaggcc gccgcacatt tgaagggtgg tgcgaagaag  360
```

-continued

```
gttattattt ctgcccccag taaggatgcc cccatgtttg ttgttggtgt taacgagcac  420
gagtacaagc cagagcttga tattatttcc aatgctagct gcacaaccaa ctgccttgct  480
ccacttgcta aggttattaa tgacaggttt ggcattgttg agggtttgat gaccactgtt  540
cattccatca ctgccaccca gaagactgtt gatgggccat cagccaaaga ctggagaggt  600
ggaagagctg cttcatttaa catcattcct agcagcactg gagctgccaa ggctgttgga  660
aaagtccttc ctgctttgaa cggaaaattg actggtatgg cattccgtgt tcccaccgtg  720
gatgtctctg ttgttgacct cacagtgagg ctggagaagg aagcctccta cgatgaaatt  780
aaaaatgcta tcaaggagga atcagagggc aagttgaagg gaattcttgg ttacactgaa  840
gatgatgtgg tctccaccga ctttgtgggc gatagcagat caagtatttt tgatgcaaag  900
gctggaattg cattgaataa gaactttgtg aagcttgtct cttggtatga caacgagtgg  960
ggatacagct cacgtgtcat tgacctgctt gtattcgttg ccaagaagtc tttt        1014

SEQ ID NO: 441          moltype = AA  length = 338
FEATURE                 Location/Qualifiers
source                  1..338
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 441
MGKVKIGING FGRIGRLVAR VALQRDDVEL VAVNDPFITT DYMTYMFKYD SVHGHWKHHD  60
VTVKDEKTLL FGDKPVTVFG HRNPEEIPWG STGADIIVES TGVFTDKDKA AAHLKGGAKK  120
VIISAPSKDA PMFVVGVNEH EYKPELDIIS NASCTTNCLA PLAKVINDRF GIVEGLMTTV  180
HSITATQKTV DGPSAKDWRG GRAASFNIIP SSTGAAKAVG KVLPALNGKL TGMAFRVPTV  240
DVSVVDLTVR LEKEASYDEI KNAIKEESEG KLKGILGYTE DDVVSTDFVG DSRSSIFDAK  300
AGIALNKNFV KLVSWYDNEW GYSSRVIDLL VFVAKKSF                          338

SEQ ID NO: 442          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 442
EASYDEIK                                                           8

SEQ ID NO: 443          moltype = AA  length = 27
FEATURE                 Location/Qualifiers
source                  1..27
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 443
GILGYTEDDV VSTDFVGDSR SSIFDAK                                      27

SEQ ID NO: 444          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 444
VIISAPSK                                                           8

SEQ ID NO: 445          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 445
VLPALNGK                                                           8

SEQ ID NO: 446          moltype = DNA  length = 1020
FEATURE                 Location/Qualifiers
source                  1..1020
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 446
atggcatcag ataagaagat taggatcggc atcaacggat ttggaaggat tggccgtttg  60
gtggccaggg ttgcccttca aaggaatgat gttgaacttg ttgctgtcaa cgatcctttc  120
atcaccactg attacatgac atacatgttc aagtacgaca ccgttcatgg ccagtggaag  180
cattttgatg tgaaggtgaa ggactctaag acccttctct ttggtgagaa gcccgtcacc  240
gtttttggaa tcaggaaccc tgaggaaatt ccatggggtg aggttggtgc tgattatgtt  300
gttgagtcaa ctggtgtatt cactgacaag gacaaggctg ctgctcactt gaagggtggt  360
gccaagaagg ttgtcatctc tgccccaagc aaagatgcac ccatgtttgt tgttggtgtt  420
aacgagaagg agtacaagcc agagcttgac attgtttcca atgctagctg cactaccaat  480
tgtcttgctc ccctcgccaa ggttatcaac gaccgatttg gaattgtgga gggtcttatg  540
accactgttc atgccatcac agctactcag aagactgttg atggtccatc aaacaaggac  600
tggaggggtg gaagagctgc ttccttcaat ataattccca gcagcactgg agctgctaag  660
gctgtaggaa aagttcttcc agcactgaat ggtaaattga ctggcatgtc cttccgagtc  720
cccactgtag atgtttcagt tgttgacctc actgtcaggc tagagaagcc agcaacctat  780
```

-continued

```
gagcaaatca aatctgctat caaggaggaa tcagagggca aattgaaggg tattttgggt  840
tacaccgaag atgatgttgt gtctactgat tttgttggtg ataacaggtc tagcatattt  900
gacgccaagg ccggaatttc tttgaacgac aactttgtaa aacttgtttc ctggtatgac  960
aacgaatggg gttacagttc ccgtgtgatt gatttgattg tccacattgc atccgtggct  1020

SEQ ID NO: 447          moltype = AA  length = 340
FEATURE                 Location/Qualifiers
source                  1..340
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 447
MASDKKIRIG INGFGRIGRL VARVALQRND VELVAVNDPF ITTDYMTYMF KYDTVHGQWK  60
HFDVKVKDSK TLLFGEKPVT VFGIRNPEEI PWGEVGADYV VESTGVFTDK DKAAAHLKGG  120
AKKVVISAPS KDAPMFVVGV NEKEYKPELD IVSNASCTTN CLAPLAKVIN DRFGIVEGLM  180
TTVHAITATQ KTVDGPSNKD WRGGRAASFN IIPSSTGAAK AVGKVLPALN GKLTGMSFRV  240
PTVDVSVVDL TVRLEKPATY EQIKSAIKEE SEGKLKGILG YTEDDVVSTD FVGDNRSSIF  300
DAKAGISLND NFVKLVSWYD NEWGYSSRVI DLIVHIASVA                        340

SEQ ID NO: 448          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 448
AGISLNDNFV K                                                       11

SEQ ID NO: 449          moltype = AA  length = 27
FEATURE                 Location/Qualifiers
source                  1..27
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 449
GILGYTEDDV VSTDFVGDNR SSIFDAK                                      27

SEQ ID NO: 450          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 450
VLPALNGK                                                           8

SEQ ID NO: 451          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 451
VVISAPSK                                                           8

SEQ ID NO: 452          moltype = DNA  length = 1050
FEATURE                 Location/Qualifiers
source                  1..1050
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 452
atggtgacct cgtcgcaaaa gattctggtt accggcggtg ccggtttcat aggcacccac  60
accgtggtcc agcttctcaa aggaggcttc agtgtttcaa taatcgacaa tttcgataac  120
tccgtcgtgg aagcagtgga ccgcgtccgc caagtggtgg gccctcagct ttctcagaac  180
ctggaattca cccagggcga tctcaggaat agggatgact tggagaaact cttctccaga  240
acaacttttg atgccgtgat ccactttgct ggcctgaaag cggttgctga aagcgtttcc  300
aagccccgtc gctattttga ttttaatttg gtcggcacca tcaacctcta ccaagttatg  360
gcaaagtata attgcagaaa gatggttttc tcatcatctg caaccgttta tggccaacct  420
gaaaagatac cgtgtgagga ggatttcagg ttacaagcta tgaatcccta tggacggacc  480
aagcttttcc tggaagaaat tgcccgagat attcagaaag ctgaaccaga atggaagatc  540
atattactga ggtacttcaa tccggttggg gctcatgaaa gtggtaaact cggtgaagat  600
cccaagggca tcccaaataa cctcatgcct tacattcagc aagtagccgt tggaagattg  660
ccggaactca atgtatatgg gcatgattat cccacaaggg atggctctgc gatccgggac  720
tatatccatg tgatggactt agcagatggt catattgctg ccttgcgaaa gctcttcaca  780
acagagaaca taggttgtac tgcttacaac ctgggaactg gtcgtggaac atccgtgctt  840
gaaatggtta cagcatttga aaaggcgtct ggcaagaaaa ttccagtaaa attatgtcca  900
agaagaccgg gagatgctac tgaggtttat gcatctacag agagagctga gaaagaactt  960
ggatggaaag caaaatatgg tgtggatgag atgtgcaggg accaatggaa ttgggcaaag  1020
aacaatcccc agggttacac ggggaagcct                                   1050

SEQ ID NO: 453          moltype = AA  length = 350
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..350
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 453
MVTSSQKILV TGGAGFIGTH TVVQLLKGGF SVSIIDNFDN SVVEAVDRVR QVVGPQLSQN  60
LEFTQGDLRN RDDLEKLFSR TTFDAVIHFA GLKAVAESVS KPRRYFDFNL VGTINLYQVM  120
AKYNCRKMVF SSSATVYGQP EKIPCEEDFR LQAMNPYGRT KLFLEEIARD IQKAEPEWKI  180
ILLRYFNPVG AHESGKLGED PKGIPNNLMP YIQQVAVGRL PELNVYGHDY PTRDGSAIRD  240
YIHVMDLADG HIAALRKLFT TENIGCTAYN LGTGRGTSVL EMVTAFEKAS GKKIPVKLCP  300
RRPGDATEVY ASTERAEKEL GWKAKYGVDE MCRDQWNWAK NNPQGYTGKP             350

SEQ ID NO: 454          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 454
LFLEEIARDI QK                                                      12

SEQ ID NO: 455          moltype = DNA  length = 1050
FEATURE                 Location/Qualifiers
source                  1..1050
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 455
atggtgtctt cctcccaaca cattctggtc accggtggtg ccggtttcat tggcacccac  60
accgtcgttc agcttctcaa agctggcttc agcgtttcaa taatcgacaa tttcgataac  120
tccgtcatgg aagcagtgga ccgcgtccgc caagtggttg gccctctgct ttctcagaac  180
ctccaattca cccagggcga tctccggaat agggatgact tggagaaact cttctccaaa  240
acaacatttg atgccgtgat ccactttgct ggcttgaaag cggttgctga aagcgttgcg  300
aagccccgtc gctattttga ttttaatttg gttggcacca tcaacctcta cgaatttatg  360
gcaaagtata attgcaaaaa gatggttttc tcatcatctg caaccgttta tggccaacct  420
gaaaagatac cgtgtgagga ggatttcaag ttacaagcta tgaatcccta tggacggacc  480
aagcttttcc tggaagaaat tgcccgagat attcagaaag ctgaaccaga atggaagatc  540
atattactga gatacttcaa tccagttggg gctcatgaaa gtggcaaact cggtgaagat  600
cccaagggca tcccaaataa cctcatgcct tacattcagc aagtagctgt tggaagattg  660
actgaactca atgtatacgg tcatgattat ccaacgaggg atggctctgc gatccgggac  720
tatatccatg tgatggactt ggcagatggc catattgctg ccctgcgaaa gctcttcaca  780
acggagaaca taggttgtac tgcttacaac ctgggaactg gtcgtggaac atctgtgctt  840
gaaatggttg cagcatttga aaaggcttct ggcaagaaaa ttccagtaaa attatgtcca  900
agaagaccgg gagatgcgac tgaggtttat gcatctacag agagagctga gaaagaactt  960
ggttggaagg caaactatgg tgtggaggag atgtgcaggg accaatggaa ttgggcaaag  1020
aacaatccct ggggttacgc ggggaagcct                                   1050

SEQ ID NO: 456          moltype = AA  length = 350
FEATURE                 Location/Qualifiers
source                  1..350
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 456
MVTSSQKILV TGGAGFIGTH TVVQLLKGGF SVSIIDNFDN SVVEAVDRVR QVVGPQLSQN  60
LEFTQGDLRN RDDLEKLFSR TTFDAVIHFA GLKAVAESVS KPRRYFDFNL VGTINLYQVM  120
AKYNCRKMVF SSSATVYGQP EKIPCEEDFR LQAMNPYGRT KLFLEEIARD IQKAEPEWKI  180
ILLRYFNPVG AHESGKLGED PKGIPNNLMP YIQQVAVGRL PELNVYGHDY PTRDGSAIRD  240
YIHVMDLADG HIAALRKLFT TENIGCTAYN LGTGRGTSVL EMVTAFEKAS GKKIPVKLCP  300
RRPGDATEVY ASTERAEKEL GWKAKYGVDE MCRDQWNWAK NNPQGYTGKP             350

SEQ ID NO: 457          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 457
LFLEEIARDI QK                                                      12

SEQ ID NO: 458          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 458
TTFDAVIHFA GLK                                                     13

SEQ ID NO: 459          moltype = DNA  length = 1341
FEATURE                 Location/Qualifiers
```

-continued

```
source                  1..1341
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 459
atgggtaagg aaaagactca catcaacatt gtcgtcattg gacatgtcga ctctgggaag  60
tcaactacca ctggtcattt gatctacaag cttggaggta ttgacaagcg tgtgattgag  120
aggttcgaga aggaggctgc tgagatgaac aagaggtcat tcaagtatgc ctgggtgctt  180
gacaagctca aggctgagcg cgaaagagga attaccattg atattgcttt gtggaagttt  240
gaaaccacca agtactactg cacagtcatc gatgcccccg gacatcgtga ctttatcaag  300
aacatgatta ctggtacctc ccaggctgat tgtgctgtcc ttatcattga ctccaccact  360
ggtggttttg aagctggtat ttctaaggat ggacagactc gtgagcatgc tcttcttgct  420
ttcacccttg gtgtgaagca gatgatctgc tgctgtaaca agatggatgc cactaccccg  480
aagtactcta aggctaggta tgatgaaatt gtgaaggaag tctcttccta cttgaagaag  540
gttggttaca acccagacaa gattccattt gtgcccatct ctggtttcga gggtgacaac  600
atgattgaga ggtccaccaa ccttgactgg tacaagggac caactctcct tgaggctctt  660
gaccaaatca atgagcccaa gaggccctca gacaagccac tcaggcttcc attgcaggat  720
gtctacaaaa ttggtggtat tggtaccgtg ccagtgggac gtgttgagac tggtgtcgtg  780
aagcccggta tggtggtgac ttttggtccc actgggctga caactgaggt taagtctgtt  840
gagatgcacc atgaggctct cacagaggct cttccaggtg acaatgttgg ctttaatgtg  900
aagaatgttg cagtcaagga tctcaagcgt ggttttgttg catccaactc caaggatgac  960
cctgccaagg aagctgccaa cttcacatct caagtcatta tcatgaacca ccctggccag  1020
attggaaatg gatacgcacc ggtgcttgat tgccacactt ctcacattgc tgtcaagttt  1080
tctgaaatct tgaccaagat tgacaggcga tctggtaagg agcttgagaa ggagcccaaa  1140
tttttgaaga atggtgatgc aggtatggtt aagatgattc caaccaagcc catggtggtt  1200
gaaactttct ctgagtatcc tccccttggt cgttttgctg tgagggacat gcgtcaaact  1260
gtggctgtcg gagtcatcaa gagtgttgag aagaaagacc ccaccggagc caaggtcacc  1320
aaggccgctg ccaagaagaa g                                            1341

SEQ ID NO: 460          moltype = AA  length = 447
FEATURE                 Location/Qualifiers
source                  1..447
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 460
MGKEKTHINI VVIGHVDSGK STTTGHLIYK LGGIDKRVIE RFEKEAAEMN KRSFKYAWVL  60
DKLKAERERG ITIDIALWKF ETTKYYCTVI DAPGHRDFIK NMITGTSQAD CAVLIIDSTT  120
GGFEAGISKD GQTREHALLA FTLGVKQMIC CCNKMDATTP KYSKARYDEI VKEVSSYLKK  180
VGYNPDKIPF VPISGFEGDN MIERSTNLDW YKGPTLLEAL DQINEPKRPS DKPLRLPLQD  240
VYKIGGIGTV PVGRVETGVV KPGMVVTFGP TGLTTEVKSV EMHHEALTEA LPGDNVGFNV  300
KNVAVKDLKR GFVASNSKDD PAKEAANFTS QVIIMNHPGQ IGNGYAPVLD CHTSHIAVKF  360
SEILTKIDRR SGKELEKEPK FLKNGDAGMV KMIPTKPMVV ETFSEYPPLG RFAVRDMRQT  420
VAVGVIKSVE KKDPTGAKVT KAAAKKK                                      447

SEQ ID NO: 461          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 461
ARYDEIVK                                                           8

SEQ ID NO: 462          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 462
DGQTREHALL AFTLGVK                                                 17

SEQ ID NO: 463          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 463
EVSSYLK                                                            7

SEQ ID NO: 464          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 464
FSEILTK                                                            7

SEQ ID NO: 465          moltype = AA  length = 10
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 465
STTTGHLIYK                                                          10

SEQ ID NO: 466          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 466
THINIVVIGH VDSGK                                                    15

SEQ ID NO: 467          moltype = DNA  length = 1341
FEATURE                 Location/Qualifiers
source                  1..1341
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 467
atgggtaaag agaaggttca cattaacatt gtggtcattg gccatgttga ctctgggaag  60
tcaaccacca ctggccactt gatttacaag cttggaggta ttgacaagcg tgttattgag  120
aggtttgaga aagaagctgc tgagatgaac aagaggtcat tcaaatacgc ctgggtgctt  180
gacaagctta aggctgagcg tgagagagga attaccattg atattgcatt atggaagttt  240
gagaccacca aatactactg cactgttatt gatgctccag gacatcgtga tttcatcaag  300
aacatgatta ctggaacttc ccaggctgat tgtgctgtcc ttatcattga ttccacaact  360
ggtggttttg aagctggtat ttccaaagat ggacagactc gcgagcatgc tctacttgcc  420
ttcacccttg gtgtcaagca aatgatctgt tgctgtaaca agatggatgc cactacccca  480
aagtactcta aggccaggta tgaagaaatt gtgaaggaag tttcttccta cttgaagaag  540
gtaggttaca acccagacaa gattccattt gttcccatct caggttttga gggtgacaac  600
atgattgaga ggtccaccaa tcttgattgg tacaagggac caacccttct cgaggctctt  660
gaccagatca atgagcccaa gagaccctca gacaagcctt tgaggcttcc cttgcaggat  720
gtctacaaga ttggtggtat tggaacagtg ccagtgggac gagttgaaac tggtgtcttg  780
aagcccggta tggtggtgac ttttgctccc actggcctga caactgaagt taagtctgtg  840
gagatgcacc atgaagctct ccaggaggcc cttccaggtg acaacgtagg atttaatgtg  900
aagaatgttg cagtcaagga tctcaagcgt ggttttgttg cctcaaactc taaggatgat  960
cctgccaagg aggctgccaa cttcacatcc caagtcatta taatgaacca tcctggccaa  1020
attggaaatg gttatgctcc tgtgcttgac tgccacactt ctcacattgc cgtcaagttt  1080
gccgaacttg tgaccaagat tgacaggcga tctggtaagg agattgagaa ggagcccaaa  1140
ttcttgaaaa atggagatgc aggttatgtt aagatgattc ccaccaagcc tatggtggtc  1200
gaaactttct ccgagtatcc tcctcttggt cgttttgccg tgagggacat gcgacaaact  1260
gtggctgttg gagtcatcaa gagtgtggag aagaaggacc caactggagc caaggtcacc  1320
aaggctgcac agaagaagaa g                                             1341

