US 20100136584A1

(19) **United States**

(12) **Patent Application Publication**

Bhatt et al.

(10) **Pub. No.: US 2010/0136584 A1**

(43) **Pub. Date: Jun. 3, 2010**

(54) **METHODS FOR USING ANTIBODIES AND ANALOGS THEREOF**

(75) Inventors: **Ram S. Bhatt**, San Diego, CA (US); **Rishi Bhatt**, San Diego, CA (US)

Correspondence Address:
FOLEY AND LARDNER LLP
SUITE 500
3000 K STREET NW
WASHINGTON, DC 20007 (US)

(73) Assignee: **ICB International, Inc.**

(21) Appl. No.: **12/606,818**

(22) Filed: **Oct. 27, 2009**

Related U.S. Application Data

(63) Continuation-in-part of application No. 12/563,330, filed on Sep. 21, 2009.

(60) Provisional application No. 61/197,601, filed on Oct. 29, 2008, provisional application No. 61/192,732, filed on Sep. 22, 2008.

Publication Classification

(51) **Int. Cl.**

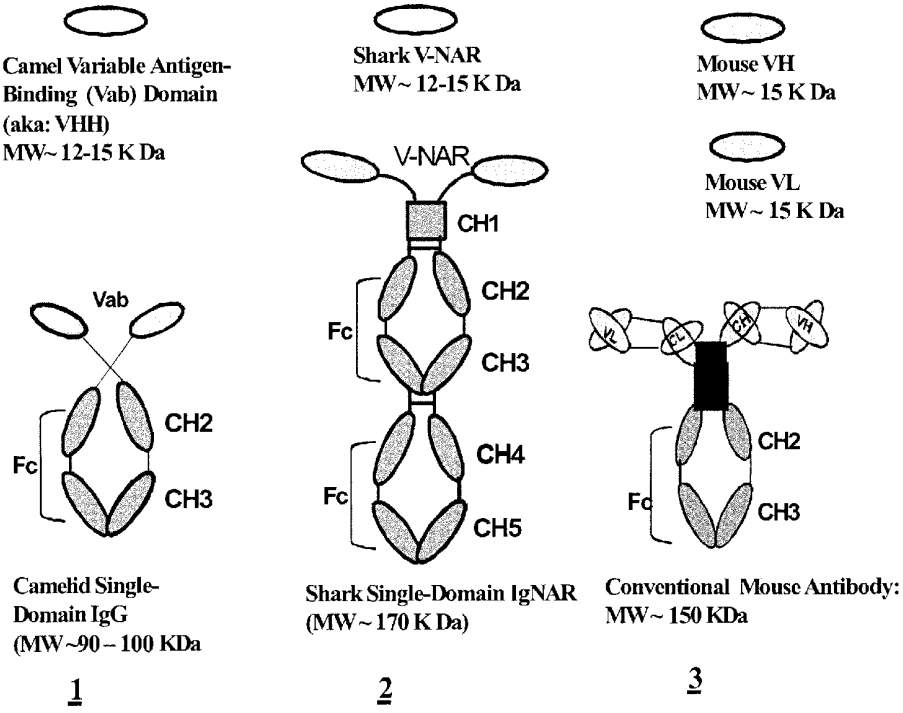
G01N 33/574	(2006.01)
G01N 33/53	(2006.01)
G01N 33/573	(2006.01)

(52) **U.S. Cl.** **435/7.23**; 435/7.94; 435/7.4; 435/7.2; 435/7.21

(57) **ABSTRACT**

Camelid and shark single-domain heavy chain antibodies lacking the light chains, and their analogs are disclosed. Methods for using the antibodies and their analogs to detect antigens are disclosed. Methods to derivatize such antibodies and analogs to develop diagnostic assays are described. Also provided are kits, and methods of using such diagnostics assays.

Comparison of Structures of Camelid, Shark and Mouse Antibodies



Comparison of Structures of Camelid, Shark and Mouse Antibodies

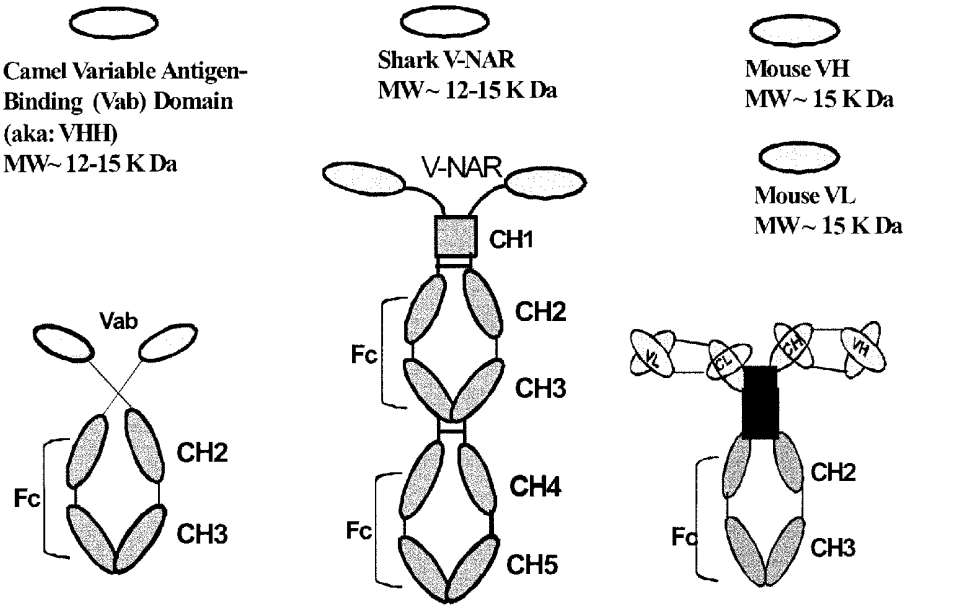


Figure 1

Novel Analogs of Heavy-chain antibodies

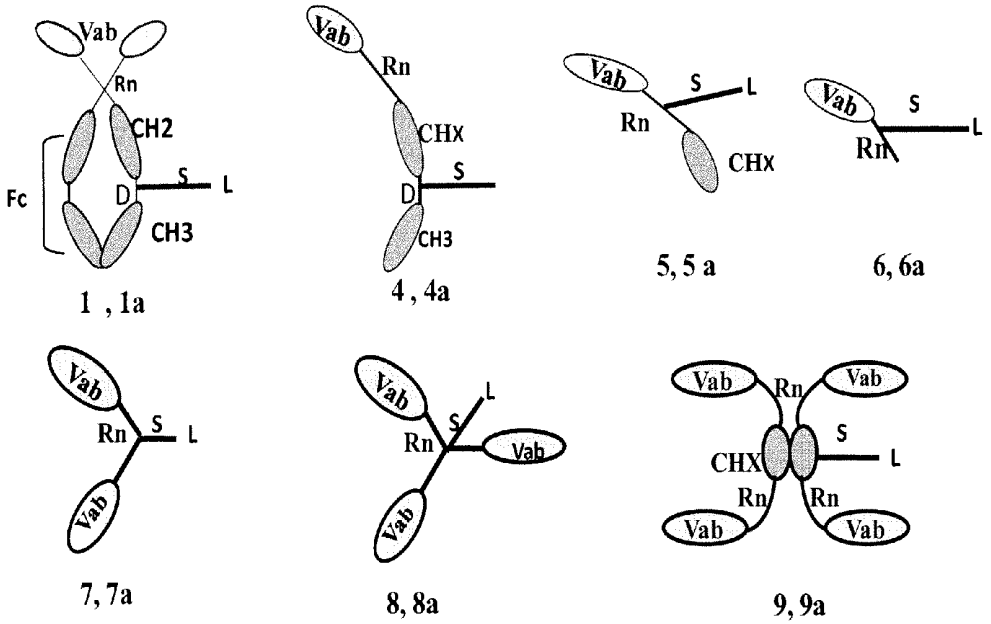
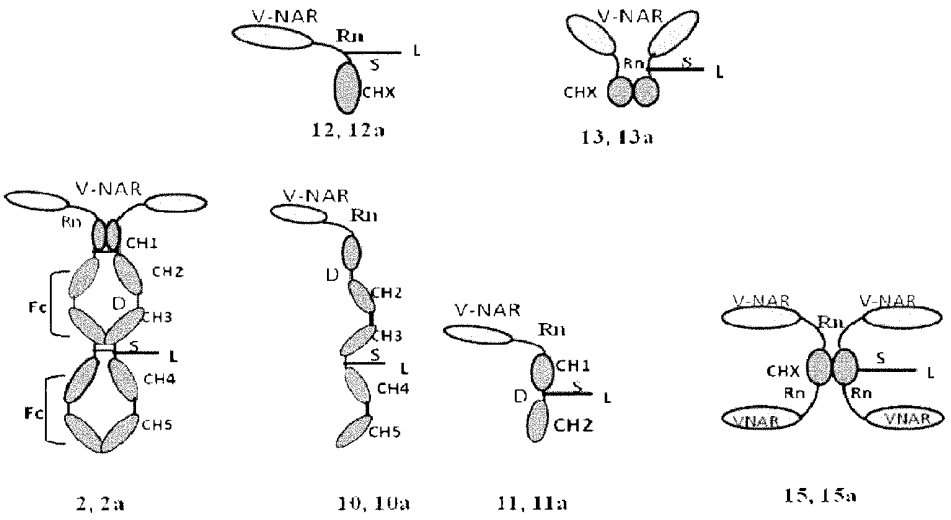


Figure 2

Novel Analogs of Single-Domain Heavy-Chain Shark Antibodies



Wherein

"a" after the digit/number indicates modified analog of the numbered native camelid and/or shark antibody. For example structure "1a" is the analog derived from native, unmodified antibody "1".

S = Heterobifunctional linker with one end being capable of conjugating with heavy-chain antibodies, and the other end with haptens, enzymes, and solid matrixes.

Generic composition of "S" is shown below:

X = NHS, Maleimide, CHO, COOH, CN, SCN, Epoxide, NH_2, Phosphate, thiophosphate, etc

"Y" Forms covalent bond with NH2, SH, CHO, NHS groups of fluorophores, haptens, enzymes, proteins, gold, magnetic particles, and solid matrixes

Figure 3

Y = Maleimide, CHO, COOH, SH, SCN, Epoxide, Phosphate, Thiophosphate

NHS (N-hydroxy-succinimide) with or without a sulfonate group.

$(CH_2CH_2O)n$ where n= 1-500

$(CH_2)n$ where n = 1-15

P = $(CH_2$-R-NHCO$)n$ and n = 1-100

R= methyl, ethyl, isopropyl, phenyl, alkylphenyl, etc.

= Nucleic acids, albumins, proteins and peptides, hydrophilic polymers

L= Biotin, Digoxegenin, Fluorescein, Cy3, Cy5, Cy7, Bodipy dyes, TAMRA, BHQ, MGBNQ, Texas Red, Alexa Fluores 350-750, Rhodamine, Oregon Green Dyes, SYTO dyes, Cascade Blue, Starbright Orange.

= Anti-camelid/shark-biotin-antibody, Streptavidin (SA), Avidin (AV)

= Horseradish Peroxidase (HRP), Alkaline Phosphatase (AP), Luciferase, □-galactosidase,

= Streptavidin-HRP, Streptavidin-AP, Streptavidin-Luciferase, Streptavidin- Galactosidase.

= Solid matrixes, such as, gold nanoparticles, magnetic particles, glass particles, glass slides, microarrays, microchannels, microfluidic devises and their appropriately modified analogs that generates reactive groups for covalent conjugation with "Y".

= Nucleic acids, 10 bases to 1000 bases long, chimeric or not, PNA, LNA, phosphorothioates, methylphosphonates, Si-RNA, micro-RNA, RNA, RNA -Analogues

= radioisotope

Figure 3 Continued

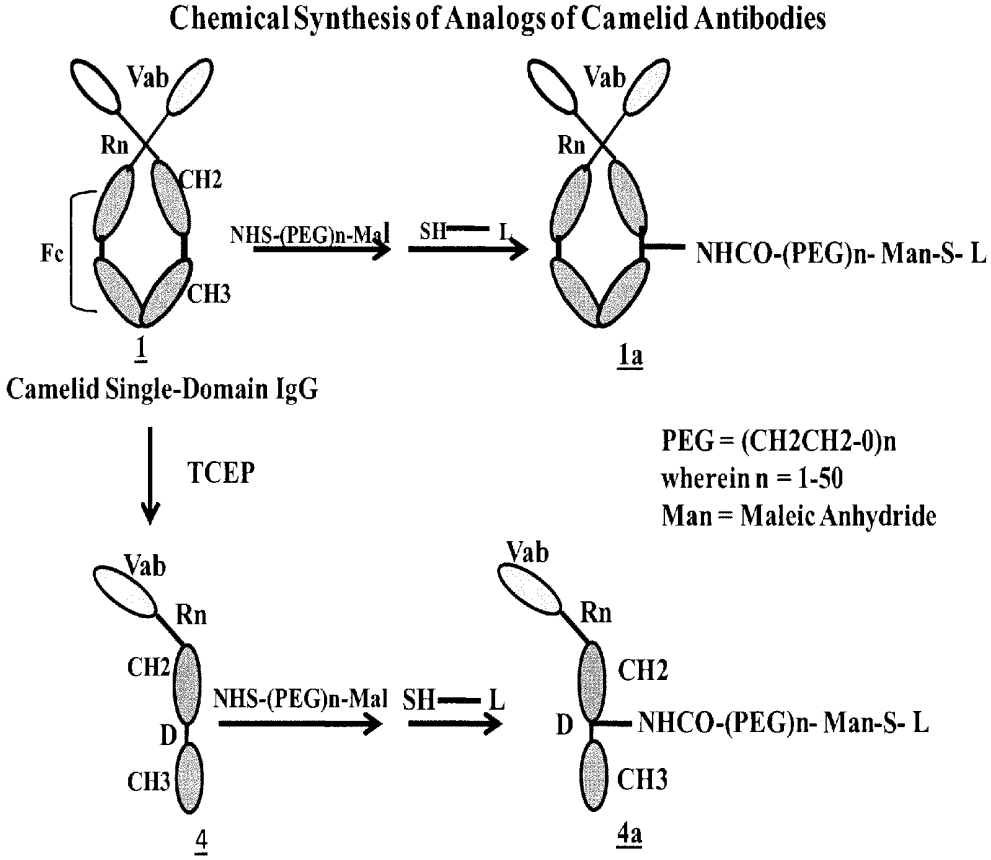

Chemical Synthesis of Analogs of Camelid Antibodies
Vab
Rn
CH2
Fc
CH3
NHS-(PEG)n-Mal
SH— L
NHCO-(PEG)n- Man-S- L
1
1a
Camelid Single-Domain IgG
TCEP
PEG = (CH2CH2-0)n
wherein n = 1-50
Man = Maleic Anhydride
D
4
4a

Figure 4

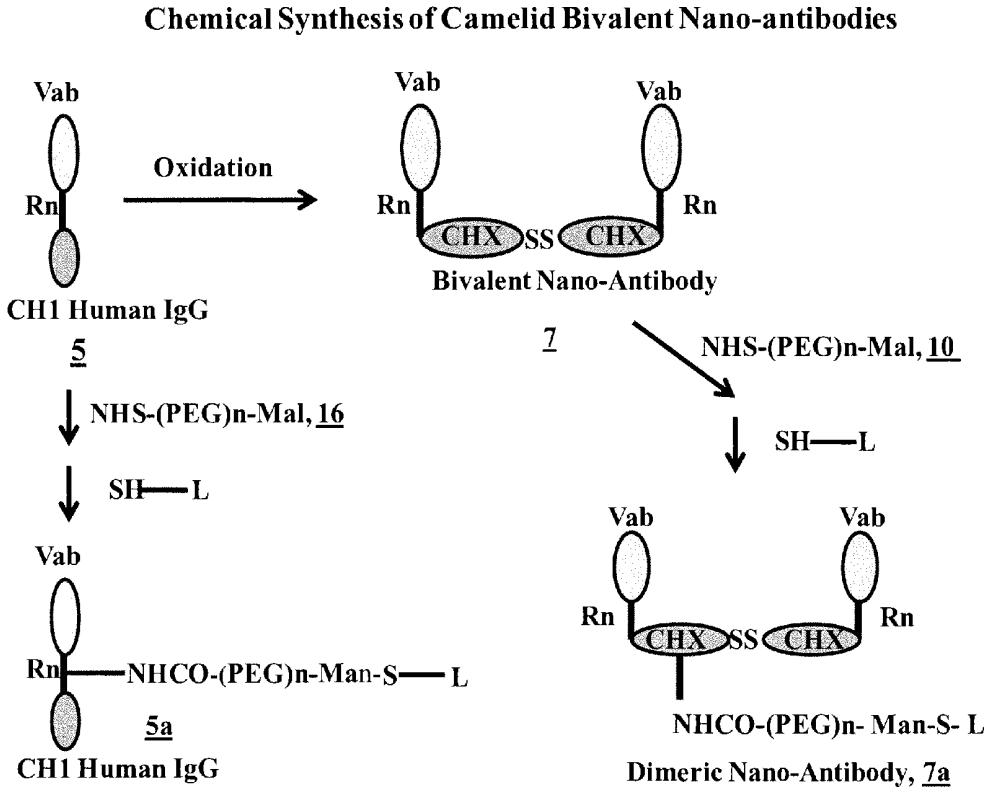

Chemical Synthesis of Camelid Bivalent Nano-antibodies
Vab
Rn
Oxidation
Vab
Vab
Rn
CHX
SS
CHX
Rn
Bivalent Nano-Antibody
7
NHS-(PEG)n-Mal, 10
SH—L
CH1 Human IgG
5
NHS-(PEG)n-Mal, 16
SH—L
Vab
Rn
NHCO-(PEG)n-Man-S—L
5a
CH1 Human IgG
Vab
Vab
Rn
CHX
SS
CHX
Rn
NHCO-(PEG)n- Man-S- L
Dimeric Nano-Antibody, 7a

Figure 5

Chemical Synthesis of Camelid Trimeric and Tetrameric Nano-antibodies

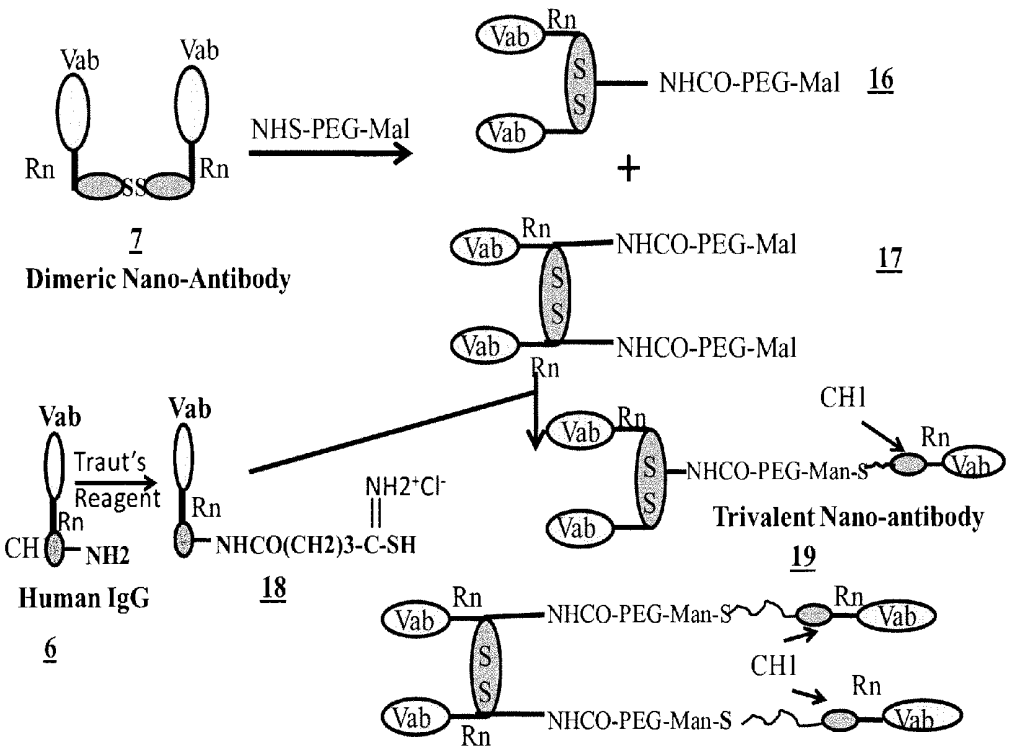


Figure 6 **Tetrameric Camelid Nano-antibody, 20**

Schematics for Cloning and Expression of Single-Domain Shark Antibodies and their Analogs

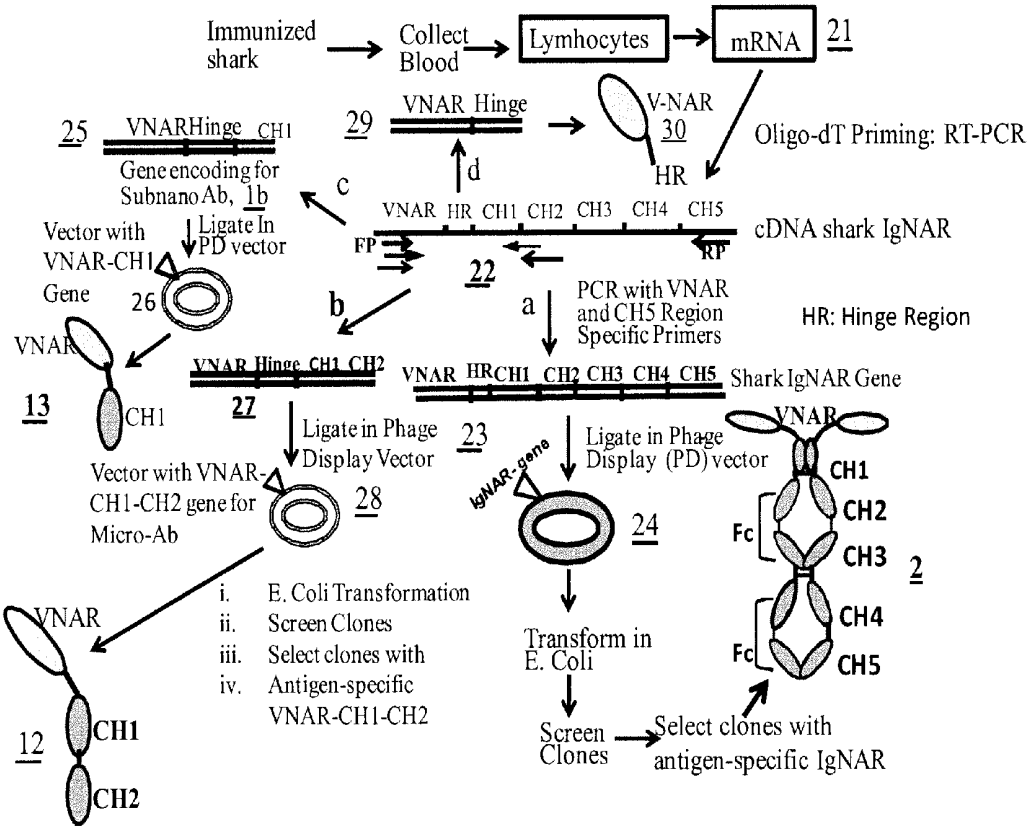


Figure 7

Analogs of Single-Domain Shark Antibodies

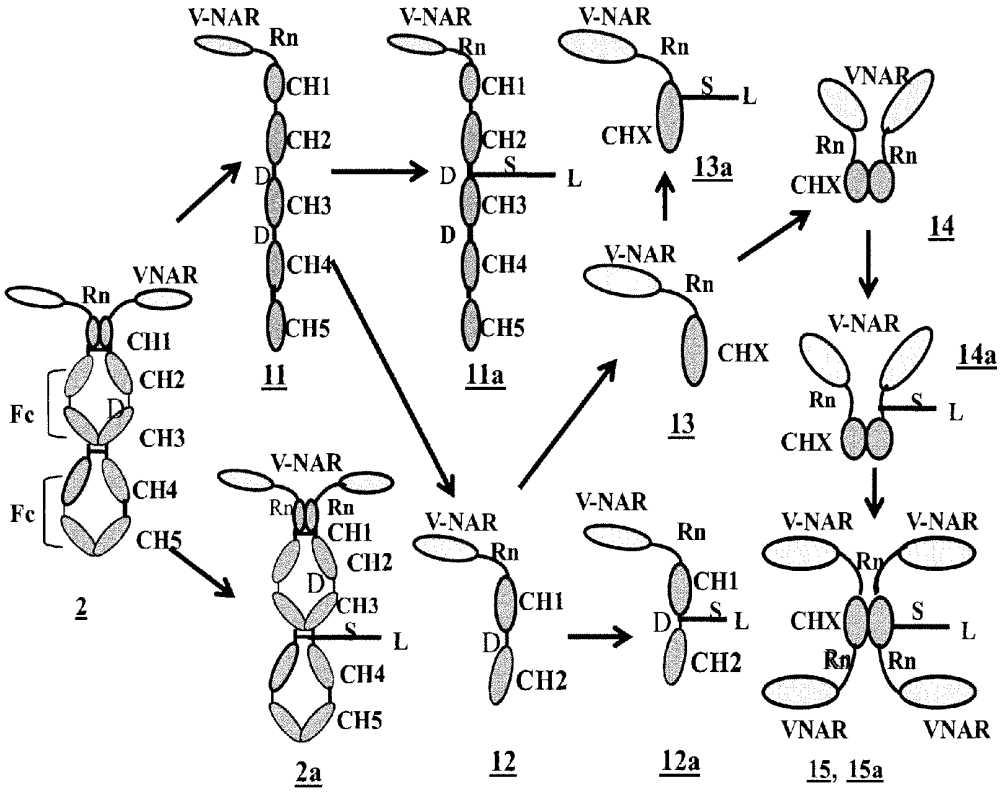


Figure 8

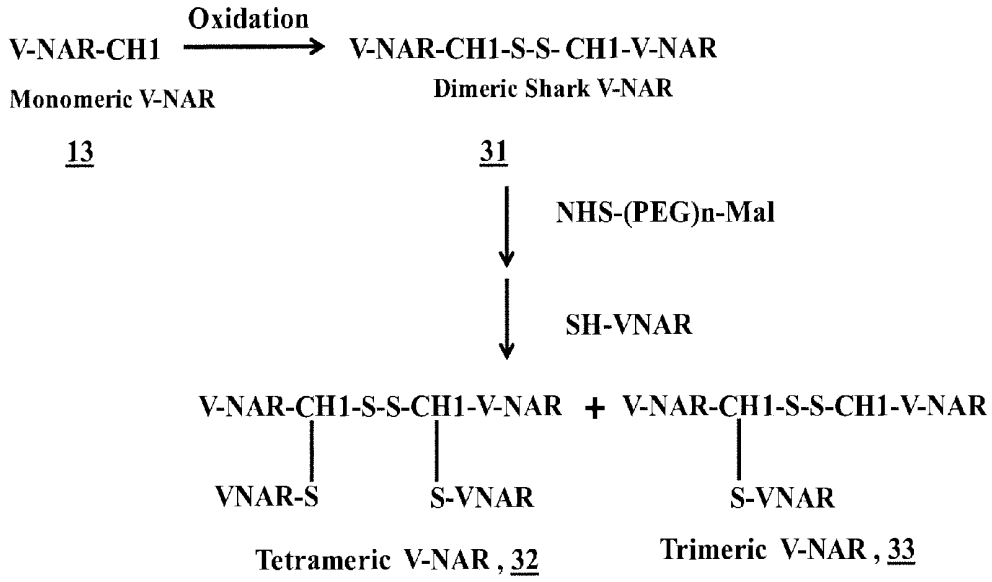

Chemical Synthesis of Shark Dimeric Nano-Antibodies and Analogs
Oxidation
V-NAR-CH1
V-NAR-CH1-S-S- CH1-V-NAR
Monomeric V-NAR
Dimeric Shark V-NAR
13
31
NHS-(PEG)n-Mal
SH-VNAR
V-NAR-CH1-S-S-CH1-V-NAR + V-NAR-CH1-S-S-CH1-V-NAR
VNAR-S
S-VNAR
S-VNAR
Tetrameric V-NAR , 32
Trimeric V-NAR , 33

Figure 9

Immobilization of Single-Domain Heavy-Chain Camelid and Shark Antibodies

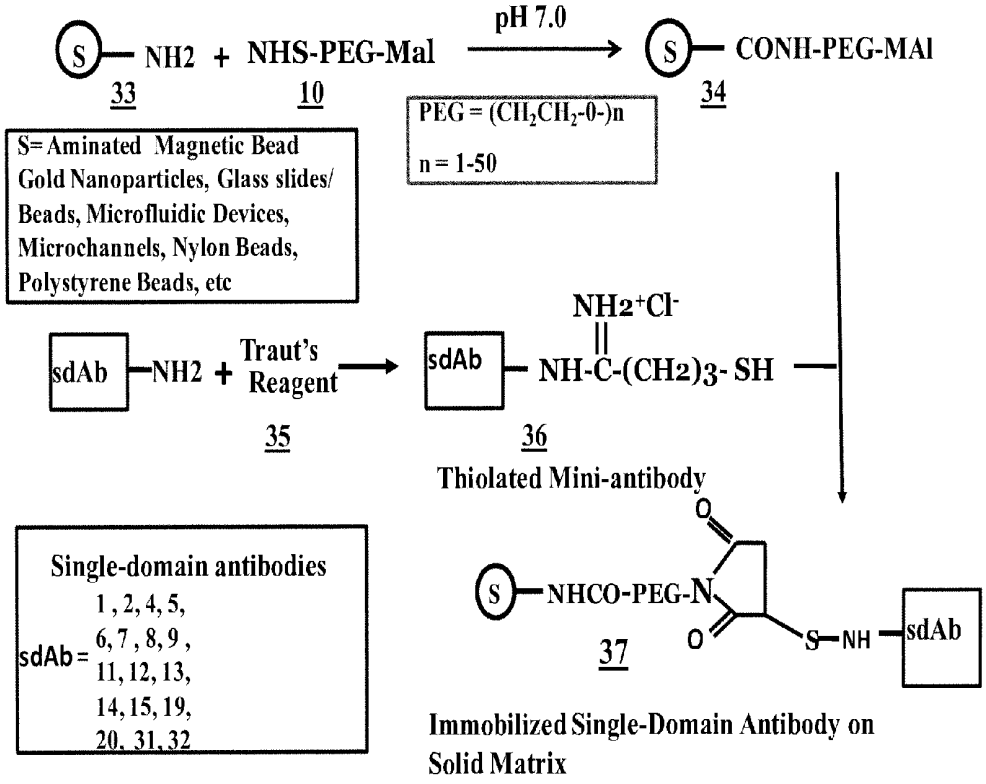
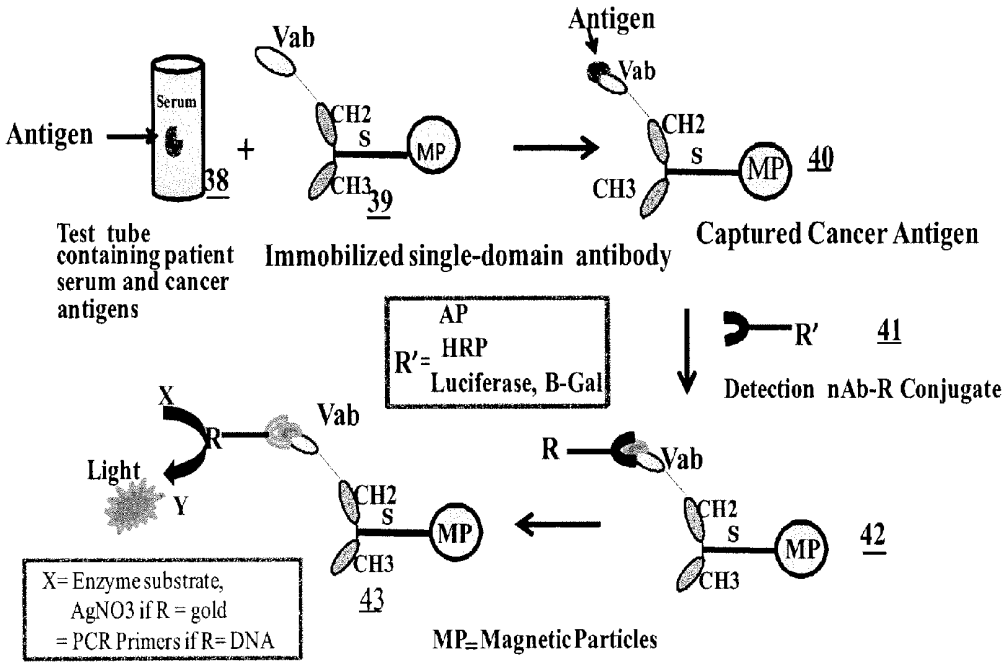


Figure 10

In-Vitro Detection of Pathological Proteins using Camelid and Shark Antibodies and Enzymatic Signal Amplification



Vab = Variable antigen-binding domain of shark and camelid single-domain antibodies

Figure 11

In-Vitro Detection of Pathological Proteins using Camelid and Shark Antibodies and Enzymatic Signal Amplification

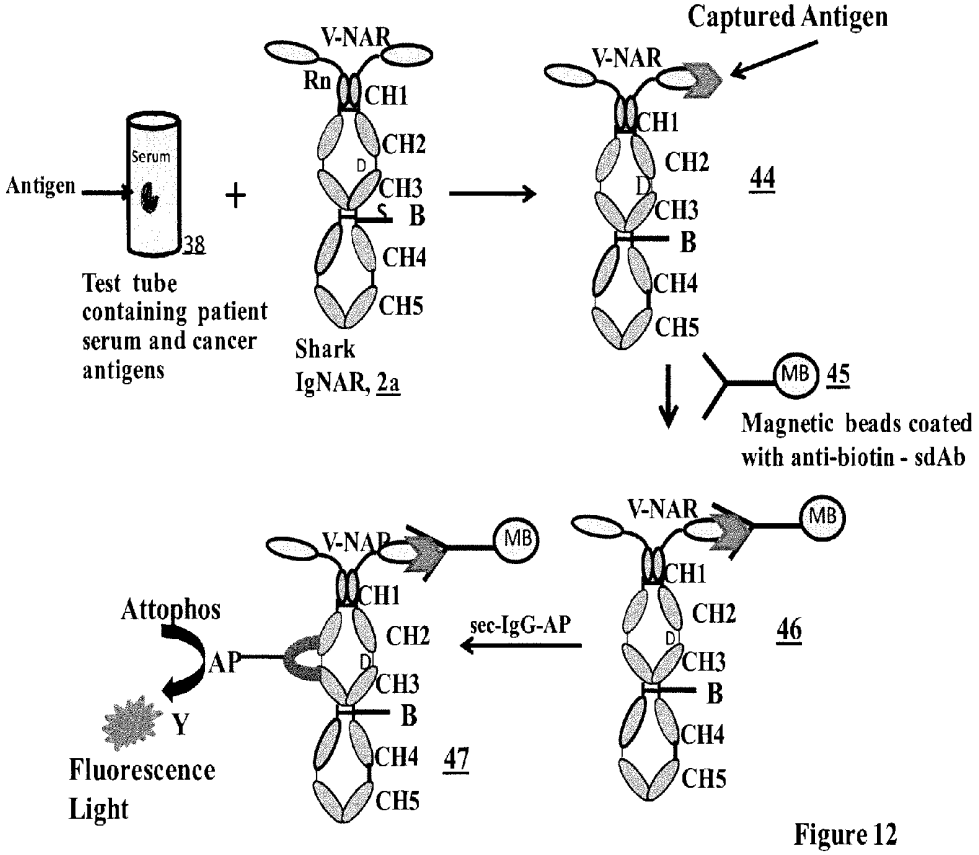


Figure 12

Use of Immuno-PCR and Single-Domain Camelid / Shark Antibodies to Develop Ultra-sensitive Diagnostic Technology to Capture and Detect < 200 molecules of Pathological Proteins

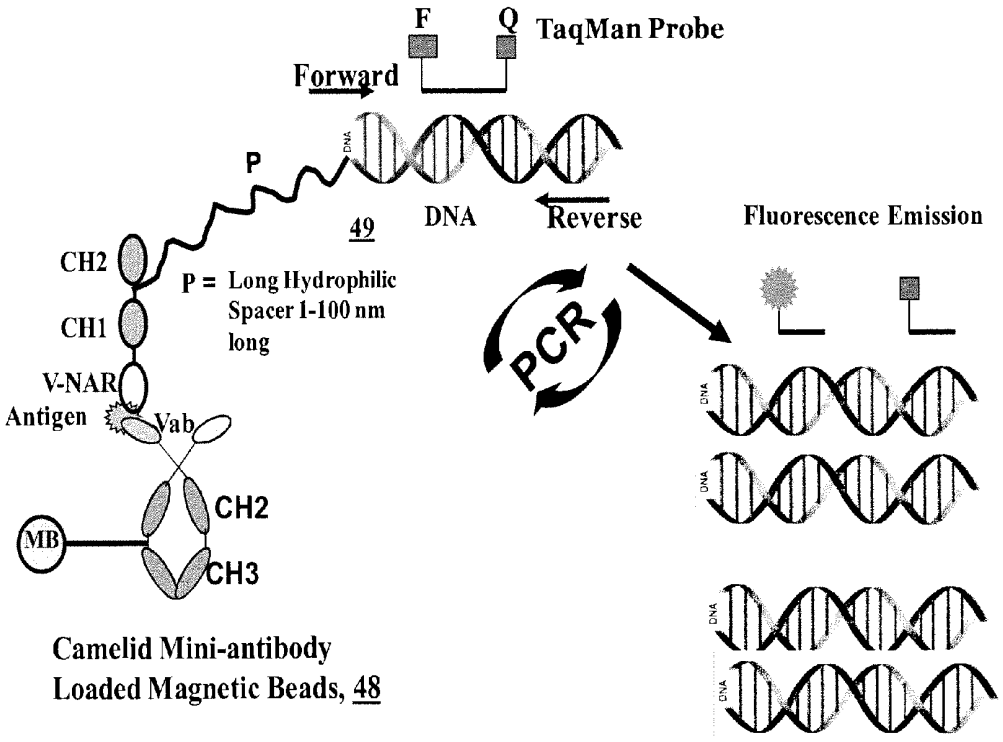


Figure 13

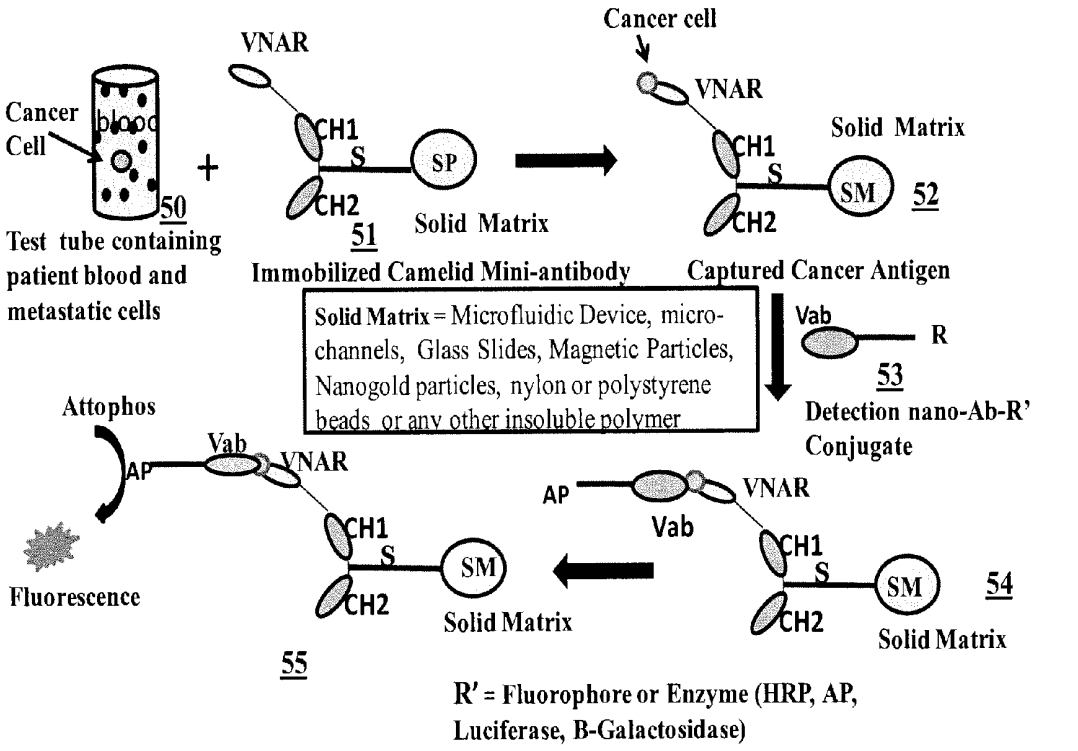

In-Vitro Capture and Detection of Circulating Tumor Cells (CTCs) Using Single-Domain Shark and / or Camel Antibodies:
Cancer Cell
blood
50
Test tube containing patient blood and metastatic cells
+
VNAR
CH1
S
SP
CH2
51
Solid Matrix
Immobilized Camelid Mini-antibody
Cancer cell
VNAR
CH1
S
SM
CH2
Solid Matrix
52
Captured Cancer Antigen
Solid Matrix = Microfluidic Device, micro-channels, Glass Slides, Magnetic Particles, Nanogold particles, nylon or polystyrene beads or any other insoluble polymer
Vab
R
53
Detection nano-Ab-R' Conjugate
Attophos
AP
Vab
VNAR
CH1
S
SM
CH2
Solid Matrix
55
Fluorescence
AP
Vab
VNAR
CH1
S
SM
CH2
Solid Matrix
54
R' = Fluorophore or Enzyme (HRP, AP, Luciferase, B-Galactosidase)

Figure 14

Non-invasive Detection of Prenatal Genetic Disorders from Captured Circulating Fetal Cells (CFCs) Using Single-Domain Shark and Camelid Heavy-Chain only Antibodies:

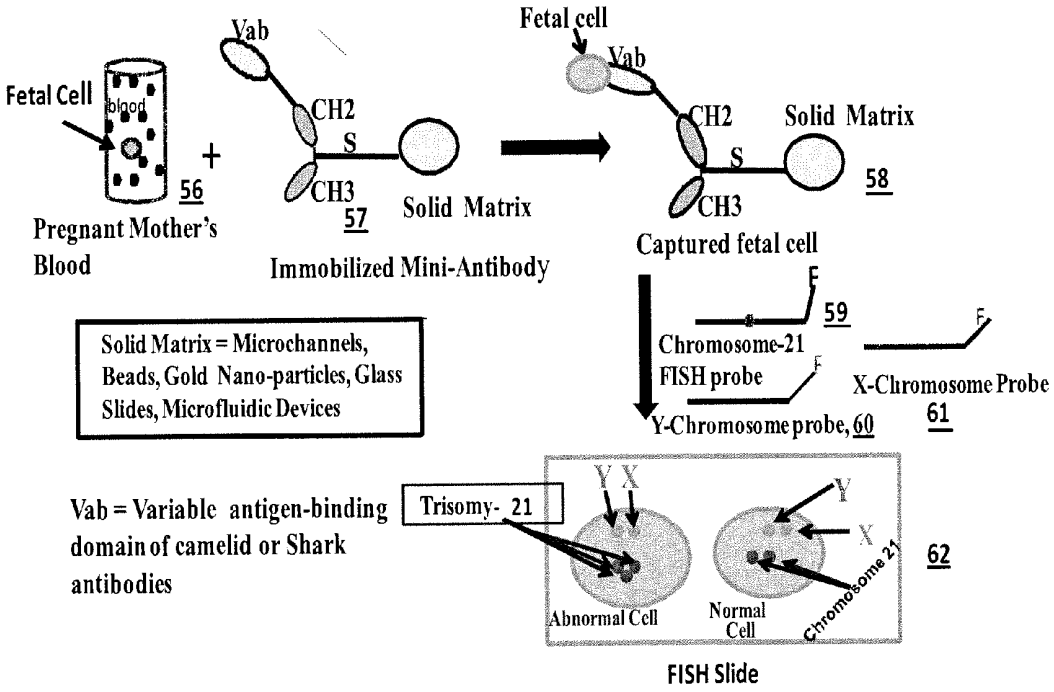


Figure 15

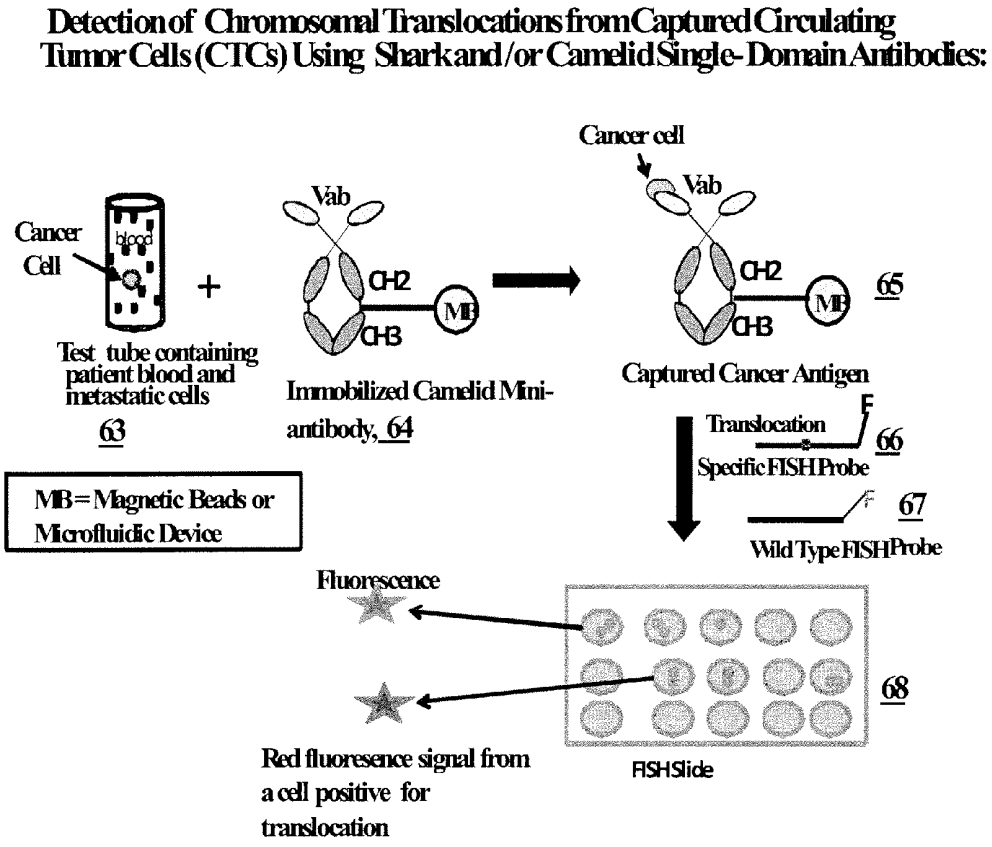

Detection of Chromosomal Translocations from Captured Circulating Tumor Cells (CTCs) Using Shark and/or Camelid Single-Domain Antibodies:
Cancer Cell
blood
Test tube containing patient blood and metastatic cells
63
Vab
CH2
CH3
MB
Immobilized Camelid Mini-antibody, 64
MB = Magnetic Beads or Microfluidic Device
Cancer cell
Captured Cancer Antigen
65
Translocation
Specific FISH Probe
66
F
67
Wild Type FISH Probe
Fluorescence
68
FISH Slide
Red fluoresence signal from a cell positive for translocation

