



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.

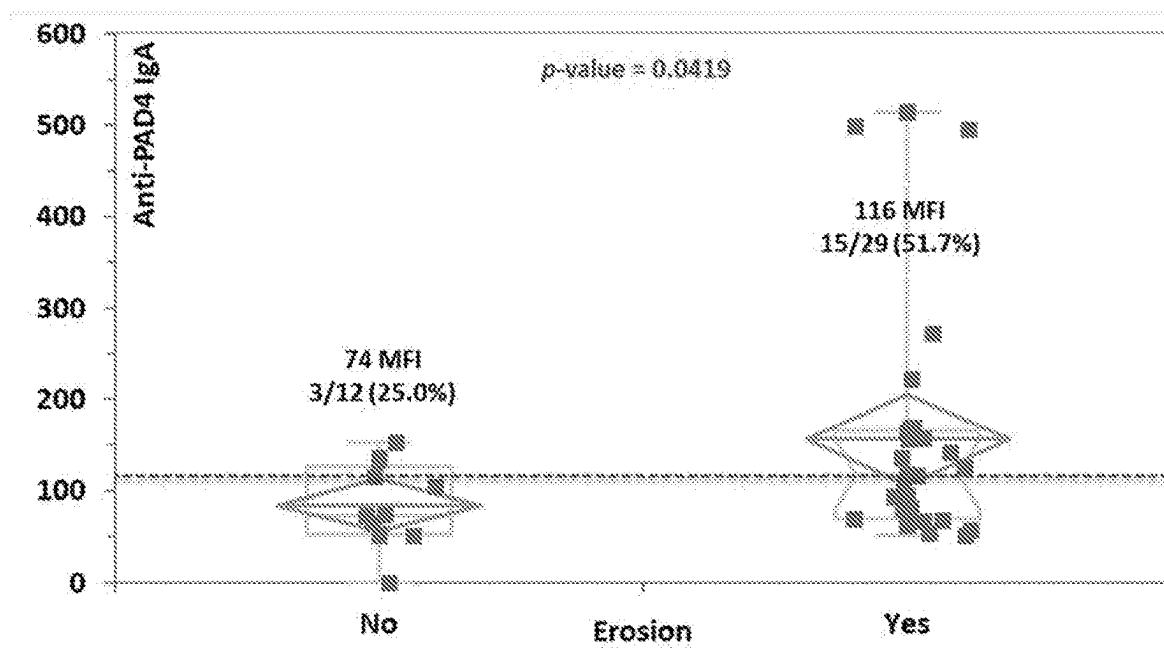


FIG. 1

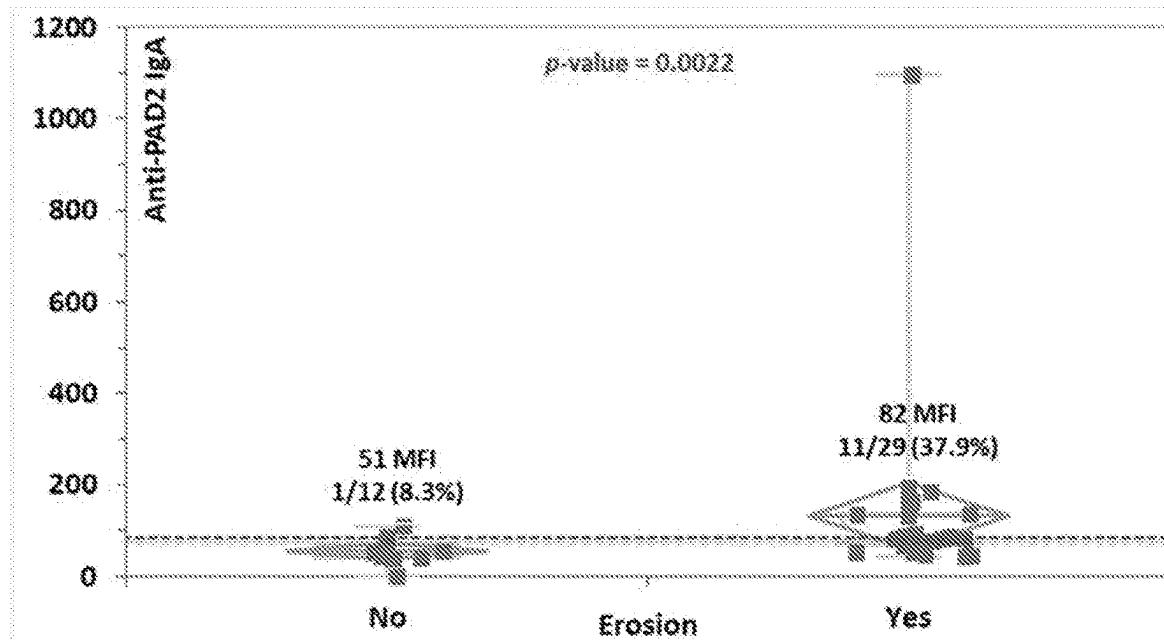


FIG. 2

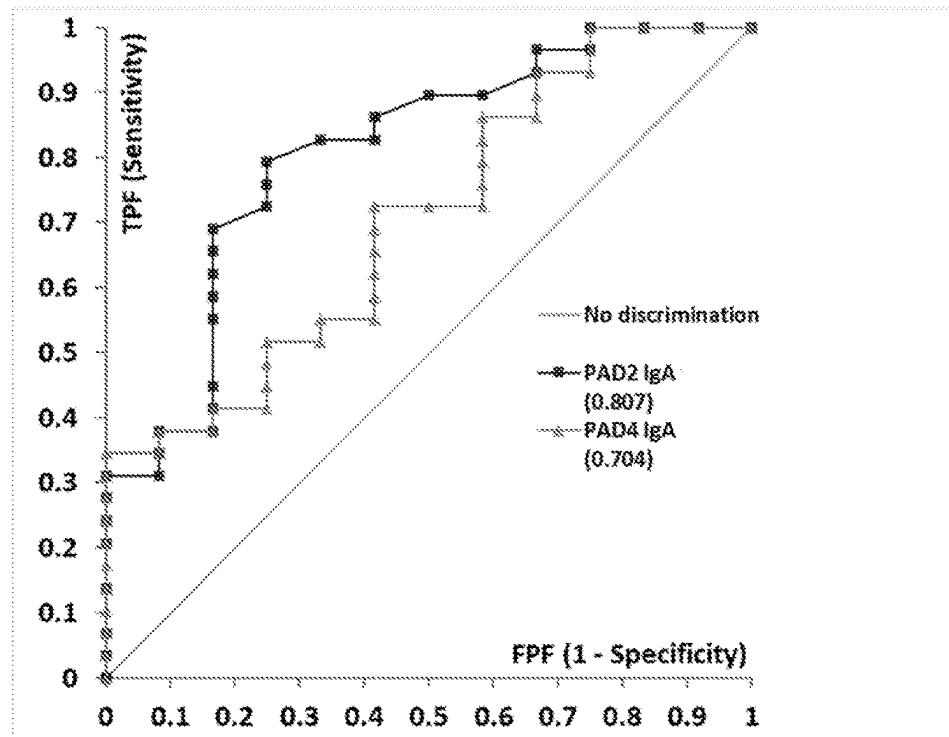


FIG. 3

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 cgggactgca ggcggggagg atgctgcgcg agcggaccgt gcggctgcag tacgggagcc
121 gcgtggaggc ggtgtacgtg ctggcacct acctctggac cgatgtctac agcgcggccc
181 cagccggggc ccaaaccctc agcctgaagc actcggaaaca cgtgtgggtg gaggtggtgc
241 gtatgggga ggctgaggag gtggccacca atggcaagc ggcgtggctt 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 tggttgcccc aggaggactg ccgtgatgag aaggctaca
601 gcaaggaaga tctcaaggac atgtccaga tgatcctgcg gaccaaaggc cccgaccggcc
661 tccccggccg 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 cccatataag gttccctgtt ggtgtggac tctccccggag
1201 atggaaacct aaaggactt cctgtgaagg agctcctgg cccagatttt ggctacgtga
1261 cccggggagcc cctctttgag tctgtcacca gccttgactc atttggaaac ctggaggtca
1321 gtcccccaact gaccgtgaac ggcaagacat accccgcttgg ccgcacatcctc atcggggagca
1381 gcttcctct gtcgtgtgtt cggaggatga ccaagggtgt 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 ttctgggg aagtccactg tggcaccaac gtccgcagga
 2041 agcccttac cttcaagtgg tggcacatgg tgccctgacc tgccaggggc cctggcggtt
 2101 gcctccttcg ctttagtttc cagaccctcc ctacacacgccc cagacccttc tgctgacatg
 2161 gactggacag cccccctggg agaccttgg gacgtggggt ggaatttggg gtatctgtgc
 2221 cttggcccttc ctgagagggg ctcagtgctc ctctgaagcc atccccactg agcctcgact
 2281 ctgtccctgc tgaaaatagc tggggccagtg tctctgtagc cctgacataa ggaacagaac
 2341 acaacaaaac acagcaaaacc atgtgccaa actgctcccc aaagaatttt gagtctctaa
 2401 tctgacactg aatgagggga gaagggaaagg agattctggg attgccagtt ctccagcag
 2461 ccatgctctg aaaatcaagg tagaatccat gggaaaggac cccaggacc cgggacccta
 2521 gacgtatctt gaactgccat cgtcatttca aatacatctc cctcagggtt tccaggtggc
 2581 caccggcaat tattcattcc ttaccaacct ctcaaattctt cttggcttcc tctctgcag
 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 cttgacttt
 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 tatttccatc aaccatttca caaacatttgg tttgtatgcca ctaagcgcca
 3421 ggcacggcgt tggctctgc aaacacagt gtttagtagca gtctggaccc ggtccctact
 3481 ggcattggaaac ccatcactcc ccaacatgca aagcccat ttaaaggcca gcctctgccc
 3541 cttcagtgat ggcgtcttta gaaatgccc tccactat tccagaaatcc 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 ggtatataccca ccgtggactc atgctcttc cccaccccca ctttctataa
 4141 atgttaggcct agaatacgtc tctctgttgc aaaactcagc taagttcctg ctccacctt
 4201 gatgttggaaa tatcttatgt aagaggccag gggatgtcgtaa agatgttta gggaaagagct
 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 kvyskedlkd
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 agtcataatc tgacatgtg cacaaggaca
61 taaaacataca agtgcactac agtgcgtc ttgttagtag caaaagattt gaaaggcaacc
121 taaaatgtcca tctgtggagg actgagtaac ggcctcccg gaagttctgt gcaaccgcta
181 aaaaggatga agccagtctc tcaatatggc tacaggatgt gcttcaagga taaaattttag
241 tttggctaca tcgaggcccc ccataaaaggc ttccccgtgg tgctgactc tccccgagat
301 gaaaaacctaa aggacttccc tgtgaaggag ctccctggccc cagattttgg ctacgtgacc
361 cgggagcccc tctttgagtc tgtcaccagc cttgactcat ttggaaacct ggaggtcagt
421 cccccagtga ccgtgaacgg caagacatac cccgttggcc gcacatctcat cgggagcagc
481 tttcctctgt ctgggtggcc gaggatgacc aagggtggtc gtgacttctt gaaggccccag
541 cagggtcagg cggccgtggc gctctactca gactggctga ctgtggccca cgtggatgag
601 ttcatgtcct ttgtccccat cccggcaca aagaaattcc tgctactcat gggcagcacc
661 tcggcctgct acaagcttt ccgagagaag cagaaggacg gcccattcatg gggccatcatg
721 ttcaaaggct tgggtgggat gagcagcaag cgaatcacca tcaacaagat tctgtccaaac
781 gagagcctt tgccaggagaa cctgtacttc cagcgctgcc tagactggaa cctgtacatc
841 ctcaagaagg agctgggact gacagagcag gacatcattt acctggccgc tctgttcaag
901 atggacgagg accaccgtgc cagaccattt ttcccaaaca ttgtgaacat gatcgctgctg
961 gacaaggacc tgggcattttt caagccattt gggccacagg ttgaggagga atgctgcctt
1021 gagatgcacg tgcgtggcct cctggagccc ctgggcctcg aatgcacattt catcgacgac
1081 atttctgcct accacaaatt tctggggaa gtccactgtg gcaccaacgt ccgcaggaaag
1141 cccttcaccc tcaagtgggt gcacatggc ccctgacactg ccagggccccc tggcgtttgc
1201 ctccttcgt tagttctcca gaccctccct cacacgcccc gggccatttctg ctgacatgg
1261 ctggacagcc cccgtggggag acctttggga cgtggggat aattttgggt atctgtgcct
1321 tgcctccct gagagggggcc 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 caggaccccg ggaccctaga
1621 cgtatcttgc actgcccattt tcatttcaaa tacatctccc tcagggtttc cagggtggcca
1681 ccccaatta ttccatttccctt accaacctt ccaatctt tggctttctc tctgcagtgt
1741 ggacactgtt ggctagtctt ccccaactccc tgagggttcca gtaaggtagc tttagaaccatt
1801 cctggaaaca tttcatctga gcaggttcc ccacgtgtgg gatgctccctt ttgcctcatc
1861 tgtctcaggatgg atgcagggtc ccccgcatgc atggggattt ctcccagac cagcataact
1921 gtgacactgag agttcaatgc gtaaagatgc ccctggcttag ccataatccat ctctcttgc
1981 ctggccttgc attctctggc cgctccctga ccttcctcct tccactgcct tgactttctt

```

2041 ccttttatt cctggtgcca tctgtccagg cagctagaca agaacttgtt cgccagcagc
2101 cagattcagg cctteccagg ggcataataa gtgaccagcc cctcctctcc ggacatcaga
2161 tccaacacat aaggaccctg gcctaccctc cagcccaaca gccagttctg gtcagctgc
2221 caacttaggg gtggtttqat tatcccattt aaattcacca gtgccttgc caaagaccct
2281 ctcatttga cataccaga ttcatccct ggctccaact gaaaagactc agtttcaatc
2341 gttaaaagtt cctttagggc cagaagaata aatgaattat aatccattt tgaagaaccg
2401 atttataacc aatgaaaagg ttataatgtt atttatattt ttggaggaac aagatttca
2461 tttgggatta ttcccttcaa ccattcaaca acatattgtt gtatgccact aagcgccagg
2521 cacggcgttgg ctgcgttgc acacagtgg tagtagcagt ctggacactgg tccctactgg
2581 catggAACCC atcaactcccc aacatgcaaa gcccacattt aaaggccagc ctctgcccct
2641 tcagtatgc gctctttaga aatgccagtc cactatattt agaaatccgc agggcacaaa
2701 acttccagca agtcaactgtt gtggtaaat gggcagtggg ggtgggggtt cttctttaaa
2761 caggccccct tcccatctac ctggccagta cccatccaat gagtccccag agcctccaga
2821 agctgttgc tcctctctgg ggacagcgc tcctgcctt ggaggccaaa gccccagatc
2881 tctccagccc cagagctgaa aacaccaagt gcctatttga gggtgtctgt ctggagactt
2941 agagtttgc atgtgtgtgt gtgtgtttgg ttaatgtggg tttatgggtt ttctttcttt
3001 tttttcttt ttttttttagt ctacattagg gggaaagttag cgcctccat gtgcagacag
3061 tgtgtctta tagatttttc taaggtttc cccatgtat tcggtaattt ctgtatgtttc
3121 tgaagttccc aggactcaca caccgttcc catctcactt gcccacccag tggacaacc
3181 ctccgtgtgg atataccccc gtggactcat gcctcttccc cacccttactt ttctataaaat
3241 gtaggccttag aatacgccttc tctgttgc aaactcagcta agttctgttccacccat
3301 tggtaataa tcttatgtaa gaggcaggg gatgtcgtga agatggcaag aagaacacag
3361 ttccaaattt ctggaaaaga gcctgtgtg gagatctaaa gatgtttagg gaagagctcg
3421 actaaaagaac aatgaaataa atggtccaag gggaaagtca

```

FIG. 6

```

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

```

```

1 msicglsng lpgssvqplk rmkpvsqygy rmcfkdeief gyieaphkgf pvvldsprdg
61 nlkdfpvkel lgpdfgyvtr epifesvtl dsfgnlevsp pvtvnngktyp lgrilicsssf
121 plsggrrmtk vvrdflkqaqq vqapvelysd wltvghvdef msfvipgk kflilmasts
181 acylfrekj kdghgeaimf klggmsskr itinkilsne slvqenlyfq rcldwnrldil
241 kkelglteqd iidlpalfkm dedhraraff pnmvmivld kdlgipkpfq pqveeeccle
301 mhvrqllepl glectfiddi sayhkflgev hcgtnvrrkp ftfkwvhmvp

```

FIG. 7

```

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

```

```

1 agtgttgggg ttggcggcca cagctaagtc caacaccagc atgtcgctgc agagaatcgt
61 gcgtgtgtcc ctggagcatc ccaccagcgc ggtgtgtgt gctggcgtgg agaccctcg
121 ggacatttat gggtcagtc ctgagggcac agaaatgttt gaggtctatg ggacgcctgg
181 cgtggacatc tacatcttc ccaacatgga gagggggccgg gagcgtgcag acaccaggcg
241 gtggcgttt gacgcgactt tggagatcat cgtggatcat aactccccca gcaatgacc
301 caacgacagc catgttcaga tttcctacca ctccagccat gagcctctgc ccctggccct
361 tgccgtgctc tacctcacct gtgtgtgacat ctctctggat tgcgaccctga actgtgagg
421 aaggcaggac aggaactttg tagacaagcg gcagtggttc tggggccca gtgggtatgg
481 cggcatcttgc ctggtaact gtgaccgtga tgatccgagc tgcgtatgtcc aggacaattt
541 tgaccagcac gtgcactgcc tgcaagacat ggaagacatg tctgtcatgg tccctgcggac
601 gcagggccct gcagccctt ttgtatgacca caaacttgtc ctccataacct ccagctatga

```

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 aggcttccg
 1501 gatgctcctg gccagccctg gggcctgctt caagctttc cagaaaaagc agaagtgtgg
 1561 ccacggggagg gcccctcctgt tccagggggt tggatgat gggcgggtca agaccatctc
 1621 catcaaccag gtgcctcca 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 gggcgtccctg 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 ttccctggta agtgcatttgc cctcacccatgtg
 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 ctgtttgggt ccaattgtctt 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 aggggtaaaa gtagttccca
 2821 taatacacat ggttgactat ggtgatccac ctgtgtatgg ttaatattag gtgtctggag
 2881 aagggttgctt cattggccct gggacttctc tctgcaggag gagagaacgc tgcctctcct
 2941 ctggattgtt ctcaggctct ctgttggct ttgtcagcg tttccacatc ctgctctgt
 3001 gcaggagagg gggctaaggg gctggatcca ccaaggcagc tcacagccgg aaaactctgg
 3061 gaatgaacca ctgaatttcg gggatgggg 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 ggagttaggt tgccagataa aataaaagat
121 gcctagttaa aattgaattt caggtaaata attaataatt ttttagtat aagtgtatcc
181 caaataactc 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 agttttccca tggtaaaca ggagtaatta ttcatccctg
601 ggctgcagt gaggatcaga tggactaatg tacaaaagg cttaggcagg ctgcctgtga
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 ctccaactcg 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 ctggctcgatg cctgaggcga cagaatgtt tggatgttat gggacccctg gctgtggacat
1801 ctacatctt cccacatgg agagggccg ggagcgtgca gacaccaggc ggtggcgctt
1861 tgacgcgact ttggagatca tcgtggatcat gaactcccc agcaatgacc tcaacgcacag
1921 ccatgttccat atttcctacc actccagcca tggcctctg cccctggcct atgcgggtgtct
1981 ctacacttcc tctctcttca ttgcacactt 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 cctccatacc tccagctatg atgccaacg
 2281 ggcacagggtc ttccacatct gcggtcctga ggatgtgtgt gaggcctata ggcattgtgt
 2341 gggccaagat aagggtgtct atgaggtacc ccgcttgcatt 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 cccctgggaa 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 tcatcaacta caataagttt gtgcagagct gcatcgactg
 3301 gaaccgtgag gtgctgaagc gggagctggg cttggcagag tggatgttca ttgacatccc
 3361 acagcttcc aagaccgaga gggaaaaagc aacggccctc ttccctgact tggtaacat
 3421 gctgggtctg gggaaagcacc tgggcattttt caagccctt gggccatca tcaatggctg
 3481 ctgctgcctg gaggagaagg tggcgttccct gctggagccg ctgggcctcc actgcacctt
 3541 cattgatgac ttcaactccat accacatgtc gcatggggag gtgcactgtg gcaccaatgt
 3601 gtgcagaaaag cccttcttcc tcaagtgggtt gacatgggtt ccctgagaca gctcccaccc
 3661 accatccctgt cccctgggg cgggcattgg cccaggtggt ggagacagag acaggcccc
 3721 gaacgataag caccaagaga ccccaaggctt ccagatggaa cactgagggt gaccgtccct
 3781 ctcagaagcc tttccctgg aagtgtccat gcctcacctt caacccatgt ggttctcaga
 3841 cttgaatctt ctcggccccc caaaaagaag gacccatattt cttatagcct tccctgtgat
 3901 tcaacacaaac ccatggagat gtccttctt cactctgaaa tcattccattt gggacaaat
 3961 ccacattggg gtctagaaac atccacgtat ctcatcagcc atcttgcctt gtgcaccccta
 4021 acagagggaaag gatccatgtat tctgtttgg tccaaattgtc tcctctctgc agaggaacaa
 4081 ccctaaaacc agaccactcc acgcaggaca ggcaggagag attcttctta aagcctccccc
 4141 cataaaaagg gagctgttga tccacttaga tcagggcggg accatcttc acccgccaa
 4201 gtcctgccttcc agatgttgc ctcaccat cgtgagctgt cacatagtag gagcttcttag
 4261 atgcattgtgg aagcaatgtg agttgttccct tagccttata aactcccat gatctgacat
 4321 gcagaaaatcc agccttgc tggaaattttt tggagacgaa agtatctggg
 4381 ggattgttgg gtacttaggg aactgggtac aagggtgaaa agtagttccc ataataacaca
 4441 tgggtgacta tgggtatcca cttgtgtatg gttaatattt ggtgtctgg gaaagggtgct
 4501 tcattggccc tggacttctt ctctgcagga ggagagaacg ctgcctctcc tctggattgg
 4561 tctcaggctc tctgttggcc tttgggtcagc gtttccat ctcgtctgc tgcaggagag
 4621 ggggctaagg ggctggatcc accaaggcag ctcacagccg gaaaactctg ggaatgaacc
 4681 actgaattca gggatgggg gtggggggc gtttctcgag 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 ccagatttgg cccacagttc ataatgtgct gatcctgacc taggcgagaa gagaaaccaa
121 atatgaaact gttgaagaac ttggactga attatgttgg aacttggtgc cctgggagtt
181 aagaggagaa ggtcgagcgg cagtggctc gggggccca tggatgtggc ggcattttgc
241 tggtaactg tgaccgtat gatccgagct gtatgttcca ggacaattgt gaccagcacg
301 tgcactgcct gcaagacctg gaagacatgt ctgtcatggt cctgcggacg cagggccctg
361 cagcccttctt tggatgaccac aaacttgtcc tccataccctc cagctatgtat gccaaacggg
421 cacaggcttcc acatcttc ggtcttgagg atgtgtgtga ggcctatagg catgtgctgg
481 gccaagataa ggtgtccat gaggatcccc gcttgcatttgg ggatgaggag cgcttcttcg
541 tggaaaggctt gtccttccctt gatccggct tcacaggact catctccctc catgtcactc
601 tgcggacga ctccaaacggag gatttctcgat catcccttat cttcaactgac actgtgggt
661 tccgagggttcc accctggatc atgacgccc gcaactctggcc accccttagag gtgtatgtgt
721 gccgtgttagt gaaacaacacg tggatgttgg atgcgggtggc agagctggcc aggaaggccg
781 gctgcaagctt gaccatcttc ccacaggccg agaaccgc当地 cgaccgctgg atccaggatg
841 agatggagctt gggctacgtt caggccgc当地 acaagaccctt cccgggtggc tttgactccc
901 caaggaatgg ggaactgc当地 gatttccctt acaaaagaat cctgggtcca gattttgggtt
961 acgtgactcg ggaaccacgc gacaggctt gtagtggccctt ggactccccc gggAACCTGG
1021 aggtcagccc tccagggtt gccaatgggaa aagagtaccc cctggggagg atcctcatttgc
1081 gggcaacctt gcctgggtca agtggccgc当地 ggttcacccca ggtgggtggc gacttccctcc
1141 atggcccttcc ggtgcagccc cccgtggagc当地 tctttgtggc ctggatgtggc gttggccatg
1201 tggatgactt tctgagctt gtccttccccc ccgtatgggaa gggcttccgg atgcttccctgg
1261 ccagccctgg ggcctgcttcc aagcttccccc aggaaaagc当地 gaagtgtggc cacgggaggg
1321 ccctccctttt ccagggggtt gttgatgatg agcaggcttcc gaccatctcc atcaaccagg
1381 tgcactccaa taaagacccatc atcaactaca ataaatgttgc gcaatgtgc atcgacttgg
1441 accgtgaggtt gctgaaagccg gaggctggcc tggcagagtg tgacatcattt gacatccccc
1501 agtcttccaa gaccgagagg aaaaaaagc当地 cggccttccctt ccctgacttg gtgaacatgc
1561 tggatgtgg gaaagcaccc ggc当地 catccccca agccctttgg gccc当地 catc aatggctgt
1621 gctgcctgg gggaaagggtt cggatccctgc tggagccgc当地 gggccctccac tgcacccctca
1681 ttgatgactt cactccatc acatgttgc atggggagggtt gcaatgtggc accaatgtgt
1741 gcaagaaagcc cttcttcc aagttgggaa acatgttgc当地 ctgagacagc tcccaccac
1801 catcccttcc ccctggggcc ggc当地 atggggatgg agacagagac agggcccttgc
1861 acgataagca ccaagagacc ccaaggcttcc agatggaaaca ctgagggtga cctgtccctct
1921 cagaaggcctt ttccctggaa gtttccatgc ctcaccccttcc acccatgtgg ttctcagact
1981 tgaatcttcc cggccccccca aaaagaagga ccttccatcc tatagccctt cctgtgattc
```