SEQ ID NO: 468          moltype = AA  length = 447
FEATURE                 Location/Qualifiers
source                  1..447
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 468
MGKEKVHINI VVIGHVDSGK STTTGHLIYK LGGIDKRVIE RFEKEAAEMN KRSFKYAWVL  60
DKLKAERERG ITIDIALWKF ETTKYYCTVI DAPGHRDFIK NMITGTSQAD CAVLIIDSTT  120
GGFEAGISKD GQTREHALLA FTLGVKQMIC CCNKMDATTP KYSKARYEEI VKEVSSYLKK  180
VGYNPDKIPF VPISGFEGDN MIERSTNLDW YKGPTLLEAL DQINEPKRPS DKPLRLPLQD  240
VYKIGGIGTV PVGRVETGVL KPGMVVTFAP TGLTTEVKSV EMHHEALQEA LPGDNVGFNV  300
KNVAVKDLKR GFVASNSKDD PAKEAANFTS QVIIMNHPGQ IGNGYAPVLD CHTSHIAVKF  360
AELVTKIDRR SGKEIEKEPK FLKNGDAGYV KMIPTKPMVV ETFSEYPPLG RFAVRDMRQT  420
VAVGVIKSVE KKDPTGAKVT KAAQKKK                                       447

SEQ ID NO: 469          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 469
DGQTREHALL AFTLGVK                                                  17

SEQ ID NO: 470          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 470
EVSSYLK                                                             7
```

-continued

```
SEQ ID NO: 471          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 471
STTTGHLIYK                                                          10

SEQ ID NO: 472          moltype = DNA  length = 1344
FEATURE                 Location/Qualifiers
source                  1..1344
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 472
atgggtaagg aaaaggttca catcagtatt gtggtcattg gccatgtcga ctctgggaaa  60
tccactacca ctggtcacct gatttacaag cttggaggca ttgacaagcg tgttattgag  120
aggtttgaga aggaagctgc tgagatgaac aagaggtctt tcaagtatgc ctgggtgctg  180
gacaaactta aggctgagcg tgaaagagga atcaccattg atattgcttt gtggaagttt  240
gaaacaacaa agtattattg cacagttatt gatgcgcctg gacataggga ttcattaag  300
aatatgatta ctgggacatc ccaagctgac tgtgctgttc ttatcattga ttcgaccact  360
ggtggttttg aagctggtat ttccaaggat ggacagactc gtgaacatgc tctgctttca  420
ttcacccttg gtgtgaaaca gatgatttgc tgctgtaaca aaatggatgc tactacaccc  480
aagtattcca aggccaggta tgatgaaatt gtgaaggaag tctcttccta cttgaagaaa  540
gtaggataca accctgacaa gattcctttt gttcctatct ctggttttga gggagacaac  600
atgattgaga ggtccacaaa ccttgactgg tacaagggtc caactctgct tgatgcactt  660
gaccagatta gtgagcccaa gaggccctct gacaagcccc tcaggcttcc ccttcaggat  720
gtgtacaaga ttggaggtat tggaactgtg ccagtgggac gtgttgagac cggtgtcttg  780
aagcctggaa tggtggtgac ttttgcacca actggactga caactgaagt caagtctgtg  840
gagatgcacc atgaatctct tacagaggca catcctggtg acaatgtggg attcaatgtt  900
aagaatgttg ctgttaagga tttgaagcgt ggttatgttg cctcaaactc aaaggatgac  960
cctgcaaagg aggctgctaa cttcacagcc caagtcatca tcatgaacca ccctggtcag  1020
attggaaatg gctatgcccc tgtcctcgac tgccacactt ctcacattgc tgtcaagttt  1080
gctgaactca tgaccaagat tgacaggcga tccggcaaag agcttgagaa ggagcccaag  1140
tttttgaaga acggtgatgc tggttttgtt aagatgattc caaccaaacc catggttgtt  1200
gaaactttct ccgagtatcc tccacttggt aggtttgctg ttagggacat gcgtcaaact  1260
gttgctgtgg gagtcatcaa gaacgttgag aagaaggatc ctaccggagc caaggtcacc  1320
aaggctgccc agaagaagaa gtga                                          1344

SEQ ID NO: 473          moltype = AA  length = 447
FEATURE                 Location/Qualifiers
source                  1..447
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 473
MGKEKVHISI VVIGHVDSGK STTTGHLIYK LGGIDKRVIE RFEKEAAEMN KRSFKYAWVL  60
DKLKAERERG ITIDIALWKF ETTKYYCTVI DAPGHRDFIK NMITGTSQAD CAVLIIDSTT  120
GGFEAGISKD GQTREHALLS FTLGVKQMIC CCNKMDATTP KYSKARYDEI VKEVSSYLKK  180
VGYNPDKIPF VPISGFEGDN MIERSTNLDW YKGPTLLDAL DQISEPKRPS DKPLRLPLQD  240
VYKIGGIGTV PVGRVETGVL KPGMVVTFAP TGLTTEVKSV EMHHESLTEA HPGDNVGFNV  300
KNVAVKDLKR GYVASNSKDD PAKEAANFTA QVIIMNHPGQ IGNGYAPVLD CHTSHIAVKF  360
AELMTKIDRR SGKELEKEPK FLKNGDAGFV KMIPTKPMVV ETFSEYPPLG RFAVRDMRQT  420
VAVGVIKNVE KKDPTGAKVT KAAQKKK                                       447

SEQ ID NO: 474          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 474
ARYDEIVK                                                            8

SEQ ID NO: 475          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 475
DGQTREHALL SFTLGVK                                                  17

SEQ ID NO: 476          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 476
EVSSYLK                                                             7
```

-continued

```
SEQ ID NO: 477          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 477
STTTGHLIYK                                                          10

SEQ ID NO: 478          moltype = DNA  length = 882
FEATURE                 Location/Qualifiers
source                  1..882
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 478
atggcttccg gtgaacagaa attccctcct caacaacaac aaacacagcc tgggaaggag  60
catgctatga ctccagtacc ccaattcact agccctgact acaagccttc aaataaactt  120
caagggaaga ttgcattagt cactgggggt gattctggga ttggacgagc ggtgtgtaac  180
ttgtttgcct tagaaggtgc taccgtggcc ttcacgtatg tgaaggggca tgaggacaag  240
gacgcgaggg acacattgga aatgatcaag agagcaaaga cttcggatgc caaggatcca  300
atggcaatag catctgattt gggttacgat gagaactgca agagggtggt tgatgaggtc  360
gtgagtgctt atggttgtat tgacattctg gtcaacaatg cagctgagca gtacgagtgt  420
ggaaccgtgg aggacataga cgagcctagg cttgagaggg tctttcgtac aaatatcttc  480
tcctatttct tcatggcgag gcatgccttg aagcacatga aggaaggaag cagcattatc  540
aacacgacat cagtgaatgc atacaaggga catgcgaaac tattggacta cacgtccacc  600
aagggggcaa ttgtggccta tacaaggggt cttgcccttc agctggtgag taagggaatt  660
cgggttaatg gggtggctcc agggcccatt tggacccctt tgataccagc ctctttcaag  720
gaggaagaaa cggcccaatt tggagcgcag gtcccaatga agagagctgg tcaacctatt  780
gaggttgctc cttcctatgt ttttcttgct tccaaccaat gctcctctta cataactgga  840
caagtccttc accccaatgg tggaaccgtt gtgaatggtt aa                     882

SEQ ID NO: 479          moltype = AA  length = 293
FEATURE                 Location/Qualifiers
source                  1..293
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 479
MASGEQKFPP QQQQTQPGKE HAMTPVPQFT SPDYKPSNKL QGKIALVTGG DSGIGRAVCN  60
LFALEGATVA FTYVKGHEDK DARDTLEMIK RAKTSDAKDP MAIASDLGYD ENCKRVVDEV  120
VSAYGCIDIL VNNAAEQYEC GTVEDIDEPR LERVFRTNIF SYFFMARHAL KHMKEGSSII  180
NTTSVNAYKG HAKLLDYTST KGAIVAYTRG LALQLVSKGI RVNGVAPGPI WTPLIPASFK  240
EEETAQFGAQ VPMKRAGQPI EVAPSYVFLA SNQCSSYITG QVLHPNGGTV VNG         293

SEQ ID NO: 480          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 480
EGSSIINTTS VNAYK                                                    15

SEQ ID NO: 481          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 481
FPPQQQQTQP GK                                                       12

SEQ ID NO: 482          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 482
GAIVAYTRGL ALQLVSK                                                  17

SEQ ID NO: 483          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 483
LLDYTSTK                                                            8

SEQ ID NO: 484          moltype = DNA  length = 882
```

-continued

```
FEATURE                 Location/Qualifiers
source                  1..882
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 484
atggcaagca ataaggagtc taagttccca gcacagagcc agaaaactca gccaggaaaa  60
gaacatgtaa tgaatccact cccacaagcc acaaatcctg atcacaaggc cgccaataaa  120
ctccagggaa aggtggcgtt ggtgacagga ggtgactcag gaattggcag agcggtttgc  180
ctgtgtttcg caaaagaggg tgcaaccgtg gcctttacat acgtaaaggg ccatgaggac  240
agggataaag atgatactct gaagatgctg cttgaagcta agacaagtgg tgcagacaat  300
ccattggcaa tagcagcgga tattggcttt gatgagaact gcaaacaggt cattgacctt  360
gttgtcaaag aatatggccg ccttgatgtt ctggtcaaca atgcagctga gcagcatttg  420
acaaactctg ttgaggaaat cacacaacag cagcttgaga gagtcttcgg aaccaacatc  480
ttttctcagt tctttttggt caagcatgct ctgaagcaca tgaaagaagg gagctgcatc  540
ataaactcta cttcagttaa tgcatacaat gggaatccag aagcgttgga ctacactgct  600
accaagggag caattgtggc cttcaccaga ggtctttctc agcagctagc gagtagggga  660
attagggtga atggtgtggc acctggccca gtttggacgc caatacaacc agcttcaaag  720
cctgctgaga tgattcagaa cttggggtgt gaggtgccaa tgaaccgcgt ggctcagcct  780
tgtgagattg caccatgtta tttgttcttg gcaacttgtc aggactcttc ctactttact  840
ggccaagtcc tccatccaaa tggtgggatg gtcgtcaacg ct                     882

SEQ ID NO: 485          moltype = AA  length = 294
FEATURE                 Location/Qualifiers
source                  1..294
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 485
MASNKESKFP AQSQKTQPGK EHVMNPLPQA TNPDHKAANK LQGKVALVTG GDSGIGRAVC  60
LCFAKEGATV AFTYVKGHED RDKDDTLKML LEAKTSGADN PLAIAADIGF DENCKQVIDL  120
VVKEYGRLDV LVNNAAEQHL TNSVEEITQQ QLERVFGTNI FSQFFLVKHA LKHMKEGSCI  180
INSTSVNAYN GNPEALDYTA TKGAIVAFTR GLSQQLASRG IRVNGVAPGP VWTPIQPASK  240
PAEMIQNLGC EVPMNRVAQP CEIAPCYLFL ATCQDSSYFT GQVLHPNGGM VVNA        294

SEQ ID NO: 486          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 486
EGATVAFTYV K                                                       11

SEQ ID NO: 487          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 487
QVIDLVVK                                                           8

SEQ ID NO: 488          moltype = DNA  length = 879
FEATURE                 Location/Qualifiers
source                  1..879
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 488
atggcttcgg gtgaaaagaa attcccacct caacaacaac aaacacagcc tgggaaggag  60
catgctatga atccagtacc ccaattcact agccctgact acaagccttc aaataaactt  120
caaggaaaga tagcattagt gactgggggt gactctggga ttggacgagc ggtgtgtaac  180
ttgtttgcct tagaaggtgc taccgtgggc ttcacatatg tgaaggggca tgaggacaag  240
gacgcgaggg acacgttgga aatgatcaag agggcaaaga cttcggatgc taaagatcca  300
atggcggtac cagctgattt gggttacgac gagaattgca agagagtggt tgatgaggtc  360
gtgaatgctt atggttgtat tgacattctg gtcaacaatg cagctgagca atacgagtgt  420
ggaacagtgg aggacattga tgagcctagg cttgagaggg tctttcgtac aaatatcttc  480
tcctatttct tcatgaccag gcatgccttg aagcacatga aggaaggaag cagcattatc  540
aacacgacat cggtgaatgc atacaaggga aatgcgaaac tattggacta cacgtccacg  600
aagggagcaa ttgtggccta tacaaggga cttgctcttc agttggtgag taagggaatt  660
cgggttaatg gggtggctcc tgggcctatt tggaccccat tgataccctc ctctttcaag  720
gaggaagaaa cggctcaatt tggtgcccag gtgccaatga agagagctgg ccagcctatt  780
gaggttgctc cgtcttatgt ttttcttgct tgcaaccaat gctcctctta cataactgga  840
caagtccttc accccaatgg tggaaccgtt gtcaatggt                         879

SEQ ID NO: 489          moltype = AA  length = 293
FEATURE                 Location/Qualifiers
source                  1..293
                        mol_type = protein
                        organism = Glycine max
```

-continued

```
SEQUENCE: 489
MASGEKKFPP QQQQTQPGKE HAMNPVPQFT SPDYKPSNKL QGKIALVTGG DSGIGRAVCN  60
LFALEGATVG FTYVKGHEDK DARDTLEMIK RAKTSDAKDP MAVPADLGYD ENCKRVVDEV  120
VNAYGCIDIL VNNAAEQYEC GTVEDIDEPR LERVFRTNIF SYFFMTRHAL KHMKEGSSII  180
NTTSVNAYKG NAKLLDYTST KGAIVAYTRG LALQLVSKGI RVNGVAPGPI WTPLIPSSFK  240
EEETAQFGAQ VPMKRAGQPI EVAPSYVFLA CNQCSSYITG QVLHPNGGTV VNG         293

SEQ ID NO: 490          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 490
EGSSIINTTS VNAYK                                                   15

SEQ ID NO: 491          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 491
FPPQQQQTQP GK                                                      12

SEQ ID NO: 492          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 492
GAIVAYTRGL ALQLVSK                                                 17

SEQ ID NO: 493          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 493
LLDYTSTK                                                           8

SEQ ID NO: 494          moltype = DNA  length = 3207
FEATURE                 Location/Qualifiers
source                  1..3207
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 494
atggcttctg agcaaatgcc tcgcagagag aacaacagca atagggtgat agagagagag  60
gttcacgttg agacacacag aaaagagtcc ctccctgata agggaagggt tgtgatagag  120
agtgagaaga aggagaaaga gagagagttg gcttctgata gagccaggtc tggaaatata  180
gttaaggagg acaaagaggt tgaagaagga agaagaggaa aagaaagtgg tggtggtggt  240
ggtgttcagc atgagaggaa gtttgaaacg aaagttgaag agaagaaggg ttttggggaa  300
aaggctgaac tcgagggaat aacaagagtg gtggtgggaa gagagtgcag caaagaggag  360
aaacaaggaa gaacaagaga gaaagtagag gcagagaaag caagatcaat ggagaaaact  420
agagatgaag ctaaaggaag agaacaagct gaaaccaaag gtcaaaccaa gaagggtggt  480
ggtgaggaac agggaagaag aagagaagaa ataggagaag aaaataaaca atccagcctc  540
gaagagattt ctaaatgcag agctgaagct gaagagaaag ttaaggagag aaaggagaga  600
tcaaaacaag cagcaagtaa tgaagctgcc aaagagagaa acaagcaggc aaataaggag  660
gcaacaaagg caaaagatgt cacaagagag aagggtcaac aaggcaatct taaaacagga  720
aatgagacac aagacattgg ccaaaagggt tatggtggag caaaggatac cattggtaga  780
gtggacaaca ttactacacc acttgctaaa aaggctcaag ttgctggtag caccactgtt  840
cagtatgctg gagagaaatc agcacaagcc aaagatgcta cggtaggatc aatggagaaa  900
actacaaagg atgctaaagg aagagaaaaa gcaggaacca aaggtcaaac aaagaagggt  960
ggtggtgagg aacaaggaaa aagaagagaa gaaacaggag aagaaaaaca atccagcctg  1020
gaggagattt ctaagcacag agctgaagct caacaaaaag caatgagttc aattgaaggt  1080
gctaaggaga gatacgagag agcaaaacaa gaagcaagta aagcttccaa agagagaaac  1140
attcaagcaa ataagagggc tgaagaagaa agaagaaaag cagaaagtgg taggggtgtt  1200
catgtggata agtttgaaga gaagaaggat tcagtgctga aagcagaact cgagagaaga  1260
acaagagagg tagagggaag agagtggaac aaagaggaga aacaaagaag ttgtccaaga  1320
gggaaagtag aggcagagaa agcaagacca atggaagaaa caaaactgga agctaaagga  1380
agagaaaaag ctgaaaccaa agagggtggt ggtgaggaaa aaggaatgag aagagaagaa  1440
ataggagaag aaaataaaca atccagcctg gaagagattt ttaagtacag agctgaagct  1500
caacagaaag ttaaggagag aaacgagaga gcaaaacaag cagcagaagc tgccaaagag  1560
agaaacaagc aggcaaataa agggaatgaa gaaggcagaa gagaaagaga aagtggtggg  1620
ggtgttcatc atgaggctaa gtttgaaact aaaggtgaaa aaaagaagga tttagtggga  1680
agagagcgga acaaagagga gaaacaaaga agcagtccaa gaggaaaagt agaggcagag  1740
aaagcaagac caatggaaga aacaaaggtg gaagctaaag gaagagaaaa agatgtaacc  1800
aaagagggtg gtgttgagga acaaggaaga ggaagaaaag aaaatgaaca atctagcttg  1860
```

-continued

