fus65a GST-fusion Rabbit Overath H2

FTKAKDGITKQLPEAKKAADDAKKLYDEVKKAAEDARGVKVSDETNSSGLYRIL

DWYCFKEGENAGQSHNCENVGFSVHHGTHKRRNVIDCGDEKANKYGDASSKTLEDTLKE

WNVKKPSAETNNNGNDMCKNGQSSESHPCTMTEEWQTHYKETVEKLRELEGAHERGKKA

HDAMLGYANTAHATNRKVEQGKPLTEVIAAAKEAGKKGAKIIIPAAAPSTPTNSTKNED

SAPTEHVDRGIATNETQVEVGIDADFDGLLEATEA

fus65b GST-fusion Rabbit Overath H5

DQEGANLLCKM

KHLADKVANKGGEDLKKKTKGFEDDVLLEVERVNNWLEKLGDRKQYNDGYAKLSDSDIE

KVKGIFTKAKDGITKQLPEAKKAADDAKKLYDEVKKAAEDARGVKVSDETNSSGLYRIL

DWYCFKEGENAGQSHNCENVGFSVHHGTHKRRNVIDCGDEKANKYGDASSKTLEDTLKE

WNVKKPSAETNNNGNDMCKNGQSSESHPCTMTEEWQTHYKETVEKLRELEGAHERGKKA

HDAMLGYANTAHATNRKVEQGKPLTEVIAAAKEAGKKGAKIIIPAAAPSTPTNSTKNED

SAPTEHVDRGIATNETQVEVGIDADFDGLLEATEA

p1271 Rabbit Doak

MATNGGDKHVKANKKLDQEGANLLCKMKHLADKVANKGGEDLKKRTKGFEDDVLLEVERV

NNWLEKLNGPKGRKDGYAKLSDSDIEKVKEIFTKAKDGIAKQLPEAKKAGENAEKLYDEV

KKAAEHARGQKLDDDRDKSTGLYRILDWYCFKEEENAGKSDNCDGVKFSEHHGTHRRRNV

IDCGDEKANKYGDASSKTLEDTLKQWESKKPQAAGASGGNDVCKATASSESHPCTMTEEW

QTPYKETVKKLKELEGAHERGKKAHDAMLGYANTAHATNRKVEQGKPLTEVIAAAKEAGK

KGAKIIIPAAAPSTPTNSTKNEDSAPTEHVDRGIATNETQVEVGIDADFDGLLEATEAAE

VKSRHQRREREFDIKLAAALEHHHHH

p1912 Rabbit 48

MDKNLTKEGAAALCKMKHLADKVAEKRSQELKDRTQNFAGYIEFELYRIDYWLEKLNGPK

GRKDGYAKLSDSDIEKVKEIFDKAKDGIAKQLPEAKKAGEDAEKLHTEVKEAAANARGQD

LDDHKQKSTGLYRVLNWYCITKEESHNATPNCDGIQFRNHYLSVNRSAIDCSSTGYEENY

DWSANALQVALNSWENVKPKKLESAGSDENCNIGQSSESHPCTMTEEWQTHYKETVKKLR

ELEDAYQRGKKAHDAMLGYANTAYAVNTKVEQEKPLAEVIAAAKEAGKKGAKIIIPAAAP

VTPTNSTKNEDSAPTEHVDRGIATNETQVEVGIDADFDSLLDATEAAEVTRRHQRLEHHHHHH

1 50

fus65a

fus65b DQEG ANLLCKMKHL ADKVANKGGE DLKKKTKGFE

p1271 MATNGGDKHV KANKKLDQEG ANLLCKMKHL ADKVANKGGE DLKKRTKGFE

p1912 MDKNLTKEG AAALCKMKHL ADKVAEKRSQ ELKDRTQNFA

51 100

fus65a FTKAKDGIT

fus65b DDVLLEVERV NNWLEKLGDR KQYNDGYAKL SDSDIEKVKG IFTKAKDGIT

p1271 DDVLLEVERV NNWLEKLNGP KGRKDGYAKL SDSDIEKVKE IFTKAKDGIA

p1912 GYIEFELYRI DYWLEKLNGP KGRKDGYAKL SDSDIEKVKE IFDKAKDGIA

101 150

fus65a KQLPEAKKAA DDAKKLYDEV KKAAEDARGV KVSDETN.SS GLYRILDWYC

fus65b KQLPEAKKAA DDAKKLYDEV KKAAEDARGV KVSDETN.SS GLYRILDWYC

p1271 KQLPEAKKAG ENAEKLYDEV KKAAEHARGQ KLDDDRDKST GLYRILDWYC

p1912 KQLPEAKKAG EDAEKLHTEV KEAAANARGQ DLDDHKQKST GLYRVLNWYC

151 200

fus65a FKEGENAGQS HNCENVGFSV HHGTHKRRNV IDCGDEKANK YGDASSKTLE

