



US 20250258175A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2025/0258175 A1**
Mahler (43) **Pub. Date:** **Aug. 14, 2025**

(54) **COMPOSITIONS AND METHODS FOR
DIAGNOSING AND ASSESSING
RHEUMATOID ARTHRITIS**

(71) Applicant: **Inova Diagnostics, Inc.**, San Diego,
CA (US)

(72) Inventor: **Michael Mahler**, Bad Neuenahr (DE)

(21) Appl. No.: **19/184,462**

(22) Filed: **Apr. 21, 2025**

Related U.S. Application Data

(63) Continuation of application No. 16/791,779, filed on
Feb. 14, 2020, now Pat. No. 12,298,304.

(60) Provisional application No. 62/806,607, filed on Feb.
15, 2019.

Publication Classification

(51) **Int. Cl.**
G01N 33/573 (2006.01)
G01N 33/58 (2006.01)
(52) **U.S. Cl.**
CPC **G01N 33/573** (2013.01); **G01N 33/58**
(2013.01); **G01N 2458/00** (2013.01); **G01N
2800/105** (2013.01); **G01N 2800/56** (2013.01)

(57) **ABSTRACT**

The present disclosure relates to the use of anti-PAD IgA as a clinical biomarker for diagnostic and prognostic information in rheumatoid arthritis (RA) patients. The disclosure further provides methods and compositions for the detection anti-PAD IgA in a biological sample.

Specification includes a Sequence Listing.

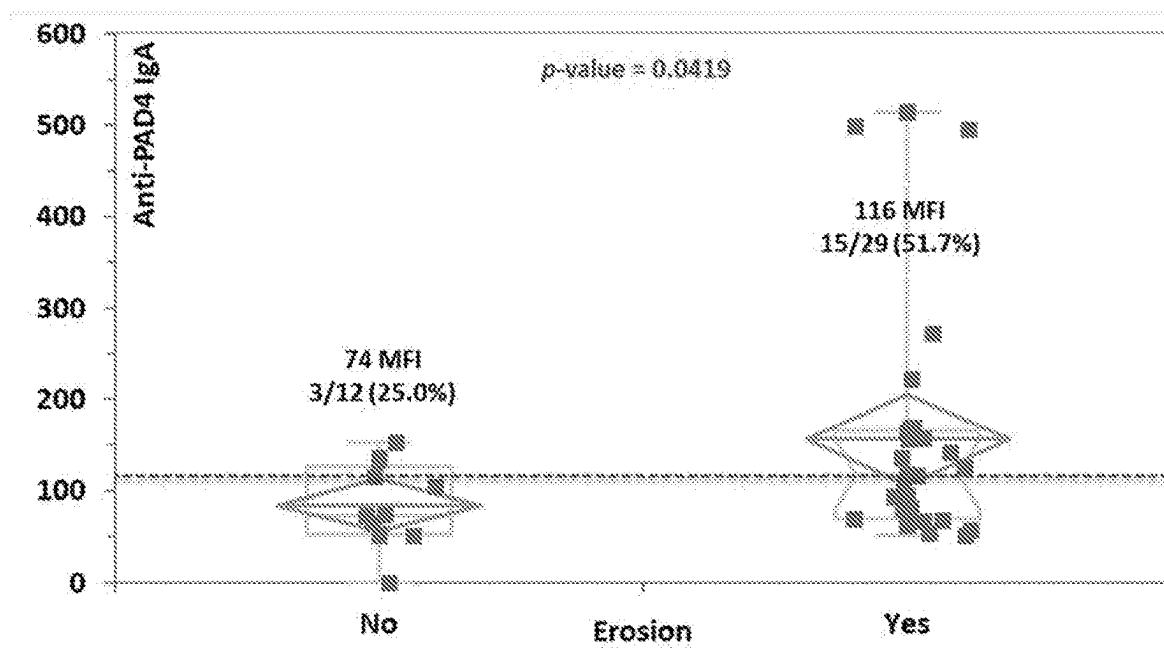


FIG. 1

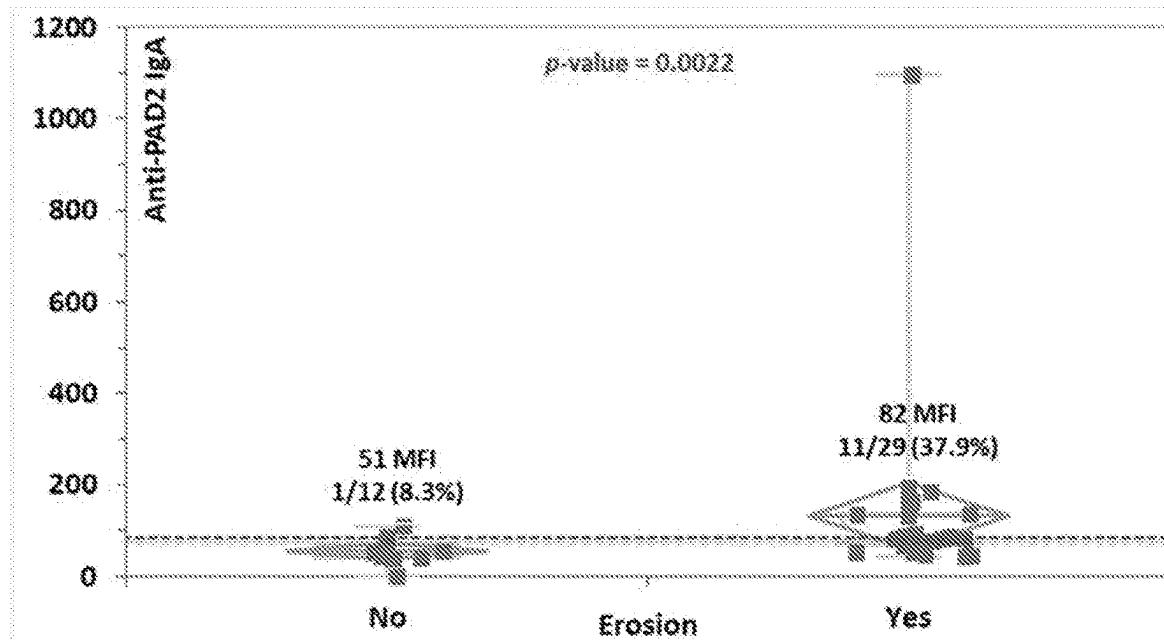


FIG. 2

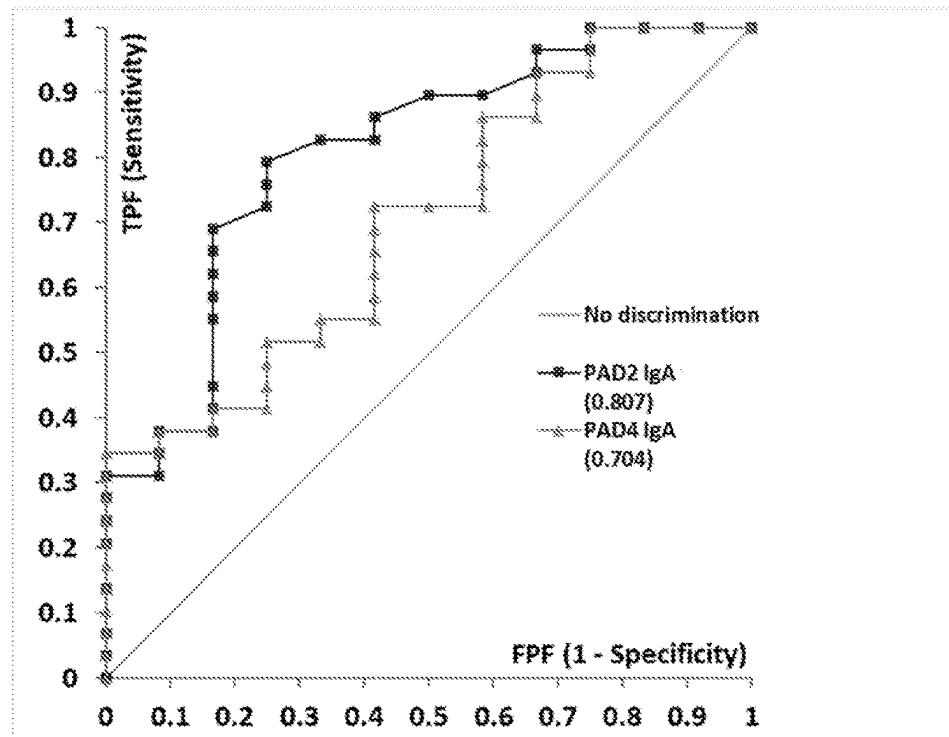


FIG. 3

DEFINITION Homo sapiens peptidyl arginine deiminase 2 (PADI2), mRNA.

ACCESSION NM_007365

VERSION NM_007365.3

sequence number-1

```

1 aggctgctgg agaaggcgca cctgctgcag gtgctcccg ccgccccgga ccagcgagcg
61 cgggactgac ggcggggagg atgctgcgcg agcggaccgt gcggctgcag tacgggagcc
121 gcgtggaggc ggtgtacgtg ctggcacct acctctggac cgatgtctac agcgcggccc
181 cagccggggc ccaaaccctc agcctgaagc actcggaaaca cgtgtgggtg gaggtggtgc
241 gtatgggga ggctgaggag gtggccacca atggcaagc gcgctggctt ctctcgcccc
301 gcaccaccct gcgggtcacc atgagccagg cgagcaccca ggccacagt gacaaggta
361 ccgtcaacta ctatgacgag gaaggagca ttcccatcga ccaggcgggg ctcttcctca
421 cagcattga gatctccctg gatgtggacg cagaccggga tggtgtggtg gagaagaaca
481 acccaaagaa ggcattcctgg acctggggcc ccgaggggca gggggccatc ctgctggtg
541 actgtgaccg agagacaccc tggttgccc aggaggactg ccgtgatgag aaggctaca
601 gcaaggaaga tctcaaggac atgtccaga tgatcctgcg gaccaaaggc cccgaccggcc
661 tccccggcc atacgagata gttctgtaca ttccatgtc agactcagac aaagtggggc
721 tgttctacgt ggagaacccg ttctcggcc aacgctatat ccacatcctg gcccggccgga
781 agcttacca tgtggtaag tacacgggtg gtcgcgcga gtcgtgttc ttctgtggaa
841 gcctctgttt ccccgacgag ggcttctcg gcctggctc catccatgtc agcctgtctg
901 agtacatggc ccaggacatt cccctgactc ccacatcctac ggacaccgtg atattccgg
961 ttgcctccgt gatcatgacc ccaacatcc tgcctccgt gtcgggttt gtgtgtctgca
1021 tgaaggataa ttacctgttc ctgaaagagg tgaagaacct tggggaaaa accaactgtg
1081 agctgaaggt ctgcttccag tacctaaacc gaggcgatcg ctggatccag gatgaaaattg
1141 agtttggcta catcgaggcc ccccataaaag gttcccccgt ggtgtggac tctccccggag
1201 atggaaacct aaaggactt cctgtgaagg agctcctgg cccagatttt ggctacgtga
1261 cccggggagcc cctctttgag tctgtcacca gccttgactc atttggaaac ctggaggtca
1321 gtccccccact gaccgtgaac ggcaagacat accccgcttgg ccgcattcctc atcggggagca
1381 gcttcctct gtcgtgtgtt cggaggatga ccaagggttgt ggctgacttc ctgaaggccc

```

1441 agcaggtgca ggcggccgtg gagctctact cagactggct gactgtggc cacgtggatg
 1501 agttcatgtc ctgttgcctcc atccccggca caaagaatt cctgctactc atggccagca
 1561 cctcggccctg ctacaagctc ttccgagaga agcagaaggaa cggccatgga gaggccatca
 1621 tggttcaaagg ctgtgggtggg atgagcagca agcgaatcac catcaacaag attctgtcca
 1681 acgagagcct tgtgcaggag aacctgtact tccagcgctg cctagactgg aaccgtgaca
 1741 tcctcaagaa ggagctggga ctgacagagc aggacatcat tgacctgccc gctctgttca
 1801 agatggacga ggaccaccgt gccagagccct tcttccaaa catggtaac atgatcgtgc
 1861 tggacaagga cctgggcatac cccaagccat tcggggcaca ggttgaggag gaatgctgcc
 1921 tggagatgca cgtgcgtggc ctcctggagc ccctgggcct cgaatgcacc ttcatcgacg
 1981 acatttctgc ctaccacaaa tttctgggg aagtccactg tggcaccaac gtccgcagga
 2041 agcccttcac cttaaagtgg tggcacatgg tgccctgacc tgccaggggc cctggcggtt
 2101 gcctccctcg cttagtttc cagaccctcc ctacacacgccc cagacccttc tgctgacatg
 2161 gactggacag cccccctggg agaccttgg gacgtggggt ggaatttggg gtatctgtgc
 2221 ctggccctcc ctgagagggg ctcagtgta ctctgaagcc atccccagtg agcctcgact
 2281 ctgtccctgc tgaaaatagc tggggccagtg tctctgtagc cctgacataa ggaacagaac
 2341 acaacaaaac acagcaaaacc atgtgccaa actgctcccc aaagaatttt gagtctctaa
 2401 tctgacactg aatgagggga gaagggaaagg agattctggg attgcccagtt ctccagcag
 2461 ccatgctctg aaaatcaagg tagaatccat gggaaaggac cccaggacc cgggacccta
 2521 gacgtatctt gaactgccc cgtcatttca aatacatctc cctcagggtt tccaggtggc
 2581 caccggcaat tattcattcc ttaccaacct ctcaaattctt cttggcttcc tctctgcagtt
 2641 gtggacactg ttggctagtc ctccccactc cctgagggc cagtaagttt gcttagaacc
 2701 ttcctggaaa catttcatct gagcaggttt cccacgtgt gggatgctcc tttgcctca
 2761 tctgtctcag ggtatgcaggc tccccccat gcatggggat ttctcccaag accagcatac
 2821 ttgtgacctg agaggtaat gctgtttttt gcccctggc agccatatcc atcttctctt
 2881 gcctggctct tgattctctg gcccgtccct gacccctccctt cttccactgc cttgactttc
 2941 ttcctttta ttcctgggtc catctgttca ggcagctaga caagaacttg ttgcagca
 3001 gccagattca ggccttccca ggggcataat aagtgaccag cccctctt cccggacatca
 3061 gatccaacac ataaggacc tggcctaccc tccagccaa cagccatgtc tgggtcagct
 3121 gccaacttag ggggtggttt attatccat tgaatttccac cagtgccctt gccaaagacc
 3181 ctctcatttg gacataccca gattcatcc ctgctccaa ctgaaaagac tcagtttcaa
 3241 tcgtttaaaag ttccttttagg gccagaagaa taaatgaatt ataatccat tttgaagaac
 3301 cgatttataa ccaatgaaaa ggttataatg taatttat tcttggagga acaagatttt
 3361 catttggat tatttcattc aaccatttca caaacatttgg ttgtatgcca ctaagcgcca
 3421 ggcacggcgt tggctctgc aaacacagt gtttagtagca gtctggaccc ggtccctact
 3481 ggcattggaaac ccatcactcc ccaacatgca aagcccat ttaaaggcca gcctctgccc
 3541 cttcagtgat ggcgtctta gaaatgccc tccactat tcaagaaatcc gcagggcaca
 3601 aaacttccag caagtcaactg ttgtggtaa atgggcagtg ggggtgggg gtttttttta
 3661 aacaggcccc cttcccatct accttagccag tacccatcca atgagttccc agagcctcca
 3721 gaagctgtg tctctctctt ggggacagca gctccctgcct ttggaggcca aagccccaga
 3781 tctctccagc cccagagctg aaaacaccaa gtgcctattt gagggtgtct gtctggagac
 3841 ttagagtttgc tcatgtgtgt gtgtgtgtt ggttaatgtg ggtttatggg ttttctttct
 3901 ttttttttta gtctacatta gggggaaatgt agcgcctccc atgtgcagac
 3961 agtgtgtctt tatagatttt tctaaggctt tcccaatga tgcggtaat ttctgtatgtt
 4021 tctgaagttc ccagactca cacacccgtt cccatctcac ttgcccaccc agtgtgacaa
 4081 ccctcggtgt ggtatatacc cctgtggactc atgctcttc cccacccca ctttctataa
 4141 atgttaggcct agaatacgtc tctctgtgc aaaactcagc taagttcctg ctccacctt
 4201 gatgttggaa tatcttatgt aagaggccag gggatgtcgta gaagatggca agaagaacac
 4261 agtttcaaat ttctggaaaa gggctgtgg tggagatcta aagatgttta gggaaagagct
 4321 cgactaaaga acaatgaaat aaatggtcca agggaaagtc a

FIG. 4

DEFINITION protein-arginine deiminase type-2 [Homo sapiens].
ACCESSION NP_031391
VERSION NP_031391.2
sequence number-2

1 mlrertvrlq ygsrveavyv lgtylwdvy saapagaqtf slkhsehvww evvrdgeaee
61 vatngkqrwl lspsttlrvt msqasteass dkvtvnyyde egsipidqag lfltaieisl
121 dvdadrdgvv eknnppkasw twgpeggai llnvncdretp wlpkedcrde kvsksedlkd
181 msqmilrtkg pdrlpagyei vlyismsdsd kvgvfyyenp ffgqryihil grrklyhvvk
241 ytggsaellf fveglcfpde gfsglvsihv slleymaqdi pltpiftdtv ifriapwimt
301 pnillppvsvf vccmkdnlylf lkevknlvek tncelkvfcq ylnrgdrwiq deiefgyiea
361 phkgfpvld sprdgnlkdf pkellgpdf gyvtreplfe svtsldsfgn levspvvtn
421 gktypngril igssfppls gg rrmrkvvrd lkaqqvqapv elysdwltvg hvdefmsfv
481 ipgtkkflll mastasykl frekqkdghg eaimfkglgg msskritink ilsneslvqe
541 nlyfqrcldw nrddilkkelg lteqdiidlp alfkmdedhr araffpnmvn mivlkdldgi
601 pkpfpqvee eclemhvrg lleplglect fiddisayhk flgevhcgtn vrkpftfk
661 whmvp

FIG. 5

DEFINITION PREDICTED: Homo sapiens peptidyl arginine deiminase 2 (PADI2),
transcript variant X2, mRNA.
ACCESSION XM_017000148
VERSION XM_017000148.2
sequence number-3

1 agcagctctg cagatgggaa ttcttctgtc agctcatatc tgcacatgtg cacaaggaca
61 taaaacataca agtgcactac agtgcgtc ttgttagtag caaaagattt gaaaggcaacc
121 taaaatgtcca tctgtggagg actgagtaac ggcctccccc gaagttctgt gcaaccgcta
181 aaaaggatga agccagtctc tcaatatggc tacaggatgt gcttcaggaa tggaaattttag
241 tttggctaca tcgaggcccc ccataaaaggc ttccccgtgg tgctggactc tcccccggat
301 gggaaacctaa aggacttccc tgtgaaggag ctccctggggc cagattttgg ctacgtgacc
361 cgggagcccc tctttgagtc tgtcaccagc cttgactcat ttggaaacct ggagggtcagt
421 cccccagtga ccgtgaacgg caagacatac cccgttggcc gcattctcat cgggagcagc
481 tttcctctgt ctgggtggcc gaggatgacc aagggtggtc gtgacttctt gaaggccccag
541 cagggtgcagg cggccgtggc gctctactca gactggctga ctgtggggca cgtggatgag
601 ttcatgtcct ttgtccccat cccggcaca aagaaattcc tgctactcat gggccagcacc
661 tcggcctgct acaagcttt ccggagagaag cagaaggacg gcccattcatg gggccatcatg
721 ttcaaaggct tgggtgggat gagcagcaag cgaatcacca tcaacaagat tctgtccaaac
781 gagagccttg tgcaggagaa cctgtacttc cagcgctgcc tagactggaa cctgtacatc
841 ctcaagaagg agctgggact gacagagcag gacatcattt acctggccgc tctgttcaag
901 atggacgagg accaccgtgc cagaccattt ttcccaaaca ttgtgaacat gatcggtctg
961 gacaaggacc tgggcattttt caagccattt gggccacagg ttgaggagga atgctgcctg
1021 gagatgcacg tgcgtggcct cctggagccc ctgggcctcg aatgcacattt catcgacgac
1081 atttctgcct accacaaatt tctggggaa gtccactgtg gcaccaacgt cccggggaa
1141 cccttcaccc tcaagtgggt gcacatggc ccctgacactg ccaggggccc tggcgtttgc
1201 ctccttcgt tagttctcca gaccctccct cacacgcccc gggccatttctg ctgacatgg
1261 ctggacagcc cccgtggggg acctttggg cgtgggggtgg aattttgggt atctgtgcct
1321 tgcctccct gagagggggc tcagtgtcct ctgaagccat ccccaacttgc cctcgactct
1381 gtcctgtc aaaatagctg ggccagtgtc tctgttagccc tgacataagg aacagaacac
1441 aacaaaacac agcaaaaccat gtgcctaaac tgctcccaa agaattttga gtctctaattc
1501 tgacactgaa tgaggggaga agggaggatttctggat tgccagttct tccagcagcc
1561 atgctctgaa aatcaaggta gaatccatgg aaaggggaccc caggccccgg ggaccctaga
1621 cgtatcttgc actgcccattt tcatttcaaa tacatctccc tcagggtttc cagggtggcca
1681 ccccaatta ttccatttccctt accaacctt ccaatctt tggctttctc tctgcagtgt
1741 ggacactgtt ggctagtctt ccccaactccc tgagggttcca gtaaggtagc ttggaaacctt
1801 cctggaaaca tttcatctga gcaggttcc ccacgtgtgg gatgctccctt ttgcctcatc
1861 tgtctcaggatgg atgcagggtc ccccgcatgc atggggattt ctcccaagac cagcataactt
1921 gtgacactgag agttcaatgc gtaaagatgc ccctggcttag ccataatccat ctctcttgc
1981 ctggccttgc attctctggc cgctccctga ccttcctcct tccactgcct tgactttctt

2041	cctttttatt	cctgggtcgc	tctgtccagg	cagctagaca	agaacttggtt	cgccagcagc
2101	cagattcagg	ccttcccagg	ggcataataa	gtgaccagcc	cctccctctcc	ggacatcaga
2161	tccaacacat	aaggaccctg	gcctaccctc	cagccccaca	gccagttctg	ggtcagctgc
2221	caacttaggg	gtggtttgat	tatcccattg	aaattcacca	gtgccttgc	caaagaccct
2281	ctcatttgg	catacccaga	ttcattccct	ggctccaact	gaaaagactc	agtttcaatc
2341	gttaaaaagt	cctttagggc	cagaagaata	aatgaattat	aatcccat	tgaagaaccgg
2401	atttataaacc	aatgaaaagg	ttataatgt	atttatattc	ttggaggaac	aagattttca
2461	tttgggatta	tttccttcaa	ccattcaaca	aacatttgtt	gtatgccact	aagcgccagg
2521	cacggcggtt	ggctctgc	acacagtgg	tagtagcagt	ctggacactgg	tccctactgg
2581	catggAACCC	atcactcccc	aacatgcaaa	gcccacattt	aaaggccagc	ctctgcccct
2641	tcagtgtatgc	gctctttaga	aatgcccagt	cactatattc	agaaatccgc	agggcacaaa
2701	acttccagca	agtcaactgtt	gtggtaaat	gggcagtg	ggtgggggg	cttcttaaa
2761	caggccccct	tcccatctac	ctagccagta	cccatccaat	gagtccccag	agcctccaga
2821	agctgttgtc	tcctctctgg	ggacagcagc	tcctgcctt	ggaggccaaa	gccccagatc
2881	tctccagccc	cagagctgaa	aacaccaagt	gcctatttga	gggtgtctgt	ctggagactt
2941	agagtttgtc	atgtgtgtgt	gtgtgtttgg	ttaatgtgg	tttatgggtt	ttctttctt
3001	tttttctttt	tttttttagt	ctacattagg	gggaagtgg	cgcctccat	gtgcagacag
3061	tgtgtcttta	tagattttc	taaggcattt	cccaatgtat	tcggtaattt	ctgatgtttc
3121	tgaagttccc	aggactcaca	caccggttcc	catctcaatt	gcccacccag	tgtgacaacc
3181	ctcgggtgtgg	atataccccc	gtggactcat	ggctcttccc	caccccccact	ttctataaat
3241	gtaggcctag	aatacgcctc	tctgttgcaa	aactcagcta	agttcctgct	tccacacttg
3301	tgttgaata	tctttagtta	gagggcaggg	gatgtcgtga	agatggcaag	aagaacacag
3361	tttcaaattt	ctggaaaaga	gcctgtgg	gagatctaaa	gatgtttagg	gaagagctcg
3421	actaaagaac	aatgaaataa	atggtccaa	gggaagtca		

FIG. 6

DEFINITION protein-arginine deiminase type-2 isoform X1 [Homo sapiens].
ACCESSION XP_016855637
VERSION XP_016855637.1
sequence number-4

1 msicgglsgn lpgssvqplk rmkpvsqygy rmcfkdeieif gyieaphkgf pvvldsprdg
61 nlkdfpvkel lgpdgvytr eplfesvtsl dsfgnlevsp pvtvngktyp lgriligssf
121 plsggrrmtk vvrdflkqaqq vqapvelysd wltvghvdef msfvppgtk kflllmasts
181 acyklfrekq kdghgeaimf kglggmsskr itinkilsne slvqenlyfq rcldwnrndl
241 kkelglteqd iidlpalfkm dedhraraff pnmvnmivld kdligpkpfg pqveeeclc
301 mhvrqllepl qlectfiddi sayhkflqev hcqtnvrrkp ftfkwvhmv

FIG. 7

DEFINITION Homo sapiens peptidyl arginine deiminase 3 (PADI3), mRNA.
ACCESSION NM_016233
VERSION NM_016233.2
sequence number-5

```
 1 agtgttgggg ttggcgccca cagctaagtc caacaccagc atgtcgctgc agagaatcg  
 61 gcgtgtgtcc ctggagcatc ccaccagcg ggtgtgtg gctggcgtgg agaccctcg  
121 ggacatttat gggtcagtgc ctgaggggac agaaatgtt gaggtctatg ggacgccttg  
181 cggtggacatc tacatcttc ccaacatgga gaggggccgg gagcgtgcag acaccaggcg  
241 gtggcgcttt gacgcgactt tggagatcat cgtggtcatg aactcccca gcaatgaccc  
301 caacgcacagc catgttcaga tttccctacca ctccagccat gagcctctgc ccctggccat  
361 tgcggtgctc tacctcacct gtgttgacat ctctctggat tgcgacctga actgtgagggg  
421 aaggcaggac aggaacttt tagacaagcg qcagtgggtc tgggggccc qttggatatgg  
481 cggcatcttg ctggtgaact gtgaccgtga tgatccgagc tgtgtatgtcc aggacaatttg  
541 tgaccagcac gtgcactgccc tgcaagaccc ggaagacatg tctgtcatgg tcctgcggac  
601 qcaggggccct qcagccccct ttgtqacca caaacttgtc ctccataacct ccagctatgaa
```

661 tgccaaacgg gcacaggctc tccacatctg cggcctgag gatgtgtgt aggccata
 721 gcatgtgcgt ggccaagata aggtgtccta tgaggtaccc cgcttgcatt gggatgagga
 781 gcgcttccttc gtggaaaggcc tgcctttccc tgcacccgc ttccacaggac tcacatccctt
 841 ccatgtcaact ctgcgtggacg actccaaacga ggatttctcg gcatccccata tccttcactga
 901 cactgtgggt ttccgagtgg caccctggat catgacgccc agcaactctgc cacccttaga
 961 ggtgtatgtg tgccgtgtga ggaacaacac gtgtttgtg gatgcgtgg cagagctggc
 1021 caggaaggcc ggctgcaagc tgaccatctg cccacaggcc gagaaccgca acgaccgctg
 1081 gatccaggat gagatggagc tggctacgt tcaggcgccg cacaagaccc tcccggtgg
 1141 ctttgactcc ccaaggaatg gggaaactgca ggatttccct tacaaaagaa tcctgggtcc
 1201 agattttgt tacgtgactc gggaaaccacg cgacaggctc gtgagtgcc tggactccctt
 1261 tgggaacctg gaggtcagcc ctccagtggt ggcataatggg aaagagtacc ccctggggag
 1321 gatccctcatt gggggcaacc tgcctgggtc aagtggccgc agggtcaccc aggtggtgcc
 1381 ggacttcctc catgcccaga aggtgcagcc cccctgtggag ctctttgtgg actgggtggc
 1441 cgtggggccat gtggatgagt ttctgagctt tgccctgtcc cccatggga aggcttcgg
 1501 gatgctcctg gccagccctg gggcctgctt caagctttc cagaaaaagc agaagtgtgg
 1561 ccacggggagg gcccctcctgt tccagggggt tggatgat gggcgggtca agaccatctc
 1621 catcaaccag gtgcctccta ataaagacct catcaactac aataagttt tgccagagctg
 1681 catcgactgg aaccgtgagg tgctgaagcg ggagctggc ctggcagagt gtgacatcat
 1741 tgacatccca cagctttca agaccgagag gaaaaaaagca acggcccttct tcctgactt
 1801 ggtgaacatg ctgggtctgg ggaagcacct gggcatcccc aagccctttg gccccatcat
 1861 caatggctgc tgctgcctgg aggagaaggt ggggtccctg ctggagccgc tggcctcca
 1921 ctgcacccctt attgtgact tcaactccata ccacatgtg catggggagg tgactgtgg
 1981 caccaatgtg tgcagaaagc ccttctctt caagtgggtgg aacatggtgc cctgagacag
 2041 ctccccaccca ccattctgtc cccctgggc gggcattggc ccagggtggg gagacagaga
 2101 caggccccctg aacgataagc accaagagac cccaaaggctc cagatggAAC actgagggtg
 2161 accgtccctc tcagaaggcct ttccctggg agtgtccatg cctcacccgc aaccatgtg
 2221 gttctcagac ttgaatcttc tcggcccccc aaaaagaagg acctcatttc ttatagcctc
 2281 tcctgtgatt caacacaacc catggagatg tcccttctc actctgaaat catccatttg
 2341 gggacaaatc cacattgggg tcttagaaaca tccacgtatc tcacagcca tcttgcctg
 2401 tgcacatccaa cagaggaagg atccatgatt ctgtttggg ccaattgtt cctctctgca
 2461 gaggaacaac cctaaaacca gaccactcca cgcaggacag gcaggagaga ttcttcctaa
 2521 agcctccccc ataaaaagg agctgtggat ccacttagat cagggcgaa ccatctttca
 2581 cccggccaag ctccgtccca gatgttgacc ctccacccagc gtgagctgtc acatagtagg
 2641 agcttctaga tgcacatgttgg agcaatgaga gttgtccctt agccttataa actccccatg
 2701 atctgacatg cagaaatcca gccttgcata gaatccttctt ggaatttctt ggagacgaaa
 2761 gatatctgggg gattgttggg tactaggag actgggtaca agggtaaaaa gtagttccca
 2821 taatacacat ggttgactat ggtgatccac ctgttgcata gttatattag gtgtctggag
 2881 aagggttgctt cattggccct gggacttctc tctgcaggag gagagaacgc tgcctctcct
 2941 ctggattgtt ctcaggctct ctgttggct ttgtcagcg tttccacatc ctgctctgt
 3001 gcaggagagg gggctaaagg gctggatcca ccaaggcgc tcacagccgg aaaactctgg
 3061 gaatgaacca ctgaatttcg gggatggggg tggggggcg gttctcgagg tgcgtgccag
 3121 ctacacgtgt gttctgtatg ggtccagctg cgttccatc actcgctaatt aaatcaacag
 3181 aaacacaaaa

FIG. 8

DEFINITION protein-arginine deiminase type-3 [Homo sapiens].
ACCESSION NP_057317
VERSION NP_057317.2
sequence number-6

1 mslqrivrvs lehptsavcv agvetlvdiy gsvpegttemf evygtpgvdi yispnmergr
61 eradtrrwrf datleiivvm nspsndlnds hvqisyhssh eplplayavl yltcvdisld
121 cdlncegrqd rnfvdkrqvw wgpsgyggil lvncdrddps cdvqdndqh vhclqdledm
181 ssvmvlrtqgp aalfddhkly lhtssyda kr aqvfhicgpe dvceayrhv l gqdkvsyevp
241 rlhgdeerff veglsfpdag ftglisfhvt llddsnedfs aspiftdtvv frvapwimtp
301 stlpplevyv crvrnnntcfv davaelarka gcklticpqa enrndrwiqd emelgyvqap
361 hktlpvvfds prngelqdfp ykrilgpdfg yvtreprdrs vsgldsfgnl evsppvvang
421 keyplgrili ggnlpgssgr rvtqvrvdfl haqkvqppve lfvdwlavgh vdeflsvfpa
481 pdkgffrmll aspgacfklf qekqkcgħgr allfqgvvdd eqvktising vlsnkdliny
541 nkfvqscidw nrevlkrelg laecdiidip qlfkterkka taffpdlvnm lvgkhlgip
601 kpfgrpiingc ccleekvrs leplglhctf iddfptpyhml hgevhcgtnv crkpfsfkww
661 nmvp

FIG. 9

DEFINITION PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X1, mRNA.
ACCESSION XM_011541571
VERSION XM_011541571.2
sequence number-21

1 tctccctgg tctgccattt cccctgtggg gggctgccc gggccaaaca gccagagg
61 ctccaccagc cctggaaatt acaggcttt ggagtaggg tgccagataa aataaaagat
121 gcctagttaa aattgaattt caggtaaata attaataatt ttttagtat aagtgtatcc
181 caaataacttc acgggagttac caaacaatga ttgactgtt acttggaaatt caagtccaaac
241 tgagctttct gtatttttat ttgctaagtc tggcagccct gcttaggagg agaaacagg
301 gtgtctgtcc ctccccactg gtgaggcac ac ctgacagggg cccagacacc agagtgggac
361 ctgcctgtat gagagaagct tggcttggt ttatccct tttcataaaaa aagctgctt
421 gagatgtaa agaatcatgt tggtaaga gtgaagagg tgttttggg gaggtctgcc
481 ttactctgag ggaagatggg aagccccccc agtggccagc tggaaatt tggggtagtc
541 ccctacgtat cctgtgcctc agttttcca tggtaaaca ggagtaatta ttcatccctg
601 ggctgcagt gaggatcaga tggactaatg tacaaaagg cttaggcagg ctgcctgt
661 taagagtgc gaggctggg gcttttatg ttatggattt gttgttata ttactccctgg
721 ggaggtggg ggcaagatgt tggactgagg agccacatgg ctgggttgg atcccgcccc
781 ggcacagac cttgggcaag ttacttaacc tggctttct catctgtata atggggagaca
841 ataatagata gtgcctccct tctagggtt tggagaggat tgcataatgg gctggcacac
901 agcatgagct cagtaagtgt cagtcactgt tatccccca gggcatgcat gccgctgcct
961 ccaggaggcc ctccaactcc cagtgccctc tccctacatg gttccact gagctgtgat
1021 cctgtggct tgagcccat ctggctgccc tggacaatgt tggctcttag ctccccagcg
1081 ggtcaacag tggggagagc agtgggtgg tctggccccc aggcatgacc tgagtgcaca
1141 gtgggcaccc cacagtcaa ggtgcatac tctggttcc aaggaggaac cccaggctag
1201 gtggaatgag gaaatgca ttcaggccc agataacatc acagataactt taaacctgca
1261 attaattttct gtttggatcc ccaaatccgg ccaagttcc cgggtgacac tgcagagctg
1321 tggAACAGCA tggcttgga gtcagacctg ttgcacatct ggctctgcca tggatgc
1381 cctcaggctt gtcacgaggc ctccctgagc ctcagttgc aatatgcca atggggatga
1441 tattgctcac cttacagctt gtaagatgtt agggagctgc gtgcggcagg tgcctggcat
1501 gcagcggca ctcaaccaga ttgaattccc ttcccaggaa agaattcaag gatatcttcc
1561 caccactaa ggagacatgt gatggatct ggggtgagaa ttgctttt taaagctctc
1621 agaactgtaa gggctccac ccaccaatgg ctgggttca ctgactctt cggtccca
1681 cctgcactgg ggatttatcg agcacatggc tggctctga cttccaga gtcgcccatt
1741 ctggctcgtg cctgaggcga cagaatgtt tggatgtctat gggacccctg gctggacat
1801 ctacatctt cccacatgg agagggccg ggagcgtgca gacaccaggc ggtggcgctt
1861 tgacgcgact ttggagatca tcgtggatca gaactcccc agcaatgacc tcaacgcacag
1921 ccatgttccat atttccatcc actccagccca tggcctctg cccctggcct atgcgggtgt
1981 ctacacttcc tggatgttca ttcctcttca ttggcgttca aactgtgagg gaaggcagga

2041 caggaacttt gtagacaaggc ggcagtgggt ctgggggccc agtgggtatg gccgcatttt
 2101 gctggtaac tgtgaccgtg atgatccgag ctgtgatgtc caggacaatt gtgaccagca
 2161 cgtgcactgc ctgcaagacc tggaaagacat gtctgtcatg gtcctgcgga cgcaaggccc
 2221 tgcagccctc tttgatgacc acaaacttgt cttccatacc tccagctatg atgccaacg
 2281 ggcacagggtc ttccacatct gcggtctga ggatgtgtgt gaggcctata ggcattgtgt
 2341 gggccaagat aagggtgtct atgaggtacc cgcgttgcatt gggatgagg agcgcttctt
 2401 cgtggaaaggc ctgtccttcc ctgatgcggg cttcacagga ctcatcttcc tccatgtcac
 2461 tctgctggac gactccaacg aggatttctc ggcattccct atcttcaactg acactgtgg
 2521 gttcccgagtgc acacccttga tcatgacgcc cagcactctg ccaccccttag aggtgtatgt
 2581 gtgccgtgtg aggaacaaca cgtgtttgtt ggtgcgggt gcagagctgg ccaggaaggc
 2641 cggctgcaag ctgaccatct gcccacagggc cgagaaccgc aacgaccgct ggtccagga
 2701 ttagatggag ctgggtctacg ttcaggcggc gcacaagacc ctcccggtgg tcttgactc
 2761 cccaaggaat ggggaaactgc aggattttcc ttacaaaaga atcctgggtc cagattttgg
 2821 ttacgtgact cgggaaaccac gcgcacagggtc tggatgtggc ctggacttcc ttgggaacct
 2881 ggaggtcagc cttccagtgg tggccaatgg gaaagagtac cccctggga ggtccctcat
 2941 tgggggcaac ctgcctgggt caagtggccg cagggtcacc caggtgggtc gggacttcct
 3001 ccatgcccac aagggtcagc ccccccgttga gcttttgg gactgggttgg ccgtgggcca
 3061 tggatgttgc tttctgagct ttgtccctgc ccccgatggg aagggttcc ggtatgttcc
 3121 ggccagccct gggccctgtc tcaagcttt ccaggaaaag cagaagtgtg gcacacgggag
 3181 gggcccttgc ttccagggggg ttgttgatgtc tgacgggtc aagaccatct ccatcaacca
 3241 ggtgccttcc aataaaagacc tcataacta caataagttt gtgcagagct gcatcgactg
 3301 gaaccgtgag gtgcgttgc gggagctggg cttggcagag tggatgttca ttgacatccc
 3361 acagcttcc aagaccgaga gggaaaaagc aacggccctc ttccctgact tggtaacat
 3421 gctgggtctg gggaaagcacc tgggcattttt caagccctt gggccatca tcaatggctg
 3481 ctgcgtgcctg gaggagaagg tgcggccct gctggagccg ctgggcctcc actgcaccc
 3541 cattgatgac ttcaactccat accacatgtc gcatggggag gtgcactgtg gcaccaatgt
 3601 gtgcagaaaag cccttcttca tcaagtgggtt gacatgggtt ccctgagaca gctcccaccc
 3661 accatccctgt cccctgggg cgggcattgg cccaggtggt ggagacagag acaggcccc
 3721 gaacgataag caccaagaga ccccaaggct ccagatggaa cactgagggt gaccgtccct
 3781 ctcagaagcc tttccctgg aagtgtccat gcctcacctg caacccatgt ggttctcaga
 3841 cttgaatctt ctcggccccc caaaaagaag gacccatattt cttatagcct ctcctgtgat
 3901 tcaacacaaac ccatggagat gtccttctt cactctgaaa tcatccattt gggacaaat
 3961 ccacattggg gtctagaaac atccacgtat ctcatcagcc atcttgcctt gtgcaccc
 4021 acagagggaaag gatccatgtat tctgtttgg tccaaattgtc tcctctctgc agaggaacaa
 4081 ccctaaaacc agaccactcc acgcaggaca ggcaggagag attcttccta aagcctccccc
 4141 cataaaaagg gagctgttgc tccacttaga tcaggcggtt accatcttc acccgccaa
 4201 gtcctgccttcc agatgttgc ctcaccat cgtgagctgt cacatagtag gagcttctag
 4261 atgcattgtgg aagcaatgtg agttgtccct tagccttata aactcccat gatctgacat
 4321 gcagaaaatcc agccttgc tggaaattttt tggagacgaa agtatctggg
 4381 ggattgttgg gtacttaggg aactgggttac aagggtgaaa agtagttccc ataataacaca
 4441 tgggtgacta tgggtatcca cttgtgtatg gttaatattt ggtgtctgg gaaagggtgct
 4501 tcattggccc tggacttctt ctctgcaggaa ggagagaacg ctgcctctcc tctggattgg
 4561 tctcaggctc tctgttggcc tttgggtcagc gttccacat ctcgtctgc tgcaggagag
 4621 ggggctaagg ggctggatcc accaaggcag ctcacagccg gaaaactctg ggaatgaacc
 4681 actgaattca gggatgggg gtggggggc gttctcgag gtgtgtgcca gctacacgtg
 4741 tggatgttgc tggatgttgc gctttccat cactcgtaa taaatcaaca gaaacacaaa

FIG. 10

DEFINITION protein-arginine deiminase type-3 isoform X1 [Homo sapiens].
ACCESSION XP_011539873
VERSION XP_011539873.1
sequence number-22

```
1 mfevygtpgv diyispnmer greradtrrw rfdatleii vmnspnsndln dshvqisyhs
61 sheplplaya vlyltcvdis ldcdlncegr qdrnfvdkrq wwgpsgygg ilivncdrdd
121 pscdvqdncd qvhclqdle dmsvmvlrtq gpaalfddh lvlhtssyda kraqvfhicg
181 pedvceayrh vlgqdkvsye vprlhdeer ffveglisfpd agftglisfh vtllddsned
241 fsaspiftdt vvfrvapwim tpstlpplev yvcrvrnntc fvdaaelar kagcklticp
301 qaenrndrwi qdemelgyvq aphkltpvf dsprngelqd fpykrilgpd fgyvtreprd
361 rsvsgldsfq nlevsppvva ngkeyplgri liggnlpqss grrvtqvrd flhaqkvqpp
421 velfvdwlav ghvdeflsvf papdgkgfrm llaspgacfk lfqekqkcg hgrallfqgvv
481 ddeqvktisi nqvlksnkli nynkfqvsci dwnrevlkre lglaecdiid ipqlfkterk
541 kataffpdlv nmlvlgkhlg ipkpfgpini gcccleekvr sileplglhc tfiddftpyh
601 mlhevchcgt nvcrkpfsfk wwnmv
```

FIG. 11

DEFINITION PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X2, mRNA.
ACCESSION XM_017001463
VERSION XM_017001463.1
sequence number-37

```
1 agaaaatggat ggatgtgact gtgtgcta ataaactttt ttatacaaac aggcagtagg
61 ccagattttgg cccacagttc ataatgtgct gatcctgacc taggcgagaa gagaaaccaa
121 atatgaaact gttgaagaac ttggactga attatgttgg aacttggtgc cctgggaggta
181 aagaggagaa ggtcgagcgg cagtgggtct gggggccca gggatgtggc ggcattttgc
241 tggtaactg tgaccgtgat gatccgagct gtgatgtcca ggacaattgt gaccagcacg
301 tgcactgcct gcaagacctg gaagacatgt ctgtcatggt cctgcggacg cagggccctg
361 cagcccttct tggatgaccac aaaccttgtcc tccataccctc cagctatgtat gccaaacggg
421 cacaggcttcc acatctgc ggtccgtggg atgtgtgtga ggcctatagg catgtgctgg
481 gccaagataa ggtgtccat gaggtaaaaa gcttgcattgg ggatgaggag cgcttcttcg
541 tggaaaggct gtccttccct gatccggct tcacaggact catctccctc catgtcactc
601 tgcggacga ctccaaacgag gatttctcgat catcccttat cttcaactgac actgtgggt
661 tccgagggtgc accctggatc atgacgccc gcaactctgcc acccccttagag gtgtatgtgt
721 gccgtgttag gaacaacacg tggatgtgg atgcgggtggc agagctggc agaaggccg
781 gctgcaagct gaccatctgc ccacaggccg agaaccgcaa cgaccgctgg atccaggatg
841 agatggagct gggctacgat caggccgc acaagaccct cccgggtggc tttgactcc
901 caaggaatgg ggaactgcag gatttccctt acaaaaagaat cctgggtcca gatttgggtt
961 acgtgactcg ggaaccacgc gacaggtctg tgagtgccct ggactccctt gggAACCTGG
1021 aggtcagccc tccagtggtg gccaatgggaa aagagtaccc cctggggagg atcctcattt
1081 gggcaacct gcctgggtca agtggccgca gggtcaccca ggtgggtggc gacttccctcc
1141 atggcccgaa ggtgcagccc cccgtggagc tctttgtgg ctgggtggc gtggggccatg
1201 tggatgagtt tctgagctt gtccttgc cccatggggaa gggctccgg atgctccctgg
1261 ccagccctgg ggcctgctt aagcttccctt aggaaaagcga gaagtggtgg cacgggaggg
1321 ccctccctttt ccagggggtt gttgatgatg agcaggtcaa gaccatctcc atcaaccagg
1381 tgctctccaa taaagacccat atcaactaca ataagttgt gcagagctgc atcgactgg
1441 accgtgaggt gctgaagccg gagctggcc tggcagagtg tgacatcatt gacatccccac
1501 agtcttcaa gaccgagagg aaaaaaagcga cggccttctt ccctgacttg gtgaacatgc
1561 tggtgctggg gaagcacctg ggcataccca agccctttgg gcccattatc aatggctgt
1621 gctgcctgg gggaaagggtg cggccctgc tggagccgt gggccctccac tgcacccatc
1681 ttgatgactt cactccatc acatgtgtgc atggggagggt gcactgtggc accaatgtgt
1741 gcagaaaagcc cttcttccat aagtgggtgg acatgggtgcc ctgagacagc tcccacccac
1801 catccctgtcc ccctggggcc ggcattggcc cagggtggtgg agacagagac agggccctgt
1861 acgataagca ccaagagacc ccaaggctcc agatggaaaca ctgagggtga cctgcccctct
1921 cagaaggcctt ttccctggaa gtgtccatgc ctcacctgca acccatgtgg ttctcagact
1981 tgaatcttctt cggccccccca aaaagaagga cctcatttct tatagcctct cctgtgattc
```

2041 aacacaaccc atggagatgt ccccttctca ctctgaatac atccatttg ggacaaaatcc
2101 acattgggt ctagaaacat ccacgtatct catcagccat cttgtctgt gcacccaac
2161 agaggaagga tccatgattc tgcttggtc caattgcttc ctctctgcag aggaacaacc
2221 ctaaaaaccag accactccac gcagacagg caggagagat tcttctaaa gcctccccca
2281 taaaaaggga gctgtggatc cacttagatc agggcggAAC catcttcac cggccaagc
2341 tcctgcccag atgttgaccc tcaccagcg tgagctgtca catagtagga gcttctagat
2401 gcatgtggaa gcaatgagag ttgtcccta gccttataaa ctccccatga tctgacatgc
2461 agaaatccag cttgtccag aatcttcctg gaatttctt gagacaaag tatctgggg
2521 atttgtgggt actagggaga ctggtacaa gggtaaaag tagttccat aatacacatg
2581 gttgactatg gtgatccacc ttgtgtatggtaatattagg tgtctggaga aggttgcttc
2641 attggccctg ggacttctct ctgcaggagg agagaacgct gccttcctc tggattggtc
2701 tcaggctctc tggcgttgc tggcgttgc ttccacatcc tgctctgctg caggagaggg
2761 ggctaagggg ctggatccac caagcagct cacagcggga aaactctggg aatgaaccac
2821 tgaattcagg ggatgggggt gggggggcgg ttctcgagggt gtgtgccagc tacacgtgt
2881 ttctgtatgg gtccagctgc gttccatca ctgcataata aatcaacaga aacacaaa

FIG. 12

DEFINITION protein-arginine deiminase type-3 isoform X2 [Homo sapiens].
ACCESSION XP_016856952
VERSION XP_016856952.1
sequence number-38

1 msvmvlrtqg paalfddhkl vlhtssydk raqvfhicgp edvceayrhv lgqdkvsyev
61 prlhdeerf fveglsflda gftglisfhv tllddsnedf saspiftdtv vfrvapwimt
121 pslpplevy vcrvrnntcf vdavaelark agcklticpq aenrnndrwiq demelgyvqa
181 phktpvvfd sprngelqdf pykrilgpdf gyvtreprdr svsgldsfgn levsppvvan
241 gkeyplgril iggnlpgsss rrvtqvrvdf lhaqkvqppv elfvdwlavg hvdeflsfv
301 apdgkgfrml laspgacfkl fqekqkcgħg ralffqgvvd deqvkritisinqvlsnkdl
361 ynkfvqscid wnrevlkrel glaedidi pqlfkterkk ataffpdln mlvlghlgi
421 pkpfgpipiing cccleekvrs lleplglhct fiddftyhm lhgevhcgtn vcrkpfssfw
481 wnmvp

FIG. 13

DEFINITION PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X2, mRNA.
ACCESSION XM_017001463
VERSION XM_017001463.1:c
sequence number-39

1 agaaatggat ggatgtgact gtgtgctaat aaaactttat ttatacaaac aggcaaggtagg
61 ccagatttgg cccacagttc ataatgtgct gatcctgacc taggcggagaa gagaaacccaa
121 atatgaaact gttgaagaac ttgggactga attatgttgg aacttgggtgc cctggggagtt
181 aagaggagaa ggtcgagcgg cagtgggtct gggggcccaag tgggtatggc ggcattttgc
241 tggtaactg tgaccgtat gatccgagct gtgtatgtcca ggacaattgt gaccagcacg
301 tgcactgcct gcaagacctg gaagacatgt ctgtcatggt cctgcggacg cagggccctg
361 cagcccttctt tgatgaccac aaacttgtcc tccatacctc cagctatgtat gccaaacggg
421 cacaggctt ccacatctgc ggtcctgagg atgtgtgtga ggcctatagg catgtgttgg
481 gccaagataa ggtgtccat gaggatcccc gcttgcattgg ggatgaggag cgcttcttcg
541 tggaaaggcct gtccttcctt gatccggct tcacaggact catcttccttc catgtcactc
601 tgctggacga ctccaaacagag gatttctcgat cattttctat cttcaactgac actgtgggt
661 tccgagtggc accctggatc atgacgcccc gcaactctgccc acccccttagag gtgtatgt
721 gccgtgttag gaaacaacacg tggtttgtgg atgcgggtggc agagctggcc aggaaggccg
781 gctgcaagct gaccatctgc ccacaggccg agaaccgcaa cgaccgctgg attcaggatg
841 agatggagct gggctacgtt caggccgc acaagaccct cccgggtggc tttactccc

901 caaggaatgg ggaactgcag gatttccctt acaaaaagaat cctgggtcca gattttgggt
961 acgtgactcg ggaaccacgc gacaggtctg ttagtgtccct ggactccttt gggAACCTGG
1021 aggtcagccc tccagtggtg gccaatggga aagagtaccc cctggggagg atcctcattg
1081 gggcaacct gcctgggtca agtggccgca gggtaaccca ggtggcgcg gacttcctcc
1141 atgcccagaa ggtgcagccc cccgtggagc tctttgtgga ctgggtggcc gtggggcatg
1201 tggatgagtt tctgagctt gtccctgccc ccgatgggaa gggctccgg atgctcctgg
1261 ccagccctgg ggcctgcttc aagcttcc aggaaaagca gaagtgtggc cacgggaggg
1321 ccctcctgtt ccagggggtt gttgatgatg aycaggtcaa gaccatctcc atcaaccagg
1381 tgctctccaa taaagaccc tcataactaca ataagttgt gcagagctgc atcgactgg
1441 accgtgaggt gctgaagcgg gagctggcc tggcagagtg tgacatcatt gacatcccac
1501 agctctcaa gaccgagagg aaaaaaagcaa cggccttctt ccctgacttg gtgaacatgc
1561 tggtgctggg gaagcacctg ggcattccccca agcccttgg gcccatcata aatggctgt
1621 gctgcctgga ggagaagggt cggtccctgc tggagccgt gggcctccac tgcacccatca
1681 ttgatgactt cactccatac cacatgctgc atggggagggt gcaactgtggc accaatgtgt
1741 gcagaaagcc ctctctttc aagtggtgga acatggtgcc ctgagacagc tcccacccac
1801 catcctgtcc ccctggggcg ggcattggcc caggtggtg agacagagac aggccccgt
1861 acgataagca ccaagagacc ccaaggctcc agatggaaaca ctgagggtga cgcgtccctct
1921 cagaagcctt ttccctggaa gtgtccatgc ctcacctgca acccatgtgg ttctcagact
1981 tgaatcttct cggccccccca aaaagaagga ctcatttttct tatagcctct cctgtgattc
2041 aacacaaccc atggagatgt ccccttctca ctctgaaatc atccatttgg ggacaatcc
2101 acattgggt ctagaaacat ccacgtatct catcagccat ctgtcctgt gcacccataac
2161 agaggaagga tccatgattc tgcttggtc caattgcttc ctctctgcag aggaacaacc
2221 ctaaaaaccag accactccac gcagacagg caggagagat tcttcctaaa gcctcccccac
2281 taaaaaggga gctgtggatc cacttagatc agggcggAAC catcttcac cccggccaagc
2341 tcctgcccag atgttgaccc tcaccagcg tgagctgtca catagtagga gcttcttagat
2401 gcatgtggaa gcaatgagag ttgtccctta gccttataaa ctcccatga tctgacatgc
2461 agaaatccag ccttgcctcg aatccctctg gaatttcttg gagacgaaag tatctgggg
2521 attgtgggt actagggaga ctgggtacaa gggtaaaaag tagttccat aatacacatg
2581 gttgactatg gtgatccacc ttgtgtggtaatattagg tgtctggaga aggttgcttc
2641 attggccctg ggacttctt ctgcaggagg agagaacgct gcctccctc tggattggtc
2701 tcaggctctc tggccctt tggtcagcgt ttccacatcc tgctctgctg caggagaggg
2761 ggctaagggg ctggatccac caaggcagct cacagcggga aaactctggg aatgaaccac
2821 tgaattcagg ggatgggggt gggggggcgg ttctcgagggt gtgtgccagc tacacgtgt
2881 ttctgtatgg gtccagctgc gttccatca ctgcataata aatcaacaga aacacaaa

FIG. 14

DEFINITION PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X2, mRNA.
ACCESSION XM_017001463
VERSION XM_017001463.1:c
sequence number-40

1 agaaatggat ggatgtgact gtgtgctaattaaaactttat ttatacaaac aggcagtagg
61 ccagatttgg cccacagttc ataatgtgct gatcctgacc taggcggagaa gagaaacccaa
121 atatgaaaact gttgaagaac ttgggactga attatgttgg aacttgggtgc cctggggagtt
181 aagaggagaa ggtcgagcgg cagttggctt gggggcccaag tgggtatggc ggcattttgc
241 tggtaactg tgaccgtat gatccgagct gtgtatgtcca ggacaattgt gaccagcacg
301 tgcactgcct gcaagacctg gaagacatgt ctgtcatggt cctgcggacg cagggccctg
361 cagcccttctt tgatgaccac aaacttgc tccataccctc cagctatgtat gccaaacggg
421 cacaggcttcc acatctgc ggtccgttggg atgtgtgtga ggcctatagg catgtgttgg
481 gccaagataa ggtgtccat gaggtaaaaa gcttgcattgg ggatgaggag cgcttcttcg
541 tggaaaggct gtcctccct gatgcggct tcacaggact catctccatc catgtcactc
601 tgctggacga ctccaaacgag gatttctcgat cattccctat ctcaactgac actgtgggt
661 tccaggtggc accctggatc atgacgcccc gcaactctgccc acccccttagag gtgtatgt
721 gccgtgttag gaaacaacacg tgggtttgtgg atgcgggtggc agagctggcc aggaaggccg
781 gctgcaagct gaccatctgc ccacaggccg agaaccgcac gacccgtgg atccaggatg
841 agatggagct gggctacgtt caggcggccgc acaagaccct cccgggtggc tttactccc

901 caaggaatgg ggaactgcag gatttccctt acaaaaagaat cctgggtcca gattttgggt
961 acgtgactcg ggaaccacgc gacaggtctg tgagtggcct ggactccttt gggAACCTGG
1021 aggtcagccc tccagtggtg gccaatggga aagagtaccc cctggggagg atcctcattg
1081 gggcaacct gcctgggtca agtggccgca gggtaaccca ggtggcgcg gacttcctcc
1141 atgcccagaa ggtgcagccc cccgtggagc tctttgtgga ctgggtggcc gtggggcatg
1201 tggatgagtt tctgagctt gtccctgccc ccgatgggaa gggctccgg atgctcctgg
1261 ccagccctgg ggcctgcttc aagcttcc aggaaaagca gaagtgtggc cacgggaggg
1321 ccctcctgtt ccagggggtt gttgatgatg aycaggtcaa gaccatctcc atcaaccagg
1381 tgctctccaa taaagaccc tcataactaca ataagttgt gcagagctgc atcgactgg
1441 accgtgaggt gctgaagcgg gagctggcc tggcagagtg tgacatcatt gacatcccac
1501 agctctcaa gaccgagagg aaaaaaagcaa cggccttctt ccctgacttg gtgaacatgc
1561 tggtgctggg gaagcacctg ggcattccccca agcccttgg gcccatcata aatggctgt
1621 gctgcctgga ggagaagggt cggtccctgc tggagccgt gggcctccac tgcacccatca
1681 ttgatgactt cactccatac cacatgctgc atggggaggt gcaactgtgg accaatgtgt
1741 gcagaaagcc ctctctttc aagtggtgga acatggtgcc ctgagacagc tcccacccac
1801 catcctgtcc ccctggggcg ggcattggcc caggtggtg agacagagac aggccccgt
1861 acgataagca ccaagagacc ccaaggctcc agatggaaaca ctgagggtga cgcgtccctct
1921 cagaagcctt ttccctggaa gtgtccatgc ctcacctgca acccatgtgg ttctcagact
1981 tgaatcttct cggccccccca aaaagaagga ctcatttttct tatagcctct cctgtgattc
2041 aacacaaccc atggagatgt ccccttctca ctctgaaatc atccatttgg ggacaatcc
2101 acattgggt ctagaaacat ccacgtatct catcagccat ctgtcctgt gcacccataac
2161 agaggaagga tccatgattc tgcttggtc caattgcttc ctctctgcag aggaacaacc
2221 ctaaaaaccag accactccac gcagacagg caggagagat tcttcctaaa gcctcccccac
2281 taaaaaggga gctgtggatc cacttagatc agggcggAAC catcttcac cccggccaagc
2341 tcctgcccag atgttgaccc tcaccagcg tgagctgtca catagtagga gcttcttagat
2401 gcatgtggaa gcaatgagag ttgtccctta gccttataaa ctcccatga tctgacatgc
2461 agaaatccag ccttgtccag aatccctctg gaatttcttg gagacagaaag tatctgggg
2521 attgtgggt actagggaga ctgggtacaa gggtaaaaag tagttccat aatacacatg
2581 gttgactatg gtatccacc ttgtgtggta taatattagg tgtctggaga aggttgcttc
2641 attggccctg ggacttctt ctgcaggagg agagaacgct gcctccctc tggattggtc
2701 tcaggctctc tggccctt tggtcagcgt ttccacatcc tgctctgctg caggagaggg
2761 ggctaagggg ctggatccac caaggcagct cacagcggga aaactctggg aatgaaccac
2821 tgaattcagg ggatgggggt gggggggcgg ttctcgagggt gtgtgccagc tacacgtgt
2881 ttctgtatgg gtccagctgc gttccatca ctgcctaata aatcaacaga aacacaaa

FIG. 15

DEFINITION PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X3, mRNA.

ACCESSION XM_011541572

VERSION XM_011541572.2

sequence number-49

1 cctaaggggc ctcagggca gtgtgggt tggcggccac agctaagtcc aacaccagca
61 tgcgtgc gagaatcgat cgtgtgtccc tggagcatcc caccagcgcg gtgtgtgtgg
121 ctggcgtgaa gaccctcgat gacatttatg gtcagtgcc tgagggcaca gaaatgtttt
181 aggtctatgg gacgcctggc gtggacatct acatctctcc caacatggag agggggccggg
241 agcgtcaga caccaggcg tggcgcttt acgcgacttt ggagatcata gtggcatga
301 actccccccag caatgaccc aacgacagcc atgttcagat ttccatccac tccagccatg
361 agcctctgc cctggccat gcggtgcctt acctcaccc tggatcata tctctggatt
421 ggcacatgaa ctgtgaggga aggacggaca ggaactttgt agacaagcgg cagtgggtct
481 gggggccca ggggtatggc ggcacatctgc tggtaactg tgaccgtat gatccgagct
541 gtatgtcca ggacaattgt gaccacgcg tgcactgcct gcaagacctg gaagacatgt
601 ctgtcatgtt cctgcggacg caggccctt cagccctt tggatgaccac aaacttgc
661 tccataccctc cagctatgtt gccaacggg cacaggctt ccacatctgc ggtccctgagg
721 atgtgtgtga ggcctatagg catgtgtgg gccaagataa ggtgtccat gaggtacccc
781 gcttgcattgg ggatggaggag cgcttctcg tggaaaggcc gtcctccct gatgcccggct
841 tcacaggact catctccctc catgtcactc tgctggacga ctccaaacgag gatttctcg

901 catcccctat cttcaactgac actgtggtgt tccgagtgcc accctggatc atgacgcccc
961 gcactctgcc accccctagag gtgtatgtgt gccgtgttag gaacaacacg tttttgtgg
1021 atgcgggtggc agagctggcc aggaaggccg gctgcaagct gaccatctgc ccacaggccg
1081 agaaccccaa cgaccgctgg atccaggatg agatggagct gggctacgtt caggcgccgc
1141 acaagaccct cccgggtggtc tttgactccc caaggaatgg ggaactgcag gatttccctt
1201 acaaaaagaat cctggtaaag tggccgcagg gtcacccagg tggtgccggat cttectccat
1261 gcccagaagg tgcaaaaaaa cgtggagctc tttgtggact ggttgccgt gggccatgtg
1321 gatgagttc tgagctttgt ccctgccccca gatgggaagg gcttcggat gtcctggcc
1381 agccctgggg cctgcttcaa gctttccag gaaaagcaga agtgtggcca cgggaggggcc
1441 ctcctgttcc ag

FIG. 16

DEFINITION protein-arginine deiminase type-3 isoform X3 [Homo sapiens].
ACCESSION XP_011539874
VERSION XP_011539874.1
sequence number-50

1 mslqrivrvs lehptsavcv agvetlvdiy gsvpegttemf evygtpgvdi yispnmergr
61 eradtrrwrf datleiivvm nspnsndlnds hvqisyhssh eplplayavl yltcvdisld
121 cdlncegrqd rnfvdkrqwy wgpsgyggil lvncdrddps cdvqdndqh vhclqdledm
181 svmvrlrtqgp aalfddhkly lhsssydakr aqvfhicgpe dvceayrhvly qdkvsyevp
241 rlhgdeerff veglsfpdag ftglisfhvt llddsnedfs aspiftdtvv frvapwimtp
301 stlpplevyy crvrnnntcfv davaelarka gcklticpqa enrndrwiqd emelgyvqap
361 hktlpvvfds prngelqdfp ykrilvkwpq ghpaggaglpp cpegaaprga lcglvgrgpc
421 g

FIG. 17

DEFINITION PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X3, mRNA.
ACCESSION XM_011541572
VERSION XM_011541572.2:c.
sequence number-59

1 cctaaggggc ctcagggca gtgtgggt tggcggccac agctaagtcc aacaccagca
61 tgtcgctgca gagaatcgta cgtgtgtccc tggagcatcc caccagcgcg gtgtgtgtgg
121 ctggcgtgg aaccctcgta gacatttatg gtcagtgcc tgagggcaca gaaatgtttg
181 aggcttatgg gacgcctggc gtggacatct acatctctcc caacatggag agggggccggg
241 agcgtcaga caccaggccc tggcgttttgc acgcgacttt ggagatcatc gtggcatga
301 actccccccag caatgaccc aacgacagcc atgttcagat ttccattaccac tccagccatg
361 agcctctgccc cctggcttat gcggtgtct acctcaccc ttttgacatc tctctggatt
421 ggcacactgaa ctgtgaggga aggccggaca ggaactttgt agacaagcgg cagtgggtct
481 gggggccca ggggtatggc ggcacatcttc tggtaactg tgaccgtat gatcccgagct
541 gtgtatgtcca ggacaatttg gaccagcacg tgcactgcct gcaagacctg gaagacatgt
601 ctgtcatgtt cctgcggacg cagggccctg cagccctt ttttgacatc aaacttgtcc
661 tccataccctc cagctatgtt gccaacggg cacaggctt ccacatctgc ggtccctgagg
721 atgtgtgtga ggcctataagg catgtgtgg gccaagataa ggtgtccat gaggtacccc
781 gcttgcattgg ggtatggagg cgcttcttc tggaaaggccct gtcctccct gatgcccggct
841 tcacaggact catctccctc catgtcactc tgctggacga ctccaaacgag gatttctcg
901 catcccctat cttcaactgac actgtggtgt tccgagtgcc accctggatc atgacgcccc
961 gcactctgcc accccctagag gtgtatgtgt gccgtgttag gaacaacacg tttttgtgg
1021 atgcgggtggc agagctggcc aggaaggccg gctgcaagct gaccatctgc ccacaggccg

1081 agaaccgcaa cgaccgctgg atccaggatg agatggagct gggctacgtt caggcgccgc
1141 acaagaccct cccgggtggc tttgactccc caaggaatgg ggaactgcag gatttccctt
1201 acaaaaagaat cctggtaaag tggccgcagg gtcacccagg tggtgccggg cttccctccat
1261 gccccagaagg tgcagccccc cgtggagctc tttgtggact gggtggccgt gggccatgtg
1321 gatgagtttc tgagctttgt ccctgcccc gatggaaagg gcttccggat gctcctggcc
1381 agccctgggg cctgcttcaa gctctccag gaaaagcaga agtgtggcca cgggaggggcc
1441 ctccctgttcc ag

FIG. 18

DEFINITION Homo sapiens peptidyl arginine deiminase 4 (PADI4), mRNA.
ACCESSION NM_012387
VERSION NM_012387.3
sequence number-61

1 agccagaggg acgagctagc ccgacgatgg cccagggac attgatccgt gtgacccag
61 agcagcccac ccatgccgtg tgtgtgctgg gcacccgtac tcagcttgc atctgcagct
121 ctgcccctga ggactgcacg tccttcagca tcaacgcctc cccaggggtg gtcgtggata
181 ttgcccacgg ccctccagcc aagaagaaat ccacagggttc ctccacatgg cccctggacc
241 ctgggttaga ggtgaccctg acgtgaaag tggccagttgg tagcacaggc gaccagaagg
301 ttcagatttc atactacgga cccaagactc caccagtcaa agctctactc tacctcaccg
361 ggggtggaaat ctccttggtc gcagacatca cccgcacccgg caaatgtgaag ccaaccagag
421 ctgtgaaaga tcagaggacc tggacctggg gcccttggtt acagggtgcc atcctgctgg
481 tgaactgtga cagagacaat ctcgaatctt ctgcccatttgc ctgcgaggat gatgaagtgc
541 ttgacagcga agacactgcag gacatgtcgc ttagtggccctt gaggcacgaag acccccaagg
601 acttcttcac aaaccataca ctggtgctcc acgtggccag gtctgagatg gacaaagtga
661 ggggtttca ggccacacgg ggcaactgtt cctccaatgtt cagctgttgc ttgggtccca
721 agtggccctc tcactacctg atggtccccg gtggaaagca caacatggac ttctacgtgg
781 aggccttcgc tttcccgac accgacttcc cggggctcat tacccctcacc atctccctgc
841 tggacacgtc caacctggag ctcccccggg ctgtgggttt ccaagacagc gtggctttcc
901 gcgtggcgcc ctggatcatg acccccaaca cccagccccc gcaggagggtg tacgcgtgca
961 gtattttga aaatgaggac ttccctgaatgtt cagtgactac tctggccatg aaagccaatgt
1021 gcaagctgac catctccctt gaggaggaga acatggatga ccagtggatg caggatgaaa
1081 tggagatcgg ctacatccaa gccccacaca aaacgctgcc cgtggcttcc gactctccaa
1141 ggaacagagg cctgaaggag ttcccatca aacgcgtgtt gggccatgtt tttggctatg
1201 taactcgagg gccccaaaca ggggttatca gtggactgaa ctcccttggg aacctggaaag
1261 tgagcccccc agtcacagtcc aggggcaagg aatacccgctt gggcaggatt ctcttcgggg
1321 acagctgtta tcccagcaat gacagccggc agatgcacca ggcctgcag gacttcctca
1381 gtgcccagca ggtgcaggcc cctgtgaagc tctattctga ctggctgtcc gtggccacag
1441 tggacgagtt ccttagctt gtgccagcac ccgacaggaa gggctccgg ctgctcctgg
1501 ccagccccag gtcctgtac aaactgttcc aggaggcagca gaatgagggc cacggggagg
1561 ccctgctgtt cgaagggttc aaaaaaaaaa aacagcagaa aataaaagaac attctgtcaa
1621 acaagacatt gagagaacat aattcatttgc tggagagatg catcgactgg aaccgcgagc
1681 tgctgaagcg ggagctggc ctggccgaga gtgacatcat tgacatcccg cagctttca
1741 agctcaaaga gttctctaag gcggaaagctt ttttcccaaa catggtaac atgctgggtc
1801 tagggaaagca cctgggcatac cccaaaggctt cggggccgtt catcaacggc cgctgctgcc
1861 tggaggagaa ggtgtgttcc ctgtggagc cactgggcctt ccagtgacc ttcatcaacg
1921 acttcttcac ctaccacatc aggcatgggg aggtgcactg cggcaccaac gtgcgcagaa
1981 agcccttcctc cttcaagttgg tggAACATGG tggccctgc gcatcttccc tggcgtcc
2041 tccctcctgg ccagatgtcg ctgggtccctc tgcagttgg caagaagag ctcttgcata
2101 tattgtggct ccctggggc ggccagccctt cccagcagtg gcttgcatttcc ttctcctgt
2161 atgtcccagt ttccactctt gaaatccca acatggtcctt agcactgcac actcagttt
2221 gctctaagaa gctgcaataa agtttttttta agtcacttttgc tacatgtt

FIG. 19

DEFINITION protein-arginine deiminase type-4 [Homo sapiens].
ACCESSION NP_036519
VERSION NP_036519.2
sequence number-62

1 maqgtlirvt peqpthavcv lgtltqlidc ssapedctsf sinaspgvvv diahgppakk
61 kstgsstwpl dpgvevtltm kvasgstgdq kvqisyygpk tppvkallyl tgveislcad
121 itrtrgkvkpt ravkdqrtwt wgpccggail lvncdrdnle ssamdcde vldsedlqdm
181 slmtlstktp kdfftnhtlv lhvarsemdk vrvfqatrgk lsskcsvvlg pkwpshylmv
241 pggkhnmfdfy vealafpdtd fpgliltis lldtsnlelp eavvfqdsvv frvapwimtp
301 ntqpqevya csifenedfl ksvttlamka kcklticpee enmddqwmqd emeigyiqaq
361 hktlpvvfd s prnrglkefp ikrvmpdfg yvtrgpqtgg isgldsfqnl evsppvtvrg
421 keyplgrilf gdscypsnds rqmhqalqdf lsqqqvqapv klysdlsvg hvdeflfsfvp
481 apdrkgfrll lasprscykl fqeinqneghg eallfegikk kkqqkiknil snktlrehs
541 fvercidwnr ellkrelgla esdiidipql fklkefskae affpnmvnml vlgkhlgipk
601 pfgpvingrc cleekvcsl1 eplglqctfi ndfftyhirh gevhcgtvnr rkpforskwwn
661 mvp

FIG. 20

DEFINITION PREDICTED: Homo sapiens peptidyl arginine deiminase 4 (PADI4), transcript variant X3, mRNA.
ACCESSION XM_011541152
VERSION XM_011541152.1:c
sequence number-67

1 tctacacctac cggggtgaa atctccttgc ggcgcagacat cacccgcacc ggcaaagtga
61 agccaaccag agctgtaaa gatcagacat gcaggacatg tcgctgatga ccctgagcac
121 gaagaccccc aaggacttc tcacaaacca tacactggg ctccacgtgg ccaggtctga
181 gatggacaaa gtgagggtgt ttccaggcc acggggcaaa ctgtcctcca agtgcagcgt
241 agtcttgggt cccaagtggc cctctacta cctgtatggc cccgggtggaa agcacaacat
301 ggacttctac gtggaggccc tcgcttccc ggacaccgac ttcccggggc tcattaccct
361 caccatctcc ctgctggaca cgtccaaacctt ggagctcccc gaggtgtgg tgttccaaga
421 cagcgtggtc ttccgcgtgg cgccctggat catgacccccc aacaccgc ccccgaggg
481 ggtgtacgcg tgcatgtttt ttgaaaaatga ggacttcctg aagtcatgtg ctactctggc
541 catgaaagcc aagtgcacgc tgaccatctg ccctgaggag gagaacatgg atgaccatgg
601 gatgcaggat gaaatggaga tcggctacat ccaagccca cacaaacgc tgcccggtgg
661 ctgcacttc ccaagggaca gaggcctgaa ggagttccc atcaaacgc tgatgggtcc
721 agattttggc tatgttaactc gaggccccca aacagggggt atcagtggac tggactccct
781 tggaaacctg gaagtggaccc ccccaactc acgtcaggggc aaggaataacc cgctggcag
841 gattctcttc gggggcaggt gttatcccac caatgcacgc cggcagatgc accaggccct
901 gcaggacttc ctcagtgccc agcagggtgca ggcccctgtg aagctctatt ctgactggc
961 gtccgtgggc cacgtggacg agttctggat ctttgtgcca gcacccgaca ggaagggctt
1021 ccggctgctc ctggccagcc ccaggtctc ctacaaactg ttccaggagc agcagaatga
1081 gggccacggg gaggccctgc tggttcaagg gatcaagaaa aaaaaacacgc agaaaataaa
1141 gaacattctg tcaaacaaga cattggagaa acataattca tttgtggaga gatgcacatcg
1201 ctggaaaccgc gagctgtca agcggggact gggcctggcc gagagtgaca tcattgacat
1261 cccgcagctc ttcaagctca aagagttctc taaggcggaa gctttttcc ccaacatgg
1321 gaacatgctg gtgcttaggaa agcacctggg catccccaaag cccttcgggc cccgtcatcaa
1381 cggccgctgc tgcctggagg agaagggtgt ttccctgtc gagccactgg gcctccagtg
1441 cacccatcata aacgacttc tcacccatca catcaggcat ggggagggtgc actgcggcac
1501 caacgtgcgc agaaaaggct ttccttcaa gtgggtggaa atgggtccct gagcccatct
1561 tccctggcgt cctctccctc ctggccagat gtgcgtgggt cctctgcagt gtggcaagca
1621 agagcttttg tgaatattgt ggctccctgg gggcggccag ccctcccaagc agtggcttgc
1681 tttcttctcc tggatgtcc cagttccca ctctgaagat cccaaacatgg tccttagcact
1741 gcacactcag ttctgctcta agaagctgca ataaagttt tttaagtccac ttgtatcatg
1801 a

FIG. 21

DEFINITION PREDICTED: Homo sapiens peptidyl arginine deiminase 4 (PADI4), transcript variant X8, mRNA.

ACCESSION XM_011541157

VERSION XM_011541157.1:c

sequence number-68

1 agctctactc tacctcaccc gggtgaaat ctccttgtgc gcagacatca cccgcaccgg
61 caaagtgaag ccaaccagag ctgtgaaaaga tcagaggacc tggacctggg gcccattgtgg
121 acagggtgcc atcctgctgg tgaactgtga cagagacaat ctcgaatctt ctgccccatgg
181 ctgcgaggat gatgaagtgc ttgacagcga agacactgcag gacatgtcgc tgatgaccct
241 gagcacgaag accccccaagg actttttcac aaaccataca ctgggtctcc acgtggccag
301 gtctgagatg gacaaagtga gggtgtttca ggccacacga gctcccccag gctgtgggt
361 tccaaagacag cgtggtcttc cgcgtggcgc cctggatcat gacccccaac acccagcccc
421 cgcaggaggt gtacgcgtgc agtatttttgg aaaatgagga cttcctgaag tcagtgacta
481 ctctggccat gaaagccaag tgcaagctga ccattctgccc tgaggaggag aacatggatg
541 accagtggat gcaaggatgaa atggagatcg gctacatcca agccccacac aaaacgctgc
601 ccgtggctt cgactctca aggaacagag gcctgaaggaa gttttccatc aaacgcgtga
661 tgggtccaga ttttggctat gtaactcgag ggcccaaacc agggggatc agtggactgg
721 actcccttgg gaacctggaa gtgagccccc cagtcacagt caggggcaag gaatacccg
781 tggcaggat tctttcggg gacagctgtt atcccaagcaa tgacagccgg cagatgcacc
841 aggccctgca ggacttcctc agtggcccaggc aggtgcagggc ccctgtgaag ctctatttg
901 actggctgtc cgtggccac gtggacgagt tcctgagct tggccagca cccgacagga
961 agggcttccg gtcgtccctg gccagccccc ggtccctgcta caaactgttc caggagcagc
1021 agaatgaggg ccacggggag gcctgtgt tcgaaggat caaaaaaaa aaacagcaga
1081 aaataaaagaa cattctgtca aacaagacat tgagagaaca taattcattt gtggagagat
1141 gcatcgactg gaaccgcgag ctgctgaagc gggagctggg cctggccgag agtgcacatca
1201 ttgacatccc gcagctctc aagctcaaag agttctctaa ggcgaaagct tttttccca
1261 acatggtgaa catgtgggt cttaggaagc acctggccat ccccaagcccc ttccgggcccc
1321 tcatcaacgg ccgctgtgc ctggaggaga aggtgtgttc cctgtggag ccactgggcc
1381 tccagtgcac cttcatcaac gacttctca cctaccacat cagggatggg gaggtgcact
1441 gcggcaccaa cgtgcgcaga aagccctct cttcaagtg gtggacatg gtgcctgag
1501 cccatcttcc ctggcgtcct ctccctcctg gccagatgtc gctgggtcct ctgcagtgt
1561 gcaagcaaga gctctgtga atattgtggc tccctgggg cggccagccc tcccagcagt
1621 ggcttgctt cttctctgt gatgtcccag tttccactc tgaagatccc aacatggtcc
1681 tagcaactgca cactcagttc tgctctaaga agctgcaata aagtttttt aagtcacttt
1741 gtacatga

FIG. 22

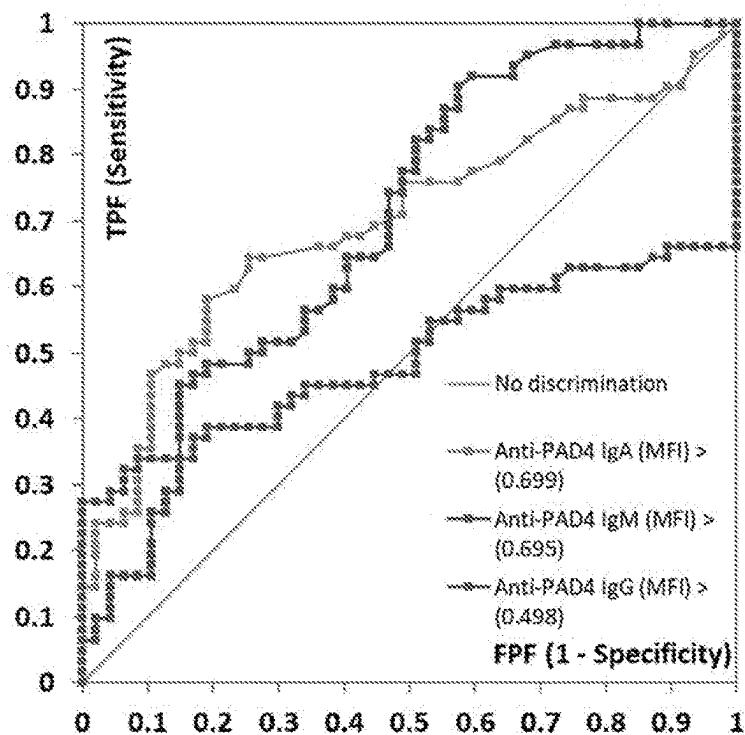


FIG. 23

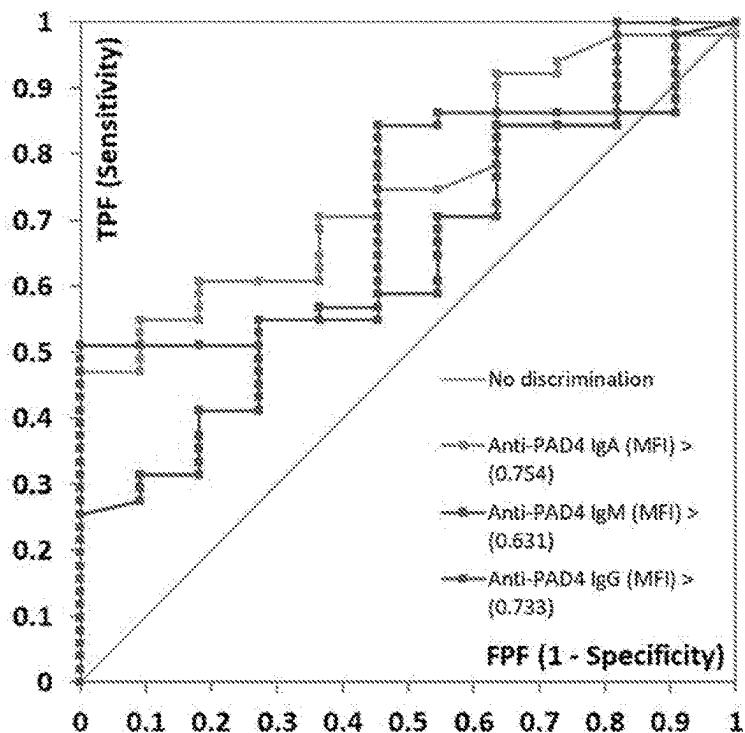


FIG. 24

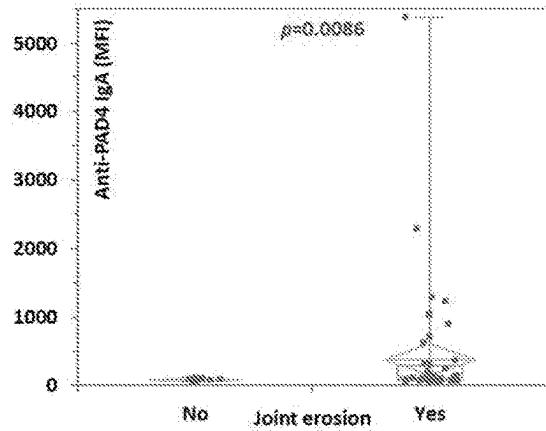


FIG. 25

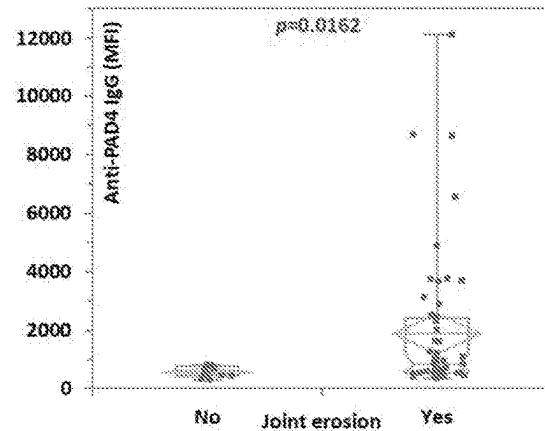


FIG. 26

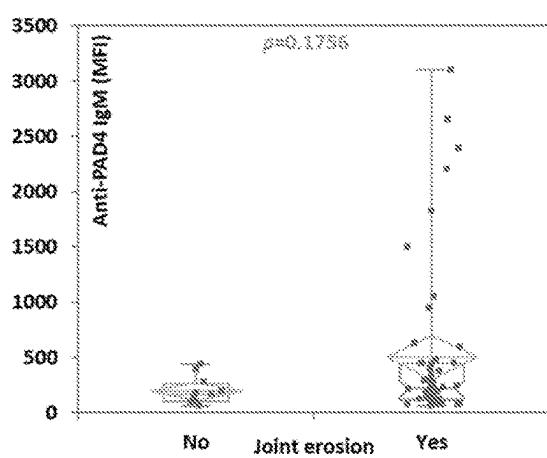


FIG. 27

COMPOSITIONS AND METHODS FOR DIAGNOSING AND ASSESSING RHEUMATOID ARTHRITIS

[0001] This application claims the benefit of U.S. Provisional Application No. 62/806,607 filed Feb. 15, 2019, the entire contents of which are incorporated herein by reference.

[0002] This application incorporates herein by reference a Sequence Listing as an ASCII text file entitled "13510-034-999_SEQ_LISTING" created on Feb. 12, 2020, and having a size of 385,711 bytes.

FIELD

[0003] The present disclosure relates to the field of molecular biology and more specifically to methods for detecting anti-PAD IgA in the serum of rheumatoid arthritis (RA) patients.

BACKGROUND

[0004] Rheumatoid Arthritis (RA) is a chronic autoimmune disease characterized by inflammation, pain and subsequent damage to synovial-lined joints. Unlike other arthritis conditions, RA is a systemic disease that can affect other organ systems including but not limited to the cardiovascular system, the respiratory system and musculature. While the exact pathogenesis of the disease is unknown, RA is characterized by the production of antibodies to self-proteins (autoantibodies) by the immune system. The most common autoantibodies implicated in RA include rheumatoid factor (RF) and anti-citrullinated protein antibodies (ACPAs), which are part of the classification criteria for this disease. ACPAs are a hallmark amongst serologic factors detected in RA patients, and as such, serve as valuable diagnostic and prognostic markers. Aletaha D. et al., Ann. Rheum. Dis. 2010, 69, 1580-1588; Taylor et al., Autoimmune Dis; 2011:815038 (2011). However, clinical heterogeneity of RA precludes the use of ACPAs and RF alone as reliable biomarkers. Patients with erosive disease require more aggressive treatment in the early phase of the disease to prevent joint damage. More precise biomarkers that specifically identify sufferers of RA and disease severity are needed.

[0005] Peptidylarginine deiminases (PAD)s are calcium-dependent enzymes that play a central role in generating autoantigens in RA through the conversion of arginine residues to citrulline, process known as citrullination. Beyond ACPA and RF, autoantibodies which target the PAD enzymes, have also been described in RA, see for example, Takizawa et al., Scand. J. Rheumatol. 3:212-215 (2005); Roth et al., Clin. Exp. Rheumatol. 1:12-18 (2006); Halvorsen et al., Ann. Rheumatol. Dis. 67:414-417 (2008); Zhao et al., J. Rheumatol., 35:969-974 (2008); Darrah et al., Sci. Trans. Med., 5 (186): 186ra65 (2013); Darrah et al., Front. Immunol., 9:2696 (2018). As such, PADs appear to play a central role in RA pathogenesis.

[0006] Thus, there exists a need for additional biomarkers for the diagnosis of RA and assessment of disease severity, including erosive conditions. The present disclosure satisfies this need and provides related advantages as well.

SUMMARY

[0007] In some embodiments, the present disclosure provides a method of diagnosing rheumatoid arthritis (RA). The

method includes: (a) contacting a biological sample from a subject suspected of having RA with a peptidyl arginine deiminase (PAD) or an antigenic fragment thereof, and (b) detecting the presence of an anti-PAD IgA in the biological sample, wherein the presence of the anti-PAD IgA is indicative of RA.

[0008] In some embodiments, the present disclosure provides a method of assessing disease severity in a subject having RA. The method includes: (a) contacting a biological sample from a subject having RA with a PAD or an antigenic fragment thereof, and (b) detecting the presence of an anti-PAD IgA in the biological sample, wherein the presence of the anti-PAD IgA is indicative of severity of RA.

[0009] In some embodiments, the biological sample includes whole blood, plasma, serum, synovial fluid or sputum. In some embodiments, the biological sample includes serum or plasma.

[0010] In some embodiments, the present disclosure provides a method of assessing disease severity, wherein the severity of RA includes the presence of joint erosion. In other embodiments, the severity of RA includes severe joint erosion. In other embodiments, anti-PAD IgA levels correlate with extent of joint erosion.

[0011] In some embodiments, the extent of joint erosion includes reduced mobility. In other embodiments, the reduced mobility includes a disability index of approximately 3.

[0012] In some embodiments, the PAD or antigenic fragment thereof used in the method of diagnosing RA or assessing disease severity is selected from the group consisting of PAD2, PAD3 and PAD4.

[0013] In some embodiments, the PAD or antigenic fragment thereof includes an amino acid sequence selected from an even numbered SEQ ID NO within SEQ ID NOS: 2-92 or an amino acid sequence comprising at least six consecutive amino acids selected from an even numbered SEQ ID NO within SEQ ID NOS: 2-92. In some embodiments, the antigenic fragment includes from 6-120, 12-100, 18-80, 24-60, 30-50 or 35-45 amino acid residues.

[0014] In some embodiments, the PAD or antigenic fragment thereof is obtained by a method comprising isolation from a natural source, chemical synthesis or recombinant expression.

[0015] In some embodiments, detection includes an immunoassay. In some embodiments, the immunoassay is selected from the group consisting of a fluorescent immunoassay (FIA), a chemiluminescent immunoassay (CIA), a radioimmunoassay (RIA), multiplex immunoassay, a protein/peptide array immunoassay, a solid phase radioimmunoassay (SPRIA), an indirect immunofluorescence assay (IIF), an enzyme linked immunosorbent assay (ELISA) and a particle based multianalyte test (PMAT), or a Dot Blot assay.

[0016] In some embodiments, the method described herein can be performed by (a) contacting the anti-PAD IgA with a detection probe specific for the anti-PAD IgA and (b) detecting specific binding of the detection probe. In some embodiments, the detection probe is specific to PAD. In other embodiments, the detection probe includes an antibody or functional fragment. In some embodiments, the functional fragment is anti-IgA. In some embodiments, the detection probe is a reporter tag.

[0017] In some embodiments, the reporter tag includes a label. In some embodiments, the label is selected from the

group consisting of a fluorophore, enzyme, chemiluminescent moiety, radioactive moiety, organic dye and small molecule. In some embodiments, the label is a fluorescent label. In some embodiments, the fluorescent label is phycoerytherin (PE).

[0018] In some embodiments, the reporter tag includes a ligand or particle. In some embodiments, the ligand includes biotin. In some embodiments, the particle includes a nanoparticle.

[0019] In other embodiments, the reporter tag is a ligand or particle. In some embodiments, the ligand is biotin and the particle is a nanoparticle.

[0020] In some embodiments, the present disclosure provides a kit. The kit includes: (a) a PAD, or antigenic fragment thereof; (b) a detection probe specific to anti-PAD IgA, and (c) a solid support.

[0021] In some embodiments, the kit further includes a label. In some embodiments, the kit includes a label selected from the group consisting of a fluorophore, enzyme, chemiluminescent moiety, radioactive moiety, organic dye and small molecule.

[0022] In some embodiments, the kit includes a positive control. In some embodiments, the positive control includes an anti-PAD IgA.

[0023] In some embodiments, the kit further includes one or more ancillary reagents. In some embodiments, the one or more ancillary reagents is selected from the group consisting of an incubation buffer, a wash buffer, a detection buffer and a detection instrument.

[0024] In some embodiments, the kit includes a PAD or antigenic fragment thereof selected from the group consisting of PAD2, PAD3 and PAD4.

[0025] In some embodiments, the PAD or antigenic fragment thereof in the kit includes an amino acid sequence selected from an even numbered SEQ ID NO within SEQ ID NOS: 2-92 or an amino acid sequence comprising at least six consecutive amino acids selected from an even numbered SEQ ID NO within SEQ ID NOS: 2-92.

[0026] In some embodiments, the kit contains an antigenic fragment including from 6-120, 12-100, 18-80, 24-60, 30-50 or 35-45 amino acid residues.

[0027] In some embodiments, the detection probe includes an antibody or functional fragment thereof. In some embodiments, the antibody or functional fragment thereof includes anti-IgA.

[0028] In some embodiments, the detection probe includes a reporter tag. In some embodiments, the reporter tag includes a label. In some embodiments, the label is selected from the group consisting of a fluorophore, enzyme, chemiluminescent moiety, radioactive moiety, organic dye and small molecule.

[0029] In some embodiments, the label is a fluorescent label. In some embodiments, the fluorescent label is (PE).

[0030] In some embodiments, the reporter tag includes a ligand or particle. In some embodiments, the ligand includes biotin. In some embodiments, the particle includes a nanoparticle.

[0031] In some embodiments, the solid support is selected from the group consisting of a bead, sphere, particle, membrane, chip, slide, plate, well and test tube. In some embodiments, the bead, sphere or particle includes micrometer or nanometer dimensions.

[0032] In some embodiments, the membrane is selected from the group consisting of nitrocellulose, nylon, polyvinylidene fluoride (PVDF) and polyvinylidene difluoride.

[0033] In some embodiments, the PAD or antigenic fragment thereof is conjugated to the solid support.

BRIEF DESCRIPTION OF THE DRAWINGS

[0034] FIG. 1 shows association of anti-PAD4 IgA with joint erosion status. Results are expressed in Median Fluorescence Intensity (MFI). P-value of the Mann-Whitney analysis is shown in red (p-value<0.05 considered significant). Median MFI for the subgroup and number of patients and % are shown. Red dashed line represents the preliminary cut-off.

[0035] FIG. 2 shows association of anti-PAD2 IgA with joint erosion status. Results are expressed in Median Fluorescence Intensity (MFI). P-value of the Mann-Whitney analysis is shown in red (p-value<0.05 considered significant). Median MFI for the subgroup and number of patients and % are shown. Red dashed line represents the preliminary cut-off.

[0036] FIG. 3 shows a receiver operating characteristic (ROC) analysis of anti-PAD2 IgA (square) and anti-PAD4 IgA (triangle) illustrating the discrimination between erosive and non-erosive disease in RA patients. Area Under the Curve (AUC) for each marker is shown in the legend.

[0037] FIG. 4 shows SEQ ID NO:1 which includes the mRNA nucleotide sequence of a human wild-type PAD2. The accession number for SEQ ID NO:1 is NM_007365.3.

[0038] FIG. 5 shows sequence number-2 which includes the amino acid sequence of a human wild-type PAD2. The accession number for SEQ ID NO:2 is NP_031391.2. SEQ ID NO:2 is the polypeptide encoded by SEQ ID NO:1.

[0039] FIG. 6 shows SEQ ID NO:3 which includes the mRNA nucleotide sequence of human PAD2 transcript variant X2. The accession number for SEQ ID NO:3 is XM_017000148.2. SEQ ID NO:3 is a transcript variant of SEQ ID NO:1.

[0040] FIG. 7 shows SEQ ID NO:4 which includes the amino acid sequence of a human PAD2 isoform X1. The accession number for SEQ ID NO:4 is XP_016855637.1. SEQ ID NO:4 is the polypeptide encoded by SEQ ID NO:3.

[0041] FIG. 8 shows SEQ ID NO:5 which includes the mRNA nucleotide sequence of a human wild-type PAD3. The accession number for SEQ ID NO:5 is NM_016233.2.

[0042] FIG. 9 shows SEQ ID NO:6 which includes the amino acid sequence of a human wild-type PAD3. The accession number for SEQ ID NO:6 is NP_057317.2. SEQ ID NO:6 is the polypeptide encoded by SEQ ID NO:5.

[0043] FIG. 10 shows SEQ ID NO:21 which includes the mRNA nucleotide sequence of a human PAD3 transcript variant X1. The accession number is XM_011541571.2. SEQ ID NO:21 is a transcript variant of SEQ ID NO:5.

[0044] FIG. 11 shows SEQ ID NO:22 which includes the amino acid sequence of a human PAD3 isoform X1. The accession number is XP_011539873.1. SEQ ID NO:22 is the polypeptide encoded by SEQ ID NO:21.

[0045] FIG. 12 shows SEQ ID NO:37 which includes the mRNA nucleotide sequence of a human PAD3 transcript variant X2. The accession number is XM_017001463.1. SEQ ID NO: 37 is a transcript variant of SEQ ID NO:5.

[0046] FIG. 13 shows SEQ ID NO:38 which includes the amino acid sequence of a human PAD3 isoform X2. The

accession number is XP_016856952.1. SEQ ID NO:38 is the polypeptide encoded by SEQ ID NO:37.

[0047] FIG. 14 shows SEQ ID NO:39 which includes the mRNA nucleotide sequence of a human PAD3 transcript variant X2. The accession number is XM_017001463.1: c. SEQ ID NO: 39 is a transcript variant of SEQ ID NO:5.

[0048] FIG. 15 shows SEQ ID NO:40 which includes the mRNA nucleotide sequence of a human PAD3 transcript variant X2. The accession number is XM_017001463.1: c. SEQ ID NO: 40 is a transcript variant of SEQ ID NO:5.

[0049] FIG. 16 shows SEQ ID NO:49 which includes the mRNA nucleotide sequence of a human PAD3 transcript variant X3. The accession number is XM_011541572.2. SEQ ID NO: 49 is a transcript variant of SEQ ID NO:5.

[0050] FIG. 17 shows SEQ ID NO:50 which includes the amino acid sequence of a human PAD3 isoform X3. The accession number is XP_011539874.1. SEQ ID NO:50 is the polypeptide encoded by SEQ ID NO:49.

[0051] FIG. 18 shows SEQ ID NO:59 which includes the mRNA nucleotide sequence of a human PAD3 transcript variant X3. The accession number is XM_011541572.2: c. SEQ ID NO: 59 is a transcript variant of SEQ ID NO:5.

[0052] FIG. 19 shows SEQ ID NO:61 which includes the mRNA nucleotide sequence of a human wild-type PAD4. The accession number is NM_012387.3.

[0053] FIG. 20 shows SEQ ID NO:62 which includes the amino acid sequence of a human wild-type PAD4. The accession number is NP_036519.2. SEQ ID NO:62 is the polypeptide encoded by SEQ ID NO:61.

[0054] FIG. 21 shows SEQ ID NO:67 which includes the mRNA nucleotide sequence of a human PAD4 transcript, variant X3. The accession number is XM_011541152.1. SEQ ID NO: 67 is a transcript variant of SEQ ID NO:61.

[0055] FIG. 22 shows SEQ ID NO:68 which includes the mRNA nucleotide sequence of a human PAD4 transcript variant X8. The accession number is XM_011541157.1. SEQ ID NO:68 is a transcript variant of SEQ ID NO:61.

[0056] FIG. 23 shows receiver operating characteristic (ROC) analysis of anti-PAD4 IgA (blue), anti-PAD4 IgG (grey) and anti-PAD4 IgM (red), illustrating the discrimination between RA and controls. Area Under the Curve (AUC) for each marker is shown in the legend.

[0057] FIG. 24 shows receiver operating characteristic (ROC) analysis of anti-PAD4 IgA (blue), anti-PAD4 IgG (grey) and anti-PAD4 IgM (red), illustrating discrimination for RA erosive disease. Area Under the Curve (AUC) for each marker is shown in the legend.

[0058] FIG. 25 shows association of anti-PAD4 IgA with joint erosion status. Results are expressed in Median Fluorescence Intensity (MFI). P-value of the Mann-Whitney analysis is shown in the graph (p-value <0.05 considered significant).

[0059] FIG. 26 shows association of anti-PAD4 IgG with joint erosion status. Results are expressed in Median Fluorescence Intensity (MFI). P-value of the Mann-Whitney analysis is shown in the graph (p-value <0.05 considered significant).

[0060] FIG. 27 shows association of anti-PAD4 IgM with joint erosion status. Results are expressed in Median Fluorescence Intensity (MFI). P-value of the Mann-Whitney analysis is shown in the graph (p-value <0.05 considered significant).

DETAILED DESCRIPTION

[0061] The present disclosure is based, in part, on the discovery that anti-PAD IgA serves as a diagnostic biomarker for RA and also as an indicator of disease severity for RA. Thus, the present disclosure benefits RA patients by providing a new biomarker which can indicate the presence of RA, disease severity, including erosive arthritis, and facilitate the early detection of RA and treatment escalation. Such benefits further enable at risk or early-stage RA patients to reduce or prevent disease progression and related erosive conditions such as joint erosion.

[0062] It must be noted that, as used in this specification and the appended claims, the singular forms "a", "an" and "the" also include plural referents unless the content clearly dictates otherwise.

[0063] It must also be noted that, as used in this specification and the appended claims, where a range of numeric values is provided, it is understood that the ranges are inclusive of the numbers defining the range. It is also understood that each intervening integer within the recited range as well as fractions thereof, including for example, every tenth of a unit of a selected intervening integer or a lower limit of the recited range is intended to be included within the disclosure, unless the context clearly dictates otherwise.

[0064] As used herein, the terms "comprises," "comprising," "includes," "including," "has," "having," "contains," "containing," and any variations thereof, are intended to cover a non-exclusive inclusion, such that a process, method, product-by-process, or composition of matter that includes has or contains an element or list of elements, does not include only those elements but can include other elements not expressly listed or inherent to such process, method, product-by-process, or composition of matter.

[0065] The present disclosure provides a method of diagnosing RA. The method includes: (a) contacting a biological sample from a subject suspected of having RA with a PAD or an antigenic fragment thereof, and (b) detecting the presence of anti-PAD IgA in the biological sample, wherein the presence of the anti-PAD IgA is indicative of RA.

[0066] The present disclosure also provides a method of assessing disease severity in a subject having RA. The method includes: (a) contacting a biological sample from a subject having RA with PAD or an antigenic fragment thereof, and (b) detecting the presence of anti-PAD IgA in the biological sample, wherein the presence of the anti-PAD IgA is indicative of severity of RA. Disease severity can be the presence of joint erosion, including assessing the extent of joint erosion.

[0067] The term "autoantibody" refers to an immunoglobulin directed against a constituent of tissue of the subject that produces the autoantibody. The term is intended to include an antibody produced by a subject's immune system that is directed against one or more of the subject's own polypeptides or antigens. Accordingly, autoantibodies can be produced by a subject's immune system when the immune system fails to distinguish, in whole or in part, between self and non-self tissue constituents. As provided herein an autoantibody directed to or specific for PAD having an IgA isotype is a beneficial biomarker for diagnosing RA, assessing disease severity of RA and/or diagnosing or determining the severity of joint erosion.

[0068] As used herein, the terms "anti-PAD IgA" when used in reference to an autoantibody is intended to mean an

IgA autoantibody directed against PAD or an antigenic fragment thereof. An IgA autoantibody is distinguishable from other antibody classes including, for example, gamma (IgG), mu (IgM), delta (IgD) and epsilon (IgE) antibody classes based in part on the constant region sequence and/or structure and other characteristics well-known in the art. IgA includes, for example, IgA1 and IgA2 subclasses as well as secretory IgA. IgA1 and IgA2 exist in monomer and dimer configurations and can form polymers with IgM. The term "anti-PAD IgA" is intended to include all IgA subclasses as well as the monomer, dimer and polymer configurations.

[0069] The presence of increased anti-PAD IgA in a subject compared to a healthy control individual can be indicative of the presence of RA, the severity of disease or the risk of developing RA. Accordingly, a measurable increase in an autoantibody to PAD having an IgA isotype, and any IgA subtypes, is used to diagnose RA, determine the severity of RA and/or diagnose or determine the severity levels of joint erosion. Exemplary methods for detection and comparison of anti-PAD IgA levels to a control are provided herein and described further below.

[0070] In some embodiments, detection of an increased level of anti-PAD IgA compared to a healthy control individual is indicative of a subject having RA. In some embodiments, following diagnosis of RA using the compositions and methods provided herein, the presence of RA can be further corroborated based on a variety of symptoms associated with the onset or presence of RA. Clinical symptoms associated with RA include, for example, pain and swelling of small and large bilateral joints, palindromic onset, mono-articular presentation, and extra-articular synovitis, like tenosynovitis and bursitis, polymyalgic-like onset and other symptoms including malaise, weight loss, fatigue, fever and disability. Grassi et al., Eur. J. Radiol., Suppl 1:S 18-24 (1998); Aletaha and Smolen, JAMA, 320(13):1360-1372 (2018).

[0071] In some embodiments, detection of an increased level of anti-PAD IgA in a subject compared to a healthy control is indicative of having severe RA. In other embodiments, detection of an increased level of anti-PAD IgA in a subject compared to an RA subject without an increased level of anti-PAD IgA, is indicative of having severe RA. In some embodiments, having severe RA is considered by the degree of joint erosion or the risk of radiographic progression as determined by methods in the art. Detection of an increased level of anti-PAD IgA in a subject compared to a healthy control or compared to an RA subject without an increased level of anti-PAD IgA is indicative that the subject has a higher probability of having more progressed RA wherein joint erosion is severe. In some embodiments, a subject having increased anti-PAD IgA can be more than 5%, more than 10%, more than 15%, more than 20%, more than 25%, more than 30%, more than 35%, more than 40%, more than 45%, more than 50%, more than 60%, more than 70%, more than 80% or more than 90% likely to have more progressed RA where severe joint erosion is present. In other embodiments, a subject having increased anti-PAD IgA can be more than 2-fold, more than 3-fold, more than 4-fold, more than 5-fold, more than 6-fold, more than 7-fold, more than 8-fold, more than 9-fold, or more than 10-fold likely to have more progressed RA where severe joint erosion is present.

[0072] In severe RA, joint erosion occurs when there is loss of bone and cartilage in the joint. Severity of joint

erosion can be determined by, for example, the Sharp score method. See Sharp, Arthritis Rheum., 32:221-229 (1989); Brower, Arthritis Rheum., 33:316-324 (1990). The Sharp score assesses joints for narrowness and erosions, based upon radiographic images. Erosion scores range from 0-3.5 and joint space narrowing scores range from 0-4. A score of 0 indicates a normal joint with no narrowing or erosions and a score of 3.5-4 indicates an abnormal joint with erosions and narrowing. In some embodiments of the present disclosure, joint erosion in a subject is determined by use of the Sharp score.

[0073] In other embodiments, having severe RA is determined by the Health Assessment Questionnaire (HAQ) Disability Index (DI). Fries et al., Arthritis Rheum., 23(2): 137-145 (1980); Bruce and Fries, Health Qual Life Outcomes, 1(1):20 (2003). The HAQ assesses physical ability in 8 sections including dressing, arising, eating, walking, hygiene, reach, grip and activities. Performing each session is allotted a score ranging from 0 (without any difficulty) to 3 (unable to do). The scores of the 8 sections are summed and divided by 8 to produce the DI. The DI, which ranges from 0 to 3, predicts disability, with a person able to complete a task without any difficulty (DI of 0), with some difficulty (DI of 1), with much difficulty (DI of 2), or unable to do (DI of 3).

[0074] Thus, in some embodiments, the presence of increased anti-PAD IgA in a subject compared to a healthy control individual is indicative of the presence of joint erosion. In other embodiments, detection of an increased level of anti-PAD IgA in a subject compared to an RA subject without an increased level of anti-PAD IgA is indicative of having joint erosion. In other embodiments, the presence of increased anti-PAD IgA compared to a healthy individual is indicative of the presence of severe joint erosion. In some embodiments, detection of an increased level of anti-PAD IgA compared to an RA subject without having an increased level of anti-PAD IgA is indicative of having severe joint erosion. In another embodiment, the presence of anti-PAD IgA is indicative of a DI of 2 or more, or of 3 or more.

[0075] In alternative embodiments, having joint erosion includes having worse radiographic joint damage. In some embodiments, having severe joint erosion includes having worse baseline radiographic joint damage. Accordingly, detection of an increased level of anti-PAD IgA in a subject compared to a healthy control individual is indicative of having worse radiographic joint damage and/or worse baseline radiographic joint damage. In other embodiments, detection of an increased level of anti-PAD IgA in a subject compared to a RA subject without having an increased level of anti-PAD IgA is indicative of having worse radiographic joint damage and/or worse baseline radiographic joint damage. One skilled in the art will recognize that methods for determining and assessing radiographic joint damage are well known in the art. Additionally, those skilled in the art will recognize and employ suitable techniques to quantify radiographic joint damage. As a non-limiting example, the Sharp scoring method can be used.

[0076] In some embodiments, detection of an increased level of anti-PAD IgA compared to a healthy control individual indicates that the subject is at risk of developing clinical symptoms of RA. In some embodiments, a subject can be at risk of developing clinical symptoms of RA within less than 3 months, less than 6 months, less than 9 months,

less than 12 months, less than 18 months, less than 2 years, less than 3 years, less than 4 years, less than 5 years, less than 6 years, less than 7 years, less than 8 years, less than 9 years, less than 10 years, less than 12 years, less than 14 years, or less than 16 years from the determination of the increased anti-PAD IgA level.

[0077] In some embodiments, the presence of an increased level of anti-PAD IgA compared to a healthy control individual indicates that the subject is more than 5%, more than 10%, more than 15%, more than 20%, more than 25%, more than 30%, more than 35%, more than 40%, more than 45%, more than 50%, more than 60%, more than 70%, or more than 80% or more than 90% likely to develop clinical symptoms of RA within 5 years following the determination of increased anti-PAD IgA. In some embodiments, the presence of an increased level of anti-PAD IgA can indicate that the subject is more than 2-fold, more than 3-fold, more than 4-fold, more than 5-fold, more than 6-fold, more than 7-fold, more than 8-fold, more than 9-fold, or more than 10-fold likely to develop clinical symptoms of RA within 5 years following determination of increased anti-PAD IgA level compared to a healthy control individual.

[0078] Anti-PAD IgA can be detected in a variety of different biological samples obtained from a subject. Such samples include, for example, solid tissue and biological fluids. As used herein, the term "biological sample" refers to any specimen from the body of an organism that can be used for analysis or diagnosis. In the context of the present disclosure, a biological sample obtained from a subject can be any sample that contains or is suspected to contain autoantibodies and encompasses any material in which an anti-PAD autoantibody can be detected. For example, a biological sample can include a liquid sample such as whole blood, plasma, serum, synovial fluid, amniotic fluid, sputum, pleural fluid, peritoneal fluid, central spinal fluid, urine, saliva, tears or other body fluid that contains autoantibodies. A biological sample can also include a solid tissue sample such as bone marrow, tissue, buccal or other solid or semi-solid aggregate of cells.

[0079] In some embodiments, anti-PAD IgA is detected in whole blood, plasma, serum, synovial fluid or sputum. In some embodiments of the present disclosure, the level of anti-PAD IgA is detected. In other embodiments, anti-PAD IgA-PAD complex can be formed using the compositions and methods described herein and an anti-PAD IgA in the complex can be detected. Accordingly, the disclosure provides compositions that include an anti-PAD IgA-PAD complex.

[0080] The biological samples of this disclosure can be obtained from any organism including, for example, mammals such as humans, primates such as monkeys, chimpanzees, orangutans and gorillas, cats, dogs, rabbits, farm animals such as cows, horses, goats, sheep and pigs, and rodents such as mice, rats, hamsters and guinea pigs.

[0081] In some embodiments, the biological sample can be a plurality of samples. In some embodiments the plurality of samples can be obtained periodically over the course of more than 12 hours, more than 1 day, more than 2 days, more than 3 days, more than 4 days, more than 5 days, more than 6 days, more than 7 days, more than 10 days, more than 14 days, more than 3 weeks, more than 1 month, more than 2 months, more than 3 months, more than 4 months, more than 5 months, more than 6 months, more than 9 months, more than 12 months, more than 18 months, more than 24 months,

more than 30 months, more than 3 years months, more than 4 years or more than 5 years.

[0082] In some embodiments, the samples of the present disclosure can be collected and processed fresh. In other embodiments, the samples of the present disclosure can be frozen, stored and processed at a later date.

[0083] In some embodiments, the present disclosure provides a method of determining the level of anti-PAD IgA in a subject to determine if that subject has RA, severe RA or joint erosion, including severe joint erosion. It is noted that, as used herein, the terms "subject," "organism," "individual" or "patient" are used as synonyms and interchangeably, and refer to a vertebrate mammal. Mammals include humans, primates such as monkeys, chimpanzees, orangutans and gorillas, cats, dogs, rabbits, farm animals such as cows, horses, goats, sheep and pigs, and rodents such as mice, rats, hamsters and guinea pigs. The subjects of this disclosure can include healthy subjects, asymptomatic subjects, and diseased subjects.

[0084] In some embodiments, the diseased subjects can suffer from any disease associated with aberrant anti-PAD IgA levels. It is noted that the term "aberrant anti-PAD IgA levels" refers to anti-PAD IgA levels in a sample that measurably deviate from the median anti-PAD IgA levels found in a population of healthy subjects. In some embodiments, the aberrant anti-PAD IgA levels can be higher than the median anti-PAD IgA levels. In some embodiments, the aberrant anti-PAD IgA levels can be lower than the median anti-PAD IgA levels.

[0085] In some embodiments, the healthy subjects can have never suffered from a certain disease. In some embodiments, the healthy subjects can be previously diseased. In some embodiments, the healthy subjects can be undergoing a routine medical checkup. In some embodiments, the healthy subjects can be members of a control group in, for example, a clinical trial. In some embodiments, the healthy subjects can be at risk of contracting a disease, as determined by the presence of certain risk factors that are well known in the art. Such risk factors include, without limitation, a genetic predisposition, a personal disease history, a familial disease history, a lifestyle factor, an environmental factor, a diagnostic indicator, and the like.

[0086] In some embodiments, the subject can be asymptomatic. Asymptomatic subjects include healthy subjects who have essentially no risk or only a low risk of developing RA (e.g., there is a less than 10%, less than 5%, less than 3%, or less than 1% probability that the asymptomatic patient will develop RA over the following five year period). Asymptomatic subjects further include healthy subjects who have a high risk of developing RA (e.g., there is a greater than 50%, greater than 70%, greater than 90%, or greater than 95% probability that the asymptomatic patient will develop RA over the following five year period). Asymptomatic subjects further include diseased subjects, who can display mild early diagnostic indicators of RA, but who are otherwise disease or complaint free (e.g., no synovial joint pain, no systemic inflammatory disorder). In some embodiments, the asymptomatic patient can be an arthralgia patient.

[0087] In some embodiments, the subject can have RA. In some embodiments, the subject can have RA with joint pain. In some embodiments, the subject can have RA with a systematic inflammatory disorder. In some embodiments, the subject can have juvenile idiopathic arthritis (JIA). In

some embodiments, the subject can have a pre-RA syndrome. In some embodiments, the pre-RA syndrome can be arthralgia.

[0088] In some embodiments, the subject can be suspected of having RA. As used herein, a subject can be “suspected of having RA” as determined by the presence of certain risk factors that are well known in the art. Such risk factors include, without limitation, a genetic predisposition, a personal disease history, a lifestyle factor, an environmental factor, a diagnostic indicator and the like.

[0089] In some embodiments, the subject can be at risk of developing RA. In some embodiments, the subject can have a genetic predisposition for developing RA or a family history of RA or other autoimmune diseases. In some embodiments, the subject can be exposed to certain lifestyle factors (e.g., smoking cigarettes) promoting the development of RA or the subject can show clinical disease manifestations of RA. In some embodiments, the subject can be a patient who is receiving a clinical workup to diagnose RA or to assess the risk of developing RA.

[0090] In some embodiments, the subjects can have anti-PAD IgA present, e.g., in their blood or another bodily tissue or fluid, (anti-PAD IgA-positive subjects). In some embodiments, the subjects can have elevated anti-PAD IgA levels, e.g., in their blood or another bodily tissue or fluid, relative to normal healthy subjects. In some embodiments, the subjects can have no anti-PAD IgA present, e.g., in their blood or another bodily tissue or fluid (anti-PAD IgA-negative subjects).

[0091] In some embodiments, the subjects can have anti-PAD IgA present, e.g., in their blood or another tissue or bodily fluid, (anti-PAD IgA-positive subjects) or the subjects can have elevated anti-PAD IgA levels, e.g., in their blood or another tissue or bodily fluid, relative to normal healthy subjects. In some embodiments, the subjects can be negative for anti-PAD IgA.

[0092] In some embodiments, the subject can be treatment naïve. In some embodiments, the subject can be undergoing treatments for RA (e.g., drug treatments). In some embodiments, the subject can be in remission. In some embodiments, the remission can be drug-induced. In some embodiments, the remission can be drug-free.

[0093] In some embodiments, the subject can be an animal model for RA. In some embodiments, the animal model can be a mouse, rabbit, or primate model of RA. In some embodiments, the animal model can involve inducing anti-PAD IgA responses by immunizing or vaccinating an animal with PAD.

[0094] It should be noted that the terms “healthy control individual,” “healthy subjects,” and grammatical equivalents herein are used interchangeably and refer to subjects who do not have increased anti-PAD IgA levels, RA or joint erosion above baseline or a standard known or determined to represent non-RA subjects.

[0095] The baseline or standard which determines or defines a subject as a non-RA subject is the reference interval. In diagnostic or health-related fields, the reference interval is a range of values observed in the reference subjects, which can be healthy control individuals, designated by specific percentiles. The reference interval can be any range of values as determined by those having skill in the art. See CLSI, “How to define and determine reference intervals in the clinical laboratory: approved guideline,” C28:A2 (2000). In some cases, the reference interval can be

stringent or less stringent depending on the specific analyte being measured or disease being studied. A person having skill in the art will understand the appropriate stringency to use when determining the reference interval. Thus, in some embodiments, the reference interval can be set at the 95th percentile. In order to increase specificity and decrease sensitivity, e.g. increase stringency, a higher cut-off can be used such as the 96th percentile or the 97th, or the 98th, or the 99th.

[0096] In the present disclosure, anti-PAD IgA can be considered increased in a subject if anti-PAD IgA levels are at least above the 95th percentile relative to anti-PAD IgA levels in healthy control subjects. In other embodiments, anti-PAD IgA can be considered increased in a subject if anti-PAD IgA levels are above the 96th, 97th, 98th or 99th percentile. A subject of the present disclosure with anti-PAD IgA levels at or above any of the disclosed reference intervals is considered to have RA and joint erosion

[0097] In some embodiments, the presence of anti-PAD IgA can be based on a comparison of signal against background in a healthy subject. In some embodiments, the presence of anti-PAD IgA can be increased or decreased relative to an average or median anti-PAD IgA level observed in a population of healthy subjects. In some embodiments, anti-PAD IgA can be absent in healthy subjects. In some embodiments, anti-PAD IgA level cannot be detected above the noise of the respective assay used to determine anti-PAD IgA level. In some embodiments, anti-PAD IgA can be considered present in a sample if an anti-PAD IgA level can be detected above the noise of the respective assay used to determine an anti-PAD IgA level. In some embodiments, anti-PAD IgA can be considered increased in a sample if the signal in an anti-PAD IgA detection assay is at least two standard deviations above noise such as the average or mean signal for control samples. In some embodiments, anti-PAD IgA can be considered present in a sample if the level of anti-PAD IgA exceeds a predetermined threshold level. An anti-PAD IgA threshold level can be determined by a skilled artisan, such as a clinical physician, based on a variety of factors, such as the specific objectives of a clinical trial or the diagnostic and prognostic significance of a certain anti-PAD IgA level or the results of another diagnostic test for RA that does not involve the detection of anti-PAD IgA levels.

[0098] In some embodiments, the present disclosure provides a polypeptide including a PAD or antigenic fragment thereof. The PAD can be used in the methods provided herein or included in the kits provided herein.

[0099] As used herein, the term “peptidyl arginine deiminase” or “PAD,” also known as PADI, refers to a family of enzymes that catalyze the post-translational modifications of protein arginine residues by deimination or demethylimation to produce citrulline (Wang and Wang, *Biochim. Biophys. Acta.*, 1829:1126-35 (2013)). Five isotypes of PADs have been identified in humans and include PAD1, PAD2, PAD3, PAD4 and PAD6. All of such PAD polypeptides, PAD1, PAD2, PAD3, PAD4 and PAD6 are included within the meaning of the term “PAD” as it is used herein.

[0100] As used herein, the term “peptidyl arginine deiminase 2” or “PAD2,” also known as PADI2, PAD-H19 and PDI2, refers to a member of the PAD family of enzymes. PAD2 is abundantly expressed in secretory glands, brain, uterus, spleen, pancreas and skeletal muscle. Known substrates of PAD2 include myelin basic protein, vimentin and

macrophages. See Vossenaar et al., Annals of the Rheumatic Diseases, 63:373-81 (2004); Watanabe et al., Biochim Biophys Acta., 966:375-383 (1988); Watanabe et al., J. Biol Chem., 264:15255-15260 (1989); Nagata et al., Experientia, 46:72-74 (1990); Urano et al., Am J Dermatopathol., 12 (3): 249-55 (1990), Vossenaar et al., Arthritis and Rheum., 48:2489-2500 (2003). Approximately 726 coding single nucleotide polymorphisms (SNP) have been identified for PAD2 and at least 184 known orthologs (see, for example, NCBI Gene ID: 11240). The term "PAD2" includes all of such PAD2 variants and PAD2 orthologs. An exemplary human PAD2 (hPAD2) nucleotide sequence can be found in GenBank under GenBank Accession number NM_007365 (SEQ ID NO: 1) and encodes an exemplary human PAD2 having the amino acid sequence found under found under GenBank Accession number NP_031391 (SEQ ID NO:2). The GenBank Accession numbers and GenBank GI numbers of this PAD2 and other exemplary PAD2 enzymes can be found below in Table 1. All of such PAD2 polypeptides and variants thereof are included within the meaning of the term "PAD2" as it is used herein.

[0101] In some embodiments, a PAD2, or antigenic fragment thereof, includes the amino acid in SEQ ID NO:2, of a mature human PAD2 (hPAD2; amino acids 25-664 of NCBI Accession Number NP_031391; GI: 122939159), or naturally occurring variants thereof:

```
SEQ ID NO: 2
MLRERTVRLQYGSRVEAVYVLGTYLWTDVYSAAPAGAQTFSLKHSEHVV
VEVVRDGEAEEVATNGKQRWLSPSTTLRVTMSQASTEASSDKVTVNYY
DEEGSIPIDQAGLFLTAIEISLDVADRDGVVEKNNPKKASWTWGPEGQ
GAILLVNCDRTPWLPKEDCRDEKVKYSKEDLKDMSSQMLRTKGPDRLLA
GYEIVLYIISMSDSKVGVFVYVENPFFCQRYIHILGRRKLYHVVKYTGGS
AELLFFVEGLCPDPDEGFSGLVSIHVSLEYMAQDIPLTPIFTDTVIFRI
APWIMTPNILPPSVFVCCMKDNYLFLKEVKNLVEKTNCELKVCFQYLN
RGDRWIQDEIEFGYIEAPHKGFPVVLSPRDGNLKDFPVKEELLGPDFGY
VTREPLFESVTSLDSFGNLEVSPPVTVNGKTYPLGRILIGSSFPLSGGR
RMTKVVDRFLKAQQVQAPVELYSDWLTVGHVDEFMSFVPIPGTKKFLLL
MASTSACYKLFRBKQKDGHGEAIMFKGLGGMSSKRITINKILSNESLVQ
ENLYFQRCLWDNRDILKKELGLTEQDIIIDLPAFLKMDEDHRARAFFPNM
VNMIIVLDKDLGIKPKPGPQVEEECLEMHVRGLLEPLGLECTFIDDISA
YHKFLGEVHCGTNVRRKPTFKWWHMV
```

[0102] As used herein, the term "peptidyl arginine deiminase 3" or "PAD3," also known as PADI3, PDI3 and UHS1, refers to a member of the PAD family of enzymes. PAD3 is detected in the epidermis where it plays a role in terminal differentiation and in hair follicles where it modulates structural proteins including filaggrin and trichoyalin. See Chavanas et al., J Dermatol Sci., 44:63-72 (2006); Kanno et al., J. Invest Dermatol. 115(5):813-23 (2000); Nachat et al., J Investig Dermatol., 125:34-41 (2005). Approximately 738 coding SNPs have been identified for PAD3 and at least 109 known orthologs (see, for example, NCBI Gene ID: 51702). The term "PAD3" includes all of such PAD3 variants and PAD3 orthologs. An exemplary human PAD3 (hPAD3) (

nucleotide sequence can be found in GenBank under GenBank Accession number NM_016233 (SEQ ID NO:5) and encodes an exemplary human PAD3 having the amino acid sequence found under found under GenBank Accession number NP_057317(SEQ ID NO:6). The GenBank Accession numbers and GenBank GI numbers of this PAD3 and other exemplary PAD3_enzymes can be found below in Table 2. All of such PAD3 polypeptides and variants thereof are included within the meaning of the term "PAD3" as it is used herein.

[0103] In some embodiments, a PAD3, or antigenic fragment thereof, includes the amino acid in SEQ ID NO:6 of a mature human PAD3 (hPAD3; amino acids 25-664 of NCBI Accession Number NP_057317; GI: 122939161), or naturally occurring variants thereof:

```
SEQ ID NO: 6
MSLQRIVRVSLEHPTSAVCVAGVETLVDIYGSPPEGTEMFEVYGTGPVGD
IYISPNNMERGRERADTRWRFDATLEIIVVMNSPSNLDNSHVGQISYHS
SHEPLPLAYAVLYLTCVDISLDCDLNCEGRQDRNFVDKRQWVWGPSSGYG
GILLVNCDRDDPSCDVQDNCDQHVHCLQDLEDMSVMVLRTQGPAALFDD
HKLVLHTSSYDAKRAQVFHICGPEDVCEAYRHVLGQDKVSYEVPRLHGD
EEFFVEGLSFPDAGFTGLISPHVTLDDSNEDFSASPIFTDTVVPRVA
PWIMTPSTLPPLEVYVCRVRNNTCFVDAVAELARKAGCKLTICPQAENR
NDRWIQDEMELGYVQAPHKTLPPVFDSPRNGELQDFPYKRILGPDPFGYV
TREPRDRRSVGLDSFGNLEVSPPVANGKEYPLGRILIGGNLPGSSGRR
VTQVVRDPLHAQKVQPPVLFVDLAVGHVDEFLSFVPAPDGKGFRMLL
ASPGACFKLFQEKKQKCGHGRALLFQGVVDDEQVKTI SINQVLSNKDLIN
YNKVFQSCIDWNREVLKRELGLAECIDIIDIPQLFKTERKKATAFFPDLV
NMLVLGKHLGIPKPFGPPIINGCCCLEEKVRSLLPGLHCTFIDDPTPY
HMLHGEVHCGTNVCRKPFPSFKWWNMV
```

[0104] As used herein, the term "peptidyl arginine deiminase 4" or "PAD4," also known as PAD, PADI4, PDI4, PADV, PDI5 and PADI5, refers to a member of the PAD family of enzymes. PAD4 can be detected in white blood cells including granulocytes and monocytes under normal physiological conditions (Vossenaar et al., Annals of the Rheumatic Diseases, 63:373-81 (2004); Asaga et al., J. Leukocyte Biology 70:46-51 (2001)) and is generally localized in the nucleus (Nakashima et al., JBC 277:49562-68 (2002)). Approximately 737 coding SNPs have been identified for PAD4 and at least 108 known orthologs (see, for example, NCBI Gene ID: 23569). The term "PAD4" includes all of such PAD4 variants and PAD4 orthologs. An exemplary human PAD4 (hPAD4) nucleotide sequence can be found in GenBank under GenBank Accession number NM_012387.2 (SEQ ID NO:61) and encodes an exemplary human PAD4 having the amino acid sequence found under found under GenBank Accession number NP_036519.2 (SEQ ID NO:62). The GenBank Accession numbers and GenBank GI numbers of this PAD4 and other exemplary PAD4 enzymes can be found below in Table 3. All of such PAD4 polypeptides and variants thereof are included within the meaning of the term "PAD4" as it is used herein.

[0105] In some embodiments, a PAD4, or antigenic fragment thereof, includes the amino acid in SEQ ID NO:62, of a mature human PAD4 (hPAD4; amino acids 25-663 of NCBI Accession Number NP_036519; GI: 216548487), or naturally occurring variants thereof:

```
SEQ ID NO: 62
MAQGTLIRVTPEQP THAVCVLGLTQL DICSSA PECTSFSINASPGVV
VDIAHGPPAKKSTGSSTWPLDPGVETLTMKVASGSTGDQKVQI SYYG
PKTPPVKALLYLTGVEISLCADI TRTGKVKPTRAVKDQRTWTGPGCGQG
A1LLVNCDRDNLESSAMDCE DEDEVLDSEDLQDMSMLTSLTKTPKDFFTN
HTLVLHVARSEM DKV RVFQATRGKLSSKCSVVLGPWKPSHYLMVPGGKH
NMDFYVEALAFFPDTDFPGLITLTISLLDTSNLELP EAVVFQDSVVPRVA
PWIMTPNTQPPQEVYAC SIFENEDFLKS VTTLAMKAKC KLTICPEEEENM
DDQWMQDEMEIG YIQAP HKTL PVV FDS PRN RGLKE FPI KRV MGP DFGYV
TRGPQTGGISGLDSFGN LEVSP PVTVRGKEYPLGRILFGDSCYP SNSDR
QMHQALQDFLSAQVQAPV KLYSDWL SVGHVDEFLSFVPAPDRKGPFRL
LASPRSCYKLFQEQQN EGHGEALLFEGIKKKQQKIKN ILSNKTLREHN
SFVERCIDWNREL KREL GLA ESDIIDIPQLFKLKEFSKA EAFPNMVN
MLVLGKHLGIPKPF GPVINGRC CLEEKVC SLEPLGLQCTFIND FFTYH
I RHGEVHCGTNVRRK PPSFKNW NMVP
```

[0106] Tables 1, 2 and 3 contain two sequence identifiers, the GI number and the GenBank accession number. The GI number and GenBank accession number run in parallel as unique identifiers to access the referenced sequence in publicly available databases. Table 1 includes GI numbers and GenBank Accession numbers for PAD2, Table 2 includes GI numbers and GenBank Accession numbers for PAD3 and Table 3 includes GI numbers and GenBank accession numbers for PAD4.

[0107] The sequence identifiers in Tables 1, 2 and 3 are provided for wild-type PAD2, 3 and 4, respectively. It should be understood that wild-type nucleic acid and amino acid sequences herein refer to those nucleic acid and amino acid sequences prevalent among a population and serve as a reference for their respective variants. As an example, SEQ ID NO:61 in Table 3 (GI number: 1519314340 and Accession number: NM_012387) identifies the wild-type nucleic acid sequence for hPAD4 while SEQ ID NO:62 (GI number: 216548487 and Accession number: NP_036519) identifies the wild-type amino acid sequence for hPAD4.

[0108] The sequence identifiers in Tables 1, 2 and 3 also are provided for variants of PAD2, 3 and 4 respectively. It

should be understood that a variant refers to a nucleic acid sequence that is similar but different from the wild-type nucleic acid sequence.

[0109] A variant in any of the Tables below can include a nucleic acid substitution that can be the result of, for example, alternative splicing (e.g. splice variant). As an example, SEQ ID NO: 69 in Table 3 (GI number: 767903519 and Accession number: XM_011541150.1: c.23G>A) is a hPAD4 nucleic acid splice variant of SEQ ID NO:61.

[0110] A variant in any of the Tables below can also include, for example, a nucleic acid substitution (e.g. SNP). As an example, SEQ ID NO:63 in Table 3 (GI number: 216548486 and Accession number: NM_012387.2:c.23G>A) is a hPAD4 nucleic acid variant of SEQ ID NO:61 and includes a single nucleic acid substitution at nucleic acid position 23, resulting in the substitution of G (guanosine) for A (adenine).

[0111] It should be understood that a variant also refers to an amino acid sequence that is similar but different to the wild-type amino acid sequence.

[0112] A variant in any of the Tables below can further include amino acid substitutions that can be the result of, for example, alternative splicing (e.g. splice variant). As an example, SEQ ID NO: 70 in Table 3 (GI number: 767903520 and Accession number: XP_011539452.1: p.Arg8His), which is encoded by SEQ ID NO:69 described above, is a hPAD4 that includes an amino acid substitution at position 8, resulting in a substitution of Arg (arginine) for His (histidine).

[0113] A variant in any of the Tables below can include, for example, amino acid substitutions that can be the result of genetic inheritance (e.g. SNP). As an example, SEQ ID NO: 64 in Table 3 (GI number: 216548487 and Accession number: NP_036519.2: p.Arg8His), which is encoded by SEQ ID NO:63 described above, is a hPAD4 that includes an amino acid substitution at position 8, resulting in a substitution of Arg (arginine) for His (histidine).

TABLE 1

Molecule Type	SEQ ID NO	GI Number	GenBank Accession Number
<i>Homo sapiens</i> peptidyl arginine deiminase 2 (PADI2), mRNA	1	1519245591	NM_007365
protein-arginine deiminase type-2 [<i>Homo sapiens</i>]	2	122939159	NP_031391
PREDICTED: <i>Homo sapiens</i> peptidyl arginine deiminase 2 (PADI2), transcript variant X2, mRNA	3	1370451734	XM_017000148
protein-arginine deiminase type-2 isoform X1 [<i>Homo sapiens</i>]	4	1034554998	XP_016855637

TABLE 2

Molecule Type	SEQ ID NO	GI Number	GenBank Accession Number	Amino Acid [Codon]	SO Term
<i>Homo sapiens</i> peptidyl arginine deiminase 3 (PADI3), mRNA	5	122939160	NM_016233	N/A	N/A
protein-arginine deiminase type-3 [<i>Homo sapiens</i>]	6	122939161	NP_057317	N/A	N/A
PADI3 transcript	7	122939160	NM_016233.2: c.154A > G	I [ATC] > V [GTC]	Coding Sequence Variant

TABLE 2-continued

Molecule Type	SEQ ID NO	GI Number	GenBank Accession Number	Amino Acid [Codon]	SO Term
protein-arginine deiminase type-3	8	122939161	NP_057317.2: p.Ile52Val	I (Ile) > V (Val)	Missense Variant
PADI3 transcript	9	122939160	NM_016233.2: c.335T > A	L [CTC] > H [CAC]	Coding Sequence Variant
protein-arginine deiminase type-3	10	122939161	NP_057317.2: p.Leu112His	L (Leu) > H (His)	Missense Variant
PADI3 transcript	11	122939160	NM_016233.2: c.511G > A	V [GTG] > M [ATG]	Coding Sequence Variant
protein-arginine deiminase type-3	12	122939161	NP_057317.2: p.Val171Met	V (Val) > M (Met)	Missense Variant
PADI3 transcript	13	122939160	NM_016233.2: c.881C > T	A [GCA] > V [GTA]	Coding Sequence Variant
protein-arginine deiminase type-3	14	122939161	NP_057317.2: p.Ala294Val	A (Ala) > V (Val)	Missense Variant
PADI3 transcript	15	122939160	NM_016233.2: c.1744G > A	A [GCC] > T [ACC]	Coding Sequence Variant
protein-arginine deiminase type-3	16	122939161	NP_057317.2: p.Ala582Thr	A (Ala) > T (Thr)	Missense Variant
PADI3 transcript	17	122939160	NM_016233.2: c.1813C > A	P [CCC] > T [ACC]	Coding Sequence Variant
protein-arginine deiminase type-3	18	122939161	NP_057317.2: p.Pro605Thr	P (Pro) > T (Thr)	Missense Variant
PADI3 transcript	19	122939160	NM_016233.2: c.1853G > A	R [CGG] > Q [CAG]	Coding Sequence Variant
protein-arginine deiminase type-3	20	122939161	NP_057317.2: p.Arg618Gln	R (Arg) > Q (Gln)	Missense Variant
Predicted: <i>Homo sapiens</i> <td>21</td> <td>1034559140</td> <td>XM_011541571</td> <td>N/A</td> <td>N/A</td>	21	1034559140	XM_011541571	N/A	N/A
protein-arginine deiminase type-3 isoform X1 [<i>Homo sapiens</i>]	22	767904616	XP_011539873	N/A	N/A
Predicted: PADI3 transcript variant X1	23	1034559140	XM_011541571.2: c.40A > G	I [ATC] > V [GTC] I (Ile) > V (Val)	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X1	24	767904616	XP_011539873.1: p.Ile14Val	L [CTC] > H [CAC]	Missense Variant
Predicted: PADI3 transcript variant X1	25	1034559140	XM_011541571.2: c.221T > A	L (Leu) > H (His)	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X1	26	767904616	XP_011539873.1: p.Leu74His	V (Val) > M (Met)	Missense Variant
Predicted: PADI3 transcript variant X1	27	1034559140	XM_011541571.2: c.397G > A	V [GTG] > M [ATG]	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X1	28	767904616	XP_011539873.1: p.Val133Met	A [GCA] > V [GTA]	Missense Variant
PADI3 transcript variant X1	29	1034559140	XM_011541571.2: c.767C > T	A (Ala) > V (Val)	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X1	30	767904616	XP_011539873.1: p.Ala256Val	A [GCC] > T [ACC]	Missense Variant
PADI3 transcript variant X1	31	1034559140	XM_011541571.2: c.1630G > A	A (Ala) > V (Val)	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X1	32	767904616	XP_011539873.1: p.Ala544Thr	A [GCC] > T [ACC]	Missense Variant
PADI3 transcript variant X1	33	1034559140	XM_011541571.2: c.1699C > A	P [CCC] > T [ACC]	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X1	34	767904616	XP_011539873.1: p.Pro567Thr	P (Pro) > T (Thr)	Missense Variant
PADI3 transcript variant X1	35	1034559140	XM_011541571.2: c.1739G > A	R [CGG] > Q [CAG]	Coding Sequence Variant

TABLE 2-continued

Molecule Type	SEQ ID NO	GI Number	GenBank Accession Number	Amino Acid [Codon]	SO Term
protein-arginine deiminase type-3 isoform X1 <i>Homo sapiens</i> peptidyl arginine deiminase 3 (PAD13), transcript variant X2, mRNA	36	767904616	XP_011539873.1: p.Arg580Gln	R (Arg) > Q (Gln)	Missense Variant
	37	1034559141	XM_017001463	N/A	N/A
protein-arginine deiminase type-3 isoform X2 [<i>Homo sapiens</i>]	38	1034559142	XP_016856952	N/A	N/A
PAD13 transcript variant X2	39	1034559141	XM_017001463.1: c	N/A	Genic Upstream Transcript Variant
PAD13 transcript variant X2	40	1034559141	XM_017001463.1: c	N/A	5 Prime UTR Variant
PAD13 transcript variant X2	41	1034559141	XM_017001463.1: c.344C > T	A [GCA] > V [GTA]	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X2	42	1034559142	XP_016856952.1: p.Ala115Val	A (Ala) > V (Val)	Missense Variant
PAD13 transcript variant X2	43	1034559141	XM_017001463.1: c.1207G > A	A [GCC] > T [ACC]	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X2	44	1034559142	XP_016856952.1: p.Ala403Thr	A (Ala) > T (Thr)	Missense Variant
PAD13 transcript variant X2	45	1034559141	XM_017001463.1: c.1276C > A	P [CCC] > T [ACC]	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X2	46	1034559142	XP_016856952.1: p.Pro426Thr	P (Pro) > T (Thr)	Missense Variant
PAD13 transcript variant X2	47	1034559141	XM_017001463.1: c.1316G > A	R [CGG] > Q [CAG]	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X2	48	1034559142	XP_016856952.1: p.Arg439Gln	R (Arg) > Q (Gln)	Missense Variant
<i>Homo sapiens</i> peptidyl arginine deiminase 3 (PAD13), transcript variant X3, mRNA	49	1034559143	XM_011541572	N/A	N/A
protein-arginine deiminase type-3 isoform X3 [<i>Homo sapiens</i>]	50	767904618	XP_011539874	N/A	N/A
PAD13 transcript variant X3	51	1034559143	XM_011541572.2: c.154A > G	I [ATC] > V [GTC]	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X3	52	767904618	XP_011539874.1: p.Ile52Val	I (Ile) > V (Val)	Missense Variant
PAD13 transcript variant X3	53	1034559143	XM_011541572.2: c.335T > A	L [CTC] > H [CAC]	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X3	54	767904618	XP_011539874.1: p.Leu112His	L (Leu) > H (His)	Missense Variant
PAD13 transcript variant X3	55	1034559143	XM_011541572.2: c.511G > A	V [GTG] > M [ATG]	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X3	56	767904618	XP_011539874.1: p.Val171Met	V (Val) > M (Met)	Missense Variant
PAD13 transcript variant X3	57	1034559143	XM_011541572.2: c.881C > T	A [GCA] > V [GTA]	Coding Sequence Variant
protein-arginine deiminase type-3 isoform X3	58	767904618	XP_011539874.1: p.Ala294Val	A (Ala) > V (Val)	Missense Variant
PAD13 transcript variant X3	59	1034559143	XM_011541572.2: c.	N/A	Genic Downstream Transcript Variant

TABLE 3

Molecule Type	SEQ ID NO	GI Number	GenBank Accession Number	Amino Acid [Codon]	SO Term
<i>Homo sapiens</i> peptidyl arginine deiminase 4 (PADI4), mRNA	61	1519314340	NM_012387	N/A	N/A
protein-arginine deiminase type-4 [<i>Homo sapiens</i>] PADI4 transcript	62	216548487	NP_036519	N/A	N/A
protein-arginine deiminase type-4 PADI4 transcript	63	216548486	NM_012387.2: c.23G > A	R [CGT] > H [CAT]	Coding Sequence Variant
protein-arginine deiminase type-4 PADI4 transcript	64	216548487	NP_036519.2: p.Arg8His	R (Arg) > H (His)	Missense Variant
protein-arginine deiminase type-4 PADI4 transcript variant X3	65	216548486	NM_012387.2: c.23G > T	R [CGT] > L [CTT]	Coding Sequence Variant
protein-arginine deiminase type-4 PADI4 transcript variant X3	66	216548487	NP_036519.2: p.Arg8Leu	R (Arg) > L (Leu)	Missense Variant
PADI4 transcript variant X8	67	767903523	XM_011541152.1: c.	N/A	Genic Upstream Transcript Variant
PADI4 transcript variant X1	68	767903533	XM_011541157.1: c.	N/A	Genic Upstream Transcript Variant
protein-arginine deiminase type-4 isoform X1 PADI4 transcript variant X1	69	767903519	XM_011541150.1: c.23G > A	R [CGT] > H [CAT]	Coding Sequence Variant
protein-arginine deiminase type-4 isoform X1 PADI4 transcript variant X1	70	767903520	XP_011539452.1: p.Arg8His	R (Arg) > H (His)	Missense Variant
protein-arginine deiminase type-4 isoform X1 PADI4 transcript variant X2	71	767903519	XM_011541150.1: c.23G > T	R [CGT] > L [CTT]	Coding Sequence Variant
protein-arginine deiminase type-4 isoform X1 PADI4 transcript variant X2	72	767903520	XP_011539452.1: p.Arg8Leu	R (Arg) > L (Leu)	Missense Variant
protein-arginine deiminase type-4 isoform X2 PADI4 transcript variant X2	73	767903521	XM_011541151.1: c.23G > A	R [CGT] > H [CAT]	Coding Sequence Variant
protein-arginine deiminase type-4 isoform X2 PADI4 transcript variant X2	74	767903522	XP_011539453.1: p.Arg8His	R (Arg) > H (His)	Missense Variant
protein-arginine deiminase type-4 isoform X2 PADI4 transcript variant X2	75	767903521	XM_011541151.1: c.23G > T	R [CGT] > L [CTT]	Coding Sequence Variant
protein-arginine deiminase type-4 isoform X2 PADI4 transcript variant X4	76	767903522	XP_011539453.1: p.Arg8Leu	R (Arg) > L (Leu)	Missense Variant
protein-arginine deiminase type-4 isoform X4 PADI4 transcript variant X4	77	767903525	XM_011541153.1: c.23G > A	R [CGT] > H [CAT]	Coding Sequence Variant
protein-arginine deiminase type-4 isoform X4 PADI4 transcript variant X4	78	767903526	XP_011539455.1: p.Arg8His	R (Arg) > H (His)	Missense Variant
protein-arginine deiminase type-4 isoform X4 PADI4 transcript variant X6	79	767903525	XM_011541153.1: c.23G > T	R [CGT] > L [CTT]	Coding Sequence Variant
protein-arginine deiminase type-4 isoform X4 PADI4 transcript variant X6	80	767903526	XP_011539455.1: p.Arg8Leu	R (Arg) > L (Leu)	Missense Variant
protein-arginine deiminase type-4 isoform X5 PADI4 transcript variant X6	81	767903529	XM_011541155.1: c.23G > A	R [CGT] > H [CAT]	Coding Sequence Variant
protein-arginine deiminase type-4 isoform X5 PADI4 transcript variant X6	82	767903530	XP_011539457.1: p.Arg8His	R (Arg) > H (His)	Missense Variant
protein-arginine deiminase type-4 isoform X5 PADI4 transcript variant X7	83	767903529	XM_011541155.1: c.23G > T	R [CGT] > L [CTT]	Coding Sequence Variant
protein-arginine deiminase type-4 isoform X5 PADI4 transcript variant X7	84	767903530	XP_011539457.1: p.Arg8Leu	R (Arg) > L (Leu)	Missense Variant
protein-arginine deiminase type-4 isoform X6 PADI4 transcript variant X7	85	767903531	XM_011541156.1: c.23G > A	R [CGT] > H [CAT]	Coding Sequence Variant
protein-arginine deiminase type-4 isoform X6 PADI4 transcript variant X7	86	767903532	XP_011539458.1: p.Arg8His	R (Arg) > H (His)	Missense Variant
protein-arginine deiminase type-4 isoform X6 PADI4 transcript variant X7	87	767903531	XM_011541156.1: c.23G > T	R [CGT] > L [CTT]	Coding Sequence Variant
protein-arginine deiminase type-4 isoform X6 PADI4 transcript variant X5	88	767903532	XP_011539458.1: p.Arg8Leu	R (Arg) > L (Leu)	Missense Variant
protein-arginine deiminase type-4 isoform X4 PADI4 transcript variant X5	89	1034557308	NM_011541154.2: c.23G > A	R [CGT] > H [CAT]	Coding Sequence Variant
protein-arginine deiminase type-4 isoform X4 PADI4 transcript variant X5	90	767903528	XP_011539456.1: p.Arg8His	R (Arg) > H (His)	Missense Variant
protein-arginine deiminase type-4 isoform X4 PADI4 transcript variant X5	91	1034557308	NM_011541154.2: c.23G > T	R [CGT] > L [CTT]	Coding Sequence Variant
protein-arginine deiminase type-4 isoform X4 PADI4 transcript variant X4	92	767903528	XP_011539456.1: p.Arg8Leu	R (Arg) > L (Leu)	Missense Variant

[0114] It should be noted that “polypeptide” includes a short oligopeptide having between 2 and 30 amino acids (e.g., 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25 or 30 amino acids) as well as longer amino acid chains, e.g., more than 30 amino acids, more than 50 amino acids, more than

100 amino acids, more than 150 amino acids, more than 200 amino acids, more than 300 amino acids, more than 400 amino acids, more than 500 amino acids, or more than 600 amino acids. In some embodiments, the PAD can be a full-length, wild-type polypeptide. PAD polypeptides can

include unnatural amino acids not encoded by the natural genetic code.

[0115] In some embodiments, the purified polypeptide includes a PAD antigenic fragment. In some embodiments, a PAD antigenic fragment includes more than 3, more than 5, more than 10, more than 15, more than 20, more than 25, more than 50, more than 75, more than 100, more than 125, more than 150, more than 200, more than 250, more than 300, more than 350, more than 400, more than 500, or more than 600 consecutive amino acids of a full-length PAD polypeptide. In some embodiments, a PAD antigenic fragment includes less than 100%, less than 95%, less than 90%, less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, less than 50%, less than 45%, less than 40%, less than 35%, less than 30%, less than 25%, less than 20%, less than 15%, less than 10%, or less than 5% of consecutive amino acids of full-length PAD. In some embodiments, a PAD antigenic fragment is a PAD peptide fragment.

[0116] In some embodiments, a PAD or antigenic fragment thereof can be a mammalian PAD. In some embodiments, a PAD or antigenic fragment thereof can be human. In some embodiments, a PAD or antigenic fragment thereof can be a PAD or antigenic fragment thereof of one of the organisms of the present disclosure. In some embodiments, a PAD or antigenic fragment thereof can include any of the variants or single nucleotide polymorphisms in Tables 1-3.

[0117] A PAD or antigenic fragment thereof can be obtained using various methods well known in the art. For example, a PAD or antigenic fragment thereof can be isolated from a natural source, produced by chemical synthesis or produced by recombinant protein expression.

[0118] Exemplary methods for expressing and purifying recombinant polypeptides, for purifying polypeptides from cells, tissues or bodily fluids, and for chemically synthesizing polypeptides are well known in the art and can be found described in Scopes R. K., Protein Purification—Principles and Practice, Springer Advanced Texts in Chemistry, 3rd Edition (1994); Simpson R. J. et al., Basic Methods in Protein Purification and Analysis: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 1st Edition (2008); Green M. R. and Sambrook J., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 4th Edition (2012); Jensen K. J. et al., Peptide Synthesis and Applications (Methods in Molecular Biology), Humana Press, 2nd Edition (2013).

[0119] Polypeptides purified or isolated from a natural source refers to the isolation and purification of a polypeptide from a source where it is naturally expressed. In some embodiments, a natural source of a PAD can be from a cell, tissue or bodily fluid of an organism. In some embodiments, the cells, tissues or bodily fluids can include, for example, whole blood, serum, plasma, synovial fluid or sputum from an organism of the present disclosure. A PAD or antigenic fragment thereof can similarly be isolated from any biological sample described and provided herein.

[0120] It should be noted that the terms “purified” or “isolated” refer to a polypeptide that is isolated, partially or completely, from a complex mixture of components, as found in nature. Thus, in some embodiments, a PAD of the present disclosure can be partially purified or substantially purified. Partial purification results in isolation from one or more components as found in nature. Substantial purifica-

tion results in isolation from all components as found in nature. Partial purification, as disclosed herein, can be achieved by the methods and compositions provided herein. In some embodiments, a partially purified PAD can be performed with a capture probe. In some embodiments, the capture probe is a polypeptide or functional fragment thereof specific to PAD. In some embodiments, the capture probe is an anti-PAD antibody. Substantial purification, as exemplified herein, can be achieved by methods known in the art. In some embodiments, a PAD is purified substantially by a process of extraction, precipitation and solubilization.

[0121] Recombinant polypeptides can be expressed in and purified from bacterial cells (e.g., *E. coli*), yeast cells (e.g., *S. cerevisiae*), insect cells (e.g., SF9), in mammalian cells (e.g., CHO) and others. Recombinant polypeptides can be expressed and purified as fusion proteins including tags for protein detection or affinity purification tags (e.g., His-tag, GST-tag, Myc-tag), including cleavable tags (e.g., tags including a TEV-cleavage site). In some embodiments, the PAD provided herein can be purified from a cell, tissue or bodily fluid obtained from an organism. Tissues or bodily fluids can include any tissue or bodily fluids obtained from the organism. In some embodiments, the tissues or bodily fluids can include blood, serum, plasma, synovial fluid, urine or milk (e.g., from goats, cows, sheep). One skilled in the art will recognize that methods for the purification of polypeptides from cells, tissues or bodily fluids are well known in the art.

[0122] In some embodiments, a PAD or antigenic fragment thereof is chemically synthesized using, for example, methods described in Jensen, K. J. (supra).

[0123] In some embodiments, a PAD antigenic fragment can be produced by enzymatically digesting full-length PAD. The full-length PAD can be obtained by, for example, any of the exemplary methods described above. The enzymatic digest can be carried out with, for example, a protease or peptidase. In some embodiments, the protease or peptidase can be an exoprotease or an exopeptidase. In some embodiments, the protease or peptidase can be an endoprotease or endopeptidase. In some embodiments, the protease or peptidase can include a serine protease, threonine protease, cysteine protease, aspartate protease, glutamic acid protease, or metalloprotease. In some embodiments, the protease or peptidase can include trypsin, chymotrypsin, pepsin, papain and any cathepsin including cathepsin B, L, D, K, or G.

[0124] In some embodiments, a PAD or antigenic fragment thereof can be a native PAD. In some embodiments, the PAD or antigenic fragment thereof can be a denatured or unfolded PAD. In some embodiments, the PAD or antigenic fragment thereof can include unnatural amino acids. In some embodiments, the unnatural amino acids can be methylated at the-amino-group to produce polypeptides with methylated backbones. In some embodiments, the unnatural amino acids can be R-amino acids. In some embodiments, the unnatural amino acids can include dyes (e.g., fluorescent dyes) or affinity tags. In some embodiments, the PAD or antigenic fragment thereof can include chemical modifications. Chemical modifications can include, e.g., chemical modifications with biotin, fluorescent dyes. A skilled artisan will recognize that methods for introducing unnatural amino acids into polypeptides and for chemically modifying polypeptides are well known in the art.

[0125] In some embodiments, an isolated, chemically synthesized or recombinant PAD or antigenic fragment thereof can be a plurality of PADs. It should be noted that the term “plurality” refers to a population of two or more members, such as polypeptide members or other referenced molecules. In some embodiments, the two or more members of a plurality of members can be the same members. For example, a plurality of polypeptides can include two or more polypeptide members having the same amino acid sequence. By way of exemplification, a plurality of members having the same amino acid sequence can include two or more members of any one of PAD exemplified in Table 1. In some embodiments, the two or more members of a plurality of members can be different members. For example, a plurality of polypeptides can include two or more polypeptide members having different amino acid sequences. By way of exemplification, a plurality of members having different amino acid sequences can include at least one member of two or more PADs exemplified in Table 1. A plurality includes 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90 or a 100 or more different members. A plurality can also include 200, 300, 400, 500, 1000, 5000, 10000, 50000, 1×10^5 , 2×10^5 , 3×10^5 , 4×10^5 , 5×10^5 , 6×10^5 , 7×10^5 , 8×10^5 , 9×10^5 , 1×10^6 , 2×10^6 , 3×10^6 , 4×10^6 , 5×10^6 , 6×10^6 , 7×10^6 , 8×10^6 , 9×10^6 or 1×10^7 or more different members. A plurality includes, for example, all integer numbers in between the above exemplary plurality numbers. In some embodiments, a PAD can be a plurality of PADs from the organisms of the present disclosure.

[0126] As provided herein, RA, RA severity and joint erosion can be determined in subjects of the present disclosure by the detection of anti-PAD2, anti-PAD3 or anti-PAD4 IgA. Detection of any of the anti-PAD IgA described herein can be performed through the use of, for example, an antibody specific to IgA or other binding molecule specific to IgA. An IgA binding molecule in the art can be used. An antibody or other binding molecule specific to any of the anti-PAD IgA described herein can also be employed.

[0127] As used herein, the term “antibody” is used interchangeably with immunoglobulin (Ig) and refers to a polypeptide product of B-cells that is able to bind to a specific molecular antigen and is composed of two heavy chains and two light chains. Each amino-terminal portion of each chain includes a variable region that confers binding specificity. See Borrebaeck (ed.) (1995) Antibody Engineering, Second Edition, Oxford University Press.; Kuby (1997) Immunology, Third Edition, W.H. Freeman and Company, New York. The term includes autoantibodies and antibodies used as detection probes in the disclosed methods. The antibody can exhibit specific binding affinity where it binds to a single molecular species or pan-specific binding where it binds selectively to more than one related molecular species. In the context of the present disclosure, the specific molecular antigen that can be bound by an antibody of the disclosure includes, for example, IgA, PAD (e.g., PAD2, PAD3 and/or PAD4), PAD: anti-PAD IgA complex (e.g., anti-PAD 2, 3 and/or 4 IgA complexes), and/or anti-PAD IgA (e.g., anti-PAD2 IgA, anti-PAD3 IgA and/or anti-PAD4 IgA). An antibody of the present disclosure can be derived from any mammalian organism, including mouse, rabbit, goat, chicken, donkey and the like. Furthermore, a primary or secondary antibody can be monoclonal, polyclonal, chimeric or humanized. The antibodies provided herein can be used in the methods and compositions of the disclosure.

[0128] As used herein, the term “functional fragment” when used in reference to an antibody is intended to refer to a portion of the antibody including heavy or light chain polypeptides that retain some or all of the binding activity as the antibody from which the fragment was derived. Such functional fragments can include, for example, an Fd, Fv, Fab, F(ab'), F(ab)2, F(ab')2, single chain Fv (scFv), diabody, triabody, tetrabody and minibody. Such antibody binding fragments can be found described in, for example, Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, New York (1989); Myers (ed.), Molec. Biology and Biotechnology: A Comprehensive Desk Reference, New York: VCH Publisher, Inc.; Huston et al., Cell Biophysics, 22:189-224 (1993); Plückthun and Skerra, Meth. Enzymol., 178:497-515 (1989) and in Day, E.D., Advanced Immunochemistry, Second Ed., Wiley-Liss, Inc., New York, NY (1990). The antibody functional fragments provided herein can be used in the methods and compositions of the disclosure. Ligands are provided herein and include any molecule having specific binding to a target. Exemplary ligands include a polypeptide, IgA binding molecules including, for example, IgA binding proteins, an affibody, an aptamer, a small molecule and the like. Specific examples of polypeptide ligands include receptors, chimeric receptors and polypeptides identified from screening of random or combinatorial libraries. Exemplary polypeptide ligands of the present disclosure include PAD (e.g., PAD2, PAD3 and/or PAD4) or an antigenic fragment thereof or an IgA binding polypeptide. An exemplary IgA binding polypeptide includes KW-0388. Other exemplary ligands that bind to IgA can be found described in Rönnmark et al., “Human immunoglobulin A (IgA)-specific ligands from combinatorial engineering of protein A,” Eur. J. Biochem. 269:2647-55 (2002) and Kruljec et al., “Alternative Affinity Ligands for Immunoglobulins,” Bioconjugate Chem. 28:2009-30 (2017). A ligand of the present disclosure can be obtained or synthesized by methods described herein or known in the art, including for example, chemically synthesized, purified from a natural source or recombinantly made. Thus, a ligand detection probe described herein can be mammalian, including mouse, rabbit, goat, chicken, donkey and the like. All of such ligands provided herein can be used in the methods and compositions of the disclosure.

[0129] As used herein, the term “detection probe” refers to a binding agent capable of specific binding to a target. Such binding agents include, for example, antibodies and ligands. Antibodies include full length antibodies as well as functional fragments such as those exemplified above. Ligands include full length polypeptides such as those exemplified above and functional binding fragments thereof. Ligands also include the non-polypeptide ligands exemplified above. When referring to specific binding to a target, a detection probe of the disclosure can bind the target directly or it can be made specific to the target by indirect means. For example, a detection probe that binds directly to anti-PAD IgA includes PAD. A direct binder also includes, for example, an antibody or other ligand that specifically recognizes a PAD: anti-PAD IgA complex as well as an antibody or other ligand that specifically binds to anti-PAD IgA. A detection probe of the disclosure that can be made specific to anti-PAD IgA by indirect means includes, for example, anti-IgA or other ligand that binds IgA. Such antibodies and ligands can be made specific to anti-PAD IgA by, for example, capturing the anti-PAD IgA with PAD and

washing away non-anti-PAD IgA prior to adding anti-IgA. Numerous other configurations for isolating such a binding complex in order to achieve specific binding to a target are well known in the art and all of which can be used as an indirect means to make a detection probe specific to a target. Thus, a “detection probe specific for anti-PAD IgA” includes, for example, PAD, a PAD: anti-PAD IgA complex binding agent, an anti-PAD IgA binding agent and an IgA binding agent. The anti-PAD IgA detection probes include binding agents to anti-PAD2 IgA, anti-PAD3 IgA and/or anti-PAD4 IgA.

[0130] Accordingly, in one embodiment an exemplary detection probe of the current disclosure which can bind anti-PAD IgA directly is a labeled PAD. A detection probe made specific to anti-PAD IgA by indirect means includes a labeled anti-IgA. These and other exemplary detection probes as well as means for capturing a PAD: anti-PAD IgA complex for specific detection are further described below.

[0131] As used herein, the term “reporter tag” refers to a molecule capable of producing a signal indicative of the detection of a biomarker. An exemplary biomarker in the present disclosure includes anti-PAD IgA. Reporter tags can be attached, for example conjugated, to the detection probe through non-covalent or covalent cross-linkage to the detection probe. Non-covalent and covalent immobilization of reporter tags to detection probes can be performed by any means known in the art described in Dennler et al., “Antibody conjugates: from heterogeneous populations to defined reagents,” *Antibodies*. 4:197-224 (2015). Reporter tags produce various signals, depending on the type of reporter tag. A person skilled in the art appreciates that there are various labels encompassed by reporter tags.

[0132] As used herein, the term “label” refers to a molecular entity that emits a signal and can be used as a readout or measurement for detection of an analyte. Various classes of labels exist. Such labels include a fluorophore, an enzyme, a chemiluminescent moiety, a radioactive moiety, an organic dye, a small molecule, a polypeptide or functional fragment thereof. Examples of fluorophores include fluorescent dyes like phycoerytherin (PE), fluorescein isothiocyanate (FITC), tetramethylrhodamine (TRITC), BODIPY and AlexaFluor® dyes. Fluorescent dyes can also include fluorescence resonance energy transfer (FRET)-dyes or time-resolved (TR)-FRET dyes. Fluorophore labels also include fluorescent proteins such as green fluorescent protein (GFP) and cyan fluorescent protein (CFP). Examples of enzyme labels include alkaline phosphatase (AP) or horseradish peroxidase (HRP). When any of the substrates 3,3'5,5'-Tetramethylbenzidine (TMB), 3,3'-Diaminobenzidene (DAB), or 2,2'-azino-bis (3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) are applied to HRP, a colored (chromogenic) or light (chemiluminescent) signal is produced. Radioactive moiety labels include carbon-14 or Tritium. Small molecule labels include biotin, resins such as agarose beads and fluorescently labeled magnetic beads, or nanoparticles such as colloidal gold. Polypeptide or functional fragment labels include Avidin, Streptavidin or NeutrAvidin which have an affinity for biotin. Polypeptide or functional fragment labels also include hemagglutinin (HA), glutathione-S-transferase (GST) or c-myc.

[0133] A label of the present disclosure can be conjugated to any of the detection probes identified herein. Conjugation can include non-covalent or covalent cross-linkage as described above. In some configurations, a label conjugated

to a detection probe requires an additional substrate or binding agent described above. As an example, an HRP label conjugated to a detection probe requires a substrate, disclosed above, to detect a detection probe. Numerous other configurations for a label are known in the art. The present disclosure includes all label configurations exemplified herein and/or known in the art. In some embodiments, a label configuration can include PE conjugated to a PAD, a PAD: anti-PAD IgA complex binding agent, an anti-PAD IgA or an anti-IgA. In alternative embodiments, a label configuration can include PE conjugated to a specific PAD including, for example, PAD2, PAD3 and/or PAD4. In further embodiments, a label configuration can include a PE conjugated to an anti-PAD IgA including, for example, anti-PAD2 IgA, anti-PAD3 IgA and/or anti-PAD4 IgA.

[0134] Methods for detecting, measuring and/or quantifying a signal produced by a label of the present disclosure are well known in the art and include detection of fluorescence, luminescence, chemiluminescence or absorbance, reflectance, transmittance, birefringence or refractive index. Optical methods include imaging methods such as confocal and non-confocal microscopy and non-imaging methods such as microplate readers. In some embodiments, methods of detecting anti-PAD IgA in biological sample can include visualization, quantification or both of a fluorescent, colorimetric or absorbance signal in a biological sample.

[0135] In some embodiments of the present disclosure, anti-PAD IgA presence can be detected by immunoassay. Methods and protocols for conducting immunoassays and biophysical protein-interaction assays are well known in the art. See, e.g., Wild D., *The Immunoassay Handbook*, Elsevier Science, 4th Edition (2013); Fu H., *Protein-Protein Interactions*, Humana Press, 4th Edition (2004). Exemplary immunoassays include fluorescent immunosorbent assay (FIA), a chemiluminescent immunoassay (CIA), a radioimmunoassay (RIA), multiplex immunoassay, a protein/peptide array immunoassay, a solid phase radioimmunoassay (SPRIA), an indirect immunofluorescence assay (IIF), an enzyme linked immunosorbent assay (ELISA) and a particle based multianalyte test (PMAT), or a Dot Blot assay.

[0136] In some embodiments, the ELISA can be a sandwich ELISA. In some embodiments, the sandwich ELISA can include the initial step of immobilizing a purified polypeptide of this disclosure on a solid support as exemplified below. For example, a PAD or antigenic fragment thereof can be immobilized on a wall of a microtiter plate well or of a cuvette. In some embodiments, contacting the sample from the subject with the PAD or antigenic fragment thereof of this disclosure can include exposing the sample to the immobilized PAD or antigenic fragment thereof.

[0137] In some embodiments, the ELISA can be a direct ELISA. In some embodiments, the direct ELISA can include the initial step of immobilizing a PAD or antigenic fragment thereof on any of the solid supports disclosed herein. For example, a PAD or antigenic fragment thereof can be immobilized to a wall of a microtiter plate well or of a cuvette. In some embodiments, contacting the sample from the subject with a PAD or antigenic fragment thereof of this disclosure can include exposing the anti-PAD IgA contained in the patient's sample to the immobilized PAD. Any of the immunoassays disclosed herein (see above) and in the art can be used, or modified to be used, in any of the methods disclosed herein.

[0138] In some embodiments, anti-PAD IgA can be detected by a particle based multianalyte test. As used herein, the term “particle based multianalyte test (PMAT)” refers to an immunoassay that allows simultaneous measurement of two or more analytes in a single assay. For example, in PMAT, different types of particles are used simultaneously, with each type having immobilized a specific binding partner for a specific molecule species on the surface of its particles. In a solution, the analyte molecules to be detected are bound to their binding partners on the corresponding particle type. The bonds are then detected optically through the addition of a secondary marker that marks all particle-bound analyte molecules of the multiplex assay. A PMAT can be performed using a variety of formats known in the art, such as flow cytometry, a capture sandwich immunoassay, or a competitive immunoassay. For example, using a dual-laser flow-based detection instrument, the binding of analyte fractions, such as autoantibodies, can be detected through the fluorescence of the secondary marker. In some embodiments, the PMAT particle can be a bead. In effecting the PMAT, the presence of one or more autoantibodies specifically associated with an autoimmune disease can be identified, and the patient can be diagnosed with the autoimmune disease that is specifically associated with the autoantibody identified by the PMAT.

[0139] In some embodiments, a Dot-Blot or line immunoassay (LIA) can be used to detect anti-PAD IgA in a biological sample. Methods and protocols for dot blot are well known in the art, including estimating polypeptide concentration. See Joint ProteomicS Laboratory (JPSL) of the Ludwig Institute for Cancer Research, Estimating protein concentration by dot blotting of multiple samples, Cold Spring Harbor Protocols, New York (2006).

[0140] In some embodiments, the immunoassay can be performed by immobilizing a capture probe to a solid support for a sufficient time to allow binding to occur. A capture probe includes a binding agent that binds to an analyte of interest. With respect to detection of an anti-PAD IgA of this disclosure, a capture probe can be any binding agent that specifically binds to anti-PAD IgA, PAD: anti-PAD IgA complex or anti-PAD IgA. Exemplary capture probes includes, PAD and/or a particular PAD such as PAD2, PAD3 and/or PAD4, as well as antigenic fragments thereof. Other exemplary capture probes include anti-IgA antibodies and functional fragments thereof, anti-IgA binding polypeptides and functional fragments thereof, anti-PAD IgA binding polypeptides, including antibodies, and functional fragments thereof and/or PAD: anti-PAD IgA complex binding polypeptides and functional fragments and binding agents.

[0141] Accordingly, in some embodiments, an immunoassay can utilize anti-IgA immobilized to a solid support to capture IgA. In some embodiments, a PAD or antigenic fragment thereof can be immobilized to a solid support to capture anti-PAD IgA. In other embodiments, an anti-PAD IgA binding polypeptide can be immobilized to a solid support to detect IgA. Anti-PAD IgA captured by the above exemplary capture probes can be detected using a detection probe provided in this disclosure.

[0142] The immunoassay can further include blocking steps, washing steps and additionally or alternatively, elution steps. Blocking steps can include contacting a solid support of the immunoassay in a blocking buffer for a sufficient time and temperature to allow blocking. Exemplary blocking

buffers are identified below as are exemplary solid supports. Washing steps include contacting a solid support of the immunoassay with a washing buffer to remove non-specific binding of polypeptides to the solid support. Exemplary washing buffers are described below. Elution buffers can include any of a variety of elution buffers known in the art or disclosed herein. Elution buffers include, for example, a 0.1 M glycine: HCl solution between pH 2.5 and 3. Polypeptide complexes can be eluted from the solid support of the immunoassay to aid in detection and measurement of, for example, PAD and anti-IgA complexes.

[0143] The present disclosure provides a kit which can be used to diagnosis RA, severity of disease and joint erosion. The kit can include a PAD of the present disclosure as exemplified in Tables 1-3 or an antigenic fragment thereof. In some embodiments, a PAD or antigenic fragment thereof can include any mammalian PAD as provided herein. Exemplary PADS include, for example, PAD2, PAD3 and PAD4.

[0144] The kit can include any of the detection probes provided herein as well as others well known in the art. For example, a detection probe can include an antibody or a ligand. A detection probe can be immobilized on a solid support. It should be noted that the term “immobilized” is used interchangeably with “attached” and both terms are intended to include covalent and non-covalent attachment, unless indicated otherwise, either explicitly or by context. In some embodiments, a PAD or antigenic fragment thereof is immobilized to a solid support. h

[0145] As provided herein and exemplified with respect to the methods of this disclosure, a kit of this disclosure can include a reporter tag. Reporter tags function to produce a signal for detection of a biomarker. Reporter tags can be attached, for example, to any of the detection probes used herein through non-covalent or covalent cross-linkage. As exemplified with respect to the methods of this disclosure, a kit can include any of the labels described or exemplified herein. For example, a label of the kit can include a fluorophore, an enzyme, a chemiluminescent moiety, a radioactive moiety, an organic dye, a small molecule, a polypeptide or functional fragment thereof. In some embodiments, a label of the kit includes PE. In some embodiments, a label of the kit includes FITC. In some embodiments, a label of the present disclosure is conjugated to a detection probe of the disclosure as exemplified above.

[0146] A kit can include any solid support provided herein or identified in the art. As used herein, the terms “solid support,” “solid surface” and other grammatical equivalents refer to any material that is appropriate for or can be modified to be appropriate for the attachment of PAD or an antigenic fragment thereof of this disclosure. Possible materials include, without limitation, glass and modified or functionalized glass, plastics (including acrylics, polystyrene, methylstyrene, polyurethanes, TeflonTM, etc.), paramagnetic materials, thoria sol, carbon graphite, titanium oxide, latex or cross-linked dextrans such as Sepharose, cellulose polysaccharides, nylon or nitrocellulose, ceramics, resins, silica or silica-based materials including silicon and modified silicon, carbon metals, inorganic glasses, optical fiber bundles, and a variety of other polymers. In some embodiments, the solid supports can be located in microtiter well plates (e.g., a 96-well, 384-well or 1536-well plate). In some embodiments, the solid supports can be located within a flow cell or flow cell apparatus (e.g., a flow cell on a BiacoreTM chip or a protein chip).

[0147] In some embodiments, the solid support can be a bead, microsphere, particle, membrane, chip, slide, well, and test tube. Beads include microspheres or particles. By “microspheres” or “particles” or grammatical equivalents herein is meant small, discrete, non-planar particles in the micrometer or nanometer dimensions. In some embodiments the bead can be spherical, in other embodiments the bead is irregular. Alternatively or additionally, the beads can be porous. The bead sizes range from nanometers to millimeters with beads from about 0.2 to about 200 microns being preferred in some embodiments. In other embodiments, bead size can range from about 0.5 to about 5 microns. In some embodiments, beads smaller than 0.2 microns and larger than 200 microns can be used. In some embodiments, the solid support can include an array of wells or depressions in a surface. This can be fabricated as is known in the art using a variety of techniques, including, photolithography, stamping techniques, molding techniques and microetching techniques. As will be appreciated by those skilled in the art, the technique used will depend on the composition and shape of the array substrate.

[0148] In some embodiments, the solid support can include a patterned surface suitable for immobilization of purified proteins in an ordered pattern (e.g., a protein chip). A “patterned surface” refers to an arrangement of different regions in or on an exposed layer of a solid support. For example, one or more of the regions can be features where one or more purified proteins are present. The features can be separated by interstitial regions where purified proteins are not present. In some embodiments, the pattern can be an x-y format of features that are in rows and columns. In some embodiments, the pattern can be a repeating arrangement of features and/or interstitial regions. In some embodiments, the pattern can be a random arrangement of features and/or interstitial regions. Exemplary patterned surfaces that can be used in the methods and compositions set forth herein are described in U.S. Pat. App. Publ. No. 2008/0280785 A1, U.S. Pat. App. Publ. No. 2004/0253640 A1, U.S. Pat. App. Publ. No. 2003/0153013 A1 and International Publication No. WO 2009/039170 A2.

[0149] In some embodiments, a solid support can have attached to its surface a PAD or an antigenic fragment thereof or anti-IgA. Any PAD exemplified by, for example, Tables 1-3, including antigenic fragments thereof can be attached to a solid support. In some embodiments, any PAD or antigenic fragment thereof of the present disclosure can be immobilized to a solid support via a linker molecule. In some embodiments, all that is required is that molecules, such as any PAD or antigenic fragment thereof of the present disclosure, remain immobilized or attached to the support under the conditions in which it is intended to use the support, for example, in applications requiring antibody binding or detection.

[0150] A kit can include a positive control. In some embodiments, a positive control can be a sample containing a detectable amount of anti-PAD IgA or levels above the threshold. In some embodiments, a positive control can be obtained from a diseased subject who has levels of anti-PAD IgA above threshold. Additionally or alternatively, a positive control can contain anti-PAD IgA synthesized in vitro using any of the methods described herein. In other embodiments, the kit can include a negative control. A negative control can be a sample containing no detectable amount of anti-PAD IgA or levels below the threshold. In some embodiments, a

negative control can be obtained from a healthy control individual or can be synthesized in vitro. For example, a negative control can include water or buffer.

[0151] The kit or the disclosure can further include one or more ancillary reagents. As used herein, “ancillary reagents” refer to a substance, mixture, material or component that is useful to carry out an intended purpose of the kit. Ancillary reagents can include a reagent, including a conjugation reagent, a buffer, standard, positive control, label, instructions and the like.

[0152] In some embodiments, a reagent of the kit of the present disclosure can include any conjugation reagent known in the art, including covalent and non-covalent conjugation reagents. Covalent conjugation reagents can include any chemical or biological reagent that can be used to covalently immobilize a polypeptide of this disclosure on a surface. Covalent conjugation reagents can include a carboxyl-to-amine reactive group such as carbodiimides such as EDC or DCC, an amine reactive group such as N-hydroxysuccinimide (NHS) ester or imidoesters, a sulfhydryl-reactive crosslinker such as maleimides, haloacetyls, or pyridyl disulfides, a carbonyl-reactive crosslinker groups such as, hydrazides or alkoxyamines, a photoreactive crosslinker such as aryl azides or diziines, or a chemoselective ligation group such as a Staudinger reaction pair. Non-covalent immobilization reagents can include any chemical or biological reagent that can be used to immobilize a polypeptide of this disclosure non-covalently on a surface, such as affinity tags such as biotin or capture reagents such as streptavidin or anti-tag antibodies, such as anti-His6 or anti-Myc antibodies.

[0153] The kits of this disclosure can include combinations of conjugation reagents. Such combinations include, e.g., EDC and NHS, which can be used, e.g., to immobilize a protein of this disclosure on a surface, such as a carboxylated dextrane matrix (e.g., on a BIACoreTM CM5 chip or a dextrane-based bead). Combinations of conjugation reagents can be stored as premixed reagent combinations or with one or more conjugation reagents of the combination being stored separately from other conjugation reagents.

[0154] In other embodiments, a reagent of the kit can include a reagent such as a coating buffer. A coating buffer can include sodium carbonate-sodium hydroxide or phosphate. In some embodiments, the coating buffer can be 0.1M NaHCO₃ (e.g., about pH 9.6).

[0155] In some embodiments, a reagent of a kit can include a washing buffer. A washing buffer can include tris (hydroxymethyl) aminomethane (Tris)-based buffers like Tris-buffered saline (TBS) or phosphate buffers like phosphate-buffered saline (PBS). Washing buffers can be composed of detergents, such as ionic or non-ionic detergents. In some embodiments, the washing buffer can be a PBS buffer at about pH 7.4 including Tween[®]20 at about 0.05%. In other embodiments, the washing buffer can be the BIO-FLASH[™] Special Wash Solution (Inova Diagnostics, Inc., San Diego, CA).

[0156] In some embodiments, a reagent of the kit can include a dilution buffer. Any dilution buffer known in the art can be included in the kit of the present disclosure. Typical dilution buffers include a carrier protein such as bovine serum albumin (BSA) and a detergent such as Tween[®]20. In some embodiments, the dilution buffer can be PBS at about pH 7.4 including BSA at about 1% BSA and Tween[®]20 at about 0.05%.

[0157] In some embodiments, a reagent can include a detection or assay buffer. Any detection or assay buffer known in the art can be included in the kit of the present disclosure. The detection or assay buffer can be a colorimetric detection or assay buffer, a fluorescent detection or assay buffer or a chemiluminescent detection or assay buffer. Colorimetric detection or assay buffers include PNPP (p-nitrophenyl phosphate), ABTS (2,2'-azino-bis (3-ethylbenzothiazoline-6-sulphonic acid)) or OPD (o-phenylenediamine). Fluorescent detection or assay buffers include QuantaBluTM or QuantaRedTM (Thermo Scientific, Waltham, MA). Chemiluminescent detection or assay buffers can include luminol or luciferin. Detection or assay buffers can also include a trigger such as H₂O₂ and a tracer such as isoluminol-conjugate. In some embodiments, the detection reagent can include one or more BIO-FLASHTM Trigger solutions (Inova Diagnostics, Inc., San Diego, CA). In some embodiments, a reagent of the kit of the present disclosure can include solutions useful for calibration or testing.

[0158] In some embodiments, a reagent of the kit can include a stop solution. Any stop solution known in the art can be included in a kit of this disclosure. The stop solutions of this disclosure terminate or delay the further development of the detection reagent and corresponding assay signals. Stop solutions can include, e.g., low-pH buffers (e.g., glycine-buffer, pH 2.0), chaotropic agents (e.g., guanidinium chloride, sodium-dodecylsulfate (SDS)) or reducing agents (e.g., dithiothreitol, β-mecaptoethanol), or the like.

[0159] In some embodiments, a reagent of the kit of this disclosure can include cleaning reagents. Cleaning reagents can include any cleaning reagent known in the art. In some embodiments, the cleaning reagents can be the cleaning reagents recommended by the manufacturers of the automated assay systems. In some embodiments, the cleaning reagents can include the BIO-FLASHTM System Rinse or the BIO-FLASHTM System Cleaning solutions (Inova Diagnostics, Inc., San Diego, CA).

[0160] A detection probe of the kit can include any of the detection probes described above. In brief, a detection probe of the kit can include antibodies and ligands. Thus, a “detection probe specific for anti-PAD IgA” includes, for example, PAD, a PAD: anti-PAD IgA complex binding agent, an anti-PAD IgA binding agent and an IgA binding agent. The anti-PAD IgA detection probes include binding agents to anti-PAD2 IgA, anti-PAD3 IgA and/or anti-PAD4 IgA.

[0161] A detection probe of the kit can be conjugated to any of the labels previously disclosed herein. For example, a detection probe can be conjugated to a fluorophore, an enzyme, a chemiluminescent moiety, a radioactive moiety, an organic dye, a small molecule, a polypeptide or functional fragment thereof. Examples of fluorophores include fluorescent dyes like phycoerytherin (PE), fluorescein isothiocyanate (FITC), tetramethylrhodamine (TRITC), BODIPY and AlexaFluor® dyes. Fluorescent dyes can also include fluorescence resonance energy transfer (FRET)-dyes or time-resolved (TR)-FRET dyes. Fluorophore labels also include fluorescent proteins such as green fluorescent protein (GFP) and cyan fluorescent protein (CFP). Examples of enzyme labels include alkaline phosphatase (AP) or horseradish peroxidase (HRP). When any of the substrates 3,3'5, 5'-Tetramethylbenzidine (TMB), 3,3'-Diaminobenzidine (DAB), or 2,2'-azino-bis (3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) are applied to HRP, a colored (chromogenic) or light (chemiluminescent) signal is produced. Radioactive moiety labels include carbon-14 or Tritium. Small molecule labels include biotin, resins such as agarose beads and fluorescently labeled magnetic beads, or nanoparticles such as colloidal gold. Polypeptide or functional fragment labels include Avidin, Streptavidin or NeutrAvidin which have an affinity for biotin. Polypeptide or functional fragment labels also include hemagglutinin (HA), glutathione-S-transferase (GST) or c-myc.

[0162] In some embodiments, the kit provided in this disclosure can include a component suitable for collecting a biological sample. A component can include collection tubes, columns, syringes, needles and the like. In some embodiments, the kit can include instructions for using the components of the kit. Instructions can be in any form, inside or outside of the kit. The instructions provide details regarding protocol and analytical techniques.

[0163] In some embodiments, a kit of the disclosure can include an instrument to an automated assay system. Automated assay systems can include systems by any manufacturer. In some embodiments, the automated assay systems can include, e.g., the BIO-FLASHTM, the BEST 2000TM, the DS2TM, the ELx50 WASHER, the ELx800 WASHER, the ELx800 READER, and the Autoblot S20TM (Inova Diagnostics, Inc., San Diego, CA). In other embodiments, an instrument of the kit can be a detection instrument. A detection instrument can include any detection instrument in the art. Detection instruments are capable of detecting or measuring a label of the reporter tags of the present disclosure. Thus, detection instruments are capable of detecting or measuring fluorescence, luminescence, chemiluminescence or absorbance, reflectance, transmittance, birefringence or refractive index. In some embodiments, detection instruments can include confocal and non-confocal microscopy, a microplate reader, a flow cytometer and the like.

[0164] Components of a kit of the disclosure can be in varying physical states. For example, some or all of the components can be lyophilized or in aqueous solution or frozen. Such components include a PAD, a detection probe, and ancillary reagents. Ancillary reagents include immobilization buffer, incubation buffer, washing buffer, dilution buffer, detection or assay buffer and blocking buffer. A person skilled in the art recognizes that there are various types of incubation, washing, detection and blocking buffers.

[0165] A kit of this disclosure can be tailored to specific assay technologies. In some embodiments, a kit can be tailored to assay technologies exemplified herein. For example, in some embodiments, the kits can be a FIA kit, a CIA kit, a RIA kit, a multiplex immunoassay kit, a protein/peptide array immunoassay kit, a SPRIA kit, an IIF kit, an ELISA, a PMAT kit, or a Dot Blot kit. In some embodiments, the ELSA kits can include a washing buffer, a sample diluents, a secondary antibody-enzyme conjugate, a detection reagent and a stop solution. In some embodiments, the Dot Blot kits can include a washing buffer, a sample diluents, a secondary antibody-enzyme conjugate, a detection reagent, and a stop solution. In some embodiments, the CIA kit can include a washing buffer, a sample diluent, a tracer (e.g., isoluminol-conjugate) and a trigger (e.g., H₂O₂). In some embodiments, the multiplex kit can include a washing buffer, a sample diluent and a secondary antibody-

enzyme conjugate. In some embodiments, the kits can be tailored to the Luminex platform and include, as an example, xMAP® beads.

[0166] A kit can be used to diagnose RA, severity of disease or joint erosion by providing a means for detecting anti-IgA bound to PAD or an antigenic fragment thereof. A kit can detect anti-IgA by any of the methods disclosed herein (see above). Complexes of anti-PAD IgA and a PAD, or antigenic fragment thereof, can have a stoichiometry of one to one or more than one to one anti-PAD IgA. In some embodiments, the complexes can have one anti-PAD IgA antibody per PAD or antigenic fragment thereof. In some embodiments, the complexes can have two anti-PAD IgA per PAD or antigenic fragment thereof. In some embodiments, the complexes can have more than two anti-PAD IgA per PAD or antigenic fragment thereof. Methods for measuring binding stoichiometries of two antigens are well known in the art and include, e.g., isothermal titration calorimetry (ITC) and ultracentrifugation.

[0167] In some embodiments, the complexes of anti-PAD IgA and PAD, or antigenic fragment thereof, can be a plurality of complexes with identical stoichiometry. For example, all complexes in the plurality of complexes have one anti-PAD IgA per purified PAD or antigenic fragment thereof. In some embodiments, the complexes of anti-PAD IgA and PAD or antigenic fragment thereof, can be a plurality of complexes with different stoichiometries. For example, some complexes in the plurality of complexes can have one anti-PAD IgA per PAD or antigenic fragment thereof and some other complexes in the plurality of complexes can have more than one anti-PAD IgA per PAD or antigenic fragment thereof.

[0168] In some embodiments, a PAD or antigenic fragment thereof can be bound by anti-PAD IgA with higher affinity. In some embodiments, anti-PAD IgA binding sites can be bound by anti-PAD IgA with more than 2-fold, more than 3-fold, more than 4-fold, more than 5-fold, more than 8-fold, more than 10-fold, more than 15-fold, more than 20-fold, more than 25-fold, more than 50-fold, more than 100-fold, more than 300-fold, more than 1,000-fold, more than 3,000-fold, more than 10,000-fold, more than 30,000-fold, or more than 100,000-fold greater binding affinity. Greater binding affinities are evidenced by lower dissociation constants (K_D s) for anti-PAD IgA-PAD complex or by higher association constants (K_A s) for the respective anti-PAD IgA and PAD. In some embodiments, the dissociation constants for (K_D s) for the anti-PAD IgA-PAD complexes can be less than 1 mM, less than 300 nM, less than 100 nM, less than 30 nM, less than 10 nM, less than 3 nM, less than 1 nM, less than 300 pM, less than 100 pM, less than 30 pM, less than 10 pM, less than 3 pM, or less than 1 pM. Methods for measuring binding affinities of antibodies to antigens are well known in the art and include ELISA, isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR).

EXAMPLE I

Detection of Anti-PAD4, Anti-PAD2 IgA and Joint Erosion in Rheumatoid Arthritis Patients

[0169] This example illustrates the use of anti-PAD4 IgA and anti-PAD2 IgA as biomarkers for the detection of joint erosion in rheumatoid arthritis (RA).

[0170] Anti-PAD4 IgA and anti-PAD2 IgA were measured using a particle-based multianalyte test (PMAT, Inova Diagnostics, San Diego, US). For this test, human recombinant full-length PAD4 polypeptide (Cayman Chemical, Ann Arbor, MI; cat no. 10500) and human recombinant full-length PAD2 polypeptide (John Hopkins University, Baltimore, MD) were coupled to paramagnetic beads with unique signatures. The coupling procedure includes bead activation, antigen coupling and bead blocking.

[0171] Bead activation was performed by incubating the beads for 30 min at room temperature with an activation buffer. Once the beads were activated, they were incubated with the antigen for 1 hour at room temperature in coupling buffer at a concentration of 22.2 µg of antigen/million of beads for PAD4 and 10 µg of antigen/million for PAD2. Finally, the beads were blocked for 1 hour at room temperature with PBS-TBN buffer. Once the beads were coupled, they were resuspended in a PBS-based assay resuspension buffer at a concentration of 1500 beads/test.

[0172] Measurement of anti-PAD4 IgA and anti-PAD2 IgA was performed as follows. First, sera from 41 RA patients with known erosion status were diluted 1:7 in Hemosil Rinse Solution (Inova Diagnostics, San Diego, CA). Next, PAD4 and PAD2 coupled beads were incubated for 9.5 min at 37° C. with patients' serum and assay buffer. After three washes with Hemosil Rinse Solution, the beads were then incubated for 9.5 min at 37° C. with a phycoerythrin (PE)-labeled anti-human IgA detector (Inova Diagnostics, San Diego, CA) at a concentration of 5µg/mL in QUANTA Flash Diluent (Inova Diagnostics, San Diego, CA). After incubation, beads were washed again in Hemosil Rinse Solution and the particles were analyzed through digital imaging technology. Finally, the Median Fluorescence Intensity (MFI) was calculated on the particles.

[0173] Anti-PAD4 IgA and anti-PAD2 IgA were significantly higher in RA patients with erosions compared to individuals without erosions ($p=0.0022$ and $p=0.0419$, respectively). See, FIG. 1 and FIG. 2. Discrimination between the populations of subjects with and without erosive disease reported an Area Under the Curve (AUC) of 0.704 (95% CI 0.529-0.879) for anti-PAD4 IgA. See FIG. 1. With an assay preliminary cut-off of 88 MFI determined by Receiver Operating Characteristic (ROC) curve analysis, anti-PAD2 IgA positive patients were 6.7 (95% CI 0.9-45.6) times more likely to have erosive disease. With an assay preliminary cut-off of 116 MFI, anti-PAD4 IgA positive patients were 3.2 (95% CI 0.8-13.4) times more likely to have erosive disease.

[0174] In conclusion, this example demonstrates that anti-PAD4 IgA and anti-PAD2 IgA are indicative of erosive disease in RA. These data further demonstrate that anti-PAD4 IgA and anti-PAD2 IgA represent useful biomarkers for patient stratification.

EXAMPLE II

Detection of Anti-PAD4 IgA, IgG and IgM in Rheumatoid Arthritis Patients

[0175] This example illustrates the use of anti-PAD4 IgA, IgG and IgM as biomarkers for the detection of joint erosion in rheumatoid arthritis (RA).

[0176] Bead activation was performed by incubating the beads for 30 min at room temperature with an activation buffer. Once the beads were activated, they were incubated

with the antigen for 1 hour at room temperature in coupling buffer at a concentration of 22.2 µg of antigen/million of beads for PAD4 and 10 µg of antigen/million for PAD2. Finally, the beads were blocked for 1 hour at room temperature with a PBS-TBN buffer. Once the beads were coupled, they were resuspended in a PBS-based assay resuspension buffer at a concentration of 1500 beads/test.

[0177] Measurement of anti-PAD4 IgA, IgG and IgM was performed as follows. First, sera from 62 RA patients with known erosion status were diluted 1:7 in Hemosil Rinse Solution (Inova Diagnostics, San Diego, CA). Next, PAD4 coupled beads were incubated for 9.5 min at 37° C. with patients' serum and assay buffer. After three washes with Hemosil Rinse Solution, the beads were then incubated for 9.5 min at 37° C. with a PE-labeled anti-human IgA, IgG or IgM detector (Inova Diagnostics, San Diego, CA) at the concentrations of 5, 1 and 5µg/mL, respectively, and diluted in QUANTA Flash Diluent (Inova Diagnostics, San Diego, CA). After incubation, beads were washed again in Hemosil Rinse Solution and the particles were analyzed through digital imaging technology. Finally, the Median Fluorescence Intensity (MFI) was calculated on the particles.

[0178] Anti-PAD4 IgA and IgM, but not IgG, were significantly higher in RA patients with erosions compared to individuals without erosions ($p=0.0004$, $p=0.0005$ and $p=0.9707$, respectively). ROC analysis showed higher AUC values for the discrimination between RA and controls for

anti-PAD4 IgA and IgM [0.70 (95% CI 0.60-0.80) and 0.70 (95% CI 0.59-0.80), respectively] than for anti-PAD4 IgG [0.50 (95% CI 0.39-0.61)]. At the relevant diagnostic area (>90% specificity), IgG outperformed the other two markers. See, FIG. 23. Discrimination for erosive disease was observed with anti-PAD4 IgA, followed by IgG and IgM. See, FIG. 24. Spearman correlation analysis showed moderate significant association between IgA and IgG (Spearman's rs-0.45, $p<0.0001$) and between IgA and IgM (Spearman's rs-0.45, $p<0.0001$), and a weak correlation between IgG and IgM (Spearman's rs-0.27, $p=0.0053$). When subjects with RA were stratified by presence or absence of erosive disease, higher titers of the three isotypes were observed in patients with erosive disease compared to individuals without erosions. However, this association was only significant for anti-PAD4 IgA and IgG ($p=0.0086$ and $p=0.0162$) (See, FIG. 25 and FIG. 26) but not anti-PAD4 IgM ($p=0.1756$) (See, FIG. 27).

[0179] In conclusion, the anti-PAD4 response in RA patients involves all three isotypes in RA. Anti-PAD4 IgA and IgG are associated with erosive disease in RA and represent useful markers for patient stratification.

[0180] It is understood that modifications, which do not substantially affect the activity of the various embodiments of this disclosure, are also included within the definition of the disclosure provided herein. Accordingly, the following examples are intended to illustrate but not limit the present disclosure.

SEQUENCE LISTING

```

Sequence total quantity: 92
SEQ ID NO: 1          moltype = DNA    length = 4361
FEATURE                Location/Qualifiers
misc_feature           1..4361
note = Homo sapiens peptidyl arginine deiminase 2 (PADI2),
          mRNA
source                 1..4361
mol_type = other DNA
organism = Homo sapiens

SEQUENCE: 1
aggctgctgg agaaggcgca cctgctgcag gtgctccgg ccgccccgga ccagcgacgc 60
cgggactgc ggccccggagg atgctgcgcg acgcggacccgt gcccgtgcag tacggagcc 120
gcgtggggc ggtgtacgtg ctgggcacct acctctggac cgatgtctac agcgcggccc 180
cagccggggc ccaaaccctt agccctgaacg actcggaaaca cgtgtgggtg gaggtgggtc 240
gtgtatgggg a ggtctgaggag gtggccacca atggcaaga ggcgtggctt ctctcgcucca 300
gcaccacccct gcgggttaccc atgacggccg cgagcacccg ggcacagca gacaaggta 360
ccgtcaacta ctatgacgag gaaggaggaca ttcccatcga ccaggccggg cttttctca 420
cagccatttg gatctccctg gatgtggacg cagacccggga tggtgtgggtg gagaagaaca 480
acccaaagaa ggcattctgg acctggggcc cccggggcca gggggccatc ctgtgggtga 540
actgtgaccg agagacacccct tggttgcctca aggaggactg ccgtgtatgaa aaggcttaca 600
gcaaggaaaga tctcaaggac atgtcccaaga tggatcctgcg gaccaaaggc cccgaccgccc 660
tccccggccg atacgagata gttctgtaca ttccatgtc agactcagac aaagtggccg 720
tgttctacgt ggagaaccccg ttacccgttcc acgcgttatccatccctg ggcggccggaa 780
agctctacca tgggttcaag tacccgggtt gtcgtggcggaa gtcgtgttcc ttcgtggaa 840
gcctctgttt ccccgacggag ggcttctcaag gtcgtgttcc catccatgtc agcctgttgg 900
agtacatggc ccaggacatt cccctgactc ccatcttccat ggacacccgtt atattccggaa 960
ttgtccctggt gatcatggcc cccaaacatcc tgcctccctgt gtcgtgttcc gtcgtgtca 1020
tgaaggatata ttacctgttcc tgcgttccatggg aacggatcatcc tggatccatgggaa accaactgtg 1080
agctgaagggt ctgttccatggg tgcgtgttcc tgcgtgttcc tgcgtgttcc ttcgtggaa 1140
agtttggcta catcgaggccc ccccttccatggg gtcgtgttcc ttcgtggaa accaactgtg 1200
atggaaacctt aaaggacttcc ttcgtgttcc ttcgtgttcc ttcgtgttcc ttcgtggaa 1260
cccgccggcc ccccttccatggg tgcgtgttcc ttcgtgttcc ttcgtgttcc ttcgtggaa 1320
gttcccttccatggg tgcgtgttcc ttcgtgttcc ttcgtgttcc ttcgtggaa accaactgtg 1380
gccttccatggg tgcgtgttcc ttcgtgttcc ttcgtgttcc ttcgtggaa accaactgtg 1440
agcagggttcc ttcgtgttcc ttcgtgttcc ttcgtgttcc ttcgtggaa accaactgtg 1500
atgttcatgtc ttttgcgttcc ttcgtgttcc ttcgtgttcc ttcgtggaa accaactgtg 1560
cctccgttccatggg tgcgtgttcc ttcgtgttcc ttcgtgttcc ttcgtggaa accaactgtg 1620
tgcgtgttcc ttcgtgttcc ttcgtgttcc ttcgtgttcc ttcgtggaa accaactgtg 1680
acggaggccct ttcgtgttcc ttcgtgttcc ttcgtgttcc ttcgtggaa accaactgtg 1740

```

-continued

```
SEQ ID NO: 2          moltype = AA  length = 665
FEATURE           Location/Qualifiers
REGION            1..665
source             note = protein-arginine deiminase type-2 [Homo sapiens]
                  1..665
                     mol_type = protein
                     organism = Homo sapiens
```

```
SEQ ID NO: 3          moltype = DNA    length = 3459
FEATURE             Location/Qualifiers
misc_feature        1..3459
                     note = PREDICTED: Homo sapiens peptidyl arginine deiminase
                           2 (PADI2), transcript variant X2, mRNA
source              1..3459
                     mol_type = unassigned DNA
                     organism = Homo sapiens
SEQUENCE: 3
accagacctg cagatggga ttcttctgtc agctcatata tgcacatgtc cacaaaadaca 60
```

-continued

taaacataca	agtgcactac	agtgtgtctt	cttgttagtag	aaaaagatg	caaagcaacc	120
taaatgtcca	tctgtggagg	actgagtaad	ggcccttccc	gaagttctgt	gcaaccgcta	180
aaaaggatgt	agecagtc	tcaatatgg	tacaggatgt	gttcraagga	tgaaatttgag	240
tttggctaca	tccggggccc	ccataaagg	ttccccgttg	tgctggactc	tcccccggat	300
ggaaaccta	aggactccc	tgtgaaggag	ctctggcc	cagatttgg	ctacgtgacc	360
cgggagcccc	tctttggatc	tgtcaccago	cttgactcat	ttggaaacct	ggaggctcgt	420
cccccaagt	cggtgaacgg	caagacata	ccgcttggcc	gtatccctat	cgggagcagc	480
tttcctctgt	ctgggtgtcg	gaggatgacc	aagggtggtc	gtgacttct	gaaggcccg	540
cagggtcagg	cgccgtgga	gtctactca	gactggctga	ctgtggcc	cgtggatgag	600
ttcatgtct	ttgtccccat	ccccggcaca	aqaaatcc	tgctactcat	ggccagcac	660
tcggcgtct	acaagctt	ccggagggaa	cagaaggacc	gcatggaga	ggccatcatg	720
ttcaaaaggct	ttgggtggat	gagcacaag	cgaaatcc	tcaacaatg	tctgtccaa	780
gagagcctt	tgccaggagaa	cctgtactt	cagcgctcc	tagactggaa	ccgtgacatc	840
ctcaagaagg	actgtggact	gacagagac	gatcatctt	acccgtcc	tctgttcaag	900
atggacgagg	accaccgtgc	cagaccctt	ttccccaaaca	tggtaacat	gtctgtctg	960
gacaaggacc	ttggcatccc	caagccat	ggggccacagg	ttgaggagga	atgtgcctg	1020
gagatgcacg	tgcggtggct	cctggagccc	ctgggctcg	aatgcaccc	categacac	1080
atttctgcct	accacaaaat	tctggggaa	gtccactgt	gcaccaacgt	ccgcaggaa	1140
cccttcacat	tcaagtggt	gcacatgtt	cctgaccc	ccaggggccc	ttggctttgc	1200
ctcttcgtct	tagtttcca	gacccttct	cacacgccc	gacccttc	ctgacatgg	1260
ctggacagcc	ccgctggag	accttggg	cgtgggggtt	aatttgggt	atctgtgcct	1320
tgccctccct	gagggggcc	tcagttgtct	ctgaaggcat	ccccactgt	cctcgactt	1380
gtccctgtgc	aaaatagctg	ggccagtg	tctgtagccc	tgacataagg	aacagaacac	1440
aacaaaaaac	agccaaacat	gtccccaa	tgctcccaa	tgcttaatc	tgcttaatc	1500
tgacactgaa	tgaggggaga	agggaggag	attctggat	tgccagttt	tccagcagcc	1560
atgctctgaa	aatcaaggt	gaatccatgg	aaagggacc	caggaccc	ggaccctaga	1620
cgtatctgt	actgcacatc	tcatttc	tcatctcc	tcagggttc	cagggtggca	1680
cccccaattt	ttcatttc	accaaccc	caaatcc	ttggcttc	tctgcgtgt	1740
ggacactgtt	ggctgttct	ccccactcc	tgagggtcc	gtaaagtt	ttaaactt	1800
cctggaaaca	tttcatctg	gcaggttcc	ccacgtgtt	gatgtctt	ttgcctcatc	1860
tgtctcagg	atgcaggatc	ccccccat	gtggggat	ctcccc	cagcatactt	1920
gtgacgttgc	agttcaatgc	gtaaagatgc	ccctggtag	ccatccat	tttcttc	1980
ctgtgttgc	attcttc	cgcccttgc	ccttcctt	tccactgct	tgactttt	2040
cctttttatt	cttgggtcc	tctgtccagg	cagctagaca	agaactt	cgccagcagc	2100
caagattcagg	ccttcccagg	ggcataataa	gtgaccagcc	cctcccttcc	ggacatcaga	2160
tccaacacat	aaggaccctg	gcctacc	gccccca	gcccatttgc	ggtcagctgc	2220
caacttaggg	gtgggttgc	tatccattt	aaatttacca	gtgccttgc	caaagaccc	2280
ctcatttgg	cataccca	ttcattcc	ggctcca	aaaaagactc	agtttcaatc	2340
gtttaaaagg	ccttttgg	cagaaga	aatgaattt	aatccat	tgaagaacc	2400
atttataacc	aatggaaagg	ttataatgt	atttattt	ttggaggaa	aagatttca	2460
tttgggat	tttccctt	ccatttca	aaatttgc	gtatgcact	aagcccccagg	2520
cacggcgtt	ggctctgc	acacagtgg	tagtagcgt	ctggacctt	tccctactt	2580
catggaccc	atactccccc	aacatgc	aaaaac	ggccacattt	aaaggcc	2640
tcagtgtgc	gtcttttgc	aatgcacgt	cactatattt	agaaatccgc	agggcaca	2700
acttccacga	agtactgtt	gtgtgtt	gggcgtgtt	gggtgggggt	tttctttt	2760
caggccccct	tcccatctac	ctagccat	ccatccat	gagtcccc	agccctccaga	2820
agctgtgtc	tcctcttgc	ggacagcgc	tcctgcctt	ggaggccaaa	gccccagatc	2880
tctccagccc	caagatgaa	aaacaccaat	gcttatttgc	gggtgtctgt	ctggagactt	2940
agagtttgc	atgtgtgt	gtgtgttgc	ttatgtgg	tttatgggtt	tttctttt	3000
ttttttttt	tttttttagt	ctacatttgc	gggaagtgc	cgccccc	gtgcagacag	3060
tgtgttttta	tagattttgc	taaggctt	ccaaatgt	tcggtaattt	ctgtatgtt	3120
tgaagtttcc	aggactcaca	caccgttcc	catctca	gcccccc	tgtgaca	3180
ctcgggttgc	atataccccc	gtgactcat	ggcttcc	cacccccc	ttctataat	3240
gtaggccat	aatacgctt	tctgttgc	aaactcagct	agttcctgt	tccaccttgc	3300
gtttaaaata	tcttatgt	gagggcagg	gtgtgtgc	agatggca	aagaacac	3360
tttcaaaattt	ctggaaaaga	gctgtgttgc	gagatctaa	gatgttttgc	gaagagctcg	3420
actaaagaac	aatgaat	atggtcca	gggaatgt			3459

SEQ ID NO: 4 moltype = AA length = 350
 FEATURE Location/Qualifiers
 REGION 1..350
 note = protein-arginine deiminase type-2 isoform X1 [Homo sapiens]
 source 1..350
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 4
 MSICGGLSNG LPGSSVQPLK RMKPVSQVGY RMCFKDEIEF GYIEAPHKGF PVVLDSRPG 60
 NLKDFPVKEL LGPDFGYVTR EPLFESVTL DSGGNLEVSP PVTVNKGKTYP LGRILIGSSF 120
 PLSGGRRMKT VVRDFLKAQQ VQAPVELYSD WLTVGHVDEF MSFVPPIPGTK KFLLLMASTS 180
 ACYKLFRKQ KDGHGEAIMF KGLGGMSSKR ITINKLISNE SLVQENLYFQ RCLDWNRDIL 240
 KKEGLLTERQD IIDLPALFKM DEDHRARAFF PNMVNMIVLD KDLGIPKPFQ PQVEECCLE 300
 MHVRGLLEPL GLECTFIDDI SAYHKFLGEV HCGTNVRRKP FTFKWWHMVP 350

SEQ ID NO: 5 moltype = DNA length = 3189
 FEATURE Location/Qualifiers

-continued

```
misc_feature          1..3189
                      note = Homo sapiens peptidyl arginine deiminase 3 (PADI3),
                           mRNA
source               1..3189
                      mol_type = unassigned DNA
                      organism = Homo sapiens
```

```
SEQ ID NO: 6          moltype = AA  length = 664
FEATURE              Location/Qualifiers
REGION               1..664
                     note = protein-arginine deiminase type-3 [Homo sapiens]
source               1..664
                     mol_type = protein
                     organism = Homo sapiens
```

SEQUENCE:	Organism	Homo sapiens
MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTGPVDI YISPNMERGR		60
ERADTRRWRF DATLEIIVVM NSPSNDLNDs HVQISYHSSH EPLPLAYAVL YLTCVDISLD		120
CDLNCEGRQD RNFVFDKRRQW WGPGSGYGGIL LVNCMDRDPs CDVQDNCDQH VHLCQLQDLEM		180
SVMVLRLTQGP AALFFDDHKLV LHTSSYDAK AQVFHICGPE DVCEARYHVL QGDVKSYEVp		240
RTHLGDEERFF VEGLSFPDAG FTGLISFHVT DDLSDNSEDs ASPIFTDTVV FRVAPWIMTP		300
STLPLPPEVYY CRVRNNTCFV DAAEALARKA GCKLTICPQA ENRNRDRWIQ EMELGYVQAP		360

-continued

HKTLPVVFDS	PRNGELQDFP	YKRILGPDFG	YVTREPRDRS	VSGLDSFGNL	EVSPVVANG	420
KEYPLGRILLI	GGNLPGSSGR	RVTQVVRDFL	HAQKVQPPVE	LFVDWLAVGH	VDEFLSFPVA	480
PDGKGKFRMLL	ASPGACFKLF	QEKKQKGHGR	ALLFQGVVDD	EQVKTISINO	VLSNKDLINY	540
NKFVQSCIDW	NREVLKRELG	LAECDIIDIP	QLFKTERKKA	TAFFPDLVNM	LVLGKHLGIP	600
KPFPGPIINGC	CCLEEKVRSLE	LEPLGLHCTF	IDDFTPYHML	HGEVHCGTNV	CRKPEPSFKWW	660
NMVP						664

SEQ ID NO: 8 moltype = AA length = 664
FEATURE Location/Qualifiers
REGION 1..664
source note = protein-arginine deiminase type-3 [Homo sapiens]
1..664

-continued

```

mol_type = protein
organism = Homo sapiens

SEQUENCE: 8
MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTGVDI YVSPNMERGR 60
ERADTRRWRF DATLEIIVVM NSPSNDLNDs HVQISYHSSH EPLPLAYAVL YLTCDISLD 120
CDLNCEGRQD RNFDVKRQWV WGPSCYGGIL LVNCDRDDPS CDVQDNCDQH VHCLQDLED 180
SVMVLRTQGP AALFDDHKLV LHTSYDAKR AQVFHICGPE DVCEAYRHVL GQDKVSYEV 240
RLHGDEERFF VEGLSFPDAG FTGLISFHVT LLDDSNEDFS ASPIFTDTVV FRVAPWIMTP 300
STLPPLLEVYV CRVRNNTCFV DAVAELARKA GCKLTICPQA ENRNRDRIQD EMELGYVQAP 360
HKTLPVVFDS PRNGELQDFP YKFLILGPDPF VYTREPRDRS VSGLDSFGNL EVSPVVANG 420
KEYPLGRILII CGNLPGSSGR RVTQVVRDPLI HAQKVQPPVE LFVDWLAVGH VDEFPLSFVPA 480
PDGKGFRMILL ASPGACFKLF QEKQKCGHGR ALLFQGVVDD EQVKTISINQ VLSNKDLINY 540
NKFVQSCIDW NREVLKRELG LAECDIIDIP QLFKTERKKA TAFFPDLVNM LVLGKHLGIP 600
KPFPGPIINGC CCLEEKVRSL LEPLGLHCTF IDDFTPYHML HGEVHCGTNV CRKPFSFKNW 660
NMVP 664

SEQ ID NO: 9
moltype = DNA length = 3189
FEATURE
misc_feature
1..3189
note = Homo sapiens peptidyl arginine deiminase 3
(PADI3) transcript, mRNA
source
1..3189
mol_type = unassigned DNA
organism = Homo sapiens

SEQUENCE: 9
agtgttgggg ttggcgccca cagtagtcgtaaaccacccgc atgtcgctgc agagaatcg 60
gcgtgttgc ctggagccat ccaccaggcggc ggtgtgtggc gttggcgctgg agaccctcg 120
ggacatttat gggtcgtgc ctggggccac agaaatgtttt gagggtctatgg gacgcctgg 180
cgtggacatc tacatcttc ccaacatggaa gaggggccggc gagcgtgcacg acaccaggcg 240
gtggcgctttt gacgcgactt tggagatcatc cgtggatcggaaactccccca gcaatgcacct 300
caacgcacgc catgttcagaat tttccatcca ctcacgcacat gggccatcgccccc 360
tgcgggtgtc taccacatcgttggatcatc tcttcggatc tgcgacccgtaaactgtgggg 420
aaggcggac aggaactttt tagacaaggcg gcagtggttgc tggggccca gtgggtatgg 480
cggcatcttgc ttggtaactt gtgaccgtgc tgatccggc tggatgttcc aggacaattt 540
tgaccgcac gtgcacgtgc ttgcacggatc ggaagacatc tctgtatgg tcctcgccgac 600
gcaggcccttc gacgccttc ttgtatggca caaacattgttc ctcacatccatc 660
tgccaaacggc gacacgttct tccacatcg cggccctcgatc gatgtgtgttggcaggctatag 720
gcatgtgtgc gggcaagata aggtgtccata tgaggatccatc cggccatcgatc gggatggg 780
gcgccttcgttgc gggccggcc tggatgttccatc tgatccggc ttcacacggatc tcatcttc 840
ccatgtacttgc tggatgtgc actccaaacgc gggatccatc tcttcgtatc 900
caactgtgttgc ttcccgatgttgc caccctggat catgacgcacgc acgacttcgc cacccttgc 960
gggtgtatgtc tgccgtgtca ggaacaacac gtgtttgttgc gatgcgggttgc cagactggc 1020
caaggaaaggcc ggtcttcgttgc tgacatccatc cccacacggcc gagaacccgc acgaccgtc 1080
gatccaggatc gagatggggc tggatgttgc tggatgttccatc cccacacggcc tccctgggtt 1140
cttgcgttccatc ccaaggaaatgc gggaaacttgc gggatccatc tccatccatc 1200
agatgtttgttgc tacgttgcactt gggaaaccacgc cgacaggatc tggatgttgc tggacttc 1260
tgggaacccgtt gggatgttgc cttccatc gggccatccatc aaaggatacc ccctggggag 1320
gtatccatcgttgc gggccatccatc tggatgttgc aatggatacc ccctggggatc 1380
ggacttccttcatc gggccatccatc aatggatccatc aatggatccatc aatggatccatc 1440
cgtggccatc gtggatgttgc ttctgtatc tggatgttgc cccatccatc 1500
gatgtgttgc tggatgttgc gggccatccatc aatggatccatc aatggatccatc 1560
ccacggggggccatc gggccatccatc tggatgttgc aatggatccatc aatggatccatc 1620
catcaacccatc gggccatccatc ataaaggatccatc aatggatccatc aatggatccatc 1680
catcgacttgc aaccgttgc tggatgttgc aatggatccatc aatggatccatc 1740
tgacatccatc cggatgttgc aatggatccatc aatggatccatc aatggatccatc 1800
gggtgaatcatc tggatgttgc gggccatccatc aatggatccatc aatggatccatc 1860
caatggatgttgc tggatgttgc aatggatccatc aatggatccatc aatggatccatc 1920
ctgcacccatc atggatgttgc tggatgttgc aatggatccatc aatggatccatc 1980
caccaatgttgc tggatgttgc aatggatccatc aatggatccatc aatggatccatc 2040
ctcccccacatc cccatccatc tggatgttgc aatggatccatc aatggatccatc 2100
caggccatccatc aacggatgttgc aatggatccatc aatggatccatc aatggatccatc 2160
acggccatccatc tggatgttgc aatggatccatc aatggatccatc aatggatccatc 2220
gttcttcgttgc tggatgttgc aatggatccatc aatggatccatc aatggatccatc 2280
tccatccatc aatggatccatc aatggatccatc aatggatccatc aatggatccatc 2340
ggggccatccatc cccatccatc tggatgttgc aatggatccatc aatggatccatc 2400
tgcacatccatc cccatccatc tggatgttgc aatggatccatc aatggatccatc 2460
gagggccatccatc aatggatccatc aatggatccatc aatggatccatc aatggatccatc 2520
agccctccatc aatggatccatc aatggatccatc aatggatccatc aatggatccatc 2580
ccggccatccatc tggatgttgc aatggatccatc aatggatccatc aatggatccatc 2640
agcttccatc tggatgttgc aatggatccatc aatggatccatc aatggatccatc 2700
atctgcacatc cccatccatc tggatgttgc aatggatccatc aatggatccatc 2760
gtatctggggatc tggatgttgc aatggatccatc aatggatccatc aatggatccatc 2820
taatacacaatc tggatgttgc aatggatccatc aatggatccatc aatggatccatc 2880
aagggttccatc tggatgttgc aatggatccatc aatggatccatc aatggatccatc 2940
ctggatgttgc tggatgttgc aatggatccatc aatggatccatc aatggatccatc 3000
cgaggagggccatc tggatgttgc aatggatccatc aatggatccatc aatggatccatc 3060

```

-continued

```
gaatgaacca ctgaattcag gggatgggg tggggggccg gttctcgagg tgtgtccag 3120
ctacacgtgt gttctgtatg ggtccagctg cgtttccatc actcgtaat aaatcaacag 3180
aaacacaaa 3189
```

```
SEQ ID NO: 10 moltype = AA length = 664
FEATURE Location/Qualifiers
REGION 1..664
note = protein-arginine deiminase type-3 [Homo sapiens]
source 1..664
mol_type = protein
organism = Homo sapiens
```

```
SEQUENCE: 10
MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTGVDI YISPNMERGR 60
ERADTTRRWF DATLEIIVNM NSPNSDLNDS HQQISYHSSS EPLPLAYAVL YHTCVDISLD 120
CDLNCEGRQD RNFVDKRQWV WGPSGYGGIL LVNCNRDDPS CDVQDNCDQH VHCLQDLED 180
SVMVLRTQGP AALFDDHKLV LHTSSYDAKR AOVFHICGPE DVCEAYRHVL QGDKVSYEVP 240
RLHGDEERFF VEGLSFPDAG FTGLISPHVT LLDDSNEDFS ASPIFTDTVV FRVAPWIMTP 300
STLPPLVEYY CRVRNNTCFV DAVIELARKA GCKLTICPQA ENRNDRWIQD EMELGYVQAP 360
HKTPLPVVFDS PRNGELQDFP YKRILGPDFG YVTREPRDRS VSGLDSFGNL EVSPPVANG 420
KEYPLGRILII CGNLPGSSGR RVTQVVRDFL HAQKVQPVVE LFVDWLAVGH VDEFLSFVPA 480
PDGKGFRMLL ASPGACFKLF QEKKQKCGHGR ALLFQGVVDD EQVKTISINQ VLSNKDLINY 540
NKEFVQSCIDW NREVLKRELG LAECDIIDIP QLFKTERKKA TAFFPDLVNM LVLGKHLGIP 600
KPFGPIINGC CCLEEKVRSL LEPLGLHCTF IDDFTPYHML HGEVHCGTNV CRKPFNSFKNW 660
NMVP 664
```

```
SEQ ID NO: 11 moltype = DNA length = 3189
FEATURE Location/Qualifiers
misc_feature 1..3189
note = Homo sapiens peptidyl arginine deiminase 3 (PADI3)
transcript, mRNA
source 1..3189
mol_type = unassigned DNA
organism = Homo sapiens
```

```
SEQUENCE: 11
agtgttgggg ttggcgccca cagctaagtc caacaccaggc atgtcgctgc agagaatcgt 60
gcgtgtgtcc ctggagcatc ccaccagcgc ggtgtgttg gctggcgctgg agaccctcgt 120
ggacatttat gggtcagtcgct ctggaggcac agaaatgttt gaggtctatg ggacgcctgg 180
cgtggacatc tacatctctc ccaacatggc gaggggccgg gaggcgtgcag acaccaggcg 240
gtggcgttgc gacgcgactt tggagatcatc ctgtgttcataa aactccccca gcaatgcacct 300
caacgcacgc catgttcaga ttcccttacca ctccagccatc gaggcgtctgc ccctggctta 360
tgcgggtgtc tacctcacct gtgttgacat ctctctggat tgcgacccgtg actgtgagggg 420
aaggcaggac aggaactttg tagacaagcg gcagtggttc tggggccca gtgggtatgg 480
cgccatcttg ctgggtqaact gtgacgtgtca tgatccgacg tggatgtgtcc aggacaattg 540
tgaccagcac atgcactgcc tcgaagacatc ggaagacatc tctgtcatgg tctctggag 600
gcaggcccctt gcaagccctt ttgtatgacca caaacatgtc ctccataactt ccagctatga 660
tgccaaaacgg gcacaggctt tccacatctg cggcgtctggat gatgtgtgtg aggctatag 720
gcatgtgtgc ggcacaaatgc aggtgttcata tgaggttccca cggctgtcaatc gggatgaggaa 780
gcgccttcttc tggttggccctt tgatggggcc ttcaaggacatc tcatctctt 840
ccatgtcaatc ctgcgtggacg acttccaaacgaa ggatccctcg gcatccccca tcttcactga 900
cactgtgtgt tcccgatgtgg caccctggat catgacgcccc agcaactctgc cacccttaga 960
gttgtatgtg tgccgtgtgc ggaacaaacatc gtgtttgtgat gatggggggc cagagctggc 1020
caggaaggcc ggcgtcaagc tgaccatctg cccacaggcc gagaaccgcg acgaccgtcg 1080
gatccaggat gagatggcgtt tggttacgt tcaggcgtcc cacaagaccc tccctgggtt 1140
ctttgactcc ccaaggaaatg gggaaactgcgaa ggatccctt taaaaaaaatc tccctgggtcc 1200
agattttgtt taatgtgtcc gggaaaccacg cgacaggatctt gtgatgtggc tggactctt 1260
tggaaacatcg gagggtcggcc ctccatgtt ggecaatggg aaagatgttcc cccctggggag 1320
gatcccttattt gggggcaacc tgcctgggtt aatggccgcg agggttacccc aggtgtgtcg 1380
ggacttcttc catggccaga aggtgcggcc cccctgggtt ctctttgtgg actgtgttgc 1440
cggtggccat gtggatgtgtt ttccatgtt ccgcgtggc ggggttcccg 1500
gatgtgttgc gccaggccctg gggccatgtt caatgttccatc cggaaaagc agaagtgtgg 1560
ccacggggagg goccttctgt tccagggggtt tggtatgtt cggatgttca agaccatctc 1620
catcaaccatcg gtgtcttcca ataaaagatctt catcaactac aataatgttgc tgcagatgt 1680
catcgactgg aaccgtgagg tgctgtcaagcg ggagctggcc ctggcgagatgttgcacatcat 1740
tgacatccca cagcttttca agacatgttccatc gggatccatc tccctgtactt 1800
ggtaacatcg ctgggtgtgg ggaagacatcg gggatccatccc aaccatcttgc gggccatcat 1860
caatggctgc tgctgtgtgg agggaaatgttccatc gggatccatccc tggatgttccatc 1920
ctgcacccatcg attgtatgttccatc ccacatgttccatc tggatgttccatc 1980
caccatgttccatc cttccatgttccatc aatggatgttccatc tggatgttccatc 2040
ctcccaaccatcg ccatccatcg tccatgttccatc aaccatcttgc cccatgttccatc 2100
caggccccctg aacatgttccatc accaagatgttccatc cccaaaggatgttccatc tggatgttccatc 2160
accgttccatc tccatgttccatc tttccatgttccatc aaccatcttgc aaccatcttgc 2220
gttcttcacatcg ttgtatgttccatc tggatgttccatc aaaaagaagg acgttccatc ttatgttccatc 2280
tccatgttccatc aacatgttccatc catggatgttccatc tccatgttccatc aaccatcttgc 2340
gggacaaatcg cacatgttccatc tccatgttccatc tccatgttccatc tccatgttccatc 2400
tgcatccatcg cagatgttccatc atccatgttccatc ctgtttgttccatc cccatgttccatc 2460
gagggacaaatcg cccatgttccatc gggatgttccatc tccatgttccatc tccatgttccatc 2520
```

-continued

agectcccccc	ataaaaaggg	agctgtggat	ccacttagat	caggcgaa	ccatcttca	2580
cccgccaaag	ctccgtccca	gatgttgacc	ctcacccagc	gtgagctgc	acatagttagg	2640
agcttctaga	tgcatgtgga	agcaatgaga	gttgccttgc	agcctataa	actccccatg	2700
atctgacatg	cagaataatcca	gccttgcac	gaatccttc	ggaattttt	ggagacgaaa	2760
gtatctgggg	gattgtggg	tactagggag	actgggtaca	agggtgaaaa	gtatgtccca	2820
taatacacat	ggttgactat	ggtgatccac	cttgtatgg	ttaatattag	gtgtctggag	2880
aagggttgc	cattggccct	gggacttctc	tctgcaggag	gagagaacgc	tgcctctcct	2940
ctggatgtgt	ctcaggctct	ctgttgcct	tttgtcaagc	tttccacatc	ctgtctgtct	3000
gcaggagagg	gggctaaggg	gctggatcca	ccaaggcgc	tcacagcggg	aaaactctgg	3060
gaatgaacca	ctgaattcag	gggatggggg	tggggggggc	gttctcgagg	tgtgtgcag	3120
ctacacgtgt	gttctgtatg	ggtccagctg	cgtttccatc	actcgtcaat	aatcaacag	3180
aaacacaaaa						3189
SEQ ID NO: 12		moltype = AA	length = 664			
FEATURE		Location/Qualifiers				
REGION		1..664				
source		note = protein-arginine deiminase type-3 [Homo sapiens]				
		1..664				
		mol_type = protein				
		organism = Homo sapiens				
SEQUENCE: 12						
MSLQRIVRVS	LEHPTSAVCV	AGVETLVDIY	GSVPEGTEMF	EVYGTGVDI	YISPNNMERGR	60
ERADTRRWRF	DATLEIIVVM	NSPSNDLND	HVQISYHSS	EPLPLAYAVL	YLTCVDISLD	120
CDLNCEGRQD	RNFVDKRQWV	WGPSCYGGIL	LVNCDRDRDS	CDVQDNCDQH	MHCLQDLED	180
SVMLVLTQGP	AALFDDHKLV	LHTSSYDAKR	AQVFHICGPE	DVCEAYRHVL	GQDKVSYEV	240
RLHGDEERFF	VEGLSPPDAG	FTGLISFHVT	LLDSNEEDFS	ASPIFTDTVV	FVAPWIMTP	300
STLPPLPEVYV	CRVLRNNTCFV	DAVAELARKA	GCKLTICPQA	ENRNRDRWIQD	EMELGYVQAP	360
HKTLPVVFDSD	PRNGELQDFP	YKRILGPDPG	YTTRPREPRDRS	VSGLDSFGNL	EVSPVVANG	420
KEYPLGRIL	GGNLPGSSGR	RVTQVVRDFL	HAQKVQPVPE	LFVDWLAVGH	VDEPLSFVPA	480
PDGKGFRM	ASPGACFKLF	QEOKQKCGHGR	ALLFQGVVDD	EQVKTISINQ	VLSNKDLINY	540
NKFVQSCIDW	NREVLKRELG	LAECDIIDIP	QLFKTERKA	TAFFPDLVNM	LVLGKHLGIP	600
KPPGIINGC	CCLEEKVRSL	LEPLGLHCTF	IDDPFTPYHML	HGEVHCTNV	CRKPFSFKWW	660
NMVP						664
SEQ ID NO: 13		moltype = DNA	length = 3189			
FEATURE		Location/Qualifiers				
misc_feature		1..3189				
source		note = Homo sapiens peptidyl arginine deiminase 3 (PADI3) transcript, mRNA				
		1..3189				
		mol_type = unassigned DNA				
		organism = Homo sapiens				
SEQUENCE: 13						
agtgttgggg	ttggccca	cagtaagtc	caacaccaggc	atgtcgctgc	agagaatcgt	60
gctgtgtc	ctggagcatc	ccaccaggcg	ggtgtgtgt	gttgcgttgg	agaccctcg	120
ggacatttat	gggtcagtgc	ctgagggcac	agaaatgtt	gagggtctatg	ggacgcctgg	180
cgtggacatc	tacatctctc	ccaaacatgg	gaggggccgg	gagcgtgcag	acaccaggcg	240
gtggcgctt	gacgcgactt	tggagatcat	cgtggatcatg	aactccccca	gcaatgcacct	300
caacgcacgc	catgttcaga	tttccatcca	cttcagccat	gagccttcgc	ccctggccta	360
tgcggtgtc	tacccatccc	gtgttgacat	ctctctggat	tgcgacactga	actgtgaggg	420
aaggcggac	aggaactttg	tagacaagcg	gcagtggggtc	tggggggccca	gtgggtatgg	480
cggcatcttgc	tgggtgact	tgacccgtg	tgtatccggc	tgtatgttcc	aggacaatttgc	540
tgaccacgc	gtgcactgc	tgcgaacactt	qaaagacatg	tctgtatcatgg	tcctcgccgac	600
gcagggccct	gcagccctct	ttgatgacca	caaacttgc	ctccataactt	ccagctatga	660
tgccaaacccgc	gcacaggcttgc	tccacatcg	cggtccctgt	gatgtgtgt	aggcctatag	720
gcatgtgtgc	ggccaaagata	agggttctca	tggatgttcc	cgcttgcatg	gggtatggaga	780
gcccgttcttgc	tggaaaggcc	tgtcttccc	tgtatccggc	ttcacaggac	tcatctcttgc	840
ccatgtca	ctgtggacgc	actccaaacga	ggatttctcg	geatccccca	tcttcactga	900
caactgtgttgc	ttcccgatgttgc	taccctgtat	catgacgcggc	agcaacttcgc	cacccttaga	960
ggtgttatgt	tgccctgtgt	ggaacaacac	gtgttttgcgt	gtatcggtgg	cagagcttgc	1020
caggaaaggcc	ggctcgaaatgc	tgaccatctg	cccacaggccc	gagaacccca	acgccccgttgc	1080
gatccaggat	gagatggagc	tgggttacgt	tcaggcgccg	cacaagaccc	tccctgggttgc	1140
cttgcgttcc	ccaaggaaatg	ggggaaactgc	ggatttccctt	tacaaaagaa	tcctgggttgc	1200
agatgttgcgttgc	tgacgttgcgttgc	ggggaaaccacg	cgacagggttgc	gtgagttggcc	ttggacttccttgc	1260
tgggaacatgc	gagggtcagcc	ctccatgtgttgc	ggccaaatgggg	aaagagtacc	ccctgggggg	1320
gatcccttc	ggggccaaatgc	tgccctgggttgc	aatgtggccgc	agggttacccc	aggtgtgttgc	1380
ggacttc	catggccca	agggtcagcc	ccccgtggat	ctctttgtgg	actgtgttgc	1440
cgtggccat	gtggatgtgttgc	ttctgtatgttgc	ttccatgtggat	agggttccat	agggttccat	1500
gatgtcttc	gocagccctgt	caagcttgc	caggaaaatgc	agaagtgttgc	1560	
ccacggggagg	ggcccttcgttgc	tccagggttgc	tgttgcgtat	gagcgttca	agaccatctc	1620
catcaacccat	gtgtcttc	ataaaacatgc	catcaactac	aataagtttgc	tgcacgttgc	1680
catcgactgg	aaacgtgttgc	tgctgttgc	ggagctggggc	ctggcagatgt	gtgacatcat	1740
tgacatccca	cagcttc	agaccgttgc	gaaaaaaatgc	acggccatcttgc	tccctgtat	1800
ggtgaacatgc	ctgggtgttgc	ggaacgttgc	ggggcatcccc	aagcccttgc	ggcccatcat	1860
caatggctgc	tgctgttgc	aggaaatgg	gcccgttgc	ctggagccgc	tgggttccca	1920
ctgcacccat	attgtatgttgc	ccacatgttgc	catggggagg	tgcactgttgc		1980

-continued

caccaatgtg	tgcagaaaagc	ccttcttttt	caagtgggtgg	aacatgggtgc	cctgagacag	2040
ctccccaccca	ccatccctgtc	ccccctggggc	gggcatttggc	ccagggtgggt	gagacagaga	2100
caggcccccc	aacgataaaagc	accaagagac	cccaaggatcc	cagatggaaad	actggagggtt	2160
accgtttccct	tcaagaaggct	tttccttggaa	agtgtccat	ccttcacccgtc	aaaccatgttg	2220
gttctcgac	tttgaatcttc	tcggccccccc	aaaaaaagg	acccatatttc	ttatagcctc	2280
tcctgtgatt	caacacaacc	catggagat	tccccttctc	actctgaaat	catccatttg	2340
gggacaaacc	cacatgggg	tctagaacaa	tccacgtatc	tcatcagcc	tcttgccttc	2400
tgcctcttaa	cagaggaagg	atccatgtt	ctgtcttggt	ccaaatgttt	ccttcgtcga	2460
gaggaaacaa	ccttaaaaacca	gaccatccca	cgccaggacag	gcaggagaga	tcttccttaa	2520
agcctcccc	ataaaaaagg	agctgtggat	ccacttagat	caggggcgaa	ccatctttca	2580
ccccggccaa	cttcgtccca	gtatgttgc	ctccacccagc	gtgactgtc	acatagtagg	2640
agctttctaga	tgcatgttgg	agcaatgaga	gttgtccctt	agccatataat	actccccat	2700
atctgacatg	cagaaaaatcc	gccttgc	gaatcttctt	gaaatttttt	ggagacggaa	2760
gtatctgggg	gattgttggg	tacttagggag	actgggtaca	aggggtaaaa	gtagttccca	2820
taatacacat	ggttgtactat	ggtgatccac	cttgcgtat	ttaaatattt	gtgtctgtt	2880
aaagggtgtt	cattggccct	ggggacttctc	tttcgcggag	gagagaacgc	tgctcttct	2940
ctggatgtt	ctcaggctct	ctgttggcc	tttgcgtat	tttccacatc	ctgtctgtgt	3000
gcaggagagg	gggctaagg	gctggatcca	ccaaaggcgc	tcacacgggg	aaaactctgg	3060
gaatgtacca	ctgaatttca	gggatgggg	tgggggggg	gttctcgagg	tgtgtccag	3120
atcacacgtt	gttctgtat	ggtccagctg	cgtttccatc	actcgctaat	aaatcaacag	3180
aaacacaaa						3189

```
SEQ ID NO: 14          moltype = AA    length = 664
FEATURE                Location/Qualifiers
REGION                 1..664
note = protein-arginine deiminase type-3 [Homo sapiens]
source                 1..664
mol_type = protein
organism = Homo sapiens
```

SEQUENCE :	14	ORGANISM	NAME	DESCRIPTION	60
MSLQRIVRVS	LEHPTSAVCV	AGVETLVDIY	GSVPGEFTMF	EVYGTGPVDI	YISPNNMERGR
ERADTRWRRE	DATLEIIVVM	NPSNPDLLND	HQVQISYHSSH	EPLPLAYAVL	YLTCVDISLD
CIDLNCERGQD	RNFVVKRQWV	WGPGSCYGGYL	LVNCDRDRDD	CDVQDNCDQH	VHCLQDLED
SVMVLRTQGP	AALFDHHKLV	LHTSSYDAKR	AQVFHICGPE	DVCEAYRHVL	GQDKVSYEV
RLHGDEERF	VEGLSLPPDAG	FTGLISFHV	LLDDSNEDFS	ASPIFTDTVV	FRVVPWIMTP
STLPLPPEVY	CRVRNNTCFV	DAEVALARKA	GCKLTICPQA	ENRNDRWQI	EMELGYVQAP
HKTLPLPVFD	PRNGELQDFD	YKIRLGPDFG	YTVTREPRDRS	VSGLDSFGNL	EVSPVVANG
KEYPLGRILLI	GGNLPGSSGE	RVTQVVRDFL	HQKVQPPBVE	LFDVWLAVGH	DEFLSFVPA
PDKGKFRMLL	ASPGACFKLF	QEKKQKGHGR	ALLFQGVVDD	EQVKTISINQ	VLSNKDLINY
NKFVQSCIDW	NREVILKRELG	LAECIDIIP	QLFKTERKKA	TAFFPDLVNM	LVLGKHLGIP
KPF GPIINGC	CCLEEKVRSL	LEPLGLHCTF	IDDFTPYHML	HGEVHCGTNV	CRKPFSFKWW
NMVP					664

```
SEQ ID NO: 15          moltype = DNA    length = 3189
FEATURE
misc_feature          Location/Qualifiers
                      1..3189
note = Homo sapiens peptidyl arginine deiminase 3 (PADI3)
       transcript, mRNA
source
                      1..3189
mol_type = unassigned DNA
organism = Homo sapiens
```

```

SEQUENCE: 15
agtgttgggg ttggcggcca cagctaagtc caacaccagc atgtcgctgc agagaatcg 60
ggctgtgtcc ctggagcatc ccacccaggc ggtgtgtgtg gtgcggctgg agaccctcg 120
ggacatttat gggtagtcgtt cttggggcac agaaatgttt gaggttctatg ggacgcttg 180
cgttgcacatc tacatcttc ccaacatggc gagggggccgg gagctgtcag acaccaggcg 240
gtggcgcttt gacgcgactt tggagatcat cgtggtcatg aactccccca gcaatgacct 300
caacgcacgc catgttcaga tttcttacca ctccaggcat gacgcctcg cccttgcctg 360
tgcgggtgtcc taccttcact gtgtgtacat ctcttcgttgc gtcgacctgtg actgtgagg 420
aaaggcaggac aggaactttt tagacaacggc gcacgtggtc tggtggggccca tggtggat 480
cggtcatctt ctggtaact gtgaccgtg tgatccggc tgcgtatgtcc aggacaattt 540
tgaccagcac gtgcactgcc tgcaagaccc ggaagacatg tctgtcatgg tccctggac 600
gcaggccccc tcagcccttc ttgtatgacca caaaatgttc ctccataccat ccagctatg 660
tgccaaacggc gcacaggctt tccacatctt cgggtcttgc gatgtgtgtg aggctatag 720
gcatgtgtcc ggccaaatggc aggtgtccca tgaggtaaccc cggttgcatg gggatggaga 780
gcgcgttctt ctggaaaggcc tgtcctttcc tgatgcggc ttcacaggac tcatctcctt 840
ccatgtcact ctgtggacg actccaaacga ggatttctcg gcacccctta tcttcactga 900
caactgtgtt ttcccgatgg cacccctggat catgaacggc agacatctgc cacccttgc 960
ggtgtatgtt tgccgtgtga ggaacaacac gtgtttgtg gatgcgggtt cagagctggc 1020
caggaaaggcc ggctcgaaacg tcgaccatctt cccacaggcc gaaacccgcg acaccgtgg 1080
gatccaggat gagatggacg tgggtacatgc tcaggccggc cacaagaccc tcccggtgtt 1140
ctttgactcc ccaaggaaatg gggaaactgca ggatttccct taaaaaaaatg tcctgggtcc 1200
agatttttgtt tacgtgtactc gggaaacccacg cgacacgtt gtgagtgccc tggactccctt 1260
tggggaaacctt gagggtcggcc cttccatgtt gggccatgtgg aaaggttaccgc cccttggggat 1320
gatccatctt cggggccaaacg tgcctgggtt aagttggccgc aggttgcacc cagggttgcg 1380
ggacttcttc catggcccaaga aggtgcggcc cccctggag ctctttgtgg actgtgtggc 1440

```

-continued

```
SEQ ID NO: 16          moltype = AA  length = 664
FEATURE                  Location/Qualifiers
REGION                   1..664
                           note = protein-arginine deiminase type-3 [Homo sapiens]
source                   1..664
                           mol_type = protein
                           organism = Homo sapiens
```

SEQUENCE :	16	60
MSLQRIVRVS	LEHPTSAVCV	AGVETLVDIY
ERADTRWRKV	DATLEIIVM	GSVPEGTEMF
CIDLNCERQD	NSPSNDLNDs	EVYGTGVDI
RNFVDRKQWY	HVQISYHSSH	YISPNMERGR
SVMVLRTQGP	WGPSPGGYL	EPLPLAYAVL
AALFDDHKLV	LVNCDRDDS	YLTCVDISLD
LHGDEERFF	CDVQDNCDQH	VHCLQDLED
VEGLSLPPDAG	GQDKVSYEV	180
FTGLISPFVHT	RL	240
LLDDSNEDFS	RLHGDEERFF	300
ASPIFTDTVV	VEGLSLPPDAG	360
FRVAPWIMTP	FTGLISPFVHT	360
ENRNRDWRQD	LLDDSNEDFS	420
EMELGYVQAP	ASPIFTDTVV	480
YTRPRDPDRS	FRVAPWIMTP	480
VSGLDSFGNL	ENRNRDWRQD	540
EVSPVVPANG	EMELGYVQAP	540
KEYPLGRILLI	YTRPRDPDRS	600
GGNLPGSSGR	VSGLDSFGNL	660
RVTQVVRDFL	EVSPVVPANG	660
HAQKVQPVPE	KEYPLGRILLI	664
LFVDWLAVGH	GGNLPGSSGR	
VDEFLSFVPA	RVTQVVRDFL	
PDGKGFRMLL	HAQKVQPVPE	
ASPGACFKLF	LFVDWLAVGH	
QEKKQKGHGR	VDEFLSFVPA	
ALLFQGVVDD	PDGKGFRMLL	
EQVKTISINQ	ASPGACFKLF	
VLSNKDLINY	QEKKQKGHGR	
NKFVQSCIDW	ALLFQGVVDD	
NREVLKRELG	EQVKTISINQ	
LAECIDIIDIP	VLSNKDLINY	
QLFKTERKKA	NKFVQSCIDW	
TTFFPDLVNM	NREVLKRELG	
LVLGKHGLP	LAECIDIIDIP	
KPFPGIINGC	QLFKTERKKA	
CCLEEKVRSL	TTFFPDLVNM	
LEPLGLHCTF	LVLGKHGLP	
IDDFTPYHML	KPFPGIINGC	
HGEVHCGTNV	CCLEEKVRSL	
CRKPFSFKWW	LEPLGLHCTF	
NMVP	IDDFTPYHML	

```
SEQ ID NO: 17          moltype = DNA    length = 3189
FEATURE                  Location/Qualifiers
misc_feature             1..3189
                           note = Homo sapiens peptidyl arginine deiminase 3
                           (PADI3)transcript, mRNA
source                   1..3189
                           mol_type = unassigned DNA
                           organism = Homo sapiens
```

SEQUENCE:	17	organism = Homo sapiens	
agtgttgggg	ttggcgccca	cagctaagtc caacaccaggc atgtcgctgc agagaatcg	60
gegtgtgtcc	ctggagcatc	ccacccaggc ggtgtgtgtg gctggggctgg agacccttcg	120
ggacattat	gggtcagtgc	ctgaggggcac aaaaatgttt gaggttctatg ggacgcctgg	180
cgttgacatc	tacatcttc	ccaaacatggc gagggggccgg gagctgtcgac acaccaggcg	240
gtggcgcttt	gacgcccactt	tggagatcat cgtggctatg aactccccca gcaatgacct	300
caacgcacgc	catgttcaga	tttccatcca ctccagccat gggcctctgc ccctggccata	360
tgccgtgtc	tacccacttgc	tgtggactatc ctccatggat tgcgacactga actgtggagg	420
aaggcaggac	aggaacttgc	tagacaagecg cggatgggtc tggggggccca gtgggtatgg	480
accatcttgc	ctggtaacttgc	gtgaccgtgc tgatccggc tgcgtatgtcc accggatcc	540
tgaccacgcac	gtgcaactgcc	tgcaagacactt ggaagacatg tctgtcatgg tcctcgccgac	600
gcaggggccct	gcagccctctt	ttgatgacca caaacttgc tcccatcacct ccagctatga	660
tgcaccaacgg	gcaccaactgc	tccacatctg cgggttctggat gatgtgtgtg aggccatgg	720
gcatgtgtcttgc	ggccaaagata	aggtgtctca tgaggatcaccc cgcttgcatg gggatgagga	780
ggccatcttc	tggttggggcc	tgtttccccc tgatccggc ttcacaggac tcatacttc	840
ccatgttact	ctgtggacgc	acttccaaacgaa ggattttcgg gcatccccca tcttcactgt	900

-continued

cactgtggtg	ttccgagtgg	caccctggat	catgacgccc	agcactctgc	cacccttaga	960
ggtgtatgtg	ttccgtgtg	ggaaacaac	gtgttttgg	gatgcgggg	cagagctggc	1020
caggaaaggcc	ggctgaagc	tgaccatctg	ccccacaggcc	gagaaccgca	acgaccgctg	1080
gatccaggat	gagatggagc	tgggtacgt	tcagggccc	cacaagaccc	tcccggtgg	1140
ctttgactcc	ccaaggaaatg	gggaaactgca	ggatttccct	tacaaaagaa	tcctgggtcc	1200
agattttgg	tacgtgactc	gggaaccac	cgacaggctc	gtgagtggcc	tggactcctt	1260
tgggaacctg	gagggtcagcc	ctccagttgt	ggccaaatggg	aaagagtacc	ccctggggag	1320
gatcctcatt	ggggggcaacc	tgcctgggtc	aagtggccgc	agggtcaccc	aggttgttgc	1380
ggacttectc	catgcccaga	aggtgcagcc	ccccgtggag	ctctttgtgg	actgtgttgc	1440
cgtggcccat	gtggatgatg	ttctgagct	tgcctctgc	cccgatggga	agggttccg	1500
gatgtctcc	ggcaggctctg	gggcgttctt	caagcttcc	caggaaaagc	agaagtgtgg	1560
ccacgggagg	gccctctgt	tccagggggt	tgttgcgtat	gagcagggtca	agaccatcc	1620
catcaaccag	gtgtctcca	ataaaagac	catcaactac	aataagtgg	tgcagagctg	1680
catcgactgg	aaccgtgg	tgcgtaaagc	ggagctggcc	ctggcaagat	gtgacatcat	1740
tgacatccca	cagcttca	agacccggag	aaaaaaagca	acggccctt	tccctgactt	1800
ggtgaacatg	ctgggtctgg	ggaaacactt	gggcatacc	aaggccctt	ggcccatcat	1860
caatggctgc	tgctgcctgg	aggagaaggt	ggggctcctg	ctggagccgc	tgggcttcca	1920
ctgcaccc	atgtgact	tcactccca	ccacatgtc	catggggagg	tgcactgtgg	1980
caccaatgt	tgcagaaagc	ccttcttctt	caagtgggtg	aacatggtc	cctgagacag	2040
ctcccccacca	ccatctgtc	ccccctgggc	gggcattggc	ccaggtgttgc	gagacagaga	2100
caggccccctg	aacgataagc	accaggagac	cccaaggctc	catatggaa	actgggggt	2160
accgtccctc	tcaagaac	tttcccttga	agtgtccat	cttcaccc	aacccatgtg	2220
gttctcagac	tgttatctc	tcggggccccc	aaaagaagg	acccat	ttatagctc	2280
tccctgtatt	caacacaacc	catggatgt	tccccc	actcttcttc	catccatttgc	2340
gggacaaatc	cacattgggg	tctagaaaca	tccacgtatc	tcatcagcca	tcttgtctgc	2400
tgatctcaa	cagaggaaagg	atccatgtt	ctgttttgtt	ccaatttgc	cctctctgc	2460
gaggaacaaac	cataaaacca	gaccactca	cgcaggac	gcaggagaga	tttccctaa	2520
agcctccccc	ataaaaagg	agctgtggat	ccactttagat	caggcggaa	ccatcttca	2580
ccoggcgaag	ctccgtccca	gatgtgacc	ctcaccc	gtgagctgc	acatagtagg	2640
agttcttaga	tgcatgttgc	agcaatgaga	gttgcctt	agccttataa	actccccatg	2700
atctgatcatg	cagaatccca	gcctgttgc	gaatccctt	ggaatttctt	ggagacgaaa	2760
gtatctgggg	gattgttgg	taactgggg	acttgggtaca	agggtgaaa	gtatgtccca	2820
taatacacat	ggttgcactat	ggttgcac	cttgcgtatgg	ttaatattag	gtgtgtggag	2880
agggttgc	cattggccct	gggacttctc	tctgcaggag	gagagaacgc	tgcctctcct	2940
ctggattgtt	ctcaggcttct	ctgttggcc	tttgcacatc	ctgtctgtct	3000	
gcaggagagg	gggtcaaggg	gctggatcca	ccaaggcgc	tcacagcggg	aaaactctgg	3060
gaatgaacca	ctgaatccag	gggtatgggg	tggggggggc	gttctcgagg	tgtgtgc	3120
ctacacgtgt	gttctgtatg	ggtccagtc	cgtttccatc	actcgtaat	aatcaacag	3180
aaacacacaa						3189

SEQ ID NO: 18	moltype = AA	length = 664
FEATURE	Location/Qualifiers	
REGION	1..664	
note = protein-arginine deiminase type-3 [Homo sapiens]		
source	1..664	
mol_type = protein		
organism = Homo sapiens		
SEQUENCE: 18		
MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTGVDI YISPNMERGR	60	
ERADTRRWRF DATLEIIVVM NSPSNDLNS HVQISYHSSH EPLPLAYAVL YLTCVDISLD	120	
CDLNCEGRQD RNFVDRKQWV WGPSCYGGIL LVNCDRDDPS CDVQDNCDQH VHCLQDLED	180	
SVMLVLTQGP AALFDDDHLKV LHTSSYDAKR AQVFHICGPE DVCEAYRHVL QDKVSYEVP	240	
RHGDEERFFF VEGLSPFDAG FTGLISFHTV LLDDSNEDFS ASPIFTDTVV FRVAPWIMTP	300	
STLPPLVEVY CRVRNNTCFV DAVAELARKA GCKLTICPQA ENRNDRWIQD EMELGYVQAP	360	
HITLPVVFDS PRNGELQDFP YKRILGPDFG YTREPRDRS VSGLDSFGNL EVSPVVANG	420	
KEYPLGRILI GGNLPGSSGR RVTQVVRDFL HAQKVQPPVE LFWDWLAVGH VDEFLSFVPA	480	
PDGKGFRMILL ASPGACFKLF QEOKQCGHGR ALLFQGVVDD EQVKTISINQ VLSNKDLINY	540	
NKFVQSCIDW NREVLKRELG LAECDIIDIP QLFKTERKKA TAFFPDVN MVLGKHLGIP	600	
KPGTIINGC CCLEEKVRSL LEPLGLHCTF IDDFTPYHML HGEVHCGTNV CRKPFSFKWW	660	
NMVP	664	

SEQ ID NO: 19	moltype = DNA	length = 3189
FEATURE	Location/Qualifiers	
misc_feature	1..3189	
note = Homo sapiens peptidyl arginine deiminase 3		
(PADI3) transcript, mRNA		
source	1..3189	
mol_type = unassigned DNA		
organism = Homo sapiens		
SEQUENCE: 19		
agtgttgggg ttggcgccca cagctaagtc caacaccaggc atgtcgctgc agagaatcgt	60	
gcgtgttcc ctggagcatc ccaccaggcg ggtgtgtgt gctggcgtgg agaccctcg	120	
ggacatttat gggtcagtgc ctgaggggcac agaaatgtt gaggtctatg ggacgcctgg	180	
cgtggacatc tacatctctc ccaacatgg aaaaaaaaaaaaaacggcgtgc acaccaggcg	240	
gtggcgcttt gacgcgactt tggagatcat cgtggatcat aactccccca gcaatgcac	300	
caacgcacagc catgttccatca ctccagccat gaggccatc ccctggctca	360	

-continued

tgccgtgctc	tacctcacct	gtgttgcacat	ctctctggat	tgccgacactga	actgtgggg	420
aaggcaggac	aggaacttgc	tagacaagcg	gcagtgggtc	tggggccca	gtgggtatgg	480
cggcatcttgc	ctggtaact	gtgaccgtga	tgatccggc	tgtgatgtcc	aggacaatttgc	540
tgaccacoc	gtgcactgcc	tgcaagacct	gaaagacatg	tctgtcatgg	tcctcgccgac	600
gcaggcccct	gcagccctct	ttgatgacca	caaacttgc	ctccatactt	ccagctatga	660
tgccaaacgg	gcacaggctc	tccacatctg	cggccctgag	gatgtgtgt	aggctatag	720
gcatgtgtcg	ggccaagata	aggtgtctca	tgagggtacc	cggtgcgt	gggtggagga	780
gegcttccttc	gtggaaaggcc	tgtcttccc	tgatggccgc	ttcacaggac	tcatctcttgc	840
ccatgtca	ctgtggacg	actccaacga	ggatttctcg	gatccccc	tcttcactga	900
caactgtgtgt	ttccgtgtgc	caccctggat	catgacgccc	agcaacttgc	cacccttaga	960
gtgtgtatgt	tgccgtgtga	ggaacaacac	gtgttttgttgc	gtgggtgg	cagacttgc	1020
caggaaaggcc	ggctcaagc	tgaccatctg	ccccacaggcc	gagaacccca	acgaccgtgc	1080
gatccaggat	gagatggagc	tgggtacgt	tcaggccgc	cacaagaccc	tccctgggttgc	1140
cttgcactcc	ccaaggaaatg	ggggaaactgc	ggatttccct	tacaaaaggaa	tcctgggttgc	1200
agattttgggt	tacgtgactc	ggggaaacccg	cgacaggctc	tgagtgccc	tggactcttgc	1260
tgccaaacccgt	gaggtcagcc	ctccatgttgc	ggccaaatgg	aaagactacc	ccctggggag	1320
gatccctatt	ggggcaacc	tgcctgggtc	aagtggccgc	agggtcaccc	agggtggcgc	1380
ggacttcctc	catgcccaga	aggtgcagcc	ccccgtggag	ctcttttgttgc	actgtgttgc	1440
cgtggcccat	ttggatgagt	ttctgagcttgc	ttgtccctgc	ccccatggga	agggtttccgc	1500
gatgctcttc	gcccggccctg	ggggctgttgc	caagcttc	caggaaaaggc	agaagtgtgg	1560
ccacggggagg	gcccctctgt	tccagggggt	tgttgcgtat	gagcagggtca	agaccatctc	1620
catcaacccag	gtgtcttcca	ataaaagacct	catcaactac	aataagtgttgc	tgcaagactgc	1680
catcgactgg	aaccgtgagg	tgctgttgc	ggagctgggc	ctggcagactgt	gtgacatcat	1740
tgacatccca	cagcttca	agacccggag	aaaaaaaggca	acggcccttc	tccctgacttgc	1800
ggtaaacatg	ctgggtgtgg	ggaagcacct	gggcateccc	aaggcccttgc	ggcccatcat	1860
caatggctgc	tgctgtctgg	aggagaagg	gcaacttgc	ctggagccgc	tggcccttcca	1920
ctgcacccatt	attgtgact	tcactccata	ccacatgttgc	catggggagg	tgcaactgttgc	1980
caccaatgt	tgcagaaagc	ccttcttctt	caagtgggttgc	aacatgggtgc	cctgagacag	2040
ctcccaaccca	ccatcttgc	ccccctgggc	gggcatttgc	ccagggtgg	gagacagagaa	2100
caggccccctg	aacgataagc	accaagagagc	ccaaaggcttc	catatggaaac	actgggtgt	2160
accgtccctc	tcaagaaac	tttccctgttgc	atgttccatg	cctcaccatgc	aaccatgttgc	2220
gttctcagac	ttgaatcttc	tcgggggggg	atgttgcgttgc	ttatagcttc	ttatagcttc	2280
tccctgttatt	caacacaacc	catggagatg	tccccttctc	actctgttgc	catccatttg	2340
gggacaaatc	cacattgggg	tctagaaaca	tccacgtatc	tcatagcc	tcttgtcttgc	2400
tgatcttcaa	cagagaaagg	atccatgttgc	ctgcttttgttgc	ccaatttgc	cctctctgc	2460
gaggaacaaac	cctaaaacca	gaccatcca	cgccaggac	cgaggagaga	ttcttcctaa	2520
agctcccccc	ataaaaagg	agctgtggat	ccatgttgc	caggccggaa	ccatcttca	2580
cccgcccaag	ctccgtccca	gatgttgc	ctcaccatgc	gtgagctgtc	acatagtagg	2640
agctttaga	tgcatgttgc	agcaactgttgc	gttgccttataa	actccccatg	2700	
atctgacatg	cagaatatcca	gccttgc	gaatcccttc	ggatatttttgc	ggagacgaaa	2760
gtatctgggg	gattgttggg	taactggag	actgggtaca	agggttggaaa	gtatgttccca	2820
taatacacat	ggttgtactat	ggtgcgttgc	tttgtgttgc	ttatatttgc	gtgtcttgc	2880
aaagtgttgc	cattggccctt	gggacttgc	tctgcaggag	gagagaacgc	tgccctcttgc	2940
ctggatttgtt	ctcaggcttc	ctgttgc	tttgttgc	ttccacatc	ctgtctgt	3000
gcaggagg	gggctaagg	gctgttgc	ccaaggccgc	tcacagccgg	aaaactctgg	3060
gaatgaacca	ctgaatttgc	gggatgggg	tggggggcc	tttctcg	ttgtgttgc	3120
ctacacgtgt	gttctgtat	gttccatgt	cgtttccatc	actcgctaat	aatcaacac	3180
aaacacacaa						3189

SEQ ID NO: 20 moltype = AA length = 664
 FEATURE Location/Qualifiers
 REGION 1..664
 note = protein-arginine deiminase type-3 [Homo sapiens]
 source 1..664
 mol_type = protein
 organism = Homo sapiens
 SEQUENCE: 20
 MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTGVDI YISPNMERGR 60
 ERADTRRWRF DATLEIIVVM NSPSNDLNDHS HVQISYHSSH EPLPLAYAVL YLTCVDISLD 120
 CDLNCEGRQD RNFVDRQWV WGPSGYGGIL NVMCDRDDPNS CDVQDNCDQH VHCLQDLED 180
 SVMVLRTQGP AALFDDHKLV LHTSSYDAKR AQVFHICGPE DVCEAYRHVL QGDKVSYEV 240
 RLHGDEERFF VEGLSFPDAG FTGLISPHVT LLDDSNEDFS ASPIFTDTVV FRVAPWIMTP 300
 STLPPLLEVYY CRVRNNTCFV DAVAELARKA GCKLTICPQA ENRNDRWIQD EMELGYVQAP 360
 HKTLPVVFDS PRNGELQDFP YKRILGPDFG YVTREPRDRS VSGLDSFGNL EVSPPVVANG 420
 KEYPLGRILII GGNLPGSSGR RVTQVVRDFL HAQKVQPPVVE LFVWDLAVGH VDEFFLSFVPA 480
 PDGKGFRMLL ASPGACFKLF QEOKKCGHGR ALLFQGVVDD EQVKTISINQ VLSNKDLINY 540
 NKFVQSCIDW NREVLKRELG LAECDIIDIP QLFKTERKKA TAFFPDLVNM LVLGKHLGIP 600
 KPPGPIINGC CCLEEKVQSL LEPLGLHCTF IDDFTPYHML HGEVHCGTNV CRKPFSFKWW 660
 NMVP 664

SEQ ID NO: 21 moltype = DNA length = 4800
 FEATURE Location/Qualifiers
 misc_feature 1..4800
 note = PREDICTED: Homo sapiens peptidyl arginine deiminase
 3 (PADI3), transcript variant X1, mRNA
 source 1..4800

-continued

mol_type = unassigned DNA
 organism = Homo sapiens
 SEQUENCE: 21
 ttcctcggt tctgcattt cccctgtggg gggctgccc gggccaaaca gccagagg 60
 ctccaccagg cctggaaatt acaggottt ggagtagggt tgccagataa aataaaatg 120
 gcctagttt aattgaattt caggtaaata attaataatt tttagtagtat aagtgtatcc 180
 caaatcttc acggggatca caacaatga ttgactgtt acttggaaat caatgtccaa 240
 tgatcttcgt gtatccccat ttgctaaatc tggcggccct ctggtaggagg agaaacagg 300
 tgctctgtcc ctcccaactt gtggagacac ctgacagggg cccagacacc agatgtggac 360
 ctgcctgtat gagagaactt tgggcttggt ttatccctt ttccataaaa aagctgttt 420
 gagatgtaa agaatcatgt tggtggaaa gtgaaggagg ttgttgggg gaggtctgg 480
 ttactctgg ggaagatggg aacgggggg agtggggcagc tggggaaat tgggttagtc 540
 ccctactgtt cctggccctt agttccccat ttgttaaaca ggatatttcttccatcttc 600
 gggctgcagt gaggatcaga tgagctaaatg tacaaaaagg cttaggcagg cttggcgtga 660
 taaggatgtca gaggctgggg gctttttatg ttatgttattt gtgttattt ttatctcttc 720
 ggaggctgggt ggcacagatc tggtactggg agccacatgg ctggggttgg atccggccc 780
 ggcacacagc ctggggcaat ttacttaacc tcgggtttt catgttata atggggacaca 840
 ataatacgata gtgcctccct tcttagggta ttgagggat tgcatgtt gctggcacac 900
 aycatgacgt cagtaatgtt cagtcactgt tateccccat gggcatgtt gccgtcttcc 960
 ccaggaggcc ctcccaacttgc ctggggccctt ccctcatgtt ggttccactt gactgtgtat 1020
 ctgtggccctt tgagccccat ttggctggcc ttggacaatgt tggtcttcgt ctccggcagg 1080
 ggttcaacacat tggggggggc agtggggggc tctggccggc aggacgttcc tgagtgcaca 1140
 gtggggccccc cacagtcaaa ggttgcatac tctgggttcc aaggagggac cccaggctgg 1200
 gtggatggggc gggaaatggca tcttcaggcc agataacatc acatgttccat taaaactgtca 1260
 ataattttctt ggttgggttca ccaatccgg ccaatgttcc cgggttgcacat tgcaagatgt 1320
 tggaaacacca tggggggggc tggtggactt ttgcacatctt ggttgcgttca tggtgttca 1380
 cctcaggcgtt gtcacgggc ctcccttggc ctcaatgtt aatatggccaa atggggatga 1440
 tattgttcacat cttacacgtt gtaagatgtt agggggatgtc tgccggccagg tgccgttccat 1500
 gcaaggggcc ctcaacccatc ttgatccccc ttcccaggggg aagatcaagc gatattttcc 1560
 caccacttacat gggagatgtt gatggatgtt ggggttggaaa ttgttccctt taaaatcttc 1620
 agaactgtaa gggggccatcc ccaccaatgg ctgggtttca ctgagcttcc cgtggccaga 1680
 ctgcactggg ggattttatcg aycacatggc ttggcttcgtt ctccggaaat gtcggccatt 1740
 ctgggtcgatc cttcggggca cggaaatgtt tgagggttcat gggacgttcc tggttgcacat 1800
 ctatcatcttcc cccaatccatc agggggggcc gggatgttcc gacaccatgg cttggcggtt 1860
 tgacggcactt ttggagatca tctgtgttcat gaaatcccc agcaatgttcc tcaacggacag 1920
 ccatgttcaatc atttcttacc actccaggca tgagcccttgc cccttggctt atgggggtgt 1980
 ctatcttccatc tgggttgcata ttcttgcggc ttggggatgtt aactgttggg gaaggccagg 2040
 cggggaaatctt gtagacaaatgg ggcagtgggg ctggggggccctt agtgggttgc tgccgttccat 2100
 gctgggtacat tggggccatgg atactccgg ctgtgtatgtt cggggatgttcc tgacccatgg 2160
 ctgtgcactgc ctggaaacacc tggaaagacat gtcgttccatg tgccgttccatc cggggggcc 2220
 tgcggcccttctt tggatggggca aaaaatgtt ccctccatacc tccggatgttcc atggccaaatcc 2280
 ggcacaggcc ttccacatctt ggggttccat ggtatgttgc gggccatgttcc cggatgtgt 2340
 ggcccaatgtt aagggttccat atgggttcc cccgttgcat ggggttgcagg agccgttcc 2400
 ctgtggaaatcc ctgttccatcc ctgtatgttcc ctteacaggat ctcatcttcc tccatgttcc 2460
 tctgtggggca gactccaaatcc agggatccatc ggcatccctt atccatgttcc acatgttgc 2520
 ttcccgatgtt gcaatgttccatc tcatgttccatc cggatgttcc cccatccatgg aggtgtatgt 2580
 tgccctgtgtt aggaacaccaatc ctgtttttt ggtatgttgc gggatgttcc cggaggccatgg 2640
 cggctgcacat ctggaccatcc gcccacatggc cgagaaccgc aacggccgtt ggttccaggaa 2700
 tgagatgtggat ctggggatgtt ttccaggccgc gcacaatggcc ctccgggttgc tttttgttcc 2760
 cccaaatgttccatggggggactgtc agggatccccc ttacaaaaaaa atccctgggttgc cagatgtttcc 2820
 ttatcgatgtt cggggggccatccatgg gtcgttccatc cggatgttcc tccatgttcc tttttgttcc 2880
 ggaggttccatgg ccccttccatgg tggccatgtt gaaagatgttcc ccccttggggatccatgttcc 2940
 tggggggccatccatgg ctggctgggtt caagtggggc cgggggttccatc cgggggttgc gggacttcc 3000
 cccatgttccatggggggccatccatgg gtcgttccatc cgggggttgc gggacttccatggggggccatccatgg 3060
 tggggggccatccatggggggccatccatgg gtcgttccatc cgggggttgc gggacttccatggggggccatccatgg 3120
 gggccatccatggggggccatccatgg tcaagttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3180
 gggccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3240
 ggttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3300
 aataaaatgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3360
 aacatgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3420
 gtcgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3480
 ctgtgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3540
 cttatgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3600
 gtcgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3660
 gtcgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3720
 gtcgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3780
 ttcacatgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3840
 cttatgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3900
 ttcacatgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 3960
 cccatgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 4020
 acatgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 4080
 cccatgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 4140
 cccatgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 4200
 gtcgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 4260
 acatgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 4320
 gtcgttccatccatggggggccatccatgg ttcacatgttccatc ccaggaaaaatgg cgggggttccatc ggtatgttcc 4380

-continued

ggattgttgg	gtactaggga	gactgggtac	aagggtaaaa	agttagtccc	ataatacacaca	4440
tgttgacta	tgtgtatcca	ccttgtatgc	gttaatatta	ggtgtctgga	gaagggttgtct	4500
tcattggccc	tgggacttct	tcctgcaggaa	ggagagaacg	ctgcctctcc	tctggatgg	4560
tctcaggotc	tctgttggcc	tttggtcago	gtttccacat	cctgccttcgc	tgcaaggagag	4620
ggggctaagg	ggctggatcc	accaaggcgag	ctcacagcgg	aaaaactctg	ggaatgaacc	4680
actgaattca	ggggatgggg	gtgggggggg	ggttctcgag	gtgtgtccca	gctacacgtg	4740
tgttctgtat	gggtccagct	cgctttccat	cactcgctaa	taaatcaaca	gaaacacaaa	4800

SEQ ID NO: 22	moltype = AA	length = 626				
FEATURE	Location/Qualifiers					
REGION	1..626					
	note = protein-arginine deiminase type-3 isoform X1 [Homo sapiens]					
source	1..626					
	mol_type = protein					
	organism = Homo sapiens					
SEQUENCE: 22						
MFEEVYGTGPV	DIYISPNNMER	GRERADTRRW	RFDATLEIIV	VNNSPSNLDN	DHVQISYHS	60
SHEPLPLAYA	TLYLTCVDIS	LDCDLNCEGR	QDRNFVDKRQ	WWVGPSGYGG	ILLVNCDRDD	120
PSCDVQDNCD	QHVHCLQDLE	DMSVMVLRTQ	GPAALFDDHK	LVLHTSSYDA	KRAQVFHICG	180
PEDVCEAYRH	VLGQDKVSYE	VPRLHGDEER	FFVEGLSPFD	AGFTGLISFH	VTLLDDSNED	240
FSASPIPTDT	VVFRVAPWIM	TPSTLPPLEV	YVCRVRNNTC	FVDAVAELAR	KAGCKLTICP	300
QAENRNDRWI	QDEMELGYVQ	APHKTLPPVF	DSPRNGELOQD	FPYKRILGPD	FGYVTREPRD	360
RSVSGLDSFG	NLEVSPPVVA	NGKEVPLGRI	LIGGNLPGSS	GRVTVQVRD	FLHAQKVQPP	420
VELFVDWLAV	GHVDEFLSFV	PAPDGKGFRM	LLASPGACFK	LFQEKEKCGH	GRALLFQGVV	480
DEBQVKTISI	NQVLSNKDLI	YNKFPVQSCI	DWNREVLIKRE	LGLAECDIID	IPQLFKTERK	540
KATAFFPDLV	NMLVLGKHLG	IPKFPGPIIN	GCCCLEEKVR	SLLEPLGLHC	TFIDDFTPYH	600
MLHGEVHCGT	NVCRKPFSFK	WWNMVP				626

SEQ ID NO: 23	moltype = DNA	length = 4800
FEATURE	Location/Qualifiers	
misc_feature	1..4800	
	note = PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X1, mRNA	
source	1..4800	
	mol_type = unassigned DNA	
	organism = Homo sapiens	
SEQUENCE: 23		

tcccccgtgt	tctggcattt	ccccctgtggg	gggctgcccc	ggccaaaca	ggccagagg	60
ctccaccacg	cttggaaattt	acaggcttgc	ggagtaggg	tgccagataa	aataaaagat	120
gcctagttaa	aattgaattt	caggttaata	attaataat	tttttagtat	aagtgtatcc	180
caaataacttc	acggggactac	caaacaatga	ttgactgttc	acttgaattt	caagtccaa	240
ttagctttct	gtatTTTAT	tttgcataatgc	tggcaggccc	gcttttaggg	agaaaacagg	300
gtgtctgtcc	ctcccaactg	gtgaggacatc	ctgacagggg	cccagacacc	agagtgggac	360
ctgcctgtat	gagagaagct	tgggttgg	ttatTTCTCT	tttcataaaa	aagctgttt	420
gagatgtaa	agaatcatgt	tgggttggaa	gtgaagagg	tgtttttggg	gaggctgtcc	480
ttactctgtat	ggaatgtatgg	aaggccccggg	atggccacgg	tgtggaaat	tgggttagt	540
ccttaactgtat	ctctgtccctat	agttttccca	tgtgtaaaca	ggagtaata	tctcatctt	600
gggctgcagtat	gaggatcaga	tgagctaaatg	tacaaaaagg	ctttaggcagg	cctgcctgt	660
taagagtgc	gaggctgggg	gtttttatg	ttatggattt	ttgttattt	ttactcttgg	720
ggaggtgtgtat	ggcaagagtat	tggactgg	agccacatgc	ctgggtttgg	atcccacccc	780
ggccacacagat	cttggggcaat	ttacttaacc	tcggcttttcc	catctgtata	atgggagaca	840
ataatagata	gtgcctccat	tcttagggta	ttgagggat	tgcataatgg	gtcgccacac	900
agcatagatgtat	cgatgtatgt	cgttcactgt	tatccccca	gggcattgtat	gccgcgtcc	960
ccaggaggccc	ctccaaactcg	cgtggccatc	tccttacatgc	ggttcccaact	gagctgtgtat	1020
cctgtggccat	ttagccccat	ctgggtggcc	tggacatgt	tgttgcctcg	ctccccacgc	1080
ggatcaacatgc	tggggagatgc	agttgggtga	tgcgtccgc	aggcgttacc	tgatgtccaa	1140
gtggggccatcc	cacagtcaaa	ggatgtatgc	tctggtttcc	aaggaggaac	cccaggctag	1200
gttggatgtat	ggaaatgtcaat	tctcaggccc	agataatccat	acagatattt	taaacctgtca	1260
attaaattttgtat	gtttgttgg	ccaaatccgg	ccaaatccgg	cggtgtgacac	tgcagatgt	1320
tggaaacacatgc	tgggttgg	gtcagacatgt	ttgcacatct	ggctctgtca	tgtatgtca	1380
cctcaggatgtat	gtcactggatc	tccttgcacat	tcctgttgc	aatatgtccaa	atggggatgt	1440
tattgtctatc	tttacatgtat	gttggatgtgg	aggggatgtc	gtgcgtccagg	tgcctggcat	1500
gcacgtggccat	cttcaacatgc	tttccatccc	tttccatccc	agaatccatgc	gatatcttcc	1560
cacccactatgc	ggagacatgt	gtatggatgt	gggggtggaaa	ttgtctttt	taaagcttcc	1620
agaactgtat	ggggctccat	ccaccaatgg	ctgggtgtca	ctgagcttcc	ctgtccccaga	1680
cctgcactgg	ggatTTATCG	agcacatgttgc	tggcttctgt	cttccgaaaga	gctcgccatt	1740
ctgggtcagtat	cttggggatgc	cagaatgttt	tgaggatctat	gggacgcctg	gcgtggacat	1800
ctacgttctat	cccaacatgc	agagggccgg	ggagctgtgc	gacacaggcg	ggtggcgctt	1860
tgacgcgtat	ttggatgtat	tggtgtatcat	gaactcccc	agcaatgtatc	tcaacgcac	1920
cctatgttccat	atttccatcc	actccatgc	tgagcttcc	ccccctggct	atgcgtgtct	1980
ctaccttaccat	tgtgttgcata	tctctcttgc	ttgcacatct	aactgttgc	gaaggccat	2040
cagggactttat	gtagacaatgc	ggcagtgggt	ctggggggcc	agtgggtatg	gccccatctt	2100
gttgggtgtat	tgtgaccgtat	atgtccatgc	ctgtgtatgc	caggacaattt	gtgaccatgc	2160
cgtgcactgtat	ctgcaagatcc	tggaaagatct	gtctgtatgc	gtcctgcgg	cgccaggcccc	2220
tgcagccatccat	tttgcattatgc	acaaaacttgc	cctccatacc	tccagatgtat	atgcacaaac	2280

-continued

```

SEQ ID NO: 24          moltype = AA   length = 626
FEATURE                Location/Qualifiers
REGION                 1..626
note = protein-arginine deiminase type-3 isoform X1 [Homo
                      sapiens]
source                1..626
                     mol_type = protein
                     organism = Homo sapiens

SEQUENCE: 24
MFEEVYGTPTGV DIFYVSPNMR GRERADTRRW RFDATLLEIIV VMNSPSNDLN DSHVQISYHS 60
SHEPLPLAYA VLYLTCVDIS LDCCDLNCERG QDRNFVDRQ WWWGPSCGYGG ILLVNCDRDD 120
PSCDVQDNCD QHVHCLQDLD DMSVMVLRTQ GPAALFDDHK LVLHTSSYDA KRAQVFHICG 180
PEDVCEAYRH VLQGDKVSYE VPRLHGDEER FFVEGLSFPD AGFTGLISFH VTLLDDSNED 240
FSASPIFTDT VVFRVAPWIM TPSTLPLLEV YVCVRVNNTC FVDAVAELAR KAGKCLTICP 300
QAENRNDRWI QDEMELGVYQ APKTLPVVF DSPRNGELQD FPYKRILGPD FGYYTVERPRD 360
RSVSGLDSFG NLEVSPPVVA NGKEYPLGR LIIGGNLPGSS GRRVTQVVRD FLHQAKVQPP 420
VELFVWDLWAV GHVDEFLSFV PAPDGKGFRM LLASPAGCFK LFQEKQKCGH GRALLFQGVV 480
DDEQVKTISI NQVLSNKLII NYNKVFVQSCI DWNREVLKRE LGLAECIDIID IPQLFKTERK 540
KATAFFPDLV NMLVLGKHLG IPKPFGPPIIN GCCCLEEKVR SLLEPLGLHC TFIDDFTPYH 600
MLHGEVHCGT NVCRKPEFSFK WWNNMPV                                         626

```

```
SEQ ID NO: 25          moltype = DNA    length = 4800
FEATURE                  Location/Qualifiers
misc_feature             1..4800
                           note = PREDICTED: Homo sapiens peptidyl arginine deiminase
                           3 (PADI3), transcript variant X1, mRNA
source                   1..4800
                           mol_type = unassigned DNA
                           organism = Homo sapiens
SEQUENCE: 25
tcccccctgg tctggcattt cccctgtggg gggctggcc a gggcaaaca ggcacagagg 60
ctccaccaggc cttggaaattt acaggctttt ggatggatgt tgccgataaa aataaaagat 120
ggcttaatgg aattgtttttt caatgtttttt attaaatattt ttttttttttt aatgttatcc 180
```

- continued

-continued

```

tggctgtat gggccagct gcgttccat cactcgtaa taaatcaaca gaaacacaaa 4800
SEQ ID NO: 26 moltype = AA length = 626
FEATURE Location/Qualifiers
REGION 1..626
note = protein-arginine deiminase type-3 isoform X1 [Homo sapiens]
source 1..626
mol_type = protein
organism = Homo sapiens
SEQUENCE: 26
MFEEVYGTGIV DIYISPNMER GRERADTRRW RFDATLEIIV VMNSPSNDLN DSHVQISYHS 60
SHEPLPLAYA VLYHTCVDIS LDCCDLNCEGR QDRNFVDKRQ WWGPSPGYGG ILLVNCRDD 120
PSCDVQDNCD QHVHCLQDLE DMSVMVLRTQ GPAALFDDHK LVLHTSSYDA KRAQVFHICG 180
PEDVCEAYRH VLQDKVSYE PVRHLGSPFD AGFTGLISFH VTLLDDSNED 240
FSASPIFTDT VVFRVAPWIM TPSTLPPLEV YVCVRVNNTC DFWAVAEELAR KAGCKLTICP 300
QAENRNDRWI QDEMEGLYVQ APHKTLPVVF DSRNGELQD FPYKRILGPD FGVTREPRD 360
RSVSGLDSFG NLEVSPPVVA NGKEVPLGRI LIGGNLPGS SGRVTQVVRD FLHQAKVQPP 420
VELFVDWLAV GHVDEFSLFV PAPDGKGFRM LLASPGACFK LFQEKKQKCGH GRALLFQGVV 480
DBBQVKTISI NQVLSNKDLI NYNKFVQSCI DWNREVLIKRE LGLAECIDI IPQLPKTERK 540
KATAFFPDLV NMLVLGKHLG IPKPFPGPIIN GCCCLEEKVR SLLEPLGLHC TFIDDDFTPYH 600
MLHGEVHCGT NVCRKPFSFK WWNNMVP 626

SEQ ID NO: 27 moltype = DNA length = 4800
FEATURE Location/Qualifiers
misc_feature 1..4800
note = PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X1, mRNA
source 1..4800
mol_type = unassigned DNA
organism = Homo sapiens
SEQUENCE: 27
tcctccctgg tctgccattt cccctgtggg gggctgccc gggccaaaca ggccagaggg 60
ctccaccagg cctggaaattt acaggctttt ggagtagggg tgccagataa aataaaagat 120
gccttagttt aattgttaattt caggtaataattt attaataat tttagtagt aagtgtatcc 180
caaataacttc acggggatc caaaatcta ttgactgttc acttggaaattt caagtccaaac 240
tgagctttct gtatTTTAT ttgtcaagtc ttggcagggc gctttagggg agaaaacagg 300
gtgtctgtcc ctcccaactg gtgaggcacac ctgacaggggg cccagacacc agagtggac 360
ctgcctgtat gagaggaaatc tgggttggg ttatTTTCTT ttccataaaa aagctgtttt 420
gagatgttaag agaatcatgt ttgttggaga atgaaagggg tggtttgggg gaggtctgccc 480
ttactctgtat ggaatgtggg aagccccgggg agtggccacg tttggaaattt tggggtagtc 540
ccctacgtat cctgtgcctc agtttccca tttgttaaca ggagtaattt tctcatctg 600
gggctgcaggat gaggatcaga tgagcttaatc tttaaaaagg cttagggcagg cctgcgtgt 660
taagagtgtca gaggtttggg gctttttat ttatgttattt ttgttattttt ttactctgg 720
ggagggttgtt ggcaagatgt tggactgagg agccacatgg ctgggtttgg atccccggcc 780
ggccacacag ctgggcaag ttacttaacc tcggctttt catctgtata atgggagaca 840
ataatagata gtggccccc tcttaggtt tttggaggat tttttttttt tttttttttt 900
agcatgcgtat ctagtgcgtt cttttttttt tttttttttt tttttttttt tttttttttt 960
ccaggaggccc ctcccaactcg cttttttttt tttttttttt tttttttttt tttttttttt 1020
cctgtggccct tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1080
ggatcaacag tggggggaggc agtgggttgg tttttttttt tttttttttt tttttttttt 1140
gtggggccccc cacatgtttt tttttttttt tttttttttt tttttttttt tttttttttt 1200
gttggatgtat gggaaatgtttt tttttttttt tttttttttt tttttttttt tttttttttt 1260
attaattttt gttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1320
tggaaacagca tgggggtttttt tttttttttt tttttttttt tttttttttt tttttttttt 1380
cctcaggccct tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1440
tattgtctatc tttagatgtt tttttttttt tttttttttt tttttttttt tttttttttt 1500
gcagcggggca cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1560
cccccaactaa ggagatgtat tttttttttt tttttttttt tttttttttt tttttttttt 1620
agaactgtttt gggggccccc ccacatgtt tttttttttt tttttttttt tttttttttt 1680
cctgcactgtt ggatTTTCTT tttttttttt tttttttttt tttttttttt tttttttttt 1740
ctgggtcaggat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1800
ctacatctctt cccaaatgtt tttttttttt tttttttttt tttttttttt tttttttttt 1860
tggatgtatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1920
ccatgttccatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1980
ctacacttacc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2040
cagggactttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2100
gctgggttggat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2160
catgcactgc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2220
tgcagccctc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2280
ggcacacatgtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2340
ggccaatgtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2400
cgtggaaatgtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2460
tctgtggac gactccaaatc tttttttttt tttttttttt tttttttttt tttttttttt 2520
gttccggatgtt gcaccctggat tttttttttt tttttttttt tttttttttt tttttttttt 2580
gtggccgtgtt aggaacaaca cttttttttt tttttttttt tttttttttt tttttttttt 2640

```

-continued

cggtctcaag	ctgaccatct	gccccacaggc	cgagaaccgc	aacgaccgc	ggatccagga	2700
ttagatggaa	ctgggctaag	ttcaggccgc	gcacaagacc	cctccgggttgc	tctttgactc	2760
cccaaggaaat	ggggaaactgc	aggattttcc	ttacaaaaga	atcccgggttgc	catagttttgg	2820
ttaatcgact	cgggaaacacc	gcccggatgttgc	tgtggatgttgc	tctggacttc	ttggggaaacct	2880
ggagggtcage	cctccatgttgc	tggccaaatgg	gaaagaggtac	ccccctgggg	ggatccatcat	2940
tggggcaac	ctgcctgggt	caagtggccg	cagggttcacc	cagggtggcc	gggacttcctc	3000
ccatggccca	aagggtgcage	ccccctgttgc	gtctttgttgc	gactgggttgc	ccgtggggcc	3060
tgtggatgttgc	ttttcgatgttgc	ttgttccctgc	ccccatgttgc	aaagggttc	ggatgttcct	3120
ggccagccct	ggggcctgttgc	tcaagcttt	ccagggaaaag	cagaagtgttgc	gccacgggg	3180
ggccatcttc	ttcccaagggg	ttgttgatgttgc	tgagccatgttgc	aaggatccat	ccataccaat	3240
gggtgttcctc	aataaaagatc	tcatcaacta	caataatgttgc	gtgcagatgttgc	gtatcgacttgc	3300
gaaccgttgc	gtgttcgttgc	ggggagtttttt	ccttggccatgttgc	tttgatcatca	tttgatcatcc	3360
acagcttttc	aagaccgaga	ggaaaaaaagc	aacggccatcc	tccctgtact	ttggtaacat	3420
gtgtgtgttgc	ggggaaagcacc	ttggccatcc	caagcccttgc	ggggccatca	tcaatggctg	3480
ctgtgtccgt	gaggagaatgttgc	ttgggttccttgc	gttggacccgc	ttggggcttc	actgcacccat	3540
cattatgttgc	ttccatccat	accacatgttgc	gtatggggat	gtctgttgc	gtaccaatgttgc	3600
gtgcagaaat	cccttttcttgc	tcaagtgggttgc	gaacatgttgc	cccttgagaca	gtctccaccc	3660
accatcttgttgc	ccccctgggg	cgggccatgg	cccaggttgc	ggagacagag	acaggcccccc	3720
gaacgataaaat	caccaagatgttgc	ccccaaatgttgc	ccatgttgc	cacttgaggat	gaccgttccct	3780
cttcggaaatgttgc	ttttcccttgc	aagtgttccat	gccttcaccc	caacccatgttgc	ttttcccttgc	3840
cttgaatcttgc	ctcgcccccc	caaaaagaaatgttgc	gacccatatttgc	tttatagttgc	ctccgtgtat	3900
tcaacacacaaat	ccatggatgttgc	gttcccttgc	cactctgtaaaatgttgc	tccatccat	ttttccatgttgc	3960
ccacattttgg	gtcttgcataatgttgc	atccatgttgc	ctatctgc	atcttgc	ttttccatgttgc	4020
acagaggaaatgttgc	gatccatgttgc	tctgttgc	tccaaatgttgc	ttttccatgttgc	ttttccatgttgc	4080
ccctaaatccatgttgc	agaccacttgc	acgcggatgttgc	ggcaggagag	attttccatgttgc	ttttccatgttgc	4140
cataaaaaatgttgc	gagctgtgttgc	tccacttgc	ttccggggat	accatcttc	ttttccatgttgc	4200
gtcttcgtccatgttgc	agatgttgc	ccttcacccatgttgc	gttggatgttgc	ttttccatgttgc	ttttccatgttgc	4260
atgtatgttgc	aagaatgttgc	agttgttccatgttgc	tccatgttgc	ttttccatgttgc	ttttccatgttgc	4320
gcagaaaatccatgttgc	agccatgttgc	agaatccatgttgc	ttggaaatgttgc	ttttccatgttgc	ttttccatgttgc	4380
ggatgttgc	gtacttagggatgttgc	gactttgttgc	aaagggttgc	ttttccatgttgc	ttttccatgttgc	4440
tgggtgttgc	tgggtgttgc	ttttccatgttgc	ttttccatgttgc	ttttccatgttgc	ttttccatgttgc	4500
tccatgttgc	ttttccatgttgc	ttttccatgttgc	ttttccatgttgc	ttttccatgttgc	ttttccatgttgc	4560
tccaggatgttgc	ttttccatgttgc	ttttccatgttgc	ttttccatgttgc	ttttccatgttgc	ttttccatgttgc	4620
ggggcttgcataatgttgc	gggttgcatccatgttgc	accatgttgc	cttcacccatgttgc	ttttccatgttgc	ttttccatgttgc	4680
actgtatgttgc	ggggatgttgc	gtggggggatgttgc	ttttccatgttgc	ttttccatgttgc	ttttccatgttgc	4740
tgttctgtatgttgc	gggttgcacgttgc	ttttccatgttgc	cactctgttgc	ttttccatgttgc	ttttccatgttgc	4800

```
SEQ ID NO: 28          moltype = AA    length = 626
FEATURE                Location/Qualifiers
REGION                 1..626
note = protein-arginine deiminase type-3 isoform X1 [Homo
                      sapiens]
source                 1..626
                      mol_type = protein
                      organism = Homo sapiens
```

SEQUENCE :	28	organism = Homo sapiens				
MPEVYGTGPMV	DIYISPNNMER	GRERADTRRW	RFDATLEIIV	VMNSPSNDLN	DSHVQISYHS	60
SHEPLPLAYA	VLYLTCVDIS	LDCDLNCEGR	QDRNRNFVDKRQ	WWWGPSPGYGG	ILLVNCDRDD	120
PSCDVQNDQC	QHMHCLQDLE	DMSWVMLRTQ	GPAALFDDHK	LVLHTSSYYDA	KRAQVFHICG	180
PEDVCEAYRH	VLGQDKVSYE	VPRLHGDER	FFVEGLSFPD	AGFTGLISFH	VTLLDDSNED	240
FSASPIFTDT	VVFRVAPWIM	TPSTLPPLEV	YXCRVRNNTC	FVDVAELAR	KAGKCLTICP	300
QAENRNRDWI	QDEMELGYVQ	APHKTLPVVF	DSPRNGLEQD	FPYKRILGPD	FGYVTREPRD	360
RSVSGLDSFG	NLEVSPPVVNA	NGKEYPLGRJ	LIGGNLPGSS	GRRVTQVQRD	FLHQAKVQQP	420
VELFVWDLAV	GHVDFELSFV	PAPDGKGFRM	LLASPAGCFK	LFQEKKQKCGH	GRALLFQGVV	480
DDEQVKTISI	NQVLSNKLID	YNKFKVQSCI	DWNREVLIKRE	LGLAECIDIID	IPQLFKTERK	540
KATAFFPDVL	NMLVKGKHLI	IPKPFCGTTIN	GCCCLEEKVR	SLLEPLGLHC	TFIDDDFTPYH	600
MLHGVEHCGT	NVCRKPFSEK	WWNNMVP				626

```
SEQ ID NO: 29          moltype = DNA    length = 4800
FEATURE                  Location/Qualifiers
misc_feature            1..4800
note = PREDICTED: Homo sapiens peptidyl arginine deiminase
                     3 (PADI3), transcript variant X1, mRNA
source                  1..4800
                      mol_type = unassigned DNA
                      organism = Homo sapiens
```

SEQUENCE: 29 Organism = HOMO sapiens

tcccccgtt	tctggcattt	ccccctgtggg	gggctgccca	gggc当地aaaca	ggccagaggg	60
ctccaccaggc	cctggaaattt	acaggcttt	ggagtaggtt	tgccagatata	aataaaaatgt	120
gectagttaa	aatttgaattt	caggtaataa	attaataattt	tttttagtat	aagtgtatcc	180
caataatcc	acggggatgtac	ccaaaatgtat	ttgtactgtt	acttgtaaaat	caagtc当地ac	240
tgagctttct	gtatttttat	ttgctaaatgc	tggcagccct	gctttaggggg	agaaaacagggt	300
gtgtctgtcc	ctcccaactgt	gtgaggacac	ctgacagggg	ccc当地acac	agagttgggac	360
ctggctgtatg	gagaatgtatgt	ttggctttgtt	ttatccctt	tttc当地aaaa	aagctgtttt	420
gagatgtatgt	agaatcatgtatgt	ttgtggaaatgt	gtaaaggggt	tgtttttttttt	gaggctgtcc	480
ttaatctgtatgt	ggaaatgtatgt	aaaggcccccc	atgtggccatgt	tgtggaaattt	ttgggttagtc	540

-continued

ccctacgtat	cctgtgcctc	agtttccccca	tgtgtaaaaca	ggagtaat	tctcatcctg	600
gggctgcagt	gaggatcaga	tgagatcgatg	tacaaaaagg	cttagggcagg	cctgcgcgtga	660
taagagtgc	gagggtggg	getttttat	ttatggatt	gttgttata	ttactccctgg	720
ggaggtggt	ggcaagagt	tggactgagg	agccacatgg	ctgggtttgg	atcccagccc	780
ggccacagac	cttggcaag	ttacttaacc	tcggcttct	catctgtata	atgggagaca	840
ataatagata	gtgcctccct	tcttagggta	tttagagggat	tgcatgtatgg	gtggcacac	900
agcgtgac	cagaatgt	cagtcactgt	tatccccat	gggcgtatgc	ggcgcgtgcct	960
ccaggagggc	ctccaactcg	cagtgccctc	tcctatcat	ggttcccaact	gagctgtgtat	1020
cctgtggct	tgagccccat	ctggctgccc	tggacaatgt	ttgtctctcg	ctccccagcg	1080
gtatcaac	tggggagagc	agtgggttgc	tctggccccc	aggcqatacc	tgagtgcaca	1140
gtgggcaccc	cacagtc	ggatgcata	tctggtttc	aaggaggaaac	cccaggctag	1200
gtggaatgag	ggaaatca	tctcaggccc	agataacatc	acagatactt	taaacctgc	1260
attaattct	ggttgttga	ccaaatccgg	ccaagtttcc	cgggtgacac	tgcagagctg	1320
tggAACAGCA	ttgggttggg	gtcagaccc	ttgcacatct	ggctctgc	tgtgtatcaa	1380
cctcaggctt	gtcacgggc	ctcactgac	tcagttgtc	aatatgc	atggggatga	1440
tattgctac	tttacagctt	gtaaagatgg	aggagatc	gtgcgcagg	tgccctggcat	1500
gcagcgggca	ctcaaccaga	ttgaattccc	ttcccaaggga	agaattcaag	gatatctcc	1560
cacccactaa	ggagacat	gtatggatc	gggggtgagaa	ttgtctttt	taaagctctc	1620
agaactgtaa	ggggctcc	ccaccaatgg	ctgggttgc	ctgagctt	cgtgcgcagg	1680
cctgcacttg	ggattatcg	agcacatgg	ttggcttgc	tttccgaaga	gtcgccatt	1740
ctgttcgtgt	cttgaggggc	cagaatgtt	tgaggttat	gggacgcctg	gcgtggacat	1800
ctacatctct	cccaacatgg	agagggggcc	ggagcgtgc	gacacaggc	ggtggcgctt	1860
tgacgcgact	ttggatgatc	tcgtgttcat	gaactcccc	agcaatggc	tcaacgacag	1920
cctatgttcc	atccatcc	ttccatggca	ttggcttgc	ccctggct	atgcgggtct	1980
ctacctcacc	tgtgttgc	tctctctgg	ttgcgacctg	aactgttgg	gaaggcagg	2040
cagaacttt	gtagaca	ggcagtggg	ctggggccc	agtgggtatg	gcccacattt	2100
gtctggtaac	tggtggctg	atgtccgg	ctgtgtatgc	caggacaaat	tgaccacgca	2160
cgtgcactgc	ctgcaagacc	ttggaa	gtctgtatgc	gtcctgggg	cgcaggggccc	2220
tgcagccctc	tttgcgtacc	acaaaactgt	cctccatacc	tccagctat	atgcacaaacg	2280
ggcacaggtc	ttccacatct	gcccgtctg	ggatgtgtgt	gaggctata	ggcatgtgt	2340
ggggcaagat	aagggttctt	atggatc	ccgcgttgc	ggggatgagg	agcgcttctt	2400
cgtggaaaggc	ctgtcttcc	ctgtggccg	tttcacagg	ttcatctct	tccatgtcac	2460
tctgctggc	gactccaa	aggatttctc	ggcatccct	atcttactg	acactgtgt	2520
gttcccgagt	gtaccctgg	tcatgacgc	cagcaactg	ccacccct	agggttatgt	2580
gtgccgtgt	aggaaacaca	cgtgttttgc	ggatgcgg	gcagagctgg	ccagggaggc	2640
cggctgc	ctccatct	gtcaccatct	gcccacatgg	cgagaaaccgc	aaacgcgtt	2700
tgcacatgg	ctgggttac	ttcaggcgc	gcacaagacc	ctcccggtt	tctttactc	2760
ccaaaggaaat	ggggaaactgc	aggatttccc	ttacaaaaga	atccctgg	cagattttgg	2820
ttacgtgact	cgggaaacc	gcccacatgg	tggtgatgc	ctggactct	ttggggac	2880
gggggttcc	ctccatct	ttggatgatc	gaaagatgtt	ccctgggg	ggatccat	2940
tggggggcaac	ctgcctgggt	caatggcc	cagggtgtc	ggggacttct	3000	
ccatggcc	aaagggtc	ccccctgtgg	gtctttgt	gactgttgg	ccgtggggca	3060
tgtggatgag	tttctgtat	ttgtccctg	ccccgtatgg	aagggtctt	ggatgtctt	3120
ggccagccct	ttggccctgt	tcataatctt	ccaggaaa	cagaatgtgt	gccacgggag	3180
ggccctct	ttccaggggg	ttgttgatg	tgacgcgtt	aagacatct	ccatcaacca	3240
gggtgtctcc	aaataagacc	tcatcaacta	caataatgtt	gtcagatgt	gcatgcact	3300
gaaccgtgt	gtgtgtc	ggggacttgc	cttgcacatgg	tgtgacatca	ttgacatccc	3360
acatcttc	aaagaccgaga	ggggaaactgc	aaagggttcc	ttccctgt	tggtaacat	3420
gtctgggtgt	ggggaaaccc	ttggcatccc	caagccctt	gggcccata	tcaatggct	3480
ctgtgtct	gaggagaagg	tgccgttcc	gtggagcc	ctggccttcc	actgcac	3540
attgtgtac	ttcaatccat	accacatgt	gtatgggg	gtgcactgt	gcaccaatgt	3600
gtcgacatgt	cccttctt	tcataatgtt	ggacatgtt	ccctgttgc	ggggacttcc	3660
accatctgt	ccccctgggg	cgggcatgg	ccccatgtt	ggagacagag	acaggccct	3720
gaacgataag	caccaagaga	ccccaaagg	ccagatgg	cactgggg	gaccgtccct	3780
ctcagaaccc	ttttccctgg	aatgttccat	gtcttcac	caacccatgt	ggttctcaga	3840
cttgaatctt	ctcggttcc	caaaaagaa	gacccat	ttttagatct	ctctgtgt	3900
tcaacacaac	ccatggat	gtccccctt	cactctaa	tcatcattt	ggggacaat	3960
ccacattggg	gtctagaa	atccacgt	tcatcagcc	atcttgt	gtgcac	4020
acagaggaa	gatccatgt	tctgtttgg	tccaaatgt	tcctcttc	agaggaacaa	4080
ccctaaacc	agaccactcc	acgcaggac	ggcaggag	atttccat	aaagccccc	4140
cataaaaaagg	gagctgtgg	tccacttgc	tcaggcg	accatctt	acccggccaa	4200
gctccctggcc	agatgttgc	cctcaccat	cgatgtgt	catatgt	gagttctt	4260
atgcatgtgg	aagcaatgag	agttgtcc	tagccttata	aactccat	gatctgtac	4320
gcagaaaatcc	acccttgtcc	tgaaatitc	tggagacgaa	agtatctgg	4380	
ggattgttgg	gtacttaggg	gactgggtt	aaagggttac	agtagttcc	ataatacaca	4440
tgtgttacta	ttgtgtatcc	ccttgtatgc	tttataat	ggtgtgttgc	4500	
tcattggccc	ttggacttct	ctctgtatgc	ggagagaac	ctgccttct	tctggatgg	4560
tctcagggtc	tctgttggcc	tttggatc	gttccat	ctgtgtatgc	tgccaggag	4620
ggggctaagg	ggctggatcc	accaaggc	tcacagcc	gaaaactct	ggaatgaac	4680
actgaattca	ggggatgggg	tttggatgggg	gggttctcg	gtgtgtgc	gctacac	4740
tgttctgtat	gggtccat	tcactcgat	taaatcaaca	gaaacaca	4800	

SEQ ID NO: 30

moltype = AA length = 626

FEATURE

Location/Qualifiers

REGION

1..626

note = protein-arginine deiminase type-3 isoform X1 [Homo]

-continued

```

source          sapiens]
1..626
mol_type = protein
organism = Homo sapiens

SEQUENCE: 30
MFEEVYGTGPGV DIYISPNNMER GRERADTRRW RFDATELIIIV VMNSPSNDLW DSHVQISYHS 60
SHEPLPLAYA VLYLTCVDIS LDCDLNCEGR QDRNFVDKRQ WWGPSPGYGG ILLVNCDRDD 120
PSCDVQDNCD QHVHCLQDLE DMSVMVLRQC GPAALFDDHK LVLHTSSYDA KRAQVFHICG 180
PEDVCEAYRH VLGQDKVSYE VPRLHGDEER FFVEGLSFPD AGFTGLISFH VTLLDDSNED 240
FSASPIFTDT VFVFRVPPWIM TPSTLPPLEV YVCRVRNNNTC FVDAVAELAR KAGCKLTICP 300
QAENRNDRWI QDEMELGYVQ APHTKLVPVF DSPRNGELOQ FPYKRILGPD FGYVTREPRD 360
RSVSGLDSFG NLEVSPPVVA NGKEYPLGRJ LIGGNLPGS SGRVTVQVRD FLHQAQKVQPP 420
VELFVWDLAV GHVDEFLSFV PAPDGKGFRM LLASPGACFK LFQEKQKCGH GRALLFQGVV 480
DEEQVKTISI NQVLSNKKDLI NYNKPFVQSCI DWNREVLKRE LGLAECDIID IPQLFKTERK 540
KATAFFPDLV NMLVLGKHG IPKPKPFIIN GCCCLEEKVR SLLEPLGLHC TFIDDFTPYH 600
MLHGEVHCGT NVCRKEFSSFWWWNMVP 626

SEQ_ID NO: 31      moltype = DNA length = 4800
FEATURE           Location/Qualifiers
misc_feature      1..4800
note = PREDICTED: Homo sapiens peptidyl arginine deiminase
                  3 (PADI3), transcript variant X1, mRNA
source            1..4800
mol_type = unassigned DNA
organism = Homo sapiens

SEQUENCE: 31
tcccccgtt tctggcattt cccctgtggg gggccaaaca ggcagagg 60
ctccacccgc cctggaaatt acaggcttg ggagtaggtg tgccagataa aataaaagat 120
gccttagttt aattgaattt caggttaata attataattt ttttagtat aagtgtatcc 180
caaataacttc acgggaaatc caaacaatga ttgactgttc acttgaattt caagtccac 240
tgagctttt ctatgtt ttgactaaggc ttggcagggc gtttagggg agaaaacagg 300
gtgtctgtcc cttccaaatc tgaggacac ctgacagggg cccagacacc agagtgggac 360
ctgcctgtat gagagaagct tgggttggg ttatccctt ttcataaaa aagctgtttt 420
gagatgtaaag aagatcatgt tgggttggg gtgaagagggt tgggttggg gaggctgtcc 480
ttactctgtgg ggaagatggg aagccgggggg agtggccacg tggaaattt tgggttagtc 540
ccctacatgtt cctgtgcctc agtttttttca tggtaaaaca gggataatata tctcatctgt 600
gggctgcagggtt gaggatccca tgagctaattt tttttttttt cttttttttt cttttttttt 660
taagagtgcg aaggctgggg gctttttatg ttatggattt gttgttattt ttactctgtt 720
ggaggtgtgtt ggcaaaatgg tggactggg aggccacatgg ctgggtttgg atccccggcc 780
ggccacacac ccgtttttttt tttttttttt tttttttttt tttttttttt tttttttttt 840
ataatagata gtcgtttttt tctttttttt tttttttttt tttttttttt tttttttttt 900
agcatgatgtt cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 960
ccaggaggccc ctccaaatcg cttttttttt tttttttttt tttttttttt tttttttttt 1020
cctgtgttgtt tgatggccat tttttttttt tttttttttt tttttttttt tttttttttt 1080
ggatcaacacac tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1140
gtggggccacc cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1200
gtgtgaatgtgg ggaatgtggc tcttcggggcc agataacatc acagataattt taaactgtca 1260
attaatttttctt gttgttgtt gttttttttt tttttttttt tttttttttt tttttttttt 1320
tggaaacacacac tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1380
cctcagggtttt gtcacggggc tttttttttt tttttttttt tttttttttt tttttttttt 1440
tattgtctacatc ttatccatgtt gtaaaggatgg agggaggtgc gtgcggccagg tgcctggcat 1500
gcagcggggca cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1560
caccctacttgg gggatccatgtt gatggatctt qgggtqaaaa ttgtctttttt taaactcttc 1620
agaactgtaa gggggccacac ccaccaatgg ctgggtttca ctgagctttt cgtgcccaga 1680
ctctgcactgg ggatttatcg agcacatgtt tggcttttgc ctggcaaga gctggccatt 1740
ctgggtctgtt cttggggggc cagaaatgtt tgaggcttat gggacggctg gctgtggacat 1800
ctacatcttc cccaaatgtt agggggccgg ggacgggtca gacacccggc ggtggccctt 1860
tgacgcgtactt ttggggatca tgggtgtcat gaactccccca agcaatgacc tcaacgcac 1920
ccatgttcacatc actccacatc tgacgttttgc cccctggccctt atgcgggtgt 1980
ctacaccttccatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2040
cagggactttt gtagacaaggc ggcggatggg ctggggggccgg agtgggtatg gccggatctt 2100
gtctgggtttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2160
cgtgcactgtt ctggcaaggacc tttttttttt tttttttttt tttttttttt tttttttttt 2220
tgcggccctt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2280
ggcacaggccctt cttccatctt ggggttttttgc ggtatgtgtt ggggttttttgc ggcacatgtt 2340
ggggccaaatgtt aagggtgttttgc atggatgttcc cccgttttgc ggggttttttgc agcgttttttgc 2400
cgtggaaatgtt ctgtccatcc tttttttttt tttttttttt tttttttttt tttttttttt 2460
tctgtgtgtt gactccatgtt ggttttttttgc tttttttttt tttttttttt tttttttttt 2520
gtttccggatgtt gcacccatgtt cttccatctt ggggttttttgc ggggttttttgc cggcaggggcc 2580
gtggccgtgtt gggatccatgtt cttccatctt ggggttttttgc ggggttttttgc cggcaggggcc 2640
ccggctgtcaatc ctggccatctt ggggttttttgc ggggttttttgc ggggttttttgc cggcaggggcc 2700
tgagatggatgtt ctgggttttgc ttttttttttgc ggggttttttgc ggggttttttgc cggcaggggcc 2760
cccaaggaaatgtt gggggaaatgtt cttccatctt ggggttttttgc ggggttttttgc cggcaggggcc 2820
ttacgtgtactt cggggaaaccatc ggcacatgttcc ttttttttttgc ggggttttttgc ggggttttttgc 2880
gggggttcacatc cttccatctt ggggttttttgc ggggttttttgc ggggttttttgc cggcaggggcc 2940
tggggccaaatgtt cttccatctt ggggttttttgc ggggttttttgc ggggttttttgc cggcaggggcc 3000

```

-continued

ccatgcccg	aagggtcagc	ccccctgtgg	gctctttgtg	gactgggtgg	ccgtggggca	3060
tgtggatgag	tttctgagct	ttgtccctgc	ccccatgggg	aagggttc	ggatgtccct	3120
ggccagccct	ggggctgtct	tcagaatctt	ccaggaaaag	cagaagtgtg	gccacgggg	3180
ggccctctg	ttccaggggg	ttgttgatga	tgagcaggtc	aagcacatct	ccatcaacca	3240
ggtgtctcc	aataaaagacc	tcatcaacta	caataagttt	gtgcagagct	gcatgcactg	3300
gaaccgtgag	gtgtgtcagc	gggagctggg	cctggcagag	tgtgacatca	ttgacatccc	3360
acagcttcc	aagaccgaga	ggaaaaaaagc	aacgacccctt	ttccctgact	ttgtgaacat	3420
gtctgggtgt	ggggaaaccc	tggcataccc	caagccctt	gggccccatca	tcaatggctg	3480
ctgtgtcgt	gaggagaagg	tgccgtccct	gctggagccg	ctgggctcc	actgcaccc	3540
cattgtatgac	ttcactccat	accacatgt	gcatgggggg	gtgcactgtg	gcaccaatgt	3600
gtgcagaaag	cccttcttctt	tcagaatggt	gaaatgggt	ccctggagaca	gtctccaccc	3660
accatcttgt	ccccctgggg	cggcattgg	ccccaggtgt	ggagacagag	acaggcccct	3720
gaacgataag	caccaagaga	ccccaaaggct	ccagatggaa	cactgggggt	gaccgtccct	3780
ctcagaagcc	tttccctgt	aagtgtccat	gcctcaccc	caacccatgt	ggtttctcaga	3840
cttgaatctt	ctcgcccccc	caaaaagaag	gacccattt	tttagatct	ctctgtgtat	3900
tcaacacaac	ccatggagat	gtccccctt	cactctaaa	tcatccattt	ggggacaat	3960
ccacattggg	gtctagaaac	atccacgtat	tcatcagcc	atcttgcct	gtgcacatct	4020
acagaggaaag	gatccatgtat	tctgtttgg	tccaattgt	tcctcttcgc	agaggaacaa	4080
ccctaaacc	agaccatcc	acgcaggaca	ggcaggagag	atttcccta	aagcctccccc	4140
cataaaaaagg	gagctgtgga	tccactttaga	tcaggcggga	accatcttcc	acccggccaa	4200
gctcctgccc	agatgtgac	cctccacccag	cgtgagctgt	cacatagtag	gagttcttag	4260
atgcatgtgg	aagcaatgag	agttgtccct	tagccttata	aactcccat	gatctgcacat	4320
gcagaaatcc	agcccttgc	agaatcttcc	tggaaatttct	tggagacgaa	agtatctggg	4380
ggattgttgg	gtactaggga	gactgggtac	aagggtgaaa	agtagtccc	ataatacaca	4440
tggttgacta	tggtgtatcca	ccttgcgtat	gttaatatta	ggtgtctgga	gaagggttgc	4500
tcatgtggcc	ttggacttct	ctctgcggaa	ggagagaacg	ctgcctctcc	tctggatgg	4560
tctcaggcgt	tctgttggcc	tttggtcagd	ttttccat	ctctgcgtgc	tgcaaggagag	4620
ggggctaaagg	ggctggatcc	accacaggcg	ctcacagcgg	aaaaactctg	ggaatgaacc	4680
actgaatcca	ggggatgggg	gtgggggggg	ggttctcgag	gtgtgtgc	gctacacgtg	4740
tgttctgtat	gggtccagct	gcgttccat	cactcgtaa	taaatcaaca	gaaacacaaa	4800

SEQ ID NO: 32 moltype = AA length = 626
 FEATURE Location/Qualifiers
 REGION 1..626
 note = protein-arginine deiminase type-3 isoform X1 [Homo sapiens]
 source 1..626
 mol_type = protein
 organism = Homo sapiens
 SEQUENCE: 32
 MFEEVYGTGPG DIYISPNNMER GRERADTRRW RFDATLEIIV VMNSPSNDLN DSHVQISYHS 60
 SHEPLPLAYA VLYLTCVDIS LDCDLNCEGR QDRNFVDKRO WWGPGSGYGG ILLVNCDRDD 120
 PSCDVQDNCD QHVHCLQDLE DMSSVMLRTQ GPAALFDDHK LVLHHTSSYDA KRAQVFHICG 180
 PEDVCEAYRH VLGQDKVSYE VPRHLGDEER FFVEGLSPFD AGFTGLISFH VTLLDDSNED 240
 FSASPIFTDT VVFRVAPWIM TPSTLPPLEV YVCRVRNNTC DSVDAVELAR KAGCKLTICP 300
 QAENRNDRWI QDEMELEYVQ APHKTLPVVF DSPrNGELQD FPYKRILGPD FGVTREPRD 360
 RSVSGLDSFG NLEVSKPVVN NGKHCPLGRI LIIGGNLPGSS GRRVTQVVRD FLHQAKVQPP 420
 VELFVWDWLAV GHVDEFLSFV PADDGKGPFM LIASPGACEK LPQEKKQKCGH GRALLFQGVV 480
 DDEQVKTISI NQVLSNKLDI NYNKFVQSCI DWNRREVLIKRE LGLAECIDIID IPQLFKTERK 540
 KATTFFPDLV NMLVLGKHLG IPKPFGPIN GCCCLEEKVR SLLEPLGLHC TFIDDFTPYH 600
 MLHGEVHCGT NVCRKPFSFK WWNMVP 626

SEQ ID NO: 33 moltype = DNA length = 4800
 FEATURE Location/Qualifiers
 misc_feature 1..4800
 note = PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X1, mRNA
 source 1..4800
 mol_type = unassigned DNA
 organism = Homo sapiens
 SEQUENCE: 33

tctccctgtt	tctggcattt	ccccctgtggg	gggctgcccc	ggccaaaca	ggccagaggg	60
ctccaccaggc	cctggaaattt	acaggcttgc	ggagtaggg	tgccagataa	aataaaagat	120
gccttagtttca	aattgttattt	caggataata	attaaataat	tttttgtat	aagtgtatcc	180
caaataacttc	acggggatc	caaaaatgt	ttgactgttc	acttggaaaatt	caagtccaa	240
ttagctttctt	gtatttttat	ttgtcaatgc	ttggcagccct	gctttaggg	agaaaacagg	300
gtgtctgtcc	ctcccaactg	gtgaggacac	ctgacaggggg	ccccagaccc	agagtgggac	360
ctgcctgtat	gagagaagct	ttgggttgg	ttatccctt	tttcataaaa	aagctgtctt	420
gagatgttaag	agaatcatgt	ttgttggaa	gtgttgggg	tggttgggg	gagggtctg	480
ttactctgt	ggaatgtggg	aaagggggggg	agtggccagc	tgtggaaattt	ttgggttagc	540
ccttacgtat	cctgtgcctc	agttttccca	ttgtttaaaa	ggagtaat	tctcatctg	600
ggggctgcagt	gaggatcaga	tgagotaatg	tacaaaaagg	cttagggagg	cctgcgcgt	660
taagagtgc	gaggctgggg	gtttttat	ttatgttatt	ttactcc	ttactcc	720
ggaggtgggt	ggcaagagt	tggactgagg	agccacatgg	ctgggttgg	atcccagccc	780
ggccacagac	cttggcag	ttacttaacc	tcggcttct	catctgtata	atgggagaca	840
ataatagata	gtgcctccct	tcctagggtt	ttgagaggat	tgcatgtat	gctggcacac	900

-continued

agcatgagct	cagtaagtgt	cagtcactgt	tatccccca	ggccatgcac	gccgcgcct	960
ccaggagccc	ctccaaactcg	cagtgcgcctc	tcctacatg	ggttccca	gagctgtgat	1020
cctgtggcct	tgagcccat	ctgggtgccc	tggacaatgt	tggctctcag	ctccccagcg	1080
ggatcaacag	tggggagagc	agtgggtga	tctgcccgc	aggcgtacc	tgatgcaca	1140
gtgggcaccc	cacagtcaa	ggatgcata	tctggttcc	aaggagaa	cccaggctag	1200
gttggaaatgg	gaaaatgcac	tctcaggccc	agataaacat	acagatactt	taaaactgca	1260
attaattttct	gggtgttga	ccaaatccgg	ccaagtttc	cggtgtcac	tgcagagctg	1320
tggAACAGCA	tgggttggc	gtcagaccc	tggcacatct	ggctctgc	tgtatgcac	1380
cctcaggctt	gtcacgagc	ctcettgago	ctcagttgt	aatatgcac	atggggatga	1440
tattgttcac	tttacatgtt	gttaaagatgg	agggagctgc	gtgcgcagg	tgcctggc	1500
gcacggggca	tcacaaaga	ttaatccgg	ttcccaaggaa	agaatcaag	gatatcttc	1560
caccactaa	ggagacatgt	gatggatct	gggggtgagaa	ttgctcttt	taaagctctc	1620
agaactgtaa	ggggctccac	ccaccaatgg	ctgggttca	ctgagcttt	cgtgcccaga	1680
cctgcactgg	ggattttatcg	agcacatgg	tggcttgc	cttccgaaga	gtcgccat	1740
ctggtcagtg	ctggggggc	cagaatgtt	tgaggcttat	gggacgectg	gcgtggacat	1800
ctacatctt	cccaacatgg	agagggggc	ggacgtgtca	gacaccaggc	gggtggcgtt	1860
tgacgcgact	ttggagatca	tcgtggcat	gaactcccc	agcaatgacc	tcaacgacag	1920
ccatgttcag	atttcttacc	actccagcc	tgacgcctc	ccccctggc	atgcgggtgt	1980
ctacatcacc	tgtgttgc	tctctgttgc	ttgcgaccc	aactgttgg	gaaggcagg	2040
caggactt	tttagacaagc	ggcgtgggt	ctggggggcc	agtgggtatg	gcccgcattt	2100
gctggtaac	tgtgaccgt	atgatccgag	ctgtatgtc	caggacaatt	gtgaccagca	2160
cgtgcactgc	ctgcaagacc	tggaaagat	gtctgtatgc	gtcctgcg	cgcaggccc	2220
tgcaagccctc	tttgatgacc	acaaatgtt	cctccatacc	tcacatgtat	atgcacaa	2280
ggcacaggtc	ttccacatct	gcccgttgc	ggatgtgt	gaggctata	ggcatgtgt	2340
ggccaagat	aagggttcc	atgaggtaco	ccgcttgc	ggggatgagg	agcgttttctt	2400
cgtggaaaggc	ctgtccctcc	ctgatgcgg	cttcacacgg	ctcatctct	tccatgtcac	2460
tctgtcgac	gacttccaa	aggatttcd	ggcatccct	atcttcaact	acactgtgg	2520
gttcccgagt	gcacccctgg	tcatgaccc	cagcactcg	ccacccctag	agggttatgt	2580
gtgccgtgt	aggaacaaca	cgtgtttgt	ggatgcgg	gcagagctgg	ccaggaaagg	2640
cggctgcaag	ctgacatcc	gcccacaggg	cgagaaccgc	aacgaccgc	ggatccagg	2700
tgagatggag	ctggggatcg	ttcaggcgc	gcaaaagacc	ctcccggtt	tctttactc	2760
cacaaggaaat	ggggaaactgc	aggatttcc	ttacaaaaga	atcttgggtc	cagattttgg	2820
ttacgtgact	cgggaaaccac	gcccacgg	tggtgagttgc	ctggactct	ttggaaacct	2880
ggaggtcagc	cctccagttg	tggcaatgg	gaaagagtac	ccccctggg	ggatccctat	2940
tgggggcaac	ctggcttgc	caaggcc	cagggtcacc	cagggtggc	gggacttcc	3000
cctatggcc	aagggtcagc	ccccctgg	gcttttgc	gactgttgg	ccgtggggca	3060
tgttgatgt	tttctgagct	tttgcctgc	ccccatgg	aaaggcttcc	ggatgtctt	3120
ggccagecc	ggggccgt	tcaagctt	ccaggaaag	cagaagtgt	gccacggag	3180
ggccctctt	ttccatgggg	ttgttgat	tgagcagg	aagaccat	ccatcaacca	3240
gggtgtctcc	tttcaaggg	tcatcaacta	caataagt	gtgcagat	gatcgact	3300
gaaccgttag	gtgtcaagc	ggggacttgg	cctggcag	tgtgacatca	ttgacatcc	3360
acagcttcc	aagaccgaga	ggaaaaaa	aacggcc	ttccctgact	tggtaacat	3420
gtctgggtcg	ggggacacc	tgggatct	caagccctt	gggaccat	tcaatggct	3480
ctgtgtctg	gaggagaagg	tgccgttcc	gctggagcc	ctggcttcc	actgcaccc	3540
cattgtgt	tttccatccat	accacatgt	gcatgggg	gtgcactgt	gcaccaatgt	3600
gtcagaaag	cccttctt	tcaagttgc	gaacatgtt	ccctggag	gtctccaccc	3660
accatctgt	ccccctgggg	cgggcatgtt	ccccatgtt	ggagacagag	acaggccct	3720
gaacgataag	caccaagaga	ccccaa	ccagatgg	cactgtgg	gaccgtcc	3780
ctcagaagcc	tttccctgg	aagtgtccat	gcctcaccc	caacccatgt	ggttctcaga	3840
tttgaatctt	ctcgcccc	caaaaagaag	gacctcattt	ttttagact	ctcctgtat	3900
tcaacacaac	ccatggat	gtccccctt	cactcttgc	aatccat	ggggacaaat	3960
ccacatggg	gtctgaaac	atccacgtat	tcatcagcc	atcttgtct	gtgcaccc	4020
acagaggaa	gatccatgt	tctgttttg	tccaattgt	tctcttgc	agaggaacaa	4080
ccctaaaacc	agaccactcc	acgcaggaca	ggcaggag	atttcc	aagccccc	4140
cataaaaaaa	gagctgtgg	tccacttga	tcaggcgg	accatctt	acccggccaa	4200
gctcttgc	agatgttgc	cctcacc	cgtgagct	cacatgttgc	gagcttctag	4260
atgcattgtt	aagcaatgt	agttgttcc	tagccttata	aactcccat	gatctgcac	4320
gcagaaatcc	agccttgc	agaatctt	tggaaattt	tggagacaa	agtatctgg	4380
ggattgttgg	gtactaggaa	gactgggt	aagggt	aaatgtccc	ataatacaca	4440
ttgttgc	ttgtgtat	ccttgc	ttatattt	gggtgttgc	gaaggttgc	4500
tcattggccc	ttggacttct	ctctgc	ggagagaac	ctgccttcc	tctggatgg	4560
tctcaggctc	tctgttgc	tttgc	tttccacat	cctgcttgc	tgccaggag	4620
ggggctaagg	ggctggatcc	accaaggc	ctcacagcg	gaaaactctg	ggaatgaacc	4680
actgaattca	ggggatgggg	gtgggggg	ggttctc	gatgcac	4740	
tgttctgtat	gggtccat	cactcg	taatcaaca	gaaacacaaa	4800	

SEQ ID NO: 34 moltype = AA length = 626
 FEATURE Location/Qualifiers
 REGION 1..626
 note = protein-arginine deiminase type-3 isoform X1 [Homo sapiens]
 source 1..626
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 34
 MFEVYGTGPGV DIYISPNMER GRERADTRRW RFDATEIIV VMNSPSNDLN DSHVQISYHS 60

-continued

SHEPLPLAYA	VLYLTCVDIS	LDCDLNCEGR	QDRNFVDKRQ	WWGPGSYGG	ILLVNCDRDD	120
PSCDVQDNCD	QHVHCLQDLE	DMSVMVLRTQ	GPAALFDDHK	LVLHTSSYDA	KRAQVFHICG	180
PEDVCEAYRH	VLGQDKVSYE	VPRLHGDEER	FFVEGLSPFD	AGFTGLISFH	VTLLDDSNED	240
FSASPIFTDT	VVFRVAPWIM	TPSTLPPLEV	YVCRVRNNTC	FVDAVAELAR	KAGKLTICP	300
QAENRNDRWI	QDEMELGYVQ	APHKTLPVVF	DSPRNGELQD	FPYKRILGPD	FGYVTREPRD	360
RSVSGLDSFG	NLEVSPPVVA	NGKEYPLGRI	LIGGNLPGSS	GRVRTQVVRD	FLHAQKVQPP	420
VELFVDWLAV	GHVDEFLSFV	PAPDGKGFRM	LLASPAGCFK	LFQEKKQKCHG	GRALLFQGVV	480
DDEQVKTISI	NQVLSNKDLI	YNKFVQSCI	DWNREVLIKRE	LGLAECDIID	IPQLFKTERK	540
KATAFFPDLV	NMLVLGKHLG	IPKPFGTIIIN	GCCCLEEKVR	SLLEPLGLHC	TFIDDDFTPYH	600
MLHGEVHCGT	NVCRKPSFK	WWNMVP				626
SEQ ID NO: 35		moltype = DNA	length = 4800			
FEATURE		Location/Qualifiers				
misc_feature		1..4800				
		note = PREDICTED: Homo sapiens peptidyl arginine deiminase				
		3 (PADI3), transcript variant X1, mRNA				
source		1..4800				
		mol_type = unassigned DNA				
		organism = Homo sapiens				
SEQUENCE: 35						
tctccctgtt	tctgccattt	ccccctgtggg	gggctgccc	ggccaaaca	ggccagagg	60
ctccaccaggc	cttggaaattt	acaggctttg	ggagtaggg	tgccagataa	aataaaagat	120
gccttagttt	aattgttattt	caggtaataa	attaataat	ttttttagtat	aagtgtatcc	180
caaataacttc	acgggaggtac	caaacaatga	ttgactgttc	acttggaaa	caagtccaa	240
tgagctttct	gtatttttat	ttgtaagtc	tggcagccct	gcttaggg	agaaaacagg	300
gtgtctgtcc	ctcccaactg	gtgaggcacd	ctgacagggg	cccagacacc	agagtgggac	360
ctgcctgtat	gagagaaatgt	tgggottgtt	ttatccccc	tttccataaa	aagtgcgtt	420
gagatgttag	aatcatgt	ttgttggaa	gtgaaagagg	tgtttttggg	gagggtctgcc	480
ttactctgtat	ggaatgtggg	aaggcggggg	agtggccacg	tgtggaaaatt	tgggttagtc	540
cctctacgtat	cttgcgcctc	agttttttca	tgtgtaaaaca	ggagtaat	tctcatctg	600
gggctgcaatgt	gaggatcaga	tgagatcaat	tacaaaaagg	cttaggcagg	cctgecggtg	660
taagagtgtca	gaggctgggg	gtctttatgt	ttaggttatt	ttactccctg	720	
ggaggtgtgt	ggcaaaatgt	tggactgagg	agccacatgg	ctgggtttgg	atcccaagcc	780
ggccacacagc	cttggcaag	ttacttaacc	tcggctttct	catctgtata	atggggagaca	840
ataaatagata	gtggccccc	tcttaggtt	tttggaggat	tgcatgtat	gtggccacac	900
agcatgactgt	cagtaatgt	cagtatgt	tatccccca	gggcatgtat	ggcgctgcct	960
ccaggaggccc	ctccaaactcg	cagtggccct	tccttacatg	gttcccaact	gagctgtat	1020
cctgtggccct	tgagcccat	ctggctgccc	tggacaatgt	tggctctcag	ctccccagcg	1080
ggatcaacag	tgggggagac	agtgggttga	tctggcccg	aggcgtacc	tgagtgccaa	1140
gtggggcaccc	cacagtcaaa	ggatgtttcc	tctggttcc	aaggaggaa	cccaggctag	1200
gtgaaatgtat	gaaatgtaa	tctcaggggc	agataaacatc	acagataactt	taaacgtca	1260
attaatttct	ggttgtttga	ccaaaatccgg	ccaagttcc	cggtgtacac	tgcagagctg	1320
tggAACACG	tggggcttgg	gtcaagaccc	tttgcacatct	ggctctgc	tgtgtatgca	1380
cctcaggcgtt	gtcagcaggc	cttcaggatc	ctcaatgtt	aatatgtccaa	atggggatgt	1440
tattgtctac	cttacagctt	ttaaatgtt	ttggatgttc	aggggatgtc	gtgcggccagg	1500
gcagcgggca	ctcaaccaga	ttgaattccc	tttcccaaggaa	agaattcaag	gatatctcc	1560
caccctactt	ggagacatgt	gtatggatct	gggggtggaaa	tttgcctttt	taaagctctc	1620
agaactgtaa	gggggtccac	ccaccaatgg	ttgggtgtca	tttgcgttcc	ctgtgtccaga	1680
cctgcactgg	ggattttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	1740
ctgggtcagt	cttggggca	cagaaatgtt	ttggatgtat	ggggacccctg	gggtggacat	1800
ctacatcttc	cccaacatcg	agaggggccg	ggggccgtgc	gacaccaggc	gggtggccgt	1860
tgacgcgtat	ttggatgtat	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	1920
ccatgttcag	atttccatcc	actttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	1980
ctacctcacc	tgtgttgc	tctttcttgc	ttttttatcg	ttttttatcg	ttttttatcg	2040
caggaaactt	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2100
gtgtgttgc	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2160
cgttgtactgc	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2220
tgtgtgttgc	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2280
ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2340
ggcacagggtt	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2400
ggggccatgg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2460
tctgtgttgc	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2520
ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2580
ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2640
ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2700
ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2760
ccaaaggaaat	ggggaaactgc	aggattttcc	tttccaaatgg	atccctgggtc	cagattttgg	2820
ttacgtgtact	ggggaaaccac	gcaacatgtt	ttttttatcg	ttttttatcg	ttttttatcg	2880
ggaggtcaggc	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	2940
tggggcaac	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	3000
ccatgttcac	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	3060
tgtggatgtat	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	3120
ggccagccctt	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	3180
ggcccttcctt	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	3240
gtgtgtctcc	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	3300
gaaccgttgc	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	ttttttatcg	3360

-continued

acagcttc aagaccgaga	ggaaaaaago aacggccttc	ttccctgaact tggtaacat	3420
gtcggtgtc gggaaagcacc	tggcatccc caagccctt	gggccccatca tcaatggctg	3480
ctgtgcctg gaggagaagg	tgcagtcctc	gctggagccg ctggcctcc	3540
cattatgac ttcaatccat	accacatgt	gcatggggag gtgcactgtg	3600
gtgcagaaag cccttcctt	tcaagtggtg	gaacatggtg	3660
accatccgt cccctgggg	cgggcatgg	cccagggtgg ggagacagag	3720
gaacgataag cacaagaga	ccccaaagg	ccagatgggg gaccgtccct	3780
ctcagaagcc ttccctgg	aagtgtccat	gcctcaccc	3840
cttgaatctt ctggcccc	caaaaagaag	gacctcattt	3900
taaacacaac ccatggatg	gtccccctt	cactctgaaa	3960
ccacattgg gtcataaac	atccacgtat	tcatacagcc	4020
acagaggaaag gatccatgt	tctgttgg	tccaatttgt	4080
ccctaaaacc agaccactc	acgcaggaca	ggcaggagag	4140
cataaaaaagg gagctgtgg	tccacttaga	tcaggcggaa	4200
gctcctgccc agatgttgc	cctcaccc	cgtgagctgt	4260
atgcatgtt aagcaatgag	agtgtccct	cacatgttag	4320
gcagaaaatcc agccttgc	agaatctcc	tggagacgaa	4380
ggattgttgg gtaactggaa	gactgggtt	aagggtgaaa	4440
tggttacta tggtgtatcca	ccttgcgtat	gttaatatta	4500
tcattggccc tggacttct	ctctgcggaa	ggagagaaacg	4560
tctcaggcct tctgttggcc	tttggtcago	gtttccacat	4620
ggggctaaagg ggctggatcc	accaggcag	ctcacagcgg	4680
actgaattca gggggatgggg	gtggggggc	ggttctcgag	4740
tgttctgtat gggctccat	cactcgtaa	taaatcaaca	4800

SEQ ID NO: 36 moltype = AA length = 626
FEATURE Location/Qualifiers
REGION 1..626
 note = protein-arginine deiminase type-3 isoform X1 [Homo sapiens]
source 1..626
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 36
MFEVYGTGPV DIYISPNNMER GRERADTRRW RFDATLEIIV VMNSPSNDLN DSHVQISYHS 60
SHEPLPLPLAYA VLYLTCVDIS LDCLDNCEGR QDRNFVDKRQ WWGPGSYGG ILLVNCDRDD 120
PSCDVQDNCD QHVHCLQDLE DMSVMVLRTQ GPAALFDDHK LVLHTSSYDA KRAQVFHICG 180
PEDVCEAYRH VLGQDKVSYE VPRLHGDEER FFVEGLSPFD AGFTGLISFH VTLLDDSNED 240
FSASPIPTDT VFVRVAPWIM TPSTLPPLEV YVCRVRNNTC FVDAVAELAR KAGCKLTICP 300
QAENRNDRWI QDEMELGYVQ APHTKLPPVF DSPRNGELOQD FPYKRILGPD FGYYTREPRD 360
RSVSGLDSFG NLEVSPPVVA NGKEYPLGRI LIGGNLPGSS GRRVTQVVRD FLHAQKVQPP 420
VELFVWDLAV GHVDEFLSFV PAPDGKGFRM LLASPGACFK LFQEKEKCGH GRALLFQGVV 480
DDEQVKTISI NQVLSNKDLI NYNKFVQSCT DWNRREVLRK LGLAECDIID IPQLFKTERK 540
KATAFFPDLV NMLVLGKHLG IPKPKGPIIN GCCCLEEKVQ SLLEPLGLHC TFIDDFTPYH 600
MLGEVHCGT NVCRKPFSFK WWNMVP 626

SEQ ID NO: 37 moltype = DNA length = 2938
FEATURE Location/Qualifiers
misc_feature 1..2938
 note = PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PAD13), transcript variant X2, mRNA
source 1..2938
 mol_type = unassigned DNA
 organism = Homo sapiens

SEQUENCE: 37
agaatggat ggtatgtact gtgtgctaataatgtgtc gatcctgacc taggcgagaa gagaaccaa 60
ccagatggat cccacagttc ataatgtgtc gatcctgacc taggcgagaa gagaaccaa 120
atatgaaact gttgaagaac ttggactga attatgtgg aacttgggtc cctggagtt 180
aaaggagaa ggtcggatgg cgatgggtt gggggcccaag tgggtatggc ggcatcttc 240
tggtaactg tgaccgtat gatcgatgtc tgatgttccca ggacaaatgt gaccagcact 300
tgcaactgcgca gaaacatgtt gatcgatgtt cttgcgtatggt cttgcgtatggc 360
cagcccttcc ttatgttccaaatgttcc tccatatacctc cagatgtatgtatggc 420
cacaggatcc ttatgttccaaatgttcc tccatatacctc cagatgtatgtatggc 480
gcccaatggat ggttccatgttcc gatcgatgttcc tccatatacctc cagatgtatgtatggc 540
tggaaaggatc gtccttcctt gatcgatgttcc tccatatacctc cagatgtatgtatggc 600
tgatgttccca gtccttcctt gatcgatgttcc tccatatacctc cagatgtatgtatggc 660
tccatgttccca gtccttcctt gatcgatgttcc tccatatacctc cagatgtatgtatggc 720
ggccgtgtgg gaaacacacg tggtttgtgg atggcgttccca gatcgatgttcc 780
gtcgcaactg gatcgatgttcc tccatatacctc cagatgtatgtatggc 840
agatggatggat ggttccatgttcc tccatatacctc cagatgtatgtatggc 900
caaggaaatgg ggaactgtc gatccatgttcc tccatatacctc cagatgtatgtatggc 960
acgtgactcg gaaaccacgc gacaggatgtc tgatgttccca gatcgatgttcc 1020
aggtcagccc tccatgttcc tccatatacctc cagatgtatgtatggc 1080
ggggcaacact ggttccatgttcc tccatatacctc cagatgtatgtatggc 1140
atgcccacaa ggttccatgttcc tccatatacctc cagatgtatgtatggc 1200
tggatgatgtt tccatgttcc tccatatacctc cagatgtatgtatggc 1260

-continued

ccagcccctgg	ggcctgttcc	aagctttcc	aggaaaagca	gaagtgtggc	cacggggagg	1320
ccctccctgtt	ccaggggggtt	gttcatgtat	agcaggtcaa	gaccatccc	atcaaccagg	1380
tgcctccaa	taaagactc	atcaactaca	ataagtttg	gcagagtc	atcgacttgg	1440
accgtgagg	gtogaagccg	gagctggcc	tggcagatgt	tgacatcatt	gacatccac	1500
agcttccaa	gaccgagagg	aaaaaaagca	cggccttctt	ccctgactt	gtgaacatgc	1560
tggtgcgtgg	gaagcacctg	ggcatccccca	agecccttgg	gcccacatc	aatggctgt	1620
getgcgtgg	ggagaagggt	cggccctgt	tggagccgt	gggcctccac	tgcacccat	1680
tttatgtact	cactccatac	cacatgtgc	atggggagg	gactgtggc	accaatgtgt	1740
gcagaaaacc	cttctttc	aagtgggtga	acatgggtgc	ctgagacagc	tcccaaccac	1800
catccctgtcc	ccctggggcc	ggcattggcc	cagggtgggg	agacagacag	aggcccctga	1860
acgataagca	ccaagagacc	ccaaggctt	agatggaaata	ctgagggtga	ccgtccccct	1920
cagaagcc	ttccctggaa	gtgtccatgc	ctcacctgc	acccatgtgg	tttcagact	1980
tgaatcttct	cgccccccca	aaaagaagga	cctcatttct	tatagectt	cctgtgttcc	2040
aacacaaccc	atggagatgt	ccccctctca	cctcgttccat	atccatgggg	ggacaaatcc	2100
acattgggg	ctagagat	ccacgtatct	catcagccat	cttgcgttgc	gcatectaacc	2160
agaggaaagg	tccatgttcc	tgctttggcc	caattgttcc	ctctgtcg	aggaacaacc	2220
ctaaaaccag	accactccac	gcaggacagg	caggagat	tcttcttaaa	gcctccccca	2280
taaaaaaggaa	gtgtgtgtat	cacttagat	agggccggac	catcttccac	ccggcccaagc	2340
tcctgcccag	atgttgcacc	tcaccccg	tgagctgtca	catatgttgg	gtttcttagat	2400
gcatgtggaa	goaatggag	ttgttccctt	gccttataaa	ctcccatgg	tctgacatgc	2460
agaaatccag	cttgcgttcc	aatcttccgt	gaatttctt	gagacgaaag	tatctgggg	2520
attgttgggt	actagggaa	ctgggttacaa	gggtgaaaag	tagtcccat	aatacacatg	2580
gttgactatg	gtgtatcc	ttgttgcatt	taatattagg	tgtctggaga	aggttgttcc	2640
atggggccctg	ggacttcttct	tcaggcagg	agagaacgc	gcctcttc	tggattggc	2700
tcaggctctc	tggttgcctt	tggttgcctt	ttccacatcc	tgtctgtgt	caggagagg	2760
ggctaaagg	ctggatccac	caaggcagct	cacagccgg	aaactctgg	aatgaaccac	2820
tgaatttgc	ggatgggggt	ggggggggcg	tttctcgagg	gtgtccagg	tacacgtgt	2880
tttctgtatgg	gtccagctgc	gttccatca	ctcgctaata	aatcaacaga	aacacaaa	2938

SEQ ID NO: 38 moltype = AA length = 485
 FEATURE Location/Qualifiers
 REGION 1..485
 note = protein-arginine deiminase type-3 isoform X2 [Homo sapiens]
 source 1..485
 mol_type = protein
 organism = Homo sapiens
 SEQUENCE: 38
 MSVMVLRTQG PAALFDDHKL VLHTSSYDAK RAQVFHICGP EDVCEAYRHV LGQDKVSYEV 60
 PRLHGDEERF FVEGLSPDPA GFTGLISFPHV TLLDDSNEDF SASPIFTDTV VFRVAPWIMT 120
 PSTLPPLEVY VCRVRNNTCF VDAVAELARK AGCKLTICPQ AENRNRDRWIQ DEMELGYVQA 180
 PHKTLPVVFDF SPRNGELQDF PYKRILGPDF GYVTREPRDR SVSGLDSFGN LEVSPPVAN 240
 GKEYPLGRIL IGGNLPGSSG RRVTVQVVRDF LHAQKVQPPV ELFVDWLAVG HVDEPLSFVP 300
 APDGKGFRML LASPGACFKL FQEKGKCGHG RALLFQGVVD DEQVKTISIN QVLSNKDLIN 360
 YNKFVQSCID WNREVLKREL GLAECDIIDI PQLFKTERKK ATAFFPDLVN MLVLGKHLGI 420
 PKPFGPIING CCCLEEKVRS LLEPLGLHCT FIDDFTPYHM LHGEVHCGTN VCRKPFSFKW 480
 WNMVP 485

SEQ ID NO: 39 moltype = DNA length = 2938
 FEATURE Location/Qualifiers
 misc_feature 1..2938
 note = PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PAD13), transcript variant X2, mRNA
 source 1..2938
 mol_type = unassigned DNA
 organism = Homo sapiens
 SEQUENCE: 39
 agaaatggat ggatgtgact gtgtgctaat aaaactttat ttatacAACAGCAGCAG 60
 ccagatggg cccacatgttcc ataatgtgt gatccatgttcc taggcggggaa gagaacccaa 120
 atatggaaact gttgaagac ttggggactt attatgttgg aacttgggtc cctggggagg 180
 aagaggagaa ggtcgagccg cagtgggtt gggggcccaag tgggtatggc ggcacatcttc 240
 tggtaactg tgaccgtat gatecgagct gtgtatgttca ggacaatgtt gaccagcagc 300
 tgcactgcct gcaagacatgt gaagacatgt ctgtcatgtt cctgcggacg cagggccctg 360
 cagcccttcc tttatgttcc tccatactcc cagctatgtt gccaacacggg 420
 cacaggctt ccacatcttc ggtcttgagg atgtgtgtt ggcctatagg catgtgttcc 480
 gccaagatata ggtgttccat gaggatcccc gtttgcgtt ggttgcgtt ccgttcttc 540
 tggaggccct gtccctccct gatgcgggt tcacaggact catctcttc catgtactc 600
 tgctggacga ctccaaacggag gatttctccg catccctat ttttactgtgtt 660
 tcccgagttgc accctggatc atgacgccc gactctggcc acccccttggat gtgtatgtt 720
 gccgtgttgag gaacaacacg ttgttttggat atgcgggtggc agagctggcc aggaaggccg 780
 gctgcgttgc gaccatcttc ccacaggccg agaaccgcac gacccgttcc atccaggatg 840
 agatggagct gggctacgtt caggggccgc acaagaccctt cccgggtggc ttttactccc 900
 caaggaaatgg ggaactgcag gatttccctt acaaaagaat cttgggttca gattttgggtt 960
 acgtgactcg ggaaccacgc gacaggtctg tgactgtggccctt gggaaacctgg 1020
 aggtcagccc tccatgttggt gccaatggga aagagtaccc cctggggagg atccttattt 1080
 gggggcaacccctt gcctgggttca agtggccgc ggttccacca ggttgggtccgg gacttccctt 1140

-continued

atgcccagaa	ggtcgcgccc	cccgctggago	tctttgtggaa	ctgggttggcc	gtggggccatg	1200
tggatgagg	tctgagcttt	gtccccgtccc	ccgatgggaa	gggcttcccg	atgtccctgg	1260
ccagccctgg	ggcctgttcc	aagetcttcc	aggaaaagca	gaagtgtggc	cacggggaggg	1320
ccttcctgtt	ccaggggggtt	gttgatgtat	agcaggtcaa	gaccatctcc	atcaaccagg	1380
tgctctccaa	taaagacctc	atacaactaca	ataagtttgt	gcagagctgc	atcgacttgg	1440
accgtgagg	gtgtggcg	gagctgggco	tggcagagtg	tgacatcatt	gacatccac	1500
agcttctcaa	gaccgagagg	aaaaaaagca	ccgcctttctt	ccctgacttg	gtgaacatgc	1560
tggtgcgtgg	gaagcacctg	ggcatcccc	agccctttgg	gcccataatc	aatggctgt	1620
gctgcctgg	ggagaagggt	cggcccttgc	tggagccgt	gggcctccac	tgcacctca	1680
ttgtatgactt	cactccat	cacatgtgc	atggggagg	gcaactgtgc	accaatgtgt	1740
gcagaaggcc	cttcttttcc	aagtggggaa	acatgggtgg	ctgagacagc	tccacccac	1800
catcctgtcc	ccctggggcg	ggcattggcc	cagggtgg	agacagagac	aggccccctga	1860
acgataagca	ccaagagacc	ccaaggctcc	agatggaaaca	ctgagggtga	ccgtccctct	1920
cagaaggcc	ttccctggaa	gtgtccatgc	ctcacctgc	accatgtgg	tttctcagact	1980
tgaatcttct	cgggccccca	aaaagaaggaa	cetcattct	tatacctct	cctgtgattc	2040
aacacaaccc	atggagatgt	cccccttc	ctctgaaatc	atccatttg	ggacaatcc	2100
acattgggg	ctagaaacat	ccacgttatct	catcagccat	cttgcctgt	gcacccat	2160
agagggaa	tccatgttcc	tgctttgggt	caattgttcc	ctctctgc	aggaaacaacc	2220
ctaaaaccag	accactccac	gcaggacagg	caggagat	tettctaaa	gcctccccca	2280
taaaaaggaa	gtgtggatc	cacttagat	agggcgaaac	catcttccac	ccggccaaagc	2340
tcctgcctcg	atgttgcacc	tcacccageg	tgagctgtca	catagtagga	gcttcagat	2400
gcatgtggaa	gaaatggag	ttgtccctta	gccttataaa	ctccccatga	tctgacatgc	2460
agaatccatcg	ccttgcaccc	aatcttctgt	gaatttcttg	gagcgaag	tatctgggg	2520
atttgtgggt	attagggaga	ctgggtacaa	gggtgaaaag	tagtcccat	aatacactat	2580
gttgactatg	tgatccacc	ttgtgtatgt	taatattagg	tgttggaga	aggttgc	2640
attggccctg	ggactctctt	ctgcaggagg	agagaacgc	gcctctcc	tggattgtc	2700
tcaggcttcc	tttgtgggtt	ttgcacatc	tgctctgtc	caggagagg	2760	
ggctaagggg	ctggatccac	caaggcaggt	cacagggaa	aaactctgg	aatgaaccac	2820
tgaattcagg	ggatgggggt	ggggggggcg	ttctcgagg	gtgtggcc	tacacgtgt	2880
ttctgtatgg	gtccagctgc	ttttccatca	ctcgctaata	aatcaacaga	aacacaaa	2938

SEQ ID NO: 40 moltype = DNA length = 2938
 FEATURE Location/Qualifiers
 misc_feature 1..2938
 note = PREDICTED: Homo sapiens peptidyl arginine deiminase
 3 (PADI3), transcript variant X2, mRNA
 source 1..2938
 mol_type = unassigned DNA
 organism = Homo sapiens

SEQUENCE: 40
 agaaatggat ggatgtact gtgtgctaat aaaactttat ttatacaaac aggcaaggtagg 60
 ccagattttg cccacacgtt ataatgtct gatccgtacc taggcgcgagaa gagaacccaa 120
 atatgaaaat ggtgaaagaa ttggacttga attatgttg aacttgggtgc ccttgggggtt 180
 aagaggagaa ggtcgacggg cttggacttgc tggggcccaag tgggtatggc ggcacatcttc 240
 tggtgaactg tgaccgtgtat gatccggact gtgtatgtca ggacaatttg gaccagcacg 300
 tgcactgtct gcaagacctg gaagacatgt ctgtcatgtt cctcgccgacg cagggccctg 360
 cagccctttt tgatgacacc aaacttgc tccataccatc cagctgtatg gccaacacggg 420
 cacagggtt ccacatctgc ggtctgttgc ggtcttgcgg atgttgcggc 480
 gccaagatataa ggtgttctat gagggtaccc ggttgcgtt ggttgcgggg cgcttctcg 540
 tggaaaggcc tgccttccct gatgcggct tcacaggact catcttcttc catgtcactc 600
 tgcgtggacca ctccaaacggat gatttctcgat catccctat ctgcactgtac actgtgtgt 660
 tcccgatggc acccttggat atgacggcca gcactctgc accccttagag gtgtatgtt 720
 gccgtgtggc gaacaaacacg tgttttgtgg atgcgtgtgc agagctggcc aggaaggccg 780
 gctgcaagct gaccatctgc ccacaggccg agaaccgcacaa cgaccgttgc atccaggatg 840
 agatggggct ggggtacatgtt caggccgcgc acaagacccct cccgggtgtc tttgactccc 900
 caaggaaatgg ggaactgtcgat gatttccctt gatgggttca gatgggtgtt 960
 acgtgacttcg gaaaccacgc gagacgttgc tgatgtggcc ggtacttccctt gggaaacctgg 1020
 aggtcagccc tccagtgtgt gccaatggga aagagtaccc cttggggagg atccttattt 1080
 ggggcaacctt gctgtggccat agtggccgcga gggttccatcca ggttgggtccgg gacttccctcc 1140
 atgcccaggaa ggtgcacggcc cccgtggccat tctttgtggc ctgggtggcc gtggccatcg 1200
 tggatgagg tttgtgttccctt gtccttgcctt cccatggggaa gggcttcccg atgtccctgg 1260
 ccagccctgg ggccctgttcc aagetcttcc agggaaaagca gaagtgtggc cacggggaggg 1320
 cccttcctgtt ccaggggggtt gttgtatgtatc agcaggtcaa gaccatctcc atcaaccagg 1380
 tgctctccaa taaagacctc atcaactaca ataagtttgtt gcaagactgc atcgacttgg 1440
 accgtgagg gtcgtggcc ggttggcc tggcagagtg tgacatcattt gacatccac 1500
 agcttctcaa gaccgagagg aaaaaaaagca cggccctttctt ccctgacttg gtgaacatgc 1560
 tggtgcgtggg gaagcacctg ggcatccccca agccctttgg gcccataatc aatggctgt 1620
 gctgcctggaa ggagaagggtt cggcccttgc tggagccgtt gggccctccac tgacccatctca 1680
 ttgtatgactt cactccat cacatgtgc atggggagg gcaactgtgc accaatgtgt 1740
 gcagaaaggcc ctcttcttcc aagtgggtggaa acatgggtggc ctgagacagc tccacccac 1800
 catcctgtcc ccctggggcg ggcattggcc cagggtgggg agacagagac aggccccctga 1860
 acgataagca ccaagagacc ccaaggctcc agatggaaaca ctgagggtga ccgtccctct 1920
 cagaaggcc ttccctggaa gtgtccatgc ctcacctgc acccatgtgg ttctcagact 1980
 tgaatcttctt cggccccccca aaaagaaggaa ctcatcttcc tatagccctt cctgtgattc 2040
 aacacaaccc atggagatgt ccccttc tca ctctgaaatc atccatttg ggacaatcc 2100
 acattggggt cttagaaacat ccacgttatctt catcagccat cttgttgcgtt gcatccat 2160

-continued

agaggaaggaa	tccatgattc	tgctttggtc	caattgtcttc	ctctctgcag	aggaacaacc	2220
ctaaaaccac	accactccac	gcaggacagg	caggagagat	tcttcttaaa	gcctccccca	2280
taaaaaggaa	gctgtggate	cacttagatc	aggcgccgaa	catctttcac	ccggccaagc	2340
tccctccccc	atgttgaccc	tcaccaggcg	tgagctgtca	catagtagga	gcttcttagat	2400
gcatgtggaa	gcaatgagag	ttgtccctta	gccttataaa	ctccccatga	tctgacatgc	2460
agaatccag	ccttgtccag	aatccctctg	gaatttcttg	gagacgaaag	tatctggggg	2520
attgttgggt	actaggaga	ctgggtacaa	gggtgaaaaa	tagtcccat	aatacacatg	2580
gttgactatg	tgatccacc	ttgtgtatgt	taatattagg	tgtctggaga	aggttgcttc	2640
attggccctg	gacttctct	ctgcaggagg	agagaacgct	gcctctcc	tggattggtc	2700
tcaggcttc	tgtggccctt	ttggtagctg	ttccacatcc	tgctctgcgt	caggagaggg	2760
ggctaagggg	ctggatccac	caaggcagot	cacagcggga	aaactctgg	aatgaaccac	2820
tgaattcagg	ggatgggggt	ggggggccgg	tttcgcagggt	gtgtgccagc	tacacgtgt	2880
ttctgtatgg	gtccagctgc	gtttccatca	ctcgctaata	aatcaacaga	aacacaaa	2938

SEQ ID NO: 41 moltype = DNA length = 2938
 FEATURE Location/Qualifiers
 misc_feature 1..2938
 note = PREDICTED: Homo sapiens peptidyl arginine deiminase
 3 (PADI3), transcript variant X2, mRNA
 source 1..2938
 mol_type = unassigned DNA
 organism = Homo sapiens

SEQUENCE: 41
 agaaatggat ggatgtgact gtgtgctaat aaaactttat ttatacaaac aggcaagtgg 60
 ccagattttg cccacagttc ataatgtgt gatcctgacc taggcggagaa gagaaccaa 120
 atatgaaact gttgaagaac ttgggactga attatgttg aacttgggtc cctggagtt 180
 aaaggagaaa ggtcgacggc cagttgggtt gggggcccaag tgggtatggc ggcacatgtc 240
 tggtaacttg tgaccgtat gatccggatc gtatgttca ggcacattgt gaccaggcacg 300
 tgcactgcct gcaagacactg ggacacatgt ctgtcatgtt cctgcggacg cagggccctg 360
 cagcccttctt tgatgaccac aaacttgc tccataccatc cagctatgtat gccaaacggg 420
 cacaggcttc ccacatggggc ggtagtgggg atgtgtgtgg ggccttatagg catgtgtgg 480
 gccaagataa ggtgtcttat gaggatcccc gtttgcgtatgg ggtatggggg cgcttcttcg 540
 tggaaaggcct gtccttcctt gatgcgggt tcacaggact catctcttc catgtcaactc 600
 tgctggacga tcccaacagag gatttctcgat catcccttat cttcaactgac actgtgtgt 660
 tcccgagttgtt atgacgttccca gcactctgccc accccatggag gtgtatgtgt 720
 gccgtgttaga gaacaacacg tgggttgggg atgcgggtggc agagatggcc aggaaggccc 780
 gctgcgaact gaccatctgc ccacaggccg agaaacccaa cgaccgttggg atccaggatg 840
 agatggagct gggctacggtt caggggccgc acaagacccct cccgggtggc tttgactccc 900
 caaggaaatgg ggaactcgac gattttccctt acaaaaatgtt cctgggttca gattttgtt 960
 acgtgacttcg ggaaccacgcg gacaggctgt tgatgtggctt ggactctttt gggaaacctgg 1020
 aggttcagccc tccagttgtt gccaatgggg aagatgtacc cttggggaggg atcccttatt 1080
 ggggcaacctt gcttgggtca agtggccgca gggtcaccca ggttgggtgg gacttccccc 1140
 atggcccaaaa ggtgcggggc cccggggggc tctttgtggc ctgggtggc gtggggccatg 1200
 tggatgggtt ttgtgtgtt gtccttcgtt ccgtatggggaa gggcttcggg atgtctctgg 1260
 ccagccctggg ggcctgttcc aagcttccca agggaaaacca gaaatgtggc cacggggaggg 1320
 ccctctgtttt ccagggggggtt gttgtatgtt agcaggatca gaccatctcc atcaaccagg 1380
 tgcctctccaa taagagactc atcaactata ataaatgttgc gcaatgtgc atcgacttgg 1440
 accctgtgggtt gctgaaggccg gagttggggc tggcaaggatc tgacatccattt gacatccac 1500
 agtcttcataa gaccggagggg aaaaaaccaaa cggcccttctt ccctgtactt gtaacatgc 1560
 tgggtgtggg ggacgacactg ggcacatccca agccctttggg gcccacatc aatggctgt 1620
 gtcgcctggg ggagaagggtt cgggtccctgc tggagccggc gggccctccac tgcacccatc 1680
 ttgtatggactt cactccatc acatgtgc atggggaggat gcaatgtggc accaactgtgt 1740
 gcaaaaaaccc ttcttttca aatgtgtggc acatgtgtcc ctggacacgc tccccccac 1800
 catccgttcc ccctggggcg ggcattggcc cagggtgggg agacagagac agggccctga 1860
 acgataacca ccaaggacacc ccaaggctcc agatggaaaca ctgggggtga ccgtccctct 1920
 cagaagectt ttccctggaa gtgtccatgc ctcacatggc accccatgtgg ttctcagact 1980
 tgaatcttctt cggccccccca aaaaaaggaa cctcattttc tatagctctt cctgtgattc 2040
 aacacaacccc atggagatgt cccttctca ctctgaaatc atccattttgg ggacaaatcc 2100
 acatgggggtt ttagaaacat ccacgttatc catcagccat cttgtctctgt gcatcttaac 2160
 agaggaaggaa tccatgattc tgctttggtc caattgtttcc ctctgtgcgg aggaacaacc 2220
 ctaaaaaccac accactccac gcaggacaggc caggagagat tcttcttaaa gcctccccca 2280
 taaaaaaggaa gctgtggatc cacttagatc agggcggaaac catctttcac ccggccaaagc 2340
 tcctgcccag atgttgaccc tcaccaggcg tgagctgtca catagtagga gcttcttagat 2400
 gcatgtggaa gcaatgagag ttgtccctta gccttataaa ctccccatga tctgacatgc 2460
 agaaatccag ctttgtccag aatccctctg gaatttcttg gagacgaaag tatctggggg 2520
 attgttgggtt actaggggaa ctgggtacaa gggtgaaaaa tagtcccat aatacacatg 2580
 gttgactatg tgatccacc ttgtgtatgtt taatattagg tgcgtggaga aggttgcttc 2640
 attggccctg gacttctctt ctgcaggagg agagaacgct gcctctcc tggattggtc 2700
 tcaggcttc tggggccctt ttccacatcc tcgtctgcgtc caggagaggg 2760
 ggctaaggggg ctggatccac caaggcagot cacagcggga aaactctgg aatgaaccac 2820
 tgaattcagg ggatgggggtt gggggccgg ttttcgcagggt gtgtgccagc tacacgtgt 2880
 ttctgtatgg gtccagctgc gtttccatca ctgcctaata aatcaacaga aacacaaa 2938

SEQ ID NO: 42 moltype = AA length = 485
 FEATURE Location/Qualifiers
 REGION 1..485

-continued

-continued

FEATURE Location/Qualifiers
REGION 1..485
note = protein-arginine deiminase type-3 isoform X2 [Homo sapiens]
source 1..485
mol_type = protein
organism = Homo sapiens

SEQUENCE: 44
MSVMVLRQQ PAALFDDHKL VLHTSSYDAK RAQVFHICGP EDVCEAYRHV LGQDKVSYEV 60
PRLHGDEERF FVEGLSFPA GFTGLISFHV TLLDDSNEDF SASPIFTDTV VFRVAPWIMT 120
PSTLPPLEVY VCRVRNNTCF VDAVAELARK AGCKLTICPQ AENRNRDRIQ DEMELGYQA 180
PHKTLPPLEVY SPRNGELQDF PYKRILGPDF GYVTREPRDR SVSGLDSFGN LEVSPPVAN 240
GKEYPLGRIL IGGNLPGSSG RRVTVQVRD LHAQKVQPPV ELFWDWLAVG HVDEFLSFVP 300
APDGKGFRML LASPGACFKL FQEOKKGHG RALLFQGVVD DEQVKTISIN QVLSNKDLIN 360
YNKFVQSCID WNREVLKREL GLAECIDI IDI PQLFKTERKK ATTFFPDLVN MLVLGKHLGI 420
PKPFGPIING CCCLEEKVRS LLEPLGLHCT FIDDPFTPYHM LHGEVHCGTN VCRKPFSEFW 480
WNMVP 485

SEQ ID NO: 45 moltype = DNA length = 2938
FEATURE Location/Qualifiers
misc_feature 1..2938
note = PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X2, mRNA
source 1..2938
mol_type = unassigned DNA
organism = Homo sapiens

SEQUENCE: 45
agaaaatggat ggatgtgact gtgtgcta ataaaactttt ttatacaa ac aggca gtagg 60
ccaggat ttttccc cccacac gtc ataaatgtgc gatcttgc acc tagggc gaa gaaaccaa 120
atatggaaact gttgaagaac ttgggactg attatgttg aa acttggc tc cttgggagg 180
aaggaggagaa ggtcgacggg cagtgggtc gggggccca g tggtatggc ggc at cttgc 240
tggtaactg tgaccgtgat gatccgagct gtgtatggc ggacaattgt gaccacgacg 300
tgcactgcgc gcaagacgtc gaagacgtc ctgtatggt ctgcggacg ca gggccct 360
cggcccttctt tgatgaccac aaactatgtcc tccatcaccc c agatcatgat gccaacacggg 420
cacagggtctt ccacatctgc ggtcttgagg atgtgtgtg a ggcctatagg catgtgtgg 480
gccaagataa ggtgttctt gaggta ccc gtttgc atgg gggatggagg cgcttcttcg 540
tggaaaggccct gtccttccctt gatccggc tca caggact catcttc tcatgtca 600
tgcgtggcga ctccaa cggc gatgttccgg catccatgtt ctatcgtac actgtgtgt 660
tccgactggc accctggatc atgacgc cca gactctgc accccctt a gatgtatgt 720
gcccgtgtgg aacaacaa cgtttttgtgg atgegggtgg agagctggc aggaaggccg 780
gtcgaacgctt gaccatctgc ccacaggccg aga acgc cca cggccgtgg atccaggat 840
agatggagctt gggcatacggtt ccggccgc aca gagactt cccgggtgg tttatgtccc 900
caaggaaatgg ggaactcgat gattttccctt aca aaaaagaat ctgggttca gat ttttgggt 960
acgtgactcg ggaaccacgc gacagggtctg tgagtggc ct g gactctttt gggaaacctgg 1020
aggttca gccc tccatgtggt gccaatggg a aagatcacc ccttggggagg atccatgg 1080
ggggcaacccctt gcttgggtca atggccgc gggtca cccca ggttggcgg gactcttc 1140
atggcccaaaa ggttgcaggcc cccgggtggatc tcttttggc tggttggc gttggccat 1200
tggatgagtt tcttgcgtt gtccttggcc cccatgggaa gggotccgg atgtcttgg 1260
ccagccctgg ggcctgttc aagcttcc ttttccca aggaaaagca gaatgtgtgc cacggggagg 1320
cccttctgtt ccagggggggtt gttatgtatc gacggatccaa gaccatctcc atcaaccagg 1380
tgcttcccaaa taagacatcc atcaactaca ataaatgtt gca gagatctgc atcaactgg 1440
accgtgaggat gcttgcaggcc gatgtggcc tggcagatgt tgacatcatt gacatcccc 1500
agcttccaa gaccggagg aaaaaaagca cggccttcc cctgtactt gtaacatgc 1560
tgggtgtgggg gaagcacctg ggcatccccca agcccttgg gaccatcatc aatggctgt 1620
gtgtccctgg gggaaagggtt cggcccttc gttggcgc ggttccca gtcacccatc 1680
ttgtatgtt cactccatc cacaatgtcc atggggaggat gcaactgtgc accaatgtgt 1740
gcaagaaagcc ttcttctt aagtgggtgg a catgtgtcc ctgagacagc tcccaaccac 1800
catccgttgc ccctggggcg ggcattggcc cagggtgtgg agacagacg aggccctgt 1860
acgataaagca ccaagagacc ccaaggctcc agatggaaaca ctgagggtgc cggcccttc 1920
caagaaggctt ttccctggaa gtttgcgtt ctcacccatc accatgtgg ttccactag 1980
tgaatttctt cggccccc aaaaagaagga cctcatttct tatagccctc cttgtgttcc 2040
aacacaaccc atggagatgtt ccccttctca ctctgaaatc atccatttgg ggacaaatcc 2100
acatgggggtt ctggatgacat ccacgttcatc atcaatgtt ctttgttgcgtt gcatcttcaac 2160
agagggatggg tccatgtatc tggatgttgc caatgttcc tc tctctgcg aggacaaacc 2220
ctaaacccac accatccac ccaggacagg caggagatgt ttttctt aatccatgtt 2280
taaaaaggaa gctgtggatc cacttagatc agggcggaaac catcttcc acggccaaagc 2340
ttctggccac atgttgaccc tcaccaggcc tgatgttgc atcatgtt gatgttgc 2400
gcatgtggaa gcaatggag ttgtccatc gcttccatc ctccatgc tctgacatgc 2460
agaaaatccac ctttgcgtt ccatgttcc tc gatgttgc gatgttgc gatgttgc 2520
atgttgggtt actggggatccaa ctggatccaa gggatggaaaag tagttccatc aatccatgtt 2580
gttgactatg gtatccacc ttgtatgtt gatgttgc gatgttgc gatgttgc 2640
atggccctgtt ggacttctt ctgcaggagg agagaacgc gcctcttc tggatgttgc 2700
tcaggcttc tggatgttgc tggatgttgc ttccatcatc tggatgttgc caggagagg 2760
ggcttaagggg ctggatccac caaggacgtt cacaaggccaa aacttggg aatggacacc 2820
tgaatccagg gggatgggggtt ggggggggggg ttttgcgtt gatgttgc tccacatgtt 2880
ttctgtatgg gttccatcatc ttcgtatgg aatcaacaga aacacaaa 2938

-continued

```

SEQ ID NO: 46          moltype = AA  length = 485
FEATURE
REGION
1..485
note = protein-arginine deiminase type-3 isoform X2 [Homo
sapiens]
1..485
source
mol_type = protein
organism = Homo sapiens
SEQUENCE: 46
MSVMVLRTOQ PAALFDDHKL VLHTSSYDAK RAQVFHICGP EDVCEAYRHV LGQDKVSYEV 60
PRLHGDEERF PVEGLSFPDA GFTGLISFHV TLLDDSNEDF SASPIFTDTV VFRVAPWIMT 120
PSTLPPLEVY VCRVRNNTCF VDAVAELARK AGCKLTICPQ AENRNRDRIQ DEMELGYVQA 180
PHKTLPVVFDF SPRNGELQDF PYKRILGPDF GYVTREPRDR SVSGLDLSFGN LEVSPPVAN 240
GKEYPLGRIL IGGNLPGSSG RRVTVQVVRDF LHAQKVQPPV ELFVDWLAVG HVDEFLSFVP 300
APDGKGFRML LASPGACFKL FQEOKCGBH RALLFQGVVD DEQVKTISIN QVLSNKDLIN 360
YNKFVQSCID WNREVLKREL GLAECIDI DI PQLFKTERKK ATAFFPDLVN MLVLGKHLGI 420
PKPFGTIING CCCCLEEKVRS LLEPLGLHCT FIDDFTPYHM LHGEVHCGTN VCRKPFSFKW 480
WNMVP
485

SEQ ID NO: 47          moltype = DNA  length = 2938
FEATURE
misc_feature
1..2938
note = PREDICTED: Homo sapiens peptidyl arginine deiminase
3 (PADI3), transcript variant X2, mRNA
source
1..2938
mol_type = unassigned DNA
organism = Homo sapiens
SEQUENCE: 47
agaaatggat ggatgtgact gtgtgctaat aaaactttat ttatacAACAG aggcaaggtagg 60
ccagatttgg cccacagttc ataatgtgtc gatcctgacc taggcggaaaa gagaacccaa 120
atatggaaat gttgaagaac ttgggactga attatgtgg aacttgggtc cctggggagtt 180
aaagaggaaa ggtcgagccg cagtgggtct gggggccccag tggatgtggc ggcacatctgc 240
tggtaactg tgaccgtgat gatccgagct gtgtatgtcca ggacaattgt gaccagcacg 300
tgactgtcctt gcaagacatgt ctgtcatgtgt cctgcggacg caggggccctg 360
cagecccttt tcatgtggac acacttgtcc tccataccctc cagctatgtat gccaaacccgg 420
cacaggcttt ccacatctgc ggtcttgagg atgtgtgtga ggcctataagg catgtgtctgg 480
gccaagataa ggtgtcttat gaggtacccc gcttgcattgg ggtatggggag cgcttcttcg 540
tggaggcgt ctccctccctt gatgcgggt tcacaggact catctcttc catgtcaactc 600
tgctggacga ctcacacggat gatttctcg catccccat ttcactgtac actgtgtgt 660
tccggatgtgc accctgtatc atgacccca gcacactgtcc accccctagag gtgtatgtgt 720
gccgtgttag gaacaacacg tgtttggatgg atgcgggtggc agagctggcc aggaaggccg 780
gtctgcacgt gaccatctgc ccacaggccg acaacccgcaaa cgaccgcgtgg atccaggatg 840
agatggagct gggctacgtt cagggccgcg acaagacccgcg cccgggtggc tttgactccc 900
caaggaatgg ggaactgtcag gatccccctt acaaaaaaat cctgggttcca gattttgggt 960
acgtgaatcg ggaaccacgc gacaggctgtc tgatgtggctt ggactctttt gggaaacctgg 1020
aggtcagccg tccatgtgtt gccaatgggg aagagtatgggg cctggggaggg atccctcattt 1080
ggggcaactt gctctgggtca agtggccgcg ggttcacccca ggtgggtgggg gacttccttc 1140
atgccccagaa ggtgcagccc cccgtggago tctttgtggta ctgggtggcc gtgggcctatg 1200
tggatgagtt tctgagttt gtccctgccc ccgtatggggaa gggcttccgg atgctcttgg 1260
ccagccctgg ggcctgttca aagcttccc agggaaacgca gaagtgtggc cacggggagg 1320
ccctctgtt ccagggggtt gttgtatgtatc agcagggtcaa gaccatctcc atcaaccagg 1380
tgcctctccaa taaagacccatc atcaactaca ataaatgttgcg ctagagctgc atgcacttgg 1440
accgtgaggat gctgaacggg gagctggggc tggcagatgt tgacatcatt gacatccac 1500
agctcttcaa gaccgtgggg aaaaaaggccaa cggccctttt ccctgtactt gtgaacatgc 1560
tggtgctggg gaaacgttgc ggcacatgttcc agccctttgg gcccacatc aatggctgt 1620
gtctgcctgg gggaaagggtt ctagccctgc tgaggccgcg gggccctccac tgccatctca 1680
tttgcgttactt cactccatc acatgtgtc atggggaggat gcaactgtggc accaatgtgt 1740
gcagaaaaggcc ctctcttttca aagtgtgtgg acatgggtggc ctggagacgc tcccaccac 1800
catccctgttcc ccctggggccg ggcattggcc cagggtgggg agacagagac agggccctgt 1860
acgataagacc ccaagagacc ccaaggcttc agatggaaacca ctgagggttgc ccgtcccttc 1920
cagaagecctt ttccctggaa gtgtccatgc ctcacactgtca accccatgtgg ttctcagact 1980
tgaatcttctt cggccccccca aaaagaaggaa cctcattttc tatagcttctt cctgtatttc 2040
aacacaaccc atggagatgtt ccccttctca ctctgaaatc atccatgggg ggacaaatcc 2100
acatgggggtt ctgatggatcat ccatggccat cttgtctgtt gcatcttcaac 2160
agagggaaaggaa tccatgtatc tgctttgttc caattgtttcc ctctgtgtc agggacaaacc 2220
ctaaaaccac accactccac gcaggacagg caggagatgt ttttcttaaa gcctccccca 2280
taaaaaggaa gctgtgttgc cacttagatc agggcggac acatcttccac cccggccaacg 2340
tccctggccatg atgtgttgcgc tcacccaggcg tgagctgtca catagtggaa gcttcttagat 2400
gcatgtggaa gcaatggagag ttgtccctta gccttataaa ctccccatgtca tctgtacatgc 2460
agaaatccac ctttgcgttca aatcttccctg gaatttcttgg gggacaaag tatctggggg 2520
attgttgggtt actaggggaa ctgggttacaa ggggtggaaag tagttcccat aatacacatgt 2580
gttgactatg gtgtatccacc ttgtgtatgtt taatattagg tgcgttggaga aggttgcgttc 2640
attggcccttgc ggacttctctt ctgcaggagg agagaacgcg gcctctcc tggattggc 2700
tcaggctctc tggtggccctt tggtcagctt ttccacatcc tgcgttgcgtc caggagagg 2760
ggctaaaggaa ctggatccac caaggcgttccac cacagccggaa aaactctgggaa atgaaccac 2820

```

-continued

tgaattcagg ggatgggggt gggggggcg ttctcgagggt gtgtccacgc tacacgtgtg 2880
 ttctgtatgg gtccagctgc gttccatca ctcgctaata aatcaacaga aacacaaa 2938

SEQ ID NO: 48
FEATURE
REGION
source
SEQUENCE: 48

moltype = AA length = 485
 Location/Qualifiers
 1..485
 note = protein-arginine deiminase type-3 isoform X2 [Homo sapiens]
 1..485
 mol_type = protein
 organism = Homo sapiens

SEQ ID NO: 49
FEATURE
misc_feature
source
SEQUENCE: 49

moltype = DNA length = 1452
 Location/Qualifiers
 1..1452
 note = PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PAD13), transcript variant X3, mRNA
 1..1452
 mol_type = unassigned DNA
 organism = Homo sapiens

SEQ ID NO: 50
FEATURE
REGION
source
SEQUENCE: 50

moltype = AA length = 421
 Location/Qualifiers
 1..421
 note = protein-arginine deiminase type-3 isoform X3 [Homo sapiens]
 1..421
 mol_type = protein
 organism = Homo sapiens

SEQ ID NO: 51
moltype = DNA length = 1452

-continued

FEATURE	Location/Qualifiers
misc_feature	1..1452
	note = PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X3, mRNA
source	1..1452
	mol_type = unassigned DNA
	organism = Homo sapiens
SEQUENCE: 51	
cctaaggggc ctcagggca gtgttgggt tggcgccac agctaagtcc aacaccagca	60
tgtcgctgca gagaatcgtg cgtgtgtccc tggagcatcc caccagcgcg gtgtgtgtgg	120
ctggcgcttgaa gaccctcgta gacattatgcgtt ggtcagtgcc tgaggcaca gaaatgtttg	180
aggctctatgg gacgccttgcgtt gttggacatctt acgtctctcc caacatggag agggggccggg	240
agcgtgcaga caccaggcggtt ggccgttgcgtt acgcgactttt ggagatcatc gtggcatga	300
actccccccatgg caatgacccca aacgcacagcc atgttcagat ttccctaccac tccagccatg	360
agectctgttgcgtt gctggccat tgggtgtctt accttcaccc tttgtgacatc tctctggatt	420
gccccatggaa ctgtgaggga aggaggaca ggaactttgtt agacaacggg cagtgggtct	480
ggggggccca gggatgttgcgtt ggcacatcttgcgtt tggtgaactt tgaccgtgtat gatccgagct	540
gtgtatgttca ggacaattgtt gaccgcacgg tgcaactgtt gcaagacatcg gaagacatgt	600
ctgtcatgtt gctggggacgtt cagggccctgtt cagcccttgcgtt tgatgaccac aaacttgc	660
tccataccctc cagctatgtt gccaacacggg cacaggtttt ccacatctgc ggtcttgagg	720
atgtgtgtga ggcctatagg catgtgttgcgtt gccaagatata ggtgttcttat gaggtacccc	780
gtttgtatgg ggtatgttgcgtt cgatcttgcgtt tggaaaggccctt gtccttcctt gatgccggct	840
tcacaggactt catcttccttc catgttacttgcgtt tgctggacatc ctccaaacggg gattttctcg	900
catccccatccatgtt gatgtgttgcgtt tccggatgttgcgtt accctggatc atgacgc	960
gcactctgtcc acccccttagag gtgtatgtgtt gccgtgttgcgtt gaacaacacgg tttttgtgg	1020
atgcgggttgcg agagctggcc aggaaggccg gctgcaacgtt gaccatctgc ccacaggccg	1080
agaacccgcaaa cgaccgttgcgtt atccaggatgtt agatggagctt gggctacgtt caggcgcgc	1140
acaagacccctt cccgggttgcgtt ttgttgcgttcc caaggaatgtt ggaactgcgtt gattttccctt	1200
acaaaagaat cctgggtcaag tggccggagggtt gtcacccagg tgggtggggat cttctccat	1260
gcccggaaagg tgcagcccccc cgtggagctt tttgtggactt ggttggccgtt gggccatgtt	1320
gatgagtttgcgttgcgtt gatggggatgtt gatggggatgttccggat gtcctgttgcgtt	1380
agccctgggg cttgttcaatgttccat gttttccat gaaaaggcaga aatgttgcgtt cggggaggcc	1440
cttctgttccat gttttccat gaaaaggcaga aatgttgcgtt cggggaggcc	1452
SEQ ID NO: 52	moltype = AA length = 421
FEATURE	Location/Qualifiers
REGION	1..421
	note = protein-arginine deiminase type-3 isoform X3 [Homo sapiens]
source	1..421
	mol_type = protein
	organism = Homo sapiens
SEQUENCE: 52	
MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTGPVDI YVSPNMERGR	60
ERADTRRWRF DATLEIIVVM NSPSNDLNDHS HVQISYHSSH EPLPLAYAVL YLTCVDISLD	120
CDLNCEGRQD RNFVDKRQWV WGPSPGYGGIL LVNCDRDDPS CDVQDNCDQH VHCLQDLED	180
SVMLVRLTQGP AALFDDHKLV LHTTSSYDAKR AQVFHICGPE DVCEAYRHVL QGDKVSYEVP	240
RHGDEERFF VEGLSPFDAG FTGLISPHVTT LLDDSNEDPS ASPIFTDTVV FRVAPWIMTP	300
STLPPLLEVYY CRVRNNTCFV DAAVELARKA GCKLTICPQA ENRNDRWIQD EMELGYVQAP	360
HKTLPVVFDS PRNGELQDFP YKRILVKWPQ GHPGGAGLPP CPEGAAPRGA LCGLVGRGPG	420
G	421
SEQ ID NO: 53	moltype = DNA length = 1452
FEATURE	Location/Qualifiers
misc_feature	1..1452
	note = PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X3, mRNA
source	1..1452
	mol_type = unassigned DNA
	organism = Homo sapiens
SEQUENCE: 53	
cctaaggggc ctcagggca gtgttgggt tggcgccac agctaagtcc aacaccagca	60
tgtcgctgca gagaatcgtg cgtgtgtccc tggagcatcc caccagcgcg gtgtgtgtgg	120
ctggcgcttgaa gaccctcgta gacattatgcgtt ggtcagtgcc tgaggcaca gaaatgtttg	180
aggctctatgg gacgccttgcgtt gttggacatctt acatctctcc caacatggag agggggccggg	240
agcgtgcaga caccaggcggtt ggccgttgcgtt acgcgactttt ggagatcatc gtggcatga	300
actccccccatgg caatgacccca aacgcacagcc atgttcagat ttccctaccac tccagccatg	360
agectctgtcc cttggccat tgggtgtctt accacacccgtt tttgtgacatc tctctggatt	420
gccccatggaa ctgtgaggga aggaggaca ggaactttgtt agacaacggg cagtgggtct	480
ggggggccca gggatgttgcgtt ggcacatcttgcgtt tggtgaactt tgaccgtgtat gatccgagct	540
gtgtatgttca ggacaattgtt gaccgcacgg tgcaactgtt gcaagacatcg gaagacatgt	600
ctgtcatgtt gctggggacgtt cagggccctgtt cagcccttgcgtt tgatgaccac aaacttgc	660
tccataccctc cagctatgtt gccaacacggg cacaggtttt ccacatctgc ggtcttgagg	720
atgtgtgtga ggcctatagg catgtgttgcgtt gccaagatata ggtgttcttat gaggtacccc	780
gtttgtatgg ggtatgttgcgtt gccaagatata ggtgttcttat gaggtacccc	840
tcacaggactt catcttccttc catgttacttgcgtt tggtggacatc ctccaaacggg gattttctcg	900

-continued

```
catccccatat cttcaactgac actgtggtgt tccgagtgcc accctggatc atgacgcca 960
gcactctgcc accccatagag gtgtatgtg gcccgtgtgg gaacaacacg tggtttgtgg 1020
atgcgggtggc agagctggcc aggaaggccg gctgcaagct gaccatctgc ccacaggccg 1080
agaaccgcoaa cgaccgctgg atccaggatg agatggagct gggctacgtt caggccgc 1140
acaagacccct cccgggtggtc tttgactccc caaggatgg ggaactgcag gatttccctt 1200
acaaaagaat cctggtaaag tggccgcagg gtcacccagg tggtggggaa ctteccat 1260
gcccagaagg tgcaaaaaaa cgtggagact tttgtggact gggtggccgt gggccatgtg 1320
gtatgagttc tgatgtttgt ccctggcccc gatggggaa gtttccggat gtcctggcc 1380
agccctgggg cctgctcaa gcttccag gaaaagcaga agtggccca cggggggcc 1440
cttcgttcc ag 1452
```

```
SEQ ID NO: 54 moltype = AA length = 421
FEATURE Location/Qualifiers
REGION 1..421
note = protein-arginine deiminase type-3 isoform X3 [Homo sapiens]
source 1..421
mol_type = protein
organism = Homo sapiens
SEQUENCE: 54
MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTPGVDI YISPNMERGR 60
ERADTRRWRF DATLEIIVVM NSPSNDLNDS HVQISYHSSH EPLPLAYAVL YHTCVDISLD 120
CDLNCEGRQD RNFVDKQWV WGPSPGYGGIL LVNCDRDDPS CDVQDNCDQH VHCLQDLED 180
SVMVLRTQGP AALFDDHKLV LHTSSYDAKR AQVFHICGPE DVCEAYRHVL GQDKVSYEVP 240
RLHGDEERFF VEGLSFPDAG FTGLISFHVT LLDDSNEDFS ASPIFTDTVV FRVAPWIMTP 300
STLPPLVEYY CRVRNNTCFV DAAVAELARKA GCKLTICPQA ENRNDRWIQD EMELGYVQAP 360
HKTLPPVFDs PRNGELQDFP YKRLILVKWPQ GHPGGAGLPP CPEGAAPRGA LCGLVGRGPC 420
G 421
```

```
SEQ ID NO: 55 moltype = DNA length = 1452
FEATURE Location/Qualifiers
misc_feature 1..1452
note = PREDICTED: Homo sapiens peptidyl arginine deiminase
            3 (PADI3), transcript variant X3, mRNA
source 1..1452
mol_type = unassigned DNA
organism = Homo sapiens
SEQUENCE: 55
cctaaggggc ctcaggggca gtgtgggggt tggccggccac agctaagtcc aacaccagca 60
tgtcgctgca gaaatcgta cgtgtgtccc tggagcatcc caccagcgcc gtgtgtgg 120
ctggcgctgg a gaccctcgta gacattatg ggtcagtggc tgagggcaca gaaatgttg 180
aggtctatgg gacgcctggc gtggacatct acatctctcc caacatggag agggggccgg 240
agcgtcaga caccaggccg tggcgcttg acgcgactt ggagatcatc gtggtcatga 300
actcccccac caatgacccg aacgcacccg atgttcagat ttcttaccac tccagccatg 360
agcctctggc cttggcttat gccgtgtctt acctcacccg tggtagacatc tctctggat 420
gcgacacctaa ctgtggggaa aggccggaca ggaactttgt agacaagccg cagtgggtct 480
ggggggcccg tgggtatggc ggcacatcttgc tggtagactt tgaccgtgtat gatccgagct 540
gtgtatgtca gcaacatgtt gaccggaca tgcactgcctt gcaagacactt gaagacatgt 600
ctgtcatgtt cttggcgacg caggccccctg cagcccttctt tgatgaccac aaacctgtcc 660
tccataccctc cagctatgtat gccaaacccgg cacaggtctt ccacatctgc ggtcttggg 720
atgtgtgtca ggcctatagg catgtgtgg gccaagatata ggtgtccat gaggttacccc 780
gcttgcattgg ggttggggat cgttcttccg tggaaaggccct gtccttccct gatgccgct 840
tcacaggact catctcttc catgtcactc tggtagggca ctccaaacgg gatttctcg 900
catccccatat cttcaactgac actgtggtgt tccgagtgcc accctggatc atgacgcca 960
gcactctgcc accccatagag gtgtatgtg gcccgtgtgg gaacaacacg tggtttgtgg 1020
atgcgggtggc agagctggcc aggaaggccg gctgcaagct gaccatctgc ccacaggccg 1080
agaaccgcoaa cgaccgctgg atccaggatg agatggagct gggctacgtt caggccgc 1140
acaagacccct cccgggtggtc tttgactccc caaggatgg ggaactgcag gatttccctt 1200
acaaaagaat cctggtaaag tggccgcagg gtcacccagg tggtggggaa ctteccat 1260
gcccagaagg tgcaaaaaaa cgtggagact tttgtggact gggtggccgt gggccatgtg 1320
gtatgagttc tgatgtttgt ccctggcccc gatggggaa gtttccggat gtcctggcc 1380
agccctgggg cctgctcaa gcttccag gaaaagcaga agtggccca cggggggcc 1440
cttcgttcc ag 1452
```

```
SEQ ID NO: 56 moltype = AA length = 421
FEATURE Location/Qualifiers
REGION 1..421
note = protein-arginine deiminase type-3 isoform X3 [Homo sapiens]
source 1..421
mol_type = protein
organism = Homo sapiens
SEQUENCE: 56
MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTPGVDI YISPNMERGR 60
ERADTRRWRF DATLEIIVVM NSPSNDLNDS HVQISYHSSH EPLPLAYAVL YLTCVDISLD 120
CDLNCEGRQD RNFVDKQWV WGPSPGYGGIL LVNCDRDDPS CDVQDNCDQH VHCLQDLED 180
```

-continued

SVMLRTQGP AALFDDHKLV LHTSSYDAKR AQVFHICGPE DVCEAYRHVL QDKVSYEVP	240
RLHGDEERFF VEGLSFPDAG FTGLISFHVT LLDDSNEDFS ASPIFTDTVV FRVAPWIMTP	300
STLPLEVYV CRVRNNTCFV DVAELARKA GCKLTICPQA ENRNDRWIQL EMELGYVQAP	360
HKTLPVVFDS PRNGELQDFP YKRILVKWPQ GHPGGAGLPP CPEGAAPRGA LCGLVGRGPGC	420
G	421
 SEQ ID NO: 57	moltype = DNA length = 1452
FEATURE	Location/Qualifiers
misc_feature	1..1452
	note = PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X3, mRNA
source	1..1452
	mol_type = unassigned DNA
	organism = Homo sapiens
SEQUENCE: 57	
cctaaggggc ctcagggca gtgttgggt tggcgccac agctaagtcc aacaccagca	60
tgtcgctgca gagaatcgta cgtgtgtccc tggagcatcc caccagcgcg gtgtgtgtgg	120
ctggcgctgaa gaccctcgta gacattttatg ggtcagtgcc tgagggcaca gaaatgtttg	180
aggtctatgg gacgcctggc gtggacatct acatctctcc caacatggag agggggccggg	240
agcgtgcaga caccaggcgg tggcgcttgc acgcgacttt ggagatcatc gtggctatga	300
actccccccag caatgacctc aacgacagec atgttcagat ttccattaccac tccagccatg	360
agcctctggc cctggctatg gccgtgtccc acctcacctg tggtgacatc tctctgttt	420
gcgacacctgaa ctgtggggaa aggccggaca ggaactttgtt agacaaggcgg cagtgggtct	480
ggggggcccg tgggtatggc ggcattttgc tggtgatggc tgaccgtat gatccggat	540
gtgtatgtcca ggacaattgtt gaccagcaag tgcaactgcct gcaagacccg gaagacatgt	600
ctgtcatgtt cctggcgacg cagggccctt cagccctttt tgatgaccac aaacttgtcc	660
tccataccctc cagctatgtat gccaaacccggg cacaggatcc ccacatctgc ggtctgtgg	720
atgtgtgtga ggcctatagg catgtgtggc gccaagataaa ggtgtcttat gaggtacccc	780
gcttgcattgg ggatggaggag cgcttcttcg tggaaaggccc gtccttcctt gatccggct	840
tcacaggact catctccctt catgtactc tgctggacga ctccaaacgg gattttctcg	900
catccctatc cttcaactgac actgtgtgtt tccgagttttt accctggatc atgaegccca	960
gcactctggc acccccttagag tggtatgtt gccgtgtggaa gacaacacgg tggtttgtgg	1020
atgcgggtggc agagctggcc aggaaggccc gctgcaagct gaccatctgc ccacaggccc	1080
agaaccgcua cgaccgctgg atccaggatg agatggatc gggctacgtt caggcgcgc	1140
acaagaccct cccgggtgtt tttgactccc caaggaatgtt ggaactgcgg gattttccctt	1200
acaaaaagat cttggtcaag tggccggcagg gtcacccagg tggtgccggaa cttccat	1260
gcccagaagg tgcagcccc cgtggagctc ttgtggact ggttggcgtt gggccatgt	1320
gatgagtttgc tgagctttgt ccctggcccc gatggggaaagg gtttccggat gtcctggcc	1380
agccctgggg cctgctcaa gctttccag gaaaaggcaga agtggccca cgggaggccc	1440
ctccctgttcc ag	1452
 SEQ ID NO: 58	moltype = AA length = 421
FEATURE	Location/Qualifiers
REGION	1..421
	note = protein-arginine deiminase type-3 isoform X3 [Homo sapiens]
source	1..421
	mol_type = protein
	organism = Homo sapiens
SEQUENCE: 58	
MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTGVDI YISPNNMERGR	60
ERADTRRWRF DATLEIIVVM NSPSNDLNDs HVQISYHSSH EPLPLAYAVL YLTCVDISLD	120
CDLNCEGRQD RNFVDKRQWV WGPSCYGGJIL LVNCDRDDPS CDVQDNCDQH VHCLQDLED	180
SVMLRTQGP AALFDDHKLV LHTSSYDAKR AQVFHICGPE DVCEAYRHVL QDKVSYEVP	240
RLHGDEERFF VEGLSFPDAG FTGLISFHVT LLDDSNEDFS ASPIFTDTVV FRVAPWIMTP	300
STLPLEVYV CRVRNNTCFV DVAELARKA GCKLTICPQA ENRNDRWIQL EMELGYVQAP	360
HKTLPVVFDS PRNGELQDFP YKRILVKWPQ GHPGGAGLPP CPEGAAPRGA LCGLVGRGPGC	420
G	421
 SEQ ID NO: 59	moltype = DNA length = 1452
FEATURE	Location/Qualifiers
misc_feature	1..1452
	note = PREDICTED: Homo sapiens peptidyl arginine deiminase 3 (PADI3), transcript variant X3, mRNA
source	1..1452
	mol_type = unassigned DNA
	organism = Homo sapiens
SEQUENCE: 59	
cctaaggggc ctcagggca gtgttgggt tggcgccac agctaagtcc aacaccagca	60
tgtcgctgca gagaatcgta cgtgtgtccc tggagcatcc caccagcgcg gtgtgtgtgg	120
ctggcgctgaa gaccctcgta gacattttatg ggtcagtgcc tgagggcaca gaaatgtttg	180
aggtctatgg gacgcctggc gtggacatct acatctctcc caacatggag agggggccggg	240
agcgtgcaga caccaggcgg tggcgcttgc acgcgacttt ggagatcatc gtggctatga	300
actccccccag caatgacctc aacgacagcc atgttcagat ttccattaccac tccagccatg	360
agcctctggc cctggctatg gccgtgtccc acctcacctg tggtgacatc tctctgttt	420
gcgacacctgaa ctgtggggaa aggccggaca ggaactttgtt agacaaggcgg cagtgggtct	480

-continued

ggggggcccg	tgggttatggc	ggcatcttgc	tggtaactg	tgaccgtat	gatccgagct	540
gtgatgtcca	ggacaattgt	gaccaggacg	tgcaactgcct	gcaagactg	gaagacatgt	600
ctgtcatgtt	cctggggacg	caggcccctg	cagecccttt	tgatgaccac	aaacttgtcc	660
tccataacct	cagctatgt	gccaaccccc	cacaggctt	ccacatctgc	ggtcctgagg	720
atgtgtgtga	ggcctatagg	catgtgttgg	gccaagataa	ggtgtcttat	gaggtacccc	780
gettgcattgg	ggatgaggag	cgttcttcgg	tggaaaggcc	gtccttcctt	gatgecggt	840
tcacaggact	catttccttc	catgtcaact	tgctggacg	cttccaaacgg	gattttctcg	900
catccccat	cttcaactgac	actgtgttgt	tccgagtggc	accctggatc	atgacgcccc	960
geactctgcc	acccttagag	gtgtatgtgt	gcccgtgtgg	gaacaacaacg	tgtttgtgg	1020
atgcgggtgc	agagctggcc	aggaaaggcc	gtgcacaaqgt	gaccatctgc	ccacaggccg	1080
agaacccgaa	cggccatgg	atccaggatg	agatggagat	gggttacgtt	caggccgcgc	1140
acaagacccct	cccggtggc	tttgcattcc	caaggatgg	ggaactcgac	gattttccctt	1200
acaaaaagaat	cctggtcaag	tggccgcagg	gtcacccagg	tggtggeggg	cttccctccat	1260
gccccagaaagg	tgcagcccc	cgtggagact	tttgtggact	ggttggccgt	gggcctatgt	1320
gatgagtttgc	tgcgtttgt	cctggggccc	gatgggaagg	ggttccggat	gtctctggcc	1380
agccctgggg	cctgttcaa	gtcttccag	gaaaagcaga	agtgtggcca	cgggaggggc	1440
ctctgttcc	ag					1452

SEQ ID NO: 60	moltype =	length =				
SEQUENCE: 60						
000						
SEQ ID NO: 61	moltype = DNA	length = 2267				
FEATURE	Location/Qualifiers					
misc_feature	1..2267					
	note = Homo sapiens peptidyl arginine deiminase 4 (PADI4),					
	mRNA					
source	1..2267					
	mol_type = unassigned DNA					
	organism = Homo sapiens					
SEQUENCE: 61						
agccagaggg	acgagcttagc	ccgcgcgttgg	cccaggggac	attgtccgt	gtgaccggc	60
agcagccccc	ccatggccgt	tgtgtgttgg	gcacccgtac	tcaagtttgc	atctgcacgt	120
ctgcccctga	ggactgcacg	tccttcagca	tcaacgcctt	cccagggggt	gtcgtggata	180
ttggcccaacgg	ccctccatgg	aagaagaaat	ccacagggtt	ctccacatgg	ccccctggacc	240
ctggggatgtt	gggtgacatgt	acgttggaa	tggccgttgg	tagcacagtc	gaccagaagg	300
ttcagatttc	atactacgg	cccaaaactc	caccaggtaa	agctctactc	tacccatccg	360
gggtggaaat	tccttgcgt	gcagacatca	cccgccacgg	caaagtgaag	ccaaaccaggag	420
ctgtggaaaga	tccagggacc	tggacttgg	ggcccttgcgg	acagggttgc	atccctgtgg	480
tgaactgtgt	ccagagaaat	ctcgaacttt	ctgcacatgg	ctgcggaggat	gtgaagtgc	540
ttgacacgca	agacctgcag	gacatgtgc	tgtatgaccct	gagcacggaa	acccccaagg	600
acttcttcac	aaaccatata	ctgggtgtcc	acgtggccag	gtctgagatg	gacaaagtga	660
gggtgtttca	ggccacacgg	ggcaaaactt	cctccaaatgt	cagcgtatgc	ttgggttccca	720
agtggccctc	tcaatcactt	atggttcccg	gtggaaagca	caacatggac	ttctactgttgg	780
aggcccttcgc	tttccggac	accgttcc	cggggctcat	tacccttacc	atctccctgc	840
tggacacgtc	caacctggag	ctcccccgg	ctgtgtgtt	ccaagacgc	gtgttctcc	900
gctgtggccgc	ctggatcatc	accccaaca	cccgacccccc	cgaggagggt	tacgcgtgc	960
gtatttttgt	aatggggac	tctctgttgc	ctgtgtactc	tctggccatg	aaagccaatg	1020
gcaagctgac	catctggccct	gaggaggaga	acatggatga	ccagtggatg	caggatgaaa	1080
tggagatcggt	ctacatccaa	gccccacaca	aaacgctgc	ctgggtcttc	gactctccaa	1140
ggaacacgggt	ctgtggagg	tttccatca	aacgcgtgt	ggttccatgt	tttggctatg	1200
taactcgagg	gcccccaaca	gggggtatca	gtggacttgg	ctcccttggg	aacctggaaag	1260
ttagcccccc	agtccatgtc	agggggcaagg	aatacccgc	gggcaggatt	ctctccgggg	1320
acagctgtta	tcccagcaat	gacagccgc	agatgcacca	ggccctgcag	gacttcccta	1380
gtgcccacgg	gtgtggggcc	cctgtgttgc	tcttattttgt	ctgggtgtcc	gtggggccacg	1440
tggacggat	cctggatctt	gtggccacgc	ccgacaggaa	ggggctccgg	ctgtctctgg	1500
ccagcccccc	gtccctgtac	aaactgttcc	aggaggacga	gaatgaggcc	cacggggagg	1560
ccctgtgtt	cgaaaggatc	aagaaaaaaa	aacagcagaa	aataaaagaac	attctgtcaa	1620
acaagacattt	aattttttgtt	tggagatgtt	catgcactgg	aaccgcgcgc	1680	
tgtgtggcg	ggggctgggc	ctggccggaa	gtgacatcat	tgacatcccg	1740	
agctcaaaaga	gttctctta	ggggaaatgtt	ttttcccaaa	catgtgttgc	atgtgtgtc	1800
tagggaaaca	cttggggatc	cccaagccct	teggggccgt	catcaacggc	cgctgtgc	1860
tggaggagaa	gtgtgtttcc	ctgtgttgc	cactggggct	ccagtgaccc	ttcatcaacg	1920
acttcttcac	ctaccatcat	aggcttgggg	agggtgtact	cggtccaccc	gtggccagaa	1980
agecccttc	cttcaatgtt	tggaaatgtt	tgcccttgc	ccatcttccc	ttgggtctcc	2040
tcctcttgg	ccagatgttgc	ctgggttcc	ctgtgttgc	caagcaaggag	ctcttgcgttgc	2100
tattgtgttgc	ccctggggcc	ggccagccct	cccaggatgt	ggttgetttc	ttctctctgt	2160
atgtcccagt	ttcccaactt	gaagatccca	acatgggtct	agcactgcac	actcgttct	2220
gtcttaaaaga	gtgtcaataa	agtttttttta	agtcacttttgc	tatcatgt		2267

SEQ ID NO: 62	moltype = AA	length = 663	
FEATURE	Location/Qualifiers		
REGION	1..663		
	note = protein-arginine deiminase type-4 [Homo sapiens]		
source	1..663		
	mol_type = protein		

-continued

organism = Homo sapiens

SEQUENCE: 62

MAQGTLIRVT PEQPTHAVCV LGTLLTQLDIC SSAPECTSF SINASPGVVV DIAHGPAAKK 60
 KSTGGSSTWPL DPGVEVLTLM KVAGSGSTDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
 ITRTGKVKP RAVKDQRTWT WGPGCGQAIL LVNCDRDNLE SSAMDCEDDE VLDSEDLQDM 180
 SMLTSLTKTP KDFFTNHHTLV LHVARSEMMDK VRVFQATRGH LSSKCSVVLG PKWPSHYLMV 240
 PGGKHNMDFY VEALAFPDTD FPGLITLTIS LLDSNTSNEELP EAVVFQDSVV FRVAPWIMTP 300
 NTQQPQEYVA CSIFENEDPL KSVTTLAMKA KCKLTCIPCE ENMDQWMOD EMEIGYIYQAP 360
 HKTLPPVVFDS PRNRGLKEFP IKRVMGPDFG YVTRGPQTGG ISGLDSFGNL EVSPPTVVRG 420
 KEYPLGRILF GDSCYPNSNDs RQMHQALQDF LSAQQVQAPV KLYSDWLSVG HVDEFLSFVP 480
 APDRKGFLRLL LASPRSCYK FQEQQNEGHG EALLFEGIKK KKQQKIKNII SNKTLREHNS 540
 FVERCIDWNR ELLKRELGLA ESDIIDIPQL FKLKEFSKAE AFFPNVNML VLGKHLGIPK 600
 PFGPVINGRC CLEEKVCSSL EPLGLQCTFI NDFFTYHIRH GEVHCGTNVR RKPFESFKWWN 660
 MVP 663

SEQ ID NO: 63

FEATURE

misc_feature

source

moltype = DNA length = 2267
 Location/Qualifiers
 1..2267
 note = Homo sapiens peptidyl arginine deiminase 4 (PADI4),
 mRNA
 1..2267
 mol_type = unassigned DNA
 organism = Homo sapiens

SEQUENCE: 63

agccagaggg acggacttagc ccgacgtatgg cccaggggac attgatccat gtgaccccag 60
 agcagccccac ccatgccgtg tttgtgtctgg gcacccgtac tcagcttgc acatcgacgt 120
 ctggccctga ggactgcacg tccttcagca tcaacgcctc cccagggtg gtcgttgata 180
 ttggccacccg cccttcacgg aagaagaat ccacaggttc ctccacatgg ccctctggacc 240
 ctggggtaga ggtgaccctg acatgtaaag tgccgtatgg tagcacaggc gaccagaagg 300
 ttccatgttc atactacggg cccaagactc caccacgtca agctcttactc taatccacgg 360
 gggtgtggaaat ctccatgtgc gcacacatca cccgcacccg caaagtgtaa ccaaccagag 420
 ctgtgtggaaaga tcacaggaccc tggacccatgg gcccctgtgg acagggtgcc acatccgtgg 480
 tgaactgtgc caagacataatcc tcgcaatctt ctgcacatgg ctgcacggat gatgaaatgc 540
 ttgcacgcga agacactgcga gacatgttc tgatgcaccc gacgcacaa accccaagg 600
 acttcttcac aaaccataca ctgggtctcc acgtggccacgt gtctggatg gacaatgtga 660
 ggggtgttca ggcacacggg gccaatgtt cctccaaatgtt cagcgtatgc ttgggtccca 720
 agtggcccttc tcaatccatgg atggccatgg gttggaaatgc caacatggac ttctatgtgg 780
 agggccctcgc ttccctggac accggacttcc cggggctatc tacccctcacc attccttcgc 840
 tggacacgtc caacccctggag ctccccctggagg ctgtgtgtt ccaagacacgc gtgggtttcc 900
 gctgtggccgc ctggatcatg acccccaaca cccagcccc gcaggagggtg taacgcgtgc 960
 gtatttttca aaatggggatc ttccctgttgc acatgtactc tctggccatgg aaacccaaatgt 1020
 gcaatgtcgc catatgcctt gaggaggaga acatggatgc ctaggtggatc caggatggaaa 1080
 tggagatcggtt ctacatccaa gccccacaca aaacgcgttcc cgtgttccact gactctccaa 1140
 ggaacagaggg cctgaaggag tttccatca aacgcgtgtat ggggtccatgg tttggctatg 1200
 taactcgagg gcccccaaca gggggatcatc gtggacttgc ctcccttggg aaacctggaaatgc 1260
 tgagcccccc acgtcacatggggcaagg aataccggcgtt gggcggatt ctcttcgggg 1320
 acagctgtta tccctggacatc gacacggccg agatgcacca ggccttgcacatgttcc 1380
 gtgcccaacg ggtgcaggcc cctgtgttgc tctatttgc acgtgttccact gttggccacg 1440
 tggccaggatc cctgtgttgc ttggccacac ccggacaggaa ggggtccatgg ctgtcttcgg 1500
 ccagccccccatc aacatgttcc agggacggca gatggggc acgggggg 1560
 ccctgtgttgc cgaaggatggaa aaaaaaaaaa aacacggcaga aataaaaaac atctgttcaa 1620
 acaagacatt gagagaacat aattcatttttgc tggagatgttgc acatgttgc aaccggcggc 1680
 tgctgtggatggc ggagatggggc ctggccggaga gtgacatcatc tgacatcccg cagcttcc 1740
 agctcaaaaatgc gtttctcaatgg gggggatgtt tttccatcaatgg gggggatgttgc 1800
 tagggaaatgc cctggggatcc cccaaatgttgc tggggccatgg catcaacggc cgtgtgtcc 1860
 tggaggagggatc ggtgtgttgc ctgtgttgc acatggccatgg ccgtgtccatc tttatgttca 1920
 acttcttcac ctaccacatc aggcatgggg aggtgtacttgc cggcaccaac gtgcgcac 1980
 agcccttcttc cttaatggggatc tggaaatgttgc tggccatggcc acatgttgc tggccatgg 2040
 tcccttcttcgg ccatgttgc ctggccatggcc tggatgttgc caacggcagatc tttatgttca 2100
 tattgtgttgc ccctggggatcc cccggccatgg gggccatggcc tggatgttgc tttatgttca 2160
 atgtcccaatgc ttccctgttgc gaagatccca acatggatgttgc agcactgtcacttgc 2220
 gctcaatggaa gctgtatggaa agtttttttgc tttatgttgc 2267

SEQ ID NO: 64

FEATURE

REGION

source

moltype = AA length = 663
 Location/Qualifiers
 1..663
 note = protein-arginine deiminase type-4 [Homo sapiens]
 1..663
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 64

MAQGTLIHVT PEQPTHAVCV LGTLLTQLDIC SSAPECTSF SINASPGVVV DIAHGPAAKK 60
 KSTGGSSTWPL DPGVEVLTLM KVAGSGSTDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
 ITRTGKVKP RAVKDQRTWT WGPGCGQAIL LVNCDRDNLE SSAMDCEDDE VLDSEDLQDM 180
 SMLTSLTKTP KDFFTNHHTLV LHVARSEMMDK VRVFQATRGH LSSKCSVVLG PKWPSHYLMV 240
 PGGKHNMDFY VEALAFPDTD FPGLITLTIS LLDSNTSNEELP EAVVFQDSVV FRVAPWIMTP 300

-continued

NTQPPQEYVA	CSIFENEDFL	KSVTTLAMKA	KCKLTICPEE	ENMDDQWMQD	EMEIGYIQQAP	360
HKTLPVVFDS	PRNRGLKEFP	IKRVMGPDPG	YVTRGPQTGG	ISGLDSFGNL	EVSPPVTVRG	420
KEYPLGRILF	GDSCYPSNDS	RQMHQALQDF	LSAQQVQAPV	KLYSDWLSVG	HVDEFLSFVP	480
APDRKGFRLL	LASPRSCYKL	FQEQQNEGHG	EALLFEGIKK	KKQQKIKNIL	SNKTLREHNS	540
FVERCIDWNR	ELLKRELGLA	ESDIIDIPQL	FKLKEFSKAE	AFFPNMVML	VLGKHLGIPK	600
PFGPVGNGRC	CLEEKVCSSL	EPLGLQCTFI	NDFFTYHIRH	GEVHCGTNVR	RKPFSFKWWN	660
MVP						663

SEQ ID NO: 65 moltype = DNA length = 2267
 FEATURE Location/Qualifiers
 misc_feature 1..2267
 note = Homo sapiens peptidyl arginine deiminase 4 (PADI4),
 mRNA
 source 1..2267
 mol_type = unassigned DNA
 organism = Homo sapiens
 SEQUENCE: 65
 agccagaggc acgagcttagc ccgcacgtatg cccaggggac attgatccctt gtgaccgg 60
 agcagcccac ccatgcgcgtg tttgtgtctgg gacacccatgtc tcacgttgc atctgcacgt 120
 ctgccccctga ggactgcacg tccttcagca tcacacgcctc cccagggggtg gtcgtggata 180
 ttggccacaggcc ccttcacaggcc aagaagaaat ccacagggttcc ctccacatgg cccctggacc 240
 ctggggtaga ggtgaccctg acgtatggaaat tgcccaagtgg tagcacaggc gaccagaagg 300
 ttcagatttc atactacggc cccaaactgc caccgttcaa agctctactc tacctcaccgg 360
 ggggtggaaat ctcccttgtc gcacatca cccgcacccgg caaagtggaaat ccaaccaggag 420
 ctgtgaaaga tcaagggacc ttggacctggg gcccctgtgg acagggtgcc atccctgtgg 480
 tgaactgtga cagagacaat ctgcacatcc ttgtccatgttgc ctgcgaggat gatgaagtgc 540
 ttgacacgcca gagacctgtc gacatgtcg tttgtggatc gggccacgg acccccaagg 600
 acttcttcac aaaccatata cttgtgtcc acgtggccacgt gtctggatg gacaaatgt 660
 ggggtgttca ggcacacacgg ggcacaaactgt cttccaaatgtc cggcgtatgc ttgggtccca 720
 agtggccctc tcaactacccgt atgggtcccccgttggaaacca caacatggac ttctacgtgg 780
 aggccctcgc tttccggac accgacttcc cgggggtccat taccctcacc atctccctgc 840
 ttggacacggtc caacctggggcc cttccggagg ctgtgggtgttgc ccaacacggc gtggcttcc 900
 ggcgtggccgc ctggatcatg acccccaaca cccagcccccc gcaggagggtg tacgggtcga 960
 gtattnnntta aataggaggac ttctctgttgc cttctgttgc cttctgttgc 1020
 gcaagcgtcgc catctgtccctt gaggaggaga acatggatgttgc cttctgttgc 1080
 ttggatgttgc ttacatccaa gcccccaacaca aacacgttgc cttctgttgc 1140
 gggaaacaggat cttccatca aacacgttgc cttctgttgc 1200
 taactcgagg gcccccaaca gggggtatca gtggactggc ctcccttggttgc aacctggaaat 1260
 ttagcccccggcc agtccacggc aataccggc aataccggc aataccggc aataccggc 1320
 acagctgttgc tccatccaa gacacccggc agatgttgc aataccggc aataccggc 1380
 gtgtcccaacgg cttccatca aacacgttgc cttctgttgc 1440
 ttggacagttt cttccatca aacacgttgc cttctgttgc 1500
 ccacggccatc gttccatca aacacgttgc cttctgttgc 1560
 ccctgtgttgc cttccatca aacacgttgc cttctgttgc 1620
 acaaaacatc gttccatca aatccatggc tttccatca aatccatggc aataccggc aataccggc 1680
 tgctgttgc gggatgttgc tttccatca aatccatggc aataccggc aataccggc aataccggc 1740
 agtctcaaaatg tttccatca aatccatggc aataccggc aataccggc aataccggc aataccggc 1800
 tagggaaatggc tttccatca aatccatggc aataccggc aataccggc aataccggc aataccggc 1860
 ttggaggaaatggc tttccatca aatccatggc aataccggc aataccggc aataccggc aataccggc 1920
 acttcttcac tttccatca aatccatggc aataccggc aataccggc aataccggc aataccggc 1980
 agcccttcac tttccatca aatccatggc aataccggc aataccggc aataccggc aataccggc 2040
 tccctcttgc cttccatca aatccatggc aataccggc aataccggc aataccggc aataccggc 2100
 tattgtggcc tttccatca aatccatggc aataccggc aataccggc aataccggc aataccggc 2160
 atgtcccaatggc tttccatca aatccatggc aataccggc aataccggc aataccggc aataccggc 2220
 gctctaataatggc tttccatca aatccatggc aataccggc aataccggc aataccggc aataccggc 2280
 gctctaataatggc tttccatca aatccatggc aataccggc aataccggc aataccggc aataccggc 2267

SEQ ID NO: 66 moltype = AA length = 663
 FEATURE Location/Qualifiers
 REGION 1..663
 note = protein-arginine deiminase type-4 [Homo sapiens]
 source 1..663
 mol_type = protein
 organism = Homo sapiens
 SEQUENCE: 66
 MAQGTLILVT PEQPHTAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60
 KSTGSSTWPL DPGEVETLTM KVAGSGTGDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
 ITRTGKVKPT RAVKDQRTWT WGPGCQGAIL LVNCDRDNLE SSAMDCEDE VLDSEDLQDM 180
 SLMTLSKTP KDFFTNHTLV LHVARSEMDK VRVFQATRGK LSSKCSVVLG PKWPSPHYLMV 240
 PGGKHNMDFY VEALAFPDTD FPGLITLTIS LLDTSNLELP EAVVFQDSVV FRVAPWIMTP 300
 NTQPPQEYVA CSIFENEDFL KSVTTLAMKA KCKLTICPEE ENMDDQWMQD EMEIGYIQQAP 360
 HKTLPVVFDS PRNRGLKEFP IKRVMGPDPG YVTRGPQTGG ISGLDSFGNL EVSPPVTVRG 420
 KEYPLGRILF GDSCYPSNDS RQMHQALQDF LSAQQVQAPV KLYSDWLSVG HVDEFLSFVP 480
 APDRKGFRLL LASPRSCYKL FQEQQNEGHG EALLFEGIKK KKQQKIKNIL SNKTLREHNS 540
 FVERCIDWNR ELLKRELGLA ESDIIDIPQL FKLKEFSKAE AFFFPMVNML VLGKHLGIPK 600
 PFGPVGNGRC CLEEKVCSSL EPLGLQCTFI NDFFTYHIRH GEVHCGTNVR RKPFSFKWWN 660
 MVP 663

-continued

```

SEQ ID NO: 67          moltype = DNA  length = 1801
FEATURE
misc_feature           Location/Qualifiers
1..1801
note = PREDICTED: Homo sapiens peptidyl arginine deiminase
4 (PADI4), transcript variant X3, mRNA
source                 1..1801
mol_type = unassigned DNA
organism = Homo sapiens

SEQUENCE: 67
ttaacatcac cgggggtggaa atctccttgat ggcagacat caccgcacc ggcagaatgt 60
agccaaacccg agctgtgaaa gatcagacct gcaggacatg tggctatgtcc ccctgagcac 120
gaagaccccc aaggacttct tcacaacca tacactgggtt ctccacgtgg ccaggctgt 180
gtatggacaaa gtgggggtt ttcaaggccac acggggcaaa ctgttcctca agtgcagcgt 240
agtcttgcgc ccaacttgcct cttcatgttcc cccgtggaa agcacaacat 300
ggacttctac gtggggccct tcgttccc ggacaccgc ttccggggc tcattaccc 360
caccatctcc ctgctggaca cgttcaaccc ggagctcccc gaggctgtgg tggttccaaga 420
cagcgtgttc ttcccggttt ccggccatgtt catgaccccc aacacccaggc ccccgccagg 480
gggtgtacggc tggatgttcc ttgaaaatgtt ggacttctgtt aagtctgttca ctactctggc 540
catgaaagcc aagtgcacgc tgaccatgtt ccgttggggg gagaacatgtt atgaccatgt 600
gatgcaggat gaaatggaga tcgggtatcat ccaaggcccc cacaacccgc tgccgtgg 660
cttcgactct ccaaggaaaca gaggctgttggaa ggagtttccca atcaaaccggc tgatgggtcc 720
agatgttggc tatgttactt cggggcccca aacaggggggat atcgtggac ttggactcttt 780
tgggaacctg gaatgtggcc ccccgatcatc agtgcaggggc aaggttaccc cgttggccag 840
gattcttcc gggggacatgtt gttatcccgtt caatgttccatc cccgtggggg accaggccct 900
gcaggacttcc tcgtgttcc acgagggttca ggccctgttgg aagcttattt ctgactgtt 960
gttccgtggc ctcgtggacg agtcttgcgtt ctttgcgttca gcaaccggaca ggaagggtt 1020
ccggctgttc ctggccaggcc ccaggctctgtt tcacaaacttgc ttccaggaggc agcagaatgt 1080
ggggccacggg gaggccctgc ttgttccaagg gatcaagaaa aaaaaacaggc agaaaaataaa 1140
gaacatttttgc tcaaaacaaaga catttggggaaa acataatttca ttgttggaga gatgttccatc 1200
cttggaaacctg gatgttccatc acggggatgttcc gggccctggcc gagatgttca tcatttgcac 1260
ccggcaggccctt tcacaaacttgc taaggccggaa gttttttccca ccaacatgtt 1320
gaacatgttgc gtgttggggaa agcacttgggg catcccaaaatccgc cccttccggc ccgttcatca 1380
ccggccgtgtc tgcctgggggg agaagggtgtt ttccctgttgc gaggccacttgc gccttccatgt 1440
caccttcatc aacgacttcttcc tcacccatca catcaggcat ggggggggttgc actggggccac 1500
caacgtgtgcg cggggggatgttcc tcttccatca ttgttggggac atgggttccatc gagcccatct 1560
tccttgggttcc tcacccatc tcgttccatc gtcgttgggtt cccttgcgtt gtttggggaa 1620
agagtttgcgtt gtaatattgtt ggcttccctgg gggccggccag cccttccatc agtggcttgc 1680
ttttttttcc tcgttgcgttcc ctttgcgttca tccttgcgtt gtttggggaa 1740
gcacacttgc ttcttgcgttca agaagtttgcataataaagttt tttaagtgcac ttgttccatc 1800
a 1801

SEQ ID NO: 68          moltype = DNA  length = 1748
FEATURE
misc_feature           Location/Qualifiers
1..1748
note = PREDICTED: Homo sapiens peptidyl arginine deiminase
4 (PADI4), transcript variant X8, mRNA
source                 1..1748
mol_type = unassigned DNA
organism = Homo sapiens

SEQUENCE: 68
agctcttactc taccttaccgc ggggtggaaat ctcttgcgttgc gcagacatca cccgcaccgg 60
caaaatgttgcg ccaaccaggat ctgttgcgttca tcagaggacc tggacctggg gcccctgtgg 120
acagggtggcc atcttgcgttca tggactgttca cagagacatc ctggatctt ctggccatgtt 180
cttgcggggat gatgttgcgtt ttggacggcc agacccgttgc gacatgttgc tgatgttcc 240
gagcacacggg acccccaagg acttcttccatc aaaccatcaatc ctgggttccatc acgttggccag 300
gttgcgttgcgtt gacaaatgttca ggggttttca ggccacacggc gtttccggggat gtttgggtt 360
tccaagacac cgttgcgttcc tcgttgcgttcc ctttgcgttcatc gaccccaac acccaggcccc 420
cgaggaggat gtacgtgttc agtgcgttgcgtt ttttttttgcgttcatc gtttggggat 480
cttgcgttgcgtt gaaaggccatc tgcaactgttcc ccatcttgcgttcc tgaggaggat aacatggat 540
accagggttgcgtt gatggatgttcc gtttgcgttcatc acggccacac aaaaacgttgcgtt 600
ccgttgcgttcc tcgttgcgttcatc aggaacacggc gtttgcgttcatc aaacgttgcgtt 660
ttttttttccatc tcgttgcgttcatc gtttggggat gtttggggat gtttggggat 720
acttgcgttgcgtt gaaacccatc tcgttgcgttcatc gtttggggat gtttggggat gtttggggat 780
ttttttttccatc tcgttgcgttcatc gtttggggat gtttggggat gtttggggat 840
aggccctgttca gggacttgcgttcatc agtgcgttgcgttcatc acggccacac aaaaacgttgcgtt 900
acttgcgttgcgtt tcgttgcgttcatc gtttggggat gtttggggat gtttggggat gtttggggat 960
aggccctgttca gggacttgcgttcatc acggccacac aaaaacgttgcgtt 1020
ttttttttccatc tcgttgcgttcatc gtttggggat gtttggggat gtttggggat gtttggggat 1080
aaataaaatgttca aacaaatgttca tcgttgcgttcatc gtttggggat gtttggggat gtttggggat 1140
gcatgttgcgttcatc gggacttgcgttcatc acggccacac aaaaacgttgcgtt 1200
ttttttttccatc tcgttgcgttcatc gtttggggat gtttggggat gtttggggat gtttggggat 1260
acatgttgcgttcatc gggacttgcgttcatc acggccacac aaaaacgttgcgtt 1320
tcgttgcgttcatc gggacttgcgttcatc acggccacac aaaaacgttgcgtt 1380
tccatgttgcgttcatc gggacttgcgttcatc acggccacac aaaaacgttgcgtt 1440
ggggccatccatc tcgttgcgttcatc gggacttgcgttcatc acggccacac aaaaacgttgcgtt 1500

```

-continued

SEQ ID NO: 69 moltype = DNA length = 2081
 FEATURE Location/Qualifiers
 misc_feature 1..2081
 note = PREDICTED: Homo sapiens peptidyl arginine deiminase 4 (PADI4), transcript variant XI, mRNA
 source 1..2081
 mol_type = unassigned DNA
 organism = Homo sapiens

SEQUENCE: 69
 agccagaggc acgacttagc ccgacgtatgg cccaggggac attgatccgt gtgaccggc
 agcagccccacccatgcgtgtgg gcacccgtac tcacgttgc atctgcacgt 60
 ctggccctgaggactgcac tccatcgaca tcaacgcctc cccagggtgt gtcgtggata 120
 ttggccacggccctccacggcc aagaagaat ccacagggtc ctccacatgg cccctggacc 180
 ctggggatgggtgacatgaaatggtgcacgtgg tagccacggc gaccgaaagg 240
 cttagatttatactacggcccaagactccacccgtacatggc accgttacttac tacctccacgg 300
 ggggtggacctgcaggacatgtggctgttgcacccatggc gaagacccca aaggacttct 360
 tcacaaaccatcacactggttcccacgtggccagggtctga gatggacaaa gtgagggtgt 420
 ttccacggccacggccggccaaatgtccatccatgtgcacgtggcgt agtctttgggtcccaatggc 480
 cctctcaactatgtggccctgtggaaagccacatcatggatcttac gtggaggccc 540
 tcgtttccggacaccggatcccccggccatccatccatgtccgtggaca 600
 cgtccaaacctggagttttccatccatgtggatccatccatgtccgtgg 660
 ggccatggatccatccatgtggatccatccatgtccgtggatccatccatgtccgtgg 720
 cggccatggatccatccatgtggatccatccatgtccgtggatccatccatgtccgtgg 780
 ttgaaatgtggacttctgttgcgtgttgc gatctgttgc catggaaatggc aagtccatgttgc 840
 tgaccatctgttgcgtgttgc gagaacatggatccatccatgtccgtggatccatccatgtccgtgg 900
 tcggctacatccaaagccatccatgtggatccatccatgtccgtggatccatccatgtccgtgg 960
 gagggcttgcggatccatccatgtggatccatccatgtccgtggatccatccatgtccgtgg 1020
 gagggccatggatccatccatgtggatccatccatgtccgtggatccatccatgtccgtgg 1080
 ccccaatgttgcgtgttgc gatctgttgc catggaaatggc aagtccatgttgc 1140
 gttatccatgttgcgtgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1200
 agcagggtgcggccctgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1260
 agtccctgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1320
 ccaggctctgttgcgtgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1380
 tggatccatgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1440
 cattggatccatgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1500
 agcggggatccatgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1560
 agaaggcttccatgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1620
 agcaccctggatccatccatgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1680
 agaagggtgttccatccatgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1740
 tccatccatgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1800
 tccatccatgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1860
 tccatccatgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1920
 ggctccctggatccatccatgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 1980
 cagtttccatccatgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 2040
 agaaggcttccatgttgc gggcagatgc accaggcccttgc gaggacttccatccatgtccgtgg 2081

SEQ ID NO: 70 moltype = AA length = 601
 FEATURE Location/Qualifiers
 REGION 1..601
 note = protein-arginine deiminase type-4 isoform XI [Homo sapiens]
 source 1..601
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 70
 MAQGTLIRVT PEQPHTAVCV LGLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60
 KSTGSSSTWP DPGVEVTLTM KVASGSTGDQ KVQISYYGP TPPVKALLYL TGVDLQDMSL 120
 MTLSTKTPKD FFTNHTLVLU VARSEMVKR VFOATRKGKLS SKCSVVLGPK WPSHYLMVPG 180
 GKHNMDFYVE ALAFPDTDPP GLITLTISLL DTSNLLELPAA VVFQDSVVPR VAPWIMPTNT 240
 QPPQEYVACSI FENEDFLKS VTTLAMAKAC KLTICPEEIN MDDQWMQDEM EIGYIQAPHK 300
 TLPVVFDSPR NRGLKEFPII RVMPDPFGYV TRGPQTGGIS GLDSFGNLEV SPPVTVRGKE 360
 YPLGRILFGD SCYPSNDSRQ MHQALQDFLS AQQVQAPVKL YSDWLSVGHV DEFLSFPVAP 420
 DRKGFRLLLA SPRSCYKLFP EQQNEGHGEA LLFEGIKKK QQKIKNIISN KTLREHNSFV 480
 ERCIDWNREL LKRELGLAES DIIDIPQLFK LKEFSKAEAF FPNNMVNMLVL GKHLGIPKPF 540
 GPVINGRCCL EEKVCSLLEP LGLQCTFIND FFTYHIRHGE VHCGTNVRK PFSFKWWNNMV 600
 P 601

SEQ ID NO: 71 moltype = DNA length = 2081
 FEATURE Location/Qualifiers
 misc_feature 1..2081
 note = PREDICTED: Homo sapiens peptidyl arginine deiminase

-continued

source 4 (PADI4), transcript variant XI, mRNA
1..2081
mol_type = unassigned DNA
organism = Homo sapiens

SEQUENCE: 71

```

agccagaggg acgagctagc ccgacgtgg cccaggggac attgatccgt gtgaccagg 60
agcagcccaat ccatgcgtg tgtgtgtgg gcacccgtac tcagttgac atctgcacgt 120
ctgccccctga ggactgcacg tccttcagca tcacacgttc cccagggtt gtcgtggata 180
ttgcccacgg ccctccagcc aagaagaat ccacagggtt ctccacatgg cccctggacc 240
ctggggtaga ggtgacccgt acgttgcggaa tggccagttg tagcacacggc gaccagaagg 300
ttcagatttcaatactacggc cccaagaactt caccatgtca agcttacttc taccttacccg 360
gggtggacct gcaggacatg tcgtgtatgc cccctgacacaa gaagacccccc aaggacttct 420
tcacaaaacca tacactgtgt tcctccatgtt ccaggtctga gatggacaaa gtgagggtgt 480
ttcaggccac acggggcaat ctgtatccca agtgcacgtc agtcttggtt cccaaagtggc 540
cctctcaactt cctgtatgtt cccgtggaa agcacaacatgg ggttgcacatgttgc 600
tcgctttccc ggacacccgc ttccccgggc tcatttaccc caccatccctc ctgtgtggaca 660
cgtnccacccctt ggagctcccc gaggctgtgg tggccatgttgc cggcgtggc ttcccggtgg 720
cgccctgtatgc catggccccc aacaccccgcc ccccgacggaa ggtgttgcgg tgcaactt 780
ttgaaaatgtt ggacttctgtt aagtcaatgtt caactctggc catgaaaggcc aagtgcacgc 840
tgaccatctt ccctgtggat gagaacatgg atgacccatgg gatgttgcggat gaaatggaga 900
tcggcttacat ccaagccccca cacaacacgc tggccgtggc ttccgtactt ccaaggaaaca 960
gaggccgttggaa ggagtttcccc atcaaacgcg tggatgggtcc agattttgc tatgtactt 1020
gaggggccca aacacgggggtt atcgatggc tggactcttctt tggggacccgtt gaaatggggcc 1080
ccccccatgtt actcgccggc aagaaatacc cggctggccat gatgttgcggc ttccgttcc 1140
gttatcccttccatgtt caatgtacgc cggcgtatgc accaggccctt gatgttgcggc 1200
agcagggttca gggccctgtt aagtttctt ctgactgttgc tggccgtggc cacgtggacg 1260
agtcttgcgtt ctggatgttgc gacacccgc ggggggggtt cccgtgttgc ctggccacggc 1320
ccaggcttgc ctacaaactgtt tccttggatg agcacaatgtt gggccacggg gaggcccttc 1380
tggatgttgcggaa gatcaagaa aaaaaacacgtt aagaaataaa aaacattctt tcaaacaaga 1440
catttgcgttccatgtt tttgtggatg gatgttgcgtt cttggacccgc gatgttgcgtt 1500
agcggggatgtt gggccgtggc gaggatgttgc tttgtggatg gatgttgcgtt cttcaagatgtt 1560
aagatgttgcgtt taaggccgtt ggtttttcc ccaacatgtt gaaatgttgc tggatgttgc 1620
agacacccgtt catcccccaag cccttcgggc cccgttcatca cggccgttgc tggccgtggagg 1680
agaaggtgtt ttccctgtt gggccacttgc gcctccatgtt caccatctt aacgttctt 1740
tcacccatgtt catcaatgtt gggccgttgc gatgttgcgtt cttggatgttgc 1800
tctcccttccatgtt tttgtggatg aatggccctt ccccttgcgtt ccttcccttc 1860
ctggccatgtt gtcgttgcgtt ccttgcgtt gttgttgcgtt agatgttgcgtt tggatgttgc 1920
gggtttccatgtt gggccgtggc gggccgtggc gatgttgcgtt ttttttccatgtt tttgtgttgc 1980
caatgttgcgtt ctctgttgcgtt cccaaatgtt ttttttccatgtt gtcgttgcgtt 2040
aagatgttgcgtt ataaatgtt tttgtgttgcgtt a 2081

```

SEQ ID NO: 72 moltype = AA length = 601
FEATURE Location/Qualifiers
REGION 1..601
note = protein-arginine deiminase type-4 isoform XI [Homo sapiens]
source 1..601
mol_type = protein
organism = Homo sapiens

SEQUENCE: 72

```

MAQGTLIRVT PEQPTHAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60
KSTGSSTWPL DPGVEVTLTM KVASGSTDQ KVQISYYGPK TPPVKALLYL TGVDLQDMSL 120
MTLSTKTPKD FFTNHTLVH VARSEMVKR VFQATRGKLS SKCSVVLGPK WPSHYLMVPG 180
GKHNMDFYVE ALAFPDTDFF GLITLTISLL DTSNLELPAA VVFQDSVVFR VAPWIMTPNT 240
QPQQEVYACIS FEFENEDFLKS VTTLAMKAKC KLTCPEEEENN MDDQWMQDEM EIGYIQAAPHK 300
TLPVVFDFSPR NRGFLKEFPIK RVMPGDPFGYV TRGPQTTGGIS GLDSFGNLLEV SPPVTVRGKE 360
YPLGRILFGD SCYPSNDSRQ MHQALQDFLSS AQQVQAPVKKL YSDWLSVGHV DEFLSFVPA 420
DRKGFRLLA SPRSCYKLFQ EQQNNEGHGEA LLFEGIKKK QQKIKNLSN KTLREHNSFV 480
ERCIDWNREL LKRELGLAES DIIDIPQLFK LKEFSKAEAF FPNMVNMLVL GKHLGIPKPF 540
GPVINGRCL EEKVCSLLEP LGLQCTTFIND FFTYHIRHGE VHCGTNVRK PFSFKWWNMV 600
P 601

```

SEQ ID NO: 73 moltype = DNA length = 1969
FEATURE Location/Qualifiers
misc_feature 1..1969
note = PREDICTED: Homo sapiens peptidyl arginine deiminase
4 (PADI4), transcript variant X2, mRNA
source 1..1969
mol_type = unassigned DNA
organism = Homo sapiens

SEQUENCE: 73

```

acagccagag ggacgagcta gcccacatggt gggccatgtt acatgttgcgtt gttgttgcgtt 60
agacgacccatgtt gggccatgtt acatgttgcgtt gttgttgcgtt gttgttgcgtt 120
ctctggccctt gggccatgtt acatgttgcgtt gttgttgcgtt gttgttgcgtt 180
tattggccatgtt gggccatgtt acatgttgcgtt gttgttgcgtt gttgttgcgtt 240
cccttggggta gggccatgtt acatgttgcgtt gttgttgcgtt gttgttgcgtt 300

```

-continued

ggttcgaggat	tcatactacg	gaccgaagac	tccaccagtc	aaagctctac	tctacacctac	360
cgggggtggaa	atctccctgt	gcgcagacat	caccggcacc	ggcaaaagtga	agccaaccag	420
agctgttgcgtt	gatcagagga	cctggactctg	ggggcccttg	ggacagggtg	ccatctctgt	480
ggtgaactgtt	gacagagacat	atctcgaaatc	ttctgcctat	gactgcgagg	atgtgaagt	540
gcttgacagc	aaagacttgc	aggacatgtc	gctgtatgacc	ctgacgacga	agaccccaa	600
ggacttcttc	acaaaccata	cactgggtct	ccacgtggcc	agggtctgaga	tggacaaaagt	660
gagggtgttt	caggccacac	ggggcaact	gtccctccaag	tgcagcgtag	tctttggctt	720
caatgtggccc	tcttcactatc	tgatgtttcc	ccgtggaaag	cacaatcttgc	actttctacgt	780
ggaggccctt	gtttccccgg	acacccgtt	cccggggtct	attaccctca	ccatctccctt	840
gctggacacg	tccaaaccttg	agctccccgg	ggctgtggtg	ttccaagaca	gcgtggctt	900
cccgctggcg	ccctggatca	tgaccccaa	cacccagccc	ccgcaggagg	tgttaacgcgtg	960
cagtatttttt	gaaaatggg	acttccgtta	gtcaatgtact	acttccggca	tgaaaagccaa	1020
gtgcaagctg	accatctgcc	ctgaggagga	gaacatggat	gaccatgttg	tgcaggatgt	1080
aatggagatc	ggctacatcc	aaggccccaa	caaaacgtctg	ccctgtgtt	tgcactctcc	1140
aaggaacaga	ggcttgaagg	agtttcccat	caaacgcgtg	atgggcttcc	gggtgtctt	1200
ggccagcccc	aggtcctgtt	acaaaacttgtt	ccaggaggcag	cagaatgggg	gcccacgggg	1260
ggccctgtgtt	ttcgaaggga	tcagaaaaaaa	aaaacgcagc	aaaataaaga	acatctgttc	1320
aaaacaagaca	tttagagaaac	ataattttatc	tgtggagaga	tgcatccact	ggaaaccgcga	1380
gctgtcaag	cgggagctgg	gcctggccga	gagtgacatc	attgcacatc	ccgcagcttt	1440
caagctcaa	gagtttctca	aggcggaagc	ttttttccca	aacatggtga	acatgttgg	1500
gctaggaaag	cacccggcga	tccccaagcc	cttcggggcc	gtcatcaact	ggccgtgtgt	1560
cctggaggagg	aagggtgttt	ccctgtgttg	ggcactgggc	cttcagtgca	ccttcatcat	1620
cgacttcttc	accttacccaa	tcaggcatgt	ggagggtgcac	tgcggccacca	acgtcgccgag	1680
aaagcccttc	tccttcagaat	ggtggacat	ggtgcctgt	gccccatcttc	cctggcggtt	1740
tctccctctt	ggccagatgt	cgctgggtcc	tctgcgtgt	ggcaagcaag	agctcttctgt	1800
aatatgttg	ctccctgggg	ggccggccac	ctcccaagcg	tgggtgttt	tctttctctgt	1860
tgatgttccca	gtttccccat	ctgaaatgtc	caacatgtc	ctagactgc	acactcgatgt	1920
ctgtctcaag	aagctgtcaat	aaagttttt	taagtcaactt	tgtacatga		1969

```

SEQ ID NO: 74 moltype = AA length = 563
FEATURE Location/Qualifiers
REGION 1..563
note = protein-arginine deiminase type-4 isoform X2 [Homo sapiens]
source 1..563
mol_type = protein
organism = Homo sapiens

SEQUENCE: 74
MAQGTLLIRVT PEQPTHAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60
KSTGTSSTWPL DPGVEVTLTM KVASGSTGDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
ITRTGKVKP T RAVKDQRTWTW WGPGCGQAIL LVNCDRDNLE SSAMCEDDE VLDSEDLQDM 180
SLMLTLSTKTP KDFFTNHTLIV LHVARSEMDK VRVFQATRGK LSSKCSCVVLG PKWPSPHYLMV 240
PGPKHHNMDFY VEALAFPPDTW PFGPLITLTIS LLLPQDGLP EAVVQFQDSVV FRVAPWIMTP 300
NTQPGKVEVA CSIFENEDFL KSVTLLAMKA KCKLTLICPEE ENMDDQWMD EMEIGYIQAP 360
HKTLPVVFDS PRNRGLENKEFP IKRVMGFRL LASPRSCYKL FQEQQNEGHG EALLFEGIJKK 420
KKQOKKINIL SNKTLREHNS FVERCIDWNR ELLKRELGLA ESDIIDIPQL FKLKEFSKAE 480
AFFPNMNLNML VLGKHLGIPK PFGPVINGRC CLEEKVCSSL EPLGLQCTFI NDFFTYHIRH 540
GEVHCGTNTVR KPKEPSFKWNN MVP 563

```

```
SEQ_ID_NO: 75          moltype = DNA  length = 1969
FEATURE                  Location/Qualifiers
misc_feature             1..1969
note = PREDICTED: Homo sapiens peptidyl arginine deiminase
                     4 (PADI4), transcript variant X2, mRNA
source                   1..1969
                         mol_type = unassigned DNA
                         organism = Homo sapiens
```

```

SEQUENCE: 75
acagcccaag ggacgagacta gcccgaatgg acattgtatcc gtgtgacccc 60
agacgacccc accatgcgcg tggctgtgtc gggcaccttg actcagcttg acatctgcag 120
cttgcgcctt gaggactgca cgtcccttcag catcaacggcc tccccagggg ttggctgttgg 180
tattggccac ggccttcag ccaagaagaaa atccacaggt tcctccacat ggccttcggaa 240
ccctggggta gaggtgaccct tgacgtatggaa agtgtggccatgt ggtagcacaacggcag 300
ggttcgatgt tcatactacg gacccaagac tccaccatgc aaagctctac tctaccctcac 360
cggggttggaa attccttcggaa gggcagacat caccggccacc ggcggaaatgtg accaaaccagg 420
agctgtgaaa gatcagaggg cctggacccgtt gggcccttgg acggacgggtt ccattctgtt 480
ggtaactgtt gacagagaca atctcgaatc ttctggccatgt gacttgcgggg attgtatggaaatgt 540
gtttgacacgc gaagacccctgc aggacatgtc gctgtatggcc ctgagcacaacggcccccaa 600
ggacttcttc acaaaccatc cacttgggtgtt ccacgtggcc aggttctggaaatgtg tggacaaatgt 660
gggggttggaa caggccacac gggggcaatgtt gtcctccaaatgt tgcacgttag tcttgggttcc 720
caagtggccct ttcacttacc ttgtatgggtttt cgggtggaaatgtt cacaacatgg actttctgtt 780
ggaggcccttc gttttccccgg acaccgttcc cccggggcttc attaccctca ccatctccct 840
gttggacacgc tccaaaccttgg agtcccccggg ggtctgtgtt ttccaaagaca ggcgtgttcc 900
ccgcgtggcc cccctggatca tgaccccccac caccggccccc cccggagggtt tgtagcgtgtt 960
cgtatttttt gaaaatgttggg acttctgttgg acgttgcgtt actctggccca tggaaatgttgg 1020
gttgcacgttcc accatctggcc ctggaggagggaa acatggat gaccatggat tgcaggatgtt 1080

```

-continued

aatgggagatc	ggctacatcc	aagccccaca	caaaacgtg	cccggtgtct	tcgactctcc	1140
aggaaacaga	ggcctgaagg	agtttcccat	caaacgcgtg	atgggcttcc	ggctgctct	1200
ggccagcccc	aggtcctgtct	acaacttgtt	ccaggagacag	cagaatagg	gccaagggggaa	1260
ggccctgtct	ttcgaaggga	tcagaaaaaaa	aaaacacgacg	aaaataaaga	acatctgtc	1320
aaaacaagaca	tttgagagacat	ataattttatc	tgtggagaga	tgcatcgact	ggaaccggca	1380
gctgtcgaa	cgggagctgg	gcctggccga	gagtgacatc	attgacatcc	cgacgacttt	1440
caagctcaa	gagttctcta	aggcggaaagc	ttttttcccc	aacatggtga	acatgtgttgt	1500
gttagggaa	cacctggcca	tccccaacgg	cttggggccc	gtcatcaacg	gcccgtgtgg	1560
cctggggaggaa	aaggtgtgtt	ccctgtgttgc	gccactgggc	ctcggatgc	ccttcatcaa	1620
cgacttttc	accttaccaca	tcaggcatgg	gggggtgcac	tgccggaccca	acgtcgccgac	1680
aaaggcccttc	tccttcaagt	ggtgaaacat	ggtgccctga	gccccatcttc	cctggcgctcc	1740
tctccctctt	ggccagatgt	cgctgggtcc	tctgcgtgtt	ggcaagcaag	agctttgtgt	1800
aatattgtgg	ctccctgggg	ggggccacgg	ctcccaacgg	tggtttgttt	tctttctctgg	1860
tgatgtccca	gtttttccat	ctgaagatcc	caacatggtc	ctagcaactgc	acactcgttt	1920
ctctcttaaq	aagctcaat	aaaatttttt	taaqtcactt	tgatcatqa		1969

SEQ_ID_NO: 76 moltype = AA length = 563
FEATURE Location/Qualifiers
REGION 1..563
note = protein-arginine deiminase type-4 isoform X2 [Homo sapiens]
source 1..563
mol_type = protein
organism = Homo sapiens

SEQUENCE: 76		Organism = HOMO sapiens
MAQGTLIRVT	PEQPPTHAVCV	LGLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60
KSTGSSTWP1	DPGVEVTLTM	KVASGSTGDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
IITR7TKVSKPT	RAVKDQRTWT	WGPGCGQAIL LVNCMDRNLSE SSAMDCEDDSE VLDSERLDQMM 180
SLMLTSLKTP1	KDFFTNHTLV	LHVARSEMDK VRVFQATRGK LSSKCSVVLG PKWPSPHYLMV 240
PGGKHNMDFY	VEALAPPDTD	FPGLLTLTIS LLDSLNEPL EAVVFQDSVV FRVAPWIMTP 300
NTQQPKEVYA	CSIFENEDFL	KSVTTLAMKA KCKLTICPEE ENMDDQWMDQ EMEIGIYIQAP 360
HKTLPVVFDS	PRNRGLKEFP	IKRVMGFRL LASPRSCYKL FQEQQNEGHG EALLFEGIJKK 420
KKQOKIKN1	SNKTLREHNS	FVERC1CDWNR ELLKRELGLA ESDIIDIPQL FKLFKEFSKAE 480
AFFPNVMNML	VLGKHLG1PK	PFGPVINGRC CLEEKVCSSL EPLGLQCTFI NDFFTYHIRH 540
GEVHCGTNVR	RKPFEKWWNN	MVP 563

```
SEQ ID NO: 77          moltype = DNA  length = 1358
FEATURE                  Location/Qualifiers
misc_feature            1..1358
note = PREDICTED: Homo sapiens peptidyl arginine deiminase
                     4 (PADI4), transcript variant X4, mRNA
source                   1..1358
                         mol_type = unassigned DNA
                         organism = Homo sapiens
```

SEQ ID NO: 78 moltype = AA length = 397
FEATURE Location/Qualifiers
REGION 1..397
note = protein-arginine deiminase type-4 isoform X4 [Homo sapiens]
source 1..397
mol type = protein

-continued

SEQUENCE: 78	organism = Homo sapiens
MAQGTLIRVT PEQPTHAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK	60
KSTGSSTWPL DPGVEVLTLM KVAGSGTGDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD	120
ITRTGKVKPT RAVKDQRTWT WGPGCGQAIL LVNCDRDNLE SSAMDCEDE VLDSEDLQDM	180
SLMLTSLTKTP KDFFTNHTLV LHVARSEMOK VRVFQATRGK LSSKCSVVLG PKWPSPHYLMV	240
PGGKHNMDFY VEALAFPDTD FPGLITLTIS LLDTSNLELP EAVVFQDSVV FRVAPWIMTP	300
NTQOPPQEYVA CSIFENEDFL KSVTTLAMKA KCKLTICPEE ENMDDQWMQD EMEIGYIQAP	360
HKTLPVVFDS PRNRGLKEFP IKRVMILSIG PFYRRRN	397
SEQ ID NO: 79	moltype = DNA length = 1358
FEATURE	Location/Qualifiers
misc_feature	1..1358
	note = PREDICTED: Homo sapiens peptidyl arginine deiminase
	4 (PADI4), transcript variant X4, mRNA
source	1..1358
	mol_type = unassigned DNA
	organism = Homo sapiens
SEQUENCE: 79	
acagccagag ggacgagcta gcccacatgggg acattgtatcc gtgtgacccc	60
agagcagccc acccatgcgg tttgtgtgtcttggcacccatgggg actcaacgcg	120
cttgcgcctt gaggactgca cgttccttcg catcaacgcg tccccaggg tggtcggtga	180
tattggccac ccgcctccatgggg ccaagaagaa atccacaggat ttctccatgggg	240
ccttggttta gaggtgaccctgacatggatggaa agtggccatggtagcacag	300
gtttagtattt tcataactacg gaccaagac tccaccatgggg aaagctctac tctacccac	360
cgggggtggaa atctccttgcgacatgggg caccgcacccggcaaaatgtaa	420
agttgtgtttt gatcggatggatggggcccttggggccatggggccatggggccatgggg	480
gggtgaactgtt gacagagaca atcttcggatccatggggccatggggccatggggccatgggg	540
gtttggacatggggccatggggccatggggccatggggccatggggccatggggccatgggg	600
ggacttccatggggccatggggccatggggccatggggccatggggccatggggccatgggg	660
gagggtgtttt caggccacacggggccatggggccatggggccatggggccatggggccatgggg	720
caagtggccc ttcactaccatggggccatggggccatggggccatggggccatggggccatgggg	780
ggaggccctt gttttcccgacatggggccatggggccatggggccatggggccatggggccatgggg	840
gctggacatggggccatggggccatggggccatggggccatggggccatggggccatggggccatgggg	900
ccggcgtggccatggggccatggggccatggggccatggggccatggggccatggggccatggggccatgggg	960
cgttattttt gaaatggggccatggggccatggggccatggggccatggggccatggggccatggggccatgggg	1020
gttcaagatggggccatggggccatggggccatggggccatggggccatggggccatggggccatggggccatgggg	1080
aatggggccatggggccatggggccatggggccatggggccatggggccatggggccatggggccatggggccatgggg	1140
aagaacatggggccatggggccatggggccatggggccatggggccatggggccatggggccatggggccatgggg	1200
attttatggggccatggggccatggggccatggggccatggggccatggggccatggggccatggggccatggggccatgggg	1260
tccggcgtttt gttccatggggccatggggccatggggccatggggccatggggccatggggccatggggccatggggccatgggg	1320
ggactggact cttttggggccatggggccatggggccatggggccatggggccatggggccatggggccatggggccatggggccatgggg	1358
SEQ ID NO: 80	moltype = AA length = 397
FEATURE	Location/Qualifiers
REGION	1..397
	note = protein-arginine deiminase type-4 isoform X4 [Homo sapiens]
source	1..397
	mol_type = protein
	organism = Homo sapiens
SEQUENCE: 80	
MAQGTLIRVT PEQPTHAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK	60
KSTGSSTWPL DPGVEVLTLM KVAGSGTGDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD	120
ITRTGKVKPT RAVKDQRTWT WGPGCGQAIL LVNCDRDNLE SSAMDCEDE VLDSEDLQDM	180
SLMLTSLTKTP KDFFTNHTLV LHVARSEMOK VRVFQATRGK LSSKCSVVLG PKWPSPHYLMV	240
PGGKHNMDFY VEALAFPDTD FPGLITLTIS LLDTSNLELP EAVVFQDSVV FRVAPWIMTP	300
NTQOPPQEYVA CSIFENEDFL KSVTTLAMKA KCKLTICPEE ENMDDQWMQD EMEIGYIQAP	360
HKTLPVVFDS PRNRGLKEFP IKRVMILSIG PFYRRRN	397
SEQ ID NO: 81	moltype = DNA length = 1842
FEATURE	Location/Qualifiers
misc_feature	1..1842
	note = PREDICTED: Homo sapiens peptidyl arginine deiminase
	4 (PADI4), transcript variant X6, mRNA
source	1..1842
	mol_type = unassigned DNA
	organism = Homo sapiens
SEQUENCE: 81	
acagccagag ggacgagcta gcccacatgggg acattgtatcc gtgtgacccc	60
agagcagccc acccatgcgg tttgtgtgtcttggcacccatgggg actcaacgcg	120
cttgcgcctt gaggactgca cgttccttcg catcaacgcg tccccaggg tggtcggtga	180
tattggccac ccgcctccatgggg ccaagaagaa atccacaggat ttctccatggggccatggggccatgggg	240
ccttggttta gaggtgaccctgacatggggccatggggccatggggccatggggccatgggg	300
gtttagtattt tcataactacg gaccaagac tccaccatgggg aaagctctac tctacccac	360
cgggggtggaa atctccttgcgacatggggccatggggccatggggccatggggccatggggccatgggg	420

-continued

actgtgtaaa	gatcagagga	cctggacctg	gggccttgc	ggacagggtg	ccatcctgct	480
gggtgaactgt	gacagagaca	atctcgaaatc	tcttgccatg	gactgcgggg	atgtatgtaa	540
gcttgacagc	ggagacactgc	aggacatgtc	gctgtatgacc	ctgagcaca	agacccccaa	600
ggacttcctc	acaaaccata	cacttgtgtc	ccacgtggcc	aggctgtgaga	tggacaaga	660
gagggtgttt	caggccacac	ggggcaaaact	gtcttccaag	tgcaagctgt	tcttgggtcc	720
caagtggccc	tctcaactacc	tgtatgttccc	cggtggaaag	cacaacatgg	acttctacgt	780
ggaggccetc	gttttcccg	acacccgactt	ccgggggttc	attaccctca	ccatctccct	840
gctggacacg	tccaaactgg	agtccttccg	ggctgtgttgc	ttccaagaca	gcgtgtgtt	900
ccgcgtggcc	ccctggatca	tgacccccc	cacccagccc	ccgcaggagg	tgtacgcgt	960
cagtatTTT	gaaaatgagg	acttcctgt	gtcagtgtact	actctggcca	tgaaaggccaa	1020
gtgcaagctg	accatctgc	ctgaggagaa	gaacatgtat	gaccgtgg	tgcaggatga	1080
aatggagatc	ggctacatcc	aagcccccaca	caaaacgcgt	cccggtgtt	tcgactctcc	1140
aaggaacaga	ggcctgaagg	agtttcccat	caaaccgcgt	atggaaagcc	ccaggccgtc	1200
gctctgagtc	ctggcagccc	cttggctcaa	ggtgccccagg	gaaaacgcacg	gaaggaggt	1260
gggggttcaag	gtttgttagg	acaaatggct	ttccgacaaag	catctgtact	gtgccccaaag	1320
ggaggggaga	gtggccctgc	ccacacacaa	ggcttcacaa	tgaaaactca	gttaaactca	1380
acaacttgc	ccgagtctcg	cttctctctg	gccaagctgt	gcgcgtggaa	gtggagatgc	1440
agagatgtgt	ttggattcac	agtgtgtact	cagccatggg	gtcgagttca	tcctgtttca	1500
gcagcgttgg	ggcgaacccag	ccagctgtca	ccgtgacttt	gccttcttcc	tgctgtttc	1560
agtccctgtc	gtttcgttgc	ctgtgttcaat	ttttatgtctc	caagataaac	aaaggccaga	1620
ttttctgtat	ccacaagttcc	tgctggacag	acagacggac	acctcgccac	cattttgttc	1680
ccctttccca	cagacaactt	tggatgttgg	caggcagcac	ttggccctca	gctaaggccaa	1740
tctgtgtatt	ccaaactcaa	tcacttcgaa	tgaggccagg	tgtgtgttgc	gacacactgt	1800
atcccagcag	tttggggaggc	caaggaggat	ggatcacatc	ag		1842

SEQ ID NO: 82 moltype = AA length = 392
 FEATURE Location/Qualifiers
 REGION 1..392
 note = protein-arginine deiminase type-4 isoform X5 [Homo sapiens]
 source 1..392
 mol_type = protein
 organism = Homo sapiens
 SEQUENCE: 82
 MAQGTLIRVT PEQPTHAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60
 KSTGSSTWPL DPGEVTLTM KVAGSSTGDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
 ITRTGKVKPT RAVKDQRTWT WGPGCGQAIL LVNCDRDNLE SSAMDCEDE VLDSEDLQDM 180
 SLMLTSLTKTP KDFFTNHTLV LHVARSEMOK VRVFQATRGK LSSKCSVVLG PKWPSPHYLMV 240
 PGGKHNMDFY VEALAFPDTD FPGLTLLTIS LLDTSNLELP EAVVFQDSVV FRVAPWIMTP 300
 NTQPPQEYVA CSIFENEDFL KSVTTLAMKA KCKLTICPPE ENMDDQWMQD EMEIGYIQAP 360
 HKTLPVVFDS PRNRGLKEFP IKRVMEDPRP AL 392

SEQ ID NO: 83 moltype = DNA length = 1842
 FEATURE Location/Qualifiers
 misc_feature 1..1842
 note = PREDICTED: Homo sapiens peptidyl arginine deiminase 4 (PADI4), transcript variant X6, mRNA
 source 1..1842
 mol_type = unassigned DNA
 organism = Homo sapiens
 SEQUENCE: 83
 acagccagag ggacgagacta gcccgcacat ggcgcagggg acattgtatcc gtgtgacccc 60
 agagcagccc accccatgcg tttgtgtgtc qggcaccttgc actcaacgtc acatctgcag 120
 ctctggccctt gaggactgca cgtcttcgcg catcaacgcg tccccagggg ttggctgttgc 180
 tattggccac gggccctccag ccaagaagaa atccacatgt tcctccacat ggccccttgg 240
 ccctgggtt gagggtgacc tgacatgtt ggtggccatg ggttagcacag ggcaccagaa 300
 ggttcagatt tcatactacg gacccaaacat tccaccatgcg aaagcttac tctaccctac 360
 cgggggtggaa atctcttgcg ggcgcacat caccgcgcac ggccaaatgtt agccaaccag 420
 agtgtgtaaa gatcagaggg cctggacatc gggcccttgc ggacagggtg ccatcctgt 480
 ggtgtactgt gacagagaca atctcgatcc ttctgcacat gactgcgggg atgtatgtaa 540
 gcttgacacg gaagacactgc aggacatgtc gctgtatgacc ctgagcaca agacccccaa 600
 ggacttcctc acaaaccata cacttgtgtc ccacgtggcc aggtctgttggacaaga 660
 gagggtgttt caggccacac gggcaaaact gtcttccaag tgcaagctgt tcttgggtcc 720
 caagtggccc tctcaactacc tttgtgttccc cggtggaaag cacaacatgg acttctacgt 780
 ggaggccetc gttttcccg aacccgactt cccgggggttc attaccctca ccatctccct 840
 gctggacacg tccaaactgg agtccttccg ggtgtgttgc ttccaagaca gcgtgtgtt 900
 ccgcgtggcc ccctggatca tgaccccccac cacccagccc ccgcaggagg tttacgcgt 960
 cagtatTTT gaaaatgagg acttcctgtt gtcagtgtact actctggcca tgaaaggccaa 1020
 gtgcaagctg accatctgcg ctgaggagaa gaacatgtt gaccgtgg tgcaagatga 1080
 aatggagatc ggttatcatcc aagccccaca caaaacgcgtc cccgtgggtt tcgactctcc 1140
 aaggaacaga ggcctgttgg agtttccat caaaacgcgtt atggaaagcc ccaggccgtc 1200
 gctctgagtc ctggcagccc ctggcttca ggtgtccagg gaaaacgcacg gaaggaggt 1260
 ggggttcaag gtttgcatttgg acaaattggct ttccgacaaag catctgtact gtgccccaaag 1320
 ggaggggaga gtggccctgc ccacacagaa ggcttcacaa tgaaaactca gttaaactca 1380
 acaacttgc tccgagtctcg cttectctg gccaagctgt gcgcgtggca gtggagatgc 1440
 agagatgtgttgcgattcac agtgtgtact cggccatggg gtcgcgttca tcctgtttca 1500

-continued

-continued

-continued

```

tccggcgtgc aatgacagcc ggcagatgc ccaggccctg caggacttc tcagtgccca 1320
gcaggtgcag gcccctgtga 1340

SEQ ID NO: 90          moltype = AA length = 397
FEATURE             Location/Qualifiers
REGION              1..397
note = protein-arginine deiminase type-4 isoform X4 [Homo
sapiens]
source              1..397
mol_type = protein
organism = Homo sapiens

SEQUENCE: 90
MAQGTLIRVT PEQPTHAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60
KSTGSSTWPL DPGVEVTLTM KVASGSTGDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
ITRTGKVKPT RAVKDQRTWT WGPGCQGAIL LVNCDRDNLE SSAMDCEDDE VLDSEDLQDM 180
SLMLTSLTKTP KDFFTNHTLV LHVARSEMDK VRVFQATRGK LSSKCSVVLG PKWPSPHYLMV 240
PGGKHNMDFY VEALAPPDTD FPGLITLTIS LLDTSNLLELP EAVVFQDSVV FRVAPWIMTP 300
NTQPPPQEYVA CSIFENEDFL KSVTTLAMKA KCKLTICPPE ENMDDQWMQD EMEIGYIQAP 360
HKTLPPVVFDS PRNRGLKEFP IKRVMILSIG PFYRRRN 397

SEQ ID NO: 91          moltype = DNA length = 1340
FEATURE             Location/Qualifiers
misc_feature        1..1340
note = PREDICTED: Homo sapiens peptidyl arginine deiminase
4 (PADI4), transcript variant X5, mRNA
source              1..1340
mol_type = unassigned DNA
organism = Homo sapiens

SEQUENCE: 91
acagccaaag ggacgagcta gccccacat gggccagggg acattgatcc gtgtgacccc 60
agagcagccc acccatgcgg tttgtgtgt gggcaccttg actcagctt acatctgcag 120
ctctgcacctt gaggactgca cgtcccttag catcaacgcg tccccagggg ttgtcggtga 180
tattgcaccc ggcctccatc ccaaaagaaa atccacatg ttctccatcat ggcctctgg 240
ccctgggtta gaggtgaccc tgacatgaa agtggccatg ggttagcacag ggcaccagaa 300
ggttcagatt tcataactcg gacccaaagac tccaccatgc aaagcttac tctaccctac 360
cggggtggaa atctccatgt gcgcagacat caccgcaccc ggccaaagtga agccaaccag 420
agctgtgaaa gatcagagga cctggacatg gggcccttgtt ggacagggtt ccattctgt 480
gttgaactt gacagagaca atctcgaatc ttctggccatg gactggcaggg atgtgaatg 540
gcttgacatc gggacatgcg acatgtgcg tctgtatggcc ctggacacaa agaccccaa 600
ggacttcttc acaaaaccata cactgggtt ccacgtggcc aggtctgaga tggacaagt 660
gagggtgttt caggccacac gggccaaact gtctccaaatg tgcagcgatg tcttgggtcc 720
caagtggccc tctcaatacc tgatggtccc cggtgaaatg cacaacatgg acttctacgt 780
ggaggccctc gttttcccg acacccactt cccggggcttc attaccctca ccattctcc 840
gctggacacg tccaaacttgc agtccccccca ggttgggtt ttccaaagaca gggtgggttt 900
ccgcgtggcg ccctggatca tgaccccaatc caccacggcc ccgcggaggg ttgtacggcg 960
cagtatTTT gaaaatgggg atcttctggaa gtcaatgtacttccatgc tggatggccaa 1020
gtgcaagatc accatctgcc ctggggggaa gaacatggat gaccatggaa tggatggatg 1080
aatggggatc ggctacatcc aagcccccaca caaaacgtgtt cccgtgggtt tcgactctcc 1140
aaggaaacaga ggcctgaagg agtttccatc caaacgtgtt atgattctat caataggccc 1200
atTTTataga agaagaaact gagggtcaga gtggagaatg tggatccaa accccggccag 1260
tccggcgtgc aatgacagcc ggcagatgc ccaggccctg caggacttc tcagtgccca 1320
gcaggtgcag gcccctgtga 1340

SEQ ID NO: 92          moltype = AA length = 397
FEATURE             Location/Qualifiers
REGION              1..397
note = protein-arginine deiminase type-4 isoform X4 [Homo
sapiens]
source              1..397
mol_type = protein
organism = Homo sapiens

SEQUENCE: 92
MAQGTLIRVT PEQPTHAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60
KSTGSSTWPL DPGVEVTLTM KVASGSTGDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
ITRTGKVKPT RAVKDQRTWT WGPGCQGAIL LVNCDRDNLE SSAMDCEDDE VLDSEDLQDM 180
SLMLTSLTKTP KDFFTNHTLV LHVARSEMDK VRVFQATRGK LSSKCSVVLG PKWPSPHYLMV 240
PGGKHNMDFY VEALAPPDTD FPGLITLTIS LLDTSNLLELP EAVVFQDSVV FRVAPWIMTP 300
NTQPPPQEYVA CSIFENEDFL KSVTTLAMKA KCKLTICPPE ENMDDQWMQD EMEIGYIQAP 360
HKTLPPVVFDS PRNRGLKEFP IKRVMILSIG PFYRRRN 397

```

1-55. (canceled)

56. A detection kit comprising:

- (a) a peptidyl arginine deiminase 4 (PAD4) or antigenic fragment thereof;
- (b) a detection probe comprising:
 - (i) an antibody or functional fragment thereof capable of binding to anti-PAD4 IgA; or
 - (ii) an antibody or functional fragment thereof capable of binding to anti-PAD4 IgG; and
- (c) a solid support;
wherein the detection kit is usable in a method for diagnosing or assessing disease severity in a subject having rheumatoid arthritis (RA), and wherein a severity of the RA is based on presence of joint erosion or severe joint erosion in the subject.

57. The detection kit of claim **56**, wherein the detection probe comprises the antibody or functional fragment thereof capable of binding to anti-PAD4 IgA.

58. The detection kit of claim **56**, wherein the detection probe comprises the antibody or functional fragment thereof capable of binding to anti-PAD4 IgG.

59. The detection kit of claim **56**, further comprising a peptidyl arginine deiminase 1 (PAD1) or antigenic fragment thereof; a peptidyl arginine deiminase 2 (PAD2) or antigenic fragment thereof; a peptidyl arginine deiminase 3 (PAD3) or antigenic fragment thereof; a peptidyl arginine deiminase 6 (PAD6) or antigenic fragment thereof; or combinations thereof.

60. The detection kit of claim **56**, wherein the detection probe further comprises a reporter tag.

61. The detection kit of claim **60**, wherein the reporter tag is selected from the group consisting of a fluorophore, an enzyme, a chemiluminescent moiety, a radioactive moiety, an organic dye, and an organic dye.

62. The detection kit of claim **56**, wherein the solid support is selected from the group consisting of a bead, a sphere, a particle, a membrane, a chip, a slide, a plate, a well, and a test tube.

63. The detection kit of claim **56**, wherein the solid support is a bead, a sphere, or a particle, and wherein the bead, the sphere, or the particle has micrometer or nanometer dimensions.

64. The detection kit of claim **56**, wherein the PAD4 or antigenic fragment thereof is conjugated to the solid support.

65. The detection kit of claim **56**, wherein the detection kit is configured for use with a biological sample, the biological sample selected from blood, plasma, serum, sputum, or saliva.

66. The detection kit of claim **56**, further comprising a positive control selected from the group consisting of a

positive control comprising anti-PAD 4 IgA and a positive control comprising anti-PAD 4 IgG.

67. The detection kit of claim **56**, further comprising: one or more ancillary reagents selected from the group consisting of immobilization buffers, incubation buffers, washing buffers, dilution buffers, detection buffers, assay buffers, and blocking buffers.

68. The detection kit of claim **56**, wherein the antibody or functional fragment thereof capable of binding to anti-PAD4 IgA comprises an anti-human IgA or fragment thereof.

69. The detection kit of claim **56**, wherein the antibody or functional fragment thereof capable of binding to anti-PAD4 IgG comprises an anti-human IgG or fragment thereof.

70. A method of diagnosing or assessing severity of rheumatoid arthritis (RA) in a subject suspected of having RA, the severity being based on presence of joint erosion or severe joint erosion in the subject, the method comprising:

- (a) contacting a biological sample from the subject with a peptidyl arginine deiminase 4 (PAD4) or antigenic fragment thereof;
- (b) contacting the biological sample with a detection probe, wherein the detection probe comprises:
 - (i) an antibody or functional fragment thereof capable of binding to anti-PAD4 IgA; or
 - (ii) an antibody or functional fragment thereof capable of binding to anti-PAD4 IgG;
- (c) detecting a presence of the detection probe, and
- (d) assessing RA severity in the subject based on the detecting the presence of the detection probe.

71. The method of claim **70**, wherein the biological sample is selected from blood, plasma, serum, sputum, or saliva.

72. The method of claim **70**, wherein the detection probe further comprises a reporter tag, and wherein detecting the presence of the detection probe comprises detecting the presence of the reporter tag of the detection probe.

73. The method of claim **71**, wherein the detection probe comprises the antibody or functional fragment thereof capable of binding to anti-PAD4 IgA.

74. The method of claim **71**, wherein detection probe comprises the antibody or functional fragment thereof capable of binding to anti-PAD4 IgG.

75. The method of claim **71**, further comprising contacting the biological sample with a peptidyl arginine deiminase 1 (PAD1) or antigenic fragment thereof; a peptidyl arginine deiminase 2 (PAD2) or antigenic fragment thereof; a peptidyl arginine deiminase 3 (PAD3) or antigenic fragment thereof; a peptidyl arginine deiminase 6 (PAD6) or antigenic fragment thereof; or combinations thereof.

* * * * *