```
aaagacattt ccaagtacag agctgaagct caacaaaatt cattcaatgc tactgaagct  1920
gctaaggaaa gatacgagag ggcaaaccaa gcgacaagtg aagctgccaa agagagaaac  1980
aagcaggcaa ataaggaggc aacaaaggca aaagatgtca caagagaaat ggatcaacaa  2040
ggcaatctta aaacagaaaa tgagacacag gacaatgacc agaagggtta tggtggagca  2100
aagtacacca ttggtagtgt ggaccacatt actacaccat ttattgtaga ggccaatctt  2160
aaatcagaaa atcagataca agacaatggc caaaagggtt atggtggagc aaaggacaca  2220
attggtagtg tagaccacat tactacacca cttgctgaaa aggccaaagc tgctggtagc  2280
accactgctc agtatgttgg ggagaaagca gcacaagcca aagatgctac tgtgcaaact  2340
ggcaagagtg ctgcaaaagt tgcctcagat ttgagggaca aggctttggc cacagggtgg  2400
agtgctgcac attattctac tgagataact gtggagggaa ccaaggctgc aacaaatgtt  2460
gttaagggag ttgctgagta tgcaggacag aaggcttctg agcttgcagc tatgtcagtg  2520
gatactgcga aaggtttagc tgctgctgca gctggggaga atgccaagga gtacactgca  2580
aggaagaagg tggaagctga gatggatttg gaggccaaaa gggaagctca aaatcaggaa  2640
tcgaaggagt ggccatcagc aaaatccacc actgagagca ttcaaagagg acaaggacaa  2700
caagaacagc atgagcaagg agaaaaatgg gaacatggtg cacaaaggac taataagcca  2760
catgaaagca acattgtgga aacctttgag gaaaacctag gtgggattgg ggaaaaaagg  2820
gaacaaaaac cacttggaag catcattgga gagagctttc aaggagttca agaaaaaggg  2880
agtgaagtga tggcaagcat tggtgaaaaa ctgggaagtg ctgcacatac attagtgaag  2940
aagccattgg ataaggccac tgagggaggg agagaggtgt tgggagctgt aggtgaaact  3000
gtggttgaga ttggtgaaaa catggtgaaa aaaccagcac agaaagtgca gcagcatcaa  3060
ggtgaaggtg gaggaggggt gctgaatgcc attggtgaaa ccatagcaga gatagcagag  3120
acaactaagg tgattgttgc tggagaggga gaaaaggaga taatagtaaa aacacatgaa  3180
tttgaatatg gagacttgaa acttgat                                      3207

SEQ ID NO: 495         moltype = AA  length = 1069
FEATURE                Location/Qualifiers
source                 1..1069
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 495
MASEQMPRRE NNSNRVIERE VHVETHRKES LPDKGRVVIE SEKKEKEREL ASDRARSGNI  60
VKEDKEVEEG RRGKESGGGG GVQHERKFET KVEEKKGFGE KAELEGITRV VVGRECSKEE  120
KQGRTREKVE AEKARSMEKT RDEAKGREQA ETKGQTKKGG GEEQGRRREE IGEENKQSSL  180
EEISKCRAEA EEKVKERKER SKQAASNEAA KERNKQANKE ATKAKDVTRE KGQQGNLKTG  240
NETQDIGQKG YGGAKDTIGR VDNITTPLAK KAQVAGSTTV QYAGEKSAQA KDATVGSMEK  300
TTKDAKGREK AGTKGQTKKG GGEEQGKRRE ETGEEKQSSL EEISKHRAEA QQKAMSSIEG  360
AKERYERAKQ EASKASKERN IQANKRAEEE RRKAESGRGV HVDKFEEKKD SVLKAELERR  420
TREVEGREWN KEEKQRSCPR GKVEAEKARP MEETKLEAKG REKAETKEGG GEEKGMRREE  480
IGEENKQSSL EEIFKYRAEA QQKVKERNER AKQAAEAAKE RNKQANKGNE EGRRERESGG  540
GVHHEAKFET KGEKKKDLVG RERNKEEKQR SSPRGKVEAE KARPMEETKV EAKGREKDVT  600
KEGGVEEQGR GRKENEQSSL KDISKYRAEA QQNSFNATEA AKERYERANQ ATSEAAKERN  660
KQANKEATKA KDVTREMDQQ GNLKTENETQ DNDQKGYGGA KYTIGSVDHI TTPFIVEANL  720
KSENQIQDNG QKGYGGAKDT IGSVDHITTP LAEKAKAAGS TTAQYVGEKA AQAKDATVQT  780
GKSAAKVASD LRDKALATGW SAAHYSTEIT VEGTKAATNV VKGVAEYAGQ KASELAAMSV  840
DTAKGLAAAA AGENAKEYTA RKKVEAEMDL EAKREAQNQE SKEWPSAKST TESIQRGQGQ  900
QEQHEQGEKW EHGAQRTNKP HESNIVETFE ENLGGIGEKR EQKPLGSIIG ESFQGVQEKG  960
SEVMASIGEK LGSAAHTLVK KPLDKATEGG REVLGAVGET VVEIGENMVK KPAQKVQQHQ  1020
GEGGGGVLNA IGETIAEIAE TTKVIVAGEG EKEIIVKTHE FEYGDLKLD              1069

SEQ ID NO: 496         moltype = AA  length = 13
FEATURE                Location/Qualifiers
source                 1..13
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 496
AAGSTTAQYV GEK                                                     13

SEQ ID NO: 497         moltype = AA  length = 21
FEATURE                Location/Qualifiers
source                 1..21
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 497
ALATGWSAAH YSTEITVEGT K                                            21

SEQ ID NO: 498         moltype = AA  length = 15
FEATURE                Location/Qualifiers
source                 1..15
                       mol_type = protein
                       organism = Glycine max
SEQUENCE: 498
AQVAGSTTVQ YAGEK                                                   15

SEQ ID NO: 499         moltype = AA  length = 15
FEATURE                Location/Qualifiers
source                 1..15
```

-continued

```
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 499
DTIGRVDNIT TPLAK                                                    15

SEQ ID NO: 500          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 500
DTIGSVDHIT TPLAEK                                                   16

SEQ ID NO: 501          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 501
ERELASDRAR SGNIVK                                                   16

SEQ ID NO: 502          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 502
ERYERANQAT SEAAK                                                    15

SEQ ID NO: 503          moltype = AA  length = 18
FEATURE                 Location/Qualifiers
source                  1..18
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 503
GGGEEQGRRR EEIGEENK                                                 18

SEQ ID NO: 504          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 504
GLAAAAAGEN AK                                                       12

SEQ ID NO: 505          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 505
GRVVIESEK                                                           9

SEQ ID NO: 506          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 506
GVAEYAGQK                                                           9

SEQ ID NO: 507          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 507
LGSAAHTLVK                                                          10

SEQ ID NO: 508          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 508
QSSLEEIFK                                                           9
```

-continued

```
SEQ ID NO: 509          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 509
QSSLEEISK                                                           9

SEQ ID NO: 510          moltype = AA  length = 20
FEATURE                 Location/Qualifiers
source                  1..20
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 510
REQKPLGSII GESFQGVQEK                                               20

SEQ ID NO: 511          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 511
TGNETQDIGQ K                                                        11

SEQ ID NO: 512          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 512
VIVAGEGEK                                                           9

SEQ ID NO: 513          moltype = AA  length = 28
FEATURE                 Location/Qualifiers
source                  1..28
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 513
VQQHQGEGGG GVLNAIGETI AEIAETTK                                      28

SEQ ID NO: 514          moltype = DNA  length = 735
FEATURE                 Location/Qualifiers
source                  1..735
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 514
atggcactgt gcaaccgcgt cagatcggga atctctctct tcaccaaatc cacgttcctc  60
acttcccagc gatccactct ccaacgatcc ctcgttgccc ctttcatttc cctcgcctcc  120
agaaattacg ccaacgtgcc ggggcaaaag gaaaccaaag ttaaggtccc tctggctttg  180
tttggaggtt caggaaacta tgcctctgct ttgtatattg catctgtgaa agccaatgca  240
gtggaaaagg ttgagactga gcttctgcaa tttgttgagg cagtaaagaa tagtcccata  300
ttttcacagt ttataaggga cttgtctgtg gctaaggatg ttagagttaa agttatccag  360
gatatttgtg gggaagccaa gttttcggat gttacaaaga acttccttgt tattgttgct  420
gagaatggaa ggcttaaaaa catagacacc attgcaaaga ggtttgcaca gttgtcaatg  480
gcatataagg gagaagtgca agcgactgtg accactgttg ttcctcttcc ccctgaagag  540
gaaaaggcat tgaaggagac tcttcaggaa ataataggtt ctggagcgaa ggttcatctt  600
gaacagaaga ttgatccaag catacttggt gggcttgtgc ttgaatttag ccagaaggtc  660
tttgacatgt caattaagac cagggcacaa cagatggaga ggatcctccg agagcccgtg  720
tccattgcca acatc                                                   735

SEQ ID NO: 515          moltype = AA  length = 245
FEATURE                 Location/Qualifiers
source                  1..245
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 515
MALCNRVRSG ISLFTKSTFL TSQRSTLQRS LVAPFISLAS RNYANVPGQK ETKVKVPLAL  60
FGGSGNYASA LYIASVKANA VEKVETELLQ FVEAVKNSPI FSQFIRDLSV AKDVRVKVIQ  120
DICGEAKFSD VTKNFLVIVA ENGRLKNIDT IAKRFAQLSM AYKGEVQATV TTVVPLPPEE  180
EKALKETLQE IIGSGAKVHL EQKIDPSILG GLVLEFSQKV FDMSIKTRAQ QMERILREPV  240
SIANI                                                              245

SEQ ID NO: 516          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
```

-continued

```
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 516
ETLQEIIGSG AK                                                        12

SEQ ID NO: 517          moltype = DNA  length = 678
FEATURE                 Location/Qualifiers
source                  1..678
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 517
atgaccacac aagtaccacc acacagtgtc caagtccaca caacaacaca ccgctacgaa  60
gctggcgtcg tccccccggg agcccgtttc gaaccgccgc gttatgaagc cggcgtcaaa  120
gcgccctcca tttaccattc cgagagaggt ccaacgactt cccaggtcct cgccgtcctc  180
gccggcctcc cggtcggtgg catcctgctc ctcctggcgg gattgaccct ggcaggaacc  240
ctcaccggac tggcagttgc cacgccgctc ttcgtcctct tcagcccggt gctggttcca  300
gccacggtcg ccatcggcct ggccgtggcc gggttcctga cgtcgggggc cttcgggctc  360
acggcgctgt cttccttctc ctggatcctg aactacatcc gggagaccca gccagcctcg  420
gagaacctcg ccgccgccgc gaagcatcac ctcgccgagg cggcggagta cgtggggcag  480
aagacgaagg aagtggggca gaagaccaag gaggttgggc aggatattca gagcaaggca  540
caggatacaa gggaggcagc tgcaagggat gcaagggagg cagctgcaag ggatgcaagg  600
gaagctgctg caagggatgc aaaggtggag gcaagggatg taaagagaac aacagtgaca  660
gcaacaaccg caaccgca                                                  678

SEQ ID NO: 518          moltype = AA  length = 226
FEATURE                 Location/Qualifiers
source                  1..226
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 518
MTTQVPPHSV QVHTTTHRYE AGVVPPGARF EPPRYEAGVK APSIYHSERG PTTSQVLAVL  60
AGLPVGGILL LLAGLTLAGT LTGLAVATPL FVLFSPVLVP ATVAIGLAVA GFLTSGAFGL  120
TALSSFSWIL NYIRETQPAS ENLAAAAKHH LAEAAEYVGQ KTKEVGQKTK EVGQDIQSKA  180
QDTREAAARD AREAAARDAR EAAARDAKVE ARDVKRTTVT ATTATA                 226

SEQ ID NO: 519          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 519
EVGQDIQSK                                                            9

SEQ ID NO: 520          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 520
RTTVTATTAT A                                                         11

SEQ ID NO: 521          moltype = DNA  length = 627
FEATURE                 Location/Qualifiers
source                  1..627
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 521
atgacaagga aacaagcaag ttccacgaga cgtggcggtc tttcaggagt cttgcattcc  60
aaatccaaat cttctcctct actctccata tcccttgtcc tatttgtagc catccttctc  120
atcctttatg cttgtatcgg gtcaggtata ttgggaggaa gaaaggatgt gattagcatg  180
gctgaagatg atttttcatg cacttttgaa gttccaagtg caattcctgt tttgaagaat  240
gcatatggtg gcagcatgaa aaatgttttg cacgtaggcc ctgagagttg ttcggtggtc  300
tccaaatttt taagagaagg ggaaactgaa gcatggggtg tggagccata tgacttagac  360
gatgctgata ggaactgcaa ggctcttgtg cagaaaggca ttatacgtgt tgcagatata  420
aaattccctt taccatacag ggtgaagtct ttttctcatg ttatagtgtc agatgctttg  480
gattacctgt ctccaaaata tataaacaaa actcttcctg agctagcaag ggtatcttct  540
gatggtatta ttattttcac aggttaccca ggtcaaccga gagctaaaat agcaccatta  600
tccaaatttg gccgtccaca taggtgc                                       627

SEQ ID NO: 522          moltype = AA  length = 209
FEATURE                 Location/Qualifiers
source                  1..209
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 522
MTRKQASSTR RGGLSGVLHS KSKSSPLLSI SLVLFVAILL ILYACIGSGI LGGRKDVISM  60
```

-continued

```
AEDDFSCTFE VPSAIPVLKN AYGGSMKNVL HVGPESCSVV SKFLREGETE AWGVEPYDLD  120
DADRNCKALV QKGIIRVADI KFPLPYRVKS FSHVIVSDAL DYLSPKYINK TLPELARVSS  180
DGIIIFTGYP GQPRAKIAPL SKFGRPHRC                                    209

SEQ ID NO: 523          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 523
NAYGGSMK                                                           8

SEQ ID NO: 524          moltype = DNA  length = 723
FEATURE                 Location/Qualifiers
source                  1..723
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 524
atggccgcgc gagctctgtt gaccagaaaa accctagcca ccgtgctccg caacgacgcg  60
aagcccataa tcggagttgg cataacagca gcggctactc attcacgcgg gttgcacgtg  120
tacacgctac ccgatctgga ttacgactat ggcgcactgg agccagccat cagcggcgac  180
atcatgcagc tgcaccacca gaagcaccac cagacttaca tcaccaacta caacaaggcc  240
ctcgagcagc tccaagacgc catcgccaag aaagattcct ccgccgtcgt taagctccag  300
ggcgccatca agttcaacgg cggaggtcat gtcaaccatt ctattttctg gaaaaatcta  360
gctcctgttc gtgaaggagg tggtgaacca cccaagggtt cactgggatg ggctattgac  420
acacattttg gttcttttga agcattaata caaaaagtta acgcagaagg tgctgcacta  480
caggggtctg gatgggtgtg gcttggtctg gacaaagagt tgaagaggct tgtagttgaa  540
accactgcca accaggaccc actggttact aagggaccaa atttggttcc attgattggt  600
attgatgttt gggagcatgc gtactactta cagtacaaga atgttagacc agactatctg  660
aagaacattt ggaaagttat taattggaaa tatgccagtg aagtgtatga gaaagagagc  720
tct                                                                723

SEQ ID NO: 525          moltype = AA  length = 241
FEATURE                 Location/Qualifiers
source                  1..241
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 525
MAARALLTRK TLATVLRNDA KPIIGVGITA AATHSRGLHV YTLPDLDYDY GALEPAISGD  60
IMQLHHQKHH QTYITNYNKA LEQLQDAIAK KDSSAVVKLQ GAIKFNGGGH VNHSIFWKNL  120
APVREGGGEP PKGSLGWAID THFGSFEALI QKVNAEGAAL QGSGWVWLGL DKELKRLVVE  180
TTANQDPLVT KGPNLVPLIG IDVWEHAYYL QYKNVRPDYL KNIWKVINWK YASEVYEKES  240
S                                                                  241

SEQ ID NO: 526          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 526
ALEQLQDAIA K                                                       11

SEQ ID NO: 527          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 527
HHQTYITNYN K                                                       11

SEQ ID NO: 528          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 528
NLAPVREGGG EPPK                                                    14

SEQ ID NO: 529          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 529
NVRPDYLK                                                           8
```

-continued

```
SEQ ID NO: 530          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 530
RLVVETTANQ DPLVTK                                                   16

SEQ ID NO: 531          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 531
YASEVYEK                                                            8

SEQ ID NO: 532          moltype = DNA  length = 720
FEATURE                 Location/Qualifiers
source                  1..720
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 532
atggccgcgc gagccctgtt gaccagaaaa accctagcca ccgttctccg caacgacgcg  60
aagcccataa tcggagctgg catagctgca gctactcatt ctcgtgggtt gcacgtgtac  120
acgctccccg atctggatta cgactatggc gctctggagc cagccatcag cggcgaaatc  180
atgcagctgc accaccagaa gcaccaccag acttacatca ccaacttcaa caaggccctc  240
gagcagctcc aagacgccgt cgccaagaaa gattcctccg ccgtcgttaa gctccagggc  300
gccatcaagt tcaacggcgg aggtcacgtc aaccattcta ttttctggaa aaatctagct  360
cctgttcgtg aaggaggtgg tgaaccaccc aagggttcac tgggatgggc cattgacaca  420
cattttggct cttttgaagc attagtacaa aaagttaatg cagaaggtgc tgcactacag  480
gggtctggat gggtgtggct tggtctggac aaagagttga agaggcttgt agttgaaacc  540
actgccaacc aggacccact ggttactaag ggaccaaatt tggttccatt gcttggtatt  600
gatgtttggg agcatgcgta ctacttacag tacaagaatg ttagaccaga ctatctgaag  660
aacatttgga aagttattaa ttggaaatat gccagtgaag tgtatgagaa agagagctct  720