fus65b FKEGENAGQS HNCENVGFSV HHGTHKRRNV IDCGDEKANK YGDASSKTLE

p1271 FKEEENAGKS DNCDGVKFSE HHGTHRRRNV IDCGDEKANK YGDASSKTLE

p1912 ITKEESHNAT PNCDGIQFRN HY.LSVNRSA IDCSSTGYEE NYDWSANALQ

201 250

fus65a DTLKEWNVKK PSAETNNNGN DMCKNGQSSE SHPCTMTEEW QTHYKETVEK

fus65b DTLKEWNVKK PSAETNNNGN DMCKNGQSSE SHPCTMTEEW QTHYKETVEK

p1271 DTLKQWESKK PQAAGASGGN DVCKATASSE SHPCTMTEEW QTPYKETVKK

p1912 VALNSWENVK PKKLESAGSD ENCNIGQSSE SHPCTMTEEW QTHYKETVKK

251 300

fus65a LRELEGAHER GKKAHDAMLG YANTAHATNR KVEQGKPLTE VIAAAKEAGK

fus65b LRELEGAHER GKKAHDAMLG YANTAHATNR KVEQGKPLTE VIAAAKEAGK

p1271 LKELEGAHER GKKAHDAMLG YANTAHATNR KVEQGKPLTE VIAAAKEAGK

p1912 LRELEDAYQR GKKAHDAMLG YANTAYAVNT KVEQEKPLAE VIAAAKEAGK

301 350

fus65a KGAKIIIPAA APSTPTNSTK NEDSAPTEHV DRGIATNETQ VEVGIDADFD

fus65b KGAKIIIPAA APSTPTNSTK NEDSAPTEHV DRGIATNETQ VEVGIDADFD

p1271 KGAKIIIPAA APSTPTNSTK NEDSAPTEHV DRGIATNETQ VEVGIDADFD

p1912 KGAKIIIPAA APVTPTNSTK NEDSAPTEHV DRGIATNETQ VEVGIDADFD

351 386

fus65a GLLEATEA

fus65b GLLEATEA

p1271 GLLEATEAAE VKSRHQRRER EFDIKLAAAL EHHHHH

p1912 SLLDATEAAE VTRRHQRLEH HHHHH

ID TB65KISG standard; RNA; INV; 2369 BP.

XX

AC M86709;

XX

SV M86709.1

XX

DT 14-JUL-1992 (Rel. 32, Created)

DT 04-MAR-2000 (Rel. 63, Last updated, Version 4)

XX

DE Trypanosoma brucei 65 kDa invariant surface glycoprotein mRNA, complete

DE cds.

XX

KW 65 kDa invariant surface glycoprotein.

XX

OS Trypanosoma brucei

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

XX

RN [1]

RP 1-2369

RX MEDLINE; 92268135.

RA Ziegelbauer K., Multhaup G., Overath P.;

RT "Molecular characterization of the invariant surface glycoproteins specific

RT for the bloodstream stage of Trypanosoma brucei";

RL J. Biol. Chem. 267:10797-10803(1992).

XX

DR SPTREMBL; Q26712; Q26712.

XX

FH Key Location/Qualifiers

FH

FT source 1..2369

FT /db\_xref="taxon:5691"

FT /organism="Trypanosoma brucei"

FT /strain="MITat"

FT /dev\_stage="bloodstream"

FT /tissue\_lib="lambda ZAPII"

FT exon 1..16

FT /standard\_name="mini exon"

FT sig\_peptide 56..115

FT CDS 56..1366

FT /codon\_start=1

FT /db\_xref="SPTREMBL:Q26712"

FT /product="65 kDa invariant surface glycoprotein"

FT /protein\_id="AAA30147.1"