Figure 16

```
   1 ggccggtcag cgtcgctgcc ggtctccggc ggagacggac tctggagttt gggcggcccg
  61 ggcggccact aggtactctg atattccgta ctaaacacgt ctgcaagtca agatgtcgca
 121 cccgtccccc caagccaagc cctccaaccc cagtaaccct cgagtcttct ttgacgtgga
 181 catcggaggg gagcgagttg gtcgaattgt cttagaattg tttgcagata tcgtacccaa
 241 aactgcggaa aattttcgtg cactgtgtac aggagaaaaa ggcattggac acacgactgg
 301 gaaacctctc catttcaaag gatgcccttt tcatcgaatt attaagaaat ttatgattca
 361 gggtggagac ttctcaaatc agaatgggac aggtggagaa agtatttatg gtgaaaaatt
 421 tgaagatgaa aatttccatt acaagcatga tcgggagggt ttactgagca tggcaaatgc
 481 aggccgcaac acaaacggtt ctcagttttt tatcacaaca gttccaactc ctcatttgga
 541 tgggaaacat gtggtgtttg gccaagtaat taaaggaata ggagtggcaa ggatattgga
 601 aaatgtggaa gtgaaaggtg aaaaacctgc taaattgtgc gttattgcag aatgtggaga
 661 attgaaggaa ggagatgacg ggggaatatt cccaaaagat ggctctggcg acagtcatcc
 721 agatttccct gaggatgcgg atatagattt aaaagatgta gataaaattt tattaataac
 781 agaagactta aaaaacattg gaaatacttt tttcaaatcc cagaactggg agatggctat
 841 taaaaaatat gcagaagttt taagatacgt ggacagttca aaggctgtta ttgagacagc
 901 agatagagcc aagctgcaac ctatagcttt aagctgtgta ctgaatattg gtgcttgtaa
 961 actgaagatg tcaaattggc agggagcaat tgacagttgt ttagaggctc ttgaactaga
1021 cccatcaaat accaaagcat tgtaccgcag agctcaagga tggcaaggat taaaagaata
1081 tgatcaagca ttggctgatc ttaagaaagc tcaggggata gcaccagaag ataaagctat
1141 ccaggcagaa ttgctgaaag tcaaacaaaa gataaaggca cagaaagata aagagaaggc
1201 agtatatgca aaaatgtttg cttagaaagg attcagtttt gcttattgtg tgttgattgt
1261 ataaatgcaa taagaaaatg taaaggtttt tgtctatgaa tatgatccct aatgtgtttc
1321 ttttgacacc ttagttcctt actgtttaca gtttaggagt actgataggg gttcatgctt
1381 aataaacatg tcacaataca gtaagtaaag tggttttgtt tgtttctttg agatggagtc
```

FIGURE 17

```
1441 ttgctctgtc acccaggctg gagtgcggtg gcgcaatctc ggctcactgc atcctctgcc
1501 tcccgggttc aagcaattct cctgcctcag cttcccaagt agctgggatt acaggcacgt
1561 gccaccacgc ccagctaatt tttgtatttt tagtagagat ggggtttcac catattggtc
1621 acgtcacgtt ggtcttgaac tcctgacctt gtgatccacc ccgccttggc ctcccaaagt
1681 gctgggatta caggtgtgag ccaccgtgcc cggccaagta aaatgttttt taaaatggtt
1741 atgtgcatta ttcataaaaa ataatggtgt ccagtctttt taaacttgta aagacacatc
1801 ttattgaata aagagatgag agcttaagtt tgtaaaaaaa aaaaaaaaaa a
```

FIGURE 17 Continued

```
   1 atccccatcc cgtggagtgg ccggcgacaa gatggcagca gcgtgtcgga gcgtgaaggg
  61 cctggtggcg gtaataaccg gaggagcctc gggcctgggc ctggccacgg cggagcgact
 121 tgtggggcag ggagcctctg ctgtgcttct ggacctgccc aactcgggtg gggaggccca
 181 agccaagaag ttaggaaaca actgcgtttt cgccccagcc gacgtgacct ctgagaagga
 241 tgtgcaaaca gctctggctc tagcaaaagg aaagtttggc cgtgtggatg tagctgtcaa
 301 ctgtgcaggc atcgcggtgg ctagcaagac gtacaactta aagaagggcc agacccatac
 361 cttggaagac ttccagcgag ttcttgatgt gaatctcatg ggcaccttca atgtgatccg
 421 cctggtggct ggtgagatgg gccagaatga accagaccag ggaggccaac gtggggtcat
 481 catcaacact gccagtgtgg ctgccttcga gggtcaggtt ggacaagctg catactctgc
 541 ttccaagggg ggaatagtgg gcatgacact gcccattgct cgggatctgg ctcccatagg
 601 tatccgggtg atgaccattg ccccaggtct gtttggcacc ccactgctga ccagcctccc
 661 agagaaagtg tgcaacttct tggccagcca agtgcccttc cctagccgac tgggtgaccc
 721 tgctgagtat gctcacctcg tacaggccat catcgagaac ccattcctca atggagaggt
 781 catccggctg gatggggcca ttcgtatgca gccttgaagg gagaaggcag agaaaacaca
 841 cgctcctctg cccttccttt ccctggggta ctactctcca gcttgggagg aagcccagta
 901 gccattttgt aactgcctac cagtcgccct ctgtgcctaa taaagtctct ttttctcaca
 961 gag
```

FIGURE 18

```
   1 tcctttccgc ttccggtgtc ccctacagtc atggctgccg ccgtcgctgc tgccggtgca
  61 ggggaacccc agtccccgga cgaattgctc ccgaaaggcg acgcggagaa gcctgaggag
 121 gagctggagg aggacgacga tgaggagcta gatgagaccc tgtcggagag actatggggc
 181 ctgacggaga tgtttccgga gagggtccgg tccgcggccg gagccacttt tgatctttcc
 241 ctctttgtgg ctcagaaaat gtacaggttt tccagggcag ccttgtggat tgggaccact
 301 tcctttatga tcctggttct tcccgttgtc tttgagacgg agaagttgca aatggagcaa
 361 cagcagcaac tgcagcagcg gcagatactt ctaggaccta acacagggct ctcaggagga
 421 atgccagggg ctctaccctc acttcctgga aagatctaga ttgttattgc tgtttgagct
 481 gtctcagtgg gataagtttg aaattcaagt gtttgaactg ctgataattt ggatttttt
 541 tttttttaa ctttggcaca ttgatctatc taaacctggt ggggagaatt atccccacat
 601 tgtctcatgg aaagactcaa cttgcaactg tgccctccac actatcctta cttctgtctc
 661 cactctgata ccagagtgca gccatgcaga tggttattcc agctctggtc acccgactcc
 721 tttcaccaaa ttgctcctaa ctggaagatc tcactttccc cttgtggggt aggaaccgat
 781 gccagtggga gggatgtgcc cctgaccatt aacgactgtt ttttttttt ttttttaaag
 841 aatggagttg ttggggcagg acatgcacac aatgtgaaac agacaaaatg cattacacct
 901 gtagtgtaaa gtggccacta tgaatcccta tgtatgagag gagggaggca ggctgcagct
 961 tcagccacag aatggggact atggaagaca gcaggagctc atttcctctg cacattttgg
1021 ctgttagacc tgtgtgtgtg tttaaaaaaa gagaagtcag tgctcacttt ttgtatttaa
1081 atattaaaaa tgattccaac tgaaagtgtc atcctaagta ccttgaaatg agaccacgtc
1141 agagacatgt actgcccctc acattttctc acctaaacca gcagcacctc catcttaaca
1201 gccataggcc caaattgttt ccaagtgaaa attcatttt agccaagtac ttcatagcaa
1261 tctttgccct gaatttaggc agtcactttg agatccatca gcctaaaaca aaggattggg
1321 tctatgtact tctttagtct ataatgacac tgtgtaatta taaagtattt gtagggaaaa
1381 aaaaaaaaaa aaaaaa
```

FIGURE 19

TGGTACCCGGGATTCGGCCATTACGGCCGGGGGGGTCTTTTCGTATTTGTCATACCGTGA

TCCGAACTTGCTGAAAACACTAGAGGTCTATGATGGGACTGCAAAGTTCCTCAGAGAACT

AGATGTAGATGATGATGCTCTCACAAAAGCTATTATTGGAACCATCGGGGATGTTGATTC

CTACCAGCTACCAGATGCTAAAGGTTACAGCAGTCTGATGCGGTATTTGTTGGGCATCAC

CGAGGAGGAACGCCAGCAAAGGCGCGAAGAGATACTCGCAACCAGCGTGAAGGATTTCAA

GGAGTTTGCCGATGCTGTCGAAACGATCAATGACAATGGGGTTGTGGTGGCTGTAGCATC

GCCTGACGACGTCGAGGCAGCAAACAAAGAGAAGTCGTTATTTTCAGACATCAAGAAGTG

CCTGTGAGGCCTTTCCATTCCCACGCCGGCAGGGTTCAGCCTGGACCTGCGCCGAGCAGA

GGTTCTTCAGCGTTTTGCTTCACTGATGAAGAATGGTCTGGCGCTGTAACATGAGTTCGC

GTGGCGAGAGCACTGATCACCTTGTCTGTAGCCCGGTTACGGTTGTTTATTTTTCGCGCC

CATTTTTAGGGGAAATAATTAGTTCAGGGTTTTGAGCGAAGCACAATAGTTAAACAGAGT

TCCATTGCTGCCTTTCCCCCTCTGCAATCAGTGCTCAAACAAGCCGGTTGTACACCAAGT

TGCTGATAGGAGATTAAGAGCTGATGAGTGATGAGAACCTGCTCAAAATTTAGCTGAATA

CTCTGTAACTTGATATACCAAAGATATAAAGGTATTGGACTACTGATT

FIGURE 20

```
   1 atgaatttac aaccaatttt ctggattgga ctgatcagtt cagtttgctg tgtgtttgct
  61 caaacagatg aaaatagatg tttaaaagca aatgccaaat catgtggaga atgtatacaa
 121 gcagggccaa attgtgggtg gtgcacaaat tcaacatttt tacaggaagg aatgcctact
 181 tctgcacgat gtgatgattt agaagcctta aaaaagaagg gttgccctcc agatgacata
 241 gaaaatccca gaggctccaa agatataaag aaaaataaaa atgtaaccaa ccgtagcaaa
 301 ggaacagcag agaagctcaa gccagaggat attactcaga tccaaccaca gcagttggtt
 361 ttgcgattaa gatcagggga gccacagaca tttacattaa aattcaagag agctgaagac
 421 tatcccattg acctctacta ccttatggac ctgtcttact caatgaaaga cgatttggag
 481 aatgtaaaaa gtcttggaac agatctgatg aatgaaatga ggaggattac ttcggacttc
 541 agaattggat ttggctcatt tgtggaaaag actgtgatgc cttacattag cacaacacca
 601 gctaagctca ggaacccttg cacaagtgaa cagaactgca ccagcccatt tagctacaaa
 661 aatgtgctca gtcttactaa taaaggagaa gtatttaatg aacttgttgg aaaacagcgc
 721 atatctggaa atttggattc tccagaaggt ggtttcgatg ccatcatgca agttgcagtt
 781 tgtggatcac tgattggctg gaggaatgtt acacggctgc tggtgttttc cacagatgcc
 841 gggtttcact ttgctggaga tgggaaactt ggtggcattg ttttaccaaa tgatggacaa
 901 tgtcacctgg aaaataatat gtacacaatg agccattatt atgattatcc ttctattgct
 961 caccttgtcc agaaactgag tgaaaataat attcagacaa tttttgcagt tactgaagaa
1021 tttcagcctg tttacaagga gctgaaaaac ttgatcccta agtcagcagt aggaacatta
1081 tctgcaaatt ctagcaatgt aattcagttg atcattgatg catacaattc cctttcctca
1141 gaagtcattt tggaaaacgg caaattgtca gaaggcgtaa caataagtta caaatcttac
1201 tgcaagaacg gggtgaatgg aacaggggaa aatggaagaa aatgttccaa tatttccatt
1261 ggagatgagg ttcaatttga aattagcata acttcaaata agtgtccaaa aaaggattct
1321 gacagcttta aaattaggcc tctgggcttt acggaggaag tagaggttat tcttcagtac
1381 atctgtgaat gtgaatgcca aagcgaaggc atcgctgaaa gtcccaagtg tcatgaagga
```

FIGURE 21

1441 aatgggacat ttgagtgtgg cgcgtgcagg tgcaatgaag ggcgtgttgg tagacattgt

1501 gaatgcagca cagatgaagt taacagtgaa gacatggatg cttactgcag gaaagaaaac

1561 agttcagaaa tctgcagtaa caatggagag tgcgtctgcg gacagtgtgt ttgtaggaag

1621 agggataata caaatgaaat ttattctggc aaattctgcg agtgtgataa tttcaactgt

1681 gatagatcca atggcttaat ttgtggagga aatggtgttt gcaagtgtcg tgtgtgtgag

1741 tgcaacccca actacactgg cagtgcatgt gactgttctt tggatactag tacttgtgaa

1801 gccagcaacg gacagatctg caatggccgg ggcatctgcg agtgtggtgt ctgtaagtgt

1861 acagatccga agtttcaagg gcaaacgtgt gagatgtgtc agacctgcct tggtgtctgt

1921 gctgagcata aagaatgtgt tcagtgcaga gccttcaata aaggagaaaa gaaagacaca

1981 tgcacacagg aatgttccta ttttaacatt accaaggtag aaagtcggga caaattaccc

2041 cagccggtcc aacctgatcc tgtgtcccat tgtaaggaga aggatgttga cgactgttgg

2101 ttctatttta cgtattcagt gaatgggaac aacgaggtca tggttcatgt tgtggagaat

2161 ccagagtgtc ccactggtcc agacatcatt ccaattgtag ctggtgtggt tgctggaatt

2221 gttcttattg gccttgcatt actgctgata tggaagcttt taatgataat tcatgacaga

2281 agggagtttg ctaaatttga aaaggagaaa atgaatgcca aatgggacac gtctctctct

2341 gtcgcccagc ctggagtgca gtggtgtgat atcagctcac tgcaacctct gacttccaga

2401 ttccagcaat tctcctgcct cagcctcccg agtacctggg attacagggt gaaaatccta

2461 tttataagag tgccgtaaca actgtggtca atccgaagta tgagggaaaa tgagtactgc

2521 ccgtgcaaat cccacaacac tgaatgcaaa gtagcaattt ccatagtcac agttaggtag

2581 ctttagggca atattgccat ggttttactc atgtgcaggt tttgaaaatg tacaatatgt

2641 ataattttta aaatgtttta ttattttgaa aataatgttg taattcatgc cagggactga

FIGURE 21 Continued

2701 caaaagactt gagacaggat ggttactctt gtcagctaag gtcacattgt gccttttga

2761 ccttttcttc ctggactatt gaaatcaagc ttattggatt aagtgatatt tctatagcga

2821 ttgaaagggc aatagttaaa gtaatgagca tgatgagagt ttctgttaat catgtattaa

2881 aactgatttt tagctttaca aatatgtcag tttgcagtta tgcagaatcc aaagtaaatg

2941 tcctgctagc tagttaagga ttgttttaaa tctgttattt tgctatttgc ctgttagaca

3001 tgactgatga catatctgaa agacaagtat gttgagagtt gctggtgtaa aatacgtttg

3061 aaatagttga tctacaaagg ccatgggaaa aattcagaga gttaggaagg aaaaaccaat

3121 agctttaaaa cctgtgtgcc attttaagag ttacttaatg tttggtaact tttatgcctt

3181 cactttacaa attcaagcct tagataaaag aaccgagcaa ttttctgcta aaaagtcctt

3241 gatttagcac tatttacata caggccatac tttacaaagt atttgctgaa tggggacctt

3301 ttgagttgaa tttatttttat tatttttatt ttgtttaatg tctggtgctt tctgtcacct

3361 cttctaatct tttaatgtat ttgtttgcaa ttttggggta agacttttt tatgagtact

3421 ttttctttga agttttagcg gtcaatttgc ctttttaatg aacatgtgaa gttatactgt

3481 ggctatgcaa cagctctcac ctacgcgagt cttactttga gttagtgcca taacagacca

3541 ctgtatgttt acttctcacc atttgagttg cccatcttgt ttcacactag tcacattctt

3601 gttttaagtg cctttagttt taacagttca ctttttacag tgctatttac tgaagttatt

3661 tattaaatat gcctaaaata cttaaatcgg atgtcttgac tctgatgtat tttatcaggt

3721 tgtgtgcatg aaatttttat agattaaaga agttgaggaa aagcaaaaaa aaaa

FIGURE 21 Continued

```
   1 gggactgagc atggatttcg gactggccct cctgctggcg gggcttctgg ggctcctcct
  61 cggccagtcc ctccaggtga agcccctgca ggtggagccc ccggagccgg tggtggccgt
 121 ggccttgggc gcctcgcgcc agctcacctg ccgcctggcc tgcgcggacc gcggggcctc
 181 ggtgcagtgg cggggcctgg acaccagcct gggcgcggtg cagtcggaca cgggccgcag
 241 cgtcctcacc gtgcgcaacg cctcgctgtc ggcggccggg acccgcgtgt gcgtgggctc
 301 ctgcgggggc cgcaccttcc agcacaccgt gcagctcctt gtgtacgcct tcccggacca
 361 gctgaccgtc tccccagcag ccctggtgcc tggtgacccg gaggtggcct gtacggccca
 421 caaagtcacg cccgtggacc ccaacgcgct ctccttctcc ctgctcgtcg ggggccagga
 481 actggagggg gcgcaagccc tgggcccgga ggtgcaggag gaggaggagg agccccaggg
 541 ggacgaggac gtgctgttca gggtgacaga gcgctggcgg ctgccgcccc tggggacccc
 601 tgtcccgccc gccctctact gccaggccac gatgaggctg cctggcttgg agctcagcca
 661 ccgccaggcc atccccgtcc tgcacagccc gacctccccg gagcctcccg acaccacctc
 721 cccggagtct cccgacacca cctccccgga gtctcccgac accacctccc aggagcctcc
 781 cgacaccacc tccccggagc ctcccgacaa gacctccccg gagcccgccc cccagcaggg
 841 ctccacacac acccccagga gcccaggctc caccaggact cgccgccctg agatctccca
 901 ggctgggccc acgcagggag aagtgatccc aacaggctcg tccaaacctg cgggtgacca
 961 gctgcccgcg gctctgtgga ccagcagtgc ggtgctggga ctgctgctcc tggccttgcc
1021 cacctatcac ctctggaaac gctgccggca cctggctgag gacgacaccc acccaccagc
1081 ttctctgagg cttctgcccc aggtgtcggc ctgggctggg ttaaggggga ccggccaggt
1141 cgggatcagc ccctcctgag tggccagcct ttccccctgt gaaagcaaaa tagcttggac
1201 cccttcaagt tgagaactgg tcagggcaaa cctgcctccc attctactca aagtcatccc
1261 tctgttcaca gagatggatg catgttctga ttgcctcttt ggagaagctc atcagaaact
1321 caaaagaagg ccactgtttg tctcacctac ccatgacctg aagcccctcc ctgagtggtc
1381 cccacctttc tggacggaac cacgtacttt ttacatacat tgattcatgt ctcacgtctc
1441 cctaaaaatg cgtaagacca agctgtgccc tgaccaccct gggcccctgt cgtcaggacc
1501 tcctgaggct ttggcaaata aacctcctaa aatgataaaa aaaaaa
```

FIGURE 22

```
  1 agattaattc acttccaggc atttcatctt cattcatttt ccaaaggggt accctgagat
 61 cacaaaggat acaaaattat gggcaaggga gttcgagtgt tgaacagcag cgagggtgtt
121 aaggggacca tcttcttcac tcaggaagga aacggtacca ccactgtgac aggaaccgtt
181 tctggcctta agcctggtct ccatggtttc catgtccatg ctcttggtga caccactaac
241 ggttgcatgt ctaccggtcc acatttcaac cctgaaggta aaacccacgg tgcacctgag
301 gatgctaatc gacatgctgg agatctagga aacatcactg ttggggatga tggaactgcc
361 accttcacaa tcactgacag ccagattcct cttgatggac caaactctat tgttggaagg
421 gctgttgttg tccacgcaga acctgatgac ctgggaaagg gaggccatga actcagcctt
481 actactggaa acgcaggtgg ccgtgttgct tgtggtatta ttggtcttca gggctaagct
541 gttgcttttc gaggacgaga gtgatgtaat aaggaggttc ttacctctag acatggctag
601 tttgtgtatt ctttggtgtg tggctgtatt aattgagctt agtggctcga tgcatttggt
661 ttaagacgga agaaaacaga aaatccaaac tttttctatt tcatgaataa cagaggacgt
721 ggttgaaaac gataaaatat tgaatatgaa aaaaaaaaaa aaaaaa
```

FIGURE 23

```
   1 ctagagatag aattgtgact agaataaagg ctataattat tatagaggtt ttaattgttt
  61 gaattgctca tggtagtgga agtagaagag caatttctag gtcaaataat agaaatgtaa
 121 ttgctaccaa gaaaaatttt attgagaatg gtagacgtgc agagcttgta gggtcgaatc
 181 cgcattcata tggatttgaa gcatggcagt gtcagcaact ttgtctagag ccttccagaa
 241 acagataatg acaaggctct aaaagggttc cagctgaagt ttctgagcgg agcacgctgt
 301 gtggctccct aggctgagtt tccaagctgc tggttcatgc cgttgacaaa ctgcaggatg
 361 gtgcccgttc gcaggccgct gtcgctgctc ctcaccttct tcctctgcgc ctgtgctgag
 421 acacccccca ggtttacacg aactccggtt gatcagacag gggtctctgg aggagttgca
 481 tcattcattt gtcaagctac aggagaccca agacctaaaa ttgtctggaa caaaaaagga
 541 aagaaagtca gcaaccagag atttgaggta atagaatttg acgatgggtc tggatcagta
 601 ctcagaatac agcccttaag gactccacgg gatgaggcca tttatgaatg tgtggcctcg
 661 aataatgtgg gagaaatcag tgtgtccaca agactcacag ttttacgtga ggatcagatt
 721 cctagagggt tccctacgat tgatatgggc ccgcagttga aggtggtgga acggacccgc
 781 accgccacca tgctgtgtgc agccagcggt aatccggatc cagaaatcac ttggtttaaa
 841 gatttcttac ctgttgacac aagcaacaac aatggtcgta ttaagcagtt acgatcagaa
 901 tctattggag ccctgcagat cgaacagagc gaagaatccg accaaggaaa atacgagtgt
 961 gttgccacca acagcgcggg cactcgctac tctgcccctg ccaatttata t
```

FIGURE 24

```
   1 gctgccggga cgggtccaag atggacggcc gctcaggttc tgcttttacc tgcggcccag
  61 agccccattc attgccccgg tgctgagcgg cgccgcgagt cggcccgagg cctccgggga
 121 ctgccgtgcc gggcgggaga ccgccatggc gaccctggaa aagctgatga aggccttcga
 181 gtccctcaag tccttccagc agcagcagca gcagcagcag cagcagcagc agcagcagca
 241 gcagcagcag cagcagcagc aacagccgcc accgccgccg ccgccgccgc cgcctcctca
 301 gcttcctcag ccgccgccgc aggcacagcc gctgctgcct cagccgcagc cgcccccgcc
 361 gccgcccccg ccgccacccg gcccggctgt ggctgaggag ccgctgcacc gaccaaagaa
 421 agaactttca gctaccaaga aagaccgtgt gaatcattgt ctgacaatat gtgaaaacat
 481 agtggcacag tctgtcagaa attctccaga atttcagaaa cttctgggca tcgctatgga
 541 acttttttctg ctgtgcagtg atgacgcaga gtcagatgtc aggatggtgg ctgacgaatg
 601 cctcaacaaa gttatcaaag ctttgatgga ttctaatctt ccaaggttac agctcgagct
 661 ctataaggaa attaaaaaga atggtgcccc tcggagtttg cgtgctgccc tgtggaggtt
 721 tgctgagctg gctcacctgg ttcggcctca gaaatgcagg ccttacctgg tgaaccttct
 781 gccgtgcctg actcgaacaa gcaagagacc cgaagaatca gtccaggaga ccttggctgc
 841 agctgttccc aaaattatgg cttcttttgg caattttgca aatgacaatg aaattaaggt
 901 tttgttaaag gccttcatag cgaacctgaa gtcaagctcc cccaccattc ggcggacagc
 961 ggctggatca gcagtgagca tctgccagca ctcaagaagg acacaatatt tctatagttg
1021 gctactaaat gtgctcttag gcttactcgt tcctgtcgag gatgaacact ccactctgct
1081 gattcttggc gtgctgctca ccctgaggta tttggtgccc ttgctgcagc agcaggtcaa
1141 ggacacaagc ctgaaaggca gcttcggagt gacaaggaaa gaaatggaag tctctccttc
1201 tgcagagcag cttgtccagg tttatgaact gacgttacat catacacagc accaagacca
1261 caatgttgtg accggagccc tggagctgtt gcagcagctc ttcagaacgc ctccacccga
```

FIGURE 25

1321 gcttctgcaa accctgaccg cagtcggggg cattgggcag ctcaccgctg ctaaggagga

1381 gtctggtggc cgaagccgta gtgggagtat tgtggaactt atagctggag ggggttcctc

1441 atgcagccct gtcctttcaa gaaaacaaaa aggcaaagtg ctcttaggag aagaagaagc

1501 cttggaggat gactctgaat cgagatcgga tgtcagcagc tctgccttaa cagcctcagt

1561 gaaggatgag atcagtggag agctggctgc ttcttcaggg gtttccactc cagggtcagc

1621 aggtcatgac atcatcacag aacagccacg gtcacagcac acactgcagg cggactcagt

1681 ggatctggcc agctgtgact tgacaagctc tgccactgat ggggatgagg aggatatctt

1741 gagccacagc tccagccagg tcagcgccgt cccatctgac cctgccatgg acctgaatga

1801 tgggacccag gcctcgtcgc ccatcagcga cagctcccag accaccaccg aagggcctga

1861 ttcagctgtt acccccttcag acagttctga aattgtgtta gacggtaccg acaaccagta

1921 tttgggcctg cagattggac agccccagga tgaagatgag gaagccacag gtattcttcc

1981 tgatgaagcc tcggaggcct tcaggaactc ttccatggcc cttcaacagg cacatttatt

2041 gaaaaacatg agtcactgca ggcagccttc tgacagcagt gttgataaat ttgtgttgag

2101 agatgaagct actgaaccgg gtgatcaaga aaacaagcct tgccgcatca aaggtgacat

2161 tggacagtcc actgatgatg actctgcacc tcttgtccat tgtgtccgcc ttttatctgc

2221 ttcgtttttg ctaacagggg gaaaaaatgt gctggttccg gacagggatg tgagggtcag

2281 cgtgaaggcc ctggccctca gctgtgtggg agcagctgtg gccctccacc cggaatcttt

2341 cttcagcaaa ctctataaag ttcctcttga caccacggaa taccctgagg aacagtatgt

2401 ctcagacatc ttgaactaca tcgatcatgg agacccacag gttcgaggag ccactgccat

2461 tctctgtggg accctcatct gctccatcct cagcaggtcc cgcttccacg tgggagattg

2521 gatgggcacc attagaaccc tcacaggaaa tacattttct ttggcggatt gcattccttt

2581 gctgcggaaa acactgaagg atgagtcttc tgttacttgc aagttagctt gtacagctgt

FIGURE 25 Continued

2641 gaggaactgt gtcatgagtc tctgcagcag cagctacagt gagttaggac tgcagctgat

2701 catcgatgtg ctgactctga ggaacagttc ctattggctg gtgaggacag agcttctgga

2761 aacccttgca gagattgact tcaggctggt gagctttttg gaggcaaaag cagaaaactt

2821 acacagaggg gctcatcatt atacagggct tttaaaactg caagaacgag tgctcaataa

2881 tgttgtcatc catttgcttg gagatgaaga ccccagggtg cgacatgttg ccgcagcatc

2941 actaattagg cttgtcccaa agctgtttta taaatgtgac caaggacaag ctgatccagt

3001 agtggccgtg gcaagagatc aaagcagtgt ttacctgaaa cttctcatgc atgagacgca

3061 gcctccatct catttctccg tcagcacaat aaccagaata tatagaggct ataacctact

3121 accaagcata acagacgtca ctatggaaaa taacctttca agagttattg cagcagtttc

3181 tcatgaacta atcacatcaa ccaccagagc actcacattt ggatgctgtg aagctttgtg

3241 tcttctttcc actgccttcc cagtttgcat ttggagttta ggttggcact gtggagtgcc

3301 tccactgagt gcctcagatg agtctaggaa gagctgtacc gttgggatgg ccacaatgat

3361 tctgaccctg ctctcgtcag cttggttccc attggatctc tcagcccatc aagatgcttt

3421 gattttggcc ggaaacttgc ttgcagccag tgctcccaaa tctctgagaa gttcatgggc

3481 ctctgaagaa gaagccaacc cagcagccac caagcaagag gaggtctggc cagccctggg

3541 ggaccgggcc ctggtgccca tggtggagca gctcttctct cacctgctga aggtgattaa

3601 catttgtgcc cacgtcctgg atgacgtggc tcctggaccc gcaataaagg cagccttgcc

3661 ttctctaaca aacccccctt ctctaagtcc catccgacga aaggggaagg agaaagaacc

3721 aggagaacaa gcatctgtac cgttgagtcc caagaaaggc agtgaggcca gtgcagcttc

3781 tagacaatct gatacctcag gtcctgttac aacaagtaaa tcctcatcac tggggagttt

3841 ctatcatctt ccttcatacc tcaaactgca tgatgtcctg aaagctacac acgctaacta

3901 caaggtcacg ctggatcttc agaacagcac ggaaaagttt ggagggtttc tccgctcagc

FIGURE 25 Continued

3961 cttggatgtt ctttctcaga tactagagct ggccacactg caggacattg ggaagtgtgt

4021 tgaagagatc ctaggatacc tgaaatcctg ctttagtcga gaaccaatga tggcaactgt

4081 ttgtgttcaa caattgttga agactctctt tggcacaaac ttggcctccc agtttgatgg

4141 cttatcttcc aaccccagca agtcacaagg ccgagcacag cgccttggct cctccagtgt

4201 gaggccaggc ttgtaccact actgcttcat ggccccgtac acccacttca cccaggccct

4261 cgctgacgcc agcctgagga acatggtgca ggcggagcag gagaacgaca cctcgggatg

4321 gtttgatgtc ctccagaaag tgtctaccca gttgaagaca aacctcacga gtgtcacaaa

4381 gaaccgtgca gataagaatg ctattcataa tcacattcgt ttgtttgaac ctcttgttat

4441 aaaagcttta aaacagtaca cgactacaac atgtgtgcag ttacagaagc aggttttaga

4501 tttgctggcg cagctggttc agttacgggt taattactgt cttctggatt cagatcaggt

4561 gtttattggc tttgtattga aacagtttga atacattgaa gtgggccagt tcagggaatc

4621 agaggcaatc attccaaaca tctttttctt cttggtatta ctatcttatg aacgctatca

4681 ttcaaaacag atcattggaa ttcctaaaat cattcagctc tgtgatggca tcatggccag

4741 tggaaggaag gctgtgacac atgccatacc ggctctgcag cccatagtcc acgacctctt

4801 tgtattaaga ggaacaaata aagctgatgc aggaaaagag cttgaaaccc aaaaagaggt

4861 ggtggtgtca atgttactga gactcatcca gtaccatcag gtgttggaga tgttcattct

4921 tgtcctgcag cagtgccaca aggagaatga agacaagtgg aagcgactgt ctcgacagat

4981 agctgacatc atcctcccaa tgttagccaa acagcagatg cacattgact ctcatgaagc

5041 ccttggagtg ttaaatacat tatttgagat tttggccccc tcctccctcc gtccggtaga

5101 catgctttta cggagtatgt tcgtcactcc aaacacaatg gcgtccgtga gcactgttca

5161 actgtggata tcgggaattc tggccatttt gagggttctg atttcccagt caactgaaga

FIGURE 25 Continued

5221 tattgttctt tctcgtattc aggagctctc cttctctccg tatttaatct cctgtacagt

5281 aattaatagg ttaagagatg gggacagtac ttcaacgcta gaagaacaca gtgaagggaa

5341 acaaataaag aatttgccag aagaaacatt ttcaaggttt ctattacaac tggttggtat

5401 tcttttagaa gacattgtta caaaacagct gaaggtggaa atgagtgagc agcaacatac

5461 tttctattgc caggaactag gcacactgct aatgtgtctg atccacatct tcaagtctgg

5521 aatgttccgg agaatcacag cagctgccac taggctgttc cgcagtgatg gctgtggcgg

5581 cagtttctac accctggaca gcttgaactt gcgggctcgt tccatgatca ccacccaccc

5641 ggccctggtg ctgctctggt gtcagatact gctgcttgtc aaccacaccg actaccgctg

5701 gtgggcagaa gtgcagcaga ccccgaaaag acacagtctg tccagcacaa agttacttag

5761 tccccagatg tctggagaag aggaggattc tgacttggca gccaaacttg gaatgtgcaa

5821 tagagaaata gtacgaagag gggctctcat tctcttctgt gattatgtct gtcagaacct

5881 ccatgactcc gagcacttaa cgtggctcat tgtaaatcac attcaagatc tgatcagcct

5941 ttcccacgag cctccagtac aggacttcat cagtgccgtt catcggaact ctgctgccag

6001 cggcctgttc atccaggcaa ttcagtctcg ttgtgaaaac ctttcaactc caaccatgct

6061 gaagaaaact cttcagtgct tggaggggat ccatctcagc cagtcgggag ctgtgctcac

6121 gctgtatgtg gacaggcttc tgtgcacccc tttccgtgtg ctggctcgca tggtcgacat

6181 ccttgcttgt cgccgggtag aaatgcttct ggctgcaaat ttacagagca gcatggccca

6241 gttgccaatg gaagaactca acagaatcca ggaatacctt cagagcagcg ggctcgctca

6301 gagacaccaa aggctctatt ccctgctgga caggtttcgt ctctccacca tgcaagactc

6361 acttagtccc tctcctccag tctcttccca cccgctggac ggggatgggc acgtgtcact

6421 ggaaacagtg agtccggaca aagactggta cgttcatctt gtcaaatccc agtgttggac

6481 caggtcagat tctgcactgc tggaaggtgc agagctggtg aatcggattc ctgctgaaga

FIGURE 25 Continued

6541 tatgaatgcc ttcatgatga actcggagtt caacctaagc ctgctagctc catgcttaag

6601 cctagggatg agtgaaattt ctggtggcca gaagagtgcc ctttttgaag cagcccgtga

6661 ggtgactctg gcccgtgtga gcggcaccgt gcagcagctc cctgctgtcc atcatgtctt

6721 ccagcccgag ctgcctgcag agccggcggc ctactggagc aagttgaatg atctgtttgg

6781 ggatgctgca ctgtatcagt ccctgcccac tctggcccgg gccctggcac agtacctggt

6841 ggtggtctcc aaactgccca gtcatttgca ccttcctcct gagaaagaga aggacattgt

6901 gaaattcgtg gtggcaaccc ttgaggccct gtcctggcat ttgatccatg agcagatccc

6961 gctgagtctg gatctccagg cagggctgga ctgctgctgc ctggccctgc agctgcctgg

7021 cctctggagc gtggtctcct ccacagagtt tgtgacccac gcctgctccc tcatctactg

7081 tgtgcacttc atcctggagg ccgttgcagt gcagcctgga gagcagcttc ttagtccaga

7141 aagaaggaca aataccccaa aagccatcag cgaggaggag gaggaagtag atccaaacac

7201 acagaatcct aagtatatca ctgcagcctg tgagatggtg gcagaaatgg tggagtctct

7261 gcagtcggtg ttggccttgg gtcataaaag gaatagcggc gtgccggcgt ttctcacgcc

7321 attgctaagg aacatcatca tcagcctggc ccgcctgccc cttgtcaaca gctacacacg

7381 tgtgccccca ctggtgtgga agcttggatg gtcacccaaa ccgggagggg attttggcac

7441 agcattccct gagatccccg tggagttcct ccaggaaaag gaagtcttta aggagttcat

7501 ctaccgcatc aacacactag gctggaccag tcgtactcag tttgaagaaa cttgggccac

7561 cctccttggt gtcctggtga cgcagccccc cgtgatggag caggaggaga gcccaccaga

7621 agaagacaca gagaggaccc agatcaacgt cctggccgtg caggccatca cctcactggt

7681 gctcagtgca atgactgtgc ctgtggccgg caacccagct gtaagctgct tggagcagca

7741 gccccggaac aagcctctga aagctctcga caccaggttt gggaggaagc tgagcattat

7801 cagagggatt gtggagcaag agattcaagc aatggtttca aagagagaga atattgccac

FIGURE 25 Continued

7861 ccatcattta tatcaggcat gggatcctgt cccttctctg tctccggcta ctacaggtgc

7921 cctcatcagc cacgagaagc tgctgctaca gatcaacccc gagcgggagc tggggagcat

7981 gagctacaaa ctcggccagg tgtccataca ctccgtgtgg ctggggaaca gcatcacacc

8041 cctgagggag gaggaatggg acgaggaaga ggaggaggag gccgacgccc ctgcaccttc

8101 gtcaccaccc acgtctccag tcaactccag gaaacaccgg gctggagttg acatccactc

8161 ctgttcgcag tttttgcttg agttgtacag ccgctggatc ctgccgtcca gctcagccag

8221 gaggaccccg gccatcctga tcagtgaggt ggtcagatcc cttctagtgg tctcagactt

8281 gttcaccgag cgcaaccagt ttgagctgat gtatgtgacg ctgacagaac tgcgaagggt

8341 gcacccttca gaagacgaga tcctcgctca gtacctggtg cctgccacct gcaaggcagc

8401 tgccgtcctt gggatggaca aggccgtggc ggagcctgtc agccgcctgc tggagagcac

8461 gctcaggagc agccacctgc ccagcagggt tggagccctg cacggcgtcc tctatgtgct

8521 ggagtgcgac ctgctggacg acactgccaa gcagctcatc ccggtcatca gcgactatct

8581 cctctccaac ctgaaaggga tcgcccactg cgtgaacatt cacagccagc agcacgtact

8641 ggtcatgtgt gccactgcgt tttacctcat tgagaactat cctctggacg tagggccgga

8701 attttcagca tcaataatac agatgtgtgg ggtgatgctg tctggaagtg aggagtccac

8761 cccctccatc atttaccact gtgccctcag aggcctggag cgcctcctgc tctctgagca

8821 gctctcccgc ctggatgcag aatcgctggt caagctgagt gtggacagag tgaacgtgca

8881 cagcccgcac cgggccatgg cggctctggg cctgatgctc acctgcatgt acacaggaaa

8941 ggagaaagtc agtccgggta gaacttcaga ccctaatcct gcagcccccg acagcgagtc

9001 agtgattgtt gctatggagc gggtatctgt tctttttgat aggatcagga aaggctttcc

9061 ttgtgaagcc agagtggtgg ccaggatcct gccccagttt ctagacgact tcttcccacc

9121 ccaggacatc atgaacaaag tcatcggaga gtttctgtcc aaccagcagc catacccccа

9181 gttcatggcc accgtggtgt ataaggtgtt tcagactctg cacagcaccg ggcagtcgtc

FIGURE 25 Continued

```
 9241 catggtccgg gactgggtca tgctgtccct ctccaacttc acgcagaggg ccccggtcgc
 9301 catggccacg tggagcctct cctgcttctt tgtcagcgcg tccaccagcc cgtgggtcgc
 9361 ggcgatcctc ccacatgtca tcagcaggat gggcaagctg gagcaggtgg acgtgaacct
 9421 tttctgcctg gtcgccacag acttctacag acaccagata gaggaggagc tcgaccgcag
 9481 ggccttccag tctgtgcttg aggtggttgc agccccagga agcccatatc accggctgct
 9541 gacttgttta cgaaatgtcc acaaggtcac cacctgctga gcgccatggt gggagagact
 9601 gtgaggcggc agctggggcc ggagcctttg gaagtctgcg cccttgtgcc ctgcctccac
 9661 cgagccagct tggtccctat gggcttccgc acatgccgcg ggcggccagg caacgtgcgt
 9721 gtctctgcca tgtggcagaa gtgctctttg tggcagtggc caggcaggga gtgtctgcag
 9781 tcctggtggg gctgagcctg aggccttcca gaaagcagga gcagctgtgc tgcaccccat
 9841 gtgggtgacc aggtcctttc tcctgatagt cacctgctgg ttgttgccag gttgcagctg
 9901 ctcttgcatc tgggccagaa gtcctccctc ctgcaggctg gctgttggcc cctctgctgt
 9961 cctgcagtag aaggtgccgt gagcaggctt tgggaacact ggcctgggtc tccctggtgg
10021 ggtgtgcatg ccacgccccg tgtctggatg cacagatgcc atggcctgtg ctgggccagt
10081 ggctgggggt gctagacacc cggcaccatt ctcccttctc tcttttcttc tcaggattta
10141 aaatttaatt atatcagtaa agagattaat tttaacgtaa ctctttctat gcccgtgtaa
10201 agtatgtgaa tcgcaaggcc tgtgctgcat gcgacagcgt ccggggtggt ggacagggcc
10261 cccggccacg ctccctctcc tgtagccact ggcatagccc tcctgagcac ccgctgacat
10321 ttccgttgta catgttcctg tttatgcatt cacaaggtga ctgggatgta gagaggcgtt
10381 agtgggcagg tggccacagc aggactgagg acaggccccc attatcctag gggtgcgctc
10441 acctgcagcc cctcctcctc gggcacagac gactgtcgtt ctccacccac cagtcaggga
10501 cagcagcctc cctgtcactc agctgagaag gccagccctc cctggctgtg agcagcctcc
```

FIGURE 25 Continued

10561 actgtgtcca gagacatggg cctcccactc ctgttccttg ctagccctgg ggtggcgtct

10621 gcctaggagc tggctggcag gtgttgggac ctgctgctcc atggatgcat gccctaagag

10681 tgtcactgag ctgtgttttg tctgagcctc tctcggtcaa cagcaaagct tggtgtcttg

10741 gcactgttag tgacagagcc cagcatccct tctgcccccg ttccagctga catcttgcac

10801 ggtgacccct tttagtcagg agagtgcaga tctgtgctca tcggagactg ccccacggcc

10861 ctgtcagagc cgccactcct atccccaggc caggtccctg gaccagcctc ctgtttgcag

10921 gcccagagga gccaagtcat taaaatggaa gtggattctg gatggccggg ctgctgctga

10981 tgtaggagct ggatttggga gctctgcttg ccgactggct gtgagacgag gcaggggctc

11041 tgcttcctca gccctagagg cgagccaggc aaggttggcg actgtcatgt ggcttggttt

11101 ggtcatgccc gtcgatgttt tgggtattga atgtggtaag tggaggaaat gttggaactc

11161 tgtgcaggtg ctgccttgag acccccaagc ttccacctgt ccctctccta tgtggcagct

11221 ggggagcagc tgagatgtgg acttgtatgc tgcccacata cgtgaggggg agctgaaagg

11281 gagcccctcc tctgagcagc ctctgccagg cctgtatgag gcttttccca ccagctccca

11341 acagaggcct cccccagcca ggaccacctc gtcctcgtgg cggggcagca ggagcggtag

11401 aaaggggtcc gatgtttgag gaggccctta agggaagcta ctgaattata acacgtaaga

11461 aaatcaccat tccgtattgg ttgggggctc ctgtttctca tcctagcttt ttcctggaaa

11521 gcccgctaga aggtttggga acgaggggaa agttctcaga actgttggct gctccccacc

11581 cgcctcccgc ctcccccgca ggttatgtca gcagctctga gacagcagta tcacaggcca

11641 gatgttgttc ctggctagat gtttacattt gtaagaaata acactgtgaa tgtaaaacag

11701 agccattccc ttggaatgca tatcgctggg ctcaacatag agtttgtctt cctcttgttt

11761 acgacgtgat ctaaaccagt ccttagcaag gggctcagaa caccccgctc tggcagtagg

FIGURE 25 Continued

11821 tgtcccccac ccccaaagac ctgcctgtgt gctccggaga tgaatatgag ctcattagta

11881 aaaatgactt cacccacgca tatacataaa gtatccatgc atgtgcatat agacacatct

11941 ataattttac acacacacct ctcaagacgg agatgcatgg cctctaagag tgcccgtgtc

12001 ggttcttcct ggaagttgac tttccttaga cccgccaggt caagttagcc gcgtgacgga

12061 catccaggcg tgggacgtgg tcagggcagg gctcattcat tgcccactag gatcccactg

12121 gcgaagatgg tctccatatc agctctctgc agaagggagg aagactttat catgttccta

12181 aaaatctgtg gcaagcaccc atcgtattat ccaaattttg ttgcaaatgt gattaatttg

12241 gttgtcaagt tttgggggtg ggctgtgggg agattgcttt tgttttcctg ctggtaatat

12301 cgggaaagat tttaatgaaa ccagggtaga attgtttggc aatgcactga agcgtgtttc

12361 tttcccaaaa tgtgcctccc ttccgctgcg ggcccagctg agtctatgta ggtgatgttt

12421 ccagctgcca agtgctcttt gttactgtcc accctcattt ctgccagcgc atgtgtcctt

12481 tcaaggggaa aatgtgaagc tgaaccccct ccagacaccc agaatgtagc atctgagaag

12541 gccctgtgcc ctaaaggaca cccctcgccc ccatcttcat ggaggggggtc atttcagagc

12601 cctcggagcc aatgaacagc tcctcctctt ggagctgaga tgagccccac gtggagctcg

12661 ggacggatag tagacagcaa taactcggtg tgtggccgcc tggcaggtgg aacttcctcc

12721 cgttgcgggg tggagtgagg ttagttctgt gtgtctggtg ggtggagtca ggcttctctt

12781 gctacctgtg agcatccttc ccagcagaca tcctcatcgg gctttgtccc tcccccgctt

12841 cctccctctg cggggaggac ccgggaccac agctgctggc cagggtagac ttggagctgt

12901 cctccagagg ggtcacgtgt aggagtgaga agaaggaaga tcttgagagc tgctgaggga

12961 ccttggagag ctcaggatgg ctcagacgag gacactcgct tgccgggcct gggcctcctg

13021 ggaaggaggg agctgctcag aatgccgcat gacaactgaa ggcaacctgg aaggttcagg

13081 ggccgctctt cccccatgtg cctgtcacgc tctggtgcag tcaaaggaac gccttccccct

FIGURE 25 Continued

```
13141 cagttgtttc taagagcaga gtctcccgct gcaatctggg tggtaactgc cagccttgga
13201 ggatcgtggc caacgtggac ctgcctacgg agggtgggct ctgacccaag tggggcctcc
13261 ttgtccaggt ctcactgctt tgcaccgtgg tcagagggac tgtcagctga gcttgagctc
13321 ccctggagcc agcagggctg tgatgggcga gtcccggagc cccacccaga cctgaatgct
13381 tctgagagca aagggaagga ctgacgagag atgtatattt aatttttaa ctgctgcaaa
13441 cattgtacat ccaaattaaa ggaaaaaaat ggaaaccatc a
```