SEQ ID NO: 533          moltype = AA  length = 240
FEATURE                 Location/Qualifiers
source                  1..240
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 533
MAARALLTRK TLATVLRNDA KPIIGAGIAA ATHSRGLHVY TLPDLDYDYG ALEPAISGEI  60
MQLHHQKHHQ TYITNFNKAL EQLQDAVAKK DSSAVVKLQG AIKFNGGGHV NHSIFWKNLA  120
PVREGGGEPP KGSLGWAIDT HFGSFEALVQ KVNAEGAALQ GSGWVWLGLD KELKRLVVET  180
TANQDPLVTK GPNLVPLLGI DVWEHAYYLQ YKNVRPDYLK NIWKVINWKY ASEVYEKESS  240

SEQ ID NO: 534          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 534
ALEQLQDAVA K                                                        11

SEQ ID NO: 535          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 535
HHQTYITNFN K                                                        11

SEQ ID NO: 536          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 536
NLAPVREGGG EPPK                                                     14

SEQ ID NO: 537          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 537
```

-continued

```
NVRPDYLK                                                            8

SEQ ID NO: 538          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 538
RLVVETTANQ DPLVTK                                                   16

SEQ ID NO: 539          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 539
YASEVYEK                                                            8

SEQ ID NO: 540          moltype = DNA  length = 486
FEATURE                 Location/Qualifiers
source                  1..486
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 540
atggctccaa ttgcagtggg agatgtcatt ccagacggca ttttggctta cttggatgaa  60
gaaaacaaac cccaaactgt ctctattcac tctctcgccg ccggcaagaa agtcatcatc  120
tttggggttc ccggtgcctt cactcctaca tgcagcttga agcacgtgcc tggcttcatt  180
gagagagcag aagagctgaa aggaaagggt gtggacgaaa taatctgtat cagcgtgaat  240
gatccctttg tgatgaactc atgggccaaa acgttcccag agaacaagca tgttaagttc  300
cttgctgatg gtgcagccaa atacaccaat gccctcggtc ttgagcttga cctcaccgac  360
aagggtctcg gcgtccgctc aaagaggttt gctttgctgg tggaagacct caaggtgaag  420
gttgcaaatg ttgaaagtgg aggagagttc accatctcca gtgctgaaga gatcatcaag  480
gccctc                                                              486

SEQ ID NO: 541          moltype = AA  length = 162
FEATURE                 Location/Qualifiers
source                  1..162
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 541
MAPIAVGDVI PDGILAYLDE ENKPQTVSIH SLAAGKKVII FGVPGAFTPT CSLKHVPGFI  60
ERAEELKGKG VDEIICISVN DPFVMNSWAK TFPENKHVKF LADGAAKYTN ALGLELDLTD  120
KGLGVRSKRF ALLVEDLKVK VANVESGGEF TISSAEEIIK AL                     162

SEQ ID NO: 542          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 542
HVPGFIERAE ELK                                                      13

SEQ ID NO: 543          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 543
RFALLVEDLK                                                          10

SEQ ID NO: 544          moltype = AA  length = 20
FEATURE                 Location/Qualifiers
source                  1..20
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 544
VANVESGGEF TISSAEEIIK                                               20

SEQ ID NO: 545          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 545
YTNALGLELD LTDK                                                     14
```

-continued

```
SEQ ID NO: 546          moltype = DNA  length = 486
FEATURE                 Location/Qualifiers
source                  1..486
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 546
atggctccaa ttgcagtggg aaacgtgatt ccagacggca ttttggctta cttggatgaa  60
gaaaacaaac cccaaactgt ctccattcac tctctcgcca ccggcaagaa agtcatcatc  120
tttggggttc ccggtgcctt cactcctact tgcagcttga agcacgttcc tggcttcatt  180
gagagagcag aagagctgaa agggaagggt gtggacgaaa taatctgtat cagtgtgaat  240
gatccatttg tgatgaactc atgggccaaa acattcccag agaacaagca tgttaagttc  300
cttgctgatg gtgcagccaa atacaccaat gcccttggtc ttgagcttga cctcactgac  360
aagggtctcg gcgtccgctc aaagaggttt gctctgctgg tggaagacct caaggtgaag  420
gttgcgaatg ttgaaagtgg aggagagttc accatctcca gtgctgaaga gatcatcaag  480
gccctt                                                             486

SEQ ID NO: 547          moltype = AA  length = 162
FEATURE                 Location/Qualifiers
source                  1..162
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 547
MAPIAVGNVI PDGILAYLDE ENKPQTVSIH SLATGKKVII FGVPGAFTPT CSLKHVPGFI  60
ERAEELKGKG VDEIICISVN DPFVMNSWAK TFPENKHVKF LADGAAKYTN ALGLELDLTD  120
KGLGVRSKRF ALLVEDLKVK VANVESGGEF TISSAEEIIK AL                     162

SEQ ID NO: 548          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 548
HVPGFIERAE ELK                                                     13

SEQ ID NO: 549          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 549
RFALLVEDLK                                                         10

SEQ ID NO: 550          moltype = AA  length = 20
FEATURE                 Location/Qualifiers
source                  1..20
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 550
VANVESGGEF TISSAEEIIK                                              20

SEQ ID NO: 551          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 551
YTNALGLELD LTDK                                                    14

SEQ ID NO: 552          moltype = DNA  length = 456
FEATURE                 Location/Qualifiers
source                  1..456
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 552
atgtcgcagt tgctggataa agccaagaac ttcgtgtcgg agaaggtaaa tgatatggcc  60
aagccggagg cgagtgtcac tgacgtggac ttcaagcgcg tgagcaaaga taatgtcgag  120
tacttggcca aggtctctgt tcgtaaccct tattcaactt ccatacccat ttgtgagatc  180
aactactctt tcaaaagcgc atccagggag atagcatcag ggaaaattcc agacccagga  240
tcgttgaagg caaaggacac aacaatggtg gatgtgccag ttaaggttcc atacagcata  300
ttgatgagct tggcaaagga cattggtgct gactgggaca tagactatca attggatctt  360
ggtctggtta ttgatgttcc tgtcattggc atcttcacca ttcctctctc tcagaaagga  420
gagatcaagc taccaaccct ctctaccatg ttcgcc                            456

SEQ ID NO: 553          moltype = AA  length = 152
FEATURE                 Location/Qualifiers
source                  1..152
```

-continued

```
                      mol_type = protein
                      organism = Glycine max
SEQUENCE: 553
MSQLLDKAKN FVSEKVNDMA KPEASVTDVD FKRVSKDNVE YLAKVSVRNP YSTSIPICEI  60
NYSFKSASRE IASGKIPDPG SLKAKDTTMV DVPVKVPYSI LMSLAKDIGA DWDIDYQLDL  120
GLVIDVPVIG IFTIPLSQKG EIKLPTLSTM FA                                152

SEQ ID NO: 554        moltype = AA  length = 8
FEATURE               Location/Qualifiers
source                1..8
                      mol_type = protein
                      organism = Glycine max
SEQUENCE: 554
DNVEYLAK                                                           8

SEQ ID NO: 555        moltype = DNA  length = 480
FEATURE               Location/Qualifiers
source                1..480
                      mol_type = other DNA
                      organism = Glycine max
SEQUENCE: 555
atgtcggacg aagagcacca tttcgagtcc aaggcagacg ccggtgcctc caaaacctac  60
cctcagcaag ccggtaccat tcgcaagaac ggctacatcg tcatcaaagg acgcccctgc  120
aaggttgtgg aagtttcgac ttccaaaact ggaaagcacg gtcacgccaa gtgccacttt  180
gttggaattg atattttcac tgccaagaaa cttgaggata ttgtgccctc ttctcacaac  240
tgtgatgttc ctcatgtgaa tcgtactgat taccagctca ttgatattgc tgaggatgga  300
tttttgagtc tgcttactga aaatgggaac accaaggatg acctcaagct tcccactgat  360
gagagtctgc tcactcagat aaaggatgga tttgctgagg gcaaggatct tgttgtgtct  420
gtcatgtctg ctatgggtga ggaacagatt tgcgccctga aggatattgg gccaaagaac  480

SEQ ID NO: 556        moltype = AA  length = 160
FEATURE               Location/Qualifiers
source                1..160
                      mol_type = protein
                      organism = Glycine max
SEQUENCE: 556
MSDEEHHFES KADAGASKTY PQQAGTIRKN GYIVIKGRPC KVVEVSTSKT GKHGHAKCHF  60
VGIDIFTAKK LEDIVPSSHN CDVPHVNRTD YQLIDIAEDG FLSLLTENGN TKDDLKLPTD  120
ESLLTQIKDG FAEGKDLVVS VMSAMGEEQI CALKDIGPKN                        160

SEQ ID NO: 557        moltype = AA  length = 12
FEATURE               Location/Qualifiers
source                1..12
                      mol_type = protein
                      organism = Glycine max
SEQUENCE: 557
LPTDESLLTQ IK                                                      12

SEQ ID NO: 558        moltype = AA  length = 7
FEATURE               Location/Qualifiers
source                1..7
                      mol_type = protein
                      organism = Glycine max
SEQUENCE: 558
NGYIVIK                                                            7

SEQ ID NO: 559        moltype = DNA  length = 480
FEATURE               Location/Qualifiers
source                1..480
                      mol_type = other DNA
                      organism = Glycine max
SEQUENCE: 559
atgtcggacg aagagcatca cttcgattct caggcggatg cgggagcgtc caagacgtat  60
cctcagcagg ctggtaccat tcgcaagaat ggctacatcg tcatcaaatc caggccttgc  120
aaggttgtag aggtctctac gtcaaagaca ggcaaacatg gacatgcaaa gtgtcacttt  180
gctgcaattg atatttttaa tgggaaaaag cttgaagata ttgttccatc atcccacaac  240
tgtgatgttc cacatgtgaa tcgcactgac tatcagctta ttgatatctc tgaagacggc  300
tttgtgagtc tgttgactga taatggaaat accaaggatg atctgaagct tcctactgat  360
gaggcattgc ttgctcagat taaagatgga ttttctgaag gaaaagacct tgtggtttct  420
gttatgtctg caatgggaga agaacagatc tgttcactga aggacatcgg tcccaaaaat  480

SEQ ID NO: 560        moltype = AA  length = 160
FEATURE               Location/Qualifiers
source                1..160
                      mol_type = protein
                      organism = Glycine max
```

-continued

```
SEQUENCE: 560
MSDEEHHFDS QADAGASKTY PQQAGTIRKN GYIVIKSRPC KVVEVSTSKT GKHGHAKCHF  60
AAIDIFNGKK LEDIVPSSHN CDVPHVNRTD YQLIDISEDG FVSLLTDNGN TKDDLKLPTD  120
EALLAQIKDG FSEGKDLVVS VMSAMGEEQI CSLKDIGPKN                        160

SEQ ID NO: 561          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 561
NGYIVIK                                                            7

SEQ ID NO: 562          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 562
TYPQQAGTIR K                                                       11

SEQ ID NO: 563          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 563
VVEVSTSK                                                           8

SEQ ID NO: 564          moltype = DNA  length = 501
FEATURE                 Location/Qualifiers
source                  1..501
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 564
atggctagcc aatcaaacac taaatcagtt cacgatttca cagttaaaga tgccagggga  60
aataatgtta atctcgccga ttacaaagga aaggtccttc tcattgtcaa tgttgcctca  120
caatgcggct tgactaattc aaactacaca gagcttaacc aattgtatga gaaatacaaa  180
gggaaaggac ttgaaattct ggcattccca tgcaatcagt ttggggctca ggagccagga  240
actaatgaag agattgtgga gtttgcttgt actcgcttca aagctgagtt tcccattttc  300
gacaaggtgg atgtgaatgg tgacaatgct gctccactgt acaagtttct gaaatcaagc  360
aaaggtggac tctttgggga tagtatcaaa tggaacttct ccaaattcct tgttgataag  420
gatgggaatg tcgttgaccg ttatgcaccc acaacttctc ctctcagcat tgagaaagac  480
ataaagaagc tgctagatgc a                                            501

SEQ ID NO: 565          moltype = AA  length = 167
FEATURE                 Location/Qualifiers
source                  1..167
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 565
MASQSNTKSV HDFTVKDARG NNVNLADYKG KVLLIVNVAS QCGLTNSNYT ELNQLYEKYK  60
GKGLEILAFP CNQFGAQEPG TNEEIVEFAC TRFKAEFPIF DKVDVNGDNA APLYKFLKSS  120
KGGLFGDSIK WNFSKFLVDK DGNVVDRYAP TTSPLSIEKD IKKLLDA                167

SEQ ID NO: 566          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 566
AEFPIFDK                                                           8

SEQ ID NO: 567          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 567
DARGNNVNLA DYK                                                     13

SEQ ID NO: 568          moltype = AA  length = 19
FEATURE                 Location/Qualifiers
source                  1..19
                        mol_type = protein
                        organism = Glycine max
```

-continued

```
SEQUENCE: 568
DGNVVDRYAP TTSPLSIEK                                                19

SEQ ID NO: 569          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 569
GGLFGDSIK                                                           9

SEQ ID NO: 570          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 570
VDVNGDNAAP LYK                                                      13

SEQ ID NO: 571          moltype = DNA  length = 357
FEATURE                 Location/Qualifiers
source                  1..357
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 571
atggcttccg aagaatccca gaacaagccc cccagcgagc agtccactac cgaactcatc  60
acgagcgcga agctggtggc ggaggcggcg cagtcggctc tgaagagcga cagcgacaaa  120
gtggacaagg ccaaggtggc tgacgctgcc ggcgatcttc tcgacgcggc ggggaagtat  180
ggaaaacttg atgacaaaca aggcataggg caatatgttg acaaggctgc tgattatctg  240
cataattacc agggtgataa taccactgct ccttcacagc ctgcagagtc taaaggtgaa  300
gaaagtggtg gtgggttagg tggtcttgca aacttggctg gaggcttctt caaaaag     357

SEQ ID NO: 572          moltype = AA  length = 119
FEATURE                 Location/Qualifiers
source                  1..119
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 572
MASEESQNKP PSEQSTTELI TSAKLVAEAA QSALKSDSDK VDKAKVADAA GDLLDAAGKY  60
GKLDDKQGIG QYVDKAADYL HNYQGDNTTA PSQPAESKGE ESGGGLGGLA NLAGGFFKK   119

SEQ ID NO: 573          moltype = AA  length = 23
FEATURE                 Location/Qualifiers
source                  1..23
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 573
AADYLHNYQG DNTTAPSQPA ESK                                           23

SEQ ID NO: 574          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 574
LVAEAAQSAL K                                                        11

SEQ ID NO: 575          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 575
QGIGQYVDK                                                           9

SEQ ID NO: 576          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 576
VADAAGDLLD AAGK                                                     14

SEQ ID NO: 577          moltype = DNA  length = 777
FEATURE                 Location/Qualifiers
source                  1..777
```

-continued

```
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 577
atggcggccg ctccctcgcc tcgcgaagag aacgtctaca tggcgaagct cgccgagcag  60
gccgagcgct acgaagagat ggtggagttc atggagaagg tctccgccgc cgccgacaac  120
gaggaactta ccgtggagga gaggaacctc ctctccgtcg cctacaagaa cgtgatcgga  180
gcgcgccgcg cctcgtggcg catcatctcc tcgatcgagc agaaggagga gagccgcggc  240
aacgaggatc acgtctccgt catccgcgac taccgatcca agatcgagtc cgagctctcc  300
aacatctgcg acggtatcct caagctcctc gactctcgcc tcattccatc cgcttcctcc  360
ggcgattcca aggtcttcta cctcaagatg aagggagatt accacagata ccttgccgag  420
ttcaagaccg gcgccgagcg taaggaggct gccgagagca ctctctctgc atacaaagcc  480
gcacaggaca ttgcgaatgc cgaactgcct cctactcatc caattaggct tggccttgct  540
cttaactttt ctgtatttta ctacgaaatc cttaattctc cggatcgtgc ctgcaacctt  600
gcaaaacagg cttttgatga agctattgct gaattggata cactgggaga ggagtcatac  660
aaggatagca ctttgatcat gcaactcctt cgtgataacc tcacactttg gacctctgac  720
atgcaggacg atggagcaga tgaaattaaa gaagcagcac cgaaaccgga tgaccag     777

SEQ ID NO: 578          moltype = AA  length = 259
FEATURE                 Location/Qualifiers
source                  1..259
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 578
MAAAPSPREE NVYMAKLAEQ AERYEEMVEF MEKVSAAADN EELTVEERNL LSVAYKNVIG  60
ARRASWRIIS SIEQKEESRG NEDHVSVIRD YRSKIESELS NICDGILKLL DSRLIPSASS  120
GDSKVFYLKM KGDYHRYLAE FKTGAERKEA AESTLSAYKA AQDIANAELP PTHPIRLGLA  180
LNFSVFYYEI LNSPDRACNL AKQAFDEAIA ELDTLGEESY KDSTLIMQLL RDNLTLWTSD  240
MQDDGADEIK EAAPKPDDQ                                               259

SEQ ID NO: 579          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 579
EAAESTLSAY K                                                       11

SEQ ID NO: 580          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 580
LLDSRLIPSA SSGDSK                                                  16

SEQ ID NO: 581          moltype = AA  length = 19
FEATURE                 Location/Qualifiers
source                  1..19
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 581
QAFDEAIAEL DTLGEESYK                                               19

SEQ ID NO: 582          moltype = AA  length = 23
FEATURE                 Location/Qualifiers
source                  1..23
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 582
VSAAADNEEL TVEERNLLSV AYK                                          23