FT /translation="MMKYLLVFAIITTRISVLLATNGGDKHVKANKKLDQEGANLLCKM

FT KHLADKVANKGGEDLKKKTKGFEDDVLLEVERVNNWLEKLGDRKQYNDGYAKLSDSDIE

FT KVKGIFTKAKDGITKQLPEAKKAADDAKKLYDEVKKAAEDARGVKVSDETNSSGLYRIL

FT DWYCFKEGENAGQSHNCENVGFSVHHGTHKRRNVIDCGDEKANKYGDASSKTLEDTLKE

FT WNVKKPSAETNNNGNDMCKNGQSSESHPCTMTEEWQTHYKETVEKLRELEGAHERGKKA

FT HDAMLGYANTAHATNRKVEQGKPLTEVIAAAKEAGKKGAKIIIPAAAPSTPTNSTKNED

FT SAPTEHVDRGIATNETQVEVGIDADFDGLLEATEAAEVKSRHQRRAMIILAVLVPAIIL

FT VVTAVAFFIMVKRRRNNSQDVDTGKAEGGVSSGKAVM"

FT mat\_peptide 116..1363

FT /product="65 kDa invariant surface glycoprotein"

XX

SQ Sequence 2369 BP; 788 A; 400 C; 585 G; 596 T; 0 other;