FIGURE 25 Continued

```
   1 atccacccgc ctctgcctcc caaagtgcta ggattacagg catgagccac catgtctggc
  61 caggaaaaat gggaagtttt aaatgctttt ccagtagcac tggagacagg gtgagatgtc
 121 tgctctcatc atttatattg tcacggtgct gggtttctat acagtcctct caggcaagaa
 181 gagaaataaa aggtggctgg gcgtggtgac tcaccctgta atcccagcac tttgggaggc
 241 cgaggtggga ggatcccttc agctcaggag tttgagacct tcaagaccag cctggacaac
 301 acagtgagat cccatctcta aaaaaaaaaa aaatacaaaa attagccggg tatggtggct
 361 tgcacctgta gtcccagata cttaggaggc tgatgtggga ggctcacctg agcctggggg
 421 gttggaactg cagttagctg agatcatgcc actgcactcc agcctgggtg atagagcaag
 481 acgctgtctc aaaaaaaaaa aaaaaaaaaa aaaggctggg cacaatgact cacacctgta
 541 atccctgcac tttgggaggc caaggcgggc agatcaactg aggtcaggag ttcaagacca
 601 acctggccaa aatggtgaaa ccctgtcttt actaaaaata caaaaattag ccgggcgtga
 661 tggcgggcac ctgtaattcc agctactcag gaggctaagg caggagaatc actcgaaccc
 721 gggaggtgga ggttgcagtg agccaagatt gcaccactgc actccagctt gggtgacagg
 781 gtgagatcct gtctaaaaaa aaaaaaaaaa atccagtcag gaacaggaat gctaaggaat
 841 caagagcagc aacaacaagc tagcagaagt aacaactaac gaacaagttc ataaagatca
 901 cagggaggac cagaagacct gcattcactg tcaattttat ttatacttat taacaataaa
 961 tggtctgaaa ataaaattaa gaagacagct tcattcacaa ttgcatagaa aaataaaata
1021 ctttagatta aaattaagaa aataaatgca aaccttgtac attaaaatct acaaaacatt
1081 gctgagagaa agcaaagaca gcctaaaaaa atggggagtg atcctatgtt cacagattag
1141 aagattcaat attgtggccg cgtgtggtga ctcatgcctg aaatcccagc actttgggag
1201 gccgggggcg gggggggggg tggattatga ggtcaggagt tcgagaccag cctggccaag
1261 agaccagcct ggccaatatg gtgaaaccct gtctccacta aaaatacaaa aattagccgg
1321 gcatgatggt gggcgcctat agtcccagct actcgggagg ctgaggaagg agaatggcgt
1381 gaacccagga agcagagctt gcagtgagcc gagattgcag cactgcactc cagcctgggc
1441 aacagagtga gactctgtct caaaaaaaaa tatatatata tatatattat ttttatttt
```

FIGURE 26

1501 agtagagaca gggtcttacc atgttggtca ggctggtctc gaactcctga cctcaggtga

1561 tcctcccatc tcggcctccc aaagtgctgg gatgacaggt gggagccgcc gcgcccggct

1621 gggaggtgtt ctttctagac ctcacctggg agtcacgcac cattacctct accacgttcc

1681 atttgtaagt gcaggccatg tatgcctgga gggaaatcaa tcttctgccg aacagggtgg

1741 tgttcaaagc accaccggct ccaccacact ctgcctttta ttctgcattc tgtttcttga

1801 gaccacgcgc cgctacgtgg aatggtctca ggagctcatg tgtttgtgct ccatgaagtc

1861 agaaagtcca gcctttgcac tgccacatac ccacccttag aacagcgtct ggtacacggt

1921 aggtgctcag tgaatgtgcc tagtggagta aacgtgcggt gcggtgctgc ctgcggtcgg

1981 atctgtgggg atcctgtgat ggggaagacg gctaccagga aaggtggaat ttggaaagtc

2041 aggacaggga gagaggaggg actttgtgga ggcaggaagg tatgggagcc agtccagcac

2101 aagggctggc gggaaagcct cgtgggtgtc agagactata cctctgtgag ggtccctggg

2161 ctcccccagg tcagggcaga ggtcctgacc ccagtctagc tctttagcgg gggccaggcc

2221 cgagatggcc agctccgttt cccctgcgtg gcctggcagc ccctccagga gctggcacgg

2281 gaagcaggcc tgtcctccac gcccttggcc ggccttggct gcgatgggtg gagacgcttc

2341 tgcccgcccc acccctgtct gtctgcctcc tcctctcaag caggcttcct ctcgacagaa

2401 atgtgtctgc aggtggctct tgagccccag ttctcaggga ttcacctgca gaggatgtag

2461 agccggggtc ccttccctgc ttctgatctg cccaaaatcc tagggagagc cctgattggc

2521 aacccctcgg ccaatggact gaggccaggg cgcgtctgca cccgggttga cagccctgac

2581 tagaaccgcg aggctggggt cagggaggag cagccgcctc ccgccggggc cgcagctgtc

2641 agagtgggca aggggatctc cacagaggga gactgtccca cagtggcccg gggtagaaca

2701 gcgcttccct ccccacccag cgcctgactg ctggccccgg gcccatccca gggcagtgcc

2761 ccaggtctct ggtcctcccc acagcctttg tctcaccagg tctcagaggg gtccgagtcc

2821 aggctcatct gttgtggtac cactgccccc tccctcatgg ttgtggaaga agcccacggg

2881 gggaaaatgc agagcagaat ccaggaaaga ctcgctcaca ggaggagctg tggcctggag

2941 acgcggcctc caccccatct ccatcatgga caacctgcca aggggtcttg aacagaagct

FIGURE 26 Continued

3001 tggggacgca catacggtcc ggtggggcag tgacccactg gcagaccagc tccctgagta

3061 gaaaagaaaa agccctggtt tgtggggttt gatgattccg tggcagcaga gataatccca

3121 ctgtggccaa tttcgaagcc tgactgtaac agctgctggc aggtgtctga acatttagcg

3181 atcagctctc cagggtcagc gggagtcgct ccagcacccc acggggtggg ggtgggtggg

3241 tggagccagg gcctctgctg cagcctccaa ccccttttctt ccacgctgat gaggcgctgt

3301 ctgggctcag gtctaccccc acgcagctgc caggctctgg cctgcctgag gcggatactt

3361 ttttttttct ttttttgat ggagtgttgc tctgtcgccc gggctggagt gcaatggtgc

3421 gatcttggct cactgcaaca tccgcctccc agattcaagc gattctcctg cctcagtctc

3481 ccaagtagct gggattacaa atgcctgccg ccacgcccgg ctaatttttg tatttttagt

3541 agagacaggg tttcaccatg ttggccaggc tggtctcgaa ctcctgacct taggtgatcc

3601 acctgccctt gcctcccaaa gtattgagat tacaggcatg agccactgtg ccctgccgag

3661 gtggaagttt ggagatgggc cacaaactcc tggagatagg gccagctcag tccccctgag

3721 cacccacaca caccctcctg agagccaaca gaaggagtga gggccccgag gcggatgacc

3781 cgtgtgcctc aacccgaacg gggagaggcg gggtctgcag cagggtacgg gcaggtgatc

3841 ccccaaggaa agatttttcct gtattgagag agaaggggcc aagaggagga gcttgtcaaa

3901 caccacagcc cctccccctc ctctcagctc caggggtcc ctggtgccag tgttcggctg

3961 atggagagaa cggcaagcgg gagagagagt gtgacccctg tgggcacatg acttcccttg

4021 ctgcactgct gcacatagca gaggtgtggt gacgaccctg ttttgtccca ttgggggcgt

4081 ttgctgttag gtctgcagaa tcctcagttg ctattggaaa tggtgacatc actggcaggg

4141 gcggagcttc agccatcctt caagttaggg aggggcacgc acactccagg ggtggagggg

4201 gacaaagaca gggtggtgtg gaccagaggg atgggtaagg ctctggaaaa gggggcgctg

4261 ggagcgcatt gcgaggggggc tggagaggga gagaggagcg gaagctgagg gtgtgaaacg

4321 gctggccccg aacacacctc gcggcgctcc agtgattcct ggtgtccgac ctcagcccca

4381 gtcagtgcgg gtccagtttc caggctctcg cggaaggcct ggctgagcac atgcggcagc

4441 cacggtcacc ctccctattc ctcttagccc gaggaggggg gtcccaagtt acatggccac

4501 gcagatgggg cctctccctc atttctgaac cttgtgggga ggggaacctt gaagggagcg

FIGURE 26 Continued

4561 ccccccagag ccatggctta gggcctcccc cacccctctg gagctccagt ctgcaagagt

4621 caggagccga aatatcgctg actgtgggtg acgactcttg cgcgcacaca cacatacaag

4681 cgggcacgac gcgttcggtc ctattaaaag gcacgcaagg gtgcgggctg cacgcggtga

4741 cacggacccc tctaacgttt ccaaactgag ctccctgcag gtccccgaca gcacaggccc

4801 ctgtcccagg acccctccag gcacgcgctc acacgcacac gcgcgctccc cggctcacgc

4861 gcgctccgac acacacgctc acgcgaacgc aggcgcacgc tctggcgcgg gaggcgcccc

4921 cttcgcctcc gtgttgggaa gcgggggcgg cgggaggggc aggagacgtt ggccccgctc

4981 gcgtttctgc agctgctgca gtcgccgcag cgtccggacc ggaaccagcg ccgtccgcgg

5041 agccgccgcc gccgccgccg ggccctttcc aagccgggcg ctcggagctg tgcccggccc

5101 cgcttcagca ccgcggacag cgccggccgc gtggggctga gccccgagcc cccgcgcacg

5161 cttcagcgcc ccttccctcg gccgacgtcc cgggaccgcc gctccggggg agacgtggcg

5221 tccgcagccc gcggggccgg gcgagcgcag gacggcccgg aagccccgcg ggggatgcgc

5281 cgagggcccc gcgttcgcgc cgcgcagagc caggcccgcg gcccgagccc atgagcacca

5341 tgcgcctgct gacgctcgcc ctgctgttct cctgctccgt cgcccgtgcc gcgtgcgacc

5401 ccaagatcgt caacattggc gcggtgctga gcacgcggaa gcacgagcag atgttccgcg

5461 aggccgtgaa ccaggccaac aagcggcacg gctcctggaa gattcagctc aatgccacct

5521 ccgtcacgca caagcccaac gccatccaga tggctctgtc ggtgtgcgag gacctcatct

5581 ccagccaggt gccctccccc acctccgcca cccacctccc ctctcctcca tcctgcaacc

5641 ccacaccccc agtttcattc catcctttcc gtgccccctt cctccctgta agacaccacc

5701 ccagagtcag ctggctgctt ccgggaggcc tcgtctcact aggaaccaaa caccagggtc

5761 tgctggctcc cctatcttgg cctgagacca gtcacctgcc accttggctg gtcctcagag

5821 ggcccctggg gctccaggcc ctgactggtg tgtgtagacg tggggctgga gtgtgtcagt

5881 gtgggggtgg gcattccggg taagagagta gaagcgcctg tccagctaca tgcccgccct

5941 gcagagcttt aaacaggacg gggcctgggg ccatctttgt ttctgcttcc aggttctcct

6001 gccctttctt tcgtcccttc cccctaccga tgggtccgcc tgggaagaga aatggctcag

6061 gtgccacggc aggacgcttt gtgggggtgg gagtgggggt gcacacgcga gaggcatcag

FIGURE 26 Continued

6121 ggcatgggag ctgtcggcag ccagcgctgc gggggaggac gtggctcctg ggattttgcc

6181 tgtcggagct gtccgccccc gggccgagcg cctgctgaat tccaatgagg ctgcaaggat

6241 ctgcaatgca gcccttttatg taagaggcaa gacagacatc cagcctagca ccgctcacac

6301 gtgcctacct gatggacaca ccacatctgt ggacacacat gctcacactc acaccaaatg

6361 ttacattagc acacactcat gcacctcagc atcacacaat caatttcata tgctcatctg

6421 cacacatgca gatccattga cacctgctca tgtgccacac acggcttggc atgcattccc

6481 agaggcacgt gcaaacatgc acatttacac acatggttcc agtcattcac acgcatgtac

6541 acgaacagac atgccagggc atgtgatgca cataaccata ccctagcaca cgcgtgaaca

6601 cctgcatggt cacacacgga cctacgggtc tttgccaagc acctctgggt gcaggctgga

6661 agcaagagct gggggaggga gaaccacttc aaacagctgc agctgcaggg cccacaccag

6721 agttttctca gaaatcctcc ctccccactt cacaagccac ccccgtgccc cagcccagga

6781 caccatggga tgggactggg gggatgcatc tgtagccagt ggctgcagtc acatattcca

6841 tctgggactg gggagggaca cggaaggtgg actcaggaaa tccaggaggg gccattcctg

6901 gggaattgct tcaactcacg cccatgttgc tgtctgtctg tgggcatggc ctctgcagca

6961 aaggcaatgc ctgcagctac cactcacgga acacaccccc ggccaggtac tgtcctgcca

7021 cgtggggcca tgcagtgcac gcccccattc gccaaagctc taagaggcac aggcagactt

7081 ggggacagac gcaggtcctt gctgtgtgaa ggtggtgtgg accaccagct gcctgcctcc

7141 ctgccttggg aggctgggga gagagggagg catcagctcc agggggctga gccgctggct

7201 ttagatctgc cccatggggc cttggtcatg ggcaggaagg ctgggctgca cccccaatgc

7261 ctcccttccc ttccttgagg atgaggccag cactcaaagt aagggcttgg tgttgttcag

7321 acagagcccg tcacaggccc tgcccctgga gacaccagca aaagggatct cggcctcttt

7381 ggcagctcct agctgcttcc ccctgaagtc cggtaccacc cttcagagct gcccgccctg

7441 tctcgggatg tgggctggcc ccacccctgg cccatcagga aggacgggtg ggttctgaga

7501 atcaaggcca tcatgatgca ggaccagcca tcctccccgg tccaccttgg gtgcgtcccg

7561 tgctccaggc ccccaggaca tcccaagggc agtccttcac ctggcccttg agcacaacac

7621 ctgcagggcc ctagtcagca gtgtgagagg agtgagggga ggtccgggtg gggtctccct

FIGURE 26 Continued

7681 cccctgccct gtgggcatgt gtgcatctgg gcctgggcat gtagcatgta cccgaatcat

7741 gcccccagcc ccccttagcc tgctgggttc agcccctgct gcttccagat ctcagcctct

7801 aacccagtgc ctggctcacc cctgactcaa gctgatcatg tctcctgtgt ccacaggtct

7861 acgccatcct agttagccat ccacctaccc ccaacgacca cttcactccc acccctgtct

7921 cctacacagc cggcttctac cgcatacccg tgctggggct gaccacccgc atgtccatct

7981 actcggacaa ggtaagcctg actgccagac caggccttcc ggccctcggc cccagggcac

8041 agcctggcca ctccaggagc agcgggccga cccgctcaca tggaactcac acaccacaaa

8101 cagccacaca gctcccccac attcatgcac gtccacacgc tctcacgtgt ccaactcaca

8161 catctgcaaa catgctcaca tgcacactca tgtgctctta cacacacaat acacactctc

8221 ttgcacatag agggctcacg tggagcccag cacgtgcccc cagcccagag caggccaaag

8281 ggaggggggca cacatcacac actcacacat cacacacaca tcacacactc acacattcat

8341 acagcaccca cacgctacac tgctatgctc accctcccca cacatgaaca ctgacacacc

8401 catggattcg cacaaagtca cacacactca ctggcacagg caccagtgac acccccctcag

8461 gacgagaggg cccgtgggct aggagaaggg atggctggga ggctttctag acaggtggac

8521 tttgaagggg agtttggaga gctgggggtt gctccaggag gaaaggggtg tgcacgcagc

8581 cagggtggtg gggccagcct ttcccactgc aggcatgggt ggagagcaat gtctgtggtt

8641 gcagctcagg gtcgggggcg ctggcctggg ggcttccagc ctctagggct gagggcacct

8701 tggcttagcc tcctgcagac cctcctggcc cacaggctat gaggagggct tctgtccaat

8761 cctggaacat cagctggaag agaggagggt catccagtca gttttgcagg aatctccaag

8821 ccagagagcc atgggggctt gctctaggtc acacagccct ccgtctaccc aggatgcaaa

8881 ctgggcactg agaggctgac caacctgggg ccactggcag acagacctgc agggccactt

8941 ggcaggggac atccagtttg gtgccagcgc tgaggagcca gagggctggg ctgtgcagcc

9001 aggcttctgg gtcccccacc ttctccaaat tcctcctgcc cagagtccac agtccttggt

9061 aacactgcct taaagcacag gggtcgcccc agccaggccc aggctcttct gggaggatgg

9121 aaggccccag aggcaggaac tgagacagag gctggaacag ccaccttcct gaggctctga

9181 aagccctggc gtgccccctc ggcacccaaa ctgctcctcc caggtgactt cacctctggc

FIGURE 26 Continued

```
 9241 catgggaagg tgggggtctc attcctgtgc cctcaggcac gacctccttc gcctctctgg
 9301 gcaccagttt cctcacatgt gaaggagaga agatggcctt acccaggaaa gcagccagtg
 9361 gtcaaatgaa aggcggggga aacggatgcc cctggcccag agcaggccaa agggaggggg
 9421 cacagctcac agctgcaccc tggccttacc acagtctcta cactacaacc aacctgtgcc
 9481 caaaacatga accggcaagg ccaggtcaga gctagtccaa gacctcaagc acagcctgcc
 9541 ttgccaccac gtcaccaggt ggatagacag aagcagggga cattttgca ccccaaggca
 9601 ctgccccagg ccacaaagag ggagcaggtg aaaaataacc tggaagcctc agaggaccac
 9661 aagatcagca agagtccaca gggacactga aggaaccagg gcttacctgg acagacacag
 9721 agaactgagg cagagggggg cagagcctgc tccactcccg gccatgccac ggcactccgt
 9781 ggcagcttga agccaggaaa agcaagccag ggcaagcaag caccacgctc tcgcctgggg
 9841 agatgaggcc tttagcccca agagtgaatt cttcttcata catagagttg tttaaatttg
 9901 ggaggactct atgggcagcc ccaggggat cttcgaggcg ctatgtgtca tcaagaattt
 9961 cctgagctca gcttgtccaa aggtggtggg ctgcagggga agaggtgagc tcaccccagg
10021 cacaattcca cagaaaccca cgtcccttag ggtgctatgg ggccaacact aaacctcctc
10081 catttccgag attatatgtg ggaggagagg ccggggtggg agagaggttc ccagggtcta
10141 aaaagtgtcc ccaggatggt ggggacaggg gtgggaaaaa ggaggggtcc cagtgtctag
10201 aaagtgtccc caggttggcc gggcgcggtg gctcacgcct gtaatcccag cactttggga
10261 ggccgaggcg ggcggatcac gaggtcagga gatcgagacc atcctggcta atacggtgaa
10321 accccatctc cactaaaaat acaaaaaaat tagccgggcg tggtggcggg cgcctgtagt
10381 ctcagctact tgggaggctg aggcaggaga atggtgtgaa cccaggaggc ggagcctgca
10441 gtgagccgag attgcactcc agcctgggta acagtgcgag actgtttaaa aaaaaaaaaa
10501 agtgtcccca gggtggtggg gacaggggtg ggagacagga gggggtccca gggtctagaa
10561 agtgtcccca ggggtgtggg gacaggagtg agaggaaggg ggtcccaggg tccagaaagt
10621 gtccccaggg tggtagggac aggggtggga gacacgaggg ggtcccaggg tctagaaagt
10681 gtccccaggg tggcagggat gggatgggag acacgagggg gtcccagagt ctagaaagtg
10741 tccccagggg tgtggggacc ggggtgagag gaaggggggtc ccagggtcca gaaagtgtcc
```

FIGURE 26 Continued

10801 ccagaggggt ggggacagga gtgagaggaa gggggtccca gggtccagaa agtgtcccca

10861 gaggggtggg gacaggagtg agaggaaggg ggtcccaggg tccagaaagt gtccccagag

10921 gggtggggac cggggtgaga ggaagggggt cccagggtct agaaagtgtc cccagagggg

10981 tggggaccgg ggtgagagga agggggtccc agggtctaga aagtgtcccc aggggtgtgg

11041 ggaccggggt gagaggaagg gggtcccagg gtccagaaag tgtccccaga ggggtgggga

11101 caggagtgag aggaaggggg tcccagggtc cagaaagtgt ccccagaggg gtggggacag

11161 gagtgagagg aaggggggtcc cagggtccag aaagtgtccc cagggtggta gggacagggg

11221 tgggagacac gaggggggtcc cagggtctag aaagtgtccc cagggtggca gggatgggat

11281 gggagacacg agggggtccc agagtctaga aagtgtcccc agcggggtgg ggacaggggt

11341 gggagacacg agggggtccc agggtctaga aagtgtcccc agggtggtag ggacggggtg

11401 ggagacacga gggggtccca gggtctagaa cgtgtcccca tgggggtggg gacaggggtg

11461 ggaggagggg ggttcccagc ctccggcggg tgttccggca gtgggaggcg ggtgggaggg

11521 cgggtccccg cgggtccacc tcagcccgcc gtgcccccgc ctcccgcaga gcatccacct

11581 gagcttcctg cgcaccgtgc cgccctactc ccaccagtcc agcgtgtggt ttgagatgat

11641 gcgtgtctac agctggaacc acatcatcct gctggtcagc gacgaccacg agggccgggc

11701 ggctcagaaa cgcctggaga cgctgctgga ggagcgtgag tccaaggtga gggtcggcgc

11761 cgcgggtggg cgcctggcgg agccgaggtg caggacgggc cgccttgtgt ctgtggctcc

11821 gtgtgtgaca ccctcttctt tccatcgtgc atggtcagca ccaccacgtc tggcgagcgc

11881 ccgccccagc ctgtcctcgg ctcatttcac tcgcttttgc cattagtcga aatctccttc

11941 gtgtcagtcc ctgcggggcg agggccagac cacctggagc tccgcaaaca cccctgcccc

12001 gcgctgccga gcgccctcgt cccctcctcc tgcccatgcc cctcgctccc tggaggcccc

12061 agccgggctt gggactcgtc acccttcccg cccaccctgt cctgagtccc cagcagcctc

12121 cctctgggca aggtctcccc tgtacacgac ccccacatac cgcccctgca ggcctgtccc

12181 ctcctggctg ggcccattcc ctgtcctccc ccgcgtggcc tcccctgaag ctctgcacct

12241 catggctcag caaagccctg tccacagaca cctgcccccc agcctggacc gcccctgtgg

12301 gctccactcc cctccttgcc cccaccacag ggctccctct gcactcctca ccaagaccca

FIGURE 26 Continued

12361 ttagccacct ggttctagga cacactgacc cccaacacag ccaggcgtcc actctgtggg

12421 gctgcagaga gatcagagct gggatttggg ggggtccgag cgaccccttt ccccttcctt

12481 accactccca tttgcagtct ggggcagagc tggttctcgg ctacagaccc ccggagccct

12541 gggtgctcag cacctgggcc agctcctgat caagcagtgg gaggaggccc aggctgagga

12601 gggccagacc tatgggtggc tgggagcatg tttcgtgtca gagctggctt catcggactt

12661 agggccaacc tagcaccccc caaggcaccc caggcccegg gagggaccag agggcatggg

12721 tcgggggtaa agccaggggc agaccagagg gtcctgggag tactgtcgtg gggggtctgc

12781 tgagtcttgg gggggagggg catgggcacc aagggcccca cccaagacag tgcccctcac

12841 cccagtgccc gacaggcccc tccccccagc gcccgacagg cccctccccc ccagcacccg

12901 acaggccccc caccccagcg ctcgacaggc ccctcacctt ctggcatgtc cacccacggc

12961 cacggactcc catcacacac tcacccctgc gcacccaact gcctgctttc actcacaggt

13021 gcatgcacac attcccatca cactccacac ctctggccac aggctcaggc ttgcctccat

13081 gcagctccag cgccacacac acacctggac acgtactcag gtgcgctcct cacacacaca

13141 cctggacacg cacaggtgcg ctcctcacac acacacctgg acacacgctc aggtgcgctc

13201 ctcatgtgca tgctcaccct tacttgcgcc agccagcaca gacacacatg cacacacgca

13261 cacacgtgca caggcacaca catgcacaca tgcacacgca cacacatgca cacacgcaca

13321 cacaagcacg cacacacgca catgcacaca tgcacacaca caagcatgct cagaccatct

13381 ggcccttccc caaccttcac aggcctttgt ggactaaccc tcccatgctg acacccacag

13441 gcgcatgcca cccctgcagg cgcacataac gcacacacac cctcttgggc gcatatgcga

13501 gccgaagggc gtgggcaccc cagttagtga ggatacctcg gtctcttgag aggcagaggg

13561 agaccaagga gagggagggg gagggacatg gggacaggcc ccgggggggc tgctcacctc

13621 ccactcagga ctgacacagg ttggagtggg cactgctggg gccacacagc aggtgcacgg

13681 cagggtgggg gcggcaggtg gggctccctc cgaacggtgg acgcggacag ggcctccttt

13741 tctcccgaga gcgaccgttt ccaagagcac agcttcgtgg cagggagcct ccacggcccc

13801 gcccctacga ccctcacccc cagctccacc ccggcccccа gccccgcccc gggcctcggt

13861 gtttgcgagc tccaggtagg agcccgtctg cagacgtgcc gaggaggtgg tgtgattgct

FIGURE 26 Continued

13921 ttagcgccgt cattttcaac cgtttataat cttcttctgt gtctgcatat tttctctgtg

13981 cacattattc atcagagtaa aaaaaggaac tatgaaaacc tcgaccaact gtcctatgac

14041 aacaagcgcg gacccaaggt atatatgcat ggacgtgcac gccacccacg gctagggagc

14101 cctggcctcg gcgcctcggc cactagggcc actgtctggc ccagccgccg agccgcaggc

14161 ccaagcaatg aggagaggca gccggcagca ggcaggagag ggcaggcagg agagggcagg

14221 cgtggcggcg tgggtgcctg aggcccaccc gccgtgaccc tagcctgcac cctcgaggca

14281 ggcgcggctg caggaggagc tgctctccgg gaagtgcatg cgaatctccg agaccccaag

14341 ggtttcctcg tggaacccgg gagggagccc gcccgcctgg ccacccactc gccggggccg

14401 ggctgctcct caggggcctg cggaatcaaa cctcagagga cctcccatgg ttttggaaaa

14461 gtcagcccca tctcttttcc tggttgcatt ccaaaactct tttctgtttc ccgtccgctg

14521 cacgcctcct gagtctgggt ccacttcagc ttgcatgctc agtgcaaaga tgaggacagg

14581 agtgaggtgg agagagagat ccagagagag cagagagagc acagaacgag cacaggtgag

14641 cgcgcaggct gaagacagga caggaccgga gaggcgaggc caggccaggc aaggactgag

14701 gaaggcaggc ggaggcgcag gtggcaggcg cagaccatgg cagccctagc taagctgcct

14761 cggggttccc agcggctccg gcccaactct caccccctgag gcgctatgtc ccctgcccca

14821 gccgcctgct aacactcttg ctcacaccgc aggcagagaa ggtgctgcag tttgacccag

14881 ggaccaagaa cgtgacggcc ctgctgatgg aggcgaaaga gctggaggcc cgggtcatca

14941 tcctttctgc caggtgaggc tgggcagggc cctacacact ccacacagga tggtacctga

15001 gccaagtacc cgccatctga gccagagctg ggacattgtt gggcacagtg accttcagct

15061 tccaaagcac cttcaccaag gacagccacc cccaccccca cccgcaccca cactcctatc

15121 ggcatggctg atgtgacacc ctccatctgt ccctcccttc ctgggccctt ccccattcca

15181 cagtcacatg ctgctgctgc cctgagctgg gctgtgggag agggatatgg aagagaccct

15241 gcccttggga agccccgcaa ctcaaggggg caaacctgtg caaacaaggc ccaggcctgg

15301 ggccagggac taggcccagg gaagtgatgt gcaggtgtgc caggcgaaaa ggcccacagc

15361 agggagaggg ccagatttcc caactgaagg attgcaacag ctgcagcagt agcttagggg

15421 gaaatcagtt aggtacccgg gaaatcaagc tgctctggac aggtccggtg ccacagagca

FIGURE 26 Continued

15481 accttggggt ggagcccact gtaaaaagct ccctatttgc aaatggctag gttctcccgg

15541 gaaggaaaag cctgggtgac tgggaatagg aaaagcaaga gtggggaagg gcagggtgag

15601 gagccttgcc ccatgggaca gccaaaaccc cactgtccct gacctaaaag ctctgctagg

15661 ctccaacaga gcggagcaaa acagcagtga aacacccgga ggaacacagc ccagccctcc

15721 agccctgcat atgggaaaga gccggcgaca ctctcagtcc cagggcagac caccatttcc

15781 acggtccatc aaatgaccct ctagcacgga gacagatgca gccccctcac cagggcagaa

15841 cgcagggtgg ggccagccag ggctcctcgg actagaaggc aggaatctcc ccagcccaag

15901 gtagggttgt tgcttagagc tgcccacggg gccttacatg ctcctcctgc agctgctata

15961 aaaaggcact aatcgtcccc cattacgccc cctgcactcg gctttaagct cacaggtcac

16021 ttgtccactc cactcatcca actgcaagcc ccagggaggg ttggcctggg ctccaggaaa

16081 aaagattttt aaagacgtga ccaggcaaag tcccaagggt atgcacaggt cccaaagaag

16141 gaatgccccg cccagggaaa gacccgccca gaaaaaaaga cccgccaagg gaaagctccg

16201 cccagaaaaa gacctgccca gggaaagccc cgcccagaaa aaaggtcctc ccagggaaag

16261 ccccacccag ggaaagctcc acccagataa aagcccgccc agggaaagcc ccgcccagaa

16321 aaaaagatcc gccaagagaa gtctccgccc agataaaagc cccgcccaga aaaagacccg

16381 ccaaaggaaa gcccagccca gaaaaaagac ctgcccaggg aaagccccac ccagggaaag

16441 ctccgcccag ataaaagccc gcccagggaa agccccgccc agaaaaaaga ccagcccaag

16501 gaaagctccg cccagggaaa gccccgccca gaaaaaaaga cctgcccagg gaaagctctg

16561 cccagataaa agcccgccca gggaaagccc cgcccagaaa aaagaccagc ccaaggaaag

16621 ccctgcccag aaaaaagacc gcccagggaa agccccgccc agaaaagacc cgcccaggaa

16681 aagctctggc cagataaaag ctcagcccag ggaaagcccc gcccataaaa aagccccgcc

16741 cagataaaag ccctgcccag cgaaagcctc gcccagggaa agccccaccc agaaaaagac

16801 ccgcccaggg aaagcccagc ccagagaaaa gacccgccca gggaaaactc tgcccagata

16861 aaagacccgc ccagggaaag ccccgcccag aaaaagtccc gcccagggaa agccccgccc

16921 agaaaaaaaa cccgcccagg gaaagccccg cccagaaaaa agtcgcgccc ggggaaagcc

16981 ctgcccagaa aaagtcccgc ccagggaaag ccccgcccag ggaaagaccc gcccagaaaa

FIGURE 26 Continued

17041 aagtcccgcc cagaaaaaag tcgcgcctgg tgaaagccct gcccagaaaa aagaccagcc

17101 cagggaaagc ccctcccaga aaaaaagacc cgcccaggaa aaagctctgg ccagataaaa

17161 gctccgccca gggaaagctc cgcccaggga aagccccgcc cagaaaaaaa gccccaccca

17221 gggaaagccc agcccagaaa aagacccgcc cagggaaaac tccacccaga aaaaagacca

17281 gcccagagaa agcccaaccc agaagaaagc cccacccagg gaaagccccg cccagggaaa

17341 gcttcaccca gagaaattcc cacttagaga aattcccact cagagaaagc cccacccaga

17401 aacgtgccac ccagggaaag ccccacccag tgaaagcccc taccagaaaa agccccgccc

17461 agaaaaagcc cctcttagag aaagccccgc cagactctca gaagttaatt tccttttttcg

17521 ttgttttgag agggagtctt gctttgtcac ccaggctgga gtgcagtagt acaatctcag

17581 ctcactgcaa cctctgcctc ctgggttcaa gcgattctcc tgcctcagcc tcccgagtag

17641 ctgggactac aggcacaagc caccacaccc ggctaatttt tgtatttta gtagagacgg

17701 ggtttcacca tgttggccag actggtctcg aacttgtggc ctctttttt ttttttttt

17761 ttttctttga gatggagtct cactctgtca cccaggctgg agtgcagtgg ctcgatctcg

17821 gctcattgca agctccacct cccgggctca cgccattctc ctgcctcagc ctcccaagta

17881 gctgggacta caggcccctg ccaccacgcc ctgctacttt tttgtatttt tagtagagac

17941 ggggtttcac catgtcagcc aggatggcct caatcttctg gcctcatgat ctgcccgcct

18001 cggcctccca aagtgctggg attacaggag tgagccaccg tgcccagcca cttgtggcct

18061 cttctgttat tttctgaatt gtttacactt cccttactca tcacagagct tgagagaaat

18121 tctgtagctg tgatgggata aaaggatgga tgggcaggtg aacaaatgga cagacagctg

18181 gctgggaagt ggaatttctt catccaagat gggtcaactc aaggaatcga ttgccctaaa

18241 acatacctgt tccactgttg gccacttttg caatagacaa gttcacatca agctagacct

18301 gcctgatccc tacattctaa ctggagtgac caacaggaca cgggagagaa aaccacatac

18361 aaaaccaatc cacagaaccc cgccatgccc aggtccccac agagagggga gggggcgttt

18421 tctccacttt ttttttctcgg cctgtgagcc cctgagccgt gcactgcggt cccacaggtt

18481 gaccccgcgt tgccagcctc agggccaagg tcacgtttcc agacatggcc ctaagaaaaa

18541 ggccagccca ggggaaggga catggtcagg gcacacagga accacgtgca taccacacat

FIGURE 26 Continued

18601 ccatgcccat gagcacacac cacacaccac atgtgtatgc acataccata cacacgtgca

18661 caaatgcatg tatacacact atagtcacac catgcataca tctctaccca cacatgcaga

18721 atcatgtaca ggtcatacac aacacacacg catgtatgca catgccatat acacaccaca

18781 tgtaccatgc atacaccata cacacatgtg cgcaaatgca tgtacacaca ccatcacacc

18841 actcatacat ctctactcac acatgcagaa tcatgtacag gtcatacaca acacatacac

18901 gtgcacacat gccatataca caccacgtgt acacacatgc accatgcata caccacacac

18961 acacgtgcac aaatgtatgt acacacacca tagtcacacc actcatagat ctctacccac

19021 acatgcacca tcatgtacag gtcacacaca acacatacac atgtgcacac atgccacata

19081 ccacgtgtac acacatgcac catgcataca ccatacacat gtgcacaaat gtatgtacac

19141 acaccatagt cacaccatgc atcatctcta cccacacatg cagaattgtg tataggtcat

19201 acacaacaca tacgcatgta tgcacatgcc acatacacac cacatgctcc atgcatacac

19261 acacaaatgt atgtacgcac catagtcaca ctacacatac atctctaccc acacatgcac

19321 catcatgtac aggtcataca caacacatac acatgtgcac acatgccata taccacatgt

19381 gtacacacat gcaccatgca tacaccatac acacgtgcat ctacacatac agatacaggc

19441 acacacacca ccatatacat cacatacaaa gacatccact gatatacgca cacctacata

19501 catgtacaca cagcacacat gcacatacat cacaacacac atgcgcacca cacaccatgt

19561 gcacacccat acacccatgc tacatgcaca ccatacccca cacatatgca tacactgacc

19621 acagcacaca tcaaccacac acccgccaac ataactcttt ctctttttg gggagataga

19681 gtcttgctct gttgcccagg ctggagtgca gtggcgtgat ctcagcttac tgcaacctct

19741 gccccctggg attcaagcga ttctcctgcc tcagcctcct gagtagcagg aattacaggt

19801 gtgtgccacc atgcccggct aatttttgta tttttagtag agatgaggtt tcaccatgtt

19861 ggccaggctg gtcttgaact cctgacctca ggtgatccgc ctgcctcagc cttccaaagt

19921 gctggcatta caggcatgag ccaccgtgcc tgcctttttt ttttttcttt ttttttgtgt

19981 gtgtgtgtga gacagtcttg ctctgtcacc caggctggag tacagtggca caatcttcgc

20041 tcactgcaac ctcggcctcc caggttcaag tgattctcgt gcctcagcct cccgagtagc

20101 tggaattaca ggtgcaggcc accatgcctg gctaagtttt gtatttttag tagagactgg

FIGURE 26 Continued

20161 gtttcgccat gttggccagg ctggtctcga actcctggcc tcaagtgatc cacccacctc

20221 ggcctcccaa agggctggga tgacaggcat cagccaccac ccccagctcc aacaactttt

20281 ttttttttt ttttttttt ttttttttt ttttaagat ggagtctcac tctgtcaccc

20341 aggctggagt gcagtggcgc gatcttggct cactgcaacc tccgcctccc agattcaacc

20401 gattctcctg ccacggcctc ccgagtagct gggattacag gcatgcgcca ccaccccggc

20461 taatttttt gtttttgtag agacggggtt tctccatgtt ggtcaggctg gtctcaaatt

20521 cctgacctca ggagacctgc ccgcctcggc ctcccaaagt gctgggatta caggcatcag

20581 ttactgcgcc tggcctcaaa atacttcaga gcaatggctc tggactgtca gggctggaaa

20641 ggaccctcct ctcacatgct gagctccttg gagtggaggc caaagtccca cagcagaggc

20701 cgactggccc acccagcatg cccatctccc aggagtgcag agaggcaagc cctgggcagt

20761 ggcaggcaca ggccttcttg accccagacc ttcagcctgt catattttgg ctcctcctta

20821 gtgaaggttg ttgagggtgt tttgcagaga gacatgacgc caatcttaat ttttgacaat

20881 tttccatagc atgcagataa tttgtttcca aaacttttca ttttcctgaa gtcatcttga

20941 ttggtatcag ctatttccat aaaacgatcg gatgagtttt gatggacaga tcaggctttt

21001 gtttacaact gttttgctcc taatcattcc accacatcac atgtcatgga cctgaattgc

21061 gtcaagaaga cgggcttgtc tgtcaggccc tggtgggcac tttgatagcg ggcatgctgt

21121 gccatgacac gtgtggtgtt gggtcttgct ggacaagctg tgctgtgttc agtgtgcgga

21181 gcctctgcta gatgctctca tttggggcac tgggccagtg ctactgggag cacttctgtt

21241 ttgtgtcact gacatccaat agcatcgtta tgtagagcaa acaccgaagg gctgcatttc

21301 tttgtgggct tattctcgag aaaactgggg gcagatccct cctcaaggag gggagggcca

21361 ccttggtttc cagtcaagta ttgtgaaaat tatccaacac tcaggcaatc cacccaaccc

21421 tgctgcccat gtctggagaa gcaaagtgtc aggggtagtc caggcccacc tggagacagg

21481 tcaggccctg cagagaaagg tctgacagac gggggtgagg gaagaccccc caaaggcctc

21541 cagagtccca ccaggtctcc aggtccttgt cataaccaga gaggccccag ccccagagga

21601 ccaggtcccc tgctccactg tccacagggg cccacctgca agcacactgg cagagctcaa

21661 gaccacacat gcctgcaagg tgaggcctgt ctgggctctg tcctcctgca ggccccaggc

FIGURE 26 Continued

```
21721 ctgggtggct gggcgaaggc agctgcttat gcagactcca gggggaaagc cgcctctcat
21781 ctctggccgt ccccaggacg ctggatccac caatatctca ccaacctgga gagccactca
21841 acccctcatt tcacatgttt gaacatagag gaccagaggg gtgtggcctg tctaggaggt
21901 cttaggagct cgggtcctga ctctgccact taccaactct tgtgtgtccc atgtgcctcc
21961 gcttcccctc gggacacaga gatattgtga aagttaaaca acataatccc cgtaaaacac
22021 ttcgagcagt gcctggtatc tggccagcaa gtgatcaatg gtgatccatt accatcctgg
22081 gaccccatca gagccttctg aggtggaggg aagggcgtgc tggggagcac aggtgcaggt
22141 cacaagaagg aagtcagtcc cataagccag gtatctaacc ccatccctgc tcccccaagg
22201 taagggccag catctaaccc cacccctgct cccccaaggt aagggccagc atctaacccc
22261 atccctgctc ccccaaggta agggggccag catctaaccc catccctgct cccccaaggt
22321 aagagccagc aacggaggcc tgggaggctc ctgggttctg ggccgcagcg cctctgcgag
22381 gtctgcaggc ttcgctctag gaggggatgg gggctgggca ggtccctgct ccagaggagg
22441 aggacctggg cctgcggagc gccgcggtgg gagtgctgga gtcctggccc gtcatccccg
22501 tctgccccac agcgaggacg atgctgccac tgtataccgc gcagccgcga tgctgaacat
22561 gacgggctcc gggtacgtgt ggctggtcgg cgagcgcgag atctcgggga acgccctgcg
22621 ctacgccccа gacggtgagt gctgggcctt ggcggggtcc ccgaacgggg aggaccccac
22681 gggctctgag tcgcatgctc gcctaggcat cctcgggctg cagctcatca acggcaagaa
22741 cgagtcggcc cacatcagcg acgccgtggg cgtggtggcc caggccgtgc acgagctcct
22801 cgagaaggag aacatcaccg acccgccgcg gggctgcgtg ggcaacacca acatctggaa
22861 gaccgggccg ctcttcaaga ggtgggcggg gcctccccgg agctgggcgg ggctgctctt
22921 ggggaggtgg gcggggtcac tccagagatg ggcggggccg ctcttgggga ggtgggcggg
22981 gccactctcc agagctgggc ggagcagctc tcaggactag gcggggccgc tcttagggag
23041 ctgggggagc gctcctcaag agatgggtgg gggcactctc ggggaggtgg gcggggtcgc
23101 ttccaggagg tgggcggcgt cgctctcagg ggtactgcag tggagcctgc tgccaacatc
23161 ctctggacac tgttacttct ctcctctccc cccacacccc cagcaccacc acatctaatg
23221 gcacaatcat ctgccctctt ctcaacactg acaccagtac ctgggccgtc actggagtgg
```

FIGURE 26 Continued

23281 ggactggctc cactgcctcc gcccctactt tccacactgc agcccaccct gaaacagcac

23341 ccctctccct gtgtggctgg cagcctttgg gaggaggctc ttgatgcaga tggggactga

23401 aagcttccag ggacccagga ggccagacaa gcagcccaag aacagcacac gagccttaga

23461 cagccagggt tggccaaggc ccagagaccc aagtgaacat ctgcagtgtg gcaggagtta

23521 gctcacagca cgcctggaca ccatgccatg ccagctcacc cccagatccc caaccactga

23581 gtagcacgtg cagagccacc atccacaacg cccacataag tgcagatgta ggcagcacgt

23641 gtgcacacac acgacacaca tacacagaac catgtgtgca cacagactca ggcacatgac

23701 acacatgtga cacaagcaca tgcatggggt gcaccccaca tagggatcac gtgtgcacac

23761 aggttcactg atgtggcccg atgggtacaa tgcacacgtg cacacacagc cggacaggac

23821 agcctggtgg ttagagctgg ctcaacctcg ccttcacttg ctggcaagag ggcaggcatc

23881 cttctgagct tctgcctccg tctctgtaag gcaggatggt tctgaggaca acgtcctaac

23941 ccacagaaag ccgggtctgg cacccataaa ccactcagct gtcatgaacc acaccgtcca

24001 tctggtgcag gcagatacca cggtgtccag ggtctggcgt ctgctgatct ttccgttctt

24061 gggactggga caagggaaaa gcccaagctg ctcagccggc aggagaagga gcagggagaa

24121 ggagcagggg gaaggagcag ggagaagcag caggggggaag gagcagggag gagcaggaga

24181 aggagcatct ctgagaagcc tcagctatgc ttccttccct agagtgctga tgtcttccaa

24241 gtatgcggat ggggtgactg gtcgcgtgga gttcaatgag gatggggacc ggaagttcgc

24301 caactacagc atcatgaacc tgcagaaccg caagctggtg caagtgggca tctacaatgg

24361 cacccacgta ggtgggggtc atgaggggtg gggggctggg gccttagggt cctggggcca

24421 agacccctgc gtggccaccc tccatctcat actcccaccc ccaggtcatc cctaatgaca

24481 ggaagatcat ctggccaggc ggagagacag agaagcctcg agggtaccag atgtccacca

24541 gactgaaggt gggggcccca cagacctccc tcagtgtccc caccccagac agcccatcca

24601 ccccctctgg cctgaaggag gagggtgcgg tgaggtcaat gaaagccact aaaggaagtg

24661 ggggtggggc ctgcttcccc tggacaccgt ccagcacacc tggcacagca caggaagcag

24721 agagaacagg agggaggaga ggaagctgcc cccatcccac agggggtctc cagtgccccc

24781 cttgacccag ccctacttaa gtctggggca gttagttgtc tgacaggacc ctgctgggga

FIGURE 26 Continued

24841 agagcagatg ggggacagca ggcagacctc agcttcagca ctcgctgtcc ccagtcctgg

24901 tcctccacac ccctcatccc tcctccagcc tgcattgctc ttgatgggac cgggtcaaac

24961 tgtcctcttc caccgtgtgg gacagccctt cctgactccc ctgggcctct gagagcctct

25021 gccctcgccg gcttcctcct ccagaacatc tttcccttgg ctccctactc cagggtgctc

25081 tcctggccat tcctccccgg gcagagccac actaccccca ctccacacac actccagtcc

25141 tggtagcatc acagaccacc aaaggcaagg acctcacagg cgacacgccc accaaccttc

25201 tctcggtcat tccaagccct caaatgtctc ttgaccctgt ctgttttctg agcccacccc

25261 tgaagcttgg tgtcagcccc tgtgacctct cacccaggct ccctcccctg ctctgcaccg

25321 gcccctgtgg cctctcaccc aagctccctt ccctgctctg cagacagggt ggggttttcc

25381 agtgccagtg tgggtttcat tgcagcccag acacctcaca ctgaaaagtc tgaaagcagc

25441 ggtcaaacgc taatggccaa aaggccccat ctaggctgtg agatggagat ggcttttac

25501 aaatttgttt ctggcctcac taatttttta aaaatactag catatatatt acctgtataa

25561 caggaaaatg taatgaaagt tttataaagc aaatcaactt cttcaatggc tcctgattcc

25621 cctggggata aaagacaaaa tgctgcctgg aggctgaggg tgggcgggcc tgccccctc

25681 tcaggctcac cctgcctagg acatgccggg agggtgcctc tcccaccacc cccacgcctc

25741 cctgcctttg cagttctgga ctgcagactc ctctgctcca cctgccctca ggcacctgct

25801 tgatccctgc ccaccttgag ggctcagctc tgacaccatc tcccctcaca gttctggcag

25861 tgtgctatgc tctattccag ccccctgtgc cccagatccc ttccccaccc ccatgccatg

25921 gtcccttgaa ggacagacag gagggcgagc ccaagcagga gtgtgggtcg aagaggccac

25981 ggcgcggtgg agcacgtaca cacgggcaag agaaaggagc cagagaccta cattcaaagc

26041 ctgagggctt cgggactggg ggccgggaca ggcagtgcgc cgggatgaag ggaggcacgg

26101 gtgggtggcc ccacgggtcc caggtcctgt gcaggtgcag ggtcggcttt gtggacatgc

26161 ccctgtcctc gtggcacagc agggtggggg tcagcctgca ggctgggctg tttctcaccc

26221 caggaagatg cctggcatac acgggacatc agcggctcct ctgctggagg gaatcatgtc

26281 ttttttttttt tgagacagag tctcgctctg tcgcccaggc tggagtgcag tggcgcggtc

26341 tccgctcact gcaagctccg cctcccgggt tcacgccatt ctcctgcctc agcctcctga

FIGURE 26 Continued

26401 gtagctggga ctacaggtgc ccaccaccac gcccggctaa ttttttgta ttttcagtag

26461 agacggggtt tcaccgtgtt agccaggatg gtctcgatct cctgacctcg tgatccggcc

26521 gcctcggcct cccaaagtgc tgggattaca ggcgtgagcc acagcacctg gcggggaatc

26581 atgtctaagc caagactgga gaaaaagtgg ccaagagaag ggtccagctc tccaggagtc

26641 ttttctgagc ccccagcccc cacccccccg gggctgcagg cagacgatgc tgacggtggc

26701 tggggaggac gtgtcctgaa cacttgggct cgtgaagaag ctccagagag gggcagtggc

26761 cggcggcgca gggcgggggg tgtgaggggt gcgtcgggga ttaagagggg cgccagggga

26821 ggctgggagc tgagaagaga ctgccgccct gggcagcctt aggtcggtgg tccaggctgg

26881 gtctcccctt ccccccccaga ttgtgacgat ccaccaggag cccttcgtgt acgtcaagcc

26941 cacgctgagt gatgggacat gcaaggagga gttcacagtc aacggcgacc cagtcaagaa

27001 ggtgatctgc accgggccca acgacacgtc gccgggcagc cgtgagtgcg cggggcaggg

27061 cgcggggcgc ggggcagggc gcggggcgtg gggcggtctg gagcccagca gttaccgccc

27121 gcacctaccc agcccgccac acggtgcctc agtgttgcta cggcttttgc atcgacctgc

27181 tcatcaagct ggcacggacc atgaacttca cctacgaggt gcacctggtg gcagatggca

27241 agttcggcac acaggagcgg gtaggctgga cggcgggggt ggggaccagc gtgagagggg

27301 cctgcaggcg cggtcggagt gggtggggca tggagtaggc ggggcttgca gatggtgggg

27361 ggtcctgggg tgagtggggc atggagtgag cggagcctgc gggctgggtc ctggcgtggg

27421 taaagcatgg ggtgggcggg gcctgagggc tgggtggggc ctgacatggg aggggcctga

27481 cgtgggggtc ggagtgggtg gggcacggag tgggcagggc ctgcaggcgg gggtctggag

27541 tgggcgggac gtggagtggg cggggcctgc tggctgtggt ggggcccgcc cggcgtggga

27601 ggggtctgcg agccagggcg gggctggagt ggggtggggc ctgcgagctg ggtagggtct

27661 tggggagaag acccccggag tgctctaggg cggcttcagt cgggggtacc tgtggcggga

27721 gctgggagga cgctgcctgc atgcccgccg gctctgtcgc ctcgcaggtg aacaacagca

27781 acaagaagga gtggaatggg atgatgggcg agctgctcag cgggcaggca gacatgatcg

27841 tggcgccgct aaccataaac aacgagcgcg cgcagtacat cgagttttcc aagcccttca

27901 agtaccaggg cctgactatt ctggtcaaga aggtgggcag gggccgggtg gcggggtggc

FIGURE 26 Continued

27961 ggcgggggga gtccctggag ggcccgggcc gcgctgacct cgcgtccctc cgcaggagat

28021 tccccggagc acgctggact cgttcatgca gccgttccag agcacactgt ggctgctggt

28081 ggggctgtcg gtgcacgtgg tggccgtgat gctgtacctg ctggaccgct tcaggtgagc

28141 gcgacccggg gctcagacac ctccatctgc ggggcgcgga gccggccagg ggcggggcag

28201 ggccgcctct cccgccctct ctcccgcccg ccctctgcgc cccgcagccc cttcggccgg

28261 ttcaaggtga acagcgagga ggaggaggag gacgcactga ccctgtcctc ggccatgtgg

28321 ttctcctggg gcgtcctgct caactccggc atcggggaag gtaaggcccc gcccggcccg

28381 cctggtcccg cctcggccct ctagggtctg acagagcccc ccgcccgccc acaggcgccc

28441 ccagaagctt ctcagcgcgc atcctgggca tggtgtgggc cggctttgcc atgatcatcg

28501 tggcctccta caccgccaac ctggcggcct tcctggtgct ggaccggccg gaggagcgca

28561 tcacgggcat caacgaccct cgggtgaggc ctggccgggc tgggggaggg aatgcgaggt

28621 gagctggggt cggcctcggt taggggcctg gggagccgcc gccgcgatcc ctgccctccg

28681 accctgcagc tgaggaaccc ctcggacaag tttatctacg ccacggtgaa gcagagctcc

28741 gtggatatct acttccggcg ccaggtggag ctgagcacca tgtaccggca tatggagaag

28801 cacaactacg agagtgcggc ggaggccatc caggccgtga gagacaagtg aggcgcgggc

28861 ggccaccctg gcggggcggg acaggtgcgg ggaggggggag ggtggcctcc accgggcagg

28921 agagcgtccg ggccgggcac cccggagggc gcgggcgtgg ggcttccagg ctggcaggac

28981 caaggccccc gtgactccgc ctctgccggc agcaagctgc atgccttcat ctgggactcg

29041 gcggtgctgg agttcgaggc ctcgcagaag tgcgacctgg tgacgactgg agagctgttt

29101 ttccgctcgg gcttcggcat aggcatgcgc aaagacagcc cctggaagca gaacgtctcc

29161 ctgtccatcc tcaagtgagt gtccgtgcgc ccgcgtcccc cctccgcccc tctccgccag

29221 aggtggacgc cctccccagt gccagaccac tccgaggcca ccactgattt cccacccagg

29281 ccgggcgctg cccactccac gccgcaccct accccgcagg ccccgccccg gccccgcccc

29341 cagcttgctc cttcccgtcc tgggcccccgc ctcactgcag gctcacttgt tcccaccgcc

29401 aggtcccacg agaatggctt catggaagac ctggacaaga cgtgggttcg gtatcaggaa

29461 tgtgactcgc gcagcaacgc ccctgcgacc cttacttttg agaacatggc cggtgcgttc

FIGURE 26 Continued

29521 tccttcatcc attctcgggt gggttctccg tgggctgcgg cctccctggc cagcaactga

29581 ggctctgggt cccggcacac aggggtcttc atgctggtag ctgggggcat cgtggccggg

29641 atcttcctga ttttcatcga gattgcctac aagcggcaca aggatgctcg ccggaagcag

29701 atgcagctgg cctttgccgc cgttaacgtg tggcggaaga acctgcaggt agggcaggcc

29761 accctccgag gcctggtgcc cagggcccgg cctggccacg gccctcctcc atccccgaag

29821 gccgtggcac tggctctggc tctggtgggc aggactggag ctaggagcca tggccagggg

29881 cagtggtgag tgctcccagg gcacgggggc agcaccggtg gggggctgcc tgcaggtggc

29941 tgcccactgc aaagccgggg ccgagggagg ccacgcaccc tgctccaagc ctccgcctgg

30001 cccctctgtc tccagagtcg cccgccggta cccattccat aggaaggcaa tcaggcaggg

30061 taagacaggg gcccgcctgt gtatggcacg tgagtccaag atgcattttg ccctccgccg

30121 acccaagccc cttgacaccc ttcggagacc ccccccttte ctgctatgtc cttgtgctcc

30181 gtgactctaa tccgaattgg gccaggtccg gtcctgcctg gtgcccaggt tgtatccatg

30241 agaatttgcc accagcaagg gcagccacgg cccacctggg acagggtggg cagtgggcct

30301 gtacaggcct aagggctcgt ggcccgcggt cgagttccgg ttcactccgt ctcttctctt

30361 tctctgggtg ccgtcctgga gcctgtgtcc tgagatgaag ccgacagtgc ggccagggct

30421 gctgggggat gggggttgct ggaggctcca cacctctcat ccgcccgctc ttgctcttgg

30481 cccccacagg tcccctgggg acctggccgc tgccagcact ggcgggcaca ggccacctgg

30541 ccatcagacc tgaggccaga gtcccgggag ctgcctctgt cactccaatt ccacctcgac

30601 acctgcctcc agccctcggc cccttcctga atcttggtgt gtgcccctlg ggggtcagtg

30661 gcctccacgc agacagctgg tgtggcctga ggggcaactc ctccagtcct cagaggactc

30721 ctcctcctcg ggacgcctgt aagccagggc cacccaggag ccagggagcc aggcggacct

30781 cccaggaaga gccagccgag agcccccaag cccagcccca gcacgagcaa ggtcaggccc

30841 gagaccccgg gcaggagaag aggccaccct cgaacgtccg ctgtcggccc gtctgtccag

30901 cacagggagg caggcaggag cgagggccca agtggccggc caggctgggc agcggcccat

30961 gcaggagcag gcgagggcag gtgtggccac caccctagcc atctaatcac ttatacatat

31021 tcattttagg atagaaagag tggtagagca gagcctgacc ctaaaaagaa agccacattt

FIGURE 26 Continued

31081 agggctatca cctccaccct ggcttccagc ttcaagaggc gtaggtcctc caaagacacg

31141 gtaaggggga gagcacccca gtcccgcgtc cgactccacc tgccctgccc tgcgtgtgtc

31201 tcccgcccca tcaccccgcc ccggaccctg ggctcctgtg gcccactctg cccctgtctc

31261 cctgtggcgg ccgctctgcc cagcccgccc atgctgctct ctctcactct ctggacctttt

31321 ctccccggcc ctcctgggtc ctcggctttc cccgtgtgtc tccgttagtc tgcccgccca

31381 cctcccctgc catgacccac acgccatctt gaagcctgtc atctcgttgg tcagtcagtc

31441 agccacacca cctctcgggg ccaggtctgg ggccctggga gcccagcgtg gccccatcct

31501 ggactcctca gctgccggga ggccacacca cttctctgtt atgtccccgt ttctctcgcc

31561 tctcccagag gggcccgccg ccctcacttc gcccctgcga cggccctgga gggggtggct

31621 gtgatgtccc atcccgtccg tctgtctggc cactggcccc gccccccaga cacctgtctc

31681 acctgtctca ccagagccat gcgtgttgca tcttcatgtg gtctctgtgt gggccggggg

31741 ctgggggccg ggcctgggtc cgtctgggtg gacggctggg gcctggagtt ggaactggcc

31801 ccggccacag gggactgtca ggcagggagt ggggtgggac caaaaggggt ggctcccacc

31861 ccaggctgag cgggggccct gcaggaggtg tggcggcagc tcccagaggg tctgagaatg

31921 ggtaggggcg gccccacaag ccctggcctg cagagcccag gacgacactg aggttcccag

31981 acagggaggc ctctggaagg gaaacgacca cctcagctcc tgaccccagc aaccccacaa

32041 ggcccacccc aaagagccag gccttctgcc ctttggagcc cagaatcccc cacctcctgc

32101 tcggggcagc ttgtccctgt agcggatatg cacactcgga ccagaggccc ccagagcgaa

32161 cccagccttg ctagaggcac cccaggccca ggcaccatgg tggggagggg ctgcccagag

32221 aggcagcgga gacctcagcc ccgtggccac cctgcagtcc agggaccagt ctggcccaca

32281 ggaagccccc agcccataag cagcatcacc agagagaagc ttacgcccgg gggaggaagt

32341 gcgatttgca gccacctgcc cctcagtgca ctggaagcgg ggcagacctc cagggcacag

32401 acaggacttg gcatcaagca agccaaatcc cgagatgaag ccaccagggt gccccaagag

32461 ggacccatga ggcctggctg ctcagcttcc tggggaaggg acttggcatg caggatgggt

32521 ggacagtgag agcctgtagg cctgggggcc actggaggct caaggagcag gtggaagcac

32581 cattcctgga gccacctctg ctgcggaaag cgggcagagc tgatgctgca aagtctgagc

FIGURE 26 Continued

32641 caggagtccc gcagggaaca gggaggggga atagcgcagg gatcgtgggc tgggcaggct

32701 ggggaagagg gggtgtccag gcagacagga gaaacagcga tttggggcag gcagccacgg

32761 ggggcaagca caaatgtcgt gcaggtgatg ggccactttc agagggtgac actgggtccc

32821 agggccctgc ctggagcgag gccaggtgca gctcagagac cctcatggtg ccctcccagg

32881 gacatgttcc cagcggaacc ctcacccgag cctctctggg caccagggac cgtcctctgg

32941 ggccagttct ggcatcacgt ggcatctggg gctggccccg ccctgcaagg ctgaactgtg

33001 gggggcactg ccagctgggg gtctgggcag gggagggcag cccagctccc acctggtctc

33061 tggggctgcg agcttattca gagggaggcg tgggtggggg gctcctttgg gtagggtggg

33121 gtcagtccgg ctgcggagat cccctgcccc tgtcctgtgg ccggtccggg ccagggcggc

33181 actgggcgct gagggctggg gtccctggcg gccggcgggg ccagcgggta ttgattgttg

33241 gttcttattt atagagcacc gggggtggac gcggcgcttt gcaaaaccaa aaagacacag

33301 tgctgccgcg acgcgctatt gagagggagg agggccagct gcagctgtgt tcccgtcata

33361 gggagagctg agactccccg cccgccctcc tctgcccccc cccccgcaga cagacagaca

33421 gacggacggg acagcggccc ggcccacgca gagccccgga gcaccacggg gtcgggggag

33481 gagcaccccc agcctccccc aggctgcgcc tgcccgcccg ccggttggcc ggctggccgg

33541 tccaccccgt cccggccccg cgcgtgcccc cagcgtgggg ctaacgggcg ccttgtctgt

33601 gtatttctat tttgcagcag taccatccca ctgatatcac gggcccgctc aacctctcag

33661 atccctcggt cagcaccgtg gtgtgaggcc cccggaggcg cccacctgcc cagttagccc

33721 ggccaaggac actgatgggt cctgctgctc gggaaggcct gagggaagcc cacccgcccc

33781 agagactgcc caccctgggc ctcccgtccg tccgcccgcc caccccgctg cctggcgggc

33841 agcccctgct ggaccaaggt gcggaccgga gcggctgagg acggggcaga gctgagtcgg

33901 ctgggcaggg ccgcagggcg ctccggcaga ggcagggccc tggggtctct gagcagtggg

33961 gagcgggggc taactggccc caggcggagg ggcttggagc agagacggca gccccatcct

34021 tcccgcagca ccagcctgag ccacagtggg gcccatggcc ccagctggct gggtcgcccc

34081 tcctcgggcg cctgcgctcc tctgcagcct gagctccacc ctcccctctt cttgcggcac

34141 cgcccaccca cacccgtct gcccttgac cccacacgcc ggggctggcc ctgccctccc

FIGURE 26 Continued

```
34201 ccacggccgt ccctgacttc ccagctggca gcgcctcccg ccgcctcggg ccgcctcctc
34261 cagactcgag agggctgagc ccctcctctc ctcgtccggc ctgcagccca gaacgggcct
34321 ccccgggggt ccccggacgc tggctcggga ctgtcttcaa ccctgccctg caccttgggc
34381 acgggagagc gccacccgcc cgcccccgcc ctcgctccgg gtgcgtgacc ggcccgccac
34441 cttgtacaga accagcactc ccagggcccg agcgcgtgcc ttccccgtgc ggcccgtgcg
34501 cagccgcgct ctgcccctcc gtccccaggg tgcaggcgcg caccgcccaa cccccacctc
34561 ccggtgtatg cagtggtgat gcctaaagga atgtcacgca gttttcggtc tgtgtcgctt
34621 gttgacgccg gcagacagtg taaagggagg gcaaaggcat gggggaagct tcgagcgctc
34681 caggcggccg cggccgctca ggcttgggcg gcagcggcgg ggctccccgg gtccgcgggc
34741 gaggcacagc cgtgggggtc gggatcgggg ttcgggtctg gcggtctcgg cgggcggagg
34801 gcggcggtgc ggaggcggcg gcggcgcgca cggcaggcgg tgagcccaga gcccagcgcc
34861 aggcaggaag ccaggctgac gaggaaggag gccggcccga gcgtgtaaac cacggccagg
34921 tcccgcaggg cgagcggctg cgcgcaatgg ctaaaggcgg cgtcggaaaa ggcagtcagg
34981 gggctgagcg tcaggcgtcc cggccacacg cagagcaccg tctcggcctc tgtgggacac
35041 agacaaaggc gcggcgtcag gtggcggctg ccgcaccgcc ctgcagaccc cgacccgcgt
35101 ccccagcagc tcacctgacg cgggcagcgg gtgccggcgc agccaggcgc agagcgggcg
35161 cagcgcgcac ccgcagcccc aagggttgcc gcgcaggtgc agcgcgtcta gagcgggcag
35221 gcggcccagc agccccggcg cgagtgccgc cagctcgttg tcctgcaggc tgagtgagcg
35281 cagcagcggg agcgcgccta gcgccgcggg ctccaggcgc gccagccggt tgccggccaa
35341 tgagaggttg cgcagcgcgc gcagcggcgc gaaagtccct ggtgccagtg cttccagctg
35401 gttggcgctc aggtccagca gctgcagcgc gcccaggccc cagaaggctc gcacatgcac
35461 cgagtgcagc ccgttctcgc gcaggtccag gcgctgtagc gcgcccgctc ccgcgaaggc
35521 acctggcggc agcgcacgga cgcggttgtg gtccagcagc agcgcgcgca ggcgcaggct
35581 caggcccggg ggcacggcgg gcagcgagag tgccgagcag ctggccaggc ctcccggcac
35641 gcacgtgcac acctcggggc agtccggggc gcccagggac cccgagggcg aggccgtggc
35701 cgacacctgg gcccagacag gccaaggcga cagcagcagc aacagcagca gcagcggccg
```

FIGURE 26 Continued

35761 aggccgcgac caggaagggc cccgcatggg ggcagccccc ccgcccccgg cacccgcggt

35821 gggaggcccg ctgcctgtgc gtccctggag cccgtcgtcc gcggagcccg ttcctgcggc

35881 gcctgcacaa atattaactc tctggcccga gctcaggcag ttcctgtccc acggctggat

35941 ccacgcttgg ggcgggggca gggcaaacag ccaacgcccg gcaccgccag ccacctgtcc

36001 gggagctcca aactactctt ggctcagcgc cggccacagc gctatcaaga ccacctcacc

36061 ccgccttgtc caccaccggg cgcccgccga gccctgacct gagcagcagg acaccgccca

36121 cctgcacccg cccagctcac ggctgcaaca ggcgcccaca cgcaatgcaa cccagaggtc

36181 ctgctcctca cccgtctcga ccccacccag gctccgcaaa gtgatcccaa cgtgcacatg

36241 cggagtggcc cccacgcagg gatggtccca ttcccatgct ggatagggcc tgggttggag

36301 acaggccttg gggtggaggg agacagcaca cccgaggcag gagaggcgtg cctgccccac

36361 ccctcccccc caggactctc ctgggataga ggtacagacc agtgcagtgg ggcaggggta

36421 tcagcccaag tcctgctgtt aaacccagcg cccttcacag ttgccagttg caggtcctgt

36481 tctaggggct ttccaaagct gggtgcagga aggagccggg ctgaccagct gtggcagaga

36541 aggtggaaac attaggggta tggcttactg ccaggggcca gaggaacagg ggaacttgct

FIGURE 26 Continued

```
   1 gaggtattgt ggccagtatt ttggttttct tatgttttgg agtcacacaa gctaaagacg
  61 ggccattttc cagacctcat ccagcttact ggcggttttg gctgtgtgtt agtgtggtct
 121 acgaattgtt tctcatcttc atccttttcc agacagtcca ggatggccga cagtttctga
 181 agtatgtgga tcccaggctg ggagtcccat tgccagagag ggactacggg ggcaactgcc
 241 tcatctatga tgctgacaac aagactgacc ctttccacaa catctgggac aagctggatg
 301 gctttgttcc tgcacacttc attggctggt atctgaagac gctcatgatc cgtgactggt
 361 ggatgtgcat gatcatcagt gtgatgttcg agttcctgga gtacagcctg gagcaccagc
 421 tgcccaactt cagcgagtgc tggtgggacc attggatcat ggacgtcctc gtctgcaacg
 481 ggctgggcat ctactgtggc atgaagaccc tcgagtggct gtccctgaag acatataagt
 541 ggcagggcct ctggaacatt ccaacctaca agggcaagat gaagaggatt gcctttcagt
 601 tcacgcctta cagctgggta cgctttgagt ggaagccagc ctccagcctg caccgctggc
 661 tggccgtgtg tggcatcatc ctggtgttcc tgctggcaga gctgaacacc ttctacctga
 721 agtttgtgct atggatgccc cctgaacact acctggtcct tctgaggctg gtcttcttcg
 781 tgaacgtggg tggtgtggcc atgcgtgaga tctacgactt catggatgaa ttgaagcccc
 841 acaggaagct gggccagcag gcctggctgg tggcagccat cacagtcaca gagcttctca
 901 tcgtggtgaa gtatgacccg cacacactca ccctgtcact gcccttctac atctcccagt
 961 gctggactct tggctccatc ctggtgctta catggactgt ctggcgcttc ttcctgcggg
1021 acatcaccat gaggtacaag gagacccggc gacagaagca gcagagtcac caggccagag
1081 ccgtcaacaa ccgggatggg caccctgggc cagatgatga cctgctaggg actggaactg
1141 cagaagaaga ggggaccacc aatgacggtg tgactgctga ggaggggacc tcagccgcgt
1201 catgagcctc accctcgtca ctggccttgt gccaaggggc tgtcccattt ctcccttcct
1261 cgtgctccct gcttcagagg caggggtggg gggggcatcc acactccagg aggggcacct
1321 gagaatacat gtttgtgtgc aggtaggtgt atgcacatat tggctcctga cactactttg
1381 gggccatgag ctgaacagtg agcctggaac ttgctctaga gcaagagtcc tttctacctg
1441 gtcaacaaag gccctggtcc atggtctccc ttgtgctcaa tctcagcacc ggctagggga
```

FIGURE 27

```
1501 ggtatcttgg tatctcggta ctctttctcc actctttatg gggtaggcag aagccccatg
1561 agaccctgtg gtcccaccca cttacagcag cataagtgaa ggatatcata ataccaacat
1621 gtctgcaaag tggtgggtct agagtcagca ctgagccatt tcctttggag ccttccttta
1681 accacgcagg actataaact gaatggtgta tacatgacag tcacagattg gccttttgcg
1741 gccagagagc ccaacttgga gacctgtacc ccaggtgcca gggtgctgtc accatggctc
1801 ccttgagcaa atggaacaaa taaagtgatg atgaaggtga aaaaaaaaaa aaaaaaaaa
```

FIGURE 27 continued

METHODS FOR USING ANTIBODIES AND ANALOGS THEREOF

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of U.S. Provisional Application 61/197,601 filed on Oct. 29, 2008 and is a continuation-in-part of U.S. application Ser. No. 12/563,330 filed Sep. 21, 2009 which claims the priority of U.S. Provisional Application No. 61/192,732 filed Sep. 22, 2008, each of which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

[0002] This invention relates to the use of single-domain heavy-chain only camelid and shark antibodies and their analogs without the light-chains.

BACKGROUND OF THE INVENTION

[0003] The occurrence of various cancers and diseases usually involves pathological antigens, altered protein expression, and/or distribution [Nature, 422, 226 (2003)]. The detection of low levels of certain protein biomarkers can be extremely useful for the diagnosis, prognosis and treatment of specific cancers and other diseases for effective disease control and/or treatment.

[0004] A class of naturally occurring antibodies have been identified from camelids and sharks. In addition to classical heterotetrameric antibodies, camelids and sharks also produce so called "incomplete antibodies" without the light-chains. Their structure is shown in FIG. **1**.

[0005] Thus, two types of antibodies exist in camels, dromedaries and llamas: one a conventional hetero-tetramer having two heavy and two light chains (MW ~160 K Da), and the other consisting of only two heavy chains, devoid of light chains (MW ~90 KDa) and also deprived of constant region CH1.

[0006] In addition to camelid antibodies having only two heavy chains and devoid of light chains, distinctly unconventional antibody isotype was identified in the serum of nurse sharks (Ginglymostoma cirratum) and wobbegong sharks (Orectolobus maculatus). The antibody was called the immunoglobulin (Ig) new antigen receptors (IgNARs). They are disulfide-bonded homodimers consisting of five constant domains (CNAR) and one variable domain (VNAR). There is no light chain, and the individual variable domains are independent in solution and do not appear to associate across a hydrophobic interface. Like the Vab domain of camelid antibodies, the variable antigen-binding domain, known as V-NAR, of single-domain shark antibodies is also stable by itself and has a molecular weight of about 15 KDa (Greenberg, A. S., Avila, D., Hughes, M., Hughes, A., McKinney, E. & Flajnik, M. F., Nature, 374, 168-173 (1995); Mol. Immunol. 38, 313-326, (2001); Comp. Biochem. Physiol. B., 15, 225 (1973)). There are three different types of IgNARs characterized by their time of appearance in shark development, and by their disulfide bond pattern [Diaz, M., Stanfield, R. L., Greenberg, A. S. & Flajnik, M. F., Immunogenetics 54, 501-512 (2002); Nuttall, S. D., Krishnan, U. V., Doughty, L., Pearson, K., Ryan, M. T., Hoogenraad, N. J., Hattarki, M., Carmichael, J. A., Irving, R. A. & Hudson, P. J., Eur. J. Biochem., 270, 3543-3554 (2003)].

RELEVANT REFERENCES

Foreign and US Patents

[0007]

U.S. Pat. Application 12/563,330 Sep. 21, 2009	Antibodies, Analogs and Uses Thereof
PCT/US2009/057681 Sep. 21, 2009	Antibodies, Analogs and Uses thereof
U.S. Pat. No. 7,371,849 (May, 2008)	Methods of constructing camel antibody libraries.
U.S. Pat. No. 6,838,254 B1 (January, 2005)	Production of antibodies or fragments thereof derived from heavy-chain immunoglobulins of camelidae.
U.S. Pat. No. 6,765,087 (July, 2004)	Immunoglobulins devoid of light chains.
U.S. Pat. No. 6,005,079 (December, 1999)	Immunoglobulins devoid of light chains.
U.S. Pat. No. 5,800,988 (September, 1998)	Immunoglobulins devoid of light chains.
WO/2002/048193 (June, 2002)	Camelidae Antibody Arrays.
EP 1264885 (December, 2002)	Antibody library.
WO/2001/090190 (November, 2001)	Single-domain antigen-binding antibody fragments derived from llama antibodies.
WO/2000/043507 (July, 2000)	Methods for producing antibody fragments.
EP 1024191 (August, 2000)	Production of chimeric antibodies from segment repertoires and displayed on phage.
WO/1999/042077 (August, 1999)	Recognition molecules interacting specifically with the active site or cleft of a target molecule.

Other References

[0008] Azwai S M et al, Serology of Orthopoxvirus Camel Infection in Dromedary camels, Comp. Immun. Microbiol. Infect. Dis., 19, 65 (1996).

[0009] Kelly P J, et al, Isolation and characterization of immunoglobulin-G and IgG Subclasses of the African elephant, Comp. Immun. Microbiol. Infect. Dis., 21, 65 (1998).

[0010] Linden Richard van der et al., Induction of immune responses and molecular cloning of the heavy-chain antibody repertoire of Lama glama, J. Immun. Methods, 240, 185 (2000).

[0011] Rivera H et al., Serological survey of viral antibodies in the Peruvian alpacas, Am. J. Vet. Res., 48, 189 (1987).

[0012] Kumar M et al., Biochemical studies on Indian Camels: Blood proteins and lipids, J. Sci. Industrial Res., 20c, 236 (1961).

[0013] Ungar-Waron H., Elias E., et al., Dromedary IgG: Purification, characterization and quantitation in Sera of Dams and newborns, Isr. J. Vet. Med., 43 (3), 198 (1987).

[0014] Azwai S M et al., The isolation and characterization of camel immunoglobulin classes and subclasses, J. Comp. Path., 109, 187 (1993).

[0015] Hamers-Casterman C., Atarhouch T., et al., Naturally occurring antibodies devoid of light chains, Nature, 363, 446 (1993).

[0016] Zhang J, Tanha J. et al., Pentamerization of single-domain antibodies from phage libraries: a novel strategy for the rapid generation of high-avidity antibody reagents, J. Mol. Biol., 335, 49 (2004).

SUMMARY OF THE INVENTION

[0017] The present invention relates to an ultrasensitive and ultraspecific method for the detection of antigens and useful for diagnosing human diseases using camelid and shark heavy chain only antibodies lacking light chain and their analogs.

[0018] In one aspect, the invention provides a method for detecting the presence or absence of an antigen in a sample. The method includes a) obtaining a sample suspected of having said antigen, b) detecting the presence or absence of the antigen in the sample utilizing a polypeptide in which the polypeptide comprises all or a portion of at least one variable antigen-binding (Vab) domain of camelid and/or shark single-domain heavy chain antibodies lacking light-chains, at least ten contiguous amino acids derived from a source other than camelid and/or shark single-domain heavy chain antibodies lacking light-chains and the polypeptide comprises at least one binding site for an antigen. The polypeptide binds specifically to the antigen and the binding is indicative of the presence of the antigen.

[0019] In another aspect, the invention provides a method for detecting the presence or absence of an antigen in a sample. The method includes a) obtaining a sample suspected of having said antigen, b) detecting the presence or absence of the antigen in the sample utilizing a composition having at least two polypeptides, in which each of the polypeptides includes all or a portion of at least one variable (Vab) domain of camelid and or shark single domain heavy chain antibody lacking light chain, all or a portion of at least one hinge region of camelid and or shark single domain heavy chain antibody lacking light chain in which at least one of the polypeptide includes at least one binding site for an antigen, and the polypeptides are linked to each other through at least one linker. The polypeptide binds specifically to the antigen and the binding is indicative of the presence or absence of the antigen. In one embodiment, at least one linker is a peptide bond. In another embodiment, at least one linker is other than a peptide bond. In one embodiment, the polypeptides of the composition include at least three, at least four, at least five or more variable antigen-binding (Vab) domains of camelid and or shark single domain heavy chain antibody. In some embodiments, the polypeptide may include one or more substitutions or deletions of the native amino acids.

[0020] In another aspect, the invention provides a method to improve the biodistribution and retention of the heavy chain only camelid and shark antibodies without light-chains and their analogs. In one embodiment, the molecular weight is greater than 15 to 17 KDa and can enter a cell or cross blood brain barrier (BBB), they are retained inside the cell to be diagnostically/therapeutically efficacious. In some embodiments, the molecular weight of the antibodies and their analogs are between ~30 to 60 KDa, more preferably 40 to 60 KDa, ideally ~55 KDa. In one embodiment, the invention encompasses the synthesis of a polypeptide with two or more variable antigen-binding domains to generate the polypeptide with a MW ~30 to 60 KDa, more preferably 40 to 60 KDa, ideally ~55 KDa. The polypeptide comprises camelid Vab domains and/or shark V-NAR domains, in which such constructions/preparations are performed either chemically and/or via recombinant DNA methods.

[0021] In another aspect, the invention provides a method for detecting an organism or a cell in a sample. The method includes a) obtaining a sample suspected of having such cell or organism, b) detecting the presence or absence of one or more antigens associated with the organism or a cell by utilizing the polypeptides or compositions of the above aspects of the invention such that the polypeptides or the compositions bind specifically to one or more antigens associated with the cell or organism and the binding is indicative of the presence or absence of a cell or organism in the sample. In some embodiments, the organism is a pathogenic organism such as bacteria or virus.

[0022] In another aspect, the invention provides a method for diagnosing an individual with one or more diseases. The method includes: a) obtaining a sample of bodily fluid from the individual; b) detecting the presence or absence of one or more biomarkers associated with the disease in which the detection comprises utilizing a polypeptide in which the polypeptide comprises all or a portion of at least one variable antigen-binding (Vab) domain of camelid and/or shark single-domain heavy chain antibodies lacking light-chains, at least ten contiguous amino acids derived from a source other than camelid and/or shark single-domain heavy chain antibodies lacking light-chains, the polypeptide binds specifically to at least one of said biomarkers and the binding of the polypeptide to one or more of the biomarkers is indicative of the presence of one or more biomarkers in the sample; c) identifying the individual as having the disease when one or more biomarkers are present in the individual's sample. In some embodiments, the method further includes determining the amount of one or more biomarkers in the sample and comparing the amount to reference values. An amount higher or lower than the reference value is indicative of a disease. In some embodiments, the reference values are the levels of the biomarkers in an individual without such one or more diseases.

[0023] In one embodiment, the polypeptide of the above aspects of the invention comprises at least two variable antigen-binding (Vab) domains of camelid and/or shark single-domain heavy-chain antibody lacking the light chains. In another embodiment, the polypeptide of the above aspects of the invention includes at least three, at least four or more variable (Vab) domains of camelid and shark heavy chain only antibody. In some embodiments, the polypeptide may include one or more substitutions or deletions of the native amino acids. In some embodiments, at least two variable antigen-binding (Vab) domains bind to two different antigens. In one embodiment of all of the above aspects of the invention, the polypeptide includes all or a portion of at least one hinge region of camelid and/or shark single domain heavy chain antibody lacking light chain.

[0024] In one embodiment of all of the above aspects of the invention, the polypeptide includes all or a portion of at least one camelid and or shark single domain heavy chain constant domain 2 (CH2). In one embodiment of all of the above aspects of the invention, the polypeptide includes all or a portion of at least one camelid and or shark single domain heavy chain constant domain 3 (CH3). In one embodiment of all of the above aspects of the invention, at least one amino acid at positions 37, 44, 45, and 47 of the Vab region is selected from the group consisting of serine, glutamine, tyrosine, histidine, asparagine, threonine, aspartic acid, glutamic acid, lysine and arginine. In some embodiments, the polypeptide may include one or more substitutions or deletions of the native amino acids.

[0025] In some embodiments of the above aspects of the invention, the polypeptide may include domains (such as variable domain or constant domain) from at least two different species such as camelid and shark, or two different camelid species such as llama, camel, alpaca and dromedaries. In some embodiments of the above aspects of the invention, the polypeptide may have improved cellular uptake, blood brain barrier permeability, biodistribution and retention.

[0026] In some embodiments the polypeptide of the above aspects of the invention is immobilized on a solid support prior to binding to said antigen. In some embodiments the polypeptide of the above aspect of the invention binds to the antigen to form a complex and the complex is immobilized on a solid support. In one embodiment, the immobilization is achieved by covalent attachment of the polypeptide to the solid surface through a spacer. In one embodiment, the length of the spacer is 1-100 nm in length. In one embodiment, the length of the spacer is 1-50 nm. In another embodiment, the length is 20 nm.

[0027] In some embodiments of the above aspects of the invention, the polypeptide is linked to at least one entity other than an antibody. In some embodiments, the entity can be solid support, radioisotope, enzyme, detectable label, ligand, fluorophore, biotin, digoxegenin, avidin, streptavidin, Fc region of IgGs, a therapeutic agent, toxin, hormone, peptide, protein, vector, siRNA, micro-RNA or nucleic acid. In some embodiments, the solid support can be beads, biosensors, nanoparticles, microchannels, microarrays, and microfluidic devices, glass slides, glass chambers, or gold particles. In some embodiments, the enzyme can be alkaline phosphatase (AP), horse-raddish-peroxidase (HRP), Luciferase, and beta-galactosidase. In some embodiments, the bead can be 1-200 micrometer in diameter, preferably 1-10 micrometer in diameter.

[0028] In some embodiments, the polypeptides or the compositions of the above aspects of the invention have structures **1**, **1***a*, **4**, **4***a*, **5**, **5***a*, **6**, **6***a*, **7**, **7***a*, **8**, **8***a*, **9**, **9***a*, **10**, **10***a* (FIG. **2**), wherein "a" represents analog of the unmodified parent antibody. For example, **1**, **1***a* represent native unmodified structure **1** without "S" and "L" whereas la contains modified structure **1** comprising of "S" and "L". Also, "CHX" in FIGS. **2** and **3** represents at least ten contiguous amino acids derived from a source other than camelid and/or shark single-domain heavy chain antibodies lacking light-chains; "S" represents a linker; "Rn" represents all or a portion of at least one camelid or shark hinge region of single domain heavy chain antibody; "L" represents an entity linked to the polypeptide, and Vab represents camelid or shark variable region of single domain heavy chain antibody, "D" represents at least two amino acids comprising at least one charged amino acid between the two domains of the camelid and shark antibodies.

[0029] In other embodiments, the polypeptides or the compositions of the above aspects of the invention have structures **2**, **2***a*, **11**, **11***a*, **12**, **12***a*, **13**, **13***a*, **14**, **14***a*, and **15**, **15***a*. (FIG. **3**), wherein "a" represents analog of the unmodified parent antibody. For example, **2**, represents native unmodified structure **2** and **2***a* represents modified structure with "S" and "L" Also, "CHX" in FIGS. **2** and **3** represents at least ten contiguous amino acids derived from a source other than camelid and/or shark single-domain heavy chain antibodies lacking light-chains; "S" represents a linker; "Rn" represents all or a portion of at least one camelid or shark hinge region of single domain heavy chain antibody; L represents an entity linked to the polypeptide, and Vab represents camelid or shark variable region of single domain heavy chain antibody, "D" represents at least two amino acids comprising at least one charged amino acid, VNAR represents shark variable antigen-binding region of single domain heavy chain only shark antibody without the light-chains. CH1, CH2, CH3, CH4 and CH5 represent five constant domains of shark antibodies.

[0030] In one embodiment, the generic composition of the antibody polypeptide is represented by:

[Vab]m-S-L

in which

Vab=Variable antigen-binding domain of camelid and/or shark single domain heavy chain antibodies;

m=1 to 10, preferably 2 to 5 such that the MW is approximately between 15 to 65 KDa for optimal biodistribution and retention in the body;

"S" is selected from the group consisting of groups I and II in which group I includes 1-20 amino acids of the hinge region of camelid and/or shark single domain heavy chain antibodies comprising at least one lysine and/or cysteine, and group II includes hetrobifunctional linker with one end being capable of covalent binding with amino- or aldehyde group of single-domain antibodies, and the other end with an entity "L";

"L" represents an entity linked to Vab domain. "L" can be a detectable label, enzyme or protein (for example, horse radish peroxidase, alkaline phosphatase, luciferase, beta-galactosidase, and streptavidin), antibody, nucleic acid (for example, DNA, Modified DNA, Locked-DNA, PNA (Peptide Nucleic Acids), RNA, Si-RNA, Micro-RNA (MiRNA), mRNA, RNA-Conjugates/Modifications), radionucleotides (for example, Fluorine-18, Gallium-67, Krypton-81m, Rubidium-82, Technetium-99m, Indium-111, Iodine-123, Xenon-133, and Thallium-201, Yttrium-90, and Iodine-131), toxins (for example, Immunotoxins, Ricin, Saporin, Maytansinoid, and Calicheamicin), solid support (for example, Microchannels, Microfluidic Device, Microarrays, Biosensors, Glass Slides, Glass Chambers, Magnetic Beads, and Gold Nanoparticles), and therapeutic agents (for example, nucleolytic enzymes, antibiotics, and chemotherapeutic agents such as Paclitaxel its derivatives).

[0031] In one embodiment, the generic composition of "S" is

S=X-P-Y in which X can be of NHS (N-Hydroxy-Succinimide), sulfo-NHS, CHO, COOH, CN, SCN, epoxide, phosphate and other moieties capable of forming covalent bond with NH2 groups of single-domain antibodies; Y can be maleimido, NHS, sulfo-NHS, SH, COOH, SCN, NH2, and epoxide, capable of forming a covalent bond with the thiol group of the detectable label; P can be $(CH_2CH_2O)n$, wherein

n=1-500; DNA, modified DNA, modified RNA; $(CH_2)n^1$, wherein n^1=1-15; $(Ra\text{-}NHCO)n^2$, wherein n^2=1-100; Ra=charged amino acid; nucleic acids; nylon, polystyrene; polypropylene; protein; and chimeric protein-nucleic acids.

[0032] In some embodiments, the disease may be cancer, Parkinson's disease, Alzheimer's disease, AIDS, Lyme disease, malaria, SARS, Down syndrome, anthrax, salmonella or bacterial botulism, staphylococcus aureus. In some embodiments, the cancer can be lung cancer, bladder cancer, gastric cancer, ovarian cancer, brain cancer, breast cancer, prostate cancer, cervical cancer, ovarian cancer, oral cancer, colorectal cancer, leukemia, childhood neuroblastoma, or Non-Hodgkin's lymphoma.

[0033] In some embodiments of the above aspects of the invention, the polypeptide can bind specifically one or more biomarkers. Exemplary biomarkers include AMACR, TMPRSS2-ERG, HAAH, APP, Aβ42, ALZAS, Tau, gamma secretase, beta secretase, PEDF, BDNF, Cystatin C, VGF nerve growth factor inducible, APO-E, GSK-3 binding protein, TEM1, PGD2, EGFR, ESR-1, HER-2/neu, P53, RAS, SMAD4, Smad7, TNF-alfa, HPV, tPA, PCA-3, Mucin, Cadherin-2, FcRn alpha chain, cytokeratins 1-20, Apo-H, Celuloplasmin, Apo AII, VGF, Vif, LEDGF/p75, TS101, gp120, CCR5, HIV protease, HIV integrase, Bacillus anthracis protein, NadD (Nicotinate Mononucletide Adenyltransferase), Plasmodium falciparum, cGMP directed phosphodiestrase, chain B of Clostridium botulinum neurotroxin type E protein, *Borrelia* VlsE protein, ACE2 receptor, SFRS4, or SAMP.

[0034] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with Alzheimer's disease. Exemplary biomarkers associated with Alzheimer's disease include but are not limited to Amyloid-beta, ALZAS, Tau, Cyclophilin-D (Cyp-D), Abeta binding alcohol dehydrogenase (ABAD), N-methyl-D-aspartate receptor (NMDAR), mSOD1, mHTT (mutant huntingtin), 3-NP, phosphatidylserine (PtDS), MPTP, integrin α4β1, integrin-α4β7, PPAR-γ, MAdCAM-1, DJ-1, Bax-1, PEDF, HPX, Cystatin-C, Beta-2-Microglobulin, BDNF, Tau-Kinase, γ-Secretase, β-Secretase, Apo-E4, and VGF-Peptide, TOM, hPReP, PLSCR1, integrin-DJ-1, and enzymes involved in the mitochondrial and myelin dysfunction.

[0035] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with Parkinson's disease. Exemplary biomarkers associated with Parkinson's disease include but are not limited to Apo-H, Cerulopasmin, Chromogranin-B, VDBP, Apo-E, Apo-AII, and alfa-Synuclein.

[0036] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with brain cancer. Exemplary biomarkers associated with brain cancer include but are not limited to TEM1, Plasmalemmal Vesicle (PV-1), Prostaglandin D Synthetase, and (PGD-S).

[0037] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with HIV/AIDS. Exemplary biomarkers associated with HIV/AIDS include but are not limited to gp120, Vif, LEDGF/p75, TS101, HIV-Integrase, HIV-Reverse Transcriptase, HIV-Protease, CCR5, and CXCR4.

[0038] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with lung cancer. Exemplary biomarkers associated with lung cancer include but are not limited to KRAS, Ki67, EGFR, KLKB1, EpCAM, CYFRA21-1, tPA, ProGRP, Neuron-specific Enolase (NSE), and hnRNP.

[0039] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with prostate cancer. Exemplary biomarkers associated with prostate cancer include but are not limited to AMACR, PCA3, TMPRSS2-ERG, HEPSIN, B7-H3, SSeCKs, EPCA-2, PSMA, BAG-1, PSA, MUC6, hK2, PCA1, PCNA, RKIP, and c-HGK.

[0040] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with breast cancer. Exemplary biomarkers associated with breast cancer include but are not limited to EGFR, EGFRT790M, HER-2, Notch-4, ALDH-1, ESR1, SBEM, HSP70, hK-10, MSA, p53, MMP-2, PTEN, Pepsinigen-C, Sigma-S, Topo-11-alfauKPA, BRCA-1, BRCA-2, SCGB2A1, and SCGB1D2.

[0041] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with colorectal cancer. Exemplary biomarkers associated with colorectal cancer include but are not limited to SMAD4, EGFR, KRAS, p53, TS, MSI-H, REGIA, EXTL3, p1K3CA, VEGF, HAAH, EpCAM, TEM8, TK1, STAT-3, SMAD-7, beta-Catenin, CK20, MMP-1, MMP-2, MMP-7,9,11, and VEGF-D.

[0042] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with ovarian cancer. Exemplary biomarkers associated with ovarian cancer include but are not limited to CD24, CD34, EpCAM, hK8, 10, 13, CKB, Cathesin B, M-CAM, c-ETS1, and EMMPRIN.

[0043] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with cervical cancer. Exemplary biomarkers associated with cervical cancer include but are not limited to HPV, CD34, ERCC1, Beta-CF, Id-1, UGF, SCC, p16, p21WAF1, PP-4, and TPS.

[0044] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with bladder cancer. Exemplary biomarkers associated with bladder cancer include but are not limited to CK18, CK20, BLCa1, BLCA-4, CYFRA21-1, TFT, BTA, Survivin, UCA1, UPII, FAS, and DD23.

[0045] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with a pathogenic bacteria. Exemplary pathogenic bacteria include but are not limited to *Clostridium Botulinum* (Bacterial Botulism), *Bacillus Anthracis* (Anthrax), *Salmonella Typhi* (Typhoid Fever), *Treponema Pallidum* (Syphilis), *Plasmodinum* (Malaria), *Chlamadyia* (STDs), *Borrelia* B (Lyme disease), *Staphyloccus Aureus*, Tetanus, Meningococcal Meningitis (Bacterial Meningitis), and *Mycobacterium tuberculosis* (Tuberculosis, TB), and NadD (Nicotinate Mononucleotide Adenyltransferase, an enzyme involved in inducing resistance to antibiotics).

[0046] In some embodiments of the above aspects of the invention, the polypeptide may specifically bind to the biomarkers associated with a pathogenic virus. Exemplary pathogenic virus include but are not limited to Pandemic Flu Virus H1N1 strain, Influenza virus H5N1 strain, Hepatitis B virus (HBV) antigen OSt-577, HBV core antigen HBcAg (HBV), HBV antigen Wnt-1, Hepatitis C Virus (HCV) antigen Wnt-1, and HCV RNA (HCV).

[0047] In some embodiments, the antibody is produced using chemical methods as described in pending U.S. patent application Ser. No. 12563330. Briefly, the method includes a chemical synthesis of a polypeptide comprising one, two, or more variable antigen-binding (Vab) domains using the parent antibody produced from camelid and/or shark as a starting material for generating the polypeptide with one or more Vab domains.

[0048] Still in another embodiment, the invention provides a method for generating polypeptides comprising multivalent variable antigen-binding domains improving binding affinity between antibody and its antigen.

[0049] In some embodiments, the antibody is produced using recombinant DNA methods as described in pending U.S. patent application Ser. No. 12563330. Briefly, the method includes isolating the RNA from lymphocytes, reverse-transcription with oligo-dT priming, amplification of the generated cDNA encoding the camelid or shark antibody, inserting the amplified DNA in a phage-display vector, transforming the *E. Coli* cells, and selecting the clones that express highly specific antibodies.

[0050] The term "antibody" as used herein refers to immunoglobulin G (IgG) having only heavy chains without the light-chain and also constant domain 1 (CH1) in case of camelid antibodies. The shark antibody without the light-chains is known in the art as shark IgNAR. An antibody of this invention can be monoclonal or polyclonal.

[0051] The term "analog" within the scope of the term "antibody" include those produced by digestion with various proteases, those produced by chemical cleavage, chemical coupling, chemical conjugation, and those produced recombinantly, so long as the fragment remains capable of specific binding to a target molecule. Analogs within the scope of the term include antibodies (or fragments thereof) that have been modified in sequence, but remain capable of specific binding to a target molecule, including: interspecies chimeric and humanized antibodies; antibody fusions; heteromeric antibody complexes and antibody fusions, such as diabodies (bispecific antibodies), single-chain diabodies, and intrabodies (see, e.g., Marasco (ed.), Intracellular Antibodies: Research and Disease Applications, Springer-Verlag New York, Inc. (1998) (ISBN: 3540641513). As used herein, antibodies can be produced by any known technique, including harvest from cell culture of native B lymphocytes, harvest from culture of hybridomas, recombinant expression systems, and phage display.

[0052] The terms "heavy chain only antibody" and "single domain heavy chain antibody" has been used herein interchangeably in the context of camelid and shark antibodies and refer to camelid immunoglobulin G (IgG) and shark IgNAR having only heavy chains without the heavy chain constant domain 1 (CH1) and further lacking the light chain such as camelids IgG2 and IgG3 and shark IgNAR. Heavy chain only antibody can be monoclonal or polyclonal.

[0053] The term "improved biodistribution and retention" as used herein in the context of polypeptides, antibodies and its analogs refers to polypeptides, antibodies and its analogs that can cross cell membrane and blood brain barrier (BBB) and have greater thermal and chemical stability than conventional immunoglobulin G with heavy and light chains. Typically such polypeptides, antibodies and its analogs have molecular weight between 25 to 90 KDa, preferably between 30 to 60 KDa. In some embodiments, the molecular weight is at least 25 KDa, 30 KDa, 35 KDa, 40 KDa, 45 KDa, 50 KDa, 55 KDa, 60 KDa, 65 KDa, 70 KDa, 75 KDa, 80 KDa, 85 KDa, or 90 KDa. Although larger and smaller molecular weights are possible.

[0054] The term "specifically binds to" as used herein in the context of an antibody or its analogs refers to binding of an antibody or its analogs specifically to an epitope such that the antibody or its analog can distinguish between two proteins with and without such epitope.

[0055] The terms "biomarker" and antigen is used interchangeably and refer to a molecule or group of molecules comprised of nucleic acids, carbohydrates, lipids, proteins, peptides, enzymes and antibodies which is associated with a disease, physiological condition, or an organism. An organism can be pathogenic or nonpathogenic. A biomarker may not necessarily be the reason for a disease or a physiological condition. An amount of a biomarker may be increased or decreased in disease or a physiological condition.

[0056] The term "camelid" as used herein refers to members of the biological family Camelidae in the Order: Artiodactyla, Suborder: Tylopoda. Exemplary members of this group include camels, dromedaries, llamas, alpacas, vicunas, and guanacos.

[0057] The term "shark" as used herein refers to members that belong to the super order Selachimorpha in the subclass Elasmobranchii in the class Chondrichthyes. There are more than 400 species of sharks known. Exemplary members of the class Chondrichthyes include great white sharks, hound-sharks, cat sharks, hammerhead sharks, blue, tiger, bull, grey reef, blacktip reef, Caribbean reef, blacktail reef, whitetip reef, oceanic whitetip sharks, zebra sharks, nurse sharks, wobbegongs, bramble sharks, dogfish, roughsharks, and prickly sharks.

[0058] The term "a portion of" in the context of antibodies such as camelid and shark heavy chain only antibodies and their analogs, or human antibodies means at least 2, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 125, 150, 200, 250, 300, 350, 400 or more amino acids.

[0059] The term "a portion of" in the context of hinge region of camelid and shark single domain heavy chain antibodies means at least 1, 2, 5, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 125, 150, 200, or more amino acids of the hinge region.

[0060] The terms "diagnose" or "diagnosis" as used herein refers to the act or process of identifying or determining a disease or condition in an organism or the cause of a disease or condition by the evaluation of the signs and symptoms of the disease or disorder. Usually, a diagnosis of a disease or disorder is based on the evaluation of one or more factors and/or symptoms that are indicative of the disease. That is, a diagnosis can be made based on the presence, absence or amount of a factor which is indicative of presence or absence of the disease or condition. Each factor or symptom that is considered to be indicative for the diagnosis of a particular disease does not need be exclusively related to the particular disease; i.e. there may be differential diagnoses that can be inferred from a diagnostic factor or symptom. Likewise, there may be instances where a factor or symptom that is indicative of a particular disease is present in an individual that does not have the particular disease.

[0061] The term "reference value" as used herein means a value which can be used for comparison with a biomarker under investigation. In one case, a reference value may be the level of a biomarker under investigation from one or more individuals without any known disease. In another case, a

reference value may be the level of the biomarker in an individual's sample collected at a different time.

[0062] "Sample" or "patient sample" as used herein includes biological samples such as cells, tissues, bodily fluids, and stool. "Bodily fluids" may include, but are not limited to, blood, serum, plasma, saliva, cerebral spinal fluid, pleural fluid, tears, lactal duct fluid, lymph, sputum, urine, amniotic fluid, and semen. A sample may include a bodily fluid that is "acellular". An "acellular bodily fluid" includes less than about 1% (w/w) whole cellular material. Plasma or serum are examples of acellular bodily fluids. A sample may include a specimen of natural or synthetic origin.

[0063] The term "body fluid" or "bodily fluid" as used herein refers to any fluid from the body of an animal. Examples of body fluids include, but are not limited to, plasma, serum, blood, lymphatic fluid, cerebrospinal fluid, synovial fluid, urine, saliva, mucous, phlegm and sputum. A body fluid sample may be collected by any suitable method. The body fluid sample may be used immediately or may be stored for later use. Any suitable storage method known in the art may be used to store the body fluid sample; for example, the sample may be frozen at about 20° C. to about −70° C. Suitable body fluids are acellular fluids. "Acellular" fluids include body fluid samples in which cells are absent or are present in such low amounts that the peptidase activity level determined reflects its level in the liquid portion of the sample, rather than in the cellular portion. Typically, an acellular body fluid contains no intact cells. Examples of acellular fluids include plasma or serum, or body fluids from which cells have been removed.

[0064] The term "enzyme linked immunosorbent assay" (ELISA) as used herein refers to an antibody-based assay in which detection of the antigen of interest is accomplished via an enzymatic reaction producing a detectable signal. ELISA can be run as a competitive or noncompetitive format. ELISA also includes a 2-site or "sandwich" assay in which two antibodies to the antigen are used, one antibody to capture the antigen and one labeled with an enzyme or other detectable label to detect captured antibody-antigen complex. In a typical 2-site ELISA, the antigen has at least one epitope to which unlabeled antibody and an enzyme-linked antibody can bind with high affinity. An antigen can thus be affinity captured and detected using an enzyme-linked antibody. Typical enzymes of choice include alkaline phosphatase or horseradish peroxidase, both of which generated a detectable product upon digestion of appropriate substrates.

[0065] The term "label" as used herein, refers to any physical molecule directly or indirectly associated with a specific binding agent or antigen which provides a means for detection for that antibody or antigen. A "detectable label" as used herein refers any moiety used to achieve signal to measure the amount of complex formation between a target and a binding agent. These labels are detectable by spectroscopic, photochemical, biochemical, immunochemical, electromagnetic, radiochemical, or chemical means, such as fluorescence, chemifluoresence, or chemiluminescence, electro-chemiluminescence or any other appropriate means. Suitable detectable labels include fluorescent dye molecules or fluorophores.

[0066] The terms "polypeptide," "protein," and "peptide" are used herein interchangeably to refer to amino acid chains in which the amino acid residues are linked by peptide bonds or modified peptide bonds. The amino acid chains can be of any length of greater than two amino acids. Unless otherwise specified, the terms "polypeptide," "protein," and "peptide" also encompass various modified forms thereof. Such modified forms may be naturally occurring modified forms or chemically modified forms. Examples of modified forms include, but are not limited to, glycosylated forms, phosphorylated forms, myristoylated forms, palmitoylated forms, ribosylated forms, acetylated forms, ubiquitinated forms, etc. Modifications also include intramolecular crosslinking and covalent attachment to various moieties such as lipids, flavin, biotin, polyethylene glycol or derivatives thereof, etc. In addition, modifications may also include cyclization, branching and cross-linking. Further, amino acids other than the conventional twenty amino acids encoded by genes may also be included in a polypeptide.

[0067] The term "detectable label" as used herein in the context of antibody or its analogs refers to a molecule or a compound or a group of molecules or a group of compounds associated with a binding agent such as an antibody or its analogs, secondary antibody and is used to identify the binding agent bound to its target such as an antigen, primary antibody. A detectable label can also be used in to detect nucleic acids. In such cases a detectable label may be incorporated into a nucleic acid during amplification reactions or a detectable label may be associated a probe to detect the nucleic acid.

[0068] The terms "ultrasensitive" or "ultrasensitivity" as used herein in the context of antibodies refers to the detection of fewer than 200 molecules of the pathogenic proteins from patient's sample.

[0069] "Detecting" as used herein in context of detecting a signal from a detectable label to indicate the presence of a nucleic acid of interest in the sample (or the presence or absence of a protein of interest in the sample) does not require the method to provide 100% sensitivity and/or 100% specificity. As is well known, "sensitivity" is the probability that a test is positive, given that the person has a genomic nucleic acid sequence, while "specificity" is the probability that a test is negative, given that the person does not have the genomic nucleic acid sequence. A sensitivity of at least 50% is preferred, although sensitivities of at least 60%, at least 70%, at least 80%, at least 90% and at least 99% are clearly more preferred. A specificity of at least 50% is preferred, although specificity of at least 60%, at least 70%, at least 80%, at least 90% and at least 99% are clearly more preferred. Detecting also encompasses assays with false positives and false negatives. False negative rates may be 1%, 5%, 10%, 15%, 20% or even higher. False positive rates may be 1%, 5%, 10%, 15%, 20% or even higher.

[0070] The term "about" as used herein in reference to quantitative measurements or values, refers to the indicated value plus or minus 10%.

[0071] "Nucleic acid" as used herein refers to an oligonucleotide, nucleotide or polynucleotide, and fragments or portions thereof, which may be single or double stranded, and represent the sense or antisense strand. A nucleic acid may include DNA or RNA, and may be of natural or synthetic origin and may contain deoxyribonucleotides, ribonucleotides, or nucleotide analogs in any combination.

[0072] Non-limiting examples of polynucleotides include a gene or gene fragment, genomic DNA, exons, introns, mRNA, tRNA, rRNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, synthetic nucleic acid, nucleic acid probes and

primers. Polynucleotides may be natural or synthetic. Polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs, uracyl, other sugars and linking groups such as fluororibose and thiolate, and nucleotide branches. A nucleic acid may be modified such as by conjugation, with a labeling component. Other types of modifications included in this definition are caps, substitution of one or more of the naturally occurring nucleotides with an analog, and introduction of chemical entities for attaching the polynucleotide to other molecules such as proteins, metal ions, labeling components, other polynucleotides or a solid support. Nucleic acid may include nucleic acid that has been amplified (e.g., using polymerase chain reaction).

[0073] A fragment of a nucleic acid generally contains at least about 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 200, 300, 400, 500, 1000 nucleotides or more. Larger fragments are possible and may include about 2,000, 2,500, 3,000, 3,500, 4,000, 5,000 7,500, or 10,000 bases.

[0074] "Gene" as used herein refers to a DNA sequence that comprises control and coding sequences necessary for the production of an RNA, which may have a non-coding function (e.g., a ribosomal or transfer RNA) or which may include a polypeptide or a polypeptide precursor. The RNA or polypeptide may be encoded by a full length coding sequence or by any portion of the coding sequence so long as the desired activity or function is retained.

[0075] "cDNA" as used herein refers to complementary or copy polynucleotide produced from an RNA template by the action of RNA-dependent DNA polymerase activity (e.g., reverse transcriptase). cDNA can be single stranded, double stranded or partially double stranded.

[0076] cDNA may contain unnatural nucleotides. cDNA can be modified after being synthesized. cDNA may comprise a detectable label.

[0077] As used herein, "subject" or "individual" is meant a human or any other animal that has cells. A subject can be a patient, which refers to a human presenting to a medical provider for diagnosis or treatment of a disease. A human includes pre and post natal forms.

[0078] The term "patient" as used herein, refers to one who receives medical care, attention or treatment. As used herein, the term is meant to encompass a person diagnosed with a disease as well as a person who may be symptomatic for a disease but who has not yet been diagnosed.

[0079] The term "vector or phagemid" as used herein refers to a recombinant DNA or RNA plasmid or virus that comprises a heterologous polynucleotide capable of being delivered to a target cell, either in vitro, in vivo or ex-vivo. The heterologous polynucleotide can comprise a sequence of interest and can be operably linked to another nucleic acid sequence such as promoter or enhancer and may control the transcription of the nucleic acid sequence of interest. As used herein, a vector need not be capable of replication in the ultimate target cell or subject. The term vector may include expression vector and cloning vector.

[0080] Suitable expression vectors are well-known in the art, and include vectors capable of expressing a polynucleotide operatively linked to a regulatory sequence, such as a promoter region that is capable of regulating expression of such DNA. Thus, an expression vector refers to a recombinant DNA or RNA construct, such as a plasmid, a phage, recombinant virus or other vector that, upon introduction into an appropriate host cell, results in expression of the inserted DNA. Appropriate expression vectors include those that are replicable in eukaryotic cells and/or prokaryotic cells and those that remain episomal or those which integrate into the host cell genome.

[0081] The term "promoter" as used herein refers to a segment of DNA that controls transcription of polynucleotide to which it is operatively linked. Promoters, depending upon the nature of the regulation, may be constitutive or regulated. Exemplary eukaryotic promoters contemplated for use in the practice of the present invention include the SV40 early promoter, the cytomegalovirus (CMV) promoter, the mouse mammary tumor virus (MMTV) steroid-inducible promoter, Moloney murine leukemia virus (MMLV) promoter. Exemplary promoters suitable for use with prokaryotic hosts include T7 promoter, beta-lactamase promoter, lactose promoter systems, alkaline phosphatase promoter, a tryptophan (trp) promoter system, and hybrid promoters such as the lac promoter.

[0082] The term "antibody" as used herein refers to immunoglobulin G (IgG) having only heavy chains without the heavy chain constant domain 1 (CH1) and also lacking the light chain such as in shark IgNAR and camelids IgG2 and IgG3. Antibody can be monoclonal or polyclonal.

[0083] The term "analog" within the scope of the term "antibody" include those produced by digestion with various proteases, those produced by chemical cleavage, chemical coupling, chemical conjugation, and those produced recombinantly, so long as the fragment remains capable of specific binding to a target molecule. Analogs within the scope of the term include antibodies (or fragments thereof) that have been modified in sequence, but remain capable of specific binding to a target molecule, including: interspecies chimeric and humanized antibodies; antibody fusions; heteromeric antibody complexes and antibody fusions, such as diabodies (bispecific antibodies), single-chain diabodies, and intrabodies (see, e.g., Marasco (ed.), Intracellular Antibodies: Research and Disease Applications, Springer-Verlag New York, Inc. (1998) (ISBN: 3540641513). As used herein, antibodies can be produced by any known technique, including harvest from cell culture of native B lymphocytes, harvest from culture of hybridomas, recombinant expression systems, and phage display.

[0084] The terms "heavy chain only antibody" and "single domain heavy chain antibody" has been used herein interchangeably in the context of camelid and shark antibodies and refer to camelid immunoglobulin G (IgG) and shark IgNAR having only heavy chains without the heavy chain constant domain 1 (CH1) and further lacking the light chain such as camelids IgG2 and IgG3 and shark IgNAR. Heavy chain only antibody can be monoclonal or polyclonal.

[0085] Unless otherwise specified, the terms "a" or "an" mean "one or more" throughout this application.

BRIEF DESCRIPTION OF THE FIGURES

[0086] FIG. **1** shows structural differences between camel, shark, and mouse immunoglobulins (IgGs). The notations CH1, CH2, CH3, CH4, CH5 represent constant domain 2, 3, 4 of single domain heavy chain antibody of the respective species. The notations Vab and VNAR represent variable domain of camelid and shark single domain heavy chain antibodies respectively.

[0087] FIG. **2** shows the structure of exemplary analogs of camelid single-domain antibodies without the light-chains: mini-antibody **1** and its analogs **1***a*; micro-antibody **4** and its

analogs **4***a*; sub-nano-antibody **5** and its analogs **5***a*; nano-antibody **6** and its analogs **6***a*; dimeric nano-antibody **7** and its analogs **7***a*; trimeri-nanoantibody **8** and its analogs; and tetrameric nano-antibody **9** and its analogs **9***a*. The notation "Rn" represents all or portion of the hinge region of camelid or shark single domain antibodies. CHX represents segment of human IgG CH1 domain or CH2 domain of camelid antibody. "S" stands for a spacer or linker. "L" is a ligand.

[0088] FIG. **3** shows the structure of exemplary analogs of shark single-domain antibodies without the light-chains: Hark IgNAR **2** and its analogs **2***a*; shark mini-antibody **11** and its analogs **11***a*; shark micro-antibody **12** and its analogs **12***a*; shark sub-nano-antibody **13** and its analogs **13***a*; shark dimeric nano-antibody **14** and its analogs **14***a*; and shark tetrameric nano-antibody **15** and **15***a*. The notation "Rn" represents all or portion of the hinge region of shark single domain antibodies. CHX represents segment of human IgG or CH1 domain of shark antibody. "S" stands for a spacer or linker. "L" is a ligand.

[0089] FIG. **4** shows the steps involved in the chemical synthesis of exemplary analogs represented by structures **1***a* and **4***a*, respectively, of camelid mini-antibody **1** and micro-antibody **4**. The notation "Rn" represents all or portion of the hinge region of camelid or shark single domain antibodies.

[0090] FIG. **5** shows the steps involved in the chemical transformation of exemplary sub-nano-antibody **5** into its analogs represented by structure **5***a*, and the synthesis of dimeric camelid nano-antibody **7** and its analogs represented by generic structure **7***a*.

[0091] FIG. **6** shows the steps involved in the transformation of exemplary camelid dimeric nano-antibody **7** into trimeric and tetrameric nano-antibodies. The notation "Rn" represents all or portion of the hinge region of camelid or shark single domain antibodies.

[0092] FIG. **7** shows the steps involved in the cloning and expression of exemplary shark IgNAR **2**, exemplary shark micro-antibody **12**, exemplary shark sub-nano-antibody **13** and shark-nano-antibody **30**. The notation "Rn" represents all or portion of the hinge region of camelid or shark single domain antibodies.

[0093] FIG. **8** shows the steps involved in the chemical synthesis of exemplary analogs of shark antibodies without the light chains: Shark IgNAR analogs represented by structure **2***a*; shark mini-antibody analogs represented by structure **11***a*; shark micro-antibody analogs represented by structure **12***a*; shark sub-nano-antibody analogs represented by **13***a*; shark dimeric nano-antibody analogs represented by **14***a*; and shark tetrameric -nano-antibody analogs represented by **15***a*. The notation "Rn" represents all or portion of the hinge region of camelid or shark single domain antibodies.

[0094] FIG. **9** shows the steps involved in the chemical synthesis of exemplary shark dimeric nano-antibody **14** and its conversion into exemplary shark trimeric and tetrameric nano-antibodies **32** and **31**.

[0095] FIG. **10** shows the steps involved in the immobilization of exemplary single-domain camelid and shark antibodies deprived of light chains having the structures **1**, **2**, **4**, **5**, **6**, **7**, **8**, **9**, **11**, **12**, **13**, **14**, **15**, **19**, **20**, **31**, and **32**.

[0096] FIG. **11** shows an exemplary scheme of capturing and detecting antigens/biomarkers associated with a disease using camelid and shark heavy chain only antibodies and their analogs.

[0097] FIG. **12** shows an exemplary scheme of capturing and detecting antigens/biomarkers associated with a disease using immobilized shark single-domain IgNAR and their analogs.

[0098] FIG. **13** shows an exemplary scheme of capturing and detecting <200 copies of antigens/biomarkers associated with a disease using camelid and shark heavy chain only antibodies and their analogs using immuno-PCR.

[0099] FIG. **14** shows an exemplary scheme of capturing and detecting circulating tumor cells from bodily fluid using camelid and shark antibodies.

[0100] FIG. **15** shows an exemplary scheme of detecting prenatal genetic disorder using captured circulating fetal cells using camelid and shark heavy chain only antibodies and their analogs.

[0101] FIG. **16** shows an exemplary scheme of detecting chromosomal translocation using captured circulating tumor cells using camelid and shark heavy chain only antibodies and their analogs.

[0102] FIG. **17** shows an exemplary nucleic acid sequence encoding human Cyclophilin D.

[0103] FIG. **18** shows an exemplary nucleic acid sequence encoding alpha beta binding Mitochondrial Alcohol Dehydrogenase (ABAD).

[0104] FIG. **19** shows an exemplary nucleic acid sequence encoding Translocase of the Outer Membrane (TOM).

[0105] FIG. **20** shows an exemplary nucleic acid sequence encoding Prosequence Protease (hPreP).

[0106] FIG. **21** shows an exemplary nucleic acid sequence encoding *Homo sapiens* integrin beta 1.

[0107] FIG. **22** shows an exemplary nucleic acid sequence encoding *Homo sapiens* mucosal vascular addressin cell adhesion molecule 1 (MADCAM1).

[0108] FIG. **23** shows an exemplary nucleic acid sequence encoding Cu/Zn-superoxide dismutase (mSOD1).

[0109] FIG. **24** shows an exemplary nucleic acid sequence encoding *Mus musculus* mRNA for MPTPdelta.

[0110] FIG. **25** shows an exemplary nucleic acid sequence encoding *Homo sapiens* huntingtin (HTT).

[0111] FIG. **26** shows an exemplary nucleic acid sequence encoding N-Methyl-D-Aspartate Receptor (NMDAR).

[0112] FIG. **27** shows an exemplary nucleic acid sequence encoding Phosphatidylserine Synthase (PTDS).

DETAILED DESCRIPTION OF THE INVENTION

[0113] The present invention discloses the use of camelid and/or shark single-domain heavy-chain only antibodies and their analogs for ultrasensitive detection of antigens. The method is useful for diagnosing human diseases at an early stage of their manifestation, when the concentration of antigens associated with such diseases is very low for example, 200 or fewer molecules in 0.1 ml of the bodily fluid. The invention also teaches methods for the development of nano-biomedical technology platforms utilizing camelid and/or shark heavy-chain only antibodies and their analogs for in-vitro diagnosis of human and animal diseases with such antibodies.

Camelid and Shark Antibodies

[0114] The hetero-tetrameric structure of antibodies exists in humans and most animals but the single-domain heavy-chain only dimeric structure, without the light-chains, is considered characteristic of camelids and sharks [Nature Bio-

technology, 23, 1126 (2005)]. These antibodies are relatively simple molecules but with unique characteristics. Their size is about ⅔rd the size of traditional antibodies, hence a lower molecular weight (about 90 KDa), with similar antigen binding affinity, but with water solubility 100 to 1000 folds higher than the conventional antibodies. Because of the lower molecular weight, the authors of this application call these antibodies as "Single-domain Mini-antibodies" (sdMnAbs) or simply "Mini-Antibodies" (MnAbs).

[0115] Another characteristic of the single-domain antibodies derived from sharks and camelids is that they have very high thermal stability compared to the conventional mAbs. For example, camel antibodies can maintain their antigen binding ability even at 90° C. [Biochim. Biophys. Acta., 141, 7 (1999)]. Furthermore, complementary determining region 3 (CDR3) of camel Vab region is longer, comprising of 16-21 amino acids, than the CDR3 of mouse VH region comprising only of 9 amino acids [Protein Engineering, 7, 1129 (1994)]. The larger length of CDR3 of camel Vab region is responsible for higher diversity of antibody repertoire of camel antibodies.

[0116] In addition to being devoid of light chains, the camel heavy-chain antibodies also lack the first domain of the constant region called CH1, though the shark antibodies do have CH1 domain and two additional constant domains CH4 and CH5 [Nature Biotech. 23, 1126 (2005)]. Furthermore, the hinge regions of camel and shark antibodies have an amino acid sequence different from that of normal heterotetrameric conventional antibodies [(S. Muyldermans, Reviews in Mol. Biotech., 74, 277 (2001)]. Without the light chain, these antibodies bind to their antigens by the variable antigen-binding domain of the heavy-chain immunoglobulin, which is referred to as Vab by the authors of this application (VHH in the literature), to distinguish it from the variable domain VH of the conventional antibodies. The single-domain Vab is amazingly stable by itself without having to be attached to the parent antibody. This smallest intact and independently functional antigen-binding fragment Vab, with a molecular weight of ~12-15 KDa, is referred to as nano-antibody by the authors of this application. In the literature, it is known as nanobody [(S. Muyldermans, Reviews in Mol. Biotech., 74, 277 (2001)].

[0117] The genes encoding these full length single-domain heavy-chain antibodies and antibody-antigen binding fragment Vab (camel and shark) can be cloned in phage display vectors, and selection of antigen binders by panning and expression of selected VHH in bacteria offer a very good alternative procedure to produce these antibodies on a large scale. Also, only one domain has to be cloned and expressed to produce in vivo an intact, matured antigen-binding fragment.

[0118] There are structural differences between the variable regions of single domain antibodies and conventional antibodies. Conventional antibodies have three constant domains while camel has two and shark has five constant domains. The largest structural difference is, however, found between a VH (conventional antibodies) and Vab (heavy-chain only antibodies of camel and shark) (see below) at the hypervariable regions. Camelid Vab and shark V-NAR domains each display surface loops which are larger than for conventional murine and human IgGs, and are able to penetrate cavities in target antigens, such as enzyme active sites and canyons in viral and infectious disease biomarkers [PNAS USA., 101, 12444 (2004); Proteins, 55, 187 (2005)]. In human and mouse, the VH loops are folded in a limited number of canonical structures. In contrast, the antigen binding loop of Vab possess many deviations of these canonical structures that specifically bind into such active sites, therefore, represent powerful tool to modulate biological activities [(K. Decanniere et al., Structure, 7, 361 (2000)]. The high incidence of amino acid insertions or deletions, in or adjacent to first and second antigen-binding loops of Vab will undoubtedly diversify, even further, the possible antigen-binding loop conformations.

[0119] Though there are structural differences between camel and shark parent heavy-chain antibodies (FIG. **1**), the antigen-antibody binding domains, Vab and V-NAR, respectively, have similar binding characteristics. The chemical and/or protease digestion of camel and shark antibodies results in Vab and V-NAR domains, with similar binding affinities to the target antigens [Nature Biotechnology, 23, 1126 (2005)].

[0120] Other structural differences are due to the hydrophilic amino acid residues which are scattered throughout the primary structure of Vab domain. These amino acid substitutions are, for example, Leu45 to R (arginine) or Leu45 to C (cysteine); Val37 to Y (Tyr); G44 to E (Glu), and W47 (Trp) to G (Gly). Therefore, the solubility of Vab is much higher than the Fab fragment of conventional mouse and human antibodies.

[0121] Another characteristic feature of the structure of camelid Vab and shark V-NAR is that it often contains a cysteine residue in the CDR3 in addition to cysteines that normally exist at positions 22 and 92 of the variable region. The cysteine residues in CDR3 form S—S bonds with other cysteines in the vicinity of CDR1 or CDR2 [Protein Engineering, 7, 1129 (1994)]. CDR1 and CDR2 are determined by the germline V gene. They play important roles together with CDR3 in antigenic binding [Nature Structural Biol., 9, 803 (1996); J. Mol. Biol., 311, 123 (2001)]. Like camel CDR3, shark also has elongated CDR3 regions comprising of 16-27 amino acids residues [Eur. J. Immunol., 35, 936 (2005)].

[0122] The germlines of dromedaries and llamas are classified according to the length of CDR2 and cysteine positions in the V region [Nguyen et al., EMBO J., 19, 921 (2000); Harmsen et al., Mol. Immun., 37, 579 (2000)].

[0123] Immunization of camels with enzymes generates heavy-chain antibodies (HCAb) significant proportions of which are known to act as competitive enzyme inhibitors that interact with the cavity of the active site [(M. Lauwereys et al., EMBO, J. 17, 3512 (1998)]. In contrast, the conventional antibodies that are competitive enzyme inhibitors cannot bind into large cavities on the antigen surface. Camel antibodies, therefore, recognize unique epitopes that are out of reach for conventional antibodies.

[0124] Production of inhibitory recombinant Vab that bind specifically into cavities on the surface of variety of enzymes, namely, lysozyme, carbonic anhydrase, alfa-amylase, and beta-lactamase has been achieved [M. Lauwereys, et al., EMBO, J. 17, 3512 (1998)]. Hepatitis C protease inhibitor from the camelised human VH has been isolated against an 11 amino acid sequence of the viral protease [F. Martin et al., Prot. Eng., 10, 607 (1997)].

Novel Analogs of Single-Domain Heavy-Chain Camelid and Shark Antibodies:

[0125] FIGS. **2** and **3** outlines the analogs of new generation of camelid and shark antibodies and their analogs, respectively, which will assist us to develop ultrasensitive and

ultraspecific diagnostic assays for the detection/identification of the pathological proteins and antigens.

Production of Parent Single-Domain Heavy-Chain Mini-Antibodies (sdmnAbs) of Structure **1**

[0126] Host animals such as camel, llama, or alpaca will be immunized with the desired antigen(s), for example HER-2 protein, a biomarker for breast cancer or Aβ42 antigenic peptide for detecting amyloid plaque, following the procedures described by Murphy et al, in 1989 [Am. J. Vet. Res., 50, 1279 (1989)], but with slight modification. Immunization of camels will be done with 250 ug antigenic peptide per injection will be used, followed by 4 booster shots every two weeks 4 weeks after the initial injection. For baby sharks, 10 ug antigen/injection will be used. One antigen per animal for immunization will be used, though it may be feasible to immunize an animal simultaneously with multiple antigens to raise an immune response to each antigen separately, which can make the production cost effective [EMBO, J., 17, 3512 (1998); J. Immunol. Methods, 240,185 (2000)].

[0127] After immunization, 100 ml camel blood (or 5 ml from shark) will be withdrawn from the animal and the total IgGs will be precipitated out using ammonium sulfate precipitation procedure. Using size exclusion chromatography over Sephadex G-25, the conventional IgGs, MW 150 KDa, will be removed from the single-domain mini-IgG, 1, with MW of 90 to 100 K Da. Affinity purification to obtain high affinity sdmnAb, I, will be done by magnetic beads coated with the antigenic peptide.

Recombinant Production of Camelid and Shark Single-Domain Antibodies:

[0128] Recombinant production of single-domain heavy-chain parent camelid antibodies **1**, **4**, **5**, **6**, **7**, **8**, **9** (FIG. **2**) and shark antibodies **2**, **11**, **12**, **13**, **14**, and **15** will be done according to protocols and procedures described in pending U.S. patent application Ser. No. 12/563,330.

Chemical Synthesis of Analogs of Single-Domain Camel Antibodies

Derivatization and Immobilization of Camelid Mini-Antibody **1**:

[0129] Schematics for derivatization and immobilization of mini-antibodies **1** are shown in FIG. **4**. Mini-antibody **1** (1 mg, 11 nmols) will be treated with commercial NHS-(PEG)$_n$-Mal (11 ul of 10 mM stock=110 nanomoles) wherein n=1-50, in 50 mM MOPS/150 mM NaCl, pH 6.8, at RT for 1 hour to obtain the pegylated conjugate of FIG. **4** (structure not shown) which will be desalted by dialysis on C-3 Amicon filters to remove excess NHS-PEG-Mal reagent.

[0130] While the pegylation is underway, the ligand will be treated with 10× folds of Traut's Reagent in MOPS buffer, pH 6.8, containing 5% EDTA, at RT for 1-2 hours. The thiolated ligand will then be purified either by dialysis (if ligands is chemical or biochemical entity) or by washing with MOPS buffer if ligand is a solid matrix.

[0131] The pegylated intermediate will be immediately conjugated with 10-20 folds excess of thiolated ligand: "SH-L" in MOPS buffer, pH 6.8 buffer containing 5 mM EDTA for 2-3 hours at room temperature (RT); where "L" may be enzyme (HRP, AP, Luciferase, galactosidase), protein, peptide, biotin, fluorophore, DNA, RNA, and solid matrix such as, magnetic beads, glass slides, gold nanoparticles, microchannels, microfluidic device.

[0132] When the ligand is a chemical or biochemical entity, for example, fluorophore, biotin, enzyme, protein, etc, the purification of the conjugate will be done by reverse-phase C8 HPLC.

[0133] Nucleic acid conjugates of mini- and nano-antibodies **1**-**15***a* will also be prepared using their pegylated conjugates followed by treatment with the thiolated-DNA/RNA molecules of interest (FIG. **4**).

[0134] When the ligand is a solid matrix, such as, magnetic beads, glass slide, microchannels, etc., which we will use to immobilize the camelid antibodies, all we need to do is to wash the excess reagent with the appropriate buffer.

[0135] The activity and the amount of camelid antibody loaded onto the solid matrix will be determined by ELISA and commercially available protein assay kits.

Single-Domain Heavy-Chain Camelid Micro-Antibody **4** and its Analogs **4***a:*

[0136] Micro-antibody, **4**, will be prepared by treating mini-antibody **1** (2 mg) with 1.0 ml of 10 mM TCEP (tris-carboxyethyl-phosphine) in 20 mM Phosphate/150 mM NaCl, pH7.4 at room temperature (RT) for one hour. The resulting micro-antibody **4** will be desalted on centricon-3 to remove the excess reagent and the buffer and stored at 4° C. in 1×PBS.

[0137] Derivatization of **4** into **4***a* will be accomplished by the method described above for conversion of **1** into **1***a*.

Single-Domain Heavy-Chain Camelid Sub-Nano-Antibody **5** and its Analogs of Structure **5***a:*

[0138] Micro-antibody **4** will be treated with trypsin or pepsin under controlled conditions at RT to cleave the CH2-CH3 domains from the antibody. After deactivation of the proteolytic enzyme with fetal calf serum, the subnano-antibody **5** will be isolated using size exclusion chromatography.

[0139] Derivatization of **5** into **5***a* will be accomplished by the method described above for conversion of **1** into **1***a*.

Single-Domain Heavy-Chain Camelid Nano-Antibody **6** and its Analogs of Structure **6***a:*

[0140] Sub-nano-antibody **5** will be treated with pepsin at a low pH of 4.5 in 2M sodium acetate buffer under mild conditions for 1-8 hours to cleave the CH2-CH3 domains from the antibody. After deactivation of the proteolytic enzyme with fetal calf serum, the nano-antibody **6** will be isolated using size exclusion chromatography.

[0141] Derivatization of **6** into **6***a* will be accomplished by the method described above for conversion of **1** into **1***a*.

Single-Domain Heavy-Chains Bivalent Nano-Antibody **7** and its Analogs of Structure **7***a:*

[0142] Schematics for the chemical synthesis of dimeric nano-antibodies and heir analogs are displayed in FIG. **5**. Nano-antibody **5** will first be oxidized with 1% iodine in 20% tetrahydrofuran/70% water/10% pyridine for 5-10 minutes to transform into the dimeric nano-antibody **7**. Treatment of **7** will be done with commercial NHS-(PEG)$_n$-Mal, wherein n=1-50, in 50 mM MOPS/150 mM NaCl, pH 6.8, at RT for 1 hour to obtain the pegylated intermediate with maleimido group (structure not shown in FIG. **5**) which, after purification by dialysis on C-3 Amicon filters, will be immediately conjugated with thiolated-ligand in MOP buffer containing 5% EDTA at pH6.8 for 2-3 hours at RT to obtain, after purifica-

tion, **7***a*. The dimeric conjugate **7***a* will then be characterized by ELISA and Western blot assays.

Trivalent and Tetravalent Camelid Nano-Antibodies and Analogs:

[0143] Schematics for the chemical synthesis of trivalent and tetravalent camelid nano-antibodies without the light-chains, and their analogs are shown in FIG. **6**.

Protocol for Developing Trivalent **8** and Tetravalent **9** Camelid Nano-Antibodies:

[0144] Bivalent nano-antibody, **7**, prepared by oxidative dimerization or chemical ligation, will be conjugated with NHS-(PEG)3-Mal (10 folds excess) in MOPS buffer at pH 7.0 for 1 hour at RT. Chemical ligation of the resulting monomeric and dimeric pegylated products **16** and **17** with the thiolated nano-antibody **18** (FIG. **6**) will be carried out by combining the two at pH 6.8 buffer containing 5 mM EDTA and allowing the reaction to occur at RT for at least 2 hours. The so formed trivalent, **19**, and tetravalent nano-antibody **20** will be purified by size exclusion chromatography and stored at 4° C. in PBS containing 0.02% NaN3.

[0145] The attachment of a ligand to **19** and **20** can be readily done by making use of the lysine(s) of the hinge region to conjugate with the NHS-L.

[0146] Pentavalent and higher analogs of nano-antibodies (Vab domains of camel antibodies) can be similarly prepared.

Production of Single-Domain Heavy-Chain Shark IgNAR (Structure **2**):

[0147] Immunization of Sharks and Isolation of Shark IgNAR: Baby sharks will be immunized with the desired antigen(s), for example ALZAS, Tau, Aβ42 peptide which are the potential biomarkers for Alzheimer's disease, following the protocol described by Suran et al [J. Immunology, 99, 679 (1967)]. Briefly, the antigen (20 ug per kg animal weight), dissolved in 20 mg/ml keyhole limpet hemocyanin (KLH) supplemented with 4 mg/ml complete Freund's adjuvant, will be injected intramuscularly. Four booster shots every two weeks four weeks after the initial injection will be administered.

[0148] After immunization, 3-5 ml shark blood will be withdrawn from the animal and the total IgGs will be precipitated out using 50% ammonium sulfate, followed by centrifugation at 2000 RPM for 10 minutes. After discarding supernatant, the precipitate will be dissolved in 20 mM PBS/150 mM NaCl containing 0.02% sodium azide and size fractionated on Sephadex G200. The conventional IgGs, MW ~230 KDa, will be separated out from the shark IgNAR with MW of ~180 K Da. Alternatively, the conventional IgG fraction will first be depleted with protein G bound to magnetic beads, followed by isolation of V-NAR protein with magnetic beads coated with protein-A. Affinity purification to obtain high affinity shark Ig-NAR, **2**, will be done by magnetic beads coated with antigenic peptide.

[0149] After determining the amino acid sequence of IgNAR, **2**, nucleic acid sequence will be derived based from the amino acid sequence and recombinant DNA protocols will be established to produce the shark single-domain antibody **2** on a large scale. Schematics for cloning and expression of IgNAR are shown in FIG. **7**.

Isolation of RNA from Immunized Shark's Lymphocytes and Cloning:

[0150] Isolation of total RNA, **21**, from immunized sharks will be done from 3-5 ml of shark blood using commercially available RNA extraction kits such as Bio-Rad's AquaPure® RNA Isolation kit. Reverse transcription using oligo-dT primer will be achieved by PCR using high fidelity DNA polymerase to obtain the IgNAR cDNA, **22**, shown in FIG. **7**.

Recombinant Production of Shark Heavy Chain Only Antibodies and their Analogs

[0151] An exemplary cloning strategy is shown in FIG. **7**. Amplicons for IgNAR cDNA, **22** and its analogs will be performed using the following protocol:

IgNAR cDNA=1.0 ug

Primers Mix=10 μmol (forward and reverse primers)

1 mM dTNPs=10 ul

10 mM MgCl2=5 ul

10×PCR Buffer=5 ul

Taq DNA Polymerase=0.6 ul

[0152] Water to=50 ul

[0153] After first denaturation round of 94° C. for 10 minutes, 35 to 36 cycles of amplification will be performed under conditions as described below:

Denaturation: 20 seconds at 94° C.

Annealing: 30 Seconds at 56° C.

[0154] Extension: 50 seconds at 72° C.

Final Extension: 10 min, 72° C.

[0155] All or portions of IgNAR cDNA using different combinations of the following forward and reverse primers.

```
Forward primers
5'-gcatgggtag accaaacaccaag-3'    (SEQ ID NO: 1)

5'-gcgtcctcagagagagtcccta-3'      (SEQ ID NO: 2)

5'-gagacggacgaatcactgaccatc-3'    (SEQ ID NO: 3)

5'-gggtagaccaaacaccaagaacagc-3'   (SEQ ID NO: 4)

Reverse primers
5'-gttctagccaataggaacgtatag-3'    (SEQ ID NO: 5)

5'-gtttgcacaagagagtagtctttac-3'   (SEQ ID NO: 6)

5'-cctaaattgtcacagcgaatcatg-3'    (SEQ ID NO: 7)

5'-gtgcagttccctagaagtcttg-3'      (SEQ ID NO: 8)
```

[0156] After amplification, the amplicon will be purified on 1.5% agarose. The amplicon will be extracted from the gel and its 5'-end kinased with gamma-ATP for blunt-end ligation with the phage-display vector using T4 DNA-ligase following standard ligation protocols.

[0157] Library or Plasmid Construction: Prior to cloning, the PCR amplicon encoding IgNAR gene will be digested with Sfi1 and Not1 (Roche) following the cocktail:

V-NAR-CH1-CH2-CH3-CH4-CH5 DNA=5 ug

10× Restriction Buffer 5 ul

Sfi1 (10 U/ul) 8 ul

Water to 50 ul

[0158] Incubate 50° C. for 8 hour

Not1 35 U

Reaction Buffer 4.5

Water to 60 ul

[0159] Incubate at 37° C. for 4-5 hours.

Ethanol Precipitate at −70° C. Pellet

Water to 50 ul

[0160] Agarose gel (1.5%) purification

Pure DNA Encoding Shark IgNAR Antibody

Vector Ligation:

IgNAR DNA=200 ng

Vector DNA=1000 ng

10× Ligase Buffer=5 ul

T4 DNA Ligase=10 U

Water to 50 ul

[0161] Incubate 15 hours at 4° C.

Ethanol Precipitate at −70° C.

[0162] Suspend pellet in 10 ul.

Electroporation:

[0163] 250 ul of *E. Coli* TG1 cells will be made electro-competent with BRL Cell-Porator® following vendor protocol.

[0164] Panning of Phage-Displayed IgNAR-Antibody **2** Library: Electroporated TG1 cells will be transfected with the phagemid-IgNAR DNA insert. Approximately, 1010 cells will be grown to mid-logarithmic phase before injection with M13K07 helper phages. Virions will be prepared as described in the literature [Andris-Widhopf J., et al, J. Immunology Methods, 242, 159 (2002)] and used for panning at a titer of 1013/ml. Specific IgNAR antibody against the antigenic peptide will be enriched by five consecutive rounds of panning using magnetic beads conjugated with antigenic peptide. Bound phage particles will be eluted with 100 mM TEA (pH 10.00), and immediately neutralized with 1M Tris.HCl (pH 7.2) and will be used to reinfect exponentially growing *E. Coli* TG1 cells.

[0165] The enrichment of phage particles carrying antigen-specific IgNAR antibody will be assessed by ELISA before and after five rounds of panning. After the fifth panning, individual colonies will be picked up to analyze the presence of the virion binding by anti-M13-HRP conjugate.

Expression and Purification of the Single Domain IgNAR **2**:

[0166] The selected positive clones will be used to infect a new bacterial strain, HB 2151, a non-suppressor strain that recognizes the amber codon as a stop codon for soluble protein production. The HB2151 cell harboring the recombinant phagemids will be grown at 28° C. in 250 ml 2×YT-ampicillin, 1% glucose in culture flasks until OD600 0.7. The cells will be washed and resuspended in 250 ml 2×YT-ampicillin, supplemented with 1 mM isopropyl beta D-thiogalactopyranoside (IPTG), and incubated over night at 22° C. to induce protein expression.

[0167] Before adding IPTG to the cultures, a portion will be spotted on an LB/ampicillin plate for future analysis of the clones. The culture will be then be centrifuged at 4000 RPM for 15 minutes to pellet the bacterial cells. The culture supernatant will then be screened by ELISA for antigen-specific IgNAR protein **2**.

Chemical Synthesis of Single-Domain Heavy-Chain Shark Only Antibodies and Their Analogs

[0168] **2***a*, **11**, **11***a*, **12**, **12***a*, **13**, **13***a*, **14**, **14***a*, **15**, **15***a*:

Derivatization of Shark IgNAR **2** to Obtain Analogs of Structure **2***a*:

[0169] Schematics for derivatization of shark IgNAR **2** are shown in FIG. **8**. Shark antibody **2** (1 mg, 6 nmols) will be treated with commercial NHS-$(PEG)_n$-Mal (6 ul of 10 mM stock=60 nanomoles) wherein n=1-50, in 50 mM MOPS/150 mM NaCl, pH 6.8, at RT for 1 hour to obtain the pegylated conjugate (structure not shown) which will be desalted by dialysis on C-3 Amicon filters to remove excess NHS-PEG-Mal reagent.

[0170] While the pegylation is underway, the ligand will be treated with 10× folds of Traut's Reagent in MOPS buffer, pH 6.8, containing 5% EDTA, at RT for 1-2 hours. The thiolated ligand will then be purified either by dialysis (if ligands is a chemical or biochemical entity) or by washing with MOPS buffer if ligand is a solid matrix.

[0171] The pegylated intermediate will be immediately conjugated with 4-5 folds excess of thiolated ligand: "SH-L" in MOPS buffer, pH 6.8 buffer containing 5 mM EDTA for 2-3 hours at room temperature (RT); where "L" may be enzyme (HRP, AP, Luciferase, galactosidase), protein, peptide, biotin, fluorophore, DNA, RNA, and solid matrix such as, magnetic beads, glass slides, gold nanoparticles, microchannels, microfluidic device.

[0172] When the ligand is a chemical or biochemical entity, for example, fluorophore, biotin, enzyme, protein, etc, the purification of the conjugate **2***a* will be done by reverse-phase C8 HPLC.

[0173] Nucleic acid conjugates of shark IgNAR **2** and analogs **11**,**12**,**13**,**14**, and **15** will also be prepared using their pegylated conjugates followed by treatment with the thiolated-DNA/RNA molecules of interest.

[0174] When the ligand is a solid matrix, such as, magnetic beads, glass slide, microchannels, etc., which we will use to immobilize the camelid antibodies, all we need to do is to wash the excess reagent with the appropriate buffer.

[0175] The activity and the amount of single-domain shark antibody loaded onto the solid matrix will be determined by ELISA and commercially available protein assay kits.

Single-domain Heavy-Chain Shark Mini-Antibody **11** and its Analogs **11***a*:

[0176] Mini-antibody, **11**, will be prepared by treating the IgNAR **2** (2 mg) with 1.0 ml of 10 mM TCEP (tris-carboxyethyl-phosphine) in 20 mM Phosphate/150 mM NaCl, pH7.4 at room temperature (RT) for one hour. The resulting micro-antibody **11** will be desalted on centricon-3 to remove the excess reagent and the buffer and stored at 4° C. in 1×PBS.

[0177] Derivatization of **11** into **11***a* will be accomplished by the method described above for conversion of **2** into **2***a*.

Single-Domain Heavy-Chain Shark Micro-antibody **12** and its Analogs of Structure **12***a*:

[0178] Mini-antibody **11** will be treated with trypsin or pepsin under controlled conditions at RT to cleave the CH3-CH4-CH5 domains from the antibody. After deactivation of the proteolytic enzyme with fetal calf serum, the shark micro-antibody **12** will be isolated using size exclusion chromatography.

[0179] Derivatization of **12** into **12***a* will be accomplished by the method described above for conversion of **2** into **2***a*.

Single-Domain Heavy-Chain Shark Sub-nano-antibody **13** and its Analogs of Structure **13***a*:

[0180] Micro-antibody **12** will be treated with trypsin or pepsin under controlled conditions at RT to cleave the CH2 domain from the antibody. After deactivation of the proteolytic enzyme with fetal calf serum, the shark sub-nano-antibody **13** will be isolated using size exclusion chromatography.

[0181] Derivatization of **13** into **13***a* will be accomplished by the method described above for conversion of **2** into **2***a*.

Single-Domain Heavy-Chain Shark Dimeric Nano-Antibody **14** and its Analogs of Structure **14***a*:

[0182] Dimeric V-NAR will be prepared by the oxidation of monomeric V-NAR, **13** to obtain **14** as described in FIG. **9**. These protocols are general and do not need detailed explanation.

[0183] Likewise, the transformation of **14** into its analogs of structure **14***a* will be accomplished as described above for the preparation of **2***a* from **2**.

Tetrameric and Trimeric V-NAR Nano-Antibodies **31** and **32**:

[0184] Dimeric V-NAR nano-antibody **14** will be treated with 4-5 molar equivalent of NHS-PEG-Mal to obtain a mixture of tri- and tetra-pegylated derivatives of dimeric V-NAR nano-antibody (FIG. **9**).

[0185] After purification, the tri- and tetrameric pegylated products will be treated with thiolated V-NAR to obtain, after purification by HPLC or just by dialysis, tetrameric and trimeric V-NAR nano-antibodies **31** and **32**.

Immobilization of Single-Domain Camelid Mini-Antibody and Shark IgNAR Antibody and Analogs onto Solid Matrixes:

[0186] Immobilization of single-domain heavy-chain only shark and camelid native antibodies and their analogs onto solid matrixes, such as gold particles, magnetic particles, microchannels, glass particles and other solid surfaces will be accomplished using the steps outlined in FIG. **10**. Aminated solid matrix **33** will first be derivatized with NHS-(PEG)n-Mal, 10, where n=20 (20 fold molar excess) at pH 7.0 for 1 hour at RT. The solid matrix will then be washed thoroughly with the same buffer (50 mM MOPS/150 mM NaCl, pH 7.0). Any unconjugated amine groups will be masked with sulfoNHS-Acetate (Pierce) by incubated the solid matrix with 40 fold excess of the reagent at pH 7.0 for 60 minutes. After washing off the excess masking reagent, the pegylated matrix **34** will then be conjugated with thiolated single-domain heavy-chain antibody, **36**, (10× excess) over the starting amine concentration. The conjugation will be performed at pH 6.5 for 2 hours at RT with gentle shaking of the matrix. The unused antibody will be recovered, and the matrix very well washed with 1×PBS/0.5% Tween-20 to obtain complex **37** in which nano-antibody is covalently bound to a solid matrix. The activity of the bound heavy-chain antibody will be measured using ELISA.

Biomarkers for Various Diseases

[0187] Single-domain heavy-chain only shark and camelid native antibodies and their analogs can be used to detect the presence or absence of one or more antigens or can be used for diagnosis of one or more diseases. Single-domain heavy-chain only shark and camelid native antibodies and their analogs can bind specifically to the antigens or one or more biomarkers for various diseases. Exemplary sequences of various biomarkers or antigens are disclosed in U.S. application Ser. No. 12/563,330 filed Sep. 21, 2009. Those sequences are incorporated by reference to its entirety. Exemplary sequences of nucleic acids encoding additional biomarkers for Alzheimer's Disease are disclosed in FIG. **17-27**.

Example 1

Capture and Detection of Pathogenic Antigens/Proteins Using Shark and Camel Single-Domain Antibodies (sdAbs)

[0188] Serum from patient blood (10 ml), collected in EDTA tubes will be treated with shark and camelid heavy chain only antibodies and their analogs coated magnetic beads for 1-2 hours on a rotator with gentle rotation to bind the antigen. The beads will be separated using a magnetic rack and subsequently washed very well with PBS/1% BSA. The antigen-microantibody complex so formed will be treated with complex, detection antibody bound to an enzyme (AP, HRP, Luciferase, beta-galactosidase, gold particles) or DNA to sandwich the antigen between the shark and camelid heavy chain only antibodies and their analogs and the detection antibody forming the complex which will be detected either using an enzyme substrate or AgNO3 if the detection antibody is conjugated to gold particles. Exemplary schematics of the process is shown in FIG. **25**. Alternatively, the detection antibody could be conjugated to DNA molecules which can then be amplified by PCR to obtain detection sensitivity equivalent to the detection of DNA by PCR as shown in FIG. **26**.

Example 2

Non-Invasive Detection of Prenatal Genetic Disorders from Captured Circulating Fetal Cells (CFCs) Using Heavy-Chain Antibodies

[0189] Blood (10 ml) from a pregnant woman will be treated at RT for 1 hour with sdAb conjugated to magnetic beads, with gentle shaking. The beads will be allowed to settle down in a magnetic rack and then subsequently washed with a wash buffer containing 20 mM PO4-2/150 mM NaCl/0.1% Triton X-100 (3×2 ml) to ensure complete removal of blood and serum. The beads will then be washed with 1×PBS to remove triton. The bound DNA will then be eluted by hot 10 mM Tris.HCl, pH7.0 or by protease digestion.

[0190] This fetal DNA will then be analyzed by real-time PCR using Y-chromosome primers to test the gender and by chromosome 21 primers to test for Down syndrome.

Example 3

In-Vitro Capture of Pathological Proteins with Single-Domain Camelid and/or Shark Antibodies and Detection by Enzymatic Signal Amplification

[0191] The high specificity of camelid and shark antibodies can be exploited to detect proteins at a much lower concentrations than what is currently possible. These antibodies are stable and functional at higher temperatures (80 to 90° C.). Also, they are stable in the presence detergents and denaturing agents. This allows us to capture the pathological antigens under stringent conditions such as performing the capture reaction at elevated temperatures and using detergents (say for an example the use of up to 10% TritonX-100-), followed by high temperature stringent washings containing detergents to minimize non-specific capture. Such use of stringent conditions is only possible in immunoassays utilizing camelid and shark antibodies. When combined with enzymatic signal as shown in FIG. **11**, the camelid and shark antibodies should be able to detect 0.1 to 1.0 attomoles of target molecules in 25 ul reaction volume, which itself is a much improvement over the existing proteomic detection technologies.

[0192] In the representative example shown in FIG. **11**, the patient serum was incubated with magnetic beads coated with camel micro-antibody **39** (0.5 ml beads containing at least 1.0 ug camelid micro-antibody) with gentle shaking of the reaction contents at RT for 45 minutes. After 45 minutes, the beads were allowed to settle down and washed with 2×SSC buffer containing 1.0% Tween-20. Detection of beads bound pathological antigen from 40 was accomplished by incubating the beads with a AP conjugate **41** of detection antibody for 1 hour at RT on a rocker. The beads were then thoroughly washed (5×2 ml) with preheated 2×SSC buffer (60° C.) containing 0.5% NP-40 to remove any non-specifically bound complex **41** (other commercially available detergents such as Triton X-100, Tween-20, SDS, LiDS, IGEPAL, Luviquat, DTPO, Antifoam 204, etc. can also be used.). The washings of the beads can also be done at temperature above 60° C. all the way up to 85° C. to remove any contaminants from the complex **42**. The detection of complex **42** was then accomplished with Attophase (100 ul of 1.0 micromolar solution, 37° C. for 30 minutes), a fluorescence substrate for AP. The liberated green fluorescence was measured using a 96 microwell plate fluorimeter. 0.1 attomole of serum PSA antigen could be readily detected with 3:1 signal to background ratio.

Example 4

In-Vitro Capture of Pathogens by Single-Domain Shark IgNAR in Solution Phase and Detection by Enzymatic Signal Amplification

[0193] In this technology format, the biotinylated shark IgNAR, **2***a* (FIG. **12**), can be added to the patient serum and allowed to react with the pathogen at 37° C. for about one hour while the reactants are gently stirred or rotated on a orbital shaker. Anti-biotin camelid antibody (or shark antibody) bound to magnetic beads **45** can be added to reaction mixture to capture the so formed shark-IgNAR-Antigen complex **44** forming a complex of structure **46**. The magnetic beads will be allowed to settle down in a magnetic rack and washed very well with a preheated (60° C.) wash buffer containing at least 1% NP-40. The complex **46** can then be detected by incubating with AP-IgG (sec) camelid complex using Attophos as a substrate as described above.

[0194] Other camelid and/or shark antibodies and their analogs can also be used the same way.

Example 5

Ultra-Sensitive Signal Amplification Using Single-Domain Heavy-Chain Only Camelid and Shark Antibodies for In-Vitro Detection of Pathogens by Immuno-PCR

[0195] The two vital components of the method of this invention are: #1) Ultra-specific capture of pathological proteins by the new generation of single-domain camelid and shark antibodies lacking the light-chains (heavy-chain only antibodies), and #2) an ultra-sensitive signal amplification technology to detect fewer than 200 molecules of protein biomarkers to diagnose diseases at a very early stage of the their manifestation. Therefore, in its preferred embodiment, this invention incorporates inherently highly specific heavy-chain antibodies for capturing the pathological proteins/antigens with high specificity with low to zero cross-reactivity, followed by detection of the captured antigen by enzymatic signal amplification, preferably immuno-PCR, to develop an ultra-sensitive, highly specific and reliable diagnostic assay to detect fewer than 200 copies of the pathological proteins from biological samples.

[0196] FIG. **13** outlines the steps of the process involved. The protocol involves capturing the antigen from bodily fluid utilizing camelid and/or shark antibody coated magnetic beads by bringing in contact the said sample containing antigen with the beads. In the example shown in FIG. **13**, camelid mini-antibody coated magnetic beads **48** are mixed with the serum for 1-2 hours with reactants constantly but slowly mixing all the time. The beads are then allowed to settle down in a magnetic rack and very well washed to ensure complete removal of the serum.

[0197] The detection of the captured antigen will be done by adding conjugate, **49**, of secondary antibody that is conjugated to 100-120 bases long DNA via a hydrophilic linker that is at least 5 nanometer long to diminish and/or remove any stearic affects in the subsequent enzymatic amplification. The reaction between the captured antigen and the conjugate **49** will be allowed to take place for 2-3 hours after which the beads will be thoroughly washed to remove any unreacted conjugate **49**.

[0198] The subsequent amplification of the attached DNA molecule by PCR using PCR kit from Applied Biosystems will allow for the indirect detection of the antigen with sensitivity almost equal to the sensitivity of detection of DNA by PCR.

[0199] There are many possible permutations and combinations of this technology. For instance, the antigen can be detected in solution phase by biotinylated or digoxigenin labeled camelid or shark antibody as described above following the steps of figure outlined in FIG. **12**. The antigen-antibody complex so formed can be immobilized onto some solid matrix using camelid or shark anti-biotin or anti-digoxigenin antibody. Detection can be done by Immuno-PCR by forming a complex of the immobilized antigen-camelid-antibody with the secondary antibody bound to DNA which can be amplified by PCR.

[0200] Alternatively, the detection antibody in FIG. **13** can be conjugated with camelid or shark anti-biotin Mini- or nano-antibody. Biotinylated DNA can be used as a detection agent which will be amplified by PCR as outlined in FIG. **13**.

Example 6

In-Vitro Capture and Detection of Rare Cells Using Shark and Camelid Heavy Chain Only

Antibodies and Their Analogs

[0201] Fresh 5 ml patient blood will be diluted with 20 ml 1×PBS/1% BSA to 25 ml. To capture circulating tumor cells (CTCs), this sample will then be passed through a microfluidic device coated with an appropriate shark and camelid heavy chain only antibodies and their analogs, such as, anti-EpCAM-micro-antibody (camelid) **51** following flow rate recommended by the manufacturer of microfluidic device. To ensure that antibodies or its analogs do not lose any activity upon conjugation, all solid matrixes will first be coated with a hydrophilic polymer, such as, NHS-PEG-Mal (MW ~5000). The conjugation of the thiolated shark and camelid heavy chain only antibodies and their analogs with maleimido-group of the polymer can be achieved at pH 6.8 in a buffer containing 5% EDTA. Exemplary schematics of the process are shown in FIG. **14**.

[0202] Alternatively, magnetic beads coated with EpCAM can be used. EpCAM (epithelial cell adhesion molecules) is frequently over expressed by carcinomas of lung, colorectal, breast, prostate, head and neck, liver, and is absent from hematological cells. The captured cells can be washed with 1% PBS (no BSA). The cell can be fixed with methanol, and then DAPI stained following CK8 or CK18 and CD45. Identification and enumeration will be done by fluorescence microscopy based upon the morphological characteristics, cell size, shape, and nuclear size. DAPI+, CK+, and CD45- cells will be classified as CTCs.

Alternative Strategies to Capture Circulating Tumor Cells (CTCs):

[0203] Patient's blood (2-3 ml) (or urine 15-20 ml after centrifugation to pellet down the cells and suspending them in 1-2 ml HBSS media) will be incubated with an appropriate biotinylated mini-sdAb (1.5 ug/ml blood sample) at RT for one hour. For example, to capture epithelial cancer cells, such as from breast, prostate, and ovarian cancers, biotinylated-anti-EpCAM-mini-antibody (camel antibody against EpCAM antigens) will be used to label the circulating cancer cells in the blood. After diluting with HBSS or RPMI-1640 media or 1×PBS/2.5% BSA to lower the sample viscosity, the diluted blood is then passed through a microfluidic device coated with antibiotin-mini-camelid or shark antibody at a flow rate allowing maximum cell capture. The captured CTCs can then be fixed by fixing with methanol, followed by fixing with 1% PFA using any standard cell fixing procedures. Enumeration will then be done by DAPI staining followed by immunohistochemical staining with commonly used mouse mHCAb such as CK-7 but more preferably mini-CK-7 for higher specificity. CTCs have to be CD45 negative.

[0204] Alternatively, most of the RBCs from the blood sample can be first lysed using ammonium chloride solution (155 mM NH4Cl/10 mM NaHCO3). After pelleting, the washed cells will be suspended in HBSS media (1-2 ml) and passed through the microfluidic device coated with heavy-chain antibody specific for the cell type one needs to capture and analyze.

[0205] Alternatively, the diluted blood sample after incubation with the biotinylated-antiEpCAM-mini-antibody or micro-antibody will be treated with the anti-biotin-mini-camelid antibody coated magnetic particles (Miltenyl) for 30 minutes while the sample is being gently rotated on a rotating wheel. After pulling down the magnetic particles with a magnet, the CTCs bound to the particles will be washed with PBS/1% BSA. The CTCs can then be enumerated by spreading them in a unilayer on a glass slide, drying them for one to two hours, followed by fixing with methanol, 1% PFA and staining the CTCs with CK-7.

[0206] Furthermore, these captured CTCs can be analyzed for the gene expression. For example, in case of prostate cancer patient, one can look for TMPRSS2-ERG translocation using PCR primers. TMPRSS2-ERG transcript is present in about 50% of the prostate cancer patients. Similarly, one can look for HER-2 expression in case of breast cancer.

Example 7

Capture and Analysis of Fetal Cells

[0207] To capture fetal cells from the blood of pregnant mothers, 5 ml blood from pregnant mothers can be diluted to 15 ml with HAM-F 12 media containing 1% BSA and passed through the microfluidic device coated with the camelid and/or shark antibodies against the fetal cell surface antigens, CD71, glycophorin-A (GPA), CD133, and CD34. The captured fetal cells be analyzed for fetal gender, and genetic abnormalities using either PCR but preferably FISH probes for chromosomes X, Y, 13, 18 and 21 as shown in FIG. **15**. For FISH analysis, the captured cells will be fixed with methanol followed by fixation with 1% PFA. After staining with epsilon-hemoglobulin, the cells be hybridized with Vysis FISH Fetal male gender can be readily detected by the appearance of XY fluorescence signal under the fluorescence microscope. Cells stained with epsilon-hemoglobulin showing XX signal will be identified as female fetal cells. Trisomies can be readily identified also based upon whether two or three chromosomes are giving the fluorescence signals.

[0208] Alternatively, most of the RBCs can either be carefully lysed using a mild treatment with ammonium chloride lysis reagent (155 mM NH4Cl/10 mM NaHCO3) to enrich for fetal nucleated red blood cells (fnRBCs) before incubating the sample with a mixture of biotinylated antibodies.

[0209] Still another option will be the use of a density gradient such as Ficol 1.073 or Percol 1.073. The buffy coat can then be processed as above to yield fetal nRBCs.

Example 8

Detection of Chromosomal Translocations from Captured Circulating Tumor Cells (CTCs) Using Shark and Camelid Heavy Chain Only Antibodies and their Analogs

[0210] Cells will be captured as described above and also shown in FIG. **16**. Enumeration of can be done using an appropriate FISH probes. For example, to test for the presence of TMPRESS2/ERG translocation in case of prostate cancer, FISH probes designed to hybridize with the junction

region will be used. Similarly, in case of CML, bcr-Abl FISH probe will be used. An exemplary schematics of the process is shown in FIG. **16**.

[0211] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. All nucleotide sequences provided herein are presented in the 5' to 3' direction.

[0212] The inventions illustratively described herein may suitably be practiced in the absence of any element or elements, limitation or limitations, not specifically disclosed herein. Thus, for example, the terms "comprising", "including," containing", etc. shall be read expansively and without limitation. Additionally, the terms and expressions employed herein have been used as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed.

[0213] Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification, improvement and variation of the inventions embodied therein herein disclosed may be resorted to by those skilled in the art, and that such modifications, improvements and variations are considered to be within the scope of this invention. The materials, methods, and examples provided here are representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the invention.

[0214] The invention has been described broadly and generically herein. Each of the narrower species and subgeneric groupings falling within the generic disclosure also form part of the invention. This includes the generic description of the invention with a proviso or negative limitation removing any subject matter from the genus, regardless of whether or not the excised material is specifically recited herein.

[0215] In addition, where features or aspects of the invention are described in terms of Markush groups, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group.

[0216] All publications, patent applications, patents, and other references mentioned herein are expressly incorporated by reference in their entirety, to the same extent as if each were incorporated by reference individually. In case of conflict, the present specification, including definitions, will control.

[0217] Other embodiments are set forth within the following claims.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 19

<210> SEQ ID NO 1
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 1

gcatgggtag accaaacacc aag 23

<210> SEQ ID NO 2
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 2

gcgtcctcag agagagtccc ta 22

<210> SEQ ID NO 3
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 3

gagacggacg aatcactgac catc 24

-continued

<210> SEQ ID NO 4
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 4

gggtagacca aacaccaaga acagc 25

<210> SEQ ID NO 5
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 5

gttctagcca ataggaacgt atag 24

<210> SEQ ID NO 6
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 6

gtttgcacaa gagagtagtc tttac 25

<210> SEQ ID NO 7
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 7

cctaaattgt cacagcgaat catg 24

<210> SEQ ID NO 8
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic primer

<400> SEQUENCE: 8

gtgcagttcc ctagaagtct tg 22

<210> SEQ ID NO 9
<211> LENGTH: 1851
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

ggccggtcag cgtcgctgcc ggtctccggc ggagacggac tctggagttt gggcggcccg 60

ggcggccact aggtactctg atattccgta ctaaacacgt ctgcaagtca agatgtcgca 120

-continued

cccgtccccc caagccaagc cctccaaccc cagtaaccct cgagtcttct ttgacgtgga 180

catcggaggg gagcgagttg gtcgaattgt cttagaattg tttgcagata tcgtacccaa 240

aactgcggaa aattttcgtg cactgtgtac aggagaaaaa ggcattggac acacgactgg 300

gaaacctctc catttcaaag gatgcccttt tcatcgaatt attaagaaat ttatgattca 360

gggtggagac ttctcaaatc agaatgggac aggtggagaa agtatttatg gtgaaaaatt 420

tgaagatgaa aatttccatt acaagcatga tcgggagggt ttactgagca tggcaaatgc 480

aggccgcaac acaaacggtt ctcagttttt tatcacaaca gttccaactc ctcatttgga 540

tgggaaacat gtggtgtttg gccaagtaat taaaggaata ggagtggcaa ggatattgga 600

aaatgtggaa gtgaaaggtg aaaaacctgc taaattgtgc gttattgcag aatgtggaga 660

attgaaggaa ggagatgacg ggggaatatt cccaaaagat ggctctggcg acagtcatcc 720

agatttccct gaggatgcgg atatagattt aaaagatgta gataaaattt tattaataac 780

agaagactta aaaaacattg gaaatacttt tttcaaatcc cagaactggg agatggctat 840

taaaaaatat gcagaagttt taagatacgt ggacagttca aaggctgtta ttgagacagc 900

agatagagcc aagctgcaac ctatagcttt aagctgtgta ctgaatattg gtgcttgtaa 960

actgaagatg tcaaattggc agggagcaat tgacagttgt ttagaggctc ttgaactaga 1020

cccatcaaat accaaagcat tgtaccgcag agctcaagga tggcaaggat taaaagaata 1080

tgatcaagca ttggctgatc ttaagaaagc tcaggggata gcaccagaag ataaagctat 1140

ccaggcagaa ttgctgaaag tcaaacaaaa gataaaggca cagaaagata aagagaaggc 1200

agtatatgca aaaatgtttg cttagaaagg attcagtttt gcttattgtg tgttgattgt 1260

ataaatgcaa taagaaaatg taaaggtttt tgtctatgaa tatgatccct aatgtgtttc 1320

ttttgacacc ttagttcctt actgtttaca gtttaggagt actgataggg gttcatgctt 1380

aataaacatg tcacaatata gtaagtaaag tggttttgtt tgtttctttg agatggagtc 1440

ttgctctgtc acccaggctg gagtgcggtg gcgcaatctc ggctcactgc atcctctgcc 1500

tcccgggttc aagcaattct cctgcctcag cttcccaagt agctgggatt acaggcacgt 1560

gccaccacgc ccagctaatt tttgtatttt tagtagagat ggggtttcac catattggtc 1620

acgtcacgtt ggtcttgaac tcctgacctt gtgatccacc ccgccttggc ctcccaaagt 1680

gctgggatta caggtgtgag ccaccgtgcc cggccaagta aaatgttttt taaaatggtt 1740

atgtgcatta ttcataaaaa ataatggtgt ccagtctttt taaacttgta aagacacatc 1800

ttattgaata aagagatgag agcttaagtt tgtaaaaaaa aaaaaaaaaa a 1851

<210> SEQ ID NO 10
<211> LENGTH: 963
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown: Alpha beta binding
Mitochondrial Alcohol Dehyrdogenase polynucleotide

<400> SEQUENCE: 10

atccccatcc cgtggagtgg ccggcgacaa gatggcagca gcgtgtcgga gcgtgaaggg 60

cctggtggcg gtaataaccg gaggagcctc gggcctgggc ctggccacgg cggagcgact 120

tgtggggcag ggagcctctg ctgtgcttct ggacctgccc aactcgggtg gggaggccca 180

-continued

agccaagaag ttaggaaaca actgcgtttt cgccccagcc gacgtgacct ctgagaagga 240

tgtgcaaaca gctctggctc tagcaaaagg aaagtttggc cgtgtggatg tagctgtcaa 300

ctgtgcaggc atcgcggtgg ctagcaagac gtacaactta aagaagggcc agacccatac 360

cttggaagac ttccagcgag ttcttgatgt gaatctcatg ggcaccttca atgtgatccg 420

cctggtggct ggtgagatgg gccagaatga accagaccag ggaggccaac gtggggtcat 480

catcaacact gccagtgtgg ctgccttcga gggtcaggtt ggacaagctg catactctgc 540

ttccaagggg ggaatagtgg gcatgacact gcccattgct cgggatctgg ctcccatagg 600

tatccgggtg atgaccattg ccccaggtct gtttggcacc ccactgctga ccagcctccc 660

agagaaagtg tgcaacttct tggccagcca agtgcccttc cctagccgac tgggtgaccc 720

tgctgagtat gctcacctcg tacaggccat catcgagaac ccattcctca atggagaggt 780

catccggctg gatggggcca ttcgtatgca gccttgaagg gagaaggcag agaaaacaca 840

cgctcctctg cccttccttt ccctggggta ctactctcca gcttgggagg aagcccagta 900

gccatttgt aactgcctac cagtcgccct ctgtgcctaa taaagtctct ttttctcaca 960

gag 963

<210> SEQ ID NO 11
<211> LENGTH: 1396
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown: Translocase of the
Outer Membrane polynucleotide

<400> SEQUENCE: 11

tcctttccgc ttccggtgtc ccctacagtc atggctgccg ccgtcgctgc tgccggtgca 60

ggggaacccc agtccccgga cgaattgctc ccgaaaggcg acgcggagaa gcctgaggag 120

gagctggagg aggacgacga tgaggagcta gatgagaccc tgtcggagag actatggggc 180

ctgacggaga tgtttccgga gagggtccgg tccgcggccg gagccacttt tgatctttcc 240

ctctttgtgg ctcagaaaat gtacaggttt tccagggcag ccttgtggat tgggaccact 300

tcctttatga tcctggttct tcccgttgtc tttgagacgg agaagttgca aatggagcaa 360

cagcagcaac tgcagcagcg gcagatactt ctaggaccta acacagggct ctcaggagga 420

atgccagggg ctctaccctc acttcctgga aagatctaga ttgttattgc tgtttgagct 480

gtctcagtgg gataagtttg aaattcaagt gtttgaactg ctgataattt ggatttttt 540

ttttttttaa ctttggcaca ttgatctatc taaacctggt ggggagaatt atccccacat 600

tgtctcatgg aaagactcaa cttgcaactg tgccctccac actatcctta cttctgtctc 660

cactctgata ccagagtgca gccatgcaga tggttattcc agctctggtc acccgactcc 720

tttcaccaaa ttgctcctaa ctggaagatc tcactttccc cttgtggggt aggaaccgat 780

gccagtggga gggatgtgcc cctgaccatt aacgactgtt ttttttttt ttttttaaag 840

aatggagttg ttggggcagg acatgcacac aatgtgaaac agacaaaatg cattacacct 900

gtagtgtaaa gtggccacta tgaatcccta tgtatgagag gagggaggca ggctgcagct 960

tcagccacag aatggggact atggaagaca gcaggagctc atttcctctg cacattttgg 1020

ctgttagacc tgtgtgtgtg tttaaaaaaa gagaagtcag tgctcacttt ttgtatttaa 1080

atattaaaaa tgattccaac tgaaagtgtc atcctaagta ccttgaaatg agaccacgtc 1140

-continued

agagacatgt actgcccctc acattttctc acctaaacca gcagcacctc catcttaaca 1200

gccataggcc caaattgttt ccaagtgaaa attcattttt agccaagtac ttcatagcaa 1260

tctttgccct gaatttaggc agtcactttg agatccatca gcctaaaaca aaggattggg 1320

tctatgtact tctttagtct ataatgacac tgtgtaatta taaagtattt gtagggaaaa 1380

aaaaaaaaaa aaaaaa 1396

<210> SEQ ID NO 12
<211> LENGTH: 828
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown: Prosequence Protease polynucleotide

<400> SEQUENCE: 12

tggtacccgg gattcggcca ttacggccgg gggggtcttt tcgtatttgt cataccgtga 60

tccgaacttg ctgaaaacac tagaggtcta tgatgggact gcaaagttcc tcagagaact 120

agatgtagat gatgatgctc tcacaaaagc tattattgga accatcgggg atgttgattc 180

ctaccagcta ccagatgcta aaggttacag cagtctgatg cggtatttgt tgggcatcac 240

cgaggaggaa cgccagcaaa ggcgcgaaga gatactcgca accagcgtga aggatttcaa 300

ggagtttgcc gatgctgtcg aaacgatcaa tgacaatggg gttgtggtgg ctgtagcatc 360

gcctgacgac gtcgaggcag caaacaaaga gaagtcgtta ttttcagaca tcaagaagtg 420

cctgtgaggc ctttccattc ccacgccggc agggttcagc ctggacctgc gccgagcaga 480

ggttcttcag cgttttgctt cactgatgaa gaatggtctg gcgctgtaac atgagttcgc 540

gtggcgagag cactgatcac cttgtctgta gcccggttac ggttgtttat ttttcgcgcc 600

catttttagg ggaaataatt agttcagggt tttgagcgaa gcacaatagt taaacagagt 660

tccattgctg cctttccccc tctgcaatca gtgctcaaac aagccggttg tacaccaagt 720

tgctgatagg agattaagag ctgatgagtg atgagaacct gctcaaaatt tagctgaata 780

ctctgtaact tgatatacca aagatataaa ggtattggac tactgatt 828

<210> SEQ ID NO 13
<211> LENGTH: 3774
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 13

atgaatttac aaccaatttt ctggattgga ctgatcagtt cagtttgctg tgtgtttgct 60

caaacagatg aaaatagatg tttaaaagca aatgccaaat catgtggaga atgtatacaa 120

gcagggccaa attgtgggtg gtgcacaaat tcaacatttt tacaggaagg aatgcctact 180

tctgcacgat gtgatgattt agaagcctta aaaaagaagg gttgccctcc agatgacata 240

gaaaatccca gaggctccaa agatataaag aaaaataaaa atgtaaccaa ccgtagcaaa 300

ggaacagcag agaagctcaa gccagaggat attactcaga tccaaccaca gcagttggtt 360

ttgcgattaa gatcagggga gccacagaca tttacattaa aattcaagag agctgaagac 420

tatcccattg acctctacta ccttatggac ctgtcttact caatgaaaga cgatttggag 480

aatgtaaaaa gtcttggaac agatctgatg aatgaaatga ggaggattac ttcggacttc 540

-continued

agaattggat ttggctcatt tgtggaaaag actgtgatgc cttacattag cacaacacca 600

gctaagctca ggaacccttg cacaagtgaa cagaactgca ccagcccatt tagctacaaa 660

aatgtgctca gtcttactaa taaaggagaa gtatttaatg aacttgttgg aaaacagcgc 720

atatctggaa atttggattc tccagaaggt ggtttcgatg ccatcatgca agttgcagtt 780

tgtggatcac tgattggctg gaggaatgtt acacggctgc tggtgttttc cacagatgcc 840

gggtttcact ttgctggaga tgggaaactt ggtggcattg ttttaccaaa tgatggacaa 900

tgtcacctgg aaaataatat gtacacaatg agccattatt atgattatcc ttctattgct 960

caccttgtcc agaaactgag tgaaaataat attcagacaa tttttgcagt tactgaagaa 1020

tttcagcctg tttacaagga gctgaaaaac ttgatcccta agtcagcagt aggaacatta 1080

tctgcaaatt ctagcaatgt aattcagttg atcattgatg catacaattc cctttcctca 1140

gaagtcattt tggaaaacgg caaattgtca gaaggcgtaa caataagtta caaatcttac 1200

tgcaagaacg gggtgaatgg aacaggggaa aatggaagaa aatgttccaa tatttccatt 1260

ggagatgagg ttcaatttga aattagcata acttcaaata agtgtccaaa aaaggattct 1320

gacagcttta aaattaggcc tctgggcttt acggaggaag tagaggttat tcttcagtac 1380

atctgtgaat gtgaatgcca aagcgaaggc atccctgaaa gtcccaagtg tcatgaagga 1440

aatgggacat ttgagtgtgg cgcgtgcagg tgcaatgaag ggcgtgttgg tagacattgt 1500

gaatgcagca cagatgaagt taacagtgaa gacatggatg cttactgcag gaaagaaaac 1560

agttcagaaa tctgcagtaa caatggagag tgcgtctgcg gacagtgtgt ttgtaggaag 1620

agggataata caaatgaaat ttattctggc aaattctgcg agtgtgataa tttcaactgt 1680

gatagatcca atggcttaat ttgtggagga aatggtgttt gcaagtgtcg tgtgtgtgag 1740

tgcaacccca actacactgg cagtgcatgt gactgttctt tggatactag tacttgtgaa 1800

gccagcaacg gacagatctg caatggccgg ggcatctgcg agtgtggtgt ctgtaagtgt 1860

acagatccga agtttcaagg gcaaacgtgt gagatgtgtc agacctgcct tggtgtctgt 1920

gctgagcata aagaatgtgt tcagtgcaga gccttcaata aaggagaaaa gaaagacaca 1980

tgcacacagg aatgttccta ttttaacatt accaaggtag aaagtcggga caaattaccc 2040

cagccggtcc aacctgatcc tgtgtcccat tgtaaggaga aggatgttga cgactgttgg 2100

ttctatttta cgtattcagt gaatgggaac aacgaggtca tggttcatgt tgtggagaat 2160

ccagagtgtc ccactggtcc agacatcatt ccaattgtag ctggtgtggt tgctggaatt 2220

gttcttattg gccttgcatt actgctgata tggaagcttt taatgataat tcatgacaga 2280

agggagtttg ctaaatttga aaaggagaaa atgaatgcca aatgggacac gtctctctct 2340

gtcgcccagc ctggagtgca gtggtgtgat atcagctcac tgcaacctct gacttccaga 2400

ttccagcaat tctcctgcct cagcctcccg agtacctggg attacagggt gaaaatccta 2460

tttataagag tgccgtaaca actgtggtca atccgaagta tgagggaaaa tgagtactgc 2520

ccgtgcaaat cccacaacac tgaatgcaaa gtagcaattt ccatagtcac agttaggtag 2580

ctttagggca atattgccat ggttttactc atgtgcaggt tttgaaaatg tacaatatgt 2640

ataatttta aaatgtttta ttattttgaa aataatgttg taattcatgc cagggactga 2700

caaaagactt gagacaggat ggttactctt gtcagctaag gtcacattgt gccttttga 2760

ccttttcttc ctggactatt gaaatcaagc ttattggatt aagtgatatt tctatagcga 2820

-continued

ttgaaagggc aatagttaaa gtaatgagca tgatgagagt ttctgttaat catgtattaa 2880

aactgatttt tagctttaca aatatgtcag tttgcagtta tgcagaatcc aaagtaaatg 2940

tcctgctagc tagttaagga ttgttttaaa tctgttattt tgctatttgc ctgttagaca 3000

tgactgatga catatctgaa agacaagtat gttgagagtt gctggtgtaa aatacgtttg 3060

aaatagttga tctacaaagg ccatgggaaa aattcagaga gttaggaagg aaaaaccaat 3120

agctttaaaa cctgtgtgcc attttaagag ttacttaatg tttggtaact tttatgcctt 3180

cactttacaa attcaagcct tagataaaag aaccgagcaa ttttctgcta aaaagtcctt 3240

gatttagcac tatttacata caggccatac tttacaaagt atttgctgaa tggggacctt 3300

ttgagttgaa tttatttttat tatttttatt ttgtttaatg tctggtgctt tctgtcacct 3360

cttctaatct tttaatgtat ttgtttgcaa ttttggggta agacttttt tatgagtact 3420

ttttctttga agttttagcg gtcaatttgc ctttttaatg aacatgtgaa gttatactgt 3480

ggctatgcaa cagctctcac ctacgcgagt cttactttga gttagtgcca taacagacca 3540

ctgtatgttt acttctcacc atttgagttg cccatcttgt ttcacactag tcacattctt 3600

gttttaagtg cctttagttt taacagttca ctttttacag tgctatttac tgaagttatt 3660

tattaaatat gcctaaaata cttaaatcgg atgtcttgac tctgatgtat tttatcaggt 3720

tgtgtgcatg aaatttttat agattaaaga agttgaggaa aagcaaaaaa aaaa 3774

<210> SEQ ID NO 14
<211> LENGTH: 1546
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 14

gggactgagc atggatttcg gactggccct cctgctggcg gggcttctgg ggctcctcct 60

cggccagtcc ctccaggtga agcccctgca ggtggagccc ccggagccgg tggtggccgt 120

ggccttgggc gcctcgcgcc agctcacctg ccgcctggcc tgcgcggacc gcggggcctc 180

ggtgcagtgg cggggcctgg acaccagcct gggcgcggtg cagtcggaca cgggccgcag 240

cgtcctcacc gtgcgcaacg cctcgctgtc ggcggccggg acccgcgtgt gcgtgggctc 300

ctgcgggggc cgcaccttcc agcacaccgt gcagctcctt gtgtacgcct tcccggacca 360

gctgaccgtc tccccagcag ccctggtgcc tggtgacccg gaggtggcct gtacggccca 420

caaagtcacg cccgtggacc ccaacgcgct ctccttctcc ctgctcgtcg ggggccagga 480

actggagggg gcgcaagccc tgggcccgga ggtgcaggag gaggaggagg agccccaggg 540

ggacgaggac gtgctgttca gggtgacaga gcgctggcgg ctgccgcccc tggggacccc 600

tgtcccgccc gccctctact gccaggccac gatgaggctg cctggcttgg agctcagcca 660

ccgccaggcc atccccgtcc tgcacagccc gacctccccg gagcctcccg acaccacctc 720

cccggagtct cccgacacca cctccccgga gtctcccgac accacctccc aggagcctcc 780

cgacaccacc tccccggagc ctcccgacaa gacctccccg gagcccgccc cccagcaggg 840

ctccacacac acccccagga gcccaggctc caccaggact cgccgccctg agatctccca 900

ggctgggccc acgcagggag aagtgatccc aacaggctcg tccaaacctg cgggtgacca 960

gctgcccgcg gctctgtgga ccagcagtgc ggtgctggga ctgctgctcc tggccttgcc 1020

cacctatcac ctctggaaac gctgccggca cctggctgag gacgacaccc acccaccagc 1080

-continued

ttctctgagg cttctgcccc aggtgtcggc ctgggctggg ttaaggggga ccggccaggt 1140

cgggatcagc ccctcctgag tggccagcct ttccccctgt gaaagcaaaa tagcttggac 1200

cccttcaagt tgagaactgg tcagggcaaa cctgcctccc attctactca aagtcatccc 1260

tctgttcaca gagatggatg catgttctga ttgcctcttt ggagaagctc atcagaaact 1320

caaaagaagg ccactgtttg tctcacctac ccatgacctg aagcccctcc ctgagtggtc 1380

cccacctttc tggacggaac cacgtacttt ttacatacat tgattcatgt ctcacgtctc 1440

cctaaaaatg cgtaagacca agctgtgccc tgaccaccct gggcccctgt cgtcaggacc 1500

tcctgaggct ttggcaaata aacctcctaa aatgataaaa aaaaaa 1546

<210> SEQ ID NO 15
<211> LENGTH: 766
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown: Cu/Zn-superoxide dismutase polynucleotide

<400> SEQUENCE: 15

agattaattc acttccaggc atttcatctt cattcatttt ccaaaggggt accctgagat 60

cacaaaggat acaaaattat gggcaaggga gttcgagtgt tgaacagcag cgagggtgtt 120

aaggggacca tcttcttcac tcaggaagga aacggtacca ccactgtgac aggaaccgtt 180

tctggcctta agcctggtct ccatggtttc catgtccatg ctcttggtga caccactaac 240

ggttgcatgt ctaccggtcc acatttcaac cctgaaggta aaacccacgg tgcacctgag 300

gatgctaatc gacatgctgg agatctagga aacatcactg ttggggatga tggaactgcc 360

accttcacaa tcactgacag ccagattcct cttgatggac caaactctat tgttggaagg 420

gctgttgttg tccacgcaga acctgatgac ctgggaaagg gaggccatga actcagcctt 480

actactggaa acgcaggtgg ccgtgttgct tgtggtatta ttggtcttca gggctaagct 540

gttgcttttc gaggacgaga gtgatgtaat aaggaggttc ttacctctag acatggctag 600

tttgtgtatt ctttggtgtg tggctgtatt aattgagctt agtggctcga tgcatttggt 660

ttaagacgga agaaaacaga aaatccaaac tttttctatt tcatgaataa cagaggacgt 720

ggttgaaaac gataaaatat tgaatatgaa aaaaaaaaaa aaaaaa 766

<210> SEQ ID NO 16
<211> LENGTH: 1011
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 16

ctagagatag aattgtgact agaataaagg ctataattat tatagaggtt ttaattgttt 60

gaattgctca tggtagtgga agtagaagag caatttctag gtcaaataat agaaatgtaa 120

ttgctaccaa gaaaaatttt attgagaatg gtagacgtgc agagcttgta gggtcgaatc 180

cgcattcata tggatttgaa gcatggcagt gtcagcaact ttgtctagag ccttccagaa 240

acagataatg acaaggctct aaaagggttc cagctgaagt ttctgagcgg agcacgctgt 300

gtggctccct aggctgagtt tccaagctgc tggttcatgc cgttgacaaa ctgcaggatg 360

gtgcccgttc gcaggccgct gtcgctgctc ctcaccttct tcctctgcgc ctgtgctgag 420

acacccccca ggtttacacg aactccggtt gatcagacag gggtctctgg aggagttgca 480

-continued

tcattcattt gtcaagctac aggagaccca agacctaaaa ttgtctggaa caaaaaagga 540

aagaaagtca gcaaccagag atttgaggta atagaatttg acgatgggtc tggatcagta 600

ctcagaatac agcccttaag gactccacgg gatgaggcca tttatgaatg tgtggcctcg 660

aataatgtgg gagaaatcag tgtgtccaca agactcacag ttttacgtga ggatcagatt 720

cctagagggt tccctacgat tgatatgggc ccgcagttga aggtggtgga acggacccgc 780

accgccacca tgctgtgtgc agccagcggt aatccggatc cagaaatcac ttggtttaaa 840

gatttcttac ctgttgacac aagcaacaac aatggtcgta ttaagcagtt acgatcagaa 900

tctattggag ccctgcagat cgaacagagc gaagaatccg accaaggaaa atacgagtgt 960

gttgccacca acagcgcggg cactcgctac tctgcccctg ccaatttata t 1011

<210> SEQ ID NO 17
<211> LENGTH: 13481
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 17

gctgccggga cgggtccaag atggacggcc gctcaggttc tgcttttacc tgcggcccag 60

agccccattc attgccccgg tgctgagcgg cgccgcgagt cggcccgagg cctccgggga 120

ctgccgtgcc gggcgggaga ccgccatggc gaccctggaa aagctgatga aggccttcga 180

gtccctcaag tccttccagc agcagcagca gcagcagcag cagcagcagc agcagcagca 240

gcagcagcag cagcagcagc aacagccgcc accgccgccg ccgccgccgc cgcctcctca 300

gcttcctcag ccgccgccgc aggcacagcc gctgctgcct cagccgcagc cgcccccgcc 360

gccgcccccg ccgccacccg gcccggctgt ggctgaggag ccgctgcacc gaccaaagaa 420

agaactttca gctaccaaga aagaccgtgt gaatcattgt ctgacaatat gtgaaaacat 480

agtggcacag tctgtcagaa attctccaga atttcagaaa cttctgggca tcgctatgga 540

acttttttctg ctgtgcagtg atgacgcaga gtcagatgtc aggatggtgg ctgacgaatg 600

cctcaacaaa gttatcaaag ctttgatgga ttctaatctt ccaaggttac agctcgagct 660

ctataaggaa attaaaaaga atggtgcccc tcggagtttg cgtgctgccc tgtggaggtt 720

tgctgagctg gctcacctgg ttcggcctca gaaatgcagg ccttacctgg tgaaccttct 780

gccgtgcctg actcgaacaa gcaagagacc cgaagaatca gtccaggaga ccttggctgc 840

agctgttccc aaaattatgg cttcttttgg caattttgca aatgacaatg aaattaaggt 900

tttgttaaag gccttcatag cgaacctgaa gtcaagctcc cccaccattc ggcggacagc 960

ggctggatca gcagtgagca tctgccagca ctcaagaagg acacaatatt tctatagttg 1020

gctactaaat gtgctcttag gcttactcgt tcctgtcgag gatgaacact ccactctgct 1080

gattcttggc gtgctgctca ccctgaggta tttggtgccc ttgctgcagc agcaggtcaa 1140

ggacacaagc ctgaaaggca gcttcggagt gacaaggaaa gaaatggaag tctctccttc 1200

tgcagagcag cttgtccagg tttatgaact gacgttacat catacacagc accaagacca 1260

caatgttgtg accggagccc tggagctgtt gcagcagctc ttcagaacgc ctccacccga 1320

gcttctgcaa accctgaccg cagtcggggg cattgggcag ctcaccgctg ctaaggagga 1380

gtctggtggc cgaagccgta gtgggagtat tgtggaactt atagctggag ggggttcctc 1440

atgcagccct gtcctttcaa gaaaacaaaa aggcaaagtg ctcttaggag aagaagaagc 1500

-continued

cttggaggat gactctgaat cgagatcgga tgtcagcagc tctgccttaa cagcctcagt 1560

gaaggatgag atcagtggag agctggctgc ttcttcaggg gtttccactc cagggtcagc 1620

aggtcatgac atcatcacag aacagccacg gtcacagcac acactgcagg cggactcagt 1680

ggatctggcc agctgtgact tgacaagctc tgccactgat ggggatgagg aggatatctt 1740

gagccacagc tccagccagg tcagcgccgt cccatctgac cctgccatgg acctgaatga 1800

tgggacccag gcctcgtcgc ccatcagcga cagctcccag accaccaccg aagggcctga 1860

ttcagctgtt acccccttcag acagttctga aattgtgtta gacggtaccg acaaccagta 1920

tttgggcctg cagattggac agccccagga tgaagatgag gaagccacag gtattcttcc 1980

tgatgaagcc tcggaggcct tcaggaactc ttccatggcc cttcaacagg cacatttatt 2040

gaaaaacatg agtcactgca ggcagccttc tgacagcagt gttgataaat ttgtgttgag 2100

agatgaagct actgaaccgg gtgatcaaga aaacaagcct tgccgcatca aaggtgacat 2160

tggacagtcc actgatgatg actctgcacc tcttgtccat tgtgtccgcc ttttatctgc 2220

ttcgtttttg ctaacagggg gaaaaaatgt gctggttccg gacagggatg tgagggtcag 2280

cgtgaaggcc ctggccctca gctgtgtggg agcagctgtg gccctccacc cggaatcttt 2340

cttcagcaaa ctctataaag ttcctcttga caccacggaa taccctgagg aacagtatgt 2400

ctcagacatc ttgaactaca tcgatcatgg agacccacag gttcgaggag ccactgccat 2460

tctctgtggg accctcatct gctccatcct cagcaggtcc cgcttccacg tgggagattg 2520

gatgggcacc attagaaccc tcacaggaaa tacattttct ttggcggatt gcattccttt 2580

gctgcggaaa acactgaagg atgagtcttc tgttacttgc aagttagctt gtacagctgt 2640

gaggaactgt gtcatgagtc tctgcagcag cagctacagt gagttaggac tgcagctgat 2700

catcgatgtg ctgactctga ggaacagttc ctattggctg gtgaggacag agcttctgga 2760

aacccttgca gagattgact tcaggctggt gagctttttg gaggcaaaag cagaaaactt 2820

acacagaggg gctcatcatt atacagggct tttaaaactg caagaacgag tgctcaataa 2880

tgttgtcatc catttgcttg gagatgaaga ccccagggtg cgacatgttg ccgcagcatc 2940

actaattagg cttgtcccaa agctgtttta taaatgtgac caaggacaag ctgatccagt 3000

agtggccgtg gcaagagatc aaagcagtgt ttacctgaaa cttctcatgc atgagacgca 3060

gcctccatct catttctccg tcagcacaat aaccagaata tatagaggct ataacctact 3120

accaagcata acagacgtca ctatggaaaa taacctttca agagttattg cagcagtttc 3180

tcatgaacta atcacatcaa ccaccagagc actcacattt ggatgctgtg aagctttgtg 3240

tcttctttcc actgccttcc cagtttgcat ttggagttta ggttggcact gtggagtgcc 3300

tccactgagt gcctcagatg agtctaggaa gagctgtacc gttgggatgg ccacaatgat 3360

tctgaccctg ctctcgtcag cttggttccc attggatctc tcagcccatc aagatgcttt 3420

gattttggcc ggaaacttgc ttgcagccag tgctcccaaa tctctgagaa gttcatgggc 3480

ctctgaagaa gaagccaacc cagcagccac caagcaagag gaggtctggc cagccctggg 3540

ggaccgggcc ctggtgccca tggtggagca gctcttctct cacctgctga aggtgattaa 3600

catttgtgcc cacgtcctgg atgacgtggc tcctggaccc gcaataaagg cagccttgcc 3660

ttctctaaca aacccccctt ctctaagtcc catccgacga aaggggaagg agaaagaacc 3720

aggagaacaa gcatctgtac cgttgagtcc caagaaaggc agtgaggcca gtgcagcttc 3780

-continued

tagacaatct gatacctcag gtcctgttac aacaagtaaa tcctcatcac tggggagttt 3840

ctatcatctt ccttcatacc tcaaactgca tgatgtcctg aaagctacac acgctaacta 3900

caaggtcacg ctggatcttc agaacagcac ggaaaagttt ggagggtttc tccgctcagc 3960

cttggatgtt ctttctcaga tactagagct ggccacactg caggacattg ggaagtgtgt 4020

tgaagagatc ctaggatacc tgaaatcctg ctttagtcga gaaccaatga tggcaactgt 4080

ttgtgttcaa caattgttga agactctctt tggcacaaac ttggcctccc agtttgatgg 4140

cttatcttcc aaccccagca agtcacaagg ccgagcacag cgccttggct cctccagtgt 4200

gaggccaggc ttgtaccact actgcttcat ggccccgtac acccacttca cccaggccct 4260

cgctgacgcc agcctgagga acatggtgca ggcggagcag gagaacgaca cctcgggatg 4320

gtttgatgtc ctccagaaag tgtctaccca gttgaagaca aacctcacga gtgtcacaaa 4380

gaaccgtgca gataagaatg ctattcataa tcacattcgt ttgtttgaac ctcttgttat 4440

aaaagcttta aaacagtaca cgactacaac atgtgtgcag ttacagaagc aggttttaga 4500

tttgctggcg cagctggttc agttacgggt taattactgt cttctggatt cagatcaggt 4560

gtttattggc tttgtattga aacagtttga atacattgaa gtgggccagt tcagggaatc 4620

agaggcaatc attccaaaca tctttttctt cttggtatta ctatcttatg aacgctatca 4680

ttcaaaacag atcattggaa ttcctaaaat cattcagctc tgtgatggca tcatggccag 4740

tggaaggaag gctgtgacac atgccatacc ggctctgcag cccatagtcc acgacctctt 4800

tgtattaaga ggaacaaata aagctgatgc aggaaaagag cttgaaaccc aaaaagaggt 4860

ggtggtgtca atgttactga gactcatcca gtaccatcag gtgttggaga tgttcattct 4920

tgtcctgcag cagtgccaca aggagaatga agacaagtgg aagcgactgt ctcgacagat 4980

agctgacatc atcctcccaa tgttagccaa acagcagatg cacattgact ctcatgaagc 5040

ccttggagtg ttaaatacat tatttgagat tttggcccct tcctccctcc gtccggtaga 5100

catgctttta cggagtatgt tcgtcactcc aaacacaatg gcgtccgtga gcactgttca 5160

actgtggata tcgggaattc tggccatttt gagggttctg atttcccagt caactgaaga 5220

tattgttctt tctcgtattc aggagctctc cttctctccg tatttaatct cctgtacagt 5280

aattaatagg ttaagagatg gggacagtac ttcaacgcta gaagaacaca gtgaagggaa 5340

acaaataaag aatttgccag aagaaacatt ttcaaggttt ctattacaac tggttggtat 5400

tcttttagaa gacattgtta caaaacagct gaaggtggaa atgagtgagc agcaacatac 5460

tttctattgc caggaactag gcacactgct aatgtgtctg atccacatct tcaagtctgg 5520

aatgttccgg agaatcacag cagctgccac taggctgttc cgcagtgatg gctgtggcgg 5580

cagtttctac accctggaca gcttgaactt gcgggctcgt tccatgatca ccacccaccc 5640

ggccctggtg ctgctctggt gtcagatact gctgcttgtc aaccacaccg actaccgctg 5700

gtgggcagaa gtgcagcaga ccccgaaaag acacagtctg tccagcacaa agttacttag 5760

tccccagatg tctggagaag aggaggattc tgacttggca gccaaacttg gaatgtgcaa 5820

tagagaaata gtacgaagag gggctctcat tctcttctgt gattatgtct gtcagaacct 5880

ccatgactcc gagcacttaa cgtggctcat tgtaaatcac attcaagatc tgatcagcct 5940

ttcccacgag cctccagtac aggacttcat cagtgccgtt catcggaact ctgctgccag 6000

cggcctgttc atccaggcaa ttcagtctcg ttgtgaaaac ctttcaactc caaccatgct 6060

-continued

gaagaaaact cttcagtgct tggaggggat ccatctcagc cagtcgggag ctgtgctcac 6120

gctgtatgtg gacaggcttc tgtgcacccc tttccgtgtg ctggctcgca tggtcgacat 6180

ccttgcttgt cgccgggtag aaatgcttct ggctgcaaat ttacagagca gcatggccca 6240

gttgccaatg gaagaactca acagaatcca ggaatacctt cagagcagcg ggctcgctca 6300

gagacaccaa aggctctatt ccctgctgga caggtttcgt ctctccacca tgcaagactc 6360

acttagtccc tctcctccag tctcttccca cccgctggac ggggatgggc acgtgtcact 6420

ggaaacagtg agtccggaca aagactggta cgttcatctt gtcaaatccc agtgttggac 6480

caggtcagat tctgcactgc tggaaggtgc agagctggtg aatcggattc ctgctgaaga 6540

tatgaatgcc ttcatgatga actcggagtt caacctaagc ctgctagctc catgcttaag 6600

cctagggatg agtgaaattt ctggtggcca gaagagtgcc ctttttgaag cagcccgtga 6660

ggtgactctg gcccgtgtga gcggcaccgt gcagcagctc cctgctgtcc atcatgtctt 6720

ccagcccgag ctgcctgcag agccggcggc ctactggagc aagttgaatg atctgtttgg 6780

ggatgctgca ctgtatcagt ccctgcccac tctggcccgg gccctggcac agtacctggt 6840

ggtggtctcc aaactgccca gtcatttgca ccttcctcct gagaaagaga aggacattgt 6900

gaaattcgtg gtggcaaccc ttgaggccct gtcctggcat ttgatccatg agcagatccc 6960

gctgagtctg gatctccagg cagggctgga ctgctgctgc ctggccctgc agctgcctgg 7020

cctctggagc gtggtctcct ccacagagtt tgtgacccac gcctgctccc tcatctactg 7080

tgtgcacttc atcctggagg ccgttgcagt gcagcctgga gagcagcttc ttagtccaga 7140

aagaaggaca aataccccaa aagccatcag cgaggaggag gaggaagtag atccaaacac 7200

acagaatcct aagtatatca ctgcagcctg tgagatggtg gcagaaatgg tggagtctct 7260

gcagtcggtg ttggccttgg gtcataaaag gaatagcggc gtgccggcgt ttctcacgcc 7320

attgctaagg aacatcatca tcagcctggc ccgcctgccc cttgtcaaca gctacacacg 7380

tgtgccccca ctggtgtgga agcttggatg gtcacccaaa ccgggagggg attttggcac 7440

agcattccct gagatccccg tggagttcct ccaggaaaag gaagtcttta aggagttcat 7500

ctaccgcatc aacacactag gctggaccag tcgtactcag tttgaagaaa cttgggccac 7560

cctccttggt gtcctggtga cgcagcccct cgtgatggag caggaggaga gcccaccaga 7620

agaagacaca gagaggaccc agatcaacgt cctggccgtg caggccatca cctcactggt 7680

gctcagtgca atgactgtgc ctgtggccgg caacccagct gtaagctgct tggagcagca 7740

gccccggaac aagcctctga aagctctcga caccaggttt gggaggaagc tgagcattat 7800

cagagggatt gtggagcaag agattcaagc aatggtttca aagagagaga atattgccac 7860

ccatcattta tatcaggcat gggatcctgt cccttctctg tctccggcta ctacaggtgc 7920

cctcatcagc cacgagaagc tgctgctaca gatcaacccc gagcgggagc tggggagcat 7980

gagctacaaa ctcggccagg tgtccataca ctccgtgtgg ctggggaaca gcatcacacc 8040

cctgagggag gaggaatggg acgaggaaga ggaggaggag gccgacgccc ctgcaccttc 8100

gtcaccaccc acgtctccag tcaactccag gaaacaccgg gctggagttg acatccactc 8160

ctgttcgcag tttttgcttg agttgtacag ccgctggatc ctgccgtcca gctcagccag 8220

gaggaccccg gccatcctga tcagtgaggt ggtcagatcc cttctagtgg tctcagactt 8280

gttcaccgag cgcaaccagt ttgagctgat gtatgtgacg ctgacagaac tgcgaagggt 8340

-continued

gcacccttca gaagacgaga tcctcgctca gtacctggtg cctgccacct gcaaggcagc 8400

tgccgtcctt gggatggaca aggccgtggc ggagcctgtc agccgcctgc tggagagcac 8460

gctcaggagc agccacctgc ccagcagggt tggagccctg cacggcgtcc tctatgtgct 8520

ggagtgcgac ctgctggacg acactgccaa gcagctcatc ccggtcatca gcgactatct 8580

cctctccaac ctgaaaggga tcgcccactg cgtgaacatt cacagccagc agcacgtact 8640

ggtcatgtgt gccactgcgt tttacctcat tgagaactat cctctggacg tagggccgga 8700

attttcagca tcaataatac agatgtgtgg ggtgatgctg tctggaagtg aggagtccac 8760

cccctccatc atttaccact gtgccctcag aggcctggag cgcctcctgc tctctgagca 8820

gctctcccgc ctggatgcag aatcgctggt caagctgagt gtggacagag tgaacgtgca 8880

cagcccgcac cgggccatgg cggctctggg cctgatgctc acctgcatgt acacaggaaa 8940

ggagaaagtc agtccgggta gaacttcaga ccctaatcct gcagcccccg acagcgagtc 9000

agtgattgtt gctatggagc gggtatctgt tctttttgat aggatcagga aaggctttcc 9060

ttgtgaagcc agagtggtgg ccaggatcct gccccagttt ctagacgact tcttcccacc 9120

ccaggacatc atgaacaaag tcatcggaga gtttctgtcc aaccagcagc catacccccа 9180

gttcatggcc accgtggtgt ataaggtgtt tcagactctg cacagcaccg ggcagtcgtc 9240

catggtccgg gactgggtca tgctgtccct ctccaacttc acgcagaggg ccccggtcgc 9300

catggccacg tggagcctct cctgcttctt tgtcagcgcg tccaccagcc cgtgggtcgc 9360

ggcgatcctc ccacatgtca tcagcaggat gggcaagctg gagcaggtgg acgtgaacct 9420

tttctgcctg gtcgccacag acttctacag acaccagata gaggaggagc tcgaccgcag 9480

ggccttccag tctgtgcttg aggtggttgc agccccagga agcccatatc accggctgct 9540

gacttgttta cgaaatgtcc acaaggtcac cacctgctga gcgccatggt gggagagact 9600

gtgaggcggc agctggggcc ggagcctttg gaagtctgcg cccttgtgcc ctgcctccac 9660

cgagccagct tggtccctat gggcttccgc acatgccgcg ggcggccagg caacgtgcgt 9720

gtctctgcca tgtggcagaa gtgctctttg tggcagtggc caggcaggga gtgtctgcag 9780

tcctggtggg gctgagcctg aggccttcca gaaagcagga gcagctgtgc tgcaccccat 9840

gtgggtgacc aggtcctttc tcctgatagt cacctgctgg ttgttgccag gttgcagctg 9900

ctcttgcatc tgggccagaa gtcctccctc ctgcaggctg gctgttggcc cctctgctgt 9960

cctgcagtag aaggtgccgt gagcaggctt tgggaacact ggcctgggtc tccctggtgg 10020

ggtgtgcatg ccacgccccg tgtctggatg cacagatgcc atggcctgtg ctgggccagt 10080

ggctgggggt gctagacacc cggcaccatt ctcccttctc tcttttcttc tcaggattta 10140

aaatttaatt atatcagtaa agagattaat tttaacgtaa ctctttctat gcccgtgtaa 10200

agtatgtgaa tcgcaaggcc tgtgctgcat gcgacagcgt ccggggtggt ggacagggcc 10260

cccggccacg ctccctctcc tgtagccact ggcatagccc tcctgagcac ccgctgacat 10320

ttccgttgta catgttcctg tttatgcatt cacaaggtga ctgggatgta gagaggcgtt 10380

agtgggcagg tggccacagc aggactgagg acaggccccc attatcctag gggtgcgctc 10440

acctgcagcc cctcctcctc gggcacagac gactgtcgtt ctccacccac cagtcaggga 10500

cagcagcctc cctgtcactc agctgagaag gccagccctc cctggctgtg agcagcctcc 10560

actgtgtcca gagacatggg cctcccactc ctgttccttg ctagccctgg ggtggcgtct 10620

-continued

gcctaggagc tggctggcag gtgttgggac ctgctgctcc atggatgcat gccctaagag 10680

tgtcactgag ctgtgttttg tctgagcctc tctcggtcaa cagcaaagct tggtgtcttg 10740

gcactgttag tgacagagcc cagcatccct tctgcccccg ttccagctga catcttgcac 10800

ggtgacccct tttagtcagg agagtgcaga tctgtgctca tcggagactg ccccacggcc 10860

ctgtcagagc cgccactcct atccccaggc caggtccctg gaccagcctc ctgtttgcag 10920

gcccagagga gccaagtcat taaaatggaa gtggattctg gatggccggg ctgctgctga 10980

tgtaggagct ggatttggga gctctgcttg ccgactggct gtgagacgag gcaggggctc 11040

tgcttcctca gccctagagg cgagccaggc aaggttggcg actgtcatgt ggcttggttt 11100

ggtcatgccc gtcgatgttt tgggtattga atgtggtaag tggaggaaat gttggaactc 11160

tgtgcaggtg ctgccttgag acccccaagc ttccacctgt ccctctccta tgtggcagct 11220

ggggagcagc tgagatgtgg acttgtatgc tgcccacata cgtgaggggg agctgaaagg 11280

gagcccctcc tctgagcagc ctctgccagg cctgtatgag gcttttccca ccagctccca 11340

acagaggcct cccccagcca ggaccacctc gtcctcgtgg cggggcagca ggagcggtag 11400

aaaggggtcc gatgtttgag gaggccctta agggaagcta ctgaattata acacgtaaga 11460

aaatcaccat tccgtattgg ttgggggctc ctgtttctca tcctagcttt ttcctggaaa 11520

gcccgctaga aggtttggga acgaggggaa agttctcaga actgttggct gctccccacc 11580

cgcctcccgc ctcccccgca ggttatgtca gcagctctga gacagcagta tcacaggcca 11640

gatgttgttc ctggctagat gtttacattt gtaagaaata acactgtgaa tgtaaaacag 11700

agccattccc ttggaatgca tatcgctggg ctcaacatag agtttgtctt cctcttgttt 11760

acgacgtgat ctaaaccagt ccttagcaag gggctcagaa caccccgctc tggcagtagg 11820

tgtcccccac ccccaaagac ctgcctgtgt gctccggaga tgaatatgag ctcattagta 11880

aaaatgactt cacccacgca tatacataaa gtatccatgc atgtgcatat agacacatct 11940

ataatttttac acacacacct ctcaagacgg agatgcatgg cctctaagag tgcccgtgtc 12000

ggttcttcct ggaagttgac tttccttaga cccgccaggt caagttagcc gcgtgacgga 12060

catccaggcg tgggacgtgg tcagggcagg gctcattcat tgcccactag gatcccactg 12120

gcgaagatgg tctccatatc agctctctgc agaagggagg aagactttat catgttccta 12180

aaaatctgtg gcaagcaccc atcgtattat ccaaattttg ttgcaaatgt gattaatttg 12240

gttgtcaagt tttgggggtg ggctgtgggg agattgcttt tgttttcctg ctggtaatat 12300

cgggaaagat tttaatgaaa ccagggtaga attgtttggc aatgcactga agcgtgtttc 12360

tttcccaaaa tgtgcctccc ttccgctgcg ggcccagctg agtctatgta ggtgatgttt 12420

ccagctgcca agtgctcttt gttactgtcc accctcattt ctgccagcgc atgtgtcctt 12480

tcaaggggaa aatgtgaagc tgaacccccct ccagacaccc agaatgtagc atctgagaag 12540

gccctgtgcc ctaaaggaca cccctcgccc ccatcttcat ggaggggggtc atttcagagc 12600

cctcggagcc aatgaacagc tcctcctctt ggagctgaga tgagccccac gtggagctcg 12660

ggacggatag tagacagcaa taactcggtg tgtggccgcc tggcaggtgg aacttcctcc 12720

cgttgcgggg tggagtgagg ttagttctgt gtgtctggtg ggtggagtca ggcttctctt 12780

gctacctgtg agcatccttc ccagcagaca tcctcatcgg gctttgtccc tcccccgctt 12840

cctccctctg cggggaggac ccgggaccac agctgctggc cagggtagac ttggagctgt 12900

-continued

cctccagagg ggtcacgtgt aggagtgaga agaaggaaga tcttgagagc tgctgaggga 12960

ccttggagag ctcaggatgg ctcagacgag gacactcgct tgccgggcct gggcctcctg 13020

ggaaggaggg agctgctcag aatgccgcat gacaactgaa ggcaacctgg aaggttcagg 13080

ggccgctctt cccccatgtg cctgtcacgc tctggtgcag tcaaaggaac gccttcccct 13140

cagttgtttc taagagcaga gtctcccgct gcaatctggg tggtaactgc cagccttgga 13200

ggatcgtggc caacgtggac ctgcctacgg agggtgggct ctgacccaag tggggcctcc 13260

ttgtccaggt ctcactgctt tgcaccgtgg tcagagggac tgtcagctga gcttgagctc 13320

ccctggagcc agcagggctg tgatgggcga gtcccggagc cccacccaga cctgaatgct 13380

tctgagagca aagggaagga ctgacgagag atgtatattt aatttttaa ctgctgcaaa 13440

cattgtacat ccaaattaaa ggaaaaaaat ggaaaccatc a 13481

<210> SEQ ID NO 18
<211> LENGTH: 36600
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown: N-Methyl-D-Aspartate Receptor polynucleotide

<400> SEQUENCE: 18

atccacccgc ctctgcctcc caaagtgcta ggattacagg catgagccac catgtctggc 60

caggaaaaat gggaagtttt aaatgctttt ccagtagcac tggagacagg gtgagatgtc 120

tgctctcatc atttatattg tcacggtgct gggtttctat acagtcctct caggcaagaa 180

gagaaataaa aggtggctgg gcgtggtgac tcaccctgta atcccagcac tttgggaggc 240

cgaggtggga ggatcccttc agctcaggag tttgagacct tcaagaccag cctggacaac 300

acagtgagat cccatctcta aaaaaaaaaa aaatacaaaa attagccggg tatggtggct 360

tgcacctgta gtcccagata cttaggaggc tgatgtggga ggctcacctg agcctggggg 420

gttggaactg cagttagctg agatcatgcc actgcactcc agcctgggtg atagagcaag 480

acgctgtctc aaaaaaaaaa aaaaaaaaaa aaaggctggg cacaatgact cacacctgta 540

atccctgcac tttgggaggc caaggcgggc agatcaactg aggtcaggag ttcaagacca 600

acctggccaa aatggtgaaa ccctgtcttt actaaaaata caaaaattag ccgggcgtga 660

tggcgggcac ctgtaattcc agctactcag gaggctaagg caggagaatc actcgaaccc 720

gggaggtgga ggttgcagtg agccaagatt gcaccactgc actccagctt gggtgacagg 780

gtgagatcct gtctaaaaaa aaaaaaaaaa atccagtcag gaacaggaat gctaaggaat 840

caagagcagc aacaacaagc tagcagaagt aacaactaac gaacaagttc ataaagatca 900

cagggaggac cagaagacct gcattcactg tcaattttat ttatacttat taacaataaa 960

tggtctgaaa ataaaattaa gaagacagct tcattcacaa ttgcatagaa aaataaaata 1020

ctttagatta aaattaagaa aataaatgca aaccttgtac attaaaatct acaaaacatt 1080

gctgagagaa agcaaagaca gcctaaaaaa atggggagtg atcctatgtt cacagattag 1140

aagattcaat attgtggccg cgtgtggtga ctcatgcctg aaatcccagc actttgggag 1200

gccgggggcg gggggggggg tggattatga ggtcaggagt tcgagaccag cctggccaag 1260

agaccagcct ggccaatatg gtgaaaccct gtctccacta aaaatacaaa aattagccgg 1320

-continued

gcatgatggt gggcgcctat agtcccagct actcgggagg ctgaggaagg agaatggcgt 1380

gaacccagga agcagagctt gcagtgagcc gagattgcag cactgcactc cagcctgggc 1440

aacagagtga gactctgtct caaaaaaaaa tatatatata tatatattat ttttatttt 1500

agtagagaca gggtcttacc atgttggtca ggctggtctc gaactcctga cctcaggtga 1560

tcctcccatc tcggcctccc aaagtgctgg gatgacaggt gggagccgcc gcgcccggct 1620

gggaggtgtt ctttctagac ctcacctggg agtcacgcac cattacctct accacgttcc 1680

atttgtaagt gcaggccatg tatgcctgga gggaaatcaa tcttctgccg aacagggtgg 1740

tgttcaaagc accaccggct ccaccacact ctgcctttta ttctgcattc tgtttcttga 1800

gaccacgcgc cgctacgtgg aatggtctca ggagctcatg tgtttgtgct ccatgaagtc 1860

agaaagtcca gcctttgcac tgccacatac ccacccttag aacagcgtct ggtacacggt 1920

aggtgctcag tgaatgtgcc tagtggagta aacgtgcggt gcggtgctgc ctgcggtcgg 1980

atctgtgggg atcctgtgat ggggaagacg gctaccagga aaggtggaat ttggaaagtc 2040

aggacaggga gagaggaggg actttgtgga ggcaggaagg tatgggagcc agtccagcac 2100

aagggctggc gggaaagcct cgtgggtgtc agagactata cctctgtgag ggtccctggg 2160

ctcccccagg tcagggcaga ggtcctgacc ccagtctagc tctttagcgg gggccaggcc 2220

cgagatggcc agctccgttt cccctgcgtg gcctggcagc ccctccagga gctggcacgg 2280

gaagcaggcc tgtcctccac gcccttggcc ggccttggct gcgatgggtg gagacgcttc 2340

tgcccgcccc acccctgtct gtctgcctcc tcctctcaag caggcttcct ctcgacagaa 2400

atgtgtctgc aggtggctct tgagccccag ttctcaggga ttcacctgca gaggatgtag 2460

agccggggtc ccttccctgc ttctgatctg cccaaaatcc tagggagagc cctgattggc 2520

aacccctcgg ccaatggact gaggccaggg cgcgtctgca cccgggttga cagccctgac 2580

tagaaccgcg aggctggggt cagggaggag cagccgcctc ccgccggggc cgcagctgtc 2640

agagtgggca aggggatctc cacagaggga gactgtccca cagtggcccg gggtagaaca 2700

gcgcttccct ccccacccag cgcctgactg ctggccccgg gcccatccca gggcagtgcc 2760

ccaggtctct ggtcctcccc acagcctttg tctcaccagg tctcagaggg gtccgagtcc 2820

aggctcatct gttgtggtac cactgccccc tccctcatgg ttgtggaaga agcccacggg 2880

gggaaaatgc agagcagaat ccaggaaaga ctcgctcaca ggaggagctg tggcctggag 2940

acgcggcctc caccccatct ccatcatgga caacctgcca aggggtcttg aacagaagct 3000

tggggacgca catacggtcc ggtggggcag tgacccactg gcagaccagc tccctgagta 3060

gaaaagaaaa agccctggtt tgtggggttt gatgattccg tggcagcaga gataatccca 3120

ctgtggccaa tttcgaagcc tgactgtaac agctgctggc aggtgtctga acatttagcg 3180

atcagctctc cagggtcagc gggagtcgct ccagcacccc acggggtggg ggtgggtggg 3240

tggagccagg gcctctgctg cagcctccaa ccccttttctt ccacgctgat gaggcgctgt 3300

ctgggctcag gtctaccccc acgcagctgc caggctctgg cctgcctgag gcggatactt 3360

ttttttttct tttttttgat ggagtgttgc tctgtcgccc gggctggagt gcaatggtgc 3420

gatcttggct cactgcaaca tccgcctccc agattcaagc gattctcctg cctcagtctc 3480

ccaagtagct gggattacaa atgcctgccg ccacgcccgg ctaatttttg tatttttagt 3540

agagacaggg tttcaccatg ttggccaggc tggtctcgaa ctcctgacct taggtgatcc 3600

-continued

acctgccctt gcctcccaaa gtattgagat tacaggcatg agccactgtg ccctgccgag 3660

gtggaagttt ggagatgggc cacaaactcc tggagatagg gccagctcag tccccctgag 3720

cacccacaca caccctcctg agagccaaca gaaggagtga gggccccgag gcggatgacc 3780

cgtgtgcctc aacccgaacg gggagaggcg gggtctgcag cagggtacgg gcaggtgatc 3840

ccccaaggaa agattttcct gtattgagag agaaggggcc aagaggagga gcttgtcaaa 3900

caccacagcc cctccccctc ctctcagctc caggggggtcc ctggtgccag tgttcggctg 3960

atggagagaa cggcaagcgg gagagagagt gtgacccctg tgggcacatg acttcccttg 4020

ctgcactgct gcacatagca gaggtgtggt gacgaccctg ttttgtccca ttgggggcgt 4080

ttgctgttag gtctgcagaa tcctcagttg ctattggaaa tggtgacatc actggcaggg 4140

gcggagcttc agccatcctt caagttaggg aggggcacgc acactccagg ggtggagggg 4200

gacaaagaca gggtggtgtg gaccagaggg atgggtaagg ctctggaaaa gggggcgctg 4260

ggagcgcatt gcgaggggggc tggagaggga gagaggagcg gaagctgagg gtgtgaaacg 4320

gctggccccg aacacacctc gcggcgctcc agtgattcct ggtgtccgac ctcagcccca 4380

gtcagtgcgg gtccagtttc caggctctcg cggaaggcct ggctgagcac atgcggcagc 4440

cacggtcacc ctccctattc ctcttagccc gaggaggggg gtcccaagtt acatggccac 4500

gcagatgggg cctctccctc atttctgaac cttgtgggga ggggaacctt gaagggagcg 4560

ccccccagag ccatggctta gggcctcccc cacccctctg gagctccagt ctgcaagagt 4620

caggagccga aatatcgctg actgtgggtg acgactcttg cgcgcacaca cacatacaag 4680

cgggcacgac gcgttcggtc ctattaaaag gcacgcaagg gtgcgggctg cacgcggtga 4740

cacggacccc tctaacgttt ccaaactgag ctccctgcag gtccccgaca gcacaggccc 4800

ctgtcccagg acccctccag gcacgcgctc acacgcacac gcgcgctccc cggctcacgc 4860

gcgctccgac acacacgctc acgcgaacgc aggcgcacgc tctggcgcgg gaggcgcccc 4920

cttcgcctcc gtgttgggaa gcgggggcgg cgggaggggc aggagacgtt ggccccgctc 4980

gcgtttctgc agctgctgca gtcgccgcag cgtccggacc ggaaccagcg ccgtccgcgg 5040

agccgccgcc gccgccgccg ggccctttcc aagccgggcg ctcggagctg tgcccggccc 5100

cgcttcagca ccgcggacag cgccggccgc gtggggctga gccccgagcc cccgcgcacg 5160

cttcagcgcc ccttccctcg gccgacgtcc cgggaccgcc gctccgggggg agacgtggcg 5220

tccgcagccc gcggggccgg gcgagcgcag gacggcccgg aagccccgcg ggggatgcgc 5280

cgagggcccc gcgttcgcgc cgcgcagagc caggcccgcg gcccgagccc atgagcacca 5340

tgcgcctgct gacgctcgcc ctgctgttct cctgctccgt cgcccgtgcc gcgtgcgacc 5400

ccaagatcgt caacattggc gcggtgctga gcacgcggaa gcacgagcag atgttccgcg 5460

aggccgtgaa ccaggccaac aagcggcacg gctcctggaa gattcagctc aatgccacct 5520

ccgtcacgca caagcccaac gccatccaga tggctctgtc ggtgtgcgag gacctcatct 5580

ccagccaggt gccctccccc acctccgcca cccacctccc ctctcctcca tcctgcaacc 5640

ccacaccccc agtttcattc catcctttcc gtgccccctt cctccctgta agacaccacc 5700

ccagagtcag ctggctgctt ccgggaggcc tcgtctcact aggaaccaaa caccagggtc 5760

tgctggctcc cctatcttgg cctgagacca gtcacctgcc accttggctg gtcctcagag 5820

ggcccctggg gctccaggcc ctgactggtg tgtgtagacg tggggctgga gtgtgtcagt 5880

-continued

gtgggggtgg gcattccggg taagagagta gaagcgcctg tccagctaca tgcccgccct 5940

gcagagcttt aaacaggacg gggcctgggg ccatctttgt ttctgcttcc aggttctcct 6000

gccctttctt tcgtcccttc cccctaccga tgggtccgcc tgggaagaga aatggctcag 6060

gtgccacggc aggacgcttt gtgggggtgg gagtgggggt gcacacgcga gaggcatcag 6120

ggcatgggag ctgtcggcag ccagcgctgc gggggaggac gtggctcctg ggattttgcc 6180

tgtcggagct gtccgccccc gggccgagcg cctgctgaat tccaatgagg ctgcaaggat 6240

ctgcaatgca gccctttatg taagaggcaa gacagacatc cagcctagca ccgctcacac 6300

gtgcctacct gatggacaca ccacatctgt ggacacacat gctcacactc acaccaaatg 6360

ttacattagc acacactcat gcacctcagc atcacacaat caatttcata tgctcatctg 6420

cacacatgca gatccattga cacctgctca tgtgccacac acggcttggc atgcattccc 6480

agaggcacgt gcaaacatgc acatttacac acatggttcc agtcattcac acgcatgtac 6540

acgaacagac atgccagggc atgtgatgca cataaccata ccctagcaca cgcgtgaaca 6600

cctgcatggt cacacacgga cctacgggtc tttgccaagc acctctgggt gcaggctgga 6660

agcaagagct gggggaggga gaaccacttc aaacagctgc agctgcaggg cccacaccag 6720

agttttctca gaaatcctcc ctccccactt cacaagccac ccccgtgccc cagcccagga 6780

caccatggga tgggactggg gggatgcatc tgtagccagt ggctgcagtc acatattcca 6840

tctgggactg gggagggaca cggaaggtgg actcaggaaa tccaggaggg gccattcctg 6900

gggaattgct tcaactcacg cccatgttgc tgtctgtctg tgggcatggc ctctgcagca 6960

aaggcaatgc ctgcagctac cactcacgga acacaccccc ggccaggtac tgtcctgcca 7020

cgtggggcca tgcagtgcac gcccccattc gccaaagctc taagaggcac aggcagactt 7080

ggggacagac gcaggtcctt gctgtgtgaa ggtggtgtgg accaccagct gcctgcctcc 7140

ctgccttggg aggctgggga gagagggagg catcagctcc agggggctga gccgctggct 7200

ttagatctgc cccatggggc cttggtcatg ggcaggaagg ctgggctgca cccccaatgc 7260

ctcccttccc ttccttgagg atgaggccag cactcaaagt aagggcttgg tgttgttcag 7320

acagagcccg tcacaggccc tgcccctgga gacaccagca aaagggatct cggcctcttt 7380

ggcagctcct agctgcttcc ccctgaagtc cggtaccacc cttcagagct gcccgccctg 7440

tctcgggatg tgggctggcc ccacccctgg cccatcagga aggacgggtg ggttctgaga 7500

atcaaggcca tcatgatgca ggaccagcca tcctccccgg tccaccttgg gtgcgtcccg 7560

tgctccaggc ccccaggaca tcccaagggc agtccttcac ctggcccttg agcacaacac 7620

ctgcagggcc ctagtcagca gtgtgagagg agtgagggga ggtccgggtg gggtctccct 7680

cccctgccct gtgggcatgt gtgcatctgg gcctgggcat gtagcatgta cccgaatcat 7740

gcccccagcc ccccttagcc tgctgggttc agcccctgct gcttccagat ctcagcctct 7800

aacccagtgc ctggctcacc cctgactcaa gctgatcatg tctcctgtgt ccacaggtct 7860

acgccatcct agttagccat ccacctaccc ccaacgacca cttcactccc acccctgtct 7920

cctacacagc cggcttctac cgcatacccg tgctggggct gaccacccgc atgtccatct 7980

actcggacaa ggtaagcctg actgccagac caggccttcc ggccctcggc cccagggcac 8040

agcctggcca ctccaggagc agcgggccga cccgctcaca tggaactcac acaccacaaa 8100

cagccacaca gctcccccac attcatgcac gtccacacgc tctcacgtgt ccaactcaca 8160

-continued

catctgcaaa catgctcaca tgcacactca tgtgctctta cacacacaat acacactctc 8220

ttgcacatag agggctcacg tggagcccag cacgtgcccc cagcccagag caggccaaag 8280

ggaggggca cacatcacac actcacacat cacacacaca tcacacactc acacattcat 8340

acagcaccca cacgctacac tgctatgctc accctcccca cacatgaaca ctgacacacc 8400

catggattcg cacaaagtca cacacactca ctggcacagg caccagtgac acccccctcag 8460

gacgagaggg cccgtgggct aggagaaggg atggctggga ggctttctag acaggtggac 8520

tttgaagggg agtttggaga gctgggggtt gctccaggag gaaaggggtg tgcacgcagc 8580

cagggtggtg gggccagcct ttcccactgc aggcatgggt ggagagcaat gtctgtggtt 8640

gcagctcagg gtcgggggcg ctggcctggg ggcttccagc ctctagggct gagggcacct 8700

tggcttagcc tcctgcagac cctcctggcc cacaggctat gaggagggct tctgtccaat 8760

cctggaacat cagctggaag agaggagggt catccagtca gttttgcagg aatctccaag 8820

ccagagagcc atgggggctt gctctaggtc acacagccct ccgtctaccc aggatgcaaa 8880

ctgggcactg agaggctgac caacctgggg ccactggcag acagacctgc agggccactt 8940

ggcaggggac atccagtttg gtgccagcgc tgaggagcca gagggctggg ctgtgcagcc 9000

aggcttctgg gtcccccacc ttctccaaat tcctcctgcc cagagtccac agtccttggt 9060

aacactgcct taaagcacag gggtcgcccc agccaggccc aggctcttct gggaggatgg 9120

aaggccccag aggcaggaac tgagacagag gctggaacag ccaccttcct gaggctctga 9180

aagccctggc gtgccccctc ggcacccaaa ctgctcctcc caggtgactt cacctctggc 9240

catgggaagg tgggggtctc attcctgtgc cctcaggcac gacctccttc gcctctctgg 9300

gcaccagttt cctcacatgt gaaggagaga agatggcctt acccaggaaa gcagccagtg 9360

gtcaaatgaa aggcgggggа aacggatgcc cctggcccag agcaggccaa agggaggggg 9420

cacagctcac agctgcaccc tggccttacc acagtctcta cactacaacc aacctgtgcc 9480

caaaacatga accggcaagg ccaggtcaga gctagtccaa gacctcaagc acagcctgcc 9540

ttgccaccac gtcaccaggt ggatagacag aagcagggga cattttgca ccccaaggca 9600

ctgccccagg ccacaaagag ggagcaggtg aaaaataacc tggaagcctc agaggaccac 9660

aagatcagca agagtccaca gggacactga aggaaccagg gcttacctgg acagacacag 9720

agaactgagg cagaggggggg cagagcctgc tccactcccg gccatgccac ggcactccgt 9780

ggcagcttga agccaggaaa agcaagccag ggcaagcaag caccacgctc tcgcctgggg 9840

agatgaggcc tttagcccca agagtgaatt cttcttcata catagagttg tttaaatttg 9900

ggaggactct atgggcagcc ccaggggggat cttcgaggcg ctatgtgtca tcaagaattt 9960

cctgagctca gcttgtccaa aggtggtggg ctgcagggga agaggtgagc tcaccccagg 10020

cacaattcca cagaaaccca cgtcccttag ggtgctatgg ggccaacact aaacctcctc 10080

catttccgag attatatgtg ggaggagagg ccggggtggg agagaggttc ccagggtcta 10140

aaaagtgtcc ccaggatggt ggggacaggg gtgggaaaaa ggaggggtcc cagtgtctag 10200

aaagtgtccc caggttggcc gggcgcggtg gctcacgcct gtaatcccag cactttggga 10260

ggccgaggcg ggcggatcac gaggtcagga gatcgagacc atcctggcta atacggtgaa 10320

accccatctc cactaaaaat acaaaaaaat tagccgggcg tggtggcggg cgcctgtagt 10380

ctcagctact tgggaggctg aggcaggaga atggtgtgaa cccaggaggc ggagcctgca 10440

-continued

gtgagccgag attgcactcc agcctgggta acagtgcgag actgtttaaa aaaaaaaaaa 10500

agtgtcccca gggtggtggg gacaggggtg ggagacagga gggggtccca gggtctagaa 10560

agtgtcccca ggggtgtggg gacaggagtg agaggaaggg ggtcccaggg tccagaaagt 10620

gtccccaggg tggtagggac aggggtggga gacacgaggg ggtcccaggg tctagaaagt 10680

gtccccaggg tggcagggat gggatgggag acacgagggg gtcccagagt ctagaaagtg 10740

tccccagggg tgtggggacc ggggtgagag gaaggggggtc ccagggtcca gaaagtgtcc 10800

ccagaggggt ggggacagga gtgagaggaa gggggtccca gggtccagaa agtgtcccca 10860

gaggggtggg gacaggagtg agaggaaggg ggtcccaggg tccagaaagt gtccccagag 10920

gggtggggac cggggtgaga ggaaggggggt cccagggtct agaaagtgtc cccagagggg 10980

tggggaccgg ggtgagagga agggggtccc agggtctaga aagtgtcccc aggggtgtgg 11040

ggaccggggt gagaggaagg gggtcccagg gtccagaaag tgtccccaga ggggtgggga 11100

caggagtgag aggaaggggg tcccagggtc cagaaagtgt ccccagaggg gtggggacag 11160

gagtgagagg aaggggggtcc cagggtccag aaagtgtccc cagggtggta gggacagggg 11220

tgggagacac gaggggggtcc cagggtctag aaagtgtccc cagggtggca gggatgggat 11280

gggagacacg agggggtccc agagtctaga aagtgtcccc agcggggtgg ggacaggggt 11340

gggagacacg agggggtccc agggtctaga aagtgtcccc agggtggtag ggacggggtg 11400

ggagacacga gggggtccca gggtctagaa cgtgtcccca tgggggtggg gacaggggtg 11460

ggaggagggg ggttcccagc ctccggcggg tgttccggca gtgggaggcg ggtgggaggg 11520

cgggtccccg cgggtccacc tcagcccgcc gtgcccccgc ctcccgcaga gcatccacct 11580

gagcttcctg cgcaccgtgc cgccctactc ccaccagtcc agcgtgtggt ttgagatgat 11640

gcgtgtctac agctggaacc acatcatcct gctggtcagc gacgaccacg agggccgggc 11700

ggctcagaaa cgcctggaga cgctgctgga ggagcgtgag tccaaggtga gggtcggcgc 11760

cgcgggtggg cgcctggcgg agccgaggtg caggacgggc cgccttgtgt ctgtggctcc 11820

gtgtgtgaca ccctcttctt tccatcgtgc atggtcagca ccaccacgtc tggcgagcgc 11880

ccgccccagc ctgtcctcgg ctcatttcac tcgcttttgc cattagtcga aatctccttc 11940

gtgtcagtcc ctgcggggcg agggccagac cacctggagc tccgcaaaca cccctgcccc 12000

gcgctgccga gcgccctcgt cccctcctcc tgcccatgcc cctcgctccc tggaggcccc 12060

agccgggctt gggactcgtc acccttcccg cccaccctgt cctgagtccc cagcagcctc 12120

cctctgggca aggtctcccc tgtacacgac ccccacatac cgcccctgca ggcctgtccc 12180

ctcctggctg ggcccattcc ctgtcctccc ccgcgtggcc tcccctgaag ctctgcacct 12240

catggctcag caaagccctg tccacagaca cctgcccccc agcctggacc gcccctgtgg 12300

gctccactcc cctccttgcc cccaccacag ggctccctct gcactcctca ccaagaccca 12360

ttagccacct ggttctagga cacactgacc cccaacacag ccaggcgtcc actctgtggg 12420

gctgcagaga gatcagagct gggatttggg ggggtccgag cgacccctttt ccccttcctt 12480

accactccca tttgcagtct ggggcagagc tggttctcgg ctacagaccc ccggagccct 12540

gggtgctcag cacctgggcc agctcctgat caagcagtgg gaggaggccc aggctgagga 12600

gggccagacc tatgggtggc tgggagcatg tttcgtgtca gagctggctt catcggactt 12660

agggccaacc tagcaccccc caaggcaccc caggccccgg gagggaccag agggcatggg 12720

-continued

tcgggggtaa agccaggggc agaccagagg gtcctgggag tactgtcgtg gggggtctgc 12780

tgagtcttgg gggggagggg catgggcacc aagggcccca cccaagacag tgcccctcac 12840

cccagtgccc gacaggcccc tccccccagc gcccgacagg cccctccccc ccagcacccg 12900

acaggcccct caccccagcg ctcgacaggc ccctcacctt ctggcatgtc cacccacggc 12960

cacggactcc catcacacac tcacccctgc gcacccaact gcctgctttc actcacaggt 13020

gcatgcacac attcccatca cactccacac ctctggccac aggctcaggc ttgcctccat 13080

gcagctccag cgccacacac acacctggac acgtactcag gtgcgctcct cacacacaca 13140

cctggacacg cacaggtgcg ctcctcacac acacacctgg acacacgctc aggtgcgctc 13200

ctcatgtgca tgctcaccct tacttgcgcc agccagcaca gacacacatg cacacacgca 13260

cacacgtgca caggcacaca catgcacaca tgcacacgca cacacatgca cacacgcaca 13320

cacaagcacg cacacacgca catgcacaca tgcacacaca caagcatgct cagaccatct 13380

ggcccttccc caaccttcac aggcctttgt ggactaaccc tcccatgctg acacccacag 13440

gcgcatgcca cccctgcagg cgcacataac gcacacacac cctcttgggc gcatatgcga 13500

gccgaagggc gtgggcaccc cagttagtga ggatacctcg gtctcttgag aggcagaggg 13560

agaccaagga gagggagggg gagggacatg gggacaggcc ccgggggggc tgctcacctc 13620

ccactcagga ctgacacagg ttggagtggg cactgctggg gccacacagc aggtgcacgg 13680

cagggtgggg gcggcaggtg gggctccctc cgaacggtgg acgcggacag ggcctccttt 13740

tctcccgaga gcgaccgttt ccaagagcac agcttcgtgg cagggagcct ccacggcccc 13800

gcccctacga ccctcacccc cagctccacc ccggccccca gccccgcccc gggcctcggt 13860

gtttgcgagc tccaggtagg agcccgtctg cagacgtgcc gaggaggtgg tgtgattgct 13920

ttagcgccgt cattttcaac cgtttataat cttcttctgt gtctgcatat tttctctgtg 13980

cacattattc atcagagtaa aaaaaggaac tatgaaaacc tcgaccaact gtcctatgac 14040

aacaagcgcg gacccaaggt atatatgcat ggacgtgcac gccacccacg gctagggagc 14100

cctggcctcg gcgcctcggc cactagggcc actgtctggc ccagccgccg agccgcaggc 14160

ccaagcaatg aggagaggca gccggcagca ggcaggagag ggcaggcagg agagggcagg 14220

cgtggcggcg tgggtgcctg aggcccaccc gccgtgaccc tagcctgcac cctcgaggca 14280

ggcgcggctg caggaggagc tgctctccgg gaagtgcatg cgaatctccg agaccccaag 14340

ggtttcctcg tggaacccgg gagggagccc gcccgcctgg ccacccactc gccggggccg 14400

ggctgctcct caggggcctg cggaatcaaa cctcagagga cctcccatgg ttttggaaaa 14460

gtcagcccca tctcttttcc tggttgcatt ccaaaactct tttctgtttc ccgtccgctg 14520

cacgcctcct gagtctgggt ccacttcagc ttgcatgctc agtgcaaaga tgaggacagg 14580

agtgaggtgg agagagagat ccagagagag cagagagagc acagaacgag cacaggtgag 14640

cgcgcaggct gaagacagga caggaccgga gaggcgaggc caggccaggc aaggactgag 14700

gaaggcaggc ggaggcgcag gtggcaggcg cagaccatgg cagccctagc taagctgcct 14760

cggggttccc agcggctccg gcccaactct cacccctgag gcgctatgtc ccctgcccca 14820

gccgcctgct aacactcttg ctcacaccgc aggcagagaa ggtgctgcag tttgacccag 14880

ggaccaagaa cgtgacggcc ctgctgatgg aggcgaaaga gctggaggcc cgggtcatca 14940

tcctttctgc caggtgaggc tgggcagggc cctacacact ccacacagga tggtacctga 15000

-continued

gccaagtacc cgccatctga gccagagctg ggacattgtt gggcacagtg accttcagct 15060

tccaaagcac cttcaccaag gacagccacc cccaccccca cccgcaccca cactcctatc 15120

ggcatggctg atgtgacacc ctccatctgt ccctcccttc ctgggccctt ccccattcca 15180

cagtcacatg ctgctgctgc cctgagctgg gctgtgggag agggatatgg aagagaccct 15240

gcccttggga agccccgcaa ctcaaggggg caaacctgtg caaacaaggc ccaggcctgg 15300

ggccagggac taggcccagg gaagtgatgt gcaggtgtgc caggcgaaaa ggcccacagc 15360

agggagaggg ccagatttcc caactgaagg attgcaacag ctgcagcagt agcttagggg 15420

gaaatcagtt aggtacccgg gaaatcaagc tgctctggac aggtccggtg ccacagagca 15480

accttggggt ggagcccact gtaaaaagct ccctatttgc aaatggctag gttctcccgg 15540

gaaggaaaag cctgggtgac tgggaatagg aaaagcaaga gtggggaagg gcagggtgag 15600

gagccttgcc ccatgggaca gccaaaaccc cactgtccct gacctaaaag ctctgctagg 15660

ctccaacaga gcggagcaaa acagcagtga aacacccgga ggaacacagc ccagccctcc 15720

agccctgcat atgggaaaga gccggcgaca ctctcagtcc cagggcagac caccatttcc 15780

acggtccatc aaatgaccct ctagcacgga gacagatgca gccccctcac cagggcagaa 15840

cgcagggtgg ggccagccag ggctcctcgg actagaaggc aggaatctcc ccagcccaag 15900

gtagggttgt tgcttagagc tgcccacggg gccttacatg ctcctcctgc agctgctata 15960

aaaaggcact aatcgtcccc cattacgccc cctgcactcg gctttaagct cacaggtcac 16020

ttgtccactc cactcatcca actgcaagcc ccagggaggg ttggcctggg ctccaggaaa 16080

aaagattttt aaagacgtga ccaggcaaag tcccaagggt atgcacaggt cccaaagaag 16140

gaatgccccg cccagggaaa gacccgccca gaaaaaaaga cccgccaagg gaaagctccg 16200

cccagaaaaa gacctgccca gggaaagccc cgcccagaaa aaaggtcctc ccagggaaag 16260

ccccacccag ggaaagctcc acccagataa aagcccgccc agggaaagcc ccgcccagaa 16320

aaaaagatcc gccaagagaa gtctccgccc agataaaagc cccgcccaga aaaagacccg 16380

ccaaaggaaa gcccagccca gaaaaaagac ctgcccaggg aaagccccac ccagggaaag 16440

ctccgcccag ataaaagccc gcccagggaa agccccgccc agaaaaaaga ccagcccaag 16500

gaaagctccg cccagggaaa gccccgccca gaaaaaaaga cctgcccagg gaaagctctg 16560

cccagataaa agcccgccca gggaaagccc cgcccagaaa aaagaccagc ccaaggaaag 16620

ccctgcccag aaaaaagacc gcccagggaa agccccgccc agaaaagacc cgcccaggaa 16680

aagctctggc cagataaaag ctcagcccag ggaaagcccc gcccataaaa aagccccgcc 16740

cagataaaag ccctgcccag cgaaagcctc gcccagggaa agccccaccc agaaaaagac 16800

ccgcccaggg aaagcccagc ccagagaaaa gacccgccca gggaaaactc tgcccagata 16860

aaagacccgc ccagggaaag ccccgcccag aaaaagtccc gcccagggaa agccccgccc 16920

agaaaaaaaa cccgcccagg gaaagccccg cccagaaaaa agtcgcgccc ggggaaagcc 16980

ctgcccagaa aaagtcccgc ccagggaaag ccccgcccag ggaaagaccc gcccagaaaa 17040

aagtcccgcc cagaaaaaag tcgcgcctgg tgaaagccct gcccagaaaa aagaccagcc 17100

cagggaaagc ccctcccaga aaaaaagacc cgcccaggaa aaagctctgg ccagataaaa 17160

gctccgccca gggaaagctc cgcccaggga aagccccgcc cagaaaaaaa gccccaccca 17220

gggaaagccc agcccagaaa aagacccgcc cagggaaaac tccacccaga aaaaagacca 17280

-continued

gcccagagaa agcccaaccc agaagaaagc cccacccagg gaaagccccg cccagggaaa 17340

gcttcaccca gagaaattcc cacttagaga aattcccact cagagaaagc cccacccaga 17400

aacgtgccac ccagggaaag ccccacccag tgaaagcccc taccagaaaa agccccgccc 17460

agaaaaagcc cctcttagag aaagccccgc cagactctca gaagttaatt tccttttttcg 17520

ttgttttgag agggagtctt gctttgtcac ccaggctgga gtgcagtagt acaatctcag 17580

ctcactgcaa cctctgcctc ctgggttcaa gcgattctcc tgcctcagcc tcccgagtag 17640

ctgggactac aggcacaagc caccacaccc ggctaatttt tgtattttta gtagagacgg 17700

ggtttcacca tgttggccag actggtctcg aacttgtggc ctcttttttt ttttttttt 17760

ttttctttga gatggagtct cactctgtca cccaggctgg agtgcagtgg ctcgatctcg 17820

gctcattgca agctccacct cccgggctca cgccattctc ctgcctcagc ctcccaagta 17880

gctgggacta caggcccctg ccaccacgcc ctgctacttt tttgtatttt tagtagagac 17940

ggggtttcac catgtcagcc aggatggcct caatcttctg gcctcatgat ctgcccgcct 18000

cggcctccca aagtgctggg attacaggag tgagccaccg tgcccagcca cttgtggcct 18060

cttctgttat tttctgaatt gtttacactt cccttactca tcacagagct tgagagaaat 18120

tctgtagctg tgatgggata aaaggatgga tgggcaggtg aacaaatgga cagacagctg 18180

gctgggaagt ggaatttctt catccaagat gggtcaactc aaggaatcga ttgccctaaa 18240

acatacctgt tccactgttg gccacttttg caatagacaa gttcacatca agctagacct 18300

gcctgatccc tacattctaa ctggagtgac caacaggaca cgggagagaa aaccacatac 18360

aaaaccaatc cacagaaccc cgccatgccc aggtccccac agagagggga gggggcgttt 18420

tctccacttt tttttctcgg cctgtgagcc cctgagccgt gcactgcggt cccacaggtt 18480

gaccccgcgt tgccagcctc agggccaagg tcacgtttcc agacatggcc ctaagaaaaa 18540

ggccagccca ggggaaggga catggtcagg gcacacagga accacgtgca taccacacat 18600

ccatgcccat gagcacacac cacacaccac atgtgtatgc acataccata cacacgtgca 18660

caaatgcatg tatacacact atagtcacac catgcataca tctctaccca cacatgcaga 18720

atcatgtaca ggtcatacac aacacacacg catgtatgca catgccatat acacaccaca 18780

tgtaccatgc atacaccata cacacatgtg cgcaaatgca tgtacacaca ccatcacacc 18840

actcatacat ctctactcac acatgcagaa tcatgtacag gtcatacaca acacatacac 18900

gtgcacacat gccatataca caccacgtgt acacacatgc accatgcata caccacacac 18960

acacgtgcac aaatgtatgt acacacacca tagtcacacc actcatagat ctctacccac 19020

acatgcacca tcatgtacag gtcacacaca acacatacac atgtgcacac atgccacata 19080

ccacgtgtac acacatgcac catgcataca ccatacacat gtgcacaaat gtatgtacac 19140

acaccatagt cacaccatgc atcatctcta cccacacatg cagaattgtg tataggtcat 19200

acacaacaca tacgcatgta tgcacatgcc acatacacac cacatgctcc atgcatacac 19260

acacaaatgt atgtacgcac catagtcaca ctacacatac atctctaccc acacatgcac 19320

catcatgtac aggtcataca caacacatac acatgtgcac acatgccata taccacatgt 19380

gtacacacat gcaccatgca tacaccatac acacgtgcat ctacacatac agatacaggc 19440

acacacacca ccatatacat cacatacaaa gacatccact gatatacgca cacctacata 19500

catgtacaca cagcacacat gcacatacat cacaacacac atgcgcacca cacaccatgt 19560

-continued

gcacacccat acacccatgc tacatgcaca ccatacccca cacatatgca tacactgacc 19620

acagcacaca tcaaccacac acccgccaac ataactcttt ctctttttg gggagataga 19680

gtcttgctct gttgcccagg ctggagtgca gtggcgtgat ctcagcttac tgcaacctct 19740

gccccctggg attcaagcga ttctcctgcc tcagcctcct gagtagcagg aattacaggt 19800

gtgtgccacc atgcccggct aatttttgta tttttagtag agatgaggtt tcaccatgtt 19860

ggccaggctg gtcttgaact cctgacctca ggtgatccgc ctgcctcagc cttccaaagt 19920

gctggcatta caggcatgag ccaccgtgcc tgccttttt ttttttctttt ttttttgtgt 19980

gtgtgtgtga gacagtcttg ctctgtcacc caggctggag tacagtggca caatcttcgc 20040

tcactgcaac ctcggcctcc caggttcaag tgattctcgt gcctcagcct cccgagtagc 20100

tggaattaca ggtgcaggcc accatgcctg gctaagtttt gtatttttag tagagactgg 20160

gtttcgccat gttggccagg ctggtctcga actcctggcc tcaagtgatc cacccacctc 20220

ggcctcccaa agggctggga tgacaggcat cagccaccac ccccagctcc aacaactttt 20280

ttttttttt ttttttttt ttttttttt tttttaagat ggagtctcac tctgtcaccc 20340

aggctggagt gcagtggcgc gatcttggct cactgcaacc tccgcctccc agattcaacc 20400

gattctcctg ccacggcctc ccgagtagct gggattacag gcatgcgcca ccaccccggc 20460

taatttttt gtttttgtag agacggggtt tctccatgtt ggtcaggctg gtctcaaatt 20520

cctgacctca ggagacctgc ccgcctcggc ctcccaaagt gctgggatta caggcatcag 20580

ttactgcgcc tggcctcaaa atacttcaga gcaatggctc tggactgtca gggctggaaa 20640

ggaccctcct ctcacatgct gagctccttg gagtggaggc caaagtccca cagcagaggc 20700

cgactggccc acccagcatg cccatctccc aggagtgcag agaggcaagc cctgggcagt 20760

ggcaggcaca ggccttcttg accccagacc ttcagcctgt catattttgg ctcctcctta 20820

gtgaaggttg ttgagggtgt tttgcagaga gacatgacgc caatcttaat ttttgacaat 20880

tttccatagc atgcagataa tttgtttcca aaacttttca ttttcctgaa gtcatcttga 20940

ttggtatcag ctatttccat aaaacgatcg gatgagtttt gatggacaga tcaggctttt 21000

gtttacaact gttttgctcc taatcattcc accacatcac atgtcatgga cctgaattgc 21060

gtcaagaaga cgggcttgtc tgtcaggccc tggtgggcac tttgatagcg ggcatgctgt 21120

gccatgacac gtgtggtgtt gggtcttgct ggacaagctg tgctgtgttc agtgtgcgga 21180

gcctctgcta gatgctctca tttggggcac tgggccagtg ctactgggag cacttctgtt 21240

ttgtgtcact gacatccaat agcatcgtta tgtagagcaa acaccgaagg gctgcatttc 21300

tttgtgggct tattctcgag aaaactgggg gcagatccct cctcaaggag gggagggcca 21360

ccttggtttc cagtcaagta ttgtgaaaat tatccaacac tcaggcaatc cacccaaccc 21420

tgctgcccat gtctggagaa gcaaagtgtc aggggtagtc caggcccacc tggagacagg 21480

tcaggccctg cagagaaagg tctgacagac gggggtgagg gaagaccccc caaaggcctc 21540

cagagtccca ccaggtctcc aggtccttgt cataaccaga gaggccccag ccccagagga 21600

ccaggtcccc tgctccactg tccacagggg cccacctgca agcacactgg cagagctcaa 21660

gaccacacat gcctgcaagg tgaggcctgt ctgggctctg tcctcctgca ggccccaggc 21720

ctgggtggct gggcgaaggc agctgcttat gcagactcca gggggaaagc cgcctctcat 21780

ctctggccgt ccccaggacg ctggatccac caatatctca ccaacctgga gagccactca 21840

-continued

```
acccctcatt tcacatgttt gaacatagag gaccagaggg gtgtggcctg tctaggaggt  21900
cttaggagct cgggtcctga ctctgccact taccaactct tgtgtgtccc atgtgcctcc  21960
gcttcccctc gggacacaga gatattgtga aagttaaaca acataatccc cgtaaaacac  22020
ttcgagcagt gcctggtatc tggccagcaa gtgatcaatg gtgatccatt accatcctgg  22080
gaccccatca gagccttctg aggtggaggg aagggcgtgc tggggagcac aggtgcaggt  22140
cacaagaagg aagtcagtcc cataagccag gtatctaacc ccatccctgc tcccccaagg  22200
taagggccag catctaaccc caccctgct ccccaaggt aagggccagc atctaacccc  22260
atccctgctc ccccaaggta agggggccag catctaaccc catccctgct cccccaaggt  22320
aagagccagc aacggaggcc tgggaggctc ctgggttctg ggccgcagcg cctctgcgag  22380
gtctgcaggc ttcgctctag gaggggatgg gggctgggca ggtccctgct ccagaggagg  22440
aggacctggg cctgcggagc gccgcggtgg gagtgctgga gtcctggccc gtcatccccg  22500
tctgccccac agcgaggacg atgctgccac tgtataccgc gcagccgcga tgctgaacat  22560
gacgggctcc gggtacgtgt ggctggtcgg cgagcgcgag atctcgggga acgccctgcg  22620
ctacgcccca gacggtgagt gctgggcctt ggcggggtcc ccgaacgggg aggaccccac  22680
gggctctgag tcgcatgctc gcctaggcat cctcgggctg cagctcatca acggcaagaa  22740
cgagtcggcc cacatcagcg acgccgtggg cgtggtggcc caggccgtgc acgagctcct  22800
cgagaaggag aacatcaccg acccgccgcg gggctgcgtg ggcaacacca acatctggaa  22860
gaccgggccg ctcttcaaga ggtgggcggg gcctccccgg agctgggcgg ggctgctctt  22920
ggggaggtgg gcggggtcac tccagagatg ggcggggccg ctcttgggga ggtgggcggg  22980
gccactctcc agagctgggc ggagcagctc tcaggactag gcggggccgc tcttagggag  23040
ctgggggagc gctcctcaag agatgggtgg gggcactctc ggggaggtgg gcggggtcgc  23100
ttccaggagg tgggcggcgt cgctctcagg ggtactgcag tggagcctgc tgccaacatc  23160
ctctggacac tgttacttct ctcctctccc cccacacccc cagcaccacc acatctaatg  23220
gcacaatcat ctgccctctt ctcaacactg acaccagtac ctgggccgtc actggagtgg  23280
ggactggctc cactgcctcc gcccctactt tccacactgc agcccaccct gaaacagcac  23340
ccctctccct gtgtggctgg cagcctttgg gaggaggctc ttgatgcaga tggggactga  23400
aagcttccag ggacccagga ggccagacaa gcagcccaag aacagcacac gagccttaga  23460
cagccagggt tggccaaggc ccagagaccc aagtgaacat ctgcagtgtg gcaggagtta  23520
gctcacagca cgcctggaca ccatgccatg ccagctcacc cccagatccc caaccactga  23580
gtagcacgtg cagagccacc atccacaacg cccacataag tgcagatgta ggcagcacgt  23640
gtgcacacac acgacacaca tacacagaac catgtgtgca cacagactca ggcacatgac  23700
acacatgtga cacaagcaca tgcatggggt gcaccccaca tagggatcac gtgtgcacac  23760
aggttcactg atgtggcccg atgggtacaa tgcacacgtg cacacacagc cggacaggac  23820
agcctggtgg ttagagctgg ctcaacctcg ccttcacttg ctggcaagag ggcaggcatc  23880
cttctgagct tctgcctccg tctctgtaag gcaggatggt tctgaggaca acgtcctaac  23940
ccacagaaag ccgggtctgg cacccataaa ccactcagct gtcatgaacc acaccgtcca  24000
tctggtgcag gcagatacca cggtgtccag ggtctggcgt ctgctgatct ttccgttctt  24060
gggactggga caagggaaaa gcccaagctg ctcagccggc aggagaagga gcagggagaa  24120
```

-continued

ggagcagggg gaaggagcag ggagaagcag caggggggaag gagcagggag gagcaggaga 24180

aggagcatct ctgagaagcc tcagctatgc ttccttccct agagtgctga tgtcttccaa 24240

gtatgcggat ggggtgactg gtcgcgtgga gttcaatgag gatggggacc ggaagttcgc 24300

caactacagc atcatgaacc tgcagaaccg caagctggtg caagtgggca tctacaatgg 24360

cacccacgta ggtgggggtc atgaggggggt gggggctggg gccttagggt cctggggcca 24420

agacccctgc gtggccaccc tccatctcat actcccaccc ccaggtcatc cctaatgaca 24480

ggaagatcat ctggccaggc ggagagacag agaagcctcg agggtaccag atgtccacca 24540

gactgaaggt gggggcccca cagacctccc tcagtgtccc caccccagac agcccatcca 24600

ccccctctgg cctgaaggag gagggtgcgg tgaggtcaat gaaagccact aaaggaagtg 24660

ggggtggggc ctgcttcccc tggacaccgt ccagcacacc tggcacagca caggaagcag 24720

agagaacagg agggaggaga ggaagctgcc cccatcccac agggggtctc cagtgcccct 24780

cttgacccag ccctacttaa gtctggggca gttagttgtc tgacaggacc ctgctgggga 24840

agagcagatg ggggacagca ggcagacctc agcttcagca ctcgctgtcc ccagtcctgg 24900

tcctccacac ccctcatccc tcctccagcc tgcattgctc ttgatgggac cgggtcaaac 24960

tgtcctcttc caccgtgtgg gacagcccttt cctgactccc ctgggcctct gagagcctct 25020

gccctcgccg gcttcctcct ccagaacatc tttcccttgg ctccctactc cagggtgctc 25080

tcctggccat tcctccccgg gcagagccac actaccccca ctccacacac actccagtcc 25140

tggtagcatc acagaccacc aaaggcaagg acctcacagg cgacacgccc accaaccttc 25200

tctcggtcat tccaagccct caaatgtctc ttgaccctgt ctgttttctg agcccacccc 25260

tgaagcttgg tgtcagcccc tgtgacctct cacccaggct ccctcccctg ctctgcaccg 25320

gcccctgtgg cctctcaccc aagctccctt ccctgctctg cagacagggt ggggttttcc 25380

agtgccagtg tgggtttcat tgcagcccag acacctcaca ctgaaaagtc tgaaagcagc 25440

ggtcaaacgc taatggccaa aaggccccat ctaggctgtg agatggagat ggctttttac 25500

aaatttgttt ctggcctcac taatttttta aaaatactag catatatatt acctgtataa 25560

caggaaaatg taatgaaagt tttataaagc aaatcaactt cttcaatggc tcctgattcc 25620

cctggggata aaagacaaaa tgctgcctgg aggctgaggg tgggcgggcc tgccccctc 25680

tcaggctcac cctgcctagg acatgccggg agggtgcctc tcccaccacc cccacgcctc 25740

cctgcctttg cagttctgga ctgcagactc ctctgctcca cctgccctca ggcacctgct 25800

tgatccctgc ccaccttgag ggctcagctc tgacaccatc tcccctcaca gttctggcag 25860

tgtgctatgc tctattccag ccccctgtgc cccagatccc ttccccaccc ccatgccatg 25920

gtcccttgaa ggacagacag gagggcgagc ccaagcagga gtgtgggtcg aagaggccac 25980

ggcgcggtgg agcacgtaca cacgggcaag agaaaggagc cagagaccta cattcaaagc 26040

ctgagggctt cgggactggg ggccgggaca ggcagtgcgc cgggatgaag ggaggcacgg 26100

gtgggtggcc ccacgggtcc caggtcctgt gcaggtgcag ggtcggcttt gtggacatgc 26160

ccctgtcctc gtggcacagc agggtggggg tcagcctgca ggctgggctg tttctcaccc 26220

caggaagatg cctggcatac acgggacatc agcggctcct ctgctggagg gaatcatgtc 26280

ttttttttttt tgagacagag tctcgctctg tcgcccaggc tggagtgcag tggcgcggtc 26340

tccgctcact gcaagctccg cctcccgggt tcacgccatt ctcctgcctc agcctcctga 26400

-continued

gtagctggga ctacaggtgc ccaccaccac gcccggctaa ttttttttgta ttttcagtag 26460

agacggggtt tcaccgtgtt agccaggatg gtctcgatct cctgacctcg tgatccggcc 26520

gcctcggcct cccaaagtgc tgggattaca ggcgtgagcc acagcacctg gcggggaatc 26580

atgtctaagc caagactgga gaaaaagtgg ccaagagaag ggtccagctc tccaggagtc 26640

ttttctgagc ccccagcccc caccccccg gggctgcagg cagacgatgc tgacggtggc 26700

tggggaggac gtgtcctgaa cacttgggct cgtgaagaag ctccagagag gggcagtggc 26760

cggcggcgca gggcgggggg tgtgaggggt gcgtcgggga ttaagagggg cgccagggga 26820

ggctgggagc tgagaagaga ctgccgccct gggcagcctt aggtcggtgg tccaggctgg 26880

gtctcccctt cccccccaga ttgtgacgat ccaccaggag cccttcgtgt acgtcaagcc 26940

cacgctgagt gatgggacat gcaaggagga gttcacagtc aacggcgacc cagtcaagaa 27000

ggtgatctgc accgggccca acgacacgtc gccgggcagc cgtgagtgcg cggggcaggg 27060

cgcggggcgc ggggcagggc gcggggcgtg gggcggtctg gagcccagca gttaccgccc 27120

gcacctaccc agcccgccac acggtgcctc agtgttgcta cggcttttgc atcgacctgc 27180

tcatcaagct ggcacggacc atgaacttca cctacgaggt gcacctggtg gcagatggca 27240

agttcggcac acaggagcgg gtaggctgga cggcgggggt ggggaccagc gtgagagggg 27300

cctgcaggcg cggtcggagt gggtggggca tggagtaggc ggggcttgca gatggtgggg 27360

ggtcctgggg tgagtggggc atggagtgag cggagcctgc gggctgggtc ctggcgtggg 27420

taaagcatgg ggtgggcggg gcctgagggc tgggtggggc ctgacatggg aggggcctga 27480

cgtgggggtc ggagtgggtg gggcacggag tgggcagggc ctgcaggcgg gggtctggag 27540

tgggcgggac gtggagtggg cggggcctgc tggctgtggt ggggcccgcc cggcgtggga 27600

ggggtctgcg agccagggcg gggctggagt ggggtggggc ctgcgagctg ggtagggtct 27660

tggggagaag acccccggag tgctctaggg cggcttcagt cgggggtacc tgtggcggga 27720

gctgggagga cgctgcctgc atgcccgccg gctctgtcgc ctcgcaggtg aacaacagca 27780

acaagaagga gtggaatggg atgatgggcg agctgctcag cgggcaggca gacatgatcg 27840

tggcgccgct aaccataaac aacgagcgcg cgcagtacat cgagttttcc aagcccttca 27900

agtaccaggg cctgactatt ctggtcaaga aggtgggcag gggccgggtg gcggggtggc 27960

ggcgggggga gtccctggag ggcccgggcc gcgctgacct cgcgtccctc cgcaggagat 28020

tccccggagc acgctggact cgttcatgca gccgttccag agcacactgt ggctgctggt 28080

ggggctgtcg gtgcacgtgg tggccgtgat gctgtacctg ctggaccgct tcaggtgagc 28140

gcgacccggg gctcagacac ctccatctgc ggggcgcgga gccggccagg ggcggggcag 28200

ggccgcctct cccgccctct ctcccgcccg ccctctgcgc cccgcagccc cttcggccgg 28260

ttcaaggtga acagcgagga ggaggaggag gacgcactga ccctgtcctc ggccatgtgg 28320

ttctcctggg gcgtcctgct caactccggc atcggggaag gtaaggcccc gcccggcccg 28380

cctggtcccg cctcggccct ctagggtctg acagagcccc ccgcccgccc acaggcgccc 28440

ccagaagctt ctcagcgcgc atcctgggca tggtgtgggc cggctttgcc atgatcatcg 28500

tggcctccta caccgccaac ctggcggcct tcctggtgct ggaccggccg gaggagcgca 28560

tcacgggcat caacgaccct cgggtgaggc ctggccgggc tgggggaggg aatgcgaggt 28620

gagctggggt cggcctcggt taggggcctg gggagccgcc gccgcgatcc ctgccctccg 28680

-continued

accctgcagc tgaggaaccc ctcggacaag tttatctacg ccacggtgaa gcagagctcc 28740

gtggatatct acttccggcg ccaggtggag ctgagcacca tgtaccggca tatggagaag 28800

cacaactacg agagtgcggc ggaggccatc caggccgtga gagacaagtg aggcgcgggc 28860

ggccaccctg gcggggcggg acaggtgcgg ggaggggggag ggtggcctcc accgggcagg 28920

agagcgtccg ggccgggcac cccggagggc gcgggcgtgg ggcttccagg ctggcaggac 28980

caaggccccc gtgactccgc ctctgccggc agcaagctgc atgccttcat ctgggactcg 29040

gcggtgctgg agttcgaggc ctcgcagaag tgcgacctgg tgacgactgg agagctgttt 29100

ttccgctcgg gcttcggcat aggcatgcgc aaagacagcc cctggaagca gaacgtctcc 29160

ctgtccatcc tcaagtgagt gtccgtgcgc ccgcgtccct cctccgcccc tctccgccag 29220

aggtggacgc cctccccagt gccagaccac tccgaggcca ccactgattt cccacccagg 29280

ccgggcgctg cccactccac gccgcaccct accccgcagg ccccgccccg gccccgcccc 29340

cagcttgctc cttcccgtcc tgggccccgc ctcactgcag gctcacttgt tcccaccgcc 29400

aggtcccacg agaatggctt catggaagac ctggacaaga cgtgggttcg gtatcaggaa 29460

tgtgactcgc gcagcaacgc ccctgcgacc cttacttttg agaacatggc cggtgcgttc 29520

tccttcatcc attctcgggt gggttctccg tgggctgcgg cctccctggc cagcaactga 29580

ggctctgggt cccggcacac aggggtcttc atgctggtag ctgggggcat cgtggccggg 29640

atcttcctga ttttcatcga gattgcctac aagcggcaca aggatgctcg ccggaagcag 29700

atgcagctgg cctttgccgc cgttaacgtg tggcggaaga acctgcaggt agggcaggcc 29760

accctccgag gcctggtgcc cagggcccgg cctggccacg gccctcctcc atccccgaag 29820

gccgtggcac tggctctggc tctggtgggc aggactggag ctaggagcca tggccagggg 29880

cagtggtgag tgctcccagg gcacgggggc agcaccggtg gggggctgcc tgcaggtggc 29940

tgcccactgc aaagccgggg ccgagggagg ccacgcaccc tgctccaagc ctccgcctgg 30000

cccctctgtc tccagagtcg cccgccggta cccattccat aggaaggcaa tcaggcaggg 30060

taagacaggg gcccgcctgt gtatggcacg tgagtccaag atgcattttg ccctccgccg 30120

acccaagccc cttgacaccc ttcggagacc ccccccttтc ctgctatgtc cttgtgctcc 30180

gtgactctaa tccgaattgg gccaggtccg gtcctgcctg gtgcccaggt tgtatccatg 30240

agaatttgcc accagcaagg gcagccacgg cccacctggg acagggtggg cagtgggcct 30300

gtacaggcct aagggctcgt ggcccgcggt cgagttccgg ttcactccgt ctcttctctt 30360

tctctgggtg ccgtcctgga gcctgtgtcc tgagatgaag ccgacagtgc ggccagggct 30420

gctgggggat gggggttgct ggaggctcca cacctctcat ccgcccgctc ttgctcttgg 30480

cccccacagg tcccctgggg acctggccgc tgccagcact ggcgggcaca ggccacctgg 30540

ccatcagacc tgaggccaga gtcccgggag ctgcctctgt cactccaatt ccacctcgac 30600

acctgcctcc agccctcggc cccttcctga atcttggtgt gtgccccttg ggggtcagtg 30660

gcctccacgc agacagctgg tgtggcctga ggggcaactc ctccagtcct cagaggactc 30720

ctcctcctcg ggacgcctgt aagccagggc cacccaggag ccagggagcc aggcggacct 30780

cccaggaaga gccagccgag agcccccaag cccagcccca gcacgagcaa ggtcaggccc 30840

gagaccccgg gcaggagaag aggccaccct cgaacgtccg ctgtcggccc gtctgtccag 30900

cacagggagg caggcaggag cgagggccca agtggccggc caggctgggc agcggcccat 30960

-continued

gcaggagcag gcgagggcag gtgtggccac caccctagcc atctaatcac ttatacatat 31020

tcattttagg atagaaagag tggtagagca gagcctgacc ctaaaaagaa agccacattt 31080

agggctatca cctccaccct ggcttccagc ttcaagaggc gtaggtcctc caaagacacg 31140

gtaaggggga gagcacccca gtcccgcgtc cgactccacc tgccctgccc tgcgtgtgtc 31200

tcccgcccca tcaccccgcc ccggaccctg ggctcctgtg gcccactctg cccctgtctc 31260

cctgtggcgg ccgctctgcc cagcccgccc atgctgctct ctctcactct ctggaccttt 31320

ctccccggcc ctcctgggtc ctcggctttc cccgtgtgtc tccgttagtc tgcccgccca 31380

cctcccctgc catgacccac acgccatctt gaagcctgtc atctcgttgg tcagtcagtc 31440

agccacacca cctctcgggg ccaggtctgg ggccctggga gcccagcgtg gccccatcct 31500

ggactcctca gctgccggga ggccacacca cttctctgtt atgtccccgt ttctctcgcc 31560

tctcccagag gggcccgccg ccctcacttc gcccctgcga cggccctgga gggggtggct 31620

gtgatgtccc atcccgtccg tctgtctggc cactggcccc gccccccaga cacctgtctc 31680

acctgtctca ccagagccat gcgtgttgca tcttcatgtg gtctctgtgt gggccggggg 31740

ctgggggccg ggcctgggtc cgtctgggtg gacggctggg gcctggagtt ggaactggcc 31800

ccggccacag gggactgtca ggcagggagt ggggtgggac caaaaggggt ggctcccacc 31860

ccaggctgag cgggggccct gcaggaggtg tggcggcagc tcccagaggg tctgagaatg 31920

ggtaggggcg gccccacaag ccctggcctg cagagcccag gacgacactg aggttcccag 31980

acagggaggc ctctggaagg gaaacgacca cctcagctcc tgaccccagc aaccccacaa 32040

ggcccacccc aaagagccag gccttctgcc ctttggagcc cagaatcccc cacctcctgc 32100

tcggggcagc ttgtccctgt agcggatatg cacactcgga ccagaggccc ccagagcgaa 32160

cccagccttg ctagaggcac cccaggccca ggcaccatgg tggggagggg ctgcccagag 32220

aggcagcgga gacctcagcc ccgtggccac cctgcagtcc agggaccagt ctggcccaca 32280

ggaagccccc agcccataag cagcatcacc agagagaagc ttacgcccgg gggaggaagt 32340

gcgatttgca gccacctgcc cctcagtgca ctggaagcgg ggcagacctc cagggcacag 32400

acaggacttg gcatcaagca agccaaatcc cgagatgaag ccaccagggt gccccaagag 32460

ggacccatga ggcctggctg ctcagcttcc tggggaaggg acttggcatg caggatgggt 32520

ggacagtgag agcctgtagg cctgggggcc actggaggct caaggagcag gtggaagcac 32580

cattcctgga gccacctctg ctgcggaaag cgggcagagc tgatgctgca aagtctgagc 32640

caggagtccc gcagggaaca gggaggggga atagcgcagg gatcgtgggc tgggcaggct 32700

ggggaagagg gggtgtccag gcagacagga gaaacagcga tttggggcag gcagccacgg 32760

ggggcaagca caaatgtcgt gcaggtgatg ggccactttc agagggtgac actgggtccc 32820

agggccctgc ctggagcgag gccaggtgca gctcagagac cctcatggtg ccctcccagg 32880

gacatgttcc cagcggaacc ctcacccgag cctctctggg caccagggac cgtcctctgg 32940

ggccagttct ggcatcacgt ggcatctggg gctggccccg ccctgcaagg ctgaactgtg 33000

gggggcactg ccagctgggg gtctgggcag gggagggcag cccagctccc acctggtctc 33060

tggggctgcg agcttattca gagggaggcg tgggtggggg gctcctttgg gtagggtggg 33120

gtcagtccgg ctgcggagat cccctgcccc tgtcctgtgg ccggtccggg ccagggcggc 33180

actgggcgct gagggctggg gtccctggcg gccggcgggg ccagcgggta ttgattgttg 33240

-continued

gttcttattt atagagcacc gggggtggac gcggcgcttt gcaaaaccaa aaagacacag 33300

tgctgccgcg acgcgctatt gagagggagg agggccagct gcagctgtgt tcccgtcata 33360

gggagagctg agactccccg cccgccctcc tctgcccccc cccccgcaga cagacagaca 33420

gacggacggg acagcggccc ggcccacgca gagccccgga gcaccacggg gtcgggggag 33480

gagcaccccc agcctccccc aggctgcgcc tgcccgcccg ccggttggcc ggctggccgg 33540

tccaccccgt cccggccccg cgcgtgcccc cagcgtgggg ctaacgggcg ccttgtctgt 33600

gtatttctat tttgcagcag taccatccca ctgatatcac gggcccgctc aacctctcag 33660

atccctcggt cagcaccgtg gtgtgaggcc cccggaggcg cccacctgcc cagttagccc 33720

ggccaaggac actgatgggt cctgctgctc gggaaggcct gagggaagcc cacccgcccc 33780

agagactgcc caccctgggc ctcccgtccg tccgcccgcc caccccgctg cctggcgggc 33840

agcccctgct ggaccaaggt gcggaccgga gcggctgagg acggggcaga gctgagtcgg 33900

ctgggcaggg ccgcagggcg ctccggcaga ggcagggccc tggggtctct gagcagtggg 33960

gagcgggggc taactggccc caggcggagg ggcttggagc agagacggca gccccatcct 34020

tcccgcagca ccagcctgag ccacagtggg gcccatggcc ccagctggct gggtcgcccc 34080

tcctcgggcg cctgcgctcc tctgcagcct gagctccacc ctcccctctt cttgcggcac 34140

cgcccaccca caccccgtct gccccttgac cccacacgcc ggggctggcc ctgccctccc 34200

ccacggccgt ccctgacttc ccagctggca gcgcctcccg ccgcctcggg ccgcctcctc 34260

cagactcgag agggctgagc ccctcctctc ctcgtccggc ctgcagccca gaacgggcct 34320

ccccgggggt ccccggacgc tggctcggga ctgtcttcaa ccctgccctg caccttgggc 34380

acgggagagc gccacccgcc cgcccccgcc ctcgctccgg gtgcgtgacc ggcccgccac 34440

cttgtacaga accagcactc ccagggcccg agcgcgtgcc ttccccgtgc ggcccgtgcg 34500

cagccgcgct ctgcccctcc gtccccaggg tgcaggcgcg caccgcccaa cccccacctc 34560

ccggtgtatg cagtggtgat gcctaaagga atgtcacgca gttttcggtc tgtgtcgctt 34620

gttgacgccg gcagacagtg taaagggagg gcaaaggcat gggggaagct tcgagcgctc 34680

caggcggccg cggccgctca ggcttgggcg gcagcggcgg ggctccccgg gtccgcgggc 34740

gaggcacagc cgtgggggtc gggatcgggg ttcgggtctg gcggtctcgg cgggcggagg 34800

gcggcggtgc ggaggcggcg gcggcgcgca cggcaggcgg tgagcccaga gcccagcgcc 34860

aggcaggaag ccaggctgac gaggaaggag gccggcccga gcgtgtaaac cacggccagg 34920

tcccgcaggg cgagcggctg cgcgcaatgg ctaaaggcgg cgtcggaaaa ggcagtcagg 34980

gggctgagcg tcaggcgtcc cggccacacg cagagcaccg tctcggcctc tgtgggacac 35040

agacaaaggc gcggcgtcag gtggcggctg ccgcaccgcc ctgcagaccc cgacccgcgt 35100

ccccagcagc tcacctgacg cgggcagcgg gtgccggcgc agccaggcgc agagcgggcg 35160

cagcgcgcac ccgcagcccc aagggttgcc gcgcaggtgc agcgcgtcta gagcgggcag 35220

gcggcccagc agccccggcg cgagtgccgc cagctcgttg tcctgcaggc tgagtgagcg 35280

cagcagcggg agcgcgccta gcgccgcggg ctccaggcgc gccagccggt tgccggccaa 35340

tgagaggttg cgcagcgcgc gcagcggcgc gaaagtccct ggtgccagtg cttccagctg 35400

gttggcgctc aggtccagca gctgcagcgc gcccaggccc cagaaggctc gcacatgcac 35460

cgagtgcagc ccgttctcgc gcaggtccag gcgctgtagc gcgcccgctc ccgcgaaggc 35520

-continued

acctggcggc agcgcacgga cgcggttgtg gtccagcagc agcgcgcgca ggcgcaggct 35580

caggcccggg ggcacggcgg gcagcgagag tgccgagcag ctggccaggc ctcccggcac 35640

gcacgtgcac acctcggggc agtccggggc gcccagggac cccgagggcg aggccgtggc 35700

cgacacctgg gcccagacag gccaaggcga cagcagcagc aacagcagca gcagcggccg 35760

aggccgcgac caggaagggc cccgcatggg ggcagccccc ccgcccccgg cacccgcggt 35820

gggaggcccg ctgcctgtgc gtccctggag cccgtcgtcc gcggagcccg ttcctgcggc 35880

gcctgcacaa atattaactc tctggcccga gctcaggcag ttcctgtccc acggctggat 35940

ccacgcttgg ggcgggggca gggcaaacag ccaacgcccg gcaccgccag ccacctgtcc 36000

gggagctcca aactactctt ggctcagcgc cggccacagc gctatcaaga ccacctcacc 36060

ccgccttgtc caccaccggg cgcccgccga gccctgacct gagcagcagg acaccgccca 36120

cctgcacccg cccagctcac ggctgcaaca ggcgcccaca cgcaatgcaa cccagaggtc 36180

ctgctcctca cccgtctcga ccccacccag gctccgcaaa gtgatcccaa cgtgcacatg 36240

cggagtggcc cccacgcagg gatggtccca ttcccatgct ggatagggcc tgggttggag 36300

acaggccttg gggtggaggg agacagcaca cccgaggcag gagaggcgtg cctgccccac 36360

ccctcccccc caggactctc ctgggataga ggtacagacc agtgcagtgg ggcaggggta 36420

tcagcccaag tcctgctgtt aaacccagcg cccttcacag ttgccagttg caggtcctgt 36480

tctaggggct ttccaaagct gggtgcagga aggagccggg ctgaccagct gtggcagaga 36540

aggtggaaac attaggggta tggcttactg ccaggggggca gaggaacagg ggaacttgct 36600

<210> SEQ ID NO 19
<211> LENGTH: 1859
<212> TYPE: DNA
<213> ORGANISM: Unknown
<220> FEATURE:
<223> OTHER INFORMATION: Description of Unknown: Phosphatidylserine Synthase polynucleotide

<400> SEQUENCE: 19

gaggtattgt ggccagtatt ttggttttct tatgttttgg agtcacacaa gctaaagacg 60

ggccattttc cagacctcat ccagcttact ggcggttttg gctgtgtgtt agtgtggtct 120

acgaattgtt tctcatcttc atccttttcc agacagtcca ggatggccga cagtttctga 180

agtatgtgga tcccaggctg ggagtcccat tgccagagag ggactacggg ggcaactgcc 240

tcatctatga tgctgacaac aagactgacc ctttccacaa catctgggac aagctggatg 300

gctttgttcc tgcacacttc attggctggt atctgaagac gctcatgatc cgtgactggt 360

ggatgtgcat gatcatcagt gtgatgttcg agttcctgga gtacagcctg gagcaccagc 420

tgcccaactt cagcgagtgc tggtgggacc attggatcat ggacgtcctc gtctgcaacg 480

ggctgggcat ctactgtggc atgaagaccc tcgagtggct gtccctgaag acatataagt 540

ggcagggcct ctggaacatt ccaacctaca agggcaagat gaagaggatt gcctttcagt 600

tcacgcctta cagctgggta cgctttgagt ggaagccagc ctccagcctg caccgctggc 660

tggccgtgtg tggcatcatc ctggtgttcc tgctggcaga gctgaacacc ttctacctga 720

agtttgtgct atggatgccc cctgaacact acctggtcct tctgaggctg gtcttcttcg 780

tgaacgtggg tggtgtggcc atgcgtgaga tctacgactt catggatgaa ttgaagcccc 840

acaggaagct gggccagcag gcctggctgg tggcagccat cacagtcaca gagcttctca 900

-continued

```
tcgtggtgaa gtatgacccg cacacactca ccctgtcact gcccttctac atctcccagt    960

gctggactct tggctccatc ctggtgctta catggactgt ctggcgcttc ttcctgcggg   1020

acatcaccat gaggtacaag gagacccggc gacagaagca gcagagtcac caggccagag   1080

ccgtcaacaa ccgggatggg caccctgggc cagatgatga cctgctaggg actggaactg   1140

cagaagaaga ggggaccacc aatgacggtg tgactgctga ggaggggacc tcagccgcgt   1200

catgagcctc accctcgtca ctggccttgt gccaaggggc tgtcccattt ctcccttcct   1260

cgtgctccct gcttcagagg caggggtggg gggggcatcc acactccagg aggggcacct   1320

gagaatacat gtttgtgtgc aggtaggtgt atgcacatat tggctcctga cactactttg   1380

gggccatgag ctgaacagtg agcctggaac ttgctctaga gcaagagtcc tttctacctg   1440

gtcaacaaag gccctggtcc atggtctccc ttgtgctcaa tctcagcacc ggctagggga   1500

ggtatcttgg tatctcggta ctctttctcc actctttatg gggtaggcag aagccccatg   1560

agaccctgtg gtcccaccca cttacagcag cataagtgaa ggatatcata ataccaacat   1620

gtctgcaaag tggtgggtct agagtcagca ctgagccatt tcctttggag ccttccttta   1680

accacgcagg actataaact gaatggtgta tacatgacag tcacagattg gccttttgcg   1740

gccagagagc ccaacttgga gacctgtacc ccaggtgcca gggtgctgtc accatggctc   1800

ccttgagcaa atggaacaaa taaagtgatg atgaaggtga aaaaaaaaaa aaaaaaaaa    1859
```

What is claimed is:

1. A method for detecting the presence or absence of an antigen in a sample comprising:

a) obtaining a sample suspected of having said antigen,

b) detecting the presence or absence of said antigen in said sample utilizing a polypeptide, wherein said polypeptide comprises all or a portion of at least one variable antigen-binding (Vab) domain of camelid and/or shark single-domain heavy chain antibodies lacking light-chains, at least ten contiguous amino acids derived from a source other than camelid and/or shark single-domain heavy chain antibodies lacking light-chains, wherein said polypeptide comprises at least one binding site for an antigen, wherein said polypeptide binds specifically to said antigen, wherein said binding is indicative of the presence of said antigen.

2. The method of claim **1**, wherein said polypeptide comprises at least two variable antigen-binding (Vab) domains of camelid and/or shark single-domain heavy-chain antibody lacking the light chains.

3. The method of claim **2**, wherein said two variable antigen-binding (Vab) domains bind to two different antigens.

4. The method of claim **1**, wherein said polypeptide has three or more variable antigen-binding (Vab) domains of camelid and/or shark single-domain heavy-chain antibody lacking the light chains.

5. The method of claim **1**, wherein said polypeptide has improved cellular uptake, blood brain barrier permeability, biodistribution and retention.

6. The method of claim **1**, wherein said polypeptide is immobilized on a solid support prior to binding to said antigen.

7