SEQ ID NO: 583          moltype = DNA  length = 993
FEATURE                 Location/Qualifiers
source                  1..993
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 583
atggcgttga acaatggcgt ccccgccacc ggcaccggcc tcatcgtcag cttcggtgag  60
atgctcatcg acttcgtccc caccgtctct ggcgtgtccc tggccgaggc ccctggcttc  120
ctcaaggccc ccggcggcgc ccccgctaac gtcgccatcg ccgtgtcgcg cctcggcggc  180
aaagccgcct tcgtcggcaa gctcggcgac gacgagttcg gccacatgct cgccggaatc  240
ctcaaggaaa acggcgttcg cgccgacggc atcaactttg accagggcgc acgcaccgcc  300
ctggccttcg tgaccctacg cgccgacggg gagcgtgagt tcatgttcta cagaaacccc  360
agcgccgaca tgctcctcaa gcccgaagaa ctcaatctcg aactcatcag atctgcaaaa  420
gttttccatt acggatcaat cagtttgatc gtggagccat gcagatcagc acacttgaag  480
gcaatggaag tagccaagga atctgggtgc ttgctctcct atgaccccaa ccttcgtcta  540
```

-continued

```
cctttgtggc cttcggctga ggaagctcgt aagcaaatac tgagcatttg ggagaaggct  600
gatttgatca aggtcagtga tgcggagctt gagttcctca caggaagtga caagattgat  660
gatgaatctg ctttgtcatt gtggcacccc aatttgaagt tgctccttgt cactcttggg  720
gaacatggtt ccagatacta caccaagagt ttcaaaggat cggtagatgc tttccatgtc  780
aatacagttg atacaactgg tgccggtgat tcctttgttg gtgctttatt ggccaagatt  840
gtcgatgatc agtccatact tgaagatgaa ccaaggttaa gagaagtact aaagtttgca  900
aatgcatgtg gagctattac aactacccaa aagggagcaa ttccggccct tcccaaagag  960
gaggctgcac tgaaactgat caaagggggg tca                               993

SEQ ID NO: 584          moltype = AA  length = 331
FEATURE                 Location/Qualifiers
source                  1..331
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 584
MALNNGVPAT GTGLIVSFGE MLIDFVPTVS GVSLAEAPGF LKAPGGAPAN VAIAVSRLGG  60
KAAFVGKLGD DEFGHMLAGI LKENGVRADG INFDQGARTA LAFVTLRADG EREFMFYRNP  120
SADMLLKPEE LNLELIRSAK VFHYGSISLI VEPCRSAHLK AMEVAKESGC LLSYDPNLRL  180
PLWPSAEEAR KQILSIWEKA DLIKVSDAEL EFLTGSDKID DESALSLWHP NLKLLLVTLG  240
EHGSRYYTKS FKGSVDAFHV NTVDTTGAGD SFVGALLAKI VDDQSILEDE PRLREVLKFA  300
NACGAITTTQ KGAIPALPKE EAALKLIKGG S                                 331

SEQ ID NO: 585          moltype = AA  length = 19
FEATURE                 Location/Qualifiers
source                  1..19
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 585
APGGAPANVA IAVSRLGGK                                               19

SEQ ID NO: 586          moltype = AA  length = 27
FEATURE                 Location/Qualifiers
source                  1..27
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 586
GSVDAFHVNT VDTTGAGDSF VGALLAK                                      27

SEQ ID NO: 587          moltype = AA  length = 19
FEATURE                 Location/Qualifiers
source                  1..19
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 587
IVDDQSILED EPRLREVLK                                               19

SEQ ID NO: 588          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 588
VSDAELEFLT GSDK                                                    14

SEQ ID NO: 589          moltype = DNA  length = 759
FEATURE                 Location/Qualifiers
source                  1..759
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 589
atgggcagaa agttcttcgt tggtggcaac tggaaatgca atgggaccac tgaggaggtg  60
aagaagattg ttactacttt aaatgaagct aaagtccctg gagaagatgt tgtagaagtt  120
gttgtgagcc ctccttttgt gttccttcct tttgtaaaaa gtttgctgcg ccctgatttc  180
catgtctcgg cccaaaattg ttgggttcgc aaaggtggtg cttatactgg agaggttagt  240
gctgaaatgc ttgttaattt gggaattcct tgggttatta ttggtcactc tgaacggagg  300
cagcttttga atgaatcaaa tgagtttgtg ggagataaag ttgcctatgc acttcaacaa  360
ggtctgaaag ttatagcatg cattggggaa actcttgaac agcgtgaagc tggtacaaca  420
acggctgttg ttgctgagca aacaaaagca attgcagcta aaatatcaaa ttgggacaat  480
gtcgttttgg cctatgagcc agtttgggcc attggaacag gaaaggttgc aactcctgct  540
caggctcaag aggttcatgc tgatttaagg aaatgggttc atgacaatgt gagtgctgaa  600
gttgctgcat ctgtaagaat tatctatgga ggctctgtaa atggaggaaa ctgcaaagaa  660
ttggcagcac agcccgatgt tgatggattt ttggttggtg gtgcatccct caaggcggaa  720
tttgtggaca tcataaacgc tgccactgtg aagaagaat                         759

SEQ ID NO: 590          moltype = AA  length = 253
FEATURE                 Location/Qualifiers
```

-continued

```
source                  1..253
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 590
MGRKFFVGGN WKCNGTTEEV KKIVTTLNEA KVPGEDVVEV VVSPPFVFLP FVKSLLRPDF  60
HVSAQNCWVR KGGAYTGEVS AEMLVNLGIP WVIIGHSERR QLLNESNEFV GDKVAYALQQ  120
GLKVIACIGE TLEQREAGTT TAVVAEQTKA IAAKISNWDN VVLAYEPVWA IGTGKVATPA  180
QAQEVHADLR KWVHDNVSAE VAASVRIIYG GSVNGGNCKE LAAQPDVDGF LVGGASLKAE  240
FVDIINAATV KKN                                                     253

SEQ ID NO: 591          moltype = AA  length = 13
FEATURE                 Location/Qualifiers
source                  1..13
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 591
AEFVDIINAA TVK                                                     13

SEQ ID NO: 592          moltype = AA  length = 19
FEATURE                 Location/Qualifiers
source                  1..19
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 592
ELAAQPDVDG FLVGGASLK                                               19

SEQ ID NO: 593          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 593
FFVGGNWK                                                           8

SEQ ID NO: 594          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 594
IVTTLNEAK                                                          9

SEQ ID NO: 595          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 595
VATPAQAQEV HADLRK                                                  16

SEQ ID NO: 596          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 596
VAYALQQGLK                                                         10

SEQ ID NO: 597          moltype = AA  length = 22
FEATURE                 Location/Qualifiers
source                  1..22
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 597
VPGEDVVEVV VSPPFVFLPF VK                                           22

SEQ ID NO: 598          moltype = DNA  length = 759
FEATURE                 Location/Qualifiers
source                  1..759
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 598
atgggcagaa aattcttcgt tggtggcaac tggaaatgca atgggaccac tgaggaggta  60
aagaagattg ttactacttt gaatgaggct aaagtccctg gagaagatgt cgtagaagtt  120
gttgtgagcc ctccttttgt gttccttcct gttgtaaaaa gtttgctgcg ccctgatttc  180
catgtttcgg cacaaaactg ttgggttcgc aaaggtggtg cttataccgg tgaggttagt  240
```

-continued

```
gctgaaatgc ttgttaattt gggaattcct tgggttatta ttggtcactc tgaacggagg  300
cagcttttaa atgaatcaaa tgagtttgtg ggagataaag ttgcctatgc acttcaacaa  360
ggtctaaaag ttattgcatg cattggggag actctcgaac agcgtgaagc tggtacaaca  420
acggctgttg tttctgagca aacaaaagca attgcagcta aaatatcaaa ttgggacaat  480
gtcgttttgg cctacgagcc agtttgggcc attggaacag gaaaggttgc tactcctgct  540
caggctcaag aggtccatgc tgatttgagg aaatgggttc atgacaatgt gagtgctgaa  600
gttgctgcat ctgtaagaat tatctatgga ggttctgtaa atggaggaaa ctgcaaagaa  660
ttggccgcac agcccgatgt tgatggattt ttggttggtg gtgcctccct gaagccggag  720
ttcgtggaca tcataaatgc tgccactgtg aagaagaat                         759

SEQ ID NO: 599          moltype = AA  length = 253
FEATURE                 Location/Qualifiers
source                  1..253
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 599
MGRKFFVGGN WKCNGTTEEV KKIVTTLNEA KVPGEDVVEV VVSPPFVFLP VVKSLLRPDF  60
HVSAQNCWVR KGGAYTGEVS AEMLVNLGIP WVIIGHSERR QLLNESNEFV GDKVAYALQQ  120
GLKVIACIGE TLEQREAGTT TAVVSEQTKA IAAKISNWDN VVLAYEPVWA IGTGKVATPA  180
QAQEVHADLR KWVHDNVSAE VAASVRIIYG GSVNGGNCKE LAAQPDVDGF LVGGASLKPE  240
FVDIINAATV KKN                                                     253

SEQ ID NO: 600          moltype = AA  length = 32
FEATURE                 Location/Qualifiers
source                  1..32
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 600
ELAAQPDVDG FLVGGASLKP EFVDIINAAT VK                                32

SEQ ID NO: 601          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 601
FFVGGNWK                                                           8

SEQ ID NO: 602          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 602
IVTTLNEAK                                                          9

SEQ ID NO: 603          moltype = AA  length = 16
FEATURE                 Location/Qualifiers
source                  1..16
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 603
VATPAQAQEV HADLRK                                                  16

SEQ ID NO: 604          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 604
VAYALQQGLK                                                         10

SEQ ID NO: 605          moltype = DNA  length = 765
FEATURE                 Location/Qualifiers
source                  1..765
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 605
atggcctcat tgggtgggtt acaaaatgtg agcggcatca attttcttat caaagagaaa  60
ggatgttgta tgctgcaggg tccaaaagtc aatgcaaagt tcgagctgaa gccgccacca  120
tatccactga atggtttgga gccggtgatg agccagcaga cacttgagtt tcactggggg  180
aagcaccaca agacttatgt ggaaaatctg aaaaaacaag ttgttgggac agagcttgat  240
gggaagtcac tagaagagat tattgtcaca tcatacaata agggtgacat tcttccagct  300
ttcaacaatg cagcacaggt atggaaccat gacttcttct gggagtgcat gaaaccaggt  360
ggaggtggaa agccatccgg ggagcttcta gaactgattg aaagagactt tggttcattt  420
gtaaaattcc ttgatgagtt caaggctgct gctgcaacac aatttggttc agggtgggct  480
```

-continued

```
tggctagcat atagagcaag aaaatttgat ggggaaaatg tagcaaatcc tccttcaccc  540
gatgaggaca acaagctagt ggtgctcaag agtcccaatg ctgtgaaccc ccttgtttgg  600
ggaggttact acccacttct taccattgat gtttgggagc atgcttacta ccttgatttt  660
cagaaccggc gtcctgatta tatatcagtg ttcatggata agcttgtttc ctgggatgca  720
gtgagctcta gacttgaaca agctaaggct ttaattacca gtgca                  765

SEQ ID NO: 606          moltype = AA  length = 255
FEATURE                 Location/Qualifiers
source                  1..255
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 606
MASLGGLQNV SGINFLIKEK GCCMLQGPKV NAKFELKPPP YPLNGLEPVM SQQTLEFHWG  60
KHHKTYVENL KKQVVGTELD GKSLEEIIVT SYNKGDILPA FNNAAQVWNH DFFWECMKPG  120
GGGKPSGELL ELIERDFGSF VKFLDEFKAA AATQFGSGWA WLAYRARKFD GENVANPPSP  180
DEDNKLVVLK SPNAVNPLVW GGYYPLLTID VWEHAYYLDF QNRRPDYISV FMDKLVSWDA  240
VSSRLEQAKA LITSA                                                   255

SEQ ID NO: 607          moltype = AA  length = 17
FEATURE                 Location/Qualifiers
source                  1..17
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 607
FDGENVANPP SPDEDNK                                                 17

SEQ ID NO: 608          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 608
LVSWDAVSSR LEQAK                                                   15

SEQ ID NO: 609          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 609
QVVGTELDGK                                                         10

SEQ ID NO: 610          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 610
SLEEIIVTSY NK                                                      12

SEQ ID NO: 611          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 611
TYVENLK                                                            7

SEQ ID NO: 612          moltype = DNA  length = 732
FEATURE                 Location/Qualifiers
source                  1..732
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 612
atggcctcat tgggtgggtt acaaaatcta agcggcatca attttcttat caaaaaagag  60
ggtccaaaag tcaatgcgaa gttcgagctg aagccgccac catatccact gaatggtttg  120
gagccggtga tgagccagca gacacttgag tttcattggg ggaaacacca caagacatat  180
gtggaaaatc tgaaaaaaca aattgttggg acagagcttg atgggaaatc actagaagag  240
attatagtga catcatataa taagggtgac attcttccag ctttcaacaa tgcagcacag  300
gtatggaacc atgacttctt ctgggagtgc atgaaaccag gtggaggtgg aaagccatcc  360
ggggagcttc tagaactgat tgaaagagac tttggctcat ttgaaaaatt ccttgatgag  420
tttaaggctg ctgctgcaac acaatttggt tcaggatggg cttggcttgc atatagagca  480
agcaaatttg atggggaaaa tgcagcaaat cctccttcac cagatgagga caacaagcta  540
gtggtgctca agagtcccaa tgctgtgaac ccccttgttt ggggaggcta ttatccactt  600
cttaccattg atgtttggga gcatgcttac aaccttgatt ttcagaaccg gcgccctgat  660
tatatatcag tgttcatgga taagcttgtt tcctgggatg cagtgagttc tagacttgaa  720
```

-continued

```
caagcaaggc tt                                                        732

SEQ ID NO: 613          moltype = AA  length = 244
FEATURE                 Location/Qualifiers
source                  1..244
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 613
MASLGGLQNL SGINFLIKKE GPKVNAKFEL KPPPYPLNGL EPVMSQQTLE FHWGKHHKTY  60
VENLKKQIVG TELDGKSLEE IIVTSYNKGD ILPAFNNAAQ VWNHDFFWEC MKPGGGGKPS  120
GELLELIERD FGSFEKFLDE FKAAAATQFG SGWAWLAYRA SKFDGENAAN PPSPDEDNKL  180
VVLKSPNAVN PLVWGGYYPL LTIDVWEHAY NLDFQNRRPD YISVFMDKLV SWDAVSSRLE  240
QARL                                                                 244

SEQ ID NO: 614          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
source                  1..10
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 614
QIVGTELDGK                                                           10

SEQ ID NO: 615          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 615
SLEEIIVTSY NK                                                        12

SEQ ID NO: 616          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
source                  1..7
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 616
TYVENLK                                                              7

SEQ ID NO: 617          moltype = DNA  length = 1419
FEATURE                 Location/Qualifiers
source                  1..1419
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 617
atggcttcta aggtggtttc ggttttggtt atagcgatga tgttgtttgc catgaactgt  60
aactgcacta gtgttggcca catgccttca accaaagagg aaggtcacga ttttcaggag  120
tctaaagcca agaccacaca aactgcgaac aaggccatgg aaactggtaa agagggacaa  180
gaagccgcag agtcatggac agaatgggct aaagagaaac tcagtgaagg tcttggcttc  240
aaacatgatc aggaatccaa agaatccacc actaacaaag tctctgatta tgccactgac  300
actgcacaga aatccaagga ttatgcaact gacactgctc agaaatcaaa ggactatgca  360
ggtgatgctg ctcagaaatc aaaggactac gcaggtgatg ctgctcagaa aaccaaagac  420
tatgccagtg acactgcaca gacatcaaag gactatgctg gtgatgctgc tcagaaatca  480
aagggctatg ttggtgatgc tgcacaaaag accaaagaat atgttggtga tgcagcacag  540
aagacaaagg attatgccac tgacgcagca cagaagacca aagattatgc tactcagaag  600
accaaggact atgcaagtga tgcaactgat gctgcaaaga agactaaaga ttatgctgct  660
cagaagacta aggactatgc aagtgaagct agtgatgttg cacagaacac aaaagattat  720
gctgctcaga agaccaagga ctatgccagc ggtggtgcgc agaagaccaa ggattatgcc  780
agcggtggtg cgcagaagac caaggactat gcaagtgatg ctgcgcagaa gaccaaggac  840
tatgcaagtg atggtgctca aaagagcaaa gaatatgcgg gtgatgtagc actgaatgcc  900
aaagattatg ctcaaaaaag caaagattac gcaggtgatg ctgcacagaa tgtcaaagat  960
tatgcaagtg atgcagtgca gaagagaaaa gagtattctg gtgatgcttc tcacaagagc  1020
aaagaggctt ctgattatgc cagtgagacg gctaagaaga ccaaagatta tgtcggtgat  1080
gctgctcaga gaagcaaagg ggctgctgaa tatgctagtg atgcggctca gaggaccaaa  1140
gagtatgctg gtgatgctac caagagaagc aaagaggctt ctaatgatca tgccaatgac  1200
atggctcaga agaccaaaga ctacgccagt gacactgctc agagaaccaa ggaaaaactc  1260
caagacattg catctgaagc tggtcaatac agtgcagaga aagctagaga aatgaaggat  1320
gcagcagcag agaaggcaag tgacatagca aaggctgcaa agcaaaaatc acaagaggtg  1380
aaagagaaat taggtggaca acaccgggat gccgagctt                          1419

SEQ ID NO: 618          moltype = AA  length = 473
FEATURE                 Location/Qualifiers
source                  1..473
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 618
MASKVVSVLV IAMMLFAMNC NCTSVGHMPS TKEEGHDFQE SKAKTTQTAN KAMETGKEGQ  60
```

-continued