tttctgtact atattgagtt cttgataaga ctgaaaggtg gaataatata aaaagatgat 60

gaagtatttg ctggtatttg caattattac cacaagaatc tctgtgttgt tggcgactaa 120

tggtggagat aaacatgtaa aagctaacaa gaaacttgat caagaaggtg ccaacttgct 180

ctgtaaaatg aagcatcttg ctgataaagt tgcgaataaa ggaggagaag acctaaagaa 240

gaaaactaaa ggttttgaag atgacgtact tttggaggtg gaaagagtga ataattggtt 300

ggaaaagctg ggagatcgga aacagtataa cgacggttat gctaagctgt ctgactctga 360

tatagaaaaa gtgaagggaa tatttactaa ggcaaaagat ggaataacta agcaactccc 420

tgaggcaaag aaggctgcgg acgacgctaa gaaactgtac gatgaagtga agaaggccgc 480

agaagatgca cgtggagtga aagtgagtga cgaaacgaac tctagtggcc tttacaggat 540

tttggattgg tattgtttta aagaagggga gaacgcgggt cagagtcata actgcgaaaa 600

cgtcgggttt agtgtgcatc acggaacgca caaaagaagg aatgttattg attgcggtga 660

cgagaaagcg aacaaatatg gggatgcttc ctcaaaaaca ttggaggaca ccttaaaaga 720

gtggaatgtc aagaagccct cagctgaaac caacaacaac ggaaatgaca tgtgcaagaa 780

tggtcaatct tccgagagcc atccgtgcac catgacagag gagtggcaga ctcactacaa 840

ggaaactgtc gaaaagctaa gggaactcga gggtgcgcac gagaggggca agaaggctca 900

tgatgctatg ttgggttacg ctaatactgc acatgcgaca aacaggaaag tggaacaggg 960

caaaccactg acagaggtga tagcggcagc taaggaagca gggaaaaagg gcgcgaaaat 1020

tataataccc gcagctgccc catcaacacc gactaacagc acgaaaaatg aagatagtgc 1080

tccaaccgag catgttgata gagggattgc aacaaatgaa acacaggtgg aagttggtat 1140

tgatgctgac tttgacggtt tgctggaggc caccgaggct gcagaagtaa aaagcagaca 1200

tcagagaaga gcaatgatta tattagcagt ccttgtacct gccattattc ttgtggtgac 1260

ggccgttgca ttcttcataa tggtgaaacg aaggaggaat aactcccagg atgtggacac 1320

cggaaaagcg gagggtgggg tttctagtgg aaaggcggta atgtagactc accttcactt 1380

ataggaatgt gcaggttttt ttcaaacagt gattgtaagt aagggaacta aagagaataa 1440

aaatgacaat gagttgaata aaagtatata gattcatgag ataaggaaga agttctagca 1500

gccttttcgt tgtttttgtt ttgtggtaat accttaccta tgctttttta aaaataaaga 1560

attgacacaa tagagcaaat attttcactt gaagaaaccg ttgctgattt tctcttccat 1620

tcgtaaatcc agctgattgt ttggagcatc catctctcac actgtacgtc atcatggcta 1680

gcatgcttca tgaaactgcc acaaccgtat catacagggg agcgctatgc aaattagcag 1740

gtgttgagct acaagggatt aatctatcac cacctcaaag gctgtgccac ttggggttgc 1800

tgatcagtcg tgagatcaac tggctggcag cagtcctcga ggaatgggaa gaaaaatata 1860

tgctcttcac atgtgtgccg taccgtggtg ttgtggacgc gtttacaacc ctgctcagga 1920

caattaagcg gcacctggag acagagaaca gttcaaagga caatgaaaag gagtgggatg 1980

cgttgctggc ttttagcatt actgaatttc aggtggaggt tatggaaatc ggaaagggta 2040

agagttacta gcattttttt cttttcttcc tacctgtaca tttaactttt tttcctcttt 2100

tatttatggt ggtatgtcct cttttttttg tcacccaatg gttaaattac catgcggtga 2160

caaaacaaag actcggtttt caaactcatt gaaagtaata gagtacgcat acctaaaaaa 2220

aagtgtatgc ggagtggatt atgtattctc atgaaatcaa tcattcaaat atctaatgca 2280

aagcgtaaac aactgtacga cggaaattat ttttattttt gcgatgccgc tataaataaa 2340

tccttcacta aaaaaaaaaa aaaaaaaaa

tttctgtactatattgagttcttgataagactgaaaggtggaataatataaaaagatgat

gaagtatttgctggtatttgcaattattaccacaagaatctctgtgttgttggcgactaa

tggtggagataaacatgtaaaagctaacaagaaacttgatcaagaaggtgccaacttgct

ctgtaaaatgaagcatcttgctgataaagttgcgaataaaggaggagaagacctaaagaa

gaaaactaaaggttttgaagatgacgtacttttggaggtggaaagagtgaataattggtt

ggaaaagctgggagatcggaaacagtataacgacggttatgctaagctgtctgactctga

tatagaaaaagtgaagggaatatttactaaggcaaaagatggaataactaagcaactccc

tgaggcaaagaaggctgcggacgacgctaagaaactgtacgatgaagtgaagaaggccgc

agaagatgcacgtggagtgaaagtgagtgacgaaacgaactctagtggcctttacaggat

tttggattggtattgttttaaagaaggggagaacgcgggtcagagtcataactgcgaaaa