```

2041 aacacaaccc atggagatgt ccccttctca ctctgaaatc atccattttgg ggacaaatcc
2101 acattgggggt ctagaaacat ccacgtatct catcagccat cttgtcctgt gcatcctaacc
2161 agaggaagga tccatgattc tgctttggtc caattgcctc ctctctgcag aggaacaacc
2221 ctaaaaccag accactccac gcaggacagg caggagagat tcttctaaa gcctccccca
2281 taaaaaggga gctgtggatc cacttagatc agggcggAAC catcttcac ccggccaagc
2341 tcctgcccag atgttgaccc tcacccagcg tgagctgtca catacgtagga gcttctagat
2401 gcatgtggaa gcaatgagag ttgtccctta gccttataaa cttcccatga tctgacatgc
2461 agaaaatccag cttgtccag aatccctctg gaatttcttgc gagacgaaag tatctggggg
2521 attgttgggt actagggaga ctgggtacaa gggtaaaaag tagtcccat aatacacatg
2581 gttgactatg gtgatccacc ttgtgatggt taatattagg tgtctggaga aggttgcttcc
2641 attggccctg ggacttctct ctgcaggagg agagaacgct gcctctccctc tggattgggt
2701 tcaggctctc tgttggccctt tggtcagcgt ttccacatcc tgctctgctg caggagagg
2761 ggctaagggg ctggatccac caaggcagct cacagcggga aaactctggg aatgaaccac
2821 tgaattcagg ggatgggggt gggggggcgg ttctcgaggt gtgtgccagc tacacgtgtg
2881 ttctgtatgg gtccagctgc gttccatca ctcgctaata 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 msmvmlrtqg paalfddhkl vlhtssybak raqvhicgp edvceayrhv lgqdkvsyev
61 prlhgdeerf fveglspda gftglisfhv tllddsnedf saspiftdtv vfrvapwint
121 pstlpplevy vcrvrnnctf vdavaelark agcklticpq aenrndrwiq demelgyvqa
181 phktlpvvfd sprngelqdf pykrilgpdf gyvtreprdr svsgldsfgn levspvvan
241 gkeyplgril iggnlpqssg rrvtqvrvdf lhaqkvqppv elfvdwlavg hvdeflfsfpv
301 apdgkgfrml laspgacfkl fqekqkcgħg rallfqgvvd deqvktisin qvlsnkdlm
361 ynkfvqscid wnrevlkrel glaecdiidi pqqlfkterkk ataffpdlnv mlvlghħlgi
421 pkpfqpiing ccclleekvrs lleplglhct fiddftpyhm lhgevhcgtn vcrkpfsfkw
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 aggcagtagg
 61 ccagatttgg cccacaggtc ataatgtct gatcctgacc taggcgagaa gagaaaccaa
121 atatgaaact gttgaagaac ttgggactga attatgttgg aacttgggtc cctgggagtt
181 aaggaggaaa ggtcgagccg cagtgggtct gggggcccaag tgggtatggc gcatcttgc
241 tggtaactg tgaccgtat gatccgagct gtatgtcca ggacaattgt gaccagcacg
301 tgcactgcct gcaagacctg gaagacatgt ctgtcatgtt cctgcggacg cagggccctg
361 cagccctctt tgatgaccac aaacttgtcc tccataccctc cagctatgtat gccaaacggg
421 cacaggtctt ccacatctgc ggtcctgagg atgtgtgtga ggcctataagg catgtcttgg
481 gccaagataa ggtgtccat gaggtacccc gcttgcattgg ggatgaggag cgcttcttcg
541 tggaaaggct gtccttcct gatgccggct tcacaggact catctccttc catgtcaactc
601 tgctggacga ctccaacggat gatttctcgat catcccctat cttcaactgac actgtggtgt
661 tccgagtgcc accctggatc atgacgcccc gcactctgcc acccctagag gtgtatgtgt
721 gccgtgttag gaacaacacg tggttggatc ggcgttggc agagctggcc aggaaggccg
781 gctgcaagct gaccatctgc ccacaggccg agaaccgc当地 cgaccgctgg atccaggatg
841 aqatqqagct qqqtacqtt cagqccccc acaaqaccct cccqqtqtc ttqactccc

```

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 tccatcacctc cagctatgtat gccaaacggg
421 cacaggctt ccacatctgc ggtccgtt gatgtgtgtga ggcctatagg catgtgttgg
481 gccaagataa ggtgtccat gaggtaaaaa gcttgcattgg ggatgaggag cgcttcttcg
541 tggaaaggct gtccttcctt gatggggct tcacaggact catctccatc catgtcactc
601 tgctggacga ctccaaacggag gattttctcgat cttccatgtac actgtgggtgt
661 tccagggtggc accctggatc atgacggccca gcaactctgtcc acccccttagag gtgtatgtgt
721 gccgtgttag gaaacaacacg tgggtttgtgg atgcgggtggc agagctggcc aggaaggccg
781 gctgcaagct gaccatctgc ccacaggccg agaaccgc当地 cgaccgcgtgg 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 atggggagggt 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 tgcaaaaaaaa cgtggagctc tttgtggact ggttgccgt gggccatgtg
1321 gatgagttc tgagctttgt ccctgtcccc gatggaaagg 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 qdkvsvyevp
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 agcctctgca 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 ttttgacacaa aacttgtcc
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 tgcagcccc 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 gtgaccccg
61 agcagcccac ccatgccgtg tgtgtgctgg gcacccgtac tcagcttgc acatctgcagct
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 caccaggtaa agctctactc tacctcaccg
361 ggggtggaaat ctccctgtgc gcagacatca cccgcacccgg caaatgtgaag ccaaccagag
421 ctgtgaaaga tcagaggacc tggacctggg gcccttggg acagggtgcc atcctgctgg
481 tgaactgtga cagagacaat ctcgaatctt ctgcccattga ctgcgaggat gatgaagtgc
541 ttgacagcga agacactgcag gacatgtcgc ttagtggccct gagcacgaag acccccaagg
601 acttcttcac aaaccataca ctggtgctcc acgtggccag gtctgagatg gacaaagtga
661 ggggtttca ggccacacgg ggcaactgtt cctccaatgt cagctgttc ttgggtccca
721 agtggccctc tcactacctg atggtccccg gtggaaagca caacatggac ttctacgtgg
781 aggccttcgc tttcccgac accgacttcc cggggctcat tacccctcacc atctccctgc
841 tggacacgtc caacctggag ctcccccggg ctgtgggttt ccaagacagc gtggcttcc
901 gcgtggcgcc ctggatcatg acccccaaca cccagccccc gcaggagggtg tacgcgtgca
961 gtattttga aaatgaggac ttccctgaatgt cagtgtactac tctggccatg aaagccaagt
1021 gcaagctgac catctccct gaggaggaga acatggatga ccagtggatg caggatgaaa
1081 tggagatcgg ctacatccaa gccccacaca aaacgctgcc cgtggcttcc gactctccaa
1141 ggaacagagg cctgaaggag ttcccatca aacgcgtgtatggccatgt tttggctatg
1201 taactcgagg gccccaaaca ggggttatca gtggactgaa ctcccttggg aacctggaaag
1261 tgagcccccc agtcacagtcc aggggcaagg aatacccgct gggcaggatt ctcttcgggg
1321 acagctgtta tcccagcaat gacagccggc agatgcacca ggcctgcag gacttcctca
1381 gtgcccagca ggtgcaggcc cctgtgaagc tctattctga ctggctgtcc gtggccacag
1441 tggacgagtt ccttagctt gtgccagcac ccgacaggaa gggcttccgg ctgctcctgg
1501 ccagccccag gtcctgtac aaactgttcc aggaggcagca gaatgagggc cacggggagg
1561 ccctgctgtt cgaagggtac aaaaaaaaaa aacagcagaa aataaaagaac attctgtcaa
1621 acaagacatt gagagaacat aattcatttg tggagagatg catcgactgg aaccgcgagc
1681 tgctgaagcg ggagctggc ctggccgaga gtgacatcat tgacatcccg cagctttca
1741 agctcaaaga gttctctaag gcggaaagctt ttttcccaa catggtaac atgctgggtgc
1801 tagggaaagca cctgggcatac cccaaaggct tcggggccgt catcaacggc cgctgctgcc
1861 tggaggagaa ggtgtgttcc ctgtggagc cactgggcct ccagtgacc ttcatcaacg
1921 acttcttcac ctaccacatc aggcatgggg aggtgcactg cggcaccaac gtgcgcagaa
1981 agcccttcctc cttcaagttgg tggAACATGG tggccctgac ccacatccccc tggcgtcc
2041 tccctcctgg ccagatgtcg ctgggtccctc tgcagtgtgg caagaagag ctcttgcata
2101 tattgtggct ccctggggc ggccagccct cccagcagtg gcttgcatttcc ttctcctgtg
2161 atgtcccagt ttcccactct gaagatccca acatggtcct agcactgcac actcagttc
2221 gctctaagaa gctgcaataa agttttta agtcactttt tacatga

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 snktlrehn
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 catgaccccc aacaccgc ccccgaggg
481 ggtgtacgcg tgcatgtatt ttgaaaatga ggacttcctg aagtcatgtg ctactctggc
541 catgaaagcc aagtgcacg tgaccatctg ccctgaggag gagaacatgg atgaccatgg
601 gatgcaggat gaaatggaga tcggctacat ccaagccca cacaaacgc tgcccggtgg
661 ctgcactt ccaagggaca gaggcctgaa ggagttccc atcaaacgc tgatgggtcc
721 agatttggc tatgttaactc gaggccccca aacagggggt atcagtggac tggactccct
781 tggaaacctg gaagtggaccc ccccaactc agtcaggggc aaggaataacc cgctggcag
841 gattcttcc ggggacagct gttatcccag caatgcacgc cggcagatgc accaggccct
901 gcaggacttc ctcagtgccc agcaggtgca ggcccctgtg aagctctatt ctgactggc
961 gtccgtgggc cacgtggacg agttctggat ctttgtgcca gcaccgc gaaaggcct
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 cctgtcatcaa
1381 cggccgctgc tgcctggagg agaagggtgtt ccctctgtc gagccactgg gcctccagtg
1441 cacccatcata aacgacttctc tcacccatca catcaggcat ggggagggtgc actgcggcac
1501 caacgtgcgc agaaaaggct tctccttcaa gtgggtggaa atgggtccct gagcccatct
1561 tccctggcgt cctctccctc ctggccagat gtgcgtgggt cctctgcagt gtggcaagca
1621 agagcttctg 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 ggcccaaaac 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 gtggacatcg gtggccctgag
1501 cccatcttcc ctggcgtcct ctccctcctg gccagatgtc gctgggtcct ctgcagtgt
1561 gcaagcaaga gctctgtga atattgtggc tccctggggg cggccagcccc tcccagcagt
1621 ggcttgctt cttctctgt gatgtcccag tttccactc tgaagatccc aacatggtcc
1681 tagcaactgca cactcagttc tgctctaaga agctgcaata aagtttttt aagtcacttt
1741 gtacatga

FIG. 22

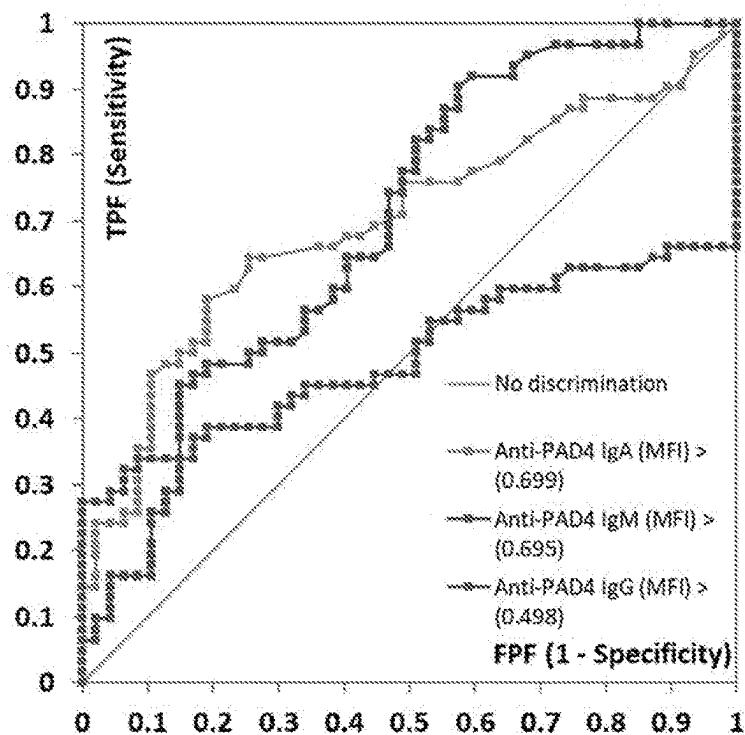


FIG. 23

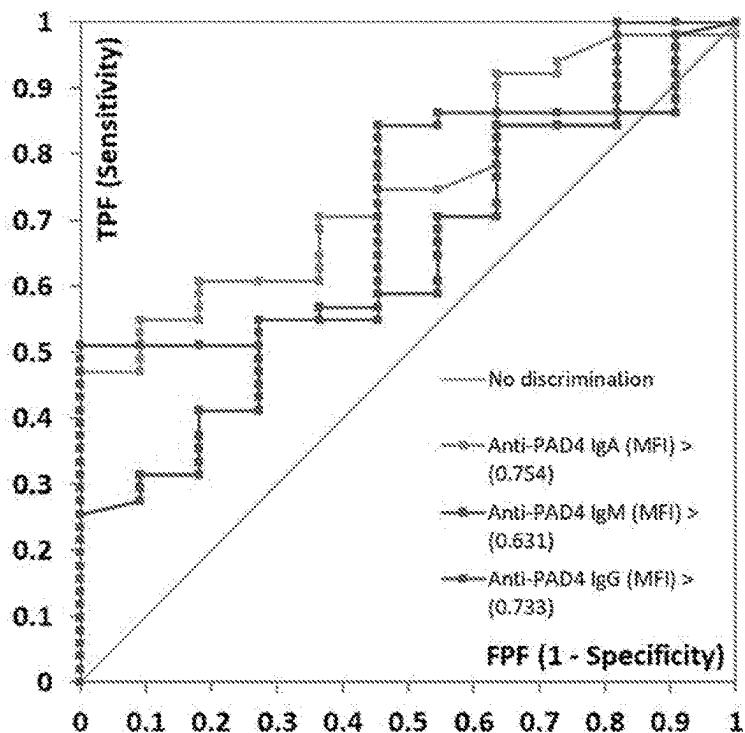


FIG. 24

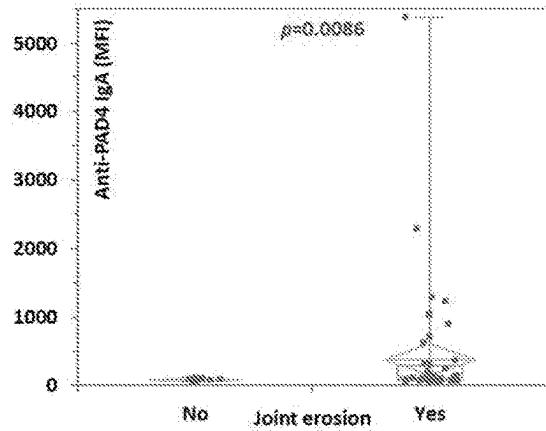


FIG. 25

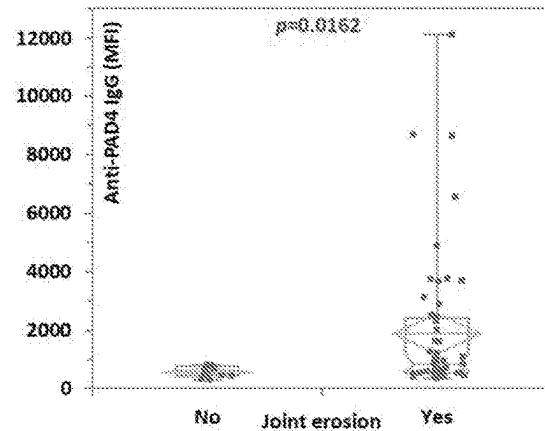


FIG. 26

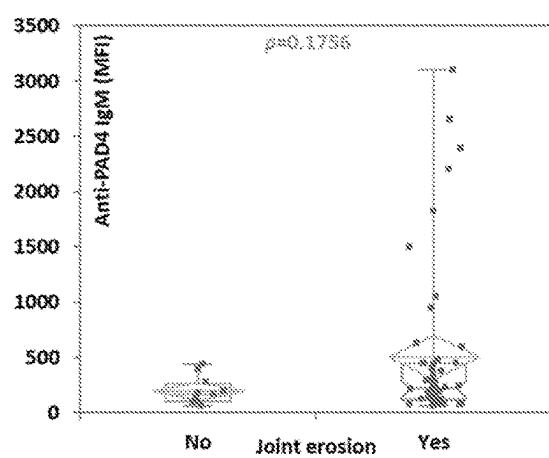


FIG. 27

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
VEVVRDGEAEEVATNGKQRWLSPSTTLRVMTMSQASTEASSDKVTVNYY
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
A1LLVNCDRNLESSAMDCE DEDVLDSEDLQDMSMLTSLTKTPKDFFTN
HTLVLHVARSEM DKV RVFQATRGKLSSKCSVVLGP KWP SHYLMVPGGKH
NMDFYVEALAFFPDTDFPGLITLTISLLDTSNLELP EAVVFQDSVVPRVA
PWIMTPNTQPPQEVYAC SIFENEDFLKS VTTLAMKAKC KL TICPEEEENM
DDQWMQDEMEIG YIQAP HKTL P VVFDSPRN RGLKE FPI KRV MGP DFGYV
TRGPQTGGISGLDSFGN LEVSPVPTVRGKEYPLGRILFGDSCYP SNSDR
QMHQALQDFLSAQVQAPV KLYSDWL SVGHVDEFLSFVPAPDRKGPFRL
LASPRSCYKLFQEQQN EGHGEALLFEGIKKKQQKIKN ILSNKTLREHN
SFVERCIDWNREL KREL GLA ESDIIDIPQLFKLKEFSKA EAFPNMVN
MLVLGKHLGIPKPF GPVINGRC CLEEKVCSLLEPLGLQCTFINDFFTYH
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 include, 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 cctgtatgag aaggcttaca 600
gcaaggaaaga tctcaaggac atgtcccaaga ttagtcctgcg gaccaaaggc cccgaccgccc 660
tccccggccg atacgagata gttctgtaca ttccatgtc agactcagac aaagtggccg 720
tgttctacgt ggagaaccccg ttacccgttcc acgcgttatccatccctg ggcggccgg 780
agctctacca tgggtcaag tccacccgggt gtcgtggcggaa gtcgtgttcc ttcgtggaa 840
gcctctgttt ccccgacggag ggcttctcaag gctgtgttcc catccatgtc agcctgtgg 900
agtacatggc ccaggacatt cccctgactc ccatcttccat ggacacccgtg atattccgg 960
ttgtccctgg gatcatgacc cccaaacatcc tgcctcccggt gtcgtgttcc gtcgtgtca 1020
tgaaggatata ttacctgttcc tccgaaaggagg tgaagaacccat tggggaaaa accaactgtg 1080
agctgaagggt ctgttcccaag tccatccaaag gatccggccgtt gtcgtgttcc gtcgtgttcc 1140
agtttggcta catcgaggccc cccctatccaaag gttcccccgtt gtcgtgttcc tccatccggag 1200
atggaaacctt aaaggacttcc cctgtgaagg agcttccctggg cccagatttt ggctacgtga 1260
cccgggggcccttccatgg gatccatggcc gatccggccgtt atttggaaaacttggggatcc 1320
gtccccccatgtt gatccatggcc gatccggccgtt acgcgttccatggccatggcc 1380
gccttcccttcttgcgttccatggcc gatccggccgtt acgcgttccatggccatggcc 1440
agcagggtca ggcggcccggtt gatccggccgtt acgcgttccatggccatggcc 1500
agttcatgttccatggcc gatccggccgtt acgcgttccatggccatggcc 1560
cctccggcccttgcgttccatggcc gatccggccgtt acgcgttccatggccatggcc 1620
tggtaaaaggcttgggtggg atggccatggcc gatccggccgtt acgcgttccatggcc 1680
acggagcccttgcgttccatggcc gatccggccgtt acgcgttccatggccatggcc 1740

```

-continued

tcctcaagaa	ggagctggga	ctgacagago	aggacatcat	tgacacctccc	gctctgttca	1800
agatggacga	ggaccacccgt	gccagagccct	tcttcccaa	catggtaac	atgatcggtc	1860
tggacaaggaa	cctggccatc	cccaagccat	tcggggccaa	gggtgaggag	gaatgtgcc	1920
tggagatgca	cgtgcgtggc	ctctggago	ccttgggcct	cgatgcacc	ttcatcgacg	1980
acattttgc	ctaccacaaa	tttctggggg	aagtccactg	tgccaccaac	gtccgcagga	2040
agcccttcac	cttcaagttgg	tggcacatgg	tgccctgacc	tgccaggggc	cctggcggtt	2100
gccttcctcg	ccttagtttc	cagacccctcc	ctcacacgc	cagacccctc	tgctgacatg	2160
gactggacag	ccccgcgtgg	agacctttgg	gacgtggggg	ggaatttggg	gtatctgtgc	2220
cttgcctcc	ctgagagggg	cctcaagtgc	ctctgaacc	atccccagtg	agcctcgact	2280
ctgtccctgc	tgaaaaataat	tggggcactgg	tctctgtacg	cctgacataa	ggaacagaac	2340
acaacaaaac	acagcaaaac	atgtggccaa	actgctcccc	aaagaattt	gagttcttaa	2400
tctgacactg	aataggggga	gaagggaagg	agattctggg	attgcaggat	cttccagcag	2460
ccatgctctg	aaaatcaagg	tagaatccat	ggaaaggagc	cccaggaccc	cggggaccta	2520
gacgttatctt	gaaactccat	cgttatcca	aatacatctt	cctcagggtt	tccagggtgc	2580
caccccaat	tattcatcc	ttaccaac	cttcaaatcc	cttgggtt	tctctgeagat	2640
qtgacactg	ttggctagtc	ctccccactc	cctgagggtc	cagtaagtt	gcttagaaacc	2700
ttctctggaaa	catttcatct	gagcagggtt	ccccacgtgt	gggatgtccc	ttttgcctca	2760
tctgtctcg	ggatgcagg	tcccccgcat	gcatggggat	tttcccccag	accaggatata	2820
tttgtacatcg	agagttcaat	gctttaaagat	gcccctgggt	agccatatacc	atcttcttt	2880
gctgtgtct	tgatttcctg	gcccgtccct	gacccttc	cttccactgc	cttgaacttc	2940
ttctctttta	ttctctgggc	catctgtcca	ggcagctaga	caagaacttg	ttcgecagca	3000
gcccagattca	ggccatccca	ggggcataaa	aagtgaccag	ccccctccct	ccggacatca	3060
gatccaaacac	ataaggacc	tggcctacc	tcggcccaa	caggccatgc	ttgggtca	3120
gccaacttag	gggtgggtt	attatccat	tgaattttac	cagtgcctt	gccaagagacc	3180
ctctcatttg	gacataccca	gattcattcc	ctggctccaa	ctgaaaagac	tcagtttcaa	3240
tcgtttaaaag	ttccctttagg	gcccagaagaa	taaatgaatt	ataatccat	tttgaagaac	3300
cgatttataa	ccatgtaaaa	ggttataat	taattttataat	ttttggggaa	acaagattt	3360
catttgggat	tattttcttc	aaccattca	caaacatttg	ttgttatgc	ctaagcgcca	3420
ggcacggcgt	ttgggtctgc	aaacacagt	gttagtagca	gtctggact	gttccctact	3480
ggcatggaa	ccatcactcc	ccaacatgc	aagcccat	ttaaaggcca	gcctctgccc	3540
cttcagtgat	gctgtgtat	gaaatggcca	tccactatata	tcgaaaattcc	gcaggggcaca	3600
aaacttccag	caagtactg	ttgtgtggaa	atggggcgtg	gggggtgggg	gtcttctta	3660
aaacaggcccc	cttccatct	acccatcc	atgagtcccc	agaggcttca	3720	
gaagctgttg	tctcctct	ggggcagca	gctccctgc	ttggaggcca	aagccccaga	3780
tctctccagc	ccacagctg	aaaacaccaa	gtgcctt	gagggtgtct	gtctggagac	3840
ttagagtttg	tcatgtgt	gtgtgttt	gtttaatgt	gttttatgg	ttttcttct	3900
tttttttttt	ttttttttta	gtctacatta	ggggggaaatg	agcgccccc	atgtgcagac	3960
agtgtgtctt	tatagattt	tctaaaggctt	tcccaatga	tgctcgtaat	ttctgtatgtt	4020
tctgaagttc	ccaggcttc	cacacccgtt	cccatctcac	ttggccaccc	agtgtgacaa	4080
ccctcgggtt	ggatatacc	ccgtggactc	cccgccccca	ctttctataaa	4140	
atgttaggc	agaatacgc	tctctgtgc	aaaactcagc	taagttctg	tttccaccc	4200
gatgttggaa	tatcttatgt	aagagggcag	gggatgtcg	gaagatggca	agaagaacac	4260
agttcaaat	ttctggaaaa	gagctgttg	tggagatcta	aagatgttta	gggaagagct	4320
cgactaaaga	acaatgaaat	aatggtcca	aggggaagtc	a		4361

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

SEQUENCE: 2
 MLRERTVRLQ YGSRVEAVYV LGTYLWTDVY SAAPAGAQTF SLKHSEHVWV EVVRDGEAEE 60
 VATNGKQRWL LSPSTTLRVF MSQASTEASS DKVTVNYYDE EGSIPIDQAG LFLTAIEISL 120
 DVADADRGKV EKNNPKKASW TWGPEGQGAI LLVNCDRDTP WLPKEDCRDE KVYSKEDLK 180
 MSQMILRTKG PDRLPAGYEI VLYISMDSD KVGVFYVENP FFGQRYIHIL GRRKLYHVK 240
 YTGGSAELLF FVEGLCFPDE GFSGLVSIHV SLLEYMAQDI PLTPIFTDTV IFRIAPWIMT 300
 PNILPPVSF VCCMKDNYLK LKEVKNLVFK TNCELKVCFQ YLNNRGRDWI Q DEIEPGYIEA 360
 PHKGFPVVLD SPRDGNLKDFV PVKELLGPDF GYVTREPLFE SVTSLDSFGN LEVSPPVTVN 420
 GKYTYPLGRIL IGSSFPFLSGG RRMKTIVRDF LKAQQVQAPV ELYSDWLTVG HVDEFMSFVP 480
 IPGTKKFLLL MASTSACYKL FREKQKDGHG EAIFMKGLGG MSSKRITINK ILSNESLVQE 540
 NLYFQRCRLDW NRDLILKKELG LTEQDIIDL P ALFKMDEDRH ARAFFPNMVN MIVLDKDLGI 600
 PKPFGPQVEE ECCLEMHVRC LLEPLGLECT FIDDISAYHK FLGEVHCGTN VRRKPFTFKW 660
 WHMVP 665

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
 agcagctctg cagatggggaa ttcttctgtc agtcataatc tgcacatgtg cacaagaca 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	gcacccctat	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	gacatctt	acccgtcc	tctgttcaag	900
atggacgagg	accaccgtgc	cagaccctt	ttccccaa	tggtaacat	gtctgtctg	960
gacaaggacc	ttggcatccc	caagccat	ggggccacagg	ttgaggagga	atgtgcctg	1020
gagatgcacg	tgcggtggct	cctggagccc	ctgggctcg	aatgcaccc	categacac	1080
atttctgcct	accacaaa	tctggggaa	gtccactgt	gcaccaacgt	ccgcaggaa	1140
cccttcac	tcaagtgg	gcacatgtt	cctgaccc	ccaggggccc	ttggctttgc	1200
ctccctcgct	tagtttcca	gaccctt	cacacgccc	gacccttc	ctgacatgg	1260
ctggacagcc	ccgctggag	accttggg	cgtgggg	aatttgggt	atctgtgcct	1320
tgccctcc	gagggggcc	tcagttgt	ctgaaggcat	ccccagtgag	cctcgactt	1380
gtccctcg	aaaatagct	ggccagtg	tctgtagccc	tgacataagg	aacagaacac	1440
aacaaaaaac	agccaaacat	gtccccaa	tgctcccaa	tgcttaatc	tgcttaatc	1500
tgacactgaa	tgaggggaga	agggaggag	attctggat	tgccagttt	tccagcagcc	1560
atgctctgaa	aatcaaggta	gaatccatgg	aaagggacc	caggaccc	ggaccctaga	1620
cgtatctg	actgcac	tcatttc	tcatctcc	tcagggtt	cagggtggca	1680
coccooaaat	ttcatttc	accaac	caatcc	ttggctt	tctgcgt	1740
ggacactt	ggctgtt	ccccactt	tgagggtcc	gtaagtt	ttaaactt	1800
cctggaaaca	tttcatctg	gcagg	ccacgtgtt	gatgtctt	ttgcctcatc	1860
tgtctcgagg	atcgagg	cccccc	ttccccc	agcag	cacatact	1920
gtgacgtt	agttcaat	gtaaaat	ccctgg	ccatccat	tttcttgc	1980
ctgtcc	attcttgc	cgcc	cccttcc	tccact	tgactt	2040
ccttttatt	cttgg	tctgtcc	cagct	agaactt	cgcc	2100
ca	tttccc	ggc	ccat	gtgacc	cccttcc	2160
tccaaacat	aaggacc	gcctt	ggcc	ccat	ggttagt	2220
caacttaggg	gttgg	tatccat	aaat	ccat	ttgc	2280
ctcatttgg	catacc	ttcattt	ggctt	ccact	tttccat	2340
gtttaaaat	ctttag	caga	aaat	ccat	tgaagaa	2400
atttataacc	aatgaa	ttataat	atttattt	ttggagg	aaat	2460
tttgggat	tttcc	ccat	aaat	ttgt	gtat	2520
cacggcgtt	ggct	ccat	acag	ttgt	ccact	2580
catggaccc	atc	ttcc	acat	ttgt	ttgg	2640
tcagtgt	gtctt	ttt	atgc	ttgt	ttgg	2700
acttccac	agtact	ttt	ttt	gggg	tttt	2760
caggcccc	ctccat	ttt	ttt	gggt	tttt	2820
agctgtgt	ttctt	ttt	ttt	gggg	tttt	2880
tcttc	ccat	ttt	ttt	tttt	tttt	2940
agagttgt	ttt	ttt	ttt	tttt	tttt	3000
tttttttt	tttttt	ttt	ttt	tttt	tttt	3060
tgtgttta	tagat	ttt	ttt	tttt	tttt	3120
tgaagtttcc	aggact	ttt	ttt	tttt	tttt	3180
ctcgggttgg	atata	ttt	ttt	tttt	tttt	3240
gtaggccat	aatac	ttt	ttt	tttt	tttt	3300
tttggaaata	tctt	ttt	ttt	tttt	tttt	3360
tttcaaaat	ctgg	ttt	ttt	tttt	tttt	3420
actaaagaac	aat	ttt	ttt	tttt	tttt	3459

```

SEQ ID NO: 4          moltype = AA  length = 350
FEATURE
REGION
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
PLSGGRRMTK 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

SEQUENCE: 5

agtgttgggg ttggggccca cagetaagtc caacaccaggc atgtcgctgc agagaatcgt	60
gcgtgttcc ctggagcatc ccaccaggcg ggtgtgtgtg gttggcggtgg agaccctcg	120
ggacatttat gggtcagtgc ctggggcac aqaaatgttt gaggtctatgg ggacgcctgg	180
cgtggacatc tacatcttc ccaacatggg gaggggccgg gacgtgcag acaccaggcg	240
gtggcgcttt gacgegactt tggagatcat cgtggatcatg aactccccca gcaatgacct	300
caacgcacage catgttcaga tttecttacca ctccagccat gagectctgc ccctggccta	360
tgcgggtgtc tacctcacct gtggatcatc ctctctggat tgcgacccgtg actgtgaggg	420
aaggcaggac aggaactttt tagacaagcg gcaatggggccca gtgggtatgg 480	
cggccatcttg ctgggttaact gtgaccgtga tgatccgagc tggatgttcc aggacattt	
tgaccagcac gtgcactgcc tgcaagaccc ggaagacatg tctgtatgg tcctgcggac	540
gcaggcccctt gtggatccca caaaacttgtc ctccacactt ccagctatga	600
tgcccaaacgg gcacaggctt tccacatctg cgggtctggat gatgtgtgtg aggccatata	660
gcatgtgtg tggccaaatg aggtgtccca tgaggatcccc ccgtgtcatg gggatgagga	720
gcccgttcc tggggaggcc tggcttccca tgatggccggc ttacacaggac tcatctcctt	780
ccatgttact ctgtggaaacg actccaaacgaa ggatccatcg ccatccccca tcttcactga	840
cactgtgttgc ttcccgatgg cacccatggat catgacgcggc agcaactctgc cacccttaga	900
ggtgttatgtg tgccgtgtga ggaacaacac gttttttgtg gatgtgtggg cagagctggc	960
caggaaaggcc ggctgtcaagc tgaccatctg cccacaggcc gagaacccgca acgaccgtgc	1020
gatccaggat gagatggagc tggggatcatgt tcaaggccggc cacaagaccc tcccggttgt	1080
ctttgacttcc ccaaggaaatg gggaaatcgtca ggatccatcg tacaatggaaa ttctgggtcc	1140
agatttttgtt tacgtgactc gggaaaccacg cgacaggctt gtgagttggcc tggactcctt	1200
tggggaaacctt ggggttccatc ctccatgtgtt ggccaatggg aaagagttacc ccctggggag	1260
gatccttattt gggggcaacc tgccctgggtc aagtggccgc agggtcaccc aggtgtgtgc	1320
ggacttccctc catggcaggc aggtgtggcggc cccctgtgggactcttttggg actgtgtggc	1380
cgtggggccat gtggatgttgc ttctgtggatc tggtccctggc cccggatggaa agggtttccg	1440
gatgtctctg gocagccctg gggccctgtt caagctttcc caggaaagac agaaatgtgg	1500
ccacggggagg gcccctctgt tccagggggt tggatgtatg gaggcaggatc agaccatctc	1560
catcaacccat gtgtgttccca ataataacatc catcaactac aataatgttgc tgcaagactg	1620
catcgactgg aaccgttggg tggctgtggc ggagctggggc ttggcaggatgt gatgtatcat	1680
tgacatccca cagctttca agacccggag gaaaaaaacgca acggcccttc tccctgactt	1740
gggtgaacatg tgggtgttgg ggaaggaccc gggcatcccc aagccctttg gcccacatcat	1800
caatggctgc tggctgttgg aggaaaggat ggggtccctg ctggagocgc tggccctccca	1860
ctgcacccctt attagtact tcaatccatc ccacatctg catggggggg tgcactgtgg	1920
caccaatgtg tgcagaaagc cttttttttt caatgtgtgc cctggagacag	1980
ctccccaccca ccattctgtc cccctggggc gggcattggc ccagggttgtg gagacagaga	2040
caggcccttc aacgataaggc accaaaggacg cccaaaggctc cagatggaaactgtgggtg	2100
accgtccctc tcaagaacctt ttcccttggg agtggatccatg ctcacatctgc aacccatgt	2160
gttctcagac ttgaatcttc tggggccccc aaaaaggaaatg accttatttc ttatagctc	2220
tcctgtgtt caacacaacc catggatgttgc tcccttttcc actctgttata catccattt	2280
gggacacatac caccatgggg tctggaaatca tccacatgttgc tcatcggccca tctgtgtctg	2340
tgcatccatc aacggatggg atccatgttgc tggatggatc ccaatgttgc cctctctgca	2400
gaggaaacac cttaaaacca gaccatcca cggaggac gcaaggagaga ttcttcctaa	2460
agccctcccc ataaaaagggg agctgtggat ccacttagat caggggggaa ccatcttca	2520
cccgcccaag ctccctccca gatgtgttgc ctcacccaggc gtgtgttgc acatagtagg	2580
agcttctaga tgcgtgttgc agcaatgaga gttgttccctt acgttataaa actccctatg	2640
atctgacatc cagaatccca gccttgc taaaatctt ggaatctttt ggagacgaaa	2700
gtatctgggg gattgttggg tactagggag actgggtaca agggtgaaaaa gtatgtccca	2760
taatacacat ggttgcactat ggtgtatccat ctgtgtatgg ttaatattag gtgtgttgtt	2820
aagggttgc tttttttttt gggacttctc tctgcaggatc gagagaacgc tggcccttcc	2880
ctggatgttgc ttcaggcttc ctgtggctt ttgttgcggg tttccacatc ctgtgttgt	2940
gcaggagagg gggctaaaggc gctggatcca ccaaggccgc tcaacagccggg aaaactctgg	3000
gaatgaacca ctgaatttgc gggatggggg tggggggccg gttctcgagg tggatgtgc	3060
ctacacgtt gttctgtatg ggttccatc actcgatata aatcaacag	3120
aaacacaaa	3180
	3189

SEQ ID NO: 6	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: 6

MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTGVDI YISPNMERGR	60
ERADTRWRWF DATLEIIVVM NSPSNDLNDs HVQISYHSSH EPLPLAYAVL YLTCDVLD	120
CDLNCEGRQD RNFVDRKQWV WGPSCYGGIL LVNCDRDDPS CDVQDNCDQH VHCLQDLED	180
SVMVLRTQGP AALFDDHKLV LHTSSYDAKR AQVFHICGPE DVCEAYRHVL QGDKVSYEV	240
RLHGDEERFF VEGLSFPDAG FTGLISFHVT LLDDSNEDFS ASPIFTDTVV FRVAPWIMTP	300
STLPPLLEVYV CRVRNNTCFV DAVAELARKA GCKLTIQPQA ENRNDRWIQD EMELGYVQAP	360

-continued

HKFLPVVFDS PRNGELQDFP YKRILGPDFG YVTREPRDRS VSGLDSFGNL EVSPPVVANG	420
KEYPLGRILC CGNLPGSSGR RVTQVVRDPL HAQKVQPVVE LFVWDLAVGH VDEFLSFVPA	480
PDGKGFRMLL ASPGACFKLF QEKQKCGHGR ALLFQGVVDD EQVKTISINQ VLSNKDLINY	540
NKEPVQSCIDW NREVLKRELG LAECDIIDIP QLFKTERKKA TAFFPDLVNM LVLGKHLGIP	600
KPFGPIINGC CCLEEKVRSL LEPLGLHCTF IDDFTPYHML HGEVHCGTNV CRKPFSFKNW	660
NMVP	664
 SEQ ID NO: 7	
FEATURE	moltype = DNA length = 3189
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: 7	
agtgttgggg ttggcggcca cagctaagtc caacaccaggc atgtcgctgc agagaatcgt	60
gcgtgttcc ctggagcatc ccacccaggc ggtgtgttgt gttggcgtgg agaccctcgt	120
ggacatttat gggtcgttgtc ctgaggccac agaaatgttt gaggctatgg ggacccctgg	180
cgtggacatc tacgttcttc ccaacatggc gaggggccggc gacgtgtcgg acaccaggcg	240
gtggcgcttt gacgcgactt tggagatcat cgtggatcat aactccccca gcaatgacct	300
caacgcacgc catgttcaga ttccatccca ctccagccat gagcctctgc ccctggctta	360
tgcgggtc taccttcaccc tggatgttgcat ctcttggat tgcgacatgtc actgtgagg	420
aaggcggaggc aggaacttgc tagacaagcg gcacgtgggtc tggggccca gtgggtatgg	480
cggcatcttg ctgggtactt gtgaccgtga tgatccgagc tggatgttgc aggacaattg	540
tgaccagcac gtgcactgcc tgcaagaccc ggaagacatgtc tggatgttgc ttctggggac	600
gcaggggccctt gcaaggccctt ttgcgttgcata caaacttgc tccttgcacactt ccagatgttgc	660
tgccaaaacggc gcacaggctt cttccatctg cgggtccctggat gatgtgttgtt aggccatata	720
gcacatgtcttgc ggccaaatggat aggtgtccca tggatgttgc cggcttgcatg gggatgagg	780
gcccatttttc ttggaaaggcc tggatgttgc tggatgttgc ttcacaggac tcatactctt	840
ccatgttactt ctgtgttgcactt gacttccaaatggat ggttttcccttgcatccatc ttcttgcacttgc	900
cactgtgttgc ttccaggatgg caccctggatcatgacgcggcc acgacttgc cacccttaga	960
gggtgtatgtt tggccgtgtga ggaacaacac gtgttttgtt gatgcgggtt cagagcttgc	1020
cagggaaaggcc ggctgtcaagc tgaccatctg cccacaggcc gagaaccgc acgaccgttgc	1080
gatccaggatg gatgtgttgc tggatgttgc tggatgttgc cacaagaccc tccctgggttgc	1140
cttgcgttccca ccaaggatgtt gggatgttgc ggttttcccttgcataaaaaaa tccctgggttgc	1200
agatgttgcgtt tacgttgcactt gggaaaccaggc cgacagggttgc tggatgttgc tggacttgc	1260
tggggaaacctt ggggttgcaggcc ctccaggatggt gggcaatgggg aaagagatacc ccctggggag	1320
gatcccttattt gggggcaacc tggccgttgggtt aagtggccggc agggttgcaccggc aggtgtgttgc	1380
ggatcttc ttcatggccaga aggtgttgc cccctgtggatcttgc tctttgtgg actgtgttggc	1440
cgtggccatc ttggatgttgc ttcttgc tggatgttgc cccatggggc tggatgttgc tggatgttgc	1500
gatgtcttc tggccatccgc tggccatgttgc tggatgttgc caccatgttgc tggatgttgc	1560
ccacggggaggc ggccttcgttccatggggatgttgc tggatgttgc tggatgttgc agaccatctt	1620
catcaacccatcgttccatggggatgttgc tggatgttgc tggatgttgc aataatgttgc tggatgttgc	1680
catcgacttgc accgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	1740
tgacatccca cagctttca agaccggatgg gaaaaaaatggatccatc tccctgtacttgc	1800
gtgttgcacatg tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	1860
caatggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	1920
ctgcacccatc attgttgc tcaatccatc cccatgttgc tggatgttgc tggatgttgc	1980
caccaatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2040
ctcccaatccca cccatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2100
caggccatccgc accaaatggatccgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2160
accgttccgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2220
gttcttcacatc ttggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2280
tccatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2340
ggggacaaatccca cccatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2400
tgcatttc tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2460
gagggaaacac cccatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2520
agcttcatccca tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2580
cccgcccaatccca cccatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2640
agcttcatccca tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2700
atctgcacatc tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2760
gtatcttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2820
taataatccatc tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2880
aagggttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	2940
ctggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	3000
gcaggagaggc gggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	3060
gaatgaacca tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	3120
ctacacgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc tggatgttgc	3180
aaacacaaaa	3189
 SEQ ID NO: 8	
FEATURE	moltype = AA length = 664
REGION	Location/Qualifiers
	1..664
	note = protein-arginine deiminase type-3 [Homo sapiens]
source	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 GQDKVSYEVP 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 cagtcagaatc caacaccaggc atgtcgctgc agagaatcg 60
gcgtgttgc ctggagccat ccaccaggcg ggtgtgtggc gttggcgctgg agaccctcg 120
ggacatttat gggtcagtgc ctggggcac agaaatgttt gagggtctatg ggacgcctgg 180
cgtggacatc tacatcttc ccaacatggg gaggggccgg gagcgtgcag acaccaggcg 240
gtggcgcttt gacgcgactt tggagatcat cgtggcattt aactccccca gcaatgacct 300
caacgcacgc catgttcaga tttccatcca ctcacgcattt gacgcatttc ccctggctta 360
tgcgggtgtc taccacatcc ttgttgcattt tgcgcattttt gatgtgtggg 420
aaggcaggac aggaactttg tagacaaggcg gcagtggttgc tggggccca gtgggtatgg 480
cggcatcttgc ttggtaactt gtgaccgtgc tgatccgagc tggatgttcc aggacaattt 540
tgaccgcac gtgcactggc ttgcacggatc ggaagacatc tttgttgcggac 600
gcaggggccctt gacgccttc ttgtatggca caaacattgtc ttccacatcc ccagctatgg 660
tgccaaacggc gcacaggctt tccacatcc ttgttgcattt gatgtgtgttggg 720
gcatgtgtgc ggccaagata aggtgtccca tgaggatccca cgcgttgcattt gggatggg 780
gcgccttc ttggaaaggcc ttgcattttcc ttgtatggccgc ttccacaggac tcatcttc 840
ccatgtcaatc ctgtgtggc acttccaaacgc ggatccatcc ttgtatggca 900
caactgtgttgc ttcccgatgttgc caccctggat catgacgcac gcacacttcg cacccttaga 960
ggtgtatgttgc ttccgtgtca ggaacaacac gtgtttgttgc gatgtgtgttgc 1020
caaggaaaggcc gcgttcacatcc ttgttgcattt gggatgggg 1080
gatccaggatc gagatggggc tgggttgcattt tcaggcgccgc cacaaggccc tccctgggtt 1140
cttgcattcc ccaaggaaatgc gggaaacttgc ggttccatcc ttgtatggcc 1200
agatttttgttgc tacgttgcattt gggaaaccacgc cgacaggatc ttgtatggcc 1260
tgggaacccgtt gagggttgcattt ctcaggatgttgc gggcaatggg aaagatgttcc ccctggggag 1320
gtatcccttgc gggggccaaatc ttgttgcattt gggatggggc 1380
ggacttccttgc catggccaga aggttgcattt cccctgtggat ttgtatggcc 1440
cgtggccatc gtggatgttgc ttcttgcattt ttgtatggcc 1500
gtatgttgc ttggatgttgc gggccatccatc ttgtatggcc 1560
ccacggggagg gccccttgc ttccagggggttgc tggatgttgc gggccatccatc 1620
catcaacccatc gtgttgcattt ataaaggatc catcaactac aataatgttgc ttgtatggcc 1680
catcgacttgc aaccgttgcattt ttgtatggcc 1740
tgacatccca cagtcatttgc ttgtatggcc 1800
ggtaacatc ttgttgcattt gggacatccatc ttgtatggcc 1860
caatggatgc ttgttgcattt gggatgttgc gggccatccatc ttgtatggcc 1920
ctgcacccatc attgtatggcc ttgtatggcc 1980
caccatgttgc ttgtatggcc 2040
ctcccccatttgc cccatccatc ttgtatggcc 2100
caggcccttgc aacgttgcattt gggccatccatc ttgtatggcc 2160
acgttgcatttgc ttgtatggcc 2220
gttcttcatc ttgtatggcc 2280
tgcggatccatc ttgtatggcc 2340
ggggccatccatc ttgtatggcc 2400
tgcatttgcatttgc ttgtatggcc 2460
gaggaaacatc ttgtatggcc 2520
agccctccatc ttgtatggcc 2580
ccggccatccatc ttgtatggcc 2640
agcttgcatttgc ttgtatggcc 2700
atgttgcatttgc ttgtatggcc 2760
gtatgttgcatttgc ttgtatggcc 2820
taatacataatc ttgtatggcc 2880
aagggttgcatttgc ttgtatggcc 2940
ctggatccatc ttgtatggcc 3000
gcaggagagg gggccatccatc ttgtatggcc 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 gtgacgtgtg 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 cacccctggat catgacgcccc agcaactctgc cacccttaga 960
gttgtatgtg tgccgtgtgc ggaacaaacatc gtgtttgtgat gatggggggc cagagctggc 1020
caggaaggcc ggcgtcaacgc tgaccatctg cccacaggcc gagaaccgc acgaccgtg 1080
gatccaggat gagatgtggc 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 tgcttgggtt aatggccgcg agggttacccc aggtgtgtcg 1380
ggacttcttc catggccaga aggtgcggcc cccctggggat ctctttgtgg actgtgttgc 1440
cggtggccat gtggatgtgtt tgccatgtt ccgcgtggc ggggttcccg 1500
gatgtgttgc gccaggccctg gggccatgtt caatgttgcg 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 gaaaaaaatgc acggccatctt tccctgtact 1800
ggtaacatcg ctgggtgtgg ggaagacatcg gggcatccccca aaccatcttgc gcccctatcat 1860
caatggctgc tgctgtgtgg agggaaatgttccatc gggatgttccatc tggatgttccatc 1920
ctgcacccatcg attgtatgttccatc ccacatgttccatc catgggggggg tgcactgtgg 1980
caccatgttccatc tgctgtgtgg aatggatgttccatc aatggatgttccatc 2040
ctcccaaccatcg ccatccatcg cccctggggat gggcatgttccatc ccggatgttccatc 2100
caggccccctg aacatgttccatc accaagatgttccatc cccaaaggatgttccatc cccatgttccatc 2160
accgttccatc tccatgttccatc tttccatgttccatc aatggatgttccatc aaccatgttccatc 2220
gttcttgcacatcg ttgtatgttccatc tggggccccc aaaaagaagg acgttccatc ttatgttccatc 2280
tccatgttccatc aacatgttccatc catggatgttccatc tccctgttccatc actctgttccatc 2340
gggacaaatcg cacatgttccatc tccatgttccatc tccatgttccatc tccatgttccatc 2400
tgcatccatcg cagatgttccatc atccatgttccatc ctgtttgttccatc ccaatgttccatc tccatgttccatc 2460
gagggacaaatcg cctaaaacatcg gaccatgttccatc cggcggatgttccatc 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	tttgtcagcg	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		molttype = 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: 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
KPFPPIINGC	CCLEEKVRSL	LEPLGLHCTF	IDDPFTPYHML	HGEVHCTNV	CRKPFSFKWW	660
NMVP						664
 SEQ ID NO: 13		molttype = 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: 13						
agtgttgggg	ttggggcca	cagtaagtc	caacaccaggc	atgtcgctgc	agagaatcgt	60
gctgtgtc	ctggagcatc	ccaccaggcg	ggtgtgtgt	gttgcgttgg	agaccctcg	120
ggacatttat	gggtcagtgc	ctgagggcac	agaaatgttt	gagggtctatg	ggacgcctgg	180
cgtggacatc	tacatctctc	ccaaacatgg	gaggggccgg	gagcgtgcag	acaccaggcg	240
gtggcgttgc	gacgcgactt	tggagatcat	cgtggatcatg	aactccccca	gcaatgacact	300
caacgcacgc	catgttcaga	tttccatcca	cttcagccat	gagccttcgc	ccctggccta	360
tgcggtgtc	tacccacat	gtgttgacat	ctctctggat	tgcgacactga	actgtgaggg	420
aaggcggac	aggaacttgc	tagacaagcg	gcagtggggtc	tggggggccca	gtgggtatgg	480
cggcatcttgc	tgggtgact	tgacccgtg	tgtatccggc	tgtatgttcc	aggacaatgg	540
tgaccacgc	gtgcactgc	tgcgaacact	qaaagacatg	tctgtatcatgg	tcctcgccgac	600
gcagggccct	gcagccctct	ttgatgacca	caaacttgc	ctccataact	ccagctatga	660
tgccaaacgg	gcacaggct	tccacatcg	cggtcctcg	gatgtgtgt	aggcctataag	720
gcatgtgtcg	ggccaaagata	agggtctca	tgagggtaccc	cgcttgcatg	ggggatggaga	780
gcccgttcttgc	tggaaaggcc	tgtcttccc	tgtatccggc	ttcacaggac	tcatctctt	840
ccatgtca	ctgtggacg	actccaaacga	ggatttctcg	geatccccca	tcttcactga	900
caactgtgttgc	ttcccgatgt	taccctggat	catgacgcggc	agcaacttcgc	cacccttaga	960
ggtgttatgt	tgccctgtgt	ggaacaacac	gtgttttgc	gtatcggtgg	cagagctggc	1020
caggaaaggcc	ggctcgaaac	tgaccatctg	cccacaggccc	gagaacccca	acgccccgt	1080
gatccaggat	gagatggagc	tgggtacatgt	tcaggcgccg	cacaagaccc	tcccggtgtt	1140
cttgcgttcc	ccaaggaaatg	ggggactgca	ggatttccctt	tacaaaagaa	tcctgggtcc	1200
agattttgtgt	tacgtgtact	ggggacccacg	cgacagggtc	gtgagtgccc	ttggactctt	1260
tgggaacatgg	gagggtcagcc	ctccatgtgt	ggccaaatgggg	aaagagtacc	ccctggggag	1320
gatccctatgg	ggggcaacc	tgccctgggtc	aatggccgc	agggttaccc	aggtgtgtcg	1380
ggacttcctc	catgccca	agggtcagcc	ccccgtggag	ctctttgtgg	actgtgttgc	1440
cgtggggccat	gtggatgagt	ttctgagctt	tgtccctgc	cccgatggga	agggttcccg	1500
gatgtcttcgt	gocagccctg	ggggctgttt	caagcttgc	caggaaaagc	agaagtgtgg	1560
ccacggggagg	ggcccttcgt	tccagggggt	tgttgatgt	gagcgttca	agaccatctc	1620
catcaacccat	gtgtcttcca	ataaaagactt	catcaactac	aataagtttg	tgcagactgt	1680
catcgactgg	aaccgtgagg	tgctgaageg	ggagctgggg	ctggcagagt	gtgacatcat	1740
tgacatccca	cagcttcca	agaccgagag	aaaaaaagca	acggcccttct	tccctgactt	1800
ggtgaacatgt	ctgggtgtgg	ggaagcacct	gggcataccc	aagcccttgc	ggcccatcat	1860
caatggctgc	tgctgtctgg	aggagaaggt	gcccgttctc	ctggagccgc	tgggcctcca	1920
ctgcacccat	attgtatgact	tcactccata	ccacatgtgt	catggggagg	tgcactgtgg	1980

-continued

caccaatgtg	tgcagaaa	g	ccttctttt	caagtgg	tg	aaacatgg	tc	cctgagac	ag	2040					
ctccccc	ccat	tcgtc	ccccctggg	gggcatt	ggc	ccagg	tg	gagac	aga	2100					
caggccc	ctg	aac	gataa	acaag	aga	ccaagg	gctc	cagat	gaa	2160					
accgtcc	ctc	tca	gaagc	ctt	ccctg	gag	tccat	actg	agggt	2220					
gttctc	agac	ttg	aatctt	teg	gcccc	aaa	agaagg	ac	ctcatt	ttatag	c	ctc	2280		
tctgtg	tatt	ca	acaca	acc	atgtg	ccccc	ttctc	act	ctgaa	at	cat	ccatt	tg	2340	
gggacaa	atc	ccat	ttggg	tct	atc	aaaca	ccac	gtat	ccat	cc	tct	ttgt	ctcg	2400	
tgcac	tctt	aa	ca	gagg	aa	atc	atc	atc	atc	atc	atc	atc	atc	2460	
gagga	aca	ac	ctaa	acca	actca	cgagg	ac	gagg	aga	ta	ttt	ctct	ctaa	2520	
agc	ctcccc	at	aaaa	aa	gg	atgt	ggat	ccact	tag	cagg	ggaa	ccat	tttca	2580	
ccggc	caag	cc	ctct	ggcc	atgt	gac	ct	accc	agg	gtg	atgt	tc	atag	tagg	2640
agctt	ctaga	tg	cat	gtg	gat	gca	g	atgt	ggat	ttgt	ccct	ttt	atc	ccat	2700
atctg	acat	g	aga	atcca	gc	ttgt	gtc	aa	atctt	ttt	ggag	acg	aaa	2760	
gtatct	gggg	g	att	ttggg	ta	ctact	gg	gggt	aaaa	gtat	tttcc	cc	tttcc	2820	
taata	acat	at	ttgt	atc	atc	atc	atc	atc	atc	atc	atc	atc	atc	atc	2880
aagg	ttgt	atc	tat	ttgg	cc	ggact	tc	tct	atc	atc	atc	atc	atc	atc	2940
ctggatt	gggt	ctc	agg	ct	ctg	ttgg	cc	ttt	ccat	tc	ctg	ct	ctg	ct	3000
gcagg	agagg	gg	ggct	taa	gg	gtt	ggat	cc	aa	ac	ag	cc	gg	aa	3060
gaat	gaac	ca	cta	att	ca	ggg	at	ggg	ggg	gg	ttt	ctc	gagg	tttgc	3120
ctacac	gt	ttt	gtat	g	gtt	ccag	ct	gtt	ccat	act	cg	cta	aat	caac	3180
aaac	acaa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	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
MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTGVDI YISPNMERGR 60
ERADTRRWRF DATLEIIVVM NSPSNDLNDs HVQISYHSSH EPLPLAYAVL YLTCVDISLD 120
CDLNCEGRQD RNFVDKRQWV WGPSPGYGGIL LVNCDRDDP CDVQDNCDQH VHCLQDLED 180
SVMVLRTQGP AALFDDHKL V LHTSSYDAKR AQVFHICGPE DVCEAYRHVL GQDKVSYEV 240
RLHGDEERFF VEGLSFPDAG FTGLISFHVT LLDDSNEDFS ASPIFTDVV FRVVPWIMTP 300
STLPPLLEVYV CRVRNNTCFV DAVAELARKA GCKLTICPQA ENRNDRWIQD EMELGYVQAP 360
HKTLPVVFDs PRNGELQDFP YKRLILGPDPF VYTREPRDRS VSGLDSFGNL EVSPPVVANG 420
KEYPLGRILI CGNLPGSSGR RVTQVVRDFL HAQKVQPPVE LFVDWLAVGH VDEFLSFVPA 480
PDGKGFRRMLL ASPGACFKLQ QEKGKCGHGR ALLFQGVND EQVKTISINQ VLSNKDLINY 540
NKFVQSCIDW NREVLKRELG LAECDIIDIP QLFKTERKKA TAFFPDVLNM LVLGKHLGIP 600
KPGPIINGC CCLEEKVRSL LEPPLGHCTF IDDFTPYHML HGEVHCCTNV CRKPFSFKWW 660
NMVP
664

```

```

SEQ ID NO: 15      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: 15
agtgttgggg ttggggccca cagctaagtc caacaccaggc atgtcgctgc agagaatcg 60
gcgtgttgc ctggagcatc ccaccaggcg ggtgtgtgt gctggcgtgg agaccctcg 120
ggacatttat gggctagtgc ctgaggcac agaaatgtt gaggtctatgg ggacgcctgg 180
cgtggacatc tacatctt ccaacatggg gaggggccgg gagegtgcac acaccaggcg 240
gtggcgcttt gacgacatc tggatcatc cgtggatcatc aactccccca gcaatgcac 300
caacgacagc catgttcaga ttctctacca ctccagccat gagectctgc ccctggctca 360
tgcgggtc tacctcacatc gtgtgtatc ctctctgtatc tgccacctga actgtgtggg 420
aggcaggac aggaactttg tagacaagcg gcagtgggtc tggggccca gtgggtatgg 480
cggcatcttg ctgggtactt tgacccgtatc tgatccggc tggatgttcc aggacaattt 540
tgaccagacatc gtgcacttgc tcgaaagacatc gtaatgtatgg tcttgcggac 600
gcagggccctt tcgatgtatc cttttttttt ccacatccatcc ccacgtatcg 660
tgccaaacccg gcacaggctt tccacatctt cgggtccatcg gatgtgtgtg aggccatag 720
gcatgtgtgc ggcacatc aggtgttccatc tgatgttccatc cgatgtatcg gggatgtatcg 780
gcgccttc ttgcgttccatc tgatgttccatc tgatgtatcg tccatccatc 840
ccatgtactt ctgtgttccatc acttccatcg ggatgttccatc gatccatccatc tcttacttgc 900
cactgtgttgc ttccatcgatc catgtatcgatc agcactctgc cacccttgc 960
gtgtgtatgttgc tccatcgatc gatgtgtgtg gatgtgtgtg aggccatag 1020
caggaaaggcc ggtgtatcgatc tgatgtatcgatc cccacaggcc gagaaccggca acgaccgtcg 1080
gtatccatcgatc gatgtgtatcgatc tgatgtatcgatc cccacaggcc tccatccatc 1140
ctttacttgc ccaaggaaatcgatc gggaaacttgcatc ggatgttccatc tacaatggaa tcttgggtcc 1200
agatgttgcatc gggaaaccacg cgacaggatc gtgtgtgttccatc tccatccatc 1260
tggaaacccg gggatgtatcgatc tgatgtatcgatc gggaaatgggg aaagatgtatcgatc ccctggggat 1320
gatccatcgatc gggatgtatcgatc tgatgtatcgatc aagtgccatcgatc aggtgtatcgatc 1380
ggacttccatc catgtatcgatc cccatcgatc tccatccatc 1440

```

-continued

```

cgtggccat gtggatgagt ttctgagctt tgcctggcc cccgatggga agggcttcgg 1500
gtatgctctg gocagccctg ggccctgtt caagctttc caggaaaagc agaagtgtgg 1560
ccacggagg gccccctgt tcacgggggt tggtgatgt gacggatca agaccatctc 1620
catcaacccat gtgcctcca ataaaagacct catcaactac aataagttt tgccagactg 1680
catcgactgg aaccgtgagg tgctgaagcg ggagctggc ctggcagagt gtgacatcat 1740
tgacatccca cagctttca agacggagag gaaaaaaaaaca acggccttc tccctgactt 1800
ggtaacatg ctgggtgtgg ggaagcacct gggcattccc aagecccttg ggcacatcat 1860
caatggctgc tgctgtctgg aggaaaggt gccgtccctg ctggagcgc tggcctcca 1920
ctgcacccctt attgtactt tcaactccata ccacatgttgc catggggggg tgcactgtgg 1980
accaatgtg tgccaaaggg ccttttctt caagtgttgg aacatgttgc cctgagacag 2040
cttccacccat ccatctgtc cccctgggg gggcatttggc ccagggtgg gagaagagaga 2100
caggccccctt aacgataaagc accaagagac cccaaaggcgc cagatgaaac actgagggtg 2160
accgtccctc tcaagaagctt ttccctggta agtgtccatg ctcacactc aaccatgtg 2220
gttctcaacat ttaatcttc tccggccccc aaaaagaagg acctcatttc ttatagctc 2280
tctgtgatt caacacaacc catggatgttcc tcccttcttc actctgaaat catccattt 2340
gggacaaaatc cacatgggg tctatggaaaatc tccacgtatc tcatcggcca tcttgtctg 2400
tgcatcttaa cagggaaagg atccatgtt ctgctttggg ccaatttgc tctctctgc 2460
gagggaaacac cttaaaaacca gaccacttca cgcaggacag gcaggagaga ttcttcctaa 2520
agecttcggcc aaaaaaaaggg agctgtggat cctgttagat caggggggaa ccatcttca 2580
ccccggcaag ctccgtccca gatgttgcctt ctcacccagc gtgagctgc acatgtagg 2640
agttctaga tgcgtgtgg agcataggaa gttgtccctt agcattataa actcccccattg 2700
atctgacatg cagaatccca gcttgcata gaatccctcc ggaattttttt ggagacggaaa 2760
gtatctgggg gatgtgggg tactggggg actgggttaca agggtgaaaaa gtatgtccca 2820
taatacacat gttgtactat ggtatccat cttgtatgg ttaatattttt gtgtctggag 2880
aagggtgtt cattttttccctt gggactttcc tctgcaggag gagagaacgc tgcctctcc 2940
ctggattgtt ctcaggctt ctgtttggct tgggtcaaggc tttccacatc ctgtctgtct 3000
gcaggggggg gggctaaaggg gctggatcat ccaaggcgc tcacaggggg aaaaactctgg 3060
gaatgaaccca ctgaattttag gggatggggg tggggggggc gttctcgagg tttgtccag 3120
ctacacgtgtt gttctgtatg ggttccagctg ctgtttccatc actcgatataa aatcacacaa 3180
aaacacaaaaa 3189

```

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
 MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTGVDI YISPNMERGR 60
 ERADTRRWRF DATLEIIVVM NSPSNDLNDs HVQISYHSSS EPLPLAYAVL YLTCVDISLD 120
 CDLNCEGRQD RNFVDKRQWV WGPSCYGGIL LVNCDRDDPS CDVQDNCDQH VHCLQDLEDm 180
 SVMVLRTQGP AALFDHHDKLV LHTSSYDAKR AQVFHICPPE DVCEAYRHVL GQDKVSYEVP 240
 RLHGDEERFF VEGLSPFPDAG FTGLISFHVT LLDDSNEDFS ASPIFTDTVV FRVAPWIMTP 300
 STLPPLEVYY CRVRNNTCFV DAAVELARKA GCKLTICPQA ENRNDRWIQD EMELGYVQAP 360
 HKTLPVVFDs PRNGELQDFP YKRLGPDFG YVTREPRDRS VSGLDSFGNL EVSPPVVANG 420
 KEYPLGRILII CGNLPGSSGR RVTQVVRDPL HAQKVQPVPVE LFVWDLAVGH VDEFLSFVPA 480
 PDKGKFRMALL ASPGACFKLF QEKQKCGHGR ALLFQGVVDD EQVKTISINQ VLSNKDLINY 540
 NKFVQSCIDW NREVLKRELG LAECDIIDIP QLFKTERKKA TFFFPLDVNM LVLGKHGIP 600
 KPFGPIINGC CCLEEKVRSL LEPLGLHCTF IDDFTPYHML HGEVHCGTNV CRKPFSFKNW 660
 NMVP 664

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
 agtgttgggg ttggcccca cagctaagtc caacaccaggc atgtcgctgc agagaatcg 60
 gctgtgttcc ctggaggatcc ccaccaggcgc ggtgtgtgg gtcggctgg agaccctcg 120
 ggacatttat gggctgtgc ctggaggcac agaaatgtt gagggtatgg ggacggctgg 180
 cgtggacatcc tacatcttc ccaacatggc gaggggccggc gagcgtgcac acaccaggcg 240
 gtggcgtttt gagcgcactt tggatcatc cttgtgtatg aactccccca gcaatgcatt 300
 caacgcacaggc catgttcaga ttccatcca ctccagccat gaggcttgc ccctggctca 360
 tggcgtgttc tacatcaccc tggatgttgcat cttctgttgc tgcgcacttgc actgttgggg 420
 aaggcaggac aggaactttt tagacaaggcg gcaatggggc tggggccca gtgggtatgg 480
 cgccatcttgc ctgggttgc gtcggccactt tgatccgc gttgtatgc agggaaattt 540
 tgaccaggacat gtcgtacttca tgcaaggactt ggaagacatg tttgtatgg tcctggcggac 600
 gcaggccccctt ctggcccttc ttgtatgcacca caaaactgttc ctccatccat ccagctatgt 660
 tggccaaacggc gtcacaggatcc tccatcatcg cggcccttgc gatgtgtgtg aggccatag 720
 gcatgtgtcg ggcacaaagata aggtgtccca tgatgttacc cgttgcacatg gggatggagga 780
 gtcgtttttc gtggaaaggcc ttttcttc tgatgccc gttcagggac tcatctctt 840
 ccatgtactt ctcgtggacg actccaaacgca ggatttctcg gcatccccca tcttcacttgc 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
ggacttctc	catgcccaga	aggtgcagcc	ccccgtggag	ctctttgtgg	actgtgttgc	1440
cgtggcccat	gtggatgatg	ttctgagct	tgcctctgc	cccgatggga	agggttccg	1500
gatgtctcc	ggccaggctg	gggcgtctt	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	acggccttct	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	aacatggtgc	cctgagacag	2040
ctcccccacca	ccatctgtc	ccccctgggc	gggcattggc	ccaggtgttgc	gagacagaga	2100
caggccccctg	aacgataagc	accaggagac	cccaaggctc	catatggaa	actgggggt	2160
accgtccctc	tcaagaac	tttcccttga	agtgtccat	cttcaccc	aacccatgtg	2220
gttctcagac	tgttaatctt	tcggggccccc	aaaagaagg	acccatattc	ttatagctc	2280
tccctgtatt	caacacaacc	catggatgt	tccccc	actcttgc	catccatttgc	2340
gggacaaatc	cacattgggg	tctagaaaca	tccacgtatc	tcatcagcca	tcttgtctgt	2400
tgatctcaa	cagaggaaagg	atccatgtt	ctgttttgtt	ccaatttgc	cctctctgc	2460
gaggaacaaac	cataaaacca	gaccactca	cgcaggacag	gcaggagaga	tttccctaa	2520
agcctccccc	ataaaaagg	agctgtggat	ccacttagat	caggccggaa	ccatcttca	2580
ccoggccaa	ctccgtccca	gatgtgacc	ctcaccc	gtgagctgc	acatagtagg	2640
agttcttaga	tgcatgttgc	agcaatgaga	gttgcctt	agccttataa	actccccatg	2700
atctgatcatg	cagaatccca	gcctgttgc	gaatccctt	ggaatttctt	ggagacgaaa	2760
gtatctgggg	gattgttgg	taactgggg	acttgggtaca	agggtgaaaa	gtagtccca	2820
taatacacat	ggttgcactat	ggttgcac	cttgcgtatgg	ttaatattag	gtgtgtggag	2880
agggttgc	cattggccct	gggacttctc	tctgcaggag	gagagaacgc	tgcctctcct	2940
ctggattgtt	ctcaggctt	ctgttggcc	tttgcacatc	ctgtctgtct	3000	
gcaggagagg	gggtcaaggg	gctggatcca	ccaaggccgc	tcacagcggg	aaaactctgg	3060
gaatgaacca	ctgaatccag	gggtatgggg	tggggggggc	gttctcgagg	tgtgtgc	3120
ctacacgtgt	gttctgtatg	ggtccagtc	cgtttccatc	actcgcta	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 aaaaaaaaaa gggcgtgcac acaccaggcg	240	
gtggcgctt gacgcgactt tggagatcat cgtggatcat aactccccca gcaatgcac	300	
caacgcacagc catgttccatca ctccagccat gggcccttc ccctggccat	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	gtatgggtgg	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	ctccatgtgt	ggccaaatgg	aaagactacc	ccctggggag	1320
gatccctatt	ggggcaacc	tgcctgggtc	aagtggccgc	agggtcaccc	agggtggcgc	1380
ggacttcctc	catgcccaga	aggtgcagcc	ccccgtggag	ctcttttgttgc	actgtgttgc	1440
cgtggcccat	ttggatgagt	ttctgagcttgc	tgtccctgc	ccccatggga	agggtttccgc	1500
gatgctctgc	gcccggccctg	ggggctgttt	caagcttc	caggaaaaggc	agaagtgtgg	1560
ccacggggagg	gcccctctgt	tccaggggg	tgttgcgtat	gagcaggtca	agaccatctc	1620
catcaacccag	gtgtcttcca	ataaaagacct	catcaactac	aataagtgttgc	tgcaagactgc	1680
catcgactgg	aaccgtgagg	tgcgttgc	ggagctgggc	ctggcagact	gtgacatcat	1740
tgacatccca	cagcttca	agacccggag	aaaaaaaggca	acggcccttc	tccctgacttgc	1800
ggtaaacatg	ctgggtgtgg	ggaagcacct	gggcateccc	aaggcccttgc	ggcccatcat	1860
caatggctgc	tgctgtctgg	aggagaaggt	gcaacttgc	ctggagccgc	tggccttcca	1920
ctgcacccatt	attgtgact	tcactccata	ccacatgtgt	catggggagg	tgcaactgtgg	1980
caccaatgt	tgcagaaagc	ccttcttctt	caagtgggttgc	aacatgggtgc	cctgagacag	2040
ctcccaaccca	ccatcttgc	ccccctgggc	gggcatttgc	ccagggtgggc	gagacagaga	2100
caggccccctg	aacgataaagc	accaagagagc	ccaaaggcttc	catatggaaac	actgggggtgc	2160
accgttccatc	tcaagaaaggc	tttccctgg	agttgtccat	cctcaccatgc	aaccatgtgc	2220
gttctcagac	ttgaatcttc	tcgggggggg	aaaaaaaggaa	acccatatttc	ttatagctc	2280
tccctgtatt	caacacaacc	catggagatg	tccccttctc	actctgaaat	catccatttg	2340
gggacaaatc	cacattgggg	tctagaaaca	tccacgtatc	tcatagcc	tcttgtctgc	2400
tgatcttcaa	cagagaaagg	atccatgttgc	ctgcttttgttgc	ccaatttgc	cctctctgc	2460
gaggaacaaac	cctaaaacca	gaccatcca	cgccaggagac	gcaggagaga	ttcttcctaa	2520
agctcccccc	ataaaaaggaa	agctgtggat	ccatcttgc	caggccggaa	ccatcttca	2580
cccgcccaag	ctccgtccca	gatgttgc	ctcaccacgc	gtgagctgtc	acatagtagg	2640
agcttctaga	tgcgttgc	agcaactgaga	gttgccttgc	agccatataa	actccccatgc	2700
atctgcacat	cagaatattca	gccttgc	gaatcccttc	ggatattttc	ggagacgaaa	2760
gtatctgggg	gattgttggg	tactaggag	actgggtaca	agggtaaaaa	gtatgtccca	2820
taatacacat	ggttgtactat	ggtgatccac	cttgcgtat	ttaatattag	gtgtctggag	2880
aaagtgttgc	cattggccct	gggacttgc	tctgcaggag	gagagaacgc	tgccctcttc	2940
ctggatttgtt	ctcaggcttc	ctgttgc	ttggtcagcg	ttccacatc	ctgtctgc	3000
gcaggagagg	gggctaaggg	gctgtatcca	ccaaaggcgc	tcacagcggg	aaaactctgg	3060
gaatgaacca	ctgaatttgc	gggatgggg	tggggggggc	tttctcgagg	tgtgtgc	3120
ctacacgtgt	gttctgtatg	ggtccagctg	cgtttccatc	actcgctaat	aatcaacag	3180
aaacacacaaa						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 QGDKVSYEVP 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 QEOKQCGHGR 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
tctccctgg tctgccattt cccctgtggg gggctgccc gggccaaaca ggccagagg 60
ctccaccagg cctggaaaattt acaggcttg ggagtagggg tgccagataa aataaaagat 120
gcctagttaa aattgaattt caggtaataa attaataat ttttagtat aagtgtatcc 180
caaataacttc acgggagttc caaacaat ttgactgttc acttggaaattt caagtccaaac 240
ttagcttttgc gtatTTTAT ttgctaaatc ttggcagccct gtttagggg agaaaacagg 300
gtgtctgtcc ctcccaactg gtgaggacac ctgacaggggg cccagacacc agagtggac 360
ctgcctgtatc gagagaaatgt tgggttggg ttatTCCTCCTC tttcataaaa aagtgtttt 420
gagatgttaag ayaatcatgt tgggttggg ggtggaaatg tgggtttggg gaggttgc 480
ttactctgag ggaagatggg aagegggggg agtggccacg tggtaatattt tgggttagtc 540
ccctacgtat cctgtgcctc agtttccccca tggtaatataa ggagtaatattt tctcatctg 600
gggctgcagt gaggatcaga ttggactata taacaaaagg cttaggcagg cctgcgggtga 660
taagagtca aaggctgggg getttttatg ttatggatattt gttgttataa ttactccctgg 720
ggaggtgtgtt ggcaaaatgt tggactgagg agccacatgg ctgggtttgg atccccaggcc 780
ggccacacacg ctggggcaag ttacttaacc tcgggttttcatctgtata atggggagaca 840
ataatagata gtgcctccctc tctgggttgc ttggagggat tgcattatgg gctggcacac 900
agcatgagct ctagtaatgtt cagtcactgtt atatccccca gggcatgtcat gggcgtgcct 960
ccaggaggcc ctcccaactcg ctaggtggcc tccctacatg ggttcccaactt gggctgtat 1020
cctgtggcct tgagccccat ctgggtggcc ttggacaatgtt tggctctcaatcccccaggcc 1080
ggatcaacacg tggggagggc agtgggggttgc tctggccggc aggcaatggc tgagtgacaca 1140
gtggggcactcc caccatgtttaa ggtatggatcatc tctgggtttcc aaggaggaaac cccagggttag 1200
gttggatgag ggaaatggaa ttctcaggcc agataacatc acagatctttaaaccctgca 1260
attaatttctt ggttggtttgc ccaatccggg ccaagtttcc cgggttgacac tgcaagatgt 1320
tggAACACGA tggggcttggg gtcaaacatgtt tggcacatctt ggctctgcata tggatgtca 1380
cttcagggtt gtcacaggcc ctccatgttgc cttaggttgcatt aatatgtccaa atggggatgt 1440
tattgttcac cttagatgtt gtaaaggatgtt tggggatgtc tggccgggg tgcctggcat 1500
gcagcggggca ctcaaccaggat ttgaaattccc ttcccaggggg agaattcaag gatatctcc 1560
caccactaa ggagatcatgtt gatggggatctt ggggtggaaat ttgtgtttttt taaagctctc 1620
agaacttgaa gggggccatcc cccaaatgttgc ctgggtttccat ctggatgttctt ctggccggaga 1680
cttcggactgg gggatattatcg agcacatggg tggcttctga ctcccaaaatgaa gctggccatt 1740
cttgcgtatcg cttaggttgc cagaatatttgc ttggatgttcat gggggccctt gctggacat 1800
ctacatctctt cccaaatcatgg agggggccggc ggagcgtgcata gacaccaggc ggtggcgctt 1860
tgcgttgcactt tggatgttgc tgcgtgttgc gaaactcccccc acgaatgttca tcaacggac 1920
ccatgttgcactt atttccatcc tcccaatggc ttggatgttgc cccatggctt atgcgggtct 1980
ctacctcacc tggatgttgc tctctctggat ttggggatgttgc aacttgttggg gaaggccaggaa 2040
caggaaactttt gtagacaaggc ggcagttgggctt ctggggggccctt agtgggtatgtt gggccatctt 2100
gtctggtaac gttggatgttgc atggatccggc tggatgttgc cggacaattt gtgaccac 2160
cgtgcgtactgc tggcaacaccatc tggatgttgc tgcgtgttgc gttccctggggc cgcaggggccc 2220
tgcggccctt tggatgttgc accaaatgttgc cccatggctt cccatggctt atgcggatgtt 2280
ggcacaggtc ttccacatgtt ggggttgc ggtatgttgc gggccatgtt ggcacatgtt 2340
ggggccaaatg aaggatgttgc atggatgttgc cccatggctt gggggatgttgc agcgttctt 2400
ctgtggatggc ctgttttccctt ctgtatggccgg cttagatggggatccatcttccatgttgcac 2460
tctgtggatggc gactccaaatgttgc agggatcccttccatcttccatcttccatgttgcac 2520
gttccggatgttgc tggccatgttgc tcatgttgc cggactcttgc ccaccccttgc aggtgtatgt 2580
gttggccatgttgc tggccatgttgc cggactcttgc tggccatgttgc ccaccccttgc aggtgtatgt 2640
cggtgtcaag ctggatgttgc tggccatgttgc cggactcttgc tggccatgttgc ccaccccttgc aggtgtatgt 2700
tggatgttgc tggccatgttgc tggccatgttgc cggactcttgc tggccatgttgc ccaccccttgc aggtgtatgt 2760
cccaaggaaatgttgc agggatcccttccatcttccatcttccatcttccatgttgcac 2820
ttacatgttgcactt cggggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 2880
ggggatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 2940
tggggccaaatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3000
ccatgttgcactt cggggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3060
tggatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3120
ggccatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3180
ggcccttccatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3240
gggtgttgcactt aataaaatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3300
gaaccatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3360
acagcttccatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3420
gttggatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3480
ctgtgttgcactt cggggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3540
cattgtatgttgc ttcactccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3600
gttggatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3660
accatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3720
gaacatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3780
cttgcgttgcactt cggggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3840
cttgcgttgcactt cggggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3900
tcaacacaacatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 3960
ccacatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 4020
acagatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 4080
ccctaaaatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 4140
cataaaaaatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 4200
gttggatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 4260
atgtgttgcactt cggggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 4320
ggagatgttgc tggccatgttgc tggccatgttgc tggccatgttgc ccaccccttgc aggtgtatgt 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	tggcagccct	gctttaggg	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	tttgcatttt	ttactcttgg	720
ggaggtgtgtat	ggcaagagtat	tggactgg	agccacatgc	ctgggtttgg	atcccacccc	780
ggccacacagat	cttggggcaat	ttacttaacc	tcggcttttct	catctgtata	atgggagaca	840
ataatagata	gtgcctccat	tcttagggta	tttgcaggat	tgcataatgg	gtcgccacac	900
agcatagatgtat	cgatgtatgt	cgttcactgt	tatccccca	gggcattgtat	gccgcgtcc	960
ccaggaggccc	ctccaaactcg	cgtggccctt	tccttacatgc	ggttcccaact	gagctgtgtat	1020
cctgtggccat	ttagccccat	ctgggtggcc	tggacatgt	tgttgcctcg	ctccccacgc	1080
ggatcaacatgc	tggggagatgc	agttgggttgc	tgcgtccgc	aggcgttacc	tgatgtccaa	1140
gtggggccatcc	cacagtcaaa	ggatgtatgc	tctggtttcc	aaggaggaac	cccaggctat	1200
gttggatgtat	ggaaatgtatgc	tctcaggccc	agataatccat	acagatattt	taaacctgc	1260
attaaattttgtat	gtttgttgc	ccaaatccgg	ccaaatccgg	cggtgtgacac	tgcagatgt	1320
tggaaacatgc	tgggttggat	gtcagatgtat	tttgcacatct	ggctctgc	tgtgtatgc	1380
cctcaggatgtat	gtcactggatc	tccttgcgc	tcctgtatgc	aatatgc	atggggatgt	1440
tattgtctatc	tttacatgtat	gttgcgtatgc	ggggatgtgc	gtgcgtccagg	tgcctggcat	1500
gcacgtggccat	cttcaacatgc	tttgcaggccc	ggatgtatgc	gatatcttc	tttgcaggccc	1560
cacccactatgc	ggagatgtatgc	gtatggatgc	gggggtgagaa	tttgcctttt	taaagcttgc	1620
agaactgtatgc	ggggctccat	ccaccaatgg	ctgggtgtca	ctgagcttgc	ctgtgtccat	1680
cctgcactggat	ggatTTATGC	agcacatgc	tggcttctgc	cttccgaaaga	gctgcgttgc	1740
ctgggtcgtatgc	cttgcaggatgc	cagaatgtat	tgaggatctat	gggacgcctg	gcgtggacat	1800
ctacgttctatc	cccaacatgc	agagggccgg	ggagctgtgc	gacacaggcg	ggtggccgtt	1860
tgacgcgtatgc	tggggatgtat	tttgcaggatgc	gggggtgagaa	tttgcctttt	taaagcttgc	1920
cctatgttgc	atttcttacc	actccatgc	tgagcttgc	ccccctgcct	atgcgtgtat	1980
ctaccttaccat	tgtgttgc	tctcttgc	tttgcacatgc	aactgttgc	ggaggccat	2040
caggaaacttgc	gtagacaatgc	ggcgttgggt	ctggggccccc	atgtggatgt	gccccatctt	2100
gttgggtgtatgc	atgtatgc	ctgtgtatgc	caggacaattt	gtgaccatgc	2160	
cgtgcactgc	ctgcaagatgc	tggaaagatct	gtctgtatgc	gtcctgcgg	cgccaggccc	2220
tgcaggccatc	tttgcattatc	acaaatcttgc	cctccatacc	tccagatgtat	atgcacaaac	2280

-continued

```

ggcacaggtc ttccacatct gccgtcttga ggatgtgtg gaggcctata ggcatgtgt 2340
ggcccaagat aaggtgtctt atgaggtaac ccgcttgcat ggggatgagg agcgtttt 2400
cgtggaaaggc ctgtcttcc ctgatggcg gttccacaggat ctcatctct tccatgtcac 2460
tctgtctggac gactccaacg aggattctc ggcattccc atcttactg acactgttgt 2520
gttcccgatg gcacccctgga tcatgaecgc cagcactctg ccacccttag aggtgtatgt 2580
gtgccgtgtg aggaacaaca cgtgttttgtt ggtatgcgggt gcagagotgg ccaggaaggc 2640
cggctcaagc ctgaccatct gcccacaggc cgagaaccgc aacgaccgtt ggatccaggat 2700
tgagatggag ctgggttacg ttcaggcgc gacacaagacc tcctcggtgg tctttactc 2760
ccccaggaaat ggggaaactgc aggatttccc ttacaaaaga atcttgggtc cagatttgg 2820
ttacgtgact cggggatggc ggcacagggt tggtgatggc ctggactctt ttgggaaacct 2880
ggaggctcgc cttccactgg tggccatgg gaagagatcc cccctgggga ggatccat 2940
tgggggcaac ctgcctgggt caagtggccg cagggtcacc caggtgggtc gggacttcc 3000
ccatggcccg aagggtcagc cccctgttg gctctttgtt gactgggtgg ccgtggggca 3060
tgtggatgg tttctgtacg ttgcggctgc cccctgtggg aagggttcc ggtatccct 3120
ggccagecctt gggggctgtc tcagaactt ccacggaaaag cagaatgtgtt gccacgggag 3180
ggccctcttgc ttcccaatggg ttgttgcata tgacggatc aagacatcc ccatcaacca 3240
ggtgctctcc aataaaagacc tcatacaacta caataagttt gtgcagatc gcatcgactg 3300
gaaccgttagt gtgtctgaaac gggagctggg ctggccggag tgcgtatca ttgatcc 3360
acagcttc aagaccggaa gggaaaaaaaac aacggcttcc ttccctgtact ttgtgaaacat 3420
gtctgggtgtc gggggatggc tggccatcc caacggccctt gggccatca tcaatggctg 3480
ctgtgtgtctt gggggatggc tggccatcc gctggagccg ctgggctcc actgcaccc 3540
cattgtatgc ttccatccat accacatgtt gcatgggggg gtcgtactgt gcaccaatgt 3600
gtgcagatgg acccttc ttcactgttggt gacatgtgtt cocttgcata gctccaccc 3660
accatctgtt cccctgggg cggcattgg cccagggtgtt ggagacagag acaggccctt 3720
gaacgataag cacaagaga ccccaagggtt ccagatggaa cactggggg gaccgtccct 3780
ctcagaagcc tttccctgg aagtgttccat gcttcacccatgtt gtttctcaga 3840
cttgaatctt ctcggccccca caaaaagaag gacttcattt ttatagctt ctctgtgtat 3900
tcaacacaaac ctcgttggat gtccttc ttcactgttgc tcatccattt gggggacaat 3960
ccacattggg gtcttggaaatc acacgttat ctcacatccat ttcattgttgc ttgcaccc 4020
acagagggaaat gatccatgtat tctgtttttgg tccaatttgc ttctcttgc agggaaacaa 4080
ccctaaaacc acggactcc acggcaggaga ggcaggagat atttccat aagccctcc 4140
cataaaaagg ggtgtgttgc tccacttgc tcaaggccggg accatccccc accccggccaa 4200
gctccctgccc agatgttgc ctcacccatgtt gtcgtactgtt cacaatgtt gagttcttag 4260
atgcatgtgg aagcaatgtat agttgtccctt tagccttata aactcccat gatctgacat 4320
gcagaaaatcc acccttgcggcc agaatccctt tggaaatttgc ttggagacgaa agtattctggg 4380
ggatttgggtt gtaactgtggaa gactgggtac aagggtgaaat agtagttccc ataataacaca 4440
tgggttactt tgggtgtatcc cctgttgcata gttatattt ggtgtctggaa gaagggttgc 4500
tcattggccc tggggacttcc tctgtcgaggaa ggagagaacg ctgccttcc tctggattgg 4560
tctcaggdtc tctgttggcc ttgcgttgcgc ttgcgttgcgc ttgcgttgcgc 4620
ggggctaaagg ggtgttgcgc accaacaatggc ctcacagccg gaaaactctg ggaatgaacc 4680
actgaattca gggggatgggg gtttttgcgg gttgtgtgcgc gttacacgttgc 4740
tgttctgtat ggggtccat cactcgatca taaaatcaaca gaaacacaaa 4800

```

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
 MFEVYGTGIV DIYVSPNMER GRERADTRRW RFDATELIVVMNSPSNDLN DSHVQISYHS 60
 SHEPLPLAYA VLYLTCVDIS LDCCDLNCEGR QDRNFVDKRQ WWGPGSYGG ILLVNCDRDD 120
 PSCDVGQDNCD QHVHCLQDLE DMSVMLVRQ GPAALFDDHK LVLHTSSYDA KRAQVFHICG 180
 PEDVCEAYRH VLQGDVKSYE VPRLHGDEER FFVEGLSFDPD AGFTGLISFH VTLLDDSNED 240
 FSASPIFTDT VVFRVAPWIM TPSTLPPELV YVCVRVNNTC FVDAVAEELAR KAGCKLTICP 300
 QAENRNRDWI QDEMEGLYVQ APHKTLPVVF DSPRNGELQD FPYKRILGPD FGYYTREPRD 360
 RSVSGLDSFG NLEVSPPVVA NGKVEPLGRI LIGGNLPGS SRRVTQVVRD FLHAQKVQPP 420
 VELFVWDLAV GHVDFLSFV PAPDGKDFRM LLASPGACFK LFQEKKQKGH GRALLFQGVV 480
 DBBQVKTISI NQVLSNLDLII NYNKFVQSCI DNWREVLKRE LGLAECIDIIP QPLFKTERK 540
 KATAFFPDLV NMLVLGKHLG IPKPKFGPIIN GCCCLEEKVR SLLEPLGLHC TFIDDFTPYH 600
 MLHGEVHCGT NVCRKPFSEFK WWNMV 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
 tctccctgtt tctgcattt cccctgtggg gggctgccc gggccaaaca ggcggagg 60
 cttccaccaggc cttggaaattt acaggcttgc gggttgcataa aataaaagat 120
 gccttagttaa aattttttttt cttttttttt aattttttttt tttttttttt 180

-continued

caaaaacttc acggggatc caaacaatga ttgactgttc acttgaattt caagtccaaac	240
ttagctttt gtattttat ttgactgttc tggcagccct gctttagggg agaaaacagg	300
gtgtctgtcc ctccccactg gtgaggacac ctgacagggg cccagacacc agagtggac	360
ctgcctgtatc gagagaagct tgggttggt ttatccctt tttcataaaa aagctgttt	420
gagatgtaa gagaatcatgt tgggttggaa gtgaagagggt tggtttgggg gaggtctgcc	480
ttactctgtatc ggaagatggg aagegggggg agtggccagc tggtaattt tgggttagt	540
ccctacgtatc ctgtgtcc tggatccca tggtaatcca ggtaatattt tctccatctgt	600
gggctgcagt gaggatcaga tggatccatg taaaaaaagg cttagccagg cctgcgttg	660
taagagtca gagggtgggg gctttttatg ttatggattt gttgttattt ttactcttgg	720
ggaggtgttgg ggcacaaatgt tggatccatg agccacatgg ctgggtttgg atcccaggcc	780
ggccacacatc ttggggcaag tttttttttt catctgtata atggggagaca	840
ataatagata gtgcctccct tcttaggtt tggagaggat tgcatgtttt gctggcacac	900
agcatgtatc cttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	960
ccaggaggccc ctccaaatctt cttttttttt tttttttttt tttttttttt tttttttttt	1020
cctgttgcctt tggatccatc tttttttttt tttttttttt tttttttttt tttttttttt	1080
ggatcaacatc tgggggggggg gttttttttt tttttttttt tttttttttt tttttttttt	1140
gtggggccatcc cacatgtttttt tttttttttt tttttttttt tttttttttt tttttttttt	1200
gttggatgtatc gggaaatgtttttt tttttttttt tttttttttt tttttttttt tttttttttt	1260
attaatattttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1320
tggaaacatc tgggtttttt tttttttttt tttttttttt tttttttttt tttttttttt	1380
cctcagggtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1440
tattgtttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1500
gcagcgccatcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1560
caccctactt gggatgtttttt tttttttttt tttttttttt tttttttttt tttttttttt	1620
agaactgtttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1680
cctgcactttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1740
cttgttcaatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1800
cttacatctt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1860
ttagctgtttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1920
ccatgtttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	1980
cttaccatcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2040
caggaaatctt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2100
gctgggttacatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2160
cgtgcacttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2220
ttagctgtttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2280
ggcacagggtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2340
ggggcaagatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2400
cgtggaaatcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2460
tctgtgtttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2520
gttcccgatgtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2580
gttccgtgtttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2640
cggctgtcaatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2700
ttagatgtttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2760
cccaaggatcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2820
ttacgttactt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2880
ggaggttccatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	2940
tggggggcaatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3000
ccatgtttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3060
tgtggatgtttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3120
ggccageccatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3180
ggcccttcctt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3240
ggtgtcttcc aataaaagacc tttttttttt tttttttttt tttttttttt tttttttttt	3300
gaaccgtttagt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3360
acagctttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3420
gctgggttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3480
ctgtgtttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3540
cattgtgtttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3600
gttcagatgtttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3660
accatcttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3720
gaacgatataatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3780
cttcagaaatcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3840
cttgaatctt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3900
tcaacacaaatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	3960
ccacatgtttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4020
acagaggatgtttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4080
ccttaaaatcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4140
cataaaaaggatcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4200
gctccctgtcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4260
atgcatgtttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4320
gcagaaaatcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4380
ggattgtttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4440
tgggttgcactt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4500
tcattgtttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4560
tctcagggttcc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4620
ggggcttcaagg tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4680
actgaattttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	4740

-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 WWGPGPSGYGG 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 DSRNGELOQD FPYKRLGPD FGYVTREPRD 360
RSVSGLDSFG NLEVSPPVVA NGKEVPLGRI LIQGNLPGSS GRRVTQVVRD FLHAQKVQPP 420
VELFVDWLAV GHVDEFSLFV PAPDGKGFRM LLASPGACFK LFQEKKQKCGH GRALLFQGVV 480
DBBQVKTISI NQVLSNKDLI NYNKFVQSCI DWNREVLIKRE LGLAECIDIID IPQLPKTERK 540
KATAFFPDLV NMLVLGKHLG IPKPKGPPIIN GCCCLEEKVR SLLEPLGLHC TFIDDDFTPYH 600
MLHGEVHCGT NVCRKPFPSFK 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 gggctgcccc gggccaaaca ggccagaggg 60
ctccaccagg cctggaaattt acaggctttt ggagtagggg tgccagataa aataaaagat 120
gccttagttt aattgttaattt caggtaataattt attaataat ttttttagtat aagtgtatcc 180
caaataacttc acggggatc caaaatcta ttgactgttc acttggaaattt caagtccaaac 240
tgagctttct gtatTTTTAT ttgtcaagtc ttggcagggct gctttaggggg agaaaacagg 300
gtgtctgtcc ctcccaactg gtgaggacac ctgacaggggg cccagacacc agagtggac 360
ctgcctgtat gagagaaatgt tgggttttgt ttatTTTCTT tttcataaaa aagctgtttt 420
gagatgttaag agaatcatgt ttgttggaaa gtgaagagggt ttgtttgggg gaggtctgccc 480
tttactctgt ggaagatggg aagcggggggg agtggccacg ttgtggaaattt tggggtagtc 540
ccctacgtat cctgtgcctc agtttccca ttgtttaaca ggagtaattt tctcatctgt 600
gggctgcagggt gaggatcaga tgaggtataa ttacaaaagg cttagggcagg cctgcgtgt 660
taagagtgtca gaggtttttt gctttttat ttatgttattt ttgttttata ttactcttt 720
ggaggtgttggt ggcaagatgt tggactgagg agccacatgg ctgggtttgg atcccaagcc 780
ggccacacag ctgggcaag ttacttaacc tcgggtttt catctgtata atgggagaca 840
ataatagata gtgcctccct tcttaggtta ttgagaggat tgcatgtatgg gctggcacac 900
agcatgcgtt cagaatgtt cagtcactgtt tttttttttt gggcatgtat gccgtgtttt 960
ccaggaggccc ctcccaactcg cttggccctc ttccatcatgtt ggtttttttt ggtgtgtat 1020
cctgtggccct tgagccccat ctgggtgtcc ttggacaatgtt ttgttccatgtt ctcccaagcg 1080
ggatcaacag tggggagatc agtgggtgttca tctggccgc accgttacc ttgtgtccaca 1140
gtgggcaccc cacatccaa ggtatgtat tctgggttttca aaggaggaaac ccagggttag 1200
gttgaatgtt gggaaatgtt tcttggccca agataacatc acagatattttt taaacctgtca 1260
attaattttt ggttgttttttccaatccgg ccaagtttcc cgggtgacac tgcagatgtt 1320
tggaaacagca tgggggttttttca gtcagacccatctt ttgcacatctt ggctctgtcca ttgtatgtca 1380
cctcagggtt gtcacaggcc ctccatgttccatgtt aatatgtccaa atggggatgtt 1440
tattgtctac tttagatgtt gtaatgttttgg agggatgttcc ttgtgttccagg tgcctggat 1500
gcagcggggca ctcaaccaga ttgtatttttcc ttcccaaggaa agaattcaag gatatttttcc 1560
ccccacttgg gggatgttccatgtt gatggatgtt ggggtgagaa ttgtttttttaaaatcttc 1620
agaactgttttgggggtttccatgtt ccacatgtt ccgtgttttca ctgtatgtt cgtggccaga 1680
cctgcacttgg ggatattttccatgtt ccacatgtt ccgtgttttca ttcccaaga gtcgtttttt 1740
ctgggttttttccatgtt ccgtgttttca cagaaatgtt ttgtgttctat gggacgttcc ggtgtggacat 1800
ctatcatctt cccaaatgtt ccgtgttttccatgtt agggggccgg gggatgttcc ggtgtggacat 1860
tgacgcgttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 1920
ccatgttccatgtt ccgtgttttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 1980
ctacacttccatgtt ccgtgttttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 2040
cagggacttttccatgtt ccgtgttttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 2100
gctgggttttccatgtt ccgtgttttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 2160
catgcacttccatgtt ccgtgttttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 2220
tgcagcccttccatgtt ccgtgttttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 2280
ggcacacaggccatgtt ccgtgttttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 2340
ggggccaaatgtt ccgtgttttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 2400
cgtggaaatgtt ccgtgttttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 2460
tctgtgttccatgtt ccgtgttttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 2520
gttccatgtt ccgtgttttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 2580
gtggccgttccatgtt ccgtgttttccatgtt ccgtgttttca ttgtgttccatgtt gatgttccatgtt gacaccaggc ggtgtggacat 2640

```

-continued

cggctgcaag	ctgaccatct	gcccacaggo	cgagaaccgc	aacgaccgct	ggatccagga	2700
ttagatggag	ctgggtacg	ttcaggcgcc	gcacaagacc	ctcccggtgg	tctttactc	2760
ccccaaagat	ggggaaactgc	aggatttccc	ttacaaaaga	atccctgggtc	cagattttgg	2820
ttacgtgact	cgggAACAC	ggcacaggbc	tgtgagtgcc	ctggactctt	ttggaaacct	2880
ggaggtcagc	cctccatgtgg	tggccaatgg	gaaagagtac	cccctgggg	ggatcctcat	2940
tggggggcac	ctgcctgggt	caagtggccg	cagggtcacc	cagggtggc	gggacttcc	3000
ccatggccag	aagggtgcgc	ccccctggga	gtctttgtg	gactgggtgg	ccgtggggca	3060
tgtggatgag	tttctgagct	ttgtccctgc	ccccatgggg	aagggttc	ggatgctctt	3120
ggccagecc	ggggcctgtc	tcaagcttt	ccaggaaag	cagaagtgtg	gccacgggag	3180
ggcccttc	ttcccaaggggg	ttgttgcata	taacaggtt	aagaccatct	ccatcaacca	3240
gtgtctcc	ataaaagacc	tatcaacta	caataagttt	gtgcagagct	gcatcgact	3300
gaaccgttag	gtgtgttagc	ggggagctggg	cctggcagag	tgtgacatca	ttgacatccc	3360
acagcttcc	aagaccgaga	ggaaaaaaago	aacggccttc	ttccctgact	tggtaacat	3420
gtctgggtc	ggggacggacc	tgggcatccc	caagccctt	ggggccatca	tcaatggctg	3480
ctgtgtcgt	gaggagaagg	tgccgttcc	gctggagccg	ctggccttcc	actgcacatt	3540
cattgtgac	ttcactccat	accacatgt	qcatggggag	tgcaactgt	qcaccaatgt	3600
gtcagaaag	cccttcctt	tcaagtgggt	gaacatgggt	cccttagagaca	gtctccacccc	3660
accatctgt	ccccctgggg	cgggcattgg	cccaagggtgt	ggagacagag	acaggccct	3720
gaacgcataag	caccaagaga	ccccaaagggt	ccagatggaa	cactgggggt	gaccgtcc	3780
ctcagaagcc	tttccctgg	aagtgtccat	gcctcacctg	caacccatgt	gttctcaga	3840
cttgaatctt	ctcgcccccc	caaaaagaag	gacctcattt	ctttagct	ctccctgtat	3900
tcaacacaac	ccatggagat	gtccccctt	cactctgaaa	tcatccattt	ggggacaat	3960
ccacatgggg	gtctgataaac	atccacgtat	ctcatcagcc	atcttgcctt	gtgcacatcta	4020
acagaggaa	gatccatgt	tctgttttt	tccaatttgt	tctcttgc	agaggaacaa	4080
ccctaaaacc	agaccactcc	acgcaggaca	ggcaggagag	attcttctca	aagccctccc	4140
cataaaaaagg	gagctgtgga	tccactttaga	tcagggcgga	accatcttc	acccggccaa	4200
gtctcgtcc	agatgtgtac	cctccatccc	cgtgagctgt	cacatagtag	gagttcttag	4260
atgcatgtgg	aagcaatgt	agttgtccct	tagccttata	aactccctat	gatctgacat	4320
gcagaaaatcc	acccttgc	agaatctcc	tggaaatttt	tggagacgaa	agtatctgg	4380
ggattgttgg	gtacttaggg	gactgggtac	aagggtaaa	agtagttccc	ataatacaca	4440
tgtttgacta	tgtgtatcca	cctgtgtat	gttaatattt	gtgtgttgg	gaagggttgc	4500
tcattggccc	tgggacttct	ctctcgagg	ggagagaa	ctgccttcc	tctggatgg	4560
tctcaggcgc	tctgttggcc	tttggtcag	gtttccat	cctgcttgc	tgcaagagag	4620
ggggctaagg	ggctggatcc	accaaggcg	ctcacagcg	gaaaactctg	ggaatgaacc	4680
actgaattca	ggggatgggg	gtgggggggg	ggttctcgag	gtgtgtc	ca gctacacgt	4740
tgttctgtat	gggtccagct	cgcttccat	cactcgtaa	taatcaaca	gaaacacaaa	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
 MFEEVYGTGV DIYISPNNMER GRERADTRRW RFDATLEIIV VMNSPSNDLN DSHVQISYHS 60
 SHPLPLPLAYA VLYLTCVDIS LDCCDLNCEGR QDRNFVDKRQ WVVWGPSCYGG ILLVNCDRDD 120
 PSCDVQDNCD QHMHCLQDLE DMSVMVLRQC GPAALFDDHK LVLHTSSYDA KRAQVFHICG 180
 PEDVCEAYRH VLGQDKVSYE VPRLHGDEER FFVEGLSFPD AGFTGLISFH VTLLDDSNED 240
 FSASPIFTDT VVFRVAPWIM TPSTLPPLEV YVCRVRNNTC FVDVAELAR KAGCKLTICP 300
 QAENRNRDRWI QDEMELGYVQ APHKTLVPVF DSPRNQELQD FPYKRLGPD FGYVTREPRD 360
 RSVSGLDSFWI NLEVSPPVVA NGKEYPLGRJ LIIGNLPGS SRRVTQVRD FLHAQKVQPP 420
 VELFVWDLAV GHVDEFPLSFV PAPDGKGFRM LLASPGACFK LFQEKKCQGH GRALLFQGVV 480
 DEQVKTISI NQVLSNKDLI NYNKFVQSCI DWNREVLKRE LGLAECDIID IPQLFKTERK 540
 KATAFFPDLV NMVLVLGKHLG IPKPFGPIIN GCCCLEEKVR SLLEPLGLHC TFIDDFTPYH 600
 MLHGEVHCGT NVCRKPFSFK WWNMVP 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
 tctccctgtt tctgccattt cccctgtggg gggctgccc gggccaaaca gcccagaggg 60
 ctccaccaggc cctggaaattt acaggctttg ggagtaggg tgccagataa aataaaagat 120
 gccttagttaa aattgtatcc aggttataa attaataattt tttttagtat aagtgtatcc 180
 caaaaacttc acgggagttc caaacaatgtt tgactgtttc acttgaattt caagtccac 240
 tgagctttct gtatttttat ttgtctttttc tggttttttcc gttttttttt gttttttttt 300
 gtgtctgtcc ctcccaactg gtgaggacac ctgacagggg cccagacacc agagtggac 360
 ctgcctgtat gagatgttgc tgggtttttt ttatccataaa aagctgtttt 420
 gagatgttgc agaatcatgtt tggtttttttt gttttttttt ggggtttttt 480
 ttactctgtat ggaatgttgc aaggccgggg agtggccagc tggtttttttt tgggtttttt 540

-continued

ccctacgtat	cctgtgcctc	agtttccccca	tgtgtaaaaca	ggagtaattta	tctcatcctg	600
gggctgcagt	gaggatcaga	tgagctaatg	tacaaaaaagg	cttagggcagg	cctgcgcgtga	660
taagagtgc	gagggtgggg	getttttatg	ttatggattt	gttggatata	ttactccctgg	720
ggaggtggt	ggcaagagt	tggactgagg	agccacatgg	ctgggtttgg	atcccagccc	780
ggccacagac	cttggcaag	ttacttaacc	tcggcttct	catctgtata	atgggagaca	840
ataatagata	gtgcctccct	tcttagggta	tttagagggat	tgcatgtatgg	gtggcacac	900
agcgtgacgt	cagaatgt	cagtcactgt	tatccccatg	gggcgtatgc	gcccgtgcct	960
ccaggaggcc	ctccaactcg	cagtgccctc	tccttacatg	gttcccaact	gagctgtgtat	1020
cctgtggct	tgagccccat	ctggctgccc	tggacaatgt	ttgtctctag	ctccccagcg	1080
gtatcaacag	tggggagagc	agtgggttgc	tctggccccc	aggcgttacc	tgagtgcaca	1140
gtgggcaccc	cacagtc	ggatgtatgc	tctggtttc	aaggaggaaac	cccaggctag	1200
gttgaatgag	ggaaatcaat	tctcaggccc	agataacatc	acagataactt	taaacctgc	1260
attaattct	ggttgttga	ccaaatccgg	ccaagtttcc	cgggtgacac	tgcagagctg	1320
tggAACAGCA	ttgggttggg	gtcagaccc	ttgcacatct	ggctctgc	tgtgtatgca	1380
cctcaggctt	gtcacgggc	cttcgttgc	tcagttgtc	aatatgc	atggggatga	1440
tattgtctac	tttacagctt	gtaaagttg	aggaggtc	gtgcgcagg	tgccctggcat	1500
gcagcgggca	ctcaaccaga	ttgaattccc	ttcccaaggga	agaattcaag	gatatctcc	1560
cacccactaa	ggagacatgt	gtatggatc	gggggtgagaa	ttgtctttt	taaagctctc	1620
agaactgtaa	gggggtccac	ccaccaatgg	ctgggttgc	ctgagcttt	cgtgcgcagg	1680
cctgcacttg	ggattttatcg	agcacatgg	ttggcttgc	tttccgaaga	gtctgcatt	1740
cttgtcgtat	cttgaggggc	cagaatgtt	tgaggttat	gggacgcctg	gcgtggacat	1800
cttacatctt	cccaacatgg	agagggggcc	ggagcgtgc	gacacaggc	ggtggcgctt	1860
tgacgcgact	ttggatgtca	tcgtgttcat	gaactcccc	agcaatggc	tcaacgacag	1920
ccatgttccat	atcccaatgg	tgacgcctcg	ccccctgc	atgcgggtct	1980	
ctacctcacc	tgtgttgc	tctctctgg	ttgcgacctg	aactgttgg	gaaggcagg	2040
cagaactttt	gtagacaagc	ggcagtggg	ctggggggcc	agtgggtatg	gcccgcattt	2100
gtctggtaac	tgtgtgcgtg	atgtccgg	ctgtgtatgc	caggacaaat	gtgaccagca	2160
cgtgcactgc	ctgcaagacc	tggaaagat	gtctgtatgc	gtcctgcgg	cgcaggggccc	2220
tgcagccctc	tttgcgtacc	acaaaatgt	cctccatacc	tccagctatg	atgcggaaacg	2280
ggcacaggtc	ttccacatct	gcccgtctg	ggatgtgtgt	gaggctata	ggcatgtgt	2340
ggggcaagat	aagggttctt	atggatgttgc	ccgcgttgc	ggggatgtgg	aggcgttctt	2400
ctgtggaaaggc	ctgtcttcc	ctgtgcgg	cttcacaggaa	cttacatctt	tccatgtcac	2460
tctgctggc	gactccaa	aggatttctc	ggcatccct	attttactg	acactgtgt	2520
gttcccgagt	gtaccctgg	tcatgacgc	cagcaactcg	ccacccct	agggttatgt	2580
gtgccgtgt	aggaaacaca	cgtgttttg	ggatgcgg	gcagagctgg	ccagggaggc	2640
cggctgc	cgttccatct	gcccacatgg	ccccctggg	ggatctctat	2700	
tgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	2760
ccatgtccac	ggggaaatgt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	2820
tttacgttgc	gggggttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	2880
gggggttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	2940
tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3000
ccatgtccac	gggggttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3060
tgtggatgt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3120
ggccaggcc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3180
ggccctctg	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3240
gggtgtctcc	aaataagacc	tcatcaacta	caataagttt	gtgcagatgt	gtatcgact	3300
gaaccgtgt	gtgtgtgt	ggggactgg	tttgcgttgc	tttgcgttgc	tttgcgttgc	3360
acatcttc	aaagccgtgt	ggggactgg	tttgcgttgc	tttgcgttgc	tttgcgttgc	3420
gtctgggtgt	ggggactgg	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3480
ctgtgtgt	ggggactgg	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3540
catgtgtgt	ggggactgg	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3600
gtctgggtgt	ggggactgg	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3660
accatctgt	ccccctgggg	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3720
gaacgataag	caccaagaga	cccccaaggct	ccagatggaa	cactgggggt	gaccgtccct	3780
ctcagaagcc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3840
tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3900
tcaacacaac	ccatgtgtgt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	3960
ccacattggg	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4020
acagaggaa	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4080
ccctaaacc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4140
cataaaaaagg	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4200
gtctgggtcc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4260
atgcgtatgtt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4320
ggagaaatcc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4380
ggatttgg	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4440
tgtgtgtact	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4500
tcattggccc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4560
tctcagggtc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4620
ggggctaaagg	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4680
actgaattca	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4740
tgttctgtat	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4800

SEQ ID NO: 30 moltype = AA length = 6260
 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 NMLVLGKHLG 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 ggcacagagg 60
ctccaccagc cctggaaatt acaggcttg ggagtagggg tgccagataa aataaaagat 120
gcctagttttaa aattgaattt caggttaata attataattt ttttttagtat aagtgtatcc 180
caaataacttc acgggaaatc caaacaatga ttgactgttc acttgaattt caagtccaa 240
tgagctttctt gtatgtttt ttgctaaatc tgccagggcctt gctttagggg agaaaacagg 300
gtgtctgtcc ctccccaaatc tgaggacac ctgacaggggg cccagacacc agagtgggac 360
ctgcctgtatc gagagaagct tgggttggg ttatccctt ttcataaaa aagctgtttt 420
gagatgttaag agaatcatgt tgggttggaa gtgaagagggt tgggtttggg gaggtctgcc 480
ttactctgtatc ggaagatggg aagecgggggg agtggccacgc tggaaattt tgggttagtc 540
ccctacatgtt cctgtgcctc agttttccca tggtaaaaca gggataatta tctcatcttg 600
gggctgcagttt gaggatccatc tgagctaaatg taaaaaaagg cttagcagg cctggctgt 660
taagagtgcg aagggtgggg gctttttatg ttatgttattt ttactcttg 720
ggaggtgtgtt ggcaaaatgtt tggttgggg aggccacatgg ctgggtttgg atcccaagccc 780
ggccacacagc ctggggcaatc ttacttacc tcggcttttc catctgtata atgggagaca 840
ataatagataatgtgccttctc tcttaggtta tggagggat tgcattatgg gctggcacac 900
agcatgagctt cagtaatgtt cagtcactgt tatcccccaatgg gggcatgtat gccgtgcct 960
ccaggaggccc ctccaaactcg cagtgcccttc tccttacatc ggttccactt gagctgtgtat 1020
cctgtgttgtt tgagcccccatt ctgggttgcctt tggttgcataatgtt tggttgcctcag ctccccacgc 1080
ggatcaacacg tggggggagac agtgggttgg tggcccccgc accggatacc tgatgtcaca 1140
gtggggccccc cacatgttttttggatccatc tgggttttc aaggaggaaac cccaggcttag 1200
gttgaatgtatc gggaaatgtatc tctcgaggcc agataacatc acagataatc taaaactgtca 1260
attaatttttctt ggttggatc ccaatgttttgc ccaatgttttgc cgggttgcacatc tgcaatgtt 1320
tggAACACGCA tgggtttggatc gtcacatctt tggcacatctt ggctctgcctt tggatgttgc 1380
cctcaggctt gtcacggggc ctcttgcgtt ctcagttgttcaatgttgccttgcctt atggggatgt 1440
tattgtctacatc ttatcgatgtt gtaaagatgttgg agggaggttgc tgccggccagg tgccctggcat 1500
gcagcggggcaatc tggatccatc ttggatccatc tgggttgggggaaatgttgccttgcctt gatatcttc 1560
caccctacttgc gggatccatc gatggatctt qgggtqaaatgg tggcttttttgc taaatgttgc 1620
agaactgttacatc ggggttccatc ccaccaatgg ctgggttgcctt ctgagcttttgc tggccacatc 1680
cttcggacttgc ggatgttgcctt gtcacatctt tggcacatctt ggctctgcctt tggatgttgc 1740
cttgggttgc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 1800
cttacatcttgc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 1860
ttggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 1920
ccatgttgcacttgc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 1980
cttacatcttgc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2040
cggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2100
gttgggttgcacttgc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2160
cgttggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2220
tggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2280
ggcggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2340
ggggccatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2400
cgttggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2460
tggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2520
gttggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2580
gttggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2640
cggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2700
ttggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2760
cccaaggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2820
ttacatcttgc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2880
ggggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 2940
tggggatccatc tggatccatc tggatccatc tggatccatc tgggttgcctt gtcacatctt gatatcttc 3000

```

-continued

ccatggcccag	aagggtcagc	ccccctgttgc	gctttttgttgc	gactgggttgg	ccgtggggcca	3060
tgtggatgag	tttcgttgc	tttgccttgc	ccccatgggg	aagggttcc	ggatgttcc	3120
ggccagccct	ggggcttgc	tcaagtttgc	ccaggaaaag	cagaagtgttgc	gccacgggg	3180
ggcccttcttgc	ttccagggggg	tttgttgc	tgagcagggttgc	aagcacatct	ccatcaacca	3240
ggtgcttcc	aataaaagacc	tcatcaacta	caataagtttgc	gtgcagaggttgc	gtatcgacttgc	3300
gaaccgttgc	gtgtgttgc	ggggacttgc	cttggcagaggttgc	tgtgacatca	tttgacatccc	3360
acagcttcc	aagaccgaga	ggaaaaaaaggg	aacgacatcc	tttccctgttgc	tttgtgaacat	3420
gtctgggttgc	ggggaaaccc	tggcattcc	caagcccttgc	gggccccatca	tcaatgttgc	3480
ctgtgttgc	gaggagaagg	tgccgttgc	gttggagcc	ctgggttgc	actgcaccc	3540
cattgtatgc	tttccatccat	accacatgttgc	gtatgggggg	gtgcactgttgc	gcaccaatgt	3600
gtgcagaaag	cccttcttgc	tcaatgttgc	gaaatgttgc	ccctgttgc	gttccccc	3660
accatctgt	ccccctgggg	cggcatttgc	ccccatgttgc	ggagacagaggttgc	acaggccc	3720
gaacgataag	caccaagaga	ccccaaagggttgc	ccagatgttgc	cactgttgc	gaccgttcc	3780
ctcagaagcc	tttcccttgc	aagtgttgc	gttccatcc	caacccatgttgc	ggtttctcaga	3840
cttgaatcttgc	ctcgcccccc	caaaaagaaag	gacccatgttgc	tttagatgttgc	ctctgttgc	3900
tcaacaacac	ccatgttgc	gttcccttgc	cactcttgc	tcatccatttgc	ggggacaat	3960
ccacattgggg	gtcttagaaac	atccacgttgc	cttcatgttgc	atcttgc	gttgcatttgc	4020
acagaggaaag	gatcttgc	tctgttgc	tccaaatgttgc	tccttgc	agaggaaaca	4080
ccctaaacc	agaccatcc	acggcaggat	ggcaggaggttgc	atttccat	aaggcttccc	4140
cataaaaaagg	gagctgttgc	tccacttgc	tcaggcgggaa	accatcc	acccggccaa	4200
gctccttgc	agatgttgc	ccttccatcc	cggttgc	cacatgttgc	gagtttgc	4260
atgcatgttgc	aagcaatgttgc	agttgttgc	tagcatttgc	aactccat	gatctgttgc	4320
gcagaaatcc	ggcccttgc	agaatcttgc	tggaaatttgc	tggagacagaa	agtatgttgc	4380
ggattgttgc	gttactgttgc	gactgttgc	aaagggttgc	agtagtccc	ataatcacaca	4440
tggttgacta	tggtgtatcc	ccttgc	tgttgc	gttaatatta	gggtgttgc	4500
tcatgttgc	ttggacttgc	ctctgttgc	ggagagaacgttgc	ctgccttcc	tctggatgttgc	4560
tctcagggttgc	tctgttgc	tttgggttgc	tttgggttgc	tttgggttgc	tttgggttgc	4620
ggggctaaagg	gggttgc	accacccatcc	ccacccatcc	ggaaaactctg	ggaatgaacc	4680
actgaatcc	ggggatgggg	gttgggggg	gttgggggg	gttgggggg	gttgggggg	4740
tgttctgttgc	gggttgc	tgcttccat	cactcg	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
 MFEEVYGTGPGV DIYISPNNMER GRERADTRRW RFDATLLEIIV VMNSPSNDLN DSHVQISYHS 60
 SHEPLPLAYA VLYLTCVDIS LDCDLNCEGR QDRNFVDKQRQ 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 APHKTLVVFV DSPrNGELQD FPYKRILGPD FGVTREPRD 360
 RSVSGLDSFG NLEVSKPVVA NGKVKLPGR LIIGGNLPGSS GRRVTQVVRD FLHQAKVQPP 420
 VELFVWDWLAV GHVDEFLSFV PADDGKGFPV LIASPGACEK LPQEKKQKCGH GRALLFQGVV 480
 DDEQVKTISI NQVLSNLDL 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
 ttccttgcattt cccctgttgc gggctgttgc gggccaaaca gggcagagg 60
 ctccaccaggc cttggaaattt acaggcttgc ggagtaggttgc tggcagatata aataaaagat 120
 gcccgttgc aattgttgc cttggaaata attaataat tttttagtat aagtgtatcc 180
 caaataacttc acggggatcc cttggaaatcc ttgtactgttgc acttggaaattt caagtccaa 240
 tggttgc tttttttat ttgttgc tttggcatttgc gtttaggggg agaaaacagg 300
 gtgtgttgc ctcccaacttgc ttgttgc tttggcatttgc gtttaggggg agaaaacagg 360
 ctgcctgtatc gagatgttgc tttggcatttgc ttatccataaa aagctgttgc 420
 gagatgttgc agaatgttgc tttggcatttgc gtttaggggg agaaaacagg 480
 ttactctgttgc gggatgttgc aatgggggggg agtggccatcc tttggaaattt tggggtagtc 540
 ccctacgtat cttgttgc tttggcatttgc gtttaggggg agaaaacagg 600
 gggctgttgc gggatgttgc tttggcatttgc gtttaggggg agaaaacagg 660
 taagatgttgc gggatgttgc tttggcatttgc gtttaggggg agaaaacagg 720
 ggaggttgc gggatgttgc tttggcatttgc gtttaggggg agaaaacagg 780
 gcccacacac cttggcatttgc tttggcatttgc gtttaggggg agaaaacagg 840
 ataataatgttgc tttggcatttgc gtttaggggg agaaaacagg 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	aggcagtaacc	tgatgcaca	1140
gtgggcaccc	cacagtcaaa	ggatgcata	tctggttcc	aaggaggaac	cccaggctag	1200
gttggaaatgg	gaaaatgcac	tctcaggccc	agataaacat	acagatactt	taaaactgca	1260
attaattttct	gggtgttga	ccaaatccgg	ccaagtttc	cggtgtcac	tgcagagctg	1320
tggAACAGCA	tgggttggc	gtcagaccc	tgcacatct	ggctctgc	tgtatgcac	1380
cctcaggctt	gtcacgagc	ctcettgago	ctcagttgt	aatatgcac	atggggatga	1440
tattgtcactc	tttacatgtt	gttaaagatgg	agggagctgc	gtgcgcagg	tgcctggcat	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	gtctgcctt	1740
ctggctcagt	ctggggggc	cagaatgtt	tgaggcttat	gggacgectg	gcgtggacat	1800
ctacatctt	cccaacatgg	agagggggc	ggacgtgtca	gacacaggc	gggtggcgtt	1860
tgacgcgact	ttggagatca	tcgtggcat	gaactcccc	agcaatgacc	tcaacgacag	1920
ccatgttcag	atttcttacc	actccagcc	tgacgcctc	ccccctggct	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	gtcctgcgga	cgcaggggcc	2220
tgcagccctc	tttgatgacc	acaaatgtt	cctccatacc	tcacatgtat	atgcacaaacg	2280
ggcacaggtc	ttccacatct	gcccgttgc	ggatgtgt	gaggctata	ggcatgtgt	2340
ggccaagat	aagggttctt	atgaggtaco	ccgcttgc	ggggatgagg	agcgttttctt	2400
cgtggaaaggc	ctgtcttcc	ctgatgcgg	cttcacacgg	ctcatctct	tccatgtcac	2460
tctgtcgac	gacttccaaac	aggatttcc	ggcatccctt	atcttcaact	acactgtgg	2520
gttcccgagt	gcacccctgg	tcatgaccc	cagcactctg	ccacccctag	agggttatgt	2580
gtgccgtgt	aggaacaaca	cgtgtttgt	ggatgcgg	gcagagctgg	ccaggaaaggc	2640
cggctgcaag	ctgacatcc	gccccacaggg	cgagaaccgc	aacgaccgc	ggatccagg	2700
tgagatggag	ctggggatcg	ttcaggcgc	gcaaaagacc	ctcccggtt	tctttgactc	2760
cacaaggaaat	ggggaaactgc	aggatttccc	ttacaaaaga	atcttgggtc	cagattttgg	2820
ttacgtgact	cgggaaaccac	gcccacgg	tgtgagttgc	ctggactctt	ttggaaacct	2880
ggaggtcagc	cctccagtt	tggccaaatgg	gaaagagtac	ccccctggga	ggatccctat	2940
tgggggcaac	ctggcttgg	caagggttgc	cagggttacc	cagggttgc	gggacttcc	3000
cctatggcc	aagggtcagc	ccccctgttgc	gcttttgc	gactgttgg	ccgtggggca	3060
tgttgatgt	tttctgagct	tttgccttgc	ccccatgtgg	aaagggttcc	ggatgtctt	3120
ggccagccct	ggggccctgt	tcaagcttt	ccaggaaag	cagaagtgt	gccacggag	3180
ggccctctt	ttccatgggg	tttgtatgc	tgagcagg	aagaccatct	ccatcaacca	3240
gggtgtctcc	tttcaatgg	tcatcaacta	caataatgtt	gtgcagatgt	gatcgact	3300
gaaccgttag	gtgtgtaa	ggggacttgg	cttgcacag	tgtgacatca	ttgacatccc	3360
acagcttcc	aagaccgaga	ggaaaaaa	aacggcc	ttccctgact	tgtgaaacat	3420
gtctgggtcg	ggggatccac	tgggcatt	caaggccctt	gggaccatca	tcaatggct	3480
ctgtgtctgt	gaggagaagg	tcgggttcc	gctggagcc	ctggcttcc	actgcaccc	3540
cattgtatgc	tttacttccat	accacatgt	gcatgggg	gtgcactgt	gcaccaatgt	3600
gtcagaaag	cccttcttcc	tcaagttgt	gaacatgtt	cccttgc	gttccacccc	3660
accatctgt	ccccctgggg	cgggcatgtt	ccccatgtt	ggagagacag	acaggccct	3720
gaacgataag	caccaagaga	ccccaaagg	ccagatgg	cactgttgc	gaccgttcc	3780
ctcagaagcc	tttcccttgg	aagtgttccat	gccttcaccc	caacccatgt	ggttctcaga	3840
tttgaatctt	ctcgcccc	aaaaaagaag	gaccttcattt	ttttagatgt	ctcctgtat	3900
tcaacacaac	ccatgttgc	gtccccctt	cactcttgc	aatccatgg	ggggacaaat	3960
ccacatggg	gtcttagaa	atccacgtat	tcatcagcc	atcttgc	gtgcaccc	4020
acagaggaa	gtatccatgt	tctgttttgc	tccaattgt	tctcttgc	agaggaacaa	4080
ccctaaaacc	agaccactcc	acgcaggaca	ggcaggag	atttccctt	aagccccc	4140
cataaaaaaa	gaggtgttgc	tccacttgc	tcaggcgg	accatcttc	acccggccaa	4200
gctcttgc	agatgttgc	cctcacc	cgtgtgt	cacatgttgc	gagttctat	4260
atgcattgtt	aagcaatgt	agttgttcc	tggcttata	aactcccat	gatgtgcac	4320
gcagaaatcc	agccttgc	agaatcttcc	tggaaattt	tggagacaa	agtatctgg	4380
ggattgttgg	gtactaggaa	gactgttgc	aagggttgc	aatatgtcc	ataatacaca	4440
ttgttgc	ttgtgtatcc	cctgtgtat	ttatatttt	gggtgttgc	gaaggttgc	4500
tcattggccc	ttggacttct	ctctgc	ggagagaac	ctgccttcc	tctggatgg	4560
tctcaggctc	tctgttgc	tttggatc	tttccat	cctgtctgc	tgccaggag	4620
ggggctaa	ggctgttgc	accaggc	ctcacagcg	gaaaactctg	ggaatgaacc	4680
actgaattca	ggggatgggg	gttgggggg	ggttctc	gtgtgtc	gctacacgt	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

-continued

acagcttc aagaccgaga ggaaaaaaago aacggcccttc ttccctgaact tggtaacat	3420
gtcggtgtc gggaaagcaccc tggcatccc caagccctt gggccatcca tcaatggctg	3480
ctgtgcctg gaggagaagg tgcatgcctt gctggagccg ctggccctt actgcaccc	3540
cattgtac ttcactccat accacatgt gcatggggag gtgcactgtg gcaccaatgt	3600
gtcagaaag cccttcctt tcaagtggtg gaacatggtg cctgagaca gctccaccc	3660
accatccgt cccccctggg cgggcatgg cccagggtgg ggagacagag acaggccc	3720
gaacgataag cacaagaga cccaaaggct ccatggggact cactggggat gaccgtccct	3780
ctcagaagcc ttccctggg aagtgtccat gcctcaccc acacccatgt gtttctcaga	3840
cttgaatctt ctggccccca caaaaagaag gacctcattt cttatagct ctcctgtat	3900
tcaacacaac ccatggatg gttcccttc cactctgaaa tcatccattt gggacaaat	3960
ccacatggg gtcttagaaac atccacgtat ctatcacggc atcttgtct gtgcacatcta	4020
acagaggaaag gatccatgtat tctgtttgg tccaattgtc tctcttcgc agaggaaaca	4080
ccctaaaacc agaccactcc acgcaggaca ggcaggagat attttccca aagcctccc	4140
cataaaaaagg gagctgtgg tccacttaga tcaggcggaa accatcttc acceggccaa	4200
gctccctggcc agatgttgc cctcacccag cgtgagctgt cacatagtag gagttctag	4260
atgcatgtgg aagcaatgat agttgtccct tagccttata aactccccc gatgtgcacat	4320
gcagaaaatcc agccttgc tggatccat tggatatttc tggagacgaa agtatctggg	4380
ggattgttgg gtactaggaa gactgggtt aagggtgaaa agtattccc ataatacaca	4440
ttggttacta ttgtgtatcca cttttgtatgtt gttatattaat ggtgtctggaa gaaggttgct	4500
tcattggccccc tggggacttc ctctgcaggaa ggagagaaacg ctgccttc tctggattgg	4560
tctcagggtc tctgttggcc tttgtgtcago gtttccacat cctgtctgc tgaggagag	4620
ggggctaaagg ggctggatcc accaaggcag ctcacagcgg gaaaactctg ggaatgaacc	4680
actgaattca gggggatgggg gtgggggggg ggttctcgag gtgtgtgcac gtcacacgt	4740
tgttctgtat gggctccat cactcgtaaa taaaatcaaca gaaacacaaa	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
 MFEVYGTGPGV DIYISPNNMER GRERADTRRW RFDATLEIIV VMNSPSNDLN DSHVQISYHS 60
 SHPLPLPLAYA VLYLTCVDIS LDCLDNCEGR QDRNFVDKRQ WWGPGSYGG ILLVNCDRDD 120
 PSCDVQDNCD QHVHCLQDLE DMSVMVLRTQ GPAALFDDHK LVLHTSSYDA KRAQVFHICG 180
 PEDVCEAYRH VLGQDKVSYE VPRLHGDEER FFVEGLSPPD AGFTGLISFH VTLLDDSNED 240
 FSASPIPTDT VFVRVAPWIM TPSTLPPLEV YVCRVRNNTC FVDAVAELAR KAGCKLTICP 300
 QAENRNDRWI QDEMELGYVQ APHKTLPPVF DSPRNGELOQD FPYKRILGPD FGYYTREPRD 360
 RSVSGLDSFG NLEVSPPVVA NGKEYPLGRI LIGGNLPGSS GRRVTQVVRD FLHAQKVQPP 420
 VELFVWDLAV GHVDEFLSFV PAPDGKGFRM LLASPGACFKL LFQEKKQKCGH GRALLFQGVV 480
 DEQVKTISI NQVLSNKDLI NYNKFVQSCI DWNREVLKRE 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
 agaaatggat ggtatgtact gtgtgtctaataatgttttataacaaac aggcagtagg 60
 ccagatgttgc cccacagttc ataatgtgttgc gatcctgacc taggcggagaa gagaacacaa 120
 atatgaaact gtttggaaac ttgggactga attatgttgg aacttgggtc cctggggat 180
 aagaggagaa ggttgcggcg cagtgggtt gggggcccaag tgggtatggc ggcacatcttc 240
 ttggtaactg tgaccgtat gatcggatgt tgatgttccca gggacatgtt gaccagcact 300
 tgcactgttgc gaaaggatgttgc gaaacatgttgc tctggggacgg caggggccct 360
 cagcccttc ttgttgcaccac aaacttgc tccataccctt cagctatgtat gccaacacggg 420
 cacaggttttccatctgc ggttctgttgc ggttctgttgc ggcctatagg catgtgttgc 480
 gccaagatgttgc ggttctgttgc ggttctgttgc ggttctgttgc ggttctgttgc 540
 tggaaaggatgttgc gtttctgttgc gatcggatgttgc tccatcttc catgttgcact 600
 tgcgttgc gtttctgttgc gatcggatgttgc tccatcttc catgttgcact 660
 tccgttgc accctggatc atgaacccca gcaactgttgc acccccttagag gtgtatgttgc 720
 gcccgttgc gaaacacacg ttttttttttgc atgggggttgc agagctggcc aggaaggccg 780
 gtcgttgc gtttctgttgc gatcggatgttgc tccatcttc catgttgcact 840
 agatggatgttgc gggctacgtt gatcggatgttgc tccatcttc catgttgcact 900
 caaggaaatgttgc gggactgttgc gatcggatgttgc tccatcttc catgttgcact 960
 acgttgcacttgc gggactgttgc gatcggatgttgc tccatcttc catgttgcact 1020
 aggttgc gggactgttgc gatcggatgttgc tccatcttc catgttgcact 1080
 gggccaaacttgc gggctggatgttgc atgggggttgc agagctggcc aggaaggccg 1140
 atggccatgttgc gggactgttgc gatcggatgttgc tccatcttc catgttgcact 1200
 tggatgttgc ttttttttttgc atgggggttgc agagctggcc aggaaggccg 1260

-continued

ccagcccctgg	ggcctgttcc	aagctttcc	aggaaaagca	gaagtgtggc	cacggggagg	1320
ccctcccttgtt	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
tttatgtactt	cactccatac	cacatgtgc	atggggagg	gactgtggc	accaatgtgt	1740
gcagaaaacc	cttctttc	aagtgggtga	acatgggtgc	ctgagacagc	tcccaaccac	1800
catccctgtcc	ccctggggcc	ggcattggcc	cagggtgg	agacagacag	aggcccctga	1860
acgataagca	ccaagagacc	ccaaggctt	agatggaa	ctgagggtga	ccgtccctt	1920
cagaagcc	ttccctggaa	gtgtccatgc	ctcacctgc	acccatgtgg	tttcagact	1980
tgaatcttct	cgccccccca	aaaagaagga	cctcatttct	tatagectt	cctgtgttcc	2040
aacacaaccc	atggagatgt	ccccctctca	cctcgttcc	atccatgg	ggacaatacc	2100
acattgggt	ctagagat	ccacgtatct	catcagccat	cttgcgttgc	gcatectaacc	2160
agaggaaagg	tccatgttcc	tgctttggc	caattgttcc	ctctgtcg	aggaacaacc	2220
ctaaaaccag	accactccac	gcaggacagg	caggagat	tcttcttaaa	gcctccccca	2280
taaaaaaggaa	gtgtgtgt	cacttagat	agggccggac	catcttccac	ccggcccaagc	2340
tcctgcccag	atgttgcacc	tcaccccg	tgagctgtca	catatgttgg	gcttcttagat	2400
gcatgtggaa	goaatggag	ttgtccctt	gccttataaa	ctcccatgg	tctgacatgc	2460
agaaatccag	cttgcgttcc	aatcttctg	gaatttctt	gagacgaaag	tatctgggg	2520
attgttgggt	actagggaa	ctgggtacaa	gggtgaaaag	tagtcccat	aatacacatg	2580
gttgactatg	gtgtatcc	ttgtgtat	taatattagg	tgtctgg	agggtgttcc	2640
atggggccct	ggacttctt	tcggcgg	agagaacgc	gcctcttc	tggattgg	2700
tcaggctctc	tggtggcc	tggcagegt	ttccacatcc	tgtctgt	caggagagg	2760
ggctaaagg	ctggatccac	caaggcag	cacagccgg	aaactctgg	aatgaaccac	2820
tgaatttgg	ggatgggggt	ggggggggcg	tttctcg	gtgtccagc	tacacgtgt	2880
tttctgtatgg	gtccagctc	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 RRVTVQVRDF 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 ggatgtactt gtgtgctaat aaaactttat ttatacAACAGCAGCAGG 60
 ccagatggg cccacatgttcc ataatgtgt gatccatgttcc taggcggagaa gagaacccaa 120
 atatggaaact gttgaagac ttggactt attatgttgg aacttgggtc cctggggagt 180
 aagaggagaa ggtcgagccg cagtgggtt gggggcccaag tgggtatggc ggcacatcttc 240
 tggtaactg tgaccgtat gatecgagat gtatgttca ggacaatgtt gaccagcagc 300
 tgcactgcct gcaagacatgt gaagacatgt ctgtcatgtt ctcggccggc cagggccctg 360
 cagcccttcc tttatgttcc tccatacttc cagctatgtt gccaacacggg 420
 cacaggctt ccacatcttc ggtcttgagg atgtgtgtt ggcctatagg catgtgttcc 480
 gccaagatataa ggtgtcttcat ggttaccc cgttgcgtt ggtatgggg cgcttcttc 540
 tggaggccct gtccctccct gatgcgggt tcacaggact catctcttc catgtactc 600
 tgctggacga ctccaaacggag gatttctcgat catccctat ctttacttgc actgtgttgc 660
 tccggatggc accctggatc atgacgccc gactctggcc acccccttggat gtgtatgtt 720
 gccgtgttgg gaaacacacg ttgttttgg atgcgggttcc agagctggcc aggaaggccg 780
 gctgcgttcc gaccatcttc ccacaggccg agaaccgc gacccgttcc atccaggatg 840
 agatggagct gggctacgtt caggggccgc acaagaccctt cccgggttcc tttacttcc 900
 caaggaaatgg ggaactgcag gatttccctt acaaaagaat cttgggttca gattttgggtt 960
 acgtgactcg ggaaccacacg gacaggttcc tgatgtggccctt gggaaaccttgg 1020
 aggtcagccc tccatgttgg gccaatggga aagagtaccc cctggggagg atccttattt 1080
 gggggcaacccctt gctgggttca agtggccgc ggttccacca ggttgggtcc gacttcc 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	gcccatacatc	aatggctgt	1620
gctgcctgg	ggagaagggt	cggccctgtc	tggagccgt	gggcctccac	tgcacctca	1680
ttgtatgactt	cactccat	cacatgtgc	atggggagg	gcaactgtgc	accaatgtgt	1740
gcagaaggcc	cttcttttc	aagtgggtg	acatgggtgc	ctgagacagc	tcccacccac	1800
catcctgtcc	ccctggggcg	ggcattggcc	cagggtgg	agacagagac	aggccccctga	1860
acgataagca	ccaagagacc	ccaaggctcc	agatggaaaca	ctgagggtg	ccgtccctct	1920
cagaaggcc	ttccctggaa	gtgtccatgc	ctcacctgc	accatgtgg	tttctcagact	1980
tgaatcttc	cgggccccca	aaaagaagg	cetcattct	tatacctct	cctgtgattc	2040
aacacaaccc	atggagatgt	cccccttc	ctctgaaatc	atccatttg	ggacaatcc	2100
acattgggg	ctagaaacat	ccacgttatct	catcagccat	cttgcctgt	gcacccat	2160
agagggagg	tccatgttcc	tgctttgg	caattgttcc	ctctctgc	aggaaacaacc	2220
ctaaaaccag	accactccac	gcaggacagg	caggagat	tettctaaa	gcctccccca	2280
taaaaaggga	gtgtggatc	cacttagat	agggcggaa	catcttccac	ccggccaaagc	2340
tcctgtccag	atgttgaccc	tcacccageg	tgagctgtca	catagtagga	gcttcagat	2400
gcatgtggaa	gaaatggag	ttgtccat	gccttataaa	ctccccatga	tctgacatgc	2460
agaatccat	cctgtccag	aatcttctt	gaatttctt	gagcgaag	tatctgggg	2520
atttgtgg	attagggaga	ctgggttacaa	gggtggaaa	tagtccat	aatacactat	2580
gttgactat	gtgatccacc	ttgtgtatgt	taatattagg	tgttggaga	aggttgc	2640
attggccctg	ggacttctct	ctgcaggagg	agagaacgc	gcctctcc	tggattgtc	2700
tcaggcttcc	tttgtggat	ttgcacatc	tgctctgtc	caggagagg	2760	
ggctaagggg	ctggatccac	caaggcag	cacagggaa	aaactctgg	aatgaaccac	2820
tgaattcagg	ggatgggggt	gggggggg	ttctcgagg	gtgtggc	tacacgtgt	2880
ttctgtatgg	gtccagctgc	ttttccat	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 ttatacaac aggcaaggtagg 60
 ccagattttg cccacagttc ataatgtct gatccgtacc taggcgcgagaa gagaacccaa 120
 atatgaaaat ggtgaaagaa ttggacttga attatgttgc aacttgggtgc ccttgggggtt 180
 aagaggagaa ggtcgacggg cttggacttgc tggggcccaag tgggtatggc ggcacatcttc 240
 tggtgaactg tgaccgtat gatccggatc gtatgttca ggacaatttg gaccagcacg 300
 tgcactgtct gcaagacccg gaagacatgt ctgtcatgtt cctcgccgacg cagggccctg 360
 cagccctttt tgatgacacc aaacttgc tccataccatc cagctgtatg gccaacacgg 420
 cacagggtt ccacatctgc ggtctgttgc ggtctatagg catgtgttgc 480
 gccaagatata ggtgttctat gagggtaccc ggttgcgtt ggttgcggag cgcttcttc 540
 tggaaaggcc tgccttccct gatgcggct tcacaggact catcttcttc catgtcactc 600
 tgcgtggacca ctccaaacggat gatttctcg catccctat cttcaactgac actgtgtgt 660
 tcccgatggc acccttggatc atgacggcca gcactctgc acccttagag gtgtatgtt 720
 gccgtgtggc gaacaaacacg tgttttgtgg atgcgtgtgc agagctggcc aggaaggccg 780
 gctgcaagct gaccatctgc ccacaggccg agaaccgcacaa cgaccgttgc atccaggatg 840
 agatggagct gggctacgtt caggccgcgc acaagacccct cccgggtgtc tttgactccc 900
 caaggaaatgg ggaactgtcgag gatttccctt acaaaagaat cttgggttca gattttgggt 960
 acgtgactctg ggaaccacgc gagacgttgc tgatgtggcc ggtacttccctt gggaaacctgg 1020
 aggtcagccc tccagttgttgc gccaatggga aagagtaccc cttggggagg atccttattt 1080
 ggggcacactt gctgtggatc agtggccgcga gggttccatcca ggttgggtccgg gacttccctcc 1140
 atgcccaggaa ggtgcacggcc cccgtggatc tctttgtggc ctgggtggcc gtggccatcg 1200
 tggatgagg tttgtatgttgc tcccttgc cccatggggaa gggcttcccg atgtccctgg 1260
 ccagccctgg ggctgttcc aagetcttcc agggaaaagca gaagtgtggc cacggggaggg 1320
 cccttcgtt ccaggggggtt gttgtatgtatc agcaggtcaa gaccatctcc atcaaccagg 1380
 tgctctccaa taaagacctc atcaactaca ataagtttgt gcaagactgc atcgacttgg 1440
 accgtgagg gtcgtggcc gagctggggc tggcagagtg tgacatcatt gacatccac 1500
 agcttctcaa gaccgagagg aaaaaaaagca cggccctttctt ccctgacttg gtgaacatgc 1560
 tggtgcgtgg gaaagcacctg ggcatcccc agccctttgg gcccatacatc aatggctgt 1620
 gctgcctgg gggaaaagggtt cggccctgtc tggagccgtt gggccctccac tgacccatctca 1680
 ttgtatgactt cactccatcc acatgtgtgc atggggagg gcaactgtgc accaatgtgt 1740
 gcagaaaggcc ctcttttcc aagtgggtgg acatgggtgc ctgagacagc tcccacccac 1800
 catcctgtcc ccctggggcg ggcattggcc cagggtggg agacagagac aggccccctga 1860
 acgataagca ccaagagacc ccaaggctcc agatggaaaca ctgagggtg ccgtccctct 1920
 cagaaggcc ttccctggaa gtgtccatgc ctcacccatc acccatgtgg ttctcagact 1980
 tgaatcttc cggcccccca aaaagaaggat ctcattttctt tatagccctt cctgtgattc 2040
 aacacaaccc atggagatgt ccccttc tca ctctgaaatc atccatttg ggacaatcc 2100
 acattggggt cttagaaacat ccacgttatct catcagccat cttgttctgt gcatccat 2160

-continued

agaggaaggaa	tccatgattc	tgctttggtc	caattgttcc	ctctctgcag	aggaacaacc	2220
ctaaaaccac	accactccac	gcaggacagg	caggagagat	tcttcttaaa	gcctccccca	2280
taaaaaggaa	gctgtggatc	cacttagatc	aggcgccgaa	catcttcac	ccggccaagc	2340
tccctccccc	atgttgaccc	tcaccaggcg	tgagctgtca	catagtagga	gcttcttagat	2400
gcatgtggaa	gcaatgagag	ttgtccctta	gccttataaa	ctccccatga	tctgacatgc	2460
agaatccag	ccttgtccag	aatccctctg	gaatttcttgc	gagacgaaag	tatctggggg	2520
attgttgggt	actaggaga	ctgggtacaa	gggtgaaaaa	tagtcccat	aatacacatg	2580
gttgactatg	gtgatccacc	ttgtgtatgt	taatattagg	tgtctggaga	aggttgcttc	2640
attggccctg	gacttctct	ctgcaggagg	agagaacgc	gcctctcc	tggattggtc	2700
tcaggcttc	tgttgtccct	ttggtagctgc	ttccacatcc	tgctctgcgt	caggagaggg	2760
ggctaagggg	ctggatccac	caaggcagot	cacagcggga	aaactctgg	aatgaaccac	2820
tgaattcagg	ggatgggggt	ggggggccgg	tttcgcagg	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
 atatgaaaact gttgaagaac ttgggactga attatgttgc aacttgggtc cctggagtt 180
 aaaggagaaa ggtcgacggc cagttgggtc gggggcccaag tgggtatggc ggcacatgtc 240
 tggtaacttg tgaccgtat gatccggatc gtatgttca ggcacatgtt gaccaggcacg 300
 tgcactgcct gcaagacactg ggacacatgt ctgtcatgtt cctgcggacg cagggccctg 360
 cagcccttcc ttatgaccac aaacttgc tccatcaccc cagctatgtat gccaaacggg 420
 cacaggcttc ccacatggc ggtagtgggg atgtgtgtgc ggcctatagg catgtgtgg 480
 gccaagataa ggtgtcttat gggatccccc gtttgcgtatgg ggtatggggg cgcttcttcg 540
 tggaaaggcct gtccttcctt gatgcgggt tcacaggact catctcttc catgtcaactc 600
 tgctggacga tcccaacggag gatttctcgat catcccttat cttcaactgac actgtgtgt 660
 tcccgagttgtt atgacgtccca gcactctgccc accccatggag gtgtatgtgt 720
 gccgtgtgag gaacaacacg tgggttggg atgcgggtgc agagctggcc aggaaggccg 780
 gctgcgaact gaccatctgc ccacaggccg agaaccgcacg cgaccgttggg atccaggatg 840
 agatggagct gggctacggtt caggggccgc acaagaccctt cccgggtggc tttgactccc 900
 caaggaaatgg ggaactcgac gattttccctt acaaaaatggat cctgggttca gattttgtt 960
 acgtgactcg ggaaccacgc gacaggctgt tgatgtggctt ggactctttt gggaaacctgg 1020
 aggttcagccc tccatgtgtt gccaatggga aagatgtacc cctggggagg atcccttatt 1080
 ggggcaacctt gcttgggtca agtggccgc gggtcacccca ggttgggtgc gacttccccc 1140
 atggcccaaaa ggtgcggggc cccggggggc tctttgtggc ctgggtggc gtggggccatg 1200
 tggatgggtt ttgtgtgtt gtccttcgtt ccgtatggggaa gggcttcggg atgtctctgg 1260
 ccagccctgg ggcctgttcc aagcttccca agggaaaacca gaaatgtggc cacggggagg 1320
 ccctctgtt ccagggggggtt gttgtatgtt agcaggatca gaccatctcc atcaaccagg 1380
 tgcctctccaa taaagacacttca atcaactataa ataaatgttgc gcaatgtgc atcgacttgg 1440
 accctgtgggtt gctgaaggccg gagttggggc tggcaaggatc tgacatccattt gacatccac 1500
 agtcttcaaa gaccggagggg aaaaaaggaa cggcccttctt ccctgtactt gtaacatgc 1560
 tgggtgtggg gaaacccatgtt ggcacatccca agccctttggg gcccacatc aatggctgt 1620
 gtcgcctgg gggaaagggtt cgggtccctgc tggagccggc gggccctccac tgcacccatc 1680
 ttgtatggactt cactccatcc acatgtgc atggggaggat gcaatgtggc accaactgtgt 1740
 gcaaaaaaccctt ctcttttccaa aatgtgttgc acatgtgtcc ctggacacgc tccaccac 1800
 catccgttcc ccctggggcg ggcattggcc cagggtgggg agacagagac agggccctgt 1860
 acgataacgc ccaaggacacc ccaaggctcc agatggaaaca ctgggggtga ccgtccctct 1920
 cagaagectt ttccctggaa gtgtccatgtt ctcacatggca accccatgtgg ttctcagact 1980
 tgaatcttcc cggccccccca aaaaaggaaat cctcattttcataatccctt cctgtgttcc 2040
 aacacaacccctt atggagatgtt cccttctca ctgttgcataatccatgggg ggacaaatcc 2100
 acatgggggtt ctggaaatccatccatgttccatgttccatgttccatgttccatgttccatgttcc 2160
 agggaaaggaaatccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttcc 2220
 ctaaaaaccaccc accactccac gcaaggacaggc caggagagat ttttcttaaa gcctccccca 2280
 taaaaaggaaatccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttcc 2340
 tcctgtccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttcc 2400
 gcatgtggaaatccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttcc 2460
 agaaatccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttcc 2520
 attgttgggtt atggggatccatgttccatgttccatgttccatgttccatgttccatgttccatgttcc 2580
 gttgactatgatcgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttcc 2640
 attggccctg gacttctctt ctgcaggagg agagaacgc gcttctcc tggattggtc 2700
 tcaggcttc tgggtccctt tggatccatgttccatgttccatgttccatgttccatgttccatgttcc 2760
 ggctaaggggg ctggatccac gcaaggcagot cacagcggga aaactctggg aatgaaccac 2820
 tgaattcagg ggtatgggggtt gggggggccgg ttttcgcagggtt gttgtccatgttccatgttcc 2880
 ttctgtatgg gttccatgttccatgttccatgttccatgttccatgttccatgttccatgttccatgttcc 2938

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

-continued

```

note = protein-arginine deiminase type-3 isoform X2 [Homo
      sapiens]
source      1..485
mol_type = protein
organism = Homo sapiens
SEQUENCE: 42
MSVMVLRTQG PAALFDDHKL VLHTSSYDAK RAQVFHICGP EDVCEAYRHV LGQDKVSYEV 60
PRLHGDEERF FVEGLSFPDA GFTGLISFHV TLLDDSNEDF SASPIFTDTV VFRVVPWIMT 120
PSTLPPLEVY VCRVRNNTCF VDAVAELARK AGCKLTICPQ AENRNRDRIQ DEMELGYVQA 180
PHKTLPPVFFD SPRNGELQDF PYKRILGPDF GYVTREPRDR SVSGLDSFGN LEVSPPVAN 240
GKEYPLGRIL IGGNLPGSSG RRVTQVVRDF LHAQKVQPPV ELFVDWLAVG HVDEFLSFVP 300
APDGKGFRML LASPGACFKL FQEOKCCHG RALLFQGVVD DEQVKTISIN QVLSNKDLIN 360
YNKFVQSCID WNREVLKREL GLAECIDIID PQLFKTERKK ATAFFPDLVN MLVLGKHLGI 420
PKPFGPIING CCCCLEEKVRS LLEPLGLHCT FIDDFTPYHM LHGEVHCGTN VCRKPFSFKW 480
WMNMVP      485

SEQ ID NO: 43      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: 43
agaatggat ggatgtgact gtgtgctaat aaaactttat ttataaaaaac aggcaaggtagg 60
ccagattgg cccacagttc ataatgtgt gatccctgacc taggcggagaa gagaaaccaa 120
atatatggaaact gtttgaagaac ttgggactga attatgttg aacttgggc cctggggattt 180
aaaggaggaaa ggttcggcg 240
tggtaactg tgaccgtatg gatccggact gtgtatgtcca ggacaatttg gaccaggc 300
tgcactggctt gcaatggacatc ctgtcatgggact cctgcggacg cagggccctg 360
cagccctttt tggatggacc ac 420
cacaggctttt ccacatctgc ggtcttgagg atgtgtgtga ggcctataagg catgtgtcg 480
gccaagataa ggtgtccat gaggtaaaaaaccc gcttgcattgg 540
tggaggcgtt gtccttcctt gatgcgggtt tcacaggact catcttcctt catgtcaactc 600
tgcggacgac 660
ctccaacagat gatcccattt catccccatcatcacttgc actgtgggt 720
tccggatggc accctgtatc atgacggccca gcaactgtcc accccctagag gtgtatgtgt 720
gcccgtgtgag gaaacaacacg tggtttgtgg atgcgggtggc agagctggcc agggaggccg 780
gctgcgaactt gaccatctgc ccacaggccg agaaccgcggca cgaccgcgtgg atccaggatg 840
agatggagat gggctacgtt caggccgcg acaagacccccc cccgggtggc tttgactccc 900
caaggaaatgg ggaactgtcag gatccccctt acggaaaatgtt cttgcggatc gatgtgggt 960
acgtgactcg ggaaccacgc gacaggctgtg tgagtggccctt ggactctttt gggaaacctgg 1020
aggtcaggccc tccatgtgtt gccaatggggaa aaqagtatggggcccttggggaggatccctt 1080
ggggcaactt gctttgttca agtggccggc gggtcacccca ggtgggtgggg gacttctcc 1140
atgccccagaa ggttcggccccccccccttggggatccttccttggggatccttggggatcctt 1200
tggatggatgtt tctgatgtttt gtcctgtccc cccatggggaa gggcttccgg atgctctgg 1260
ccacgtcttggc ggcctgttccca aagctcttccca aggaaatggccca gaatgtggcc cacggggagg 1320
ccctctgtttt caagggggtt gttgtatgttggccca gggcccttccatcacttccatcactt 1380
tgtctccaa taaagacccatc atcaactaca ataaatgtgtt gcaatgttgc atcgacttgg 1440
accgtgaggat gctgtggcc gatgtggcccttggcagatgttgc tggatccatc gacatccac 1500
agctcttcaaa gaccgtggggaaaatggccca cggatcttccatcctgttgcgttgc 1560
tggatgtggggaa gaagcgttccatc ggcatggggaa agecccttggccatccatc aatggctgt 1620
gctgcctgtgg gggaaagggtt cgggtttttccatcacttccatcacttccatcacttccatc 1680
tttgcgttccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 1740
gcagaaaaggcccttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 1800
catccctgttccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 1860
acgataagccca ccaagggatccatcacttccatcacttccatcacttccatcacttccatcacttccatc 1920
cagaageccatc ttccctgtggaa gtgtccatcacttccatcacttccatcacttccatcacttccatc 1980
tggatcttccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2040
aacacaacccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2100
acatggggatccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2160
agaggaaggatccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2220
ctaaaaccatc accactccac gcaggacaggc caggagagat tttccatcaa gcccctccatcacttccatc 2280
taaaaaaggaa gctgtgttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2340
tctctgtggccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2400
gcatgtgtggaa gcaatgtggccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2460
agaaatccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2520
attgtggggatccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2580
gttgactatgttccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2640
attggccctgtggatccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2700
tcaggcttcatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2760
ggctaaaggggccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2820
tgaattcaggatccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2880
ttctgtatgg gttccatcacttccatcacttccatcacttccatcacttccatcacttccatcacttccatc 2938

```

SEQ ID NO: 44 moltype = AA length = 485

-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 gtatgtccca ggacaattgt gaccagcacg 300
tgactgtcctt gcaagacatgt ctgtcatgtgt cctgcggacg caggggccctg 360
cageccctttt tgatgaccac aaacttgtcc tccataccctc cagctatgtat gccaacacggg 420
cacaggctttt ccacatctgc ggtcttgagg atgtgtgtg ggcctataagg catgtgtctgg 480
gccaagataa ggtgtcttat gaggtacccc gcttgcattgg ggtatggggag cgcttcttcg 540
tggaggcgt ctccctccctt gatgcgggtc tcacaggact catctcttc catgtcaactc 600
tgctggacga ctcacacggat gatttctcg catccccat ttcactgtac actgtgtgt 660
tccggatgtgc accctgtatc atgacccca gcactctgc accccctagag gtgtatgtgt 720
gccgtgttag gaacaacacgg tggtttgtgg atgcgggtggc agagctggcc aggaaggccg 780
gtctgcacgt gaccatctgc ccacaggccg acaacccgcaaa cgaccgcgtgg atccaggatg 840
agatggagct gggctacgtt cagggccgcg acaagacccgcg cccgggtggc tttgactccc 900
caaggaatgg ggaactcgat gatccccat tccatggggat ctttgcgtt gatgtgtgtt 960
acgtgaatcg ggaaccacgc gacaggctgt tgatgtggcctt ggactctttt gggaaacctgg 1020
aggtcagccg tccatgtgtg gccaatgggg aagagatgggg cctggggagg atccctcattt 1080
ggggcaactt gctctgggtca agtggccgcg ggttcacccca ggtgggtgggg gacttccttc 1140
atgccccagaa ggtgcagccc cccgtggago tctttgtggc ctggatggcc gttggccatcg 1200
tggatgagtt tctgagttt gtccctgccc ccgtatggggaa gggcttccgg atgctctgg 1260
ccagccctgg ggcctgttca aagcttccctt aggaaaggca gaatgtggc cacggggagg 1320
ccctctgtt ccagggggtt gttgtatgtatc agcagggtcaa gaccatctcc atcaaccagg 1380
tgcctctccaa taaagacccatc atcaactaca ataaatgttgcg ctagagctgc atgcacttgaa 1440
accgtgaggat gctgaacggg gagctggggc tggcagatgt tgacatcatt gacatccac 1500
agctcttcaa gaccgtgggg aaaaaggccaa cggccctttt ccctgtactt gtgaacatgc 1560
tggtgctggg gaaacgttcc ggcacatgttcc agccctttgg gcccacatc aatggctgt 1620
gtctgcctgg gggaaagggtt ctagccctgc tgtagccgcgtt gggccctccac tgccatctca 1680
tttgcgttccatc cactccatc acatgtgtc atggggaggat gcaactgtggc accaatgtgt 1740
gcagaaaaggcc ctcttcttcc aagtgggtggc acatgggtggc ctggagacgc tccaccac 1800
catccctgttcc ccctggggccg ggcattggcc cagggtgggg agacagagac agggccctga 1860
acgataagacc ccaagagacc ccaaggctcc agatggaaaca ctgagggttgc ccgtcccttc 1920
cagaageccctt tccctggaa gtgtccatgc ctcacactgc accccatgtgg ttctcagact 1980
tgaatcttccatc cggccccccca aaaagaaggaa cctcattttt tatagcttctt cctgtatttc 2040
aacacaaccc atggagatgtt ccccttctca ctctgaaatc atccatgggg ggacaaatcc 2100
acatgggggtt ctgatggatcat ccatggccat cttgtctgtt gcatcttaac 2160
agagggaaaggaa tccatgttccatc tgctttgtggc caattgtttcc ttctctgtcag aggaaccaacc 2220
ctaaaaccac accactccac gcaggacagg caggagatgt tttctcaaa gcctccccc 2280
taaaaaggaa gctgtgttgc cacttagatc agggcggacac catctttcac ccggccaaacg 2340
tccctggccatc atgttgcaccc tcacccaggcg tgagatgtca catagtggaa gcttcttagat 2400
gcatgtggaa gcaatggggat ttgtccctta gccttataaa ctccccatga tctgacatgc 2460
agaaatccatc ctttgcgttca aatcttccctg gaatttcttgg gagacaaag tatctggggg 2520
attgttgggtt actaggggaa ctgggttacaa ggggtggaaat tagttccat aatacacatg 2580
gttgactatg gtgtatccacc ttgtgtatgtt taatattagg tgcgttggaga aggttgcctc 2640
attggcccttc ggacttcttc ctgcaggagg agagaacgcg gcctctcc tggattggc 2700
tcaggcttc tggatggccctt tggtcagctt ttccacatcc tgcgttgc caggagagg 2760
ggctaaaggaa ctggatccac caaggcagct cacagccggaa aaactctgggaa atgaaccac 2820

```

-continued

```
tgaattcagg ggatgggggt gggggggcggt ttctcgaggt gtgtgccac tacacgtgtg 2880
ttctgtatgg gtccagctgc gttccatca ctgcataata aatcaacaga aacacaaa 2938
```

```
SEQ ID NO: 48 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: 48
MSVMVLRTQG PAALFDDHKL VLHTSSYDAK RAQVFHICGP EDVCEAYRHV LGQDKVSYEV 60
PRLHGDEERF FVEGLSFPD GFTGLISFHV TLLDDSNEDF SASPIFTDTV VFRVAPWIMT 120
PSTLPPLEVY VCRVRNNTCF DAVAAELARK AGCKLTICPQ AENRNDRWIQ DEMELGYVQA 180
PHKTLPVFSPD PRNGELQDF PYKRILGPDF GYVTREPRDR SVSGLDSFGN LEVSPPVAN 240
GKEYPLGRIL IGGNLPGSSG RRVTVQVRDF LHAQKVQPPV ELFVDWLAVG HVDEFLSFVP 300
ADPGKGFRML LASPGACFKL FQEOKCCHG RALLFQGVVD DEQVKTISIN QVLSNKDLIN 360
YNKFVQSCID WNREVLKREL GLAECIDI DI PQLFKTERKK ATAFFPDLVN MLVLGKHG 420
PKPFGPIING CCLEEKVQS LLEPLGLHCT FIDDFTPYHM LHGEVHCGTN VCRKPFSFKW 480
WNMVP 485
```

```
SEQ ID NO: 49 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: 49
ctaaggggc ctcagggca gtgttgggg tggcgccac agctaagtcc aacaccagca 60
tgtcgctca gagaatcggt cggtgtccc tgtagcatcc caccagcgcc gtgtgtgtgg 120
ctggcggtgg aaccctcggt gacatttatg ggtcagtgc tgagggcaca gaaatgttg 180
aggctcgatcc gaccgctggc gtggacatcc acatctctcc caacatggag agggggccgg 240
agcgtgcaga caccaggccg tgccgtttt acgcgactt ggagatcatc gtgtgtatga 300
actccccccat caatgaccc aacgacagcc atgttacatgat ttcctaccac tccagccatg 360
agcctctgcg cctggcttat gccggctct acctcacccg tggatggatc tctctggatt 420
ggcacccatgg ctgtgtggg aggccaggaa ggaactttt agacaacggg cagtgggtct 480
ggggggccatgg tgggtatggc ggcacatcc tggatggact tgaccgtat gatccgagct 540
gtgatgtcca ggacaattgt gaccacacg tggactgtcc gcaagactg gaagacatgt 600
ctgtcatgtt cctggcgacg cagggccctg cagcccttcc tggatggaccac aaacctgtcc 660
tccataaccctc cagctatgtat gccaaacccggg cacaggctt ccacatctgc ggtccgtggg 720
atgtgtgtga ggcctatagg catgtgtccatgg gcaactatggat ggtgtctat gaggtacccc 780
gtttgcatgg ggtatgggggg cgttctccg tggaggccct gtccttcct gatccggct 840
tcacaggact catctccctt catgtcaactc tggatggacg ctccaaacgg gatttctccg 900
catccccatcc ttcactgtac actgtgtgt tccggatggcc accctggatc atgacgccc 960
gcactctggc accctgtatgg gtgtatgtt gccgtgtggaa gaaacacacg tggatgtgg 1020
atgcgggtggc agagctggcc aggaaggccg gtcgcaacgtt gaccatctgc ccacaggccg 1080
agaaccgcaa cgaccgtgtt atccaggatg agatggatc gggctacgtt caggccgcgc 1140
acaagacccct cccgggtgtc ttgtactccc caagggatgg ggaactgtccg gatttccct 1200
acaaaaaaaat cctggatgtt gtcggccagg gtcacccagg tggatggggaa cttcctccat 1260
gccccagaaagg tgcagcccccc cgtggatgtt tttgtggact ggttggccgtt gggccatgtg 1320
gatgagtttcc tggatgtttt gtcggccccc gatggaaagg gttccggat gtcctggcc 1380
agccctgggg cctgctcaa gtcggatgtt gaaaaggcaga agtggccca cggggggcc 1440
ctctgttcc ag 1452
```

```
SEQ ID NO: 50 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: 50
MSLQRIVRVS LEHPTSAVCV AGVETLVDIY GSVPEGTEMF EVYGTPGVDI YISPNMERGR 60
ERADTRRWRF DATLEIIIVVM NSPSNDLNDs HVQISYHSSH EPLPLAYAVL YLTCVDISLD 120
CDLNCEGRQD RNFDVKRQWV WGPSCYGGIL LVNCDRDDPS CDVQDNCDQH VHCLQDLED 180
SVMVLRTQGP AALFDDHKLV LHTSSYDAKR AQVFHICGP DVCEAYRHVL GQDKVSYEV 240
RLHGDEERFF VEGLSPFDAG FTGLISFHVLL DSDNEDFS ASPIFTDTVV FRVAPWIMT 300
STLPPLLEVYV CRVRNNTCFV DAVAAELARKA GCKLTICPQA ENRNDRWIQC EMELEYVQAP 360
HKTLPVFVDS PRNGELQDFP YKRILVKWPQ GHPGGAGLPP CPEGAAPRGA LCGLVGRGPC 420
G 421
```

```
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 gtcacccagggtt tggtggggat 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 LCGLVGRG	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 QGDKVSYEV 240
RLHGDEERFF VEGLSFPDAG FTGLISFHVT LLDDSNEDFS ASPIFTDTVV FRVAPWIMTP 300
STLPPLEVVY CRVRNNTCFV DAAEALARKA 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 caatgacccct aacgcacccg atgttcagat ttcttaccac tccagccatg 360
agcctctggc cttggcttat gccgtgtctt acctcaccc tttttggatt 420
gcgacacctaa ctgtggggaa aggccggaca ggaactttgt agacaagccg cagtgggtct 480
ggggggcccg tgggtatggc ggcacatcttgc tggtaacttgc tgaccgtgtat gatccgagct 540
gtgtatgttca gcaacatgtt gaccggaca tgcactgcctt gcaagacactt gaaagacatgt 600
ctgtcatgtt cttggcgacg caggccccctg cagcccttctt tgatgaccac aaacctgtcc 660
tccataccctc cagctatgtat gccaaacccgg cacaggtctt ccacatctgc ggtcttggagg 720
atgtgtgtga ggcctatagg catgtgtgg gccaagatata ggtgtccat gaggttacccc 780
gcttgcattgg ggttggggat cgttcttgc tggaaaggccct gtccttccct gatgccgct 840
tcacaggact catctcttc catgtcactc tggtaacttgc ctccaaacgg gattttctcg 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 gagaatcgtg cgtgtgtccc tggagcatcc caccagcgcg gtgtgtgtgg	120
ctggcgtgaa gaccctcgta gacattttatg ggtcagtgc tgagggcaca gaaatgtttg	180
aggtctatgg gacgcgtggc gtggacatct acatctctcc caacatggag agggggccggg	240
agcgtgcaga caccaggcgg tggcgtttt acgcgcattt ggagatcatc gtggcatga	300
actccccccag caatgacctc aacgacagec atgttcagat ttccattaccac tccagccatg	360
agcctctggc cctggctat gccgtgtctt acctcacctt tggtgacatc tctctgttt	420
gcgacacctgaa ctgtggggg aggccggaca ggaactttgtt agacaaggccg cagtgggtct	480
ggggggcccg tgggtatggc ggcattttgc tggtgatggc tgaccgtat gatccggat	540
gtgtatgtcca ggacaatttg gaccagcaag tgcaactgcct gcaagacccg gaagacatgt	600
ctgtcatgtt cctggcgacg cagggccctt tgatgaccac aaacttgtcc	660
tccataccctc cagctatgtat gccaaacccggg cacaggatcc ccacatctgc ggtctgtgg	720
atgtgtgtga ggcctatagg catgtgtggc gccaagataaa ggtgtcttat gaggtacccc	780
gcttgcattgg ggatgggggg cgcttcttcg tggaaaggccc gtccttcctt gatccgggt	840
tcacaggact catctccctt catgtactc tgctggacga ctccaaacgg gattttctcg	900
catccccctt cttcaactgac actgtgtgtt tccgagtttggt accctggatc atgaegccca	960
gcactctggc acccccttagag tggtatgtt gccgtgtgg gacaacaacccg tggtttgtgg	1020
atgcgggtggc agagctggcc agggaggccc gctgcaagct gaccatctgc ccacaggccc	1080
agaaccgcua cgaccgctgg atccaggatg agatggatc gggctacgtt caggcgcgc	1140
acaagaccct cccgggtgtc ttgtactccc caaggaatgg ggaactgcag gattttccctt	1200
acaaaaagat cttgttcaag tggccggcagg gtcacccagg tggtgccggaa cttccat	1260
gccccagaaagg tgcagccccc cgtggagctc ttgtggact ggttggcgtt gggccatgt	1320
gatgagtttgc tgagctttgt ccctggcccc gatggggagg gttccggat gtcctggcc	1380
agccctgggg cctgctcaa gctttccag gaaaaggcga agtggccca cgggaggggcc	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 DATLEIIIVVM 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 gagaatcgtg cgtgtgtccc tggagcatcc caccagcgcg gtgtgtgtgg	120
ctggcgtgaa gaccctcgta gacattttatg ggtcagtgc tgagggcaca gaaatgtttg	180
aggtctatgg gacgcgtggc gtggacatct acatctctcc caacatggag agggggccggg	240
agcgtgcaga caccaggcgg tggcgtttt acgcgcattt ggagatcatc gtggcatga	300
actccccccag caatgacctc aacgacagcc atgttcagat ttccattaccac tccagccatg	360
agcctctggc cctggctat gccgtgtctt acctcacctt tggtgacatc tctctgttt	420
gcgacacctgaa ctgttgaggaa aggccggaca ggaactttgtt agacaaggccg cagtgggtct	480

-continued

ggggggcccg	tgggtatggc	ggcatcttgc	tggtaactg	tgaccgtgat	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	gggtggatgt	acgttggaa	tggccgttgg	tagcacaggc	gaccagaagg	300
ttcagatttc	atactacgg	cccaaaactc	caccaggtaa	agctctactc	tacccatccg	360
gggtggaaat	tccttgc	gcagacatca	cccgccacgg	caaagtgaag	ccaaaccaggag	420
ctgtggaaaga	tccagggacc	tggacttgg	ggcccttgc	acagggttgc	atccctgttgg	480
tgaactgttgc	cgagacaaat	ctcgaatctt	ctgcacatgg	ctgcggaggat	gtgaagtgc	540
ttgacacgca	agacctgcag	gacatgttgc	tgtatgttgc	gagcacaagg	acccccaagg	600
acttcttcac	aaaccatata	ctgggtgttcc	acgtggccag	gtctgttgc	gacaaagtga	660
gggtgtttca	ggccacacgg	ggcaactgt	cctccaaatgt	cagcgtatgc	ttgggttccca	720
agtggcccttc	tcaatctatc	atggttcccg	gtggaaaggc	caacatggac	ttctatgttgg	780
aggcccttcgc	tttccggac	accgttcc	cggggctcat	tacccttacc	atctccctgc	840
tggacacgtc	caacctggag	ctcccccgg	ctgtgtgtt	ccaagacgc	gtgttcttcc	900
gctgtggccgc	ctggatcatc	accccaaca	cccgccccc	cgaggagggt	tacgtgttgc	960
gtatttttgtt	aatggggac	tctctgttgc	ctgtgttgc	aaagccatgt	1020	
gcaagctgac	catctggccct	gaggaggaga	acatggatgc	ccagtggatg	caggatgaaa	1080
tggagatcg	ctacatccaa	gccccacaca	aaacgctgc	ctgggtcttc	gactctccaa	1140
ggaacacgg	ctgtggagg	tttccatca	aacgcgttgc	gggttccatg	tttggctatg	1200
taactcgagg	gcccccaaca	gggggtatca	gtggacttgg	ctcccttggg	aacctggaaag	1260
ttagcccccc	agtccatgc	agggggcaagg	aatacccgc	gggcaggatt	ctctccgggg	1320
acagctgttca	tcccacat	gacagccgc	agatgcacca	ggccctgcag	gacttccatc	1380
gtgcccacgg	gtgtggggcc	cctgtgttgc	tcttattttgc	ctgggtgtcc	gtggggccacg	1440
tggacggat	cctggatctt	gtggccacgc	ccgacaggaa	ggggctccgg	ctgtcttgg	1500
ccagcccccc	gtccctgttac	aaactgttcc	aggaggacgc	gaatggggc	cacggggagg	1560
ccctgttgc	cgaaaggatc	aagaaaaaaa	aacagcagaa	aataaaagaa	attctgtcaa	1620
acaagacatt	aattttttgtt	tggagatgtt	catgcactgg	aaacccgcac	1680	
tgctgttgc	ggggctgggc	ctggccatgc	tgacatccgc	cagctcttca	1740	
agctcaaaaga	gttcttcata	ggggaaatgtt	tttccatca	catgtgttgc	atgtgttgc	1800
tagggaa	cttggggatc	cccaagccct	teggggccgt	catcaacggc	cgctgttgc	1860
tggaggagaa	gtgtgttcc	ctgtgttgc	cactgggcct	ccagtgatcc	ttcatcaacg	1920
acttcttcac	ctaccatcatc	aggcttgggg	agggttgc	gggcacccaa	gtggccagaa	1980
agecccttc	cttcaatgttgc	tggaaatgttgc	ttggccatgc	ccatcttcc	ttgggttgc	2040
tccttcctgg	ccatgtgttgc	ctgggttcc	ctgtgttgc	caagcaagg	ctcttgc	2100
tattgtgttgc	ccctggggcc	ggccagccct	cccaggatgt	ggttgcatttgc	tttccatgttgc	2160
atgtcccaatgttgc	ttccatcttgc	gaatgttgc	acatgggttgc	agcactgcac	actcgttgc	2220
gtcttaaaatgttgc	gtgttgcataaa	agtttttttgc	agtcacttttgc	tatcatgttgc		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

```

SEQUENCE: 62
organism = Homo sapiens

MAQGTLIRVT PEQPTHAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60
KSTGSSTWPL DPGVEVTLTM KVAGSGTGQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
ITRTGKVKPT RAVKDQRTWT WGPGCGQAIL LVNCDRDNLE SSAMDCEDDE VLDSEDLQDM 180
SLMTLSTKTP KDFFTNHTLV LHVARSEMOK VRVFQATRGK LSSKCSVVLG PKWPSPHYLMV 240
PGGKHNMDFY VEALAPPDTD FPGLITLTIS LLDTSNLELP EAVVFQDSVV FRVAPWIMTP 300
NTOPPQEVEYA CSIFENEDFL KSVTTLAMKA KCKLTICPEE ENMDDQWMQD EMEIGYIQAP 360
HKTLPVVFDs PRNRGLKEFP IKRVMGPDFG YVTRGPQTGG ISGLDSFGNL EVSPPVTVRG 420
KEYPLGRILF GDSCYPSNDS RQMHOALQDF LSAQQVQAPV KLYSDWLSVG HVDEPLSFVP 480
APDRKGFRLL LASPRCYSKYL FQEQQNEGHG EALLFEGIKK KKQKIKNIL SNKTLREHNS 540
FVERCIDWNR ELLKRELGLA ESDIDIDIPOL FKLFKEFSKAE AFFPNMVNML VLGKHGLGIPK 600
PFGPVINGRC CLEEKVCSSL EPLGLQCTFI NDFFTYHIRH GEVHCNTNVR RKPFSSFKWWN 660
MVP 663

SEQ_ID NO: 63 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: 63
agccagaggc acgagcttagc ccgacgatgg cccaggggac attgatccat gtgaccagg 60
agcagccac ccatgcctgt tggtgtctgg gcacccgtac tcagtttgc atctgcagct 120
ctggccctga ggactgcacg tccttcacgc tcaacgcctc cccagggggt gtctggata 180
ttggcccacgg ccctccacgg aagaagaaat ccacaggttc ctccacatgg cccctggacc 240
ctggggtaga ggtgaccctg acgtgaaaag tggccagtgg tagcacaggc gaccagaagg 300
tttagatctt taactacggc cccaaacactc caccatcaa agctctactc taccttacccg 360
gggtggaaat ctcccttggtc gcagacatca cccgcacccg caaagtggaa ccaaccagg 420
ctgtggaaaga tcaagggacc tggacccatggggcc gcccattggg acagggtggc atccctgtgg 480
tgaactgtga cagagacaat ctgcattttt ctgcacatggc ctgcggaggat gatgaatgtc 540
tttgcacatcgca gagactgcacg gacatgtcgc tgatgaccctt gggcacgaa acccccaagg 600
acttcttcac aaaccatata ctgtggctcc acgtggccgg gtctggatg gacaaatgtg 660
gggtgtttca ggcacacgg ggcacaaatgtt cctccatggc cagcgtatgc ttgggtcccc 720
agtggccctc tcaactacctg atggtcccccgtt gttggaaatgc caacatggc ttctacatgtt 780
agggccctcgc tttccggac accgcattcc cggggctcat taccctcacc atctccctgc 840
tggacacatgc caaccatggcgttccggatgg ctgtggatgtt ccaagacgc gtggcttc 900
gcgtggccgc cttggatcatc aaaaaaaaacca cccaggcccccc gcaaggatgtg tacgcgtgca 960
gtatatttttta aatgtggac ttccatggatgtt cttccatggc tctggccatgtt aatggcaatgt 1020
gcaagctgac catctggccctt gaggaggaga acatggatgc ccagtggatg caggatgaaa 1080
tggagatccgg ctacatccaa gccccacaca aaacgtgtcc cgtggcttc gactctccaa 1140
ggaaacagagg ctgtggatgttcccccata aacgcgtgtt ggggtccatgtt tttggatgt 1200
taactcggc gcccacaaaatc ggggttatacgtt gttggactggc ctcccttggg aacctggaaag 1260
tgagcccccc agtcacatgc agggggcaagg aataccggcgtt gggcaggattt ctcttgggg 1320
acagctgttta tcccaatgcacatgc acacggccgg agatgcacca gggccctggc gactccctca 1380
gtggccacggc cttgtggcgttcccccata aacgcgtgtt gttggatgtt ggggtccatgtt 1440
tggacgttcccccata aacgcgtgtt ggggtccatgtt tttggatgt 1500
ccagccccccatc aacgcgtgtt ggggtccatgtt tttggatgt 1560
ccctgtgttcccccata aacgcgtgtt ggggtccatgtt tttggatgt 1620
acaagacatc gggggccggc aatccatgtt gggcaggatgtt ggggtccatgtt tttggatgt 1680
tgcgtgttcccccata aacgcgtgtt ggggtccatgtt tttggatgt 1740
agctcaaaga gttctctaag gggggccggc aatccatgtt ggggtccatgtt tttggatgt 1800
tagggaaatgc cttgggtcccccata aacgcgtgtt ggggtccatgtt tttggatgt 1860
tggggggatgc gttgtgttcccccata aacgcgtgtt ggggtccatgtt tttggatgt 1920
acttcttcac ctaccacatgc aatccatgtt ggggtccatgtt tttggatgt 1980
agecccttc tcccaatgtt gggggccggc aatccatgtt ggggtccatgtt tttggatgt 2040
tccctgtgttcccccata aacgcgtgtt ggggtccatgtt tttggatgt 2100
tattgtgttcccccata aacgcgtgtt ggggtccatgtt tttggatgt 2160
atgtcccaatgc aatccatgtt ggggtccatgtt tttggatgt 2220
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 2280
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 2340
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 2400
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 2460
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 2520
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 2580
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 2640
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 2700
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 2760
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 2820
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 2880
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 2940
gtctcaaaatgc aatccatgtt ggggtccatgtt tttggatgt 3000

```

-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 ctccctgtgc ggcacatca cccgcacccgg caaagtggaaat ccaaccaggag 420
 ctgtgaaaga tcaagggacc ttggacctggg gcccctgtgg acagggtgcc atccctgtgg 480
 tgaactgtga cagagacaat ctgcacatcc ttgtccatgttgc ctgcgaggat gatgaagtgc 540
 ttgacacgcca gagacctgtgc gacatgtcgcc tgatgacccat gggccacgg acccccaagg 600
 acttcttcac aaaccatata cttgtgtcc acgtggccatg gtcttgatg gacaaatgt 660
 ggggtgttca ggcacacacgg ggcacaaatgt cttcccaatgtc cagcgtatgc ttgggtccca 720
 agtggccctc tcaactacccgt atggcccccgg gttggaaacca caacatggac ttctacgtgg 780
 aggccctcgc tttccggac accgacttcc cgggggtccat taccctcacc atctccctgc 840
 ttggacacgctc caacctggggcc cttccggagg ctgtgggtgttgc ccaacacggc gtggcttcc 900
 ggcgtggccgg ctggatcatg acccccaaca cccagcccccc gcaggagggtg tacgggtgc 960
 gtattnnntta aataggaggac ttctctgttgc cagtgtactac tctggccatg aaaggccaatg 1020
 gcaagcgtgc acatgtccctc gaggaggaga acatggatggc ccagggtggatg caggatggaaa 1080
 ttggatggccgtt cttccatccaa gcccccaacaca aacacgttgc cttgggttcc gacttccaa 1140
 gggaaacaggat cttccatccaa aacacgttgc ggggtccatgat tttggctatg 1200
 taactcgagg gcccccaaca gggggtatca gtggactggc ctcccttggg aacctggaaag 1260
 ttagcccccggc agtccacggc aataccggc aatggggatggc gggcaggatt ctcttcgggg 1320
 acagctgttca tccccatccaa gacacccggc agatgtccaa gggccctgtc gacttccca 1380
 gtgccccacgg cttgtgttgc tctatccatgttgc ctgggttcc gttggccacgg 1440
 ttggacagttt cttgtgttgc tttccatccaa aacacgttgc ggggtccatgat tttggctatg 1500
 ccacggccatc ttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 1560
 ccctgtgttgc cttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 1620
 acaaaacatc tttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 1680
 tgctgttgc tttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 1740
 agtctcaaaatg tttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 1800
 tagggaaatggc tttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 1860
 ttggaggaaatggc tttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 1920
 acttcttcac tttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 1980
 agcccttcac tttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 2040
 tccctcttgc tttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 2100
 tatttgtggcc tttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 2160
 atgtcccactt tttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 2220
 gctctaataatggc tttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 2280
 gctctaataatggc tttccatccaa aatccatccaa aacacgttgc ggggtccatgat tttggctatg 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 KVAGSGSTGQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
 ITRTGKVPT RAVKDQRTWT WGPGCQGAIL LVNCDRDNLE SSAMDCEDE VLDSEDLQDM 180
 SLMTLSTKTP 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
agccaaacctt agctgtgaaa gatcggacat gcaggacatg tggctatgt ccctgaggac 120
gaagaccccc aaggacttct tcacaaacca tacactgggt ctccacgtgg ccaggctgt 180
gtatggacaaa gtgggggtt ttcaaggccac acggggcaaa ctgtcttcca agtgcagcgt 240
agtcttcgtt ccaactgtgc ctttactata cttgtatggc cccgggtggaa agcacaacat 300
ggacttctac gtggggggcc tcgttcccg ggacacccgc ttccggggc tcattaccc 360
caccatctcc ctgtggaca cgttccaaacctt ggagctcccc gaggctgtgg tggttccaaga 420
cagcgtgttc ttccgtgtt ccggccgtt catgaccccc aacacccaggc cccggcagg 480
gggtgtacggc tggcgttggat ttgaaaatgt ggacttctgtt aagtctgtt ctactctggc 540
catgaaagcc aagtgtcaagc tgaccatctg cttgtggagg gagaacatgtt atgaccatgt 600
gatgcaggat gaaatggaga tcgggtatcat ccaaggcccc cacaacacgc tgccgtgtt 660
cttcgactct ccaaggaaaca gaggctgtt ggagtttccc atcaaaccggc tgatgggtcc 720
agatgttggc tatgttactc gaggggccca aacagggggtt atcgtggac ttgactcttt 780
tgggaacctg gaagtgtggcc ccccgatcatc agtgcaggggc aaggtatacc cgttggccag 840
gattcttc gggggacagct gttatcccgg caatgttccatc cttgtggaggat accaggccct 900
gcaggactt ctcgtgtcc acgagggttca ggccctgttgg aagcttattt ctgactgtt 960
gttccgtggc ctcgtgtggc agtctgttgc ctttggcgc gcaaccggaca ggaagggtt 1020
ccggctgttc ctggccaggcc ccaggctctg tcacaaacttgc ttccggggc agcagaatgt 1080
ggggccacggg gaggccctgc ttgttggcgg gatcaagaaa aaaaaacacgc agaaaaataaa 1140
gaacattctg tcaaacaaga cattgtggggca acataatttca ttgtggaga gatgtatcg 1200
cttggaaacctc gtcgtgttgc acgagggttgc gggccctggcc gagatgttca tcattgtt 1260
ccggcaggctc ttcaaggatc aaggttcttca taaggcggaa gttttttcc ccaacatgtt 1320
gaacatgttc gtgttggggc agcacttggg catcccaaaatcc ccccttgggc ccgttcatca 1380
ccggccgtgtc tgcgttggggc agaagggtgtt ttccctgtt gggccacttgc gcctccatgt 1440
caccttcatc aacgacttcttcc tcacccatca catcaggcat ggggggggttgc actggggcac 1500
caacgtgtgc gaaaggccct tcttccatca ttgtggggac atgggtccatc gagccatct 1560
tccctgggtt ccttccctc ctggccaggat gtcgtgtgggtt cctctgttgcgtt gttggcaagca 1620
agagcttctt gtaatattgt ggctccctgg gggccggccag ccctcccaatc agtggcttgc 1680
tttcttcctc tttgtatgttcc cagtttccatca ctctgttgcgtt cccaaatgtt ttcttagact 1740
gcacacttgc ttctgttgcata agaagctgtca ataaatgtt tttaatgttca ttgttacatc 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
agctctactc tacctcaccgc ggggtggaaat ctcttgcgc gacacatca cccgcacccg 60
caaaatgttca ccaaccatgc ctgttgcata tcacggatcc tggacctggg gcccctgtgg 120
acagggtggcc atctctgtgg tgaactgttgc cagagacat ctcgaatctt ctggccatgtt 180
cttcgtggatgttgc ttggcggccatc agacgttgcgac agacgttgcgac tggatgttgc 240
gagcacacgg acccccaagg acttcttcac aaaccatata ctgggttccatc acgtggccag 300
gttctgttgc gacaaatgttca ggggtttca gggccacacgc gttttttccatc gttgtgttgc 360
tccaaaggatc cttgtgttgc cttgtgttgc gttttttccatc gaccccaac acccaggcccc 420
ccggggatgttgc tttttttccatc gttttttccatc gttttttccatc gttttttccatc 480
ctctgttgc tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 540
accaggatgttgc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 600
ccgtgttgc tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 660
tgggttccatc tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 720
acttcgttgc tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 780
tgggttccatc tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 840
aggccctgtca gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 900
actgggttgc tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 960
agggttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 1020
agggttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 1080
aaataaaatgttca tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 1140
gcatgttgc tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 1200
tttgcgttgc tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 1260
acatgttgc tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 1320
tcatcaacgg cttgtgttgc tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 1380
tccaggatgttgc tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 1440
ggggccatca cttgtgttgc tttttttccatc gttttttccatc gttttttccatc gttttttccatc gttttttccatc 1500

```

-continued

cccatcttcc	ctggcgctct	ctcccttcgt	gccagatgtc	gctgggtcct	ctgcagtgtg	1560
gcaagcaaga	gotcttgta	atatttgtgc	tccctgggg	cggccacccc	tcccacagt	1620
ggtttgttt	cttcttcgt	gatgtccag	tttccccactc	tgaagatccc	aacatggcc	1680
tagcaactgca	cactcaagttc	tgctctaaga	agctgcaata	aagttttttt	aagtcaactt	1740
gtacatga						1748

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 X1, mRNA
source	1..2081
	mol_type = unassigned DNA
	organism = Homo sapiens

SEQUENCE: 69						
agccagaggc	acgagctagc	ccgacgatgg	cccagggcag	attgatccgt	gtgaccagg	60
ageagccac	ccatggcggt	tgtgtgtgg	gcacccgtac	tcagcttgcac	atctgcacgt	120
ctgccccgtg	gaaactgcacg	tccttcgtac	tcaacgcctc	cccagggtg	gtcgtggata	180
ttggccacgg	ccctccagcc	aagaagaaat	ccacagggtt	ctccacatgg	ccccgggacc	240
ctggggtaga	ggtagccctg	acatgtaaaag	tggccagtttgc	tagcagggc	gaccggagg	300
ttcagatttc	atactacggg	cccaagactc	caccaggtaa	agcttactc	tacccatccg	360
gggttggact	gcaggacatg	tcgctgtatg	ccctgagccac	gaagaccccc	aaggactt	420
tcacaaacca	tacactgggt	ctccacgtgg	ccagggtctg	gatggacaaa	gtgagggtt	480
ttcaggccac	acggggccaaa	ctgtcttca	agtgcacgt	agtcttgggt	cccaagttggc	540
cctctcaact	ctgtatggc	cccggtggaa	agcacaacat	ggacttctac	gtggggggcc	600
tcgctttccc	ggacaccggc	ttccggggg	tcattaccct	caccatccc	ctgtggaca	660
cgtccaaacct	ggagatcccc	gagggtgtgg	tgttccaaagg	cagcgtggc	ttccgggtgg	720
cgcctctggat	catggccccc	aacaccggc	ccccgggg	gggtgtacgc	tgcagtattt	780
ttgaaaatgt	ggacttctg	aagtcaatgt	ctactctggc	catgaaaaggc	aagtgcacgc	840
tgaccatctg	coctgaggag	gagaacatgg	atgaccatgg	gatgcaggat	gaaatggaga	900
tcggctcac	ccaaggcccc	cacaacaaacg	tgcccggtgg	cttgcactt	ccaaaggaaaca	960
gaggcctgaa	ggagttttcc	atcaaaacgg	tgtatgggtc	agatttggc	tatgttaactc	1020
gaggggccca	aacagggggt	atcaatggac	tggacttctt	tggggactc	gaatgtggcc	1080
ccccagtcac	agtcaaggggc	aaggaatacc	cgctggggcag	gattcttc	ggggacagct	1140
gttatcccg	caatgacagc	cgccatgtca	accaggccct	gcaggactc	ctcaatgtccc	1200
agcaggatgc	ggcccccgtg	aaatcttca	ctgacttggc	gtccgtggc	cacgtggacg	1260
agttccctgag	ctttgtgtca	gcacccgaca	ggaaagggtt	ccggctgtc	ctggccagcc	1320
ccaggccttg	ctacaaactg	ttccaggago	agcagaatga	gggcccacgg	gaggccctgc	1380
tgttcgaagg	gatcaagaaa	aaaaaacacg	agaaaataaa	gaacatttc	tcaaaacaaga	1440
cattgagaga	acataatca	tttggggaga	gatgcacgtca	ctggggacccg	gagctgtca	1500
agcggggact	ggccctggcc	gagagtaca	tcattgcacat	cccgcaactc	ttcaagctca	1560
aagagttctc	taaggcgaa	gctttttcc	ccaaacatgtt	gaacatgtg	gtgctaggga	1620
agcacccatgg	catccccaa	cccttcgggg	ccgtcatcaa	ccggccgtgc	tgccctggagg	1680
agaagggttg	ttccctgtgt	gagccactgg	gcctccatgg	tcacccatc	aacgactt	1740
tcacccatca	catcggat	ggggagggtc	actgcggcac	caacgtgcgc	agaaaaggcc	1800
tctcccttca	gtggggaaac	atgggtccct	gagcccat	tccttgggt	cctctccctc	1860
ctggccatgt	gtcgctgggt	ccctgtgtgt	gtggcaacag	agagcttctt	tgaatattgt	1920
ggtcccttgg	ggggggccag	ccctccatgg	atgggttgc	tttcttcc	tgtgtatgtcc	1980
cagtccatca	ctctgtatgt	cccaacatgg	tccttagact	gcacactcag	ttctgtctca	2040
agaaggtgtca	ataaaatgtt	tttaagtatgt	a			2081

SEQ ID NO: 70	moltype = AA length = 601
FEATURE	Location/Qualifiers
REGION	1..601
	note = protein-arginine deiminase type-4 isoform X1 [Homo sapiens]
source	1..601
	mol_type = protein
	organism = Homo sapiens

SEQUENCE: 70						
MAQGTLIRVT	PEQPTTHAVCV	LGTLTQLDIC	SSAPEDCTSF	SINASPGVVV	DIAHGPPAKK	60
KSTGSSTWPL	DPGVEVTLTM	KVASGSTGDQ	KVQISYYGPK	TPPVKALLYL	TGVDLQDMSL	120
MTLSTKTPKD	FPTNHTLVLU	VARSEMVKR	VFQATRKGKLS	SKCSVVLGPK	WPSHYLMVPG	180
GKHNMDFYVE	ALAFPDTDFF	GLITLTISLL	DTSNLELPFE	VVFQDSVVR	VAPWIMTPNT	240
QPQKEVYACS	IFENEDFLKS	VTTLAMKAKC	KLTICPEEEN	MDDQWMQDEM	EIGYIOPAPHK	300
TLPVVFDSPR	NRGLKBFPII	RVMGPDFGVY	TRGPQTGGGIS	GLDSFGNLEV	SPPVTVRGKE	360
YPLGRILFGD	SCYPSNDSRQ	MHQALQDFLS	AQQVQAPVKL	YSDWLSVGHV	DEFLSFPAP	420
DRKGFRLLA	SPRSCYKLFQ	EQQNENGHGEA	LLFEGIKKK	QOKIKNLNSN	KTLREHNSFV	480
ERCIDWNREL	LKRELGLAES	DIIDIPQLFK	LKEFSKAEEAF	FPNMVNMLVL	GKHLGIPKPF	540
GPVINGRCL	EEKVCSLLEP	LGLQCTFIND	FFTYPHIRHGE	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 ccagggtctga gatggacaaa gtgagggtgt 480
ttcaggccac acggggcaat ctgtatccca agtgcacgtc agtcttggtt cccaaagtggc 540
cctctcaactt cctgtatgtt cccgggtggaa agcacaacatgg ggttgcacatgttgc 600
tcgctttcccg ggacacccgc ttccccgggc tcatttaccc caccatcccttc ctgtgtggaca 660
cgtnccacccgtt ggagctccccc gagggtgtgg tggccatgttgc cggccgtggc ttcccggtgg 720
cgccctgtatgc catgatccccc aacacccccc ccccccggaa ggtgttgcgg tgcaatgtt 780
ttgaaaatgtt ggacttctgtt aagtcaatgtt caactctggc catgaaaggcc aagtgcacgc 840
tgaccatctt ccctgtggggaa gagaacatgg atgacccatgg gatggcggat gaaatggaga 900
tcggcttacat ccaagccccca cacaacacgc tggccgtggc ttccgtactt ccaaggaaaca 960
gaggccgttggaa ggaggttccccccatcaacacgc tggccgtggc agattttgc tatgtactt 1020
gaggggcccccac acgggggggtt atcgatgggg tggactcttctt tggggacccgtt gaaatggggcc 1080
ccccccatgtt acgttccggc aagaaatacc cggccgtggc acgttcccttc gggggacccgtt 1140
gttatccctgtt caatgtacacgc cggccgtggc accaggccctt gggggacccgtt ctcgtggccc 1200
agcagggtgtt gggccctgtt aagtttctt ctgactgtggc gtccgtggc cacgtggacg 1260
agtccctgtt ctggccgtggc gcaacccgc ggggggggtt cccgtgtgtc ctggccacgc 1320
ccaggccgtt ctggccgtggc tttccggggc acggccatgtt gggccacccggg gaggcccttc 1380
tggtcaagg gatcaagaa aaaaaacacgc aagaaataaa aaacattctt tcaaaacaaga 1440
cattggaga acataattca tttgtggaga gatgcacatgg tcggccatgc gggccatgtt 1500
agcggggatgtt gggccgtggc gggccgtggc gggccatgtt gatgcacatgg tcggccatgc 1560
aagagtttccatggccaa gttttttccatggc gcaacatgtt gggccatgtt 1620
agcacctggg catcccccaag cccttcggggc cccgtcatcaaa cggccgtgtc tgccctggagg 1680
agaaggtgtt tttccctgtt gagccactgg gcctccatgtt caccatctt aacgttctt 1740
tacccatccatccatggccatgtt gggccgtggc acgttccggc acacgttgcgc agaaaacccct 1800
tctcccttccatggccatgtt gggccgtggc acgttccggc acacgttgcgc agaaaacccct 1860
ctggccatgtt gggccgtggc acgttccggc acacgttgcgc agaaaacccct 1920
ggcccccgtt gggccggccag ccctccatggccatgtt gggccgtggc acgttccggc acacgttgcgc 1980
caatggccatgtt gggccgtggc acgttccggc acacgttgcgc agaaaacccct 2040
agaaggtgtt ataaatgttccatggccaa gttttttccatggc gcaacatgtt gggccatgtt 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 EIGYIQAQPHK 300
TLPVVFDFSPR NRGFLKEFPIK RVMPGDPFGYV TRGPQTTGGIS GLDSFGNLLEV SPPVTVRGKE 360
YPLGRILFGD SCYPSNDSRQ MHQALQDFLS 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 gcccacatgg gggccatgggg acattgtatcc gtgtgacccc 60
agagcagcccaatccatgtt gggccatgtt acgttgcgtt acatctgtt 120
ctctggccctt gggccatgtt gggccatgtt acgttgcgtt acatctgtt 180
tattggccatgtt gggccatgtt acgttgcgtt acatctgtt 240
ccctggggta gggccatgtt gggccatgtt acgttgcgtt acatctgtt 300

```

-continued

ggttcagatt	tcatactacg	gacccaagac	tccaccagtc	aaagctctac	tctacacctac	360
cgggggtggaa	atctccttg	gcccggacat	cacccgcacc	ggcaaaagtga	agccaaaccag	420
agctgtgaaa	gatcagagga	cctggacctg	ggggcccttgc	ggcacagggt	ccatctctgc	480
ggtgaactgt	gacagagaca	atctcgatc	tctgccatg	gactgcgggg	atgatgaagt	540
gcttgacagc	gaagacactgc	aggacatgtc	gctgtatgacc	ctgagcaca	agacccccaa	600
ggacttcttc	acaaaccata	cactggtgt	ccacgtggcc	aggtctgaga	tggacaaga	660
gagggtgttt	caggccacac	ggggcaaaact	gtcctccaag	tgcagegtag	tcttgggttcc	720
caagtggccc	tccactacc	tgtatggccc	cgttggaaag	cacaacatgg	acttctacgt	780
ggaggccctc	gctttcccg	acaccgactt	ccgggggctc	attacectca	ccatctccct	840
gtctggacacg	tccaaacctgg	agtcctccca	ggctgtgttt	ttccaaagaca	gcgttgctt	900
ccgcgtggcg	ccctggatca	tgcacccca	caccaggccc	ccgcaggagg	tgtacgcgt	960
cagtattttt	gaaaatgagg	acttcctgaa	gtcagtgtact	actctggcca	tgaaagccaa	1020
gtgcaagctg	accatctgc	ctgaggagg	gaacatggat	gaccagtgg	tgcaggatga	1080
aatggagatc	ggctcatatc	aagccccaca	caaaacgcgtt	ccctgtgtt	tcgactctcc	1140
aaggaacaga	ggccctgagg	agttttccat	caaacgcgtt	atgggcttcc	ggctgtctt	1200
ggccagcccc	aggctctgt	acaaactgtt	ccaggacac	cagaatgggg	gccacgggaa	1260
ggccctgtct	ttcgaaggga	tcaagaaaa	aaaacagcac	aaaataaga	acattctgc	1320
aaacaagaca	tttagagaca	ataattttt	tgtggagaga	tgcactcgact	ggAACCCGCGA	1380
gctgtctgg	cggggatctgg	gcctggccga	gagtgcacat	attgcacatcc	cgccagcttt	1440
caagctcaa	gagtttctta	aggccgaa	tttttttccca	aacatgttga	acatgtctgt	1500
gctagggaa	cacccggcca	tcccaaaagcc	cttggggccc	gtcataac	gccgtgtct	1560
cctggaggag	aagggtgttt	ccctgcttgc	gcctactggc	ctccactatca	1620	
cgacttcttc	acacttaccaca	tcaggcatgg	ggagggtgcac	tgcggcacca	acgtgcgcag	1680
aaagcccttc	tccttcactgt	ggtggaaat	ggtgccttgc	ggccatcttc	cctggcgttcc	1740
tctcccttct	ggccagatgt	cgctgggtt	tctgcactgt	ggcaagcaag	agcttttgtt	1800
aatattgtgg	ctccctgggg	ggggccagcc	ctcccaagca	tggcttgc	tcttcctctg	1860
tgtatgttca	gtttccact	ctgaaatccc	caacatggc	ctagcaactgc	acactcagtt	1920
ctgctctaag	aagttttt	aaagttttt	taagtcaat	tgtatcat		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
 MAQGTLIRVT PEPPTHAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60
 KSTGSSTWPL DPGVEVTLTM KVAGSSTGDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
 ITRTGKVKPT RAVKDQRTWT WGPGCQGAIL LVNCDRDNLE SSAMDCEDE VLDSEDLQDM 180
 SLMLTSLTKTP KDFFTNHTLV LHVARSEMOK VRVFQATRGK LSSKCSVVLG PKWPSPHYLMV 240
 PGGKHNMDFY VEALAFFPDTD FPGLITLTIS LLDTSNLELP EAVVFQDSVV FRVAPWIMTP 300
 NTQPQPEVYA CSIFENEDFL KSVTTLAMKA KCKLTICPEE ENMDDQWMQD EMEIGYIQAP 360
 HKTLPVVFDS PRNRLKEFP IKRVMGFRL LASPRSCYKL FQEQQNEGHG EALLFEGIKK 420
 KKQQKIKNL SNKTLREHNS FVERCIDWNR ELLKRELGLA ESDIIDIPQL FKLKEFSKAE 480
 AFFFPMVNMVL VLGKHLGIPI PFGFVINGRC CLEEKVCSLL EPLGLQCTFI NDFFTYHIRH 540
 GEVHCGTNVR RKPFSFKWWN 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
 acagccagac ggacgagact gcccgacat gggccagggg acatttgcattcc gtgtgacccc 60
 agaggcggcc accccatggc tggtgtgtt gggcaccttg actcagcttgc acatctgcac 120
 ctctggccct gaggacttgc cgtcttcatc catcaacggcc tcccccagggg ttggctgttga 180
 tattggccac ggccctccag ccaagaagaa atccacaggat ttctccacat ggccttgc 240
 ccctggggta gaggtgaccc tgacgtatc gttggccatgt ggtagcacag ggcggccat 300
 gggtcagatt tcatactacg gacccaaagac tccaccagtc aaagcttgcac tctacttgc 360
 cgggggtggaa atctccctgt ggcgcacat cacccggcacc ggcggaaatgtga agccaaaccag 420
 agctgtgaaa gatcagagga ctttgcaccc tggggcccttg ggacagggtt ccatctctgc 480
 ggtgaactgt gacagagaca atctcgatc ttctgccatg gactgcgggg atgatgaagt 540
 gcttgacagc gaagacactgc aggacatgtc gctgtatgacc ctgagcaca agacccccaa 600
 ggacttcttc acaaaccata cactggtgtt ccacgtggcc aggtctgaga tggacaaga 660
 gagggtgttt caggccacac gggggcaaaact gtccctccaag tgcagegtag tcttgggttcc 720
 caagtggccc tccactacc tttatggccc cgttggaaag cacaacatgg acttctacgt 780
 ggaggccctc gctttcccg acaccgactt ccgggggctt attacectca ccatctccct 840
 gcttggacacg tccaaacctgg agctcccaaa ggctgtgttt ttccaaagaca gcgttgctt 900
 ccgcgtggcg ccctggatca tgaccccaaa cacccggccccc ccgcaggagg tggacgcgt 960
 cagtattttt gaaaatgagg acttcctgaa gtcagtgtact actctggcca tggaaagccaa 1020
 gtgcgaagctg accatctgcc ctgaggaggaa gaacatggat gaccaggccat tgcaaggat 1080

-continued

```

aatggagatc ggctacatcc aagccccaca caaaacgctg cccgtggct tcgactctcc 1140
aaggaacaga ggcctgagg agtttcccat caaacgcgtg atgggttcc ggctgtccct 1200
ggccagcccc aggtctcgct acaaactgtt ccaggagca cagaatgggg gccacgggga 1260
ggccctgtgt ttcgaaggaa tcaagaaaaa aaaacacgca aaaaataaga acatctgtc 1320
aaacaagaca ttgagagaac ataattcatt tggagaga tgcacgact ggaaccgcga 1380
gtgtgtcaag cgggagctgg gcctggccca gagtgcacat attgcacatcc cgcaactt 1440
caagctcaaa gagttctcta agggggaaat tttttccccc aacatggta acatgttgt 1500
gttagggaaat caccggca tcccaagcc ctccggggcc gtcatcaacg gccgtgtgt 1560
cctggaggaaat aagggtgttt ccgtgtgg gcaactgggg ctccagtgtc cttcatcaa 1620
cgacttcc acctacccat tcggcatgg ggaggtgcac tgccggcacca acgtgcgcag 1680
aaagcccttc tccctcaat ggtggatctt ggtggccatgg gcccatttcc ctgggtgtc 1740
tcccttcctt gggcagatgt cgctgggtcc tctgcgtgtt ggcaagcaag agcttttgt 1800
aatattgtgg ctccctgggg gggccagcc ctccagca gttggcttctt tctctctgt 1860
tgatgtccca gttttccact ctgaagatcc caacatggtc ctgcactgc acactcgtt 1920
ctgtctcaag aagctgcaat aaagttttt taagtcaat tgcacatgt 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
MAAGTLLIRVT PEQPHTAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60
KSTGSSSTWPL DPGVEVTLM KVASGSTDQ KVQISYYGPK TPPVKALLYL TGVEISLCD 120
ITRTGKVKPT RAVKDQRTWT WGPGCQGAIL LVNCDRDNLE SSAMDCEDE VLDSEDLQDM 180
SMLTSLTKTP KDFFTNHTLV LHVARSEMDK VRVFQATRGK LSSKCSVVLG PKWPSHYLMV 240
PGGKHNNDFY VEALAFPDTD FPGLITLTIS LLDTSNLELP EAVVFQDSVV FRVAPWIMTP 300
NTQPPQEYVA CSIFENEDFL KSVTTLAMKA KCKLTICPEE ENMDDQWMQD EMEIGYIQAP 360
HKTLPVVPDS PRNRLKEFP IKRVMGFLRLL LASPRSCYKL FQEQQNEGHG EALLFEGIKK 420
KKQQKIKNL SNKTLREHNS FVERCIDWNR ELLKRELGLA ESDIIDIPQL FKLKEFSKAE 480
AFFPNMVNML VLGKHLGIPK PFGPVPINGRC CLEEKVCSSL EPLGLQCTFI NDFFTYHIRH 540
GEVHCGTNVR RKPFSFKWWN 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

```

```

SEQUENCE: 77
acagccagag ggacgagcta gcccgcacat gggccagggg acattgtatcc gtgtgacccc 60
agagcagccc acccatgccg tggtgtgtt gggcacccctt actcagatgg acatctgcgg 120
ctctggccctt gggactgcg cgtcttcacat catcaacgcg tccccagggg tggtgttgg 180
tattggccac gggccctccag ccaagaagaa atccacatgg tcttcacat gggcccttgg 240
ccctggggta gagggtgaccc tgacatggaa agtggccatgg ggttagcacag ggcaccaaga 300
ggtgtcaggatt tcaatactcg gacccaaatcg tccaccaggatc aaagcttacatc tctacccatc 360
cgggggtggaa atcttcattgt ggcgcacat caccgcaccc ggcacaaatgtg agccaaccac 420
agctgtgaaa gatcaagatc cttggacatgg qggcccttgg ggcacagggtt ccatctgt 480
gggtgaaatgt gacagagaca atctcgaaatc ttctggcatg gactgcgggg atgatgaatg 540
gttgcacacg gaagacatgc aggacatgtc gctgtatggcc ctggacacca agaccccaaa 600
ggacttcttc acaaaatcgca cactggatgtt ccacgtggcc aggtctgtgaa tggacaaatgt 660
gagggtgtttt cggccacac gggccaaatctt gtccttcaatg tgcagcgtatgg tctttgggtcc 720
caagtggccc tctcaatcc ttgtgttccca cggtggaaatg cacaacatgg acttctatgt 780
ggaggccctt gttttcccg acacccgactt cccggggctt attacccatc ccatctccct 840
gtctggacacg tccaaatcgatgg agtccccccg ggtctgtgggtt ttccaaagaca ggcgtggctt 900
cccgctggcc ccctggatgtt tgaccccaaaatc caccacggcc cccgcgggggg tgcacgtgt 960
cagtatattttt gaaaatgggg acttcttgc gtcagtgactt actctggccca tggaaatgt 1020
gtgcacatcg accatctgcgatcc ctggaggatggaa gacatggatgg gaccaggatgg 1080
aatggagatc ggctacatcc aagccccaca caaaacgctg cccgtggct tcgactctcc 1140
aaggaacaga ggcctgagg agtttcccat caaacgcgtt atgattctat caataggcccc 1200
attttatata agaagaaatctt gggctcagaatgttggagatgtt tggaaatgtt accccgcgcgt 1260
tcggcgtttt gtcccgatgtt tggctatgtt actcgaggggc cccaaacagg gggatgtt 1320
ggactggact cttttggaaatc ctttggatgtt agccccccc 1358

```

```

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

organism = Homo sapiens

SEQUENCE: 78

MAQGTLIRVT PEQPHTAVCV LGLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPAAKK 60
 KSTGSSTWPL DPGVEVTLTM KVASGSTGDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
 ITRTGKVKPRT RAVKDQRTWT WGPGCGQAIL LVNCDRDNLE SSAMDCEDE VLDSEDLQDM 180
 SLMLTSLTKTP KDFFTNHTLV LHVARSEMDK VRVFQATRGK LSSKCSVVLG PKWPSHYLMV 240
 PGGKHNMDFY VEALAFPDTD FPGLITLTIS LLDTSNLIELP EAVVFQDSVV FRVAPWIMTP 300
 NTQQPQEYVA CSIFENEDFL KSVTTLAMKA KCKLTIPEE ENMDDQWMQD EMEIGYIQAP 360
 HKTLPPVVFDS PRNRGLKEFP IKRVMILSIG PFYRRRN 397

SEQ ID NO: 79

moltype = DNA length = 1358
 Location/Qualifiers
 misc_feature 1..1358
 note = PREDICTED: Homo sapiens peptidyl arginine deiminase type-4 (PADI4), transcript variant X4, mRNA
 source 1..1358
 mol_type = unassigned DNA
 organism = Homo sapiens

SEQUENCE: 79

acagccagag ggacgagcta gcccgaatgggg acattgtatcc gtgtgacccc 60
 agagcagccc acccatgcgg tttgtgtgtctt gggcacccctt actcagcttg acatctgcag 120
 ctctgccttc gaggactgca cgtccttcag catcaacggcc tccccaggggg ttgtcggtta 180
 tattgccccac gggccctccag ccaagaagaa atccacaggt tcctccacat ggccctctgg 240
 ccctggggta gagggtgacc tgacatgaa atgtggccagt ggttagcacag ggcgacccagaa 300
 ggttcagatt tcatactacg gaccggacat tccacccatc aaagctctac ttcacctcaca 360
 cgggggtggaa atctccttgtt ggcgcagacat caccggcacc ggcggaaatgtaa agccaaaccag 420
 agctgtgaaa gatcagaggg cctggacccctt gggcccttgtt ggacagggtt ccatctgtct 480
 ggttcagactt gacacgacatc attcgcatac ttctgcacat gacttcggggg atgtatgtaa 540
 gtttcacacgg gaagacccatc aggcacatgtc getgtatggcc ctgacccacgaa agaccccca 600
 ggacttccttc acacccatcacttggccat ccaacgtggcc aggtttcgatgg tggaccaaatg 660
 gaggggtgttt caggccacac gggggcaaaact gtccctccaaatg tgcacgttgcgtatc tcttgggtcc 720
 caaagtggccc tcttcactactt tgatgggtttt cgggtggaaatg cacaacatgg acttccttcgt 780
 gggggccctt gttttccccc acacccggatcc cccggggctt attaccctca ccatctccct 840
 gtgttcacacgg tccaaaccttgg agtcctccggaa ggctgtgtttt ttcaaaagacaa gctgtgtttt 900
 cccgtggccg cccctggatca tgaccccaatgg caccggccccc ccggcaggagg tttttttttt 960
 cagttttttt gaaaatggggg acttcctgtt gtcacgttactt acttcggccca tgaaagccaa 1020
 gtgttcacacgg accatctgtt cttggggggaa gaaatggggatgg gaccatgttgg tttttttttt 1080
 atttggggatgg ggttacatccaa aagggcccaatgg ccaacgtgtt ccctgtgtttt tttttttttt 1140
 aagggacacgg ggcctggatgg agttttccatc ccaacgtgtt atgatccatcaatggccca 1200
 atttttataga agaagaaactt gaggctcaga gtggagaatgtt tggaaatcaa accccggccatgg 1260
 tccgggtttt gtcacatgtt acttcggggcccccacccacgggg ggttacatgtt 1320
 ggacttcgttcccttggggaa cttttggggaaatgg gggggcccccacccacgggg ggttacatgtt 1388

SEQ ID NO: 80

moltype = AA length = 397
 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 PEQPHTAVCV LGLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPAAKK 60
 KSTGSSTWPL DPGVEVTLTM KVASGSTGDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120
 ITRTGKVKPRT RAVKDQRTWT WGPGCGQAIL LVNCDRDNLE SSAMDCEDE VLDSEDLQDM 180
 SLMLTSLTKTP KDFFTNHTLV LHVARSEMDK VRVFQATRGK LSSKCSVVLG PKWPSHYLMV 240
 PGGKHNMDFY VEALAFPDTD FPGLITLTIS LLDTSNLIELP EAVVFQDSVV FRVAPWIMTP 300
 NTQQPQEYVA CSIFENEDFL KSVTTLAMKA KCKLTIPEE ENMDDQWMQD EMEIGYIQAP 360
 HKTLPPVVFDS PRNRGLKEFP IKRVMILSIG PFYRRRN 397

SEQ ID NO: 81

moltype = DNA length = 1842
 Location/Qualifiers
 misc_feature 1..1842
 note = PREDICTED: Homo sapiens peptidyl arginine deiminase type-4 (PADI4), transcript variant X6, mRNA
 source 1..1842
 mol_type = unassigned DNA
 organism = Homo sapiens

SEQUENCE: 81

acagccagag ggacgagcta gcccgaatgggg acattgtatcc gtgtgacccc 60
 agagcagccc acccatgcgg tttgtgtgtctt gggcacccctt actcagcttg acatctgcag 120
 ctctgccttc gaggactgca cgtccttcag catcaacggcc tccccaggggg ttgtcggtta 180
 tattgccccac gggccctccag ccaagaagaa atccacaggt tcctccacat ggccctctgg 240
 ccctggggta gagggtgacc tgacatgaa atgtggccagt ggttagcacag ggcgacccagaa 300
 ggttcagatt tcatactacg gaccggacat tccacccatc aaagctctac ttcacctcaca 360
 cgggggtggaa atctccttgtt ggcgcagacat caccggcacc ggcggaaatgtaa agccaaaccag 420

-continued

actgtgtaaa	gatcagagga	cctggacctg	gggccttgc	ggacagggtg	ccatcctgct	480
gggtgaactgt	gacagagaca	atctcgaaatc	tcttgccatg	gactgcgggg	atgtatgtaa	540
gcttgacagc	gaagacactgc	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	ctgaggagga	gaacatgtat	gaccgttgg	tgcaggatga	1080
aatggagatc	ggctacatcc	aagcccccaca	caaaacgcgt	cccggtgtt	tcgactctcc	1140
aaggaacaga	ggcctgaagg	agtttcccat	caaaccgcgt	atggaaagcc	ccaggccgtc	1200
gctctgagtc	ctggcagccc	cttggctcaa	ggtgccccagg	gaaaacgcacg	gaaggaggt	1260
gggggttcaag	gtttgtttagg	acaaatggct	ttccgacaaag	catctgtact	gtgccccaaag	1320
ggaggggaga	gtggccctgc	ccacacacaa	ggcttcacaa	tgaaaactca	gttaaactca	1380
acaacttgc	ccgagtctcg	cttctctctg	gccaagctgt	gcgcgtggca	gtggagatgc	1440
agagatgtgt	ttggattcac	agtgtgtact	cagccatggg	gtcgagttca	tcctgtttca	1500
gcagcgttgg	ggcgaacccag	ccagctgtca	ccgtgacttt	gccttcttcc	tgctgtttc	1560
agtccctgtc	gtttcgttgc	ctgtgttcaat	tttgcgttca	caagataaac	aaaggccaga	1620
ttttctgtat	ccacaagttcc	tgctggacag	acagacggac	acctcgccac	cattttgttc	1680
cccttccca	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
 SMLTSLTKTP 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 atccacatgg tcctccacat ggccccttgg 240
 ccctgggtt gagggttgcacc tgacatgtt ggtggccatg ggttagcacag ggcaccagaa 300
 ggttcagatt tcataactacg gacccaaacat tccaccatgcg aaagcttac tctaccctac 360
 cgggggttggaa atctcttgcg ggcgcacat caccgcgcac ggccaaatgtt agccaaccag 420
 agtgtgtaaa gatcagagga cctggacatc gggcccttgc ggacagggtg ccatcctgt 480
 ggtgtactgt gacagacaca atctcgatcc ttctgcgcatt 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 ctgaggagga gaacatgtt gaccgttgg tgcaggatga 1080
 aatggagatc ggttatcatcc aagccccaca caaaacgcgtc cccgtgggtt tcgactctcc 1140
 aaggaacaga ggcctgttgg agtttccat caaaacgcgtt atggaaagcc ccaggccgtc 1200
 gctctgagtc ctggcagccc ctggcttca ggtgttccagg 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

gcagcgctgg ggccgaaccag ccagctgcta ccgtgacttt gcccctttc tgctgcttc	1560
agtccctgctt ggttcagttt ctgtgtctac ttgtatgtcc caagataaac aaaggccaga	1620
ttttctgttg ccacaatgtcc tgctggacag acagacggac acctccggcac cattttgttc	1680
ccctttccca cagacaactt tggatgaagg caggcagcac tggccctca gctaaagcaca	1740
tctgtgtatt ccaactcaaa tcactcagaa tgaggccagg tgggtggct gacacctgta	1800
atccccagcag ttggggaggc caaggaggat ggatcacctg ag	1842
 SEQ ID NO: 84 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: 84	
MAQGTLIRVT PEQPTHAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60	
KSTGSSTWPL DPGVEVTLTM KVAGSGTDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120	
IITRTGKVKPT RAVKDQRTWT WGPGCQGAIL LVNCDRDNLE SSAMDCEDE VLDSEDLQDM 180	
SLMLTSLTKP KDFFTNHTLV LHVARSEMDK VRVFQATRGK LSSKCSVVLG PKWPSPHYLMV 240	
PGGKHNMDFY VEALAFPDTD FPGLITLTIS LLDTSNLELP EAVVFQDSVV FRVAPWIMTP 300	
NTQOPPQEYVA CSIFENEDFL KSVTTLAMKA KCKLTICPEE ENMDDQWMQD EMEIGYIQAP 360	
HKTLPVVFDS PRNRGLKEFP IKRVMEDPRP AL 392	
 SEQ ID NO: 85 moltype = DNA length = 1275	
FEATURE Location/Qualifiers	
misc_feature 1..1275	
note = PREDICTED: Homo sapiens peptidyl arginine deiminase 4 (PADI4), transcript variant X7, mRNA	
source 1..1275	
mol_type = unassigned DNA	
organism = Homo sapiens	
 SEQUENCE: 85	
acagccagag ggacgagcta gcccgcacat gggccagggg acattgtatcc gtgtgacccc 60	
agagcagccg acccatgcgc tggatgtgtt gggcacccctt actcagttt acatctgcag 120	
ctctggccctt gggactgcg cgtcccttag catcaacgcgc tccccagggg tggtcgttga 180	
tatttgcggcc gggccctccag ccaaaagaaa atccacaggat tcctccacat gggccctgg 240	
ccctggggta ggggtgaccc tgacgatgaa agtggccagt ggttagcacag ggcacccaga 300	
gggttcagatt tctatccatc gacccaaatgc tccaccaggat aaagcttac tctacccatc 360	
cgggggtggaa atctccctgt ggcgcacat caccgcgc ggcacaaatgtga agccaaacccag 420	
agctgtgaaa gatcaagatc cttggacccg gggcccttgc ggacagggtt ccatctgttgc 480	
gggtgaactgt gacagagaca atctcgaatc ttctggccatg gactgcgggg atgtatgttgc 540	
gttttgcacgc ggaaatgcgc gggatgtttt gctgtatgttgc ctggacacgc agaccccaa 600	
ggacttcttc aaaaaatccatc cacttgcgtt ccacgtggcc aggtctgaga tggacaaatgt 660	
gagggtgtttt caggccacac gggccaaatctt gtcttccaaatg tgcagccgttgc tcttgggtcc 720	
caagtggccc ttcactaccatc tgatggtccc cggtggaaatg cacaacatgg acttctacgt 780	
ggaggcccttc gttttcccgatc acacccactt cccggggcttc attacccatc ccatctccct 840	
gtcttgcacgc tccaaatccatc agtccatccatc ggttgcgttgc ttccaaatgc gctgtgttgc 900	
ccggcgtggcc ccttggatca tgaccccaatc caccgcgc ccgcgggggg tgcacccgttgc 960	
cagtatccatc gaaaatgtttt gttttccatc acacccatccatc gttttccatc tgaaagccaa 1020	
gtgtcaagcttccatc accatccatc ctggggatca gaaatgtttt gttttccatc tgaaagccaa 1080	
aatggatcatc ggttacatccatc aacccatccatc gttttccatc cccggggcttc tgcactctcc 1140	
aaggaaacaga ggccttgcaggatc tttttccatc caaaatgcgttccatc tgaaatgc gttttccatc 1200	
tgcaccaggcc ccttggatca tttttccatc caaaatgcgttccatc tgaaatgc gttttccatc 1260	
attctgtatc gttttccatc 1275	
 SEQ ID NO: 86 moltype = AA length = 386	
FEATURE Location/Qualifiers	
REGION 1..386	
note = protein-arginine deiminase type-4 isoform X6 [Homo sapiens]	
source 1..386	
mol_type = protein	
organism = Homo sapiens	
 SEQUENCE: 86	
MAQGTLIRVT PEQPTHAVCV LGTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60	
KSTGSSTWPL DPGVEVTLTM KVAGSGTDQ KVQISYYGPK TPPVKALLYL TGVEISLCAD 120	
IITRTGKVKPT RAVKDQRTWT WGPGCQGAIL LVNCDRDNLE SSAMDCEDE VLDSEDLQDM 180	
SLMLTSLTKP KDFFTNHTLV LHVARSEMDK VRVFQATRGK LSSKCSVVLG PKWPSPHYLMV 240	
PGGKHNMDFY VEALAFPDTD FPGLITLTIS LLDTSNLELP EAVVFQDSVV FRVAPWIMTP 300	
NTQOPPQEYVA CSIFENEDFL KSVTTLAMKA KCKLTICPEE ENMDDQWMQD EMEIGYIQAP 360	
HKTLPVVFDS PRNRGLKEFP IKRVMEDPRP AL 386	
 SEQ ID NO: 87 moltype = DNA length = 1275	
FEATURE Location/Qualifiers	
misc_feature 1..1275	

-continued

```

note = PREDICTED: Homo sapiens peptidyl arginine deiminase
        4 (PADI4), transcript variant X7, mRNA
source      1..1275
mol_type = unassigned DNA
organism = Homo sapiens
SEQUENCE: 87
acagccagag ggacgagcta gcccacat ggcggggg acattgtatc gtgtgacccc 60
agagcagccc acccatgccc tttgtgtgt gggcacctt actcagatc acatctcgag 120
ctctgcctt caggactgca cgtcttcag catcaacgcc tccccgggg ttgtcggtga 180
tattggccac ggcctccac ccaagaagaa atccacatc ttctccat ggcctcgga 240
ccctgggata gagggtaccc tgacatgaa agtggccatg ggttagcaca ggcggggaa 300
ggttcagatt tcatactacg gacccaaagac tccaccatgc aaagcttac tctacccatc 360
cgggggtggaa atctccttgt ggcggacat caccggcacc ggcaaaatgtga agccaaccag 420
agctgtgaaa gatcagagga cctggacatg gggcccttg ggacagggtg ccatctcgat 480
ggttgaactgt gagacagaca atctcgaatc ttctggccatg gactgcgggg atgtgaatg 540
ggcttgacatc ggacacctgc aggcacatgc gctgtatggc ctgagcaca agaccccaa 600
ggacttcttc acaaaccata cactgggtgtt ccacgtggcc aggtctgtg tggacaatgt 660
gagggtgttt caggccacac gggccaaact gtcttccaaatg tgcagcgatg tcttgggtcc 720
caagtggccc tctcactacc tgatgttcccg cgggtggaaatg cacaacatgg acttctacgt 780
ggaggccctc gtttcccg gacccgactt cccggggctc attacccatc ccatctccat 840
gctggacacg tccaaacctgg agtccccca ggtggccatg ttccaaagaca ggcgggttt 900
ccgcgtggcc ccctggatca tgaccccaatc caccggggcc cccggggatg tgcggcgatg 960
cagtatgtt gaaaatgagg acttctgtt gtcagtgact actctggccca tgaaagccaa 1020
gtgcaagctg accatctgcc ctggggggaaatg gacatgtt gaccgttgg tgcaggatgt 1080
aatggagatc ggctacatcc aagccccaca caaaacgtt ggcgtggctc tcgactctcc 1140
aaggaacaga ggcctgaaatg agttccat caaaacgtt gtcggcgatg atgcaatgac agccggcaga 1200
tgcggccatc ctggcaggac ttcttcgtt cccggggatg gacggccatc gtgaagctt 1260
attctgactg gctgtt 1275

SEQ ID NO: 88      moltype = AA length = 386
FEATURE          Location/Qualifiers
REGION           1..386
note = protein-arginine deiminase type-4 isoform X6 [Homo
            sapiens]
source      1..386
mol_type = protein
organism = Homo sapiens
SEQUENCE: 88
MAQGTLIRVT PEQPTHAVCV LGTTLTQLDIC SSAPEDCTSF SINASPGVVV DIAHGPPAKK 60
KSTGSSTWPL DPGVEVTLTM KVAGSSTGDQ KVQISYYGPQ TPPVKALLYL TGVEISLCAD 120
IITRTGKVKPT RAVKDQRTWT WGPGCGQAIL LVNCDRDNLE SSAMDCEDE VLDSEDLQDM 180
SLMLTSLTKTP KDFFTNHTLV LHVARSEMOK VRVFQATRGK LSSKCSVVLG PKWPSPHYLMV 240
PGGKHNMDFY VEALAPPDTD FPGLITLTIS LLDTSNLELP EAVVFQDSVV FRVAPWIMTP 300
NTQPPQEYVA CSIFENEDFL KSVTALMKA KCKLATICPEE ENMDQWMQD EMEIGYIQAP 360
HKTLPVVFDS PRNRGLKEFP IKRVMQ 386

SEQ ID NO: 89      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: 89
acagccagag ggacgagcta gcccacat ggcggggg acattgtatc gtgtgacccc 60
agagcagccc acccatgccc tttgtgtgt gggcacctt actcagatc acatctcgag 120
ctctgcctt caggactgca cgtcttcag catcaacgcc tccccgggg ttgtcggtga 180
tattggccac ggcctccac ccaagaagaa atccacatc ttctccat ggcctcgga 240
ccctgggata gagggtaccc tgacatgaa agtggccatg ggttagcaca ggcggggaa 300
ggttcagatt tcatactacg gacccaaagac tccaccatgc aaagcttac tctacccatc 360
cgggggtggaa atctccttgt ggcggacat caccggcacc ggcaaaatgtga agccaaccag 420
agctgtgaaa gatcagagga cctggacatg gggcccttg ggacagggtg ccatctcgat 480
ggttgaactgt gagacagaca atctcgaatc ttctggccatg gactgcgggg atgtgaatg 540
ggcttgacatc ggacacctgc aggcacatgc gctgtatggc ctgagcaca agaccccaa 600
ggacttcttc acaaaccata cactgggtgtt ccacgtggcc aggtctgtg tggacaatgt 660
gagggtgttt caggccacac gggccaaact gtcttccaaatg tgcagcgatg tcttgggtcc 720
caagtggccc tctcactacc tgatgttcccg cgggtggaaatg cacaacatgg acttctacgt 780
ggaggccctc gtttcccg gacccgactt cccggggctc attacccatc ccatctccat 840
gctggacacg tccaaacctgg agtccccca ggtggccatg ttccaaagaca ggcgggttt 900
ccgcgtggcc ccctggatca tgaccccaatc caccggggcc cccggggatg tgcggcgatg 960
cagtatgtt gaaaatgagg acttctgtt gtcagtgact actctggccca tgaaagccaa 1020
gtgcaagctg accatctgcc ctggggggaaatg gacatgtt gaccgttgg tgcaggatgt 1080
aatggagatc ggctacatcc aagccccaca caaaacgtt ggcgtggctc tcgactctcc 1140
aaggaacaga ggcctgaaatg agttccat caaaacgtt gtcggcgatg atgattctat caataggccc 1200
attttataga aagaaactt gagggtcaga gtggagaatg tggaaatcaa accccggccag 1260

```

-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
ctctggccctt gaggactgca cgtcccttag catcaacgcg tccccagggg ttgtcggtta 180
tattggccac ggcgcctccg ccaagaagaa atccacatgt ttctccatat ggcgcctgg 240
ccctgggtaa gagggtaccc tgacatggaa agtggccatg ggttagcacag ggcaccagaa 300
ggttcagattt tcataactacg gacccaaagac tccaccatgc aaagcttac tctaccatc 360
cggggtggaa atctccatgt gcgcagacat caccgcaccc ggccaaagtga agccaaccag 420
agctgtgaaa gatcagagga cctggacatg gggcccttgtt ggacagggtt ccattctgt 480
gttgaactt gacagagaca atctcgaatc ttctggccatg gactggcagg atgtgaatg 540
gcttgacatc gggacatgc acatgtgc gctgtatggcc ctgacacaa agaccccaa 600
ggacttccatc acaaaccata cactgggtt ccacgtggcc aggtctgaga tggacaatg 660
gagggtgtttt caggccacac gggccaaactt gtctccaaatg tgccagcgatg tcttgggtcc 720
caagtggccc tctcaatacc tgatggtccc cggtgaaatg cacaacatgg acttctacgt 780
ggaggccctt gttttcccg acacccactt cccggggctt attaccctca ccattccct 840
gctggacacg tccaaacctt gacttcccgaa ggttgggtt ttccaaagaca gggtgggttt 900
ccgcgtggcg ccctggatca tgaccccaatc caccacggcc ccgcggagg ttgtacggcg 960
cgtatattttt gaaaatgagg atctccatgg gtcagtgtact actctggcca tgaaagccaa 1020
gtgcaagatc accatctgcc ctgaggaggaa gaacatggat gaccatggta tgccaggatga 1080
aatggagatc ggctacatcc aagcccccaca caaaacgtctt cccgtggctt tcgactctcc 1140
aaggaacaga ggcctgaagg agtttccat caaacgcgtt atgattctat caataggccc 1200
attttataga agaagaaactt gaggctcaga gtggagaatg tggaaatccaa 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.

* * * * *