```
EAAESWTEWA KEKLSEGLGF KHDQESKEST TNKVSDYATD TAQKSKDYAT DTAQKSKDYA  120
GDAAQKSKDY AGDAAQKTKD YASDTAQTSK DYAGDAAQKS KGYVGDAAQK TKEYVGDAAQ  180
KTKDYATDAA QKTKDYATQK TKDYASDATD AAKKTKDYAA QKTKDYASEA SDVAQNTKDY  240
AAQKTKDYAS GGAQKTKDYA SGGAQKTKDY ASDAAQKTKD YASDGAQKSK EYAGDVALNA  300
KDYAQKSKDY AGDAAQNVKD YASDAVQKRK EYSGDASHKS KEASDYASET AKKTKDYVGD  360
AAQRSKGAAE YASDAAQRTK EYAGDATKRS KEASNDHAND MAQKTKDYAS DTAQRTKEKL  420
QDIASEAGQY SAEKAREMKD AAAEKASDIA KAAKQKSQEV KEKLGGQHRD AEL         473

SEQ ID NO: 619          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 619
DYAGDAAQNV K                                                        11

SEQ ID NO: 620          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 620
DYASDAAQK                                                           9

SEQ ID NO: 621          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 621
DYASDATDAA K                                                        11

SEQ ID NO: 622          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 622
DYASDAVQK                                                           9

SEQ ID NO: 623          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 623
DYASDTAQRT K                                                        11

SEQ ID NO: 624          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 624
DYASDTAQTS K                                                        11

SEQ ID NO: 625          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 625
DYASEASDVA QNTK                                                     14

SEQ ID NO: 626          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 626
DYATDAAQK                                                           9

SEQ ID NO: 627          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
```

-continued

```
                        organism = Glycine max
SEQUENCE: 627
DYVGDAAQRS K                                                    11

SEQ ID NO: 628          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 628
EASDYASETA K                                                    11

SEQ ID NO: 629          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 629
EYAGDVALNA K                                                    11

SEQ ID NO: 630          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 630
EYVGDAAQK                                                       9

SEQ ID NO: 631          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 631
GAAEYASDAA QRTK                                                 14

SEQ ID NO: 632          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 632
GYVGDAAQK                                                       9

SEQ ID NO: 633          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 633
LQDIASEAGQ YSAEK                                                15

SEQ ID NO: 634          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 634
LSEGLGFK                                                        8

SEQ ID NO: 635          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 635
VSDYATDTAQ K                                                    11

SEQ ID NO: 636          moltype = DNA  length = 1374
FEATURE                 Location/Qualifiers
source                  1..1374
                        mol_type = other DNA
                        organism = Glycine max
SEQUENCE: 636
atggcttcga gggtggtttc ggttttggtg atagcgatga tgttgtttgc catgaactgt  60
aactgcacta gtgttggcca catgccttca accaaagagg aaggtcacga ttttcaggag  120
```

-continued

```
gctaaagcca agacaacaca aactgcgaac aaggccatgg aaactggtaa agatgggaaa  180
gaagccgcag agtcatggac agaatgggct aaagagaaac tcactgaagg tcttggcttc  240
aaacatgatc aagaatccac cactaagaaa gtctctgact atgccggtga tgctgctcag  300
aaatcaaagg actatgccgg tgacactgct cagaaaacca aggactatgc cggtgacact  360
gctcagaaaa ccaaggacta tgccggtgac actgcacaga catcaaagga ctatgccggt  420
gatgctgctc agaaatcaaa gggctatgtc ggtgatgctg cacaaaagac caaagaatat  480
gttggtgatg ctgctcagaa gaccatggat tatgccactg atgcagcaca gaagaccaaa  540
gattatgcta ctcagaagac taaggactat gcaagtgatg caactgatgc tgcaaaaaag  600
actaaagatt atgctgctca gaagaccaag gattatgcaa gtgatgcaac tgatgctgca  660
aaaaagacta aggattatgc tgcgcagaag accaaggact atgccagtgg tggtgcgcag  720
aagaccaagg actatgccag tggtgctgcg cagaagacca aggactatgc aagtgatgct  780
gcgcagagga ccaaggacca tgcaagtgat ggtgctcaaa agagcaaaga atatgcgggt  840
gatgtagcac agaatgccaa agattatgct caaaagagca aagattatgc gggtgatgct  900
gtacagaata tcaaagatta tgcaaatgat gcagcgcaca agagcaaaga ttatgctggt  960
gatgcttctc acaagagcaa agaagcttct gattatgcca gtgagacggc tcagaagacc  1020
aaagactatg ttggtgatgc tgctcagaaa agcaaagagg cttctgaata tgctagtgat  1080
gcggctcaga ggaccaaaga gtatgctggt gatgccacca agagaagcaa agaggcttct  1140
gatcatgcca atgacatggc tcggaagacc aaagactacg ccagtgacac tgctcagaga  1200
accaaggaaa aactccaaga cattgcatct gaggctggtc aatacagtac agagaaggcc  1260
agagaaatga aggatgcagc agcagagaag gctagtgaca ttgcaaaggc tgcaaagcaa  1320
aaatcacaag agctgaaaga caaattaggt ggacaacacc gggatgctga gctt        1374

SEQ ID NO: 637          moltype = AA  length = 458
FEATURE                 Location/Qualifiers
source                  1..458
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 637
MASRVVSVLV IAMMLFAMNC NCTSVGHMPS TKEEGHDFQE AKAKTTQTAN KAMETGKDGK  60
EAAESWTEWA KEKLTEGLGF KHDQESTTKK VSDYAGDAAQ KSKDYAGDTA QKTKDYAGDT  120
AQKTKDYAGD TAQTSKDYAG DAAQKSKGYV GDAAQKTKEY VGDAAQKTMD YATDAAQKTK  180
DYATQKTKDY ASDATDAAKK TKDYAAQKTK DYASDATDAA KKTKDYAAQK TKDYASGGAQ  240
KTKDYASGAA QKTKDYASDA AQRTKDHASD GAQKSKEYAG DVAQNAKDYA QKSKDYAGDA  300
VQNIKDYAND AAHKSKDYAG DASHKSKEAS DYASETAQKT KDYVGDAAQK SKEASEYASD  360
AAQRTKEYAG DATKRSKEAS DHANDMARKT KDYASDTAQR TKEKLQDIAS EAGQYSTEKA  420
REMKDAAAEK ASDIAKAAKQ KSQELKDKLG GQHRDAEL                          458

SEQ ID NO: 638          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 638
DYAGDAVQNI K                                                       11

SEQ ID NO: 639          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 639
DYASDAAQRT K                                                       11

SEQ ID NO: 640          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 640
DYASDATDAA K                                                       11

SEQ ID NO: 641          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 641
DYASDTAQRT K                                                       11

SEQ ID NO: 642          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 642
DYVGDAAQK                                                          9
```

-continued

```
SEQ ID NO: 643          moltype = AA  length = 12
FEATURE                 Location/Qualifiers
source                  1..12
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 643
EASDYASETA QK                                                    12

SEQ ID NO: 644          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
source                  1..14
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 644
EASEYASDAA QRTK                                                  14

SEQ ID NO: 645          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 645
EYAGDVAQNA K                                                     11

SEQ ID NO: 646          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
source                  1..9
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 646
GYVGDAAQK                                                        9

SEQ ID NO: 647          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
source                  1..15
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 647
LQDIASEAGQ YSTEK                                                 15

SEQ ID NO: 648          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
source                  1..8
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 648
LTEGLGFK                                                         8

SEQ ID NO: 649          moltype = AA  length = 11
FEATURE                 Location/Qualifiers
source                  1..11
                        mol_type = protein
                        organism = Glycine max
SEQUENCE: 649
VSDYAGDAAQ K                                                     11

SEQ ID NO: 650          moltype =   length =
SEQUENCE: 650
000

SEQ ID NO: 651          moltype =   length =
SEQUENCE: 651
000

SEQ ID NO: 652          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 652
STKSEGKL                                                         8

SEQ ID NO: 653          moltype = AA  length = 9
```

-continued

```
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
VARIANT                 1..3
                        note = any amino acid
VARIANT                 8..9
                        note = any amino acid
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 653
XXXKDNVXX                                                       9

SEQ ID NO: 654          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 3
                        note = any amino acid
VARIANT                 8..9
                        note = any amino acid
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 654
XLXKDNVXX                                                       9

SEQ ID NO: 655          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 655
ALSKDNVDD                                                       9

SEQ ID NO: 656          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 6..10
                        note = any amino acid
VARIANT                 12..15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 656
EEEDDXXXXX EXXXX                                                15

SEQ ID NO: 657          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 657
EEEDDTEVEA ESDAK                                                15

SEQ ID NO: 658          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 658
LHLKEEAGEY LEESK                                                15

SEQ ID NO: 659          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
```

-continued

```
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 659
SLDISPEATV EEEDD                                                15

SEQ ID NO: 660          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 660
ESTDKEKKDA                                                      10

SEQ ID NO: 661          moltype =   length =
SEQUENCE: 661
000

SEQ ID NO: 662          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
VARIANT                 2..3
                        note = any amino acid
VARIANT                 6
                        note = any amino acid
VARIANT                 8..9
                        note = any amino acid
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 662
DXXDKXRXX                                                       9

SEQ ID NO: 663          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 663
DALDKIRFE                                                       9

SEQ ID NO: 664          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..4
                        note = any amino acid
VARIANT                 9..15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 664
XXXXKSKLXX XXXXX                                                15

SEQ ID NO: 665          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..4
                        note = any amino acid
VARIANT                 9..11
                        note = any amino acid
VARIANT                 13..15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 665
XXXXKSKLXX XPXXX                                                15
```

-continued

```
SEQ ID NO: 666          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 666
SLTDKSKLDA QPELF                                                15

SEQ ID NO: 667          moltype =   length =
SEQUENCE: 667
000

SEQ ID NO: 668          moltype = AA  length = 5
FEATURE                 Location/Qualifiers
REGION                  1..5
                        note = epitope
VARIANT                 5
                        note = any amino acid
source                  1..5
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 668
IPDKX                                                           5

SEQ ID NO: 669          moltype = AA  length = 5
FEATURE                 Location/Qualifiers
REGION                  1..5
                        note = epitope
source                  1..5
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 669
IPDKT                                                           5

SEQ ID NO: 670          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 670
DDEDEEEKKD EEGKV                                                15

SEQ ID NO: 671          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 671
PEYLSFVKGI VDSED                                                15

SEQ ID NO: 672          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 672
VDSEDLPLNI SREML                                                15

SEQ ID NO: 673          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..2
                        note = any amino acid
VARIANT                 6..12
                        note = any amino acid
VARIANT                 15
```

-continued

```
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 673
XXFSKXXXXX XXEDX                                                15

SEQ ID NO: 674          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..2
                        note = any amino acid
VARIANT                 6..12
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 674
XXFSKXXXXX XXEDS                                                15

SEQ ID NO: 675          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 675
EAFSKNLKLG IHEDS                                                15

SEQ ID NO: 676          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 676
LGDKVEKVVV SDRVV                                                15

SEQ ID NO: 677          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 677
DEDAGEADAD MPPLE                                                15

SEQ ID NO: 678          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 678
MPPLEDADAD AEGSK                                                15

SEQ ID NO: 679          moltype =   length =
SEQUENCE: 679
000

SEQ ID NO: 680          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..4
                        note = any amino acid
VARIANT                 9
                        note = any amino acid
VARIANT                 11..13
                        note = any amino acid
source                  1..15
```

-continued

```
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 680
XXXXDKIRXE XXXDK                                                15

SEQ ID NO: 681      moltype = AA  length = 15
FEATURE             Location/Qualifiers
REGION              1..15
                    note = epitope
source              1..15
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 681
SDALDKIRFE SLTDK                                                15

SEQ ID NO: 682      moltype = AA  length = 15
FEATURE             Location/Qualifiers
REGION              1..15
                    note = epitope
VARIANT             1
                    note = any amino acid
VARIANT             3..4
                    note = any amino acid
VARIANT             6..9
                    note = any amino acid
VARIANT             12..15
                    note = any amino acid
source              1..15
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 682
XYXXRXXXXQ NXXXX                                                15

SEQ ID NO: 683      moltype = AA  length = 15
FEATURE             Location/Qualifiers
REGION              1..15
                    note = epitope
VARIANT             1
                    note = any amino acid
VARIANT             3..4
                    note = any amino acid
VARIANT             6..9
                    note = any amino acid
VARIANT             12..15
                    note = any amino acid
source              1..15
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 683
XYXXRXXXXQ NXXXX                                                15

SEQ ID NO: 684      moltype = AA  length = 15
FEATURE             Location/Qualifiers
REGION              1..15
                    note = epitope
source              1..15
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 684
DYVTRMKEGQ NDIYY                                                15

SEQ ID NO: 685      moltype = AA  length = 15
FEATURE             Location/Qualifiers
REGION              1..15
                    note = epitope
VARIANT             2..4
                    note = any amino acid
VARIANT             6..9
                    note = any amino acid
VARIANT             12..15
                    note = any amino acid
source              1..15
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 685
DXXXRXXXXQ NXXXX                                                15
```

-continued

```
SEQ ID NO: 686          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 2..4
                        note = any amino acid
VARIANT                 6..9
                        note = any amino acid
VARIANT                 12..15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 686
DXXXRXXXXQ NXXXX                                                15

SEQ ID NO: 687          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 687
DYVTRMKEGQ NDIYY                                                15

SEQ ID NO: 688          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 4..11
                        note = any amino acid
VARIANT                 14..15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 688
RKPXXXXXXX XAAXX                                                15

SEQ ID NO: 689          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 689
RKPEEITKEE YAAFY                                                15

SEQ ID NO: 690          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 690
SLTDKSKLD                                                       9

SEQ ID NO: 691          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 691
FIHIIPDK                                                        8

SEQ ID NO: 692          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
```

-continued

```
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 692
VTRMKEGQ                                                        8

SEQ ID NO: 693          moltype =   length =
SEQUENCE: 693
000

SEQ ID NO: 694          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..3
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 694
XXXLKREDEE RLDEV                                                15

SEQ ID NO: 695          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 695
GEPLKREDEE RLDEV                                                15

SEQ ID NO: 696          moltype =   length =
SEQUENCE: 696
000

SEQ ID NO: 697          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 697
ILERKKSPNR LVVDE                                                15

SEQ ID NO: 698          moltype =   length =
SEQUENCE: 698
000

SEQ ID NO: 699          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 699
FPESGDR                                                         7

SEQ ID NO: 700          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 700
SDPKSGK                                                         7

SEQ ID NO: 701          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
```

-continued

```
SEQUENCE: 701
QTMEKLQL                                                        8

SEQ ID NO: 702          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 702
LGDVVSV                                                         7

SEQ ID NO: 703          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 703
GGMRSVEFKV VETDP                                                15

SEQ ID NO: 704          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 704
RKAFEEAEKN APSII                                                15

SEQ ID NO: 705          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 705
KREKTHGE                                                        8

SEQ ID NO: 706          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 706
RIVSQLL                                                         7

SEQ ID NO: 707          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 707
SRAHVIVIGA TNRPN                                                15

SEQ ID NO: 708          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 708
LRRFGRFDR                                                       9

SEQ ID NO: 709          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
```

-continued

```
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 709
MDVIDLEDET IDAEV                                                15

SEQ ID NO: 710          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 710
VTNEHFQTAL GTSNP                                                15

SEQ ID NO: 711          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 711
VEHPEKFEKF GMSPS                                                15

SEQ ID NO: 712          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 712
EIFDKARQ                                                        8

SEQ ID NO: 713          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 713
DIIDPALLRP GRLDQ                                                15

SEQ ID NO: 714          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 714
IRENIEKDIE                                                      10

SEQ ID NO: 715          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 715
IERERKSREN PEAMD                                                15

SEQ ID NO: 716          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 716
PEAMDEDTVD DEVAE                                                15
```

-continued

```
SEQ ID NO: 717          moltype =   length =
SEQUENCE: 717
000

SEQ ID NO: 718          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 1..5
                        note = any amino acid
VARIANT                 7
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 718
XXXXXXEXKED                                                     10

SEQ ID NO: 719          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 719
LDKFAENKED                                                      10

SEQ ID NO: 720          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
VARIANT                 2..3
                        note = any amino acid
VARIANT                 7
                        note = any amino acid
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 720
KXXLGDX                                                         7

SEQ ID NO: 721          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
VARIANT                 3
                        note = any amino acid
VARIANT                 7
                        note = any amino acid
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 721
KIXLGDX                                                         7

SEQ ID NO: 722          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 722
KIRLGDT                                                         7

SEQ ID NO: 723          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 1..3
                        note = any amino acid
VARIANT                 5
                        note = any amino acid
VARIANT                 7
```

-continued

```
                        note = any amino acid
VARIANT                 9
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 723
XXXDXFXKXE                                                      10

SEQ ID NO: 724          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 1..3
                        note = any amino acid
VARIANT                 5
                        note = any amino acid
VARIANT                 7
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 724
XXXDXFXKIE                                                      10

SEQ ID NO: 725          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 725
GDTDLFAKIE                                                      10

SEQ ID NO: 726          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 726
IDSCLTVADQ YDIQI                                                15

SEQ ID NO: 727          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 727
NPTRPLTLNT IDEHL                                                15

SEQ ID NO: 728          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 728
CHHLNREIPE DLAFA                                                15

SEQ ID NO: 729          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..3
                        note = any amino acid
VARIANT                 5
                        note = any amino acid
VARIANT                 7..9
                        note = any amino acid
VARIANT                 12..13
```

-continued

```
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 729
XXXHXLXXXN KXXEA                                                15

SEQ ID NO: 730          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..2
                        note = any amino acid
VARIANT                 5
                        note = any amino acid
VARIANT                 7..9
                        note = any amino acid
VARIANT                 12..13
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 730
XXVHXLXXXN KXXEA                                                15

SEQ ID NO: 731          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 731
QRVHALYGLN KRVEA                                                15

SEQ ID NO: 732          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..6
                        note = any amino acid
VARIANT                 9
                        note = any amino acid
VARIANT                 11
                        note = any amino acid
VARIANT                 13
                        note = any amino acid
VARIANT                 15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 732
XXXXXXGLXK XRXEX                                                15