cgtcgggtttagtgtgcatcacggaacgcacaaaagaaggaatgttattgattgcggtga

cgagaaagcgaacaaatatggggatgcttcctcaaaaacattggaggacaccttaaaaga

gtggaatgtcaagaagccctcagctgaaaccaacaacaacggaaatgacatgtgcaagaa

tggtcaatcttccgagagccatccgtgcaccatgacagaggagtggcagactcactacaa

ggaaactgtcgaaaagctaagggaactcgagggtgcgcacgagaggggcaagaaggctca

tgatgctatgttgggttacgctaatactgcacatgcgacaaacaggaaagtggaacaggg

caaaccactgacagaggtgatagcggcagctaaggaagcagggaaaaagggcgcgaaaat

tataatacccgcagctgccccatcaacaccgactaacagcacgaaaaatgaagatagtgc

tccaaccgagcatgttgatagagggattgcaacaaatgaaacacaggtggaagttggtat

tgatgctgactttgacggtttgctggaggccaccgaggctgcagaagtaaaaagcagaca

tcagagaagagcaatgattatattagcagtccttgtacctgccattattcttgtggtgac

ggccgttgcattcttcataatggtgaaacgaaggaggaataactcccaggatgtggacac

cggaaaagcggagggtggggtttctagtggaaaggcggtaatgtagactcaccttcactt

ataggaatgtgcaggtttttttcaaacagtgattgtaagtaagggaactaaagagaataa

aaatgacaatgagttgaataaaagtatatagattcatgagataaggaagaagttctagca

gccttttcgttgtttttgttttgtggtaataccttacctatgcttttttaaaaataaaga

attgacacaatagagcaaatattttcacttgaagaaaccgttgctgattttctcttccat

tcgtaaatccagctgattgtttggagcatccatctctcacactgtacgtcatcatggcta

gcatgcttcatgaaactgccacaaccgtatcatacaggggagcgctatgcaaattagcag

gtgttgagctacaagggattaatctatcaccacctcaaaggctgtgccacttggggttgc

tgatcagtcgtgagatcaactggctggcagcagtcctcgaggaatgggaagaaaaatata

tgctcttcacatgtgtgccgtaccgtggtgttgtggacgcgtttacaaccctgctcagga

caattaagcggcacctggagacagagaacagttcaaaggacaatgaaaaggagtgggatg

cgttgctggcttttagcattactgaatttcaggtggaggttatggaaatcggaaagggta

agagttactagcatttttttcttttcttcctacctgtacatttaactttttttcctcttt

tatttatggtggtatgtcctcttttttttgtcacccaatggttaaattaccatgcggtga

caaaacaaagactcggttttcaaactcattgaaagtaatagagtacgcatacctaaaaaa

aagtgtatgcggagtggattatgtattctcatgaaatcaatcattcaaatatctaatgca

aagcgtaaacaactgtacgacggaaattatttttatttttgcgatgccgctataaataaa

tccttcactaaaaaaaaaaaaaaaaaaaa

tttctgtactatattgagttcttgataagactgaaaggtggaataatataaaaagatgat

10 20 30 40 50 60

M M

gaagtatttgctggtatttgcaattattaccacaagaatctctgtgttgttggcgactaa

K Y L L V F A I I T T R I S V L L A T N

AFLIII BCLI BANI

tggtggagataaacatgtaaaagctaacaagaaacttgatcaagaaggtgccaacttgct

G G D K H V K A N K K L D Q E G A N L L

SFANI MBOII

ctgtaaaatgaagcatcttgctgataaagttgcgaataaaggaggagaagacctaaagaa

190 200 210 220 230 240

C K M K H L A D K V A N K G G E D L K K

MBOII MBOII

gaaaactaaaggttttgaagatgacgtacttttggaggtggaaagagtgaataattggtt

250 260 270 280 290 300

K T K G F E D D V L L E V E R V N N W L

ESPI

ggaaaagctgggagatcggaaacagtataacgacggttatgctaagctgtctgactctga

310 320 330 340 350 360

E K L G D R K Q Y N D G Y A K L S D S D

SSPI

tatagaaaaagtgaagggaatatttactaaggcaaaagatggaataactaagcaactccc

370 380 390 400 410 420

I E K V K G I F T K A K D G I T K Q L P

MSTII

.BBVI HGAI

tgaggcaaagaaggctgcggacgacgctaagaaactgtacgatgaagtgaagaaggccgc

430 440 450 460 470 480

E A K K A A D D A K K L Y D E V K K A A

MBOII MBOII

. SFANI

agaagatgcacgtggagtgaaagtgagtgacgaaacgaactctagtggcctttacaggat

490 500 510 520 530 540

E D A R G V K V S D E T N S S G L Y R I

DRAI

tttggattggtattgttttaaagaaggggagaacgcgggtcagagtcataactgcgaaaa

550 560 570 580 590 600

L D W Y C F K E G E N A G Q S H N C E N

SFANI

cgtcgggtttagtgtgcatcacggaacgcacaaaagaaggaatgttattgattgcggtga

610 620 630 640 650 660

V G F S V H H G T H K R R N V I D C G D

HPHI SFANI

. . FOKI

cgagaaagcgaacaaatatggggatgcttcctcaaaaacattggaggacaccttaaaaga

670 680 690 700 710 720

E K A N K Y G D A S S K T L E D T L K E

NSPBII

PVUII AFLIII

gtggaatgtcaagaagccctcagctgaaaccaacaacaacggaaatgacatgtgcaagaa