SEQ ID NO: 733          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..6
                        note = any amino acid
VARIANT                 9
                        note = any amino acid
VARIANT                 11
                        note = any amino acid
VARIANT                 13
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 733
XXXXXXGLXK XRXEG                                                15

SEQ ID NO: 734          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
```

-continued

```
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 734
KAHEIAGLRK ERAEG                                                15

SEQ ID NO: 735          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 735
ERAEGKFVPD PRFEE                                                15

SEQ ID NO: 736          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 736
RSGIFGSYNY DELMG                                                15

SEQ ID NO: 737          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 737
DELMGSLEGN EGFGR                                                15

SEQ ID NO: 738          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 738
LNTAGSYKFS SDRTI                                                15

SEQ ID NO: 739          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 739
AAEKYKSEGH DMIIL                                                15

SEQ ID NO: 740          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 740
DSLGLTGHER YTIDL                                                15

SEQ ID NO: 741          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 741
IRKVVELQEE LDIDV                                                15
```

-continued

```
SEQ ID NO: 742          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..7
                        note = any amino acid
VARIANT                 10
                        note = any amino acid
VARIANT                 14
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 742
XXXXXXXKDX VEDXE                                                15

SEQ ID NO: 743          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..7
                        note = any amino acid
VARIANT                 14
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 743
XXXXXXXKDE VEDXE                                                15

SEQ ID NO: 744          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 744
YQIALSIKDE VEDLE                                                15

SEQ ID NO: 745          moltype =   length =
SEQUENCE: 745
000

SEQ ID NO: 746          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 2
                        note = any amino acid
VARIANT                 4..7
                        note = any amino acid
VARIANT                 10
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 746
NXEXXXXKDX                                                      10

SEQ ID NO: 747          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 747
NPETYILKDQ                                                      10

SEQ ID NO: 748          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
VARIANT                 6
```

-continued

```
                        note = any amino acid
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 748
LDFLEXET                                                        8

SEQ ID NO: 749          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 749
LDFLEVET                                                        8

SEQ ID NO: 750          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 750
QYTEGDKRV                                                       9

SEQ ID NO: 751          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 3
                        note = any amino acid
VARIANT                 6..10
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 751
PLXQQXXXXX                                                      10

SEQ ID NO: 752          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 3
                        note = any amino acid
VARIANT                 7..10
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 752
PLXQQGXXXX                                                      10

SEQ ID NO: 753          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 753
PLVQQGQQGQ                                                      10

SEQ ID NO: 754          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 754
LGEMNGKM                                                        8

SEQ ID NO: 755          moltype = AA  length = 15
```

-continued

```
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 755
KVITLSVEDQ DYMGR                                                    15

SEQ ID NO: 756          moltype =   length =
SEQUENCE: 756
000

SEQ ID NO: 757          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..2
                        note = any amino acid
VARIANT                 4
                        note = any amino acid
VARIANT                 6..9
                        note = any amino acid
VARIANT                 11
                        note = any amino acid
VARIANT                 14
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 757
XXGXLXXXXG XRTXE                                                    15

SEQ ID NO: 758          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 758
ASGKLSQYEG GRTKE                                                    15

SEQ ID NO: 759          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..5
                        note = any amino acid
VARIANT                 8..11
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 759
XXXXXLLXXX XEEDN                                                    15

SEQ ID NO: 760          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..5
                        note = any amino acid
VARIANT                 9..11
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 760
XXXXXLLEXX XEEDN                                                    15

SEQ ID NO: 761          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
```

-continued

```
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 761
EKSDNLLEGG GEEDN                                                15

SEQ ID NO: 762          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 3..11
                        note = any amino acid
VARIANT                 13
                        note = any amino acid
VARIANT                 15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 762
RKXXXXXXXX XYXNX                                                15

SEQ ID NO: 763          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 6..10
                        note = any amino acid
VARIANT                 15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 763
RKDHSXXXXX LYANX                                                15

SEQ ID NO: 764          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 764
RKDHSDVSNQ LYANY                                                15

SEQ ID NO: 765          moltype =   length =
SEQUENCE: 765
000

SEQ ID NO: 766          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 766
QGKEYVFVAN SDNLG                                                15

SEQ ID NO: 767          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 767
IAKYGLDDNT IDFIG                                                15

SEQ ID NO: 768          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
```

-continued

```
SEQUENCE: 768
EETREKKEEE SRERE                                                15

SEQ ID NO: 769          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..7
                        note = any amino acid
VARIANT                 9..10
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 769
XXXXXXXNXX IFEED                                                15

SEQ ID NO: 770          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..7
                        note = any amino acid
VARIANT                 9
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 770
XXXXXXXNXY IFEED                                                15

SEQ ID NO: 771          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 771
QHEEQDENPY IFEED                                                15

SEQ ID NO: 772          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 772
SIFAISREQV RALAP                                                15

SEQ ID NO: 773          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 7
                        note = any amino acid
VARIANT                 9..10
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 773
XELPREXRXX                                                      10

SEQ ID NO: 774          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 9..10
                        note = any amino acid
source                  1..10
```

-continued

```
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 774
XELPREERXX                                                          10

SEQ ID NO: 775      moltype = AA  length = 10
FEATURE             Location/Qualifiers
REGION              1..10
                    note = epitope
source              1..10
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 775
FELPREERGR                                                          10

SEQ ID NO: 776      moltype = AA  length = 15
FEATURE             Location/Qualifiers
REGION              1..15
                    note = epitope
VARIANT             1..5
                    note = any amino acid
VARIANT             9..12
                    note = any amino acid
VARIANT             15
                    note = any amino acid
source              1..15
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 776
XXXXXPLYXX XXDEX                                                    15

SEQ ID NO: 777      moltype = AA  length = 15
FEATURE             Location/Qualifiers
REGION              1..15
                    note = epitope
VARIANT             1..5
                    note = any amino acid
VARIANT             10..12
                    note = any amino acid
VARIANT             15
                    note = any amino acid
source              1..15
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 777
XXXXXPLYIX XXDEX                                                    15

SEQ ID NO: 778      moltype = AA  length = 15
FEATURE             Location/Qualifiers
REGION              1..15
                    note = epitope
source              1..15
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 778
IPAGTPLYIV NRDEN                                                    15

SEQ ID NO: 779      moltype =   length =
SEQUENCE: 779
000

SEQ ID NO: 780      moltype = AA  length = 15
FEATURE             Location/Qualifiers
REGION              1..15
                    note = epitope
VARIANT             1..5
                    note = any amino acid
VARIANT             7..8
                    note = any amino acid
VARIANT             10
                    note = any amino acid
VARIANT             13..15
                    note = any amino acid
source              1..15
                    mol_type = protein
                    organism = synthetic construct
```

-continued

```
SEQUENCE: 780
XXXXXSXXRX RRXXX                                                15

SEQ ID NO: 781          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 781
PFELPSEERG RRAVA                                                15

SEQ ID NO: 782          moltype =   length =
SEQUENCE: 782
000

SEQ ID NO: 783          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 9..10
                        note = any amino acid
VARIANT                 15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 783
XKEEEHQEXX EEEQX                                                15

SEQ ID NO: 784          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 784
EKEEEHQEQH EEEQD                                                15

SEQ ID NO: 785          moltype =   length =
SEQUENCE: 785
000

SEQ ID NO: 786          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 786
FAGKDNIVSS LDNVA                                                15

SEQ ID NO: 787          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 787
SNLALGKLKE                                                      10

SEQ ID NO: 788          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 788
EGGEKGNKDR PELKT                                                15
```

-continued

```
SEQ ID NO: 789          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 789
LNTTTQTA                                                          8

SEQ ID NO: 790          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 790
QQGYAVTKDT ISSAA                                                  15

SEQ ID NO: 791          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 791
KKPSQPQE                                                          8

SEQ ID NO: 792          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 792
IGEKLGDAKQ REELL                                                  15

SEQ ID NO: 793          moltype = AA  length = 5
FEATURE                 Location/Qualifiers
REGION                  1..5
                        note = epitope
source                  1..5
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 793
ETKDT                                                             5

SEQ ID NO: 794          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 794
VAYPLDLF                                                          8

SEQ ID NO: 795          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 795
KDDENVNSQ                                                         9

SEQ ID NO: 796          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
```

-continued

```
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 796
CAEAIFKSQA ETGEI                                                15

SEQ ID NO: 797          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 797
GHYLNATAG                                                       9

SEQ ID NO: 798          moltype = AA  length = 5
FEATURE                 Location/Qualifiers
REGION                  1..5
                        note = epitope
source                  1..5
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 798
LHIHR                                                           5

SEQ ID NO: 799          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 799
DDFVEKDRSR GIYFT                                                15

SEQ ID NO: 800          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 800
WGNAPGA                                                         7

SEQ ID NO: 801          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 801
RNEGRDLARE GNEII                                                15

SEQ ID NO: 802          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 802
VWKEIKFEFE AMDTL                                                15

SEQ ID NO: 803          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 803
HPGDGASGRL SNATG                                                15

SEQ ID NO: 804          moltype = AA  length = 15
```

-continued

```
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 804
VIRIVSEQTD ILEPF                                                15

SEQ ID NO: 805          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 805
ATSGDRDGKT FVSTA                                                15

SEQ ID NO: 806          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 806
PVTVNLWQPE KNAFE                                                15

SEQ ID NO: 807          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 807
AIRVTQS                                                         7

SEQ ID NO: 808          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 808
SEARKSTNTP DAQKV                                                15

SEQ ID NO: 809          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 809
ILWHLTKEDK TYGVP                                                15

SEQ ID NO: 810          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 810
VAFSIDNRQI VSASR                                                15

SEQ ID NO: 811          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
```

-continued

```
SEQUENCE: 811
RLYSLDAGS                                                    9

SEQ ID NO: 812          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 812
DLKVDLKTE                                                    9

SEQ ID NO: 813          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 813
KSGEERAQEE AIDID                                             15

SEQ ID NO: 814          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 814
DNQLVAADAK LNFDD                                             15

SEQ ID NO: 815          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 815
FLDVGGNASE NQVVE                                             15

SEQ ID NO: 816          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 816
VRFYKEEEKN DPEFA                                             15

SEQ ID NO: 817          moltype =   length =
SEQUENCE: 817
000

SEQ ID NO: 818          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
VARIANT                 5..7
                        note = any amino acid
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 818
FDKFXXX                                                      7

SEQ ID NO: 819          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
```

-continued

```
SEQUENCE: 819
FDKFATG                                                          7

SEQ ID NO: 820          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 820
GKPLVIEEVE VAPPQ                                                 15

SEQ ID NO: 821          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 3..6
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 821
XQXXXXENKP                                                       10

SEQ ID NO: 822          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 3
                        note = any amino acid
VARIANT                 5..6
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 822
EQXQXXENKP                                                       10

SEQ ID NO: 823          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 823
EQHQHDENKP                                                       10

SEQ ID NO: 824          moltype =   length =
SEQUENCE: 824
000

SEQ ID NO: 825          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 4..6
                        note = any amino acid
VARIANT                 9..13
                        note = any amino acid
VARIANT                 15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 825
XKDXXXEKXX XXXDX                                                 15

SEQ ID NO: 826          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
```

-continued

```
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 826
YKDYTAEKAI EGKDS                                                15

SEQ ID NO: 827          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..3
                        note = any amino acid
VARIANT                 5..6
                        note = any amino acid
VARIANT                 10..13
                        note = any amino acid
VARIANT                 15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 827
XXXDXXREEX XXXRX                                                15

SEQ ID NO: 828          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..3
                        note = any amino acid
VARIANT                 5..6
                        note = any amino acid
VARIANT                 12..13
                        note = any amino acid
VARIANT                 15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 828
XXXDXXREEE EXXRX                                                15

SEQ ID NO: 829          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 829
GNKDQTREEE EETRR                                                15

SEQ ID NO: 830          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 830
KETKMGEYSD YATQK                                                15

SEQ ID NO: 831          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 831
QKAGEYTDYA AQKAK                                                15

SEQ ID NO: 832          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
```

-continued

```
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 832
TTVQKAGEYK DYTAE                                                15

SEQ ID NO: 833          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 833
KDTTVQKAGE YKDYT                                                15

SEQ ID NO: 834          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 834
DGDMEEEGVL HVERR                                                15

SEQ ID NO: 835          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 835
EEEGVLHVER RREKM                                                15

SEQ ID NO: 836          moltype =   length =
SEQUENCE: 836
000

SEQ ID NO: 837          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 1..5
                        note = any amino acid
VARIANT                 7
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 837
XXXXXFXDKP                                                      10

SEQ ID NO: 838          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 838
EKTLLFGDKP                                                      10

SEQ ID NO: 839          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 839
KLKGILGYTE DDVVS                                                15

SEQ ID NO: 840          moltype =   length =
```

-continued

```
SEQUENCE: 840
000

SEQ ID NO: 841          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 2..11
                        note = any amino acid
VARIANT                 14
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 841
LXXXXXXXXX XDDXE                                                15

SEQ ID NO: 842          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 842
LEFTQGDLRN RDDLE                                                15

SEQ ID NO: 843          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 843
SVSIIDNFDN SVVEA                                                15

SEQ ID NO: 844          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 844
EKIPCEEDFR LQAMN                                                15

SEQ ID NO: 845          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 845
PELNVYGHDY PTRDG                                                15

SEQ ID NO: 846          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 846
PTRDGSAIRD YIHVM                                                15

SEQ ID NO: 847          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..4
                        note = any amino acid
VARIANT                 7..9
                        note = any amino acid
VARIANT                 12..14
```

-continued

```
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 847
XXXXDKXXXE RXXXE                                                15

SEQ ID NO: 848          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..4
                        note = any amino acid
VARIANT                 9
                        note = any amino acid
VARIANT                 12..14
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 848
XXXXDKRVXE RXXXE                                                15

SEQ ID NO: 849          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 849
LGGIDKRVIE RFEKE                                                15

SEQ ID NO: 850          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 850
GGFEAGISKD GQTRE                                                15

SEQ ID NO: 851          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
VARIANT                 4
                        note = any amino acid
VARIANT                 7..9
                        note = any amino acid
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 851
PISXFEXXX                                                       9

SEQ ID NO: 852          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
VARIANT                 7..9
                        note = any amino acid
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 852
PISGFEXXX                                                       9

SEQ ID NO: 853          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
```

-continued

```
SEQUENCE: 853
PISGFEGDN                                                          9

SEQ ID NO: 854          moltype =   length =
SEQUENCE: 854
000

SEQ ID NO: 855          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 4
                        note = any amino acid
VARIANT                 7..10
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 855
LLEXLDXXXX                                                         10

SEQ ID NO: 856          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 856
LLEALDQINE                                                         10

SEQ ID NO: 857          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 857
GTVEDIDEPR LERVF                                                   15

SEQ ID NO: 858          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 858
WTPLIPASFK EEETA                                                   15

SEQ ID NO: 859          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
VARIANT                 2..5
                        note = any amino acid
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 859
KXXXXQQQ                                                           8

SEQ ID NO: 860          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
VARIANT                 2..4
                        note = any amino acid
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 860
KXXXQQQQ                                                           8

SEQ ID NO: 861          moltype = AA  length = 8
```

-continued

```
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 861
KFPPQQQQ                                                        8

SEQ ID NO: 862          moltype =   length =
SEQUENCE: 862
000

SEQ ID NO: 863          moltype =   length =
SEQUENCE: 863
000

SEQ ID NO: 864          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 864
RGKVEAE                                                         7

SEQ ID NO: 865          moltype =   length =
SEQUENCE: 865
000

SEQ ID NO: 866          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 5..10
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 866
RRGKXXXXXX                                                      10

SEQ ID NO: 867          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 867
RRGKESGGGG                                                      10

SEQ ID NO: 868          moltype =   length =
SEQUENCE: 868
000

SEQ ID NO: 869          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 869
RNKQANKGNE EGRRE                                                15

SEQ ID NO: 870          moltype =   length =
SEQUENCE: 870
000

SEQ ID NO: 871          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
```

-continued

```
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 871
MASEQMPRRE NNSNR                                                    15

SEQ ID NO: 872          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..3
                        note = any amino acid
VARIANT                 9
                        note = any amino acid
VARIANT                 11..15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 872
XXXKGRVVXE XXXXX                                                    15

SEQ ID NO: 873          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 9
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 873
LPDKGRVVXE SEKKE                                                    15

SEQ ID NO: 874          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 874
LPDKGRVVIE SEKKE                                                    15

SEQ ID NO: 875          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 7..8
                        note = any amino acid
VARIANT                 11..15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 875
ATKAKDXXRE XXXXX                                                    15

SEQ ID NO: 876          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 11..15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 876
ATKAKDVTRE XXXXX                                                    15

SEQ ID NO: 877          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
```

-continued

```
SEQUENCE: 877
ATKAKDVTRE KGQQG                                                15

SEQ ID NO: 878          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 3
                        note = any amino acid
VARIANT                 7
                        note = any amino acid
VARIANT                 10..15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 878
XGXKGQXKKX XXXXX                                                15

SEQ ID NO: 879          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 3
                        note = any amino acid
VARIANT                 10..11
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 879
XGXKGQTKKX XGEEQ                                                15

SEQ ID NO: 880          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 880
AGTKGQTKKG GGEEQ                                                15

SEQ ID NO: 881          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 2..4
                        note = any amino acid
VARIANT                 10
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 881
RXXXETKEGX                                                      10

SEQ ID NO: 882          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 882
REKAETKEGG                                                      10

SEQ ID NO: 883          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 1..4
```

-continued

```
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 883
XXXXNKQSSL                                                      10

SEQ ID NO: 884          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 884
IGEENKQSSL                                                      10

SEQ ID NO: 885          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 3
                        note = any amino acid
VARIANT                 6..13
                        note = any amino acid
VARIANT                 15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 885
XKXVAXXXXX XXXEX                                                15

SEQ ID NO: 886          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 3
                        note = any amino acid
VARIANT                 7..9
                        note = any amino acid
VARIANT                 11
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 886
XKXVAEXXXQ XASEL                                                15

SEQ ID NO: 887          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 887
VKGVAEYAGQ KASEL                                                15

SEQ ID NO: 888          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 1..2
                        note = any amino acid
VARIANT                 8..10
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 888
XXKVEAEXXX                                                      10
```

-continued

```
SEQ ID NO: 889          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 889
RKKVEAEMDL                                                          10

SEQ ID NO: 890          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 1..4
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 890
XXXXIAEIAE                                                          10

SEQ ID NO: 891          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
VARIANT                 1..2
                        note = any amino acid
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 891
XXETIAEIAE                                                          10

SEQ ID NO: 892          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 892
IGETIAEIAE                                                          10

SEQ ID NO: 893          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 3..4
                        note = any amino acid
VARIANT                 6
                        note = any amino acid
VARIANT                 8..11
                        note = any amino acid
VARIANT                 15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 893
XYXXSXKXXX XEKVX                                                    15

SEQ ID NO: 894          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 4
                        note = any amino acid
VARIANT                 6
                        note = any amino acid
VARIANT                 8..11
```

-continued

```
                        note = any amino acid
VARIANT                 15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 894
XYIXSXKXXX XEKVX                                                15

SEQ ID NO: 895          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 895
LYIASVKANA VEKVE                                                15

SEQ ID NO: 896          moltype = AA  length = 5
FEATURE                 Location/Qualifiers
REGION                  1..5
                        note = epitope
VARIANT                 3
                        note = any amino acid
source                  1..5
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 896
YEXGV                                                           5

SEQ ID NO: 897          moltype = AA  length = 5
FEATURE                 Location/Qualifiers
REGION                  1..5
                        note = epitope
VARIANT                 3
                        note = any amino acid
source                  1..5
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 897
YEXGV                                                           5

SEQ ID NO: 898          moltype = AA  length = 5
FEATURE                 Location/Qualifiers
REGION                  1..5
                        note = epitope
source                  1..5
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 898
YEAGV                                                           5

SEQ ID NO: 899          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 899
AREAAARDAR EAAAR                                                15

SEQ ID NO: 900          moltype =   length =
SEQUENCE: 900
000

SEQ ID NO: 901          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
VARIANT                 2..5
                        note = any amino acid
VARIANT                 9
                        note = any amino acid
source                  1..9
```

-continued

```
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 901
VXXXXVTAX                                                       9

SEQ ID NO: 902          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 902
VKRTTVTAT                                                       9

SEQ ID NO: 903          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 903
AWGVEPYDLD DADRN                                                15

SEQ ID NO: 904          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..2
                        note = any amino acid
VARIANT                 4..10
                        note = any amino acid
VARIANT                 14..15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 904
XXWXXXXXXX EKEXX                                                15

SEQ ID NO: 905          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..2
                        note = any amino acid
VARIANT                 4..7
                        note = any amino acid
VARIANT                 9..10
                        note = any amino acid
VARIANT                 15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 905
XXWXXXXEXX EKESX                                                15

SEQ ID NO: 906          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 906
INWKYASEVY EKESS                                                15

SEQ ID NO: 907          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..4
                        note = any amino acid
VARIANT                 12
```

-continued

```
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 907
XXXXEEEREH AXQLI                                                15

SEQ ID NO: 908          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 908
KESSEEEREH AEQLI                                                15

SEQ ID NO: 909          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 909
KESSEEEREH AEQLI                                                15

SEQ ID NO: 910          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 910
PDGILAYLDE ENKPQ                                                15

SEQ ID NO: 911          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 911
CSLKHVPGFI ERAEE                                                15

SEQ ID NO: 912          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 912
ALGLELDLTD KGLGV                                                15

SEQ ID NO: 913          moltype =   length =
SEQUENCE: 913
000

SEQ ID NO: 914          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
VARIANT                 1
                        note = any amino acid
VARIANT                 5..7
                        note = any amino acid
VARIANT                 9
                        note = any amino acid
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 914
XDIDXXXDX                                                       9
```

-continued

```
SEQ ID NO: 915          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 915
WDIDYQLDL                                                       9

SEQ ID NO: 916          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 916
KADAGASKTY PQQAG                                                15

SEQ ID NO: 917          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 917
CDVPHVNRTD YQLID                                                15

SEQ ID NO: 918          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 918
FAEGKDLVVS VMSAM                                                15

SEQ ID NO: 919          moltype =   length =
SEQUENCE: 919
000

SEQ ID NO: 920          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
VARIANT                 1..4
                        note = any amino acid
VARIANT                 6
                        note = any amino acid
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 920
XXXXPXFDK                                                       9

SEQ ID NO: 921          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 921
KAEFPIFDK                                                       9

SEQ ID NO: 922          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
```

-continued

```
SEQUENCE: 922
WNFSKFLVDK DGNVV                                                15

SEQ ID NO: 923          moltype = AA  length = 5
FEATURE                 Location/Qualifiers
REGION                  1..5
                        note = epitope
source                  1..5
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 923
KSRDD                                                           5

SEQ ID NO: 924          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 924
PSQPAESKGE ESGGG                                                15

SEQ ID NO: 925          moltype =   length =
SEQUENCE: 925
000

SEQ ID NO: 926          moltype =   length =
SEQUENCE: 926
000

SEQ ID NO: 927          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 927
KLDDKQGIG                                                       9

SEQ ID NO: 928          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 928
SIEQKEESRG NEDHV                                                15

SEQ ID NO: 929          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 929
FKTGAERKEA AESTL                                                15

SEQ ID NO: 930          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 930
ELDTLGEESY KDSTL                                                15

SEQ ID NO: 931          moltype =   length =
SEQUENCE: 931
000

SEQ ID NO: 932          moltype = AA  length = 15
```

-continued

```
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..2
                        note = any amino acid
VARIANT                 4..6
                        note = any amino acid
VARIANT                 8..9
                        note = any amino acid
VARIANT                 11
                        note = any amino acid
VARIANT                 13..15
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 932
XXGXXXSXXE XLXXX                                                15

SEQ ID NO: 933          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 933
GTGLIVSFGE MLIDF                                                15

SEQ ID NO: 934          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 934
DEFGHMLAGI LKENG                                                15

SEQ ID NO: 935          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 935
SADMLLKPEE LNLEL                                                15

SEQ ID NO: 936          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 936
DLIKVSDAEL EFLTG                                                15

SEQ ID NO: 937          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 937
EFLTGSDKID DESAL                                                15

SEQ ID NO: 938          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
```

-continued

```
SEQUENCE: 938
VDDQSILEDE PRLRE                                                15

SEQ ID NO: 939          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 939
KGGAYTGEVS AEMLV                                                15

SEQ ID NO: 940          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 940
GGGKPSGELL ELIER                                                15

SEQ ID NO: 941          moltype =   length =
SEQUENCE: 941
000

SEQ ID NO: 942          moltype =   length =
SEQUENCE: 942
000

SEQ ID NO: 943          moltype = AA  length = 7
FEATURE                 Location/Qualifiers
REGION                  1..7
                        note = epitope
source                  1..7
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 943
KAAAATQ                                                         7

SEQ ID NO: 944          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
VARIANT                 3..5
                        note = any amino acid
VARIANT                 8..9
                        note = any amino acid
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 944
TKXXXSDXX                                                       9

SEQ ID NO: 945          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
VARIANT                 4..5
                        note = any amino acid
VARIANT                 8..9
                        note = any amino acid
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 945
TKDXXSDXX                                                       9

SEQ ID NO: 946          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
```

-continued

```
SEQUENCE: 946
TKDYASDAA                                                          9

SEQ ID NO: 947          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
VARIANT                 4..6
                        note = any amino acid
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 947
VKDXXXDA                                                           8

SEQ ID NO: 948          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
VARIANT                 5..6
                        note = any amino acid
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 948
VKDYXXDA                                                           8

SEQ ID NO: 949          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 949
VKDYASDA                                                           8

SEQ ID NO: 950          moltype =   length =
SEQUENCE: 950
000

SEQ ID NO: 951          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
VARIANT                 3..7
                        note = any amino acid
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 951
KDXXXXXAQ                                                          9

SEQ ID NO: 952          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 952
KDYASDTAQ                                                          9

SEQ ID NO: 953          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..7
                        note = any amino acid
VARIANT                 9
                        note = any amino acid
VARIANT                 11
                        note = any amino acid
VARIANT                 13
                        note = any amino acid
source                  1..15
```

-continued

```
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 953
XXXXXXXEXY XLXHP                                                15

SEQ ID NO: 954          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
VARIANT                 1..6
                        note = any amino acid
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 954
XXXXXXIEAY LLAHP                                                15

SEQ ID NO: 955          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 955
DALFKAIEAY LLAHP                                                15

SEQ ID NO: 956          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 956
QKESEERKQ                                                       9

SEQ ID NO: 957          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 957
SYRLQSGD                                                        8

SEQ ID NO: 958          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 958
DALRVPSGT                                                       9

SEQ ID NO: 959          moltype = AA  length = 9
FEATURE                 Location/Qualifiers
REGION                  1..9
                        note = epitope
source                  1..9
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 959
RKTISSEDK                                                       9

SEQ ID NO: 960          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 960
QHPEREPQQP GEKEE                                                15
```

-continued

```
SEQ ID NO: 961          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 961
GHQSQKGKHQ QEEEN                                                15

SEQ ID NO: 962          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 962
LLNALPEEVI QHTFN                                                15

SEQ ID NO: 963          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 963
MRGRINYIRR NEGKD                                                15

SEQ ID NO: 964          moltype = AA  length = 20
FEATURE                 Location/Qualifiers
REGION                  1..20
                        note = epitope candidate
source                  1..20
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 964
SSPRGKVEAE KARPMEETKV                                           20

SEQ ID NO: 965          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope candidate
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 965
SSPRGKVEAE                                                      10

SEQ ID NO: 966          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope candidate
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 966
SPRGKVEAEK                                                      10

SEQ ID NO: 967          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope candidate
source                  1..10
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 967
PRGKVEAEKA                                                      10

SEQ ID NO: 968          moltype = AA  length = 10
FEATURE                 Location/Qualifiers
REGION                  1..10
                        note = epitope candidate
source                  1..10
```

-continued

```
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 968
RGKVEAEKAR                                                          10

SEQ ID NO: 969      moltype = AA  length = 10
FEATURE             Location/Qualifiers
REGION              1..10
                    note = epitope candidate
source              1..10
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 969
GKVEAEKARP                                                          10

SEQ ID NO: 970      moltype = AA  length = 10
FEATURE             Location/Qualifiers
REGION              1..10
                    note = epitope candidate
source              1..10
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 970
KVEAEKARPM                                                          10

SEQ ID NO: 971      moltype = AA  length = 10
FEATURE             Location/Qualifiers
REGION              1..10
                    note = epitope candidate
source              1..10
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 971
VEAEKARPME                                                          10

SEQ ID NO: 972      moltype = AA  length = 10
FEATURE             Location/Qualifiers
REGION              1..10
                    note = epitope candidate
source              1..10
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 972
EAEKARPMEE                                                          10

SEQ ID NO: 973      moltype = AA  length = 10
FEATURE             Location/Qualifiers
REGION              1..10
                    note = epitope candidate
source              1..10
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 973
AEKARPMEET                                                          10

SEQ ID NO: 974      moltype = AA  length = 10
FEATURE             Location/Qualifiers
REGION              1..10
                    note = epitope candidate
source              1..10
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 974
EKARPMEETK                                                          10

SEQ ID NO: 975      moltype = AA  length = 10
FEATURE             Location/Qualifiers
REGION              1..10
                    note = epitope candidate
source              1..10
                    mol_type = protein
                    organism = synthetic construct
SEQUENCE: 975
KARPMEETKV                                                          10

SEQ ID NO: 976      moltype = AA  length = 8
```

-continued

```
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = modified epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 976
ARGKVEAE                                                        8

SEQ ID NO: 977          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = modified epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 977
PAGKVEAE                                                        8

SEQ ID NO: 978          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = modified epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 978
PRAKVEAE                                                        8

SEQ ID NO: 979          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = modified epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 979
PRGAVEAE                                                        8

SEQ ID NO: 980          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = modified epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 980
PRGKAEAE                                                        8

SEQ ID NO: 981          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = modified epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 981
PRGKVAAE                                                        8

SEQ ID NO: 982          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = modified epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 982
PRGKVEAE                                                        8

SEQ ID NO: 983          moltype = AA  length = 8
FEATURE                 Location/Qualifiers
REGION                  1..8
                        note = modified epitope
source                  1..8
                        mol_type = protein
                        organism = synthetic construct
```

-continued

```
SEQUENCE: 983
PRGKVEAA                                                    8

SEQ ID NO: 984          moltype = AA  length = 15
FEATURE                 Location/Qualifiers
REGION                  1..15
                        note = epitope
source                  1..15
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 984
MLSKQMRRRL SADER                                            15

SEQ ID NO: 985          moltype = AA  length = 14
FEATURE                 Location/Qualifiers
REGION                  1..14
                        note = epitope
source                  1..14
                        mol_type = protein
                        organism = synthetic construct
SEQUENCE: 985
ASDHVPRRSK TSLR                                             14
```

1-11. (canceled)

12. A method for detecting binding between an antigen and an antibody from a sample obtained from a subject,

wherein the antigen is a polypeptide selected from one of the following:

(1β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NO: 660;

(2β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 663, 666, 669-671, 675-678 and 681, 684, 689-691;

(3β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 695, 697 and 699-716;

(9β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 747-749;

(14β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 761, 764;

(25β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 814 and 815; and

(26β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 816, 819;

wherein the polypeptide as defined in (1β), (2β), (3β), (9β), (14β), (25β), and (26β) specifically binds to an IgE antibody from a patient allergic to soybean, and

wherein the method comprises contacting the antibody in a sample obtained from the subject that binds to the antigen with the antigen from a kit.

13. The method according to claim **12**, wherein the antigen from the kit is immobilized to a carrier or a surface usable for detection of binding between the antibody and the polypeptide.

14. The method according to claim **12**, wherein the kit further comprises a reagent that detects binding between said antigen and said antibody in the sample obtained from the subject.

15. The method according to claim **12**, wherein the polypeptide being any one of the following:

(2β-A) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 666, 670-671, 675-678, 681, 684 and 689;

(3-A) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 695, 697, 703, 704, 707, 709-711, 713, 715 and 716;

(14β-A) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 761, 764;

(25β-A) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 814 and 815; and

(26β-A) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NO: 816.

16. The method according to claim **12**, wherein the polypeptide being any one of the following:

(1β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 70 or 89 and further comprising at least one amino acid sequence selected from the group consisting of SEQ ID NO: 660;

(2β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 106 or 139 or consisting of an amino acid sequence having at least 75% identity to the amino acid sequence of SEQ ID NO: 122, and further comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 663, 666, 669-671, 675-678 and 681, 684, 689-691;

(3β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 16 and further comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 695, 697 and 699-716;

(9β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 201 and further comprising at

least one amino acid sequence selected from the group consisting of SEQ ID NOs: 747-749;

(14β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 280 or 286 and further comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 761, 764;

(25β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 348 or 358 and further comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 814 and 815; and

(26β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 367, 384 or 394 and further comprising at least one amino acid sequence selected from the group consisting of SEQ ID Nos 816, 819.

17. The method according to claim **12**, wherein the sample is obtained from the subject, wherein the method further comprises detecting binding between an IgE antibody present in the sample and the antigen from the kit, and wherein the antigen from the kit is immobilized to a carrier or a surface.

18. The method according to claim **12**, wherein the sample is obtained from the subject, wherein the method further comprises detecting binding between an IgE antibody present in the sample and the antigen from the kit, and wherein the kit further comprises a reagent that detects binding between said antigen and said antibody.

19. A method for diagnosing and treating an allergy to soybean in a subject, wherein the method comprises:

(i) detecting binding between an antibody in a sample obtained from the subject with and an antigen from a kit,

wherein the antigen is a polypeptide selected from one of the following:

(1β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NO: 660;

(2β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 663, 666, 669-671, 675-678 and 681, 684, 689-691;

(3β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 695, 697 and 699-716;

(9β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 747-749;

(14β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 761, 764;

(25β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 814 and 815; and

(26β) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 816, 819;

wherein the polypeptide as defined in (1β), (2β), (3β), (9β), (14), (25β), and (26β) specifically binds to an IgE antibody from a patient allergic to soybean, and

(ii) administering hyposensitization therapy to the subject when binding between said antibody and the antigen is detected.

20. The method according to claim **19**, wherein the antigen from the kit is immobilized to a carrier or a surface usable for detection of binding between the antibody and the polypeptide.

21. The method according to claim **19**, wherein the kit further comprises a reagent that detects binding between said antigen and said antibody in the sample obtained from the subject.

22. The method according to claim **19**, wherein the polypeptide being any one of the following:

(2β-A) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 666, 670-671, 675-678, 681, 684 and 689;

(3-A) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 695, 697, 703, 704, 707, 709-711, 713, 715 and 716;

(14β-A) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 761, 764;

(25β-A) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 814 and 815; and

(26β-A) a polypeptide comprising at least one amino acid sequence selected from the group consisting of SEQ ID NO: 816.

23. The method according to claim **19**, wherein the polypeptide being any one of the following:

(1β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 70 or 89 and further comprising at least one amino acid sequence selected from the group consisting of SEQ ID NO: 660;

(2β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 106 or 139 or consisting of an amino acid sequence having at least 75% identity to the amino acid sequence of SEQ ID NO: 122, and further comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 663, 666, 669-671, 675-678 and 681, 684, 689-691;

(3β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 16 and further comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 695, 697 and 699-716;

(9β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 201 and further comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 747-749;

(14β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 280 or 286 and further comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 761, 764;

(25β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 348 or 358 and further

comprising at least one amino acid sequence selected from the group consisting of SEQ ID NOs: 814 and 815; and

(26β-B1) a polypeptide consisting of an amino acid sequence having at least 90% identity to the amino acid sequence of SEQ ID NO: 367, 384 or 394 and further comprising at least one amino acid sequence selected from the group consisting of SEQ ID Nos 816, 819.

* * * * *