730 740 750 760 770 780

W N V K K P S A E T N N N G N D M C K N

MBOII FOKI APALI

. . . BSP1286

. . . HGIAI

tggtcaatcttccgagagccatccgtgcaccatgacagaggagtggcagactcactacaa

790 800 810 820 830 840

G Q S S E S H P C T M T E E W Q T H Y K

ECOK

. AVAI

. XHOI FSPI

ggaaactgtcgaaaagctaagggaactcgagggtgcgcacgagaggggcaagaaggctca

850 860 870 880 890 900

E T V E K L R E L E G A H E R G K K A H

SFANI

tgatgctatgttgggttacgctaatactgcacatgcgacaaacaggaaagtggaacaggg

910 920 930 940 950 960

D A M L G Y A N T A H A T N R K V E Q G

HPHI BBVI

caaaccactgacagaggtgatagcggcagctaaggaagcagggaaaaagggcgcgaaaat

970 980 990 1000 1010 1020

K P L T E V I A A A K E A G K K G A K I

BBVI NSPBII

. PVUII BBVI

tataatacccgcagctgccccatcaacaccgactaacagcacgaaaaatgaagatagtgc

1030 1040 1050 1060 1070 1080

I I P A A A P S T P T N S T K N E D S A

BSP1286

HGIAI

.MBOII

tccaaccgagcatgttgatagagggattgcaacaaatgaaacacaggtggaagttggtat

1090 1100 1110 1120 1130 1140

P T E H V D R G I A T N E T Q V E V G I

SFANI BBVI BGLI PSTI

tgatgctgactttgacggtttgctggaggccaccgaggctgcagaagtaaaaagcagaca

1150 1160 1170 1180 1190 1200

D A D F D G L L E A T E A A E V K S R H

MBOII BSPMI

. . XCMI

tcagagaagagcaatgattatattagcagtccttgtacctgccattattcttgtggtgac

1210 1220 1230 1240 1250 1260

Q R R A M I I L A V L V P A I I L V V T

EAEI

XMAIII

. MBOII

. . HPHI

. . .BSMI HPHI BSTNI

. . .. . . PFIMI

ggccgttgcattcttcataatggtgaaacgaaggaggaataactcccaggatgtggacac

1270 1280 1290 1300 1310 1320

A V A F F I M V K R R R N N S Q D V D T

FOKI HPHI

. . ACCI

cggaaaagcggagggtggggtttctagtggaaaggcggtaatgtagactcaccttcactt

1330 1340 1350 1360 1370 1380

G K A E G G V S S G K A V M \*

BSPMI

ataggaatgtgcaggtttttttcaaacagtgattgtaagtaagggaactaaagagaataa

1390 1400 1410 1420 1430 1440

XMNI MBOII

aaatgacaatgagttgaataaaagtatatagattcatgagataaggaagaagttctagca

1450 1460 1470 1480 1490 1500

BBVI DRAI

gccttttcgttgtttttgttttgtggtaataccttacctatgcttttttaaaaataaaga

1510 1520 1530 1540 1550 1560

SSPI MBOII

. . MBOII

attgacacaatagagcaaatattttcacttgaagaaaccgttgctgattttctcttccat

1570 1580 1590 1600 1610 1620

NSPBII

PVUII

.FOKI

.. SFANI ECOB NHEI

tcgtaaatccagctgattgtttggagcatccatctctcacactgtacgtcatcatggcta

1630 1640 1650 1660 1670 1680

SPHI HAEII

. . BSPMI

gcatgcttcatgaaactgccacaaccgtatcatacaggggagcgctatgcaaattagcag

1690 1700 1710 1720 1730 1740

HPHI

gtgttgagctacaagggattaatctatcaccacctcaaaggctgtgccacttggggttgc

1750 1760 1770 1780 1790 1800

BCLI AVAI

. XHOI

. . BBVI MBOII

tgatcagtcgtgagatcaactggctggcagcagtcctcgaggaatgggaagaaaaatata

1810 1820 1830 1840 1850 1860

MBOII AFLIII AFLIII

. . MLUI HGAI

tgctcttcacatgtgtgccgtaccgtggtgttgtggacgcgtttacaaccctgctcagga

1870 1880 1890 1900 1910 1920

BANI BSTNI

caattaagcggcacctggagacagagaacagttcaaaggacaatgaaaaggagtgggatg

1930 1940 1950 1960 1970 1980

SFANI

. FOKI

cgttgctggcttttagcattactgaatttcaggtggaggttatggaaatcggaaagggta

1990 2000 2010 2020 2030 2040

MBOII

agagttactagcatttttttcttttcttcctacctgtacatttaactttttttcctcttt

2050 2060 2070 2080 2090 2100

HPHI

tatttatggtggtatgtcctcttttttttgtcacccaatggttaaattaccatgcggtga

2110 2120 2130 2140 2150 2160

HPHI

caaaacaaagactcggttttcaaactcattgaaagtaatagagtacgcatacctaaaaaa

2170 2180 2190 2200 2210 2220

aagtgtatgcggagtggattatgtattctcatgaaatcaatcattcaaatatctaatgca

2230 2240 2250 2260 2270 2280

SFANI

aagcgtaaacaactgtacgacggaaattatttttatttttgcgatgccgctataaataaa

2290 2300 2310 2320 2330 2340

tccttcactaaaaaaaaaaaaaaaaaaaa

2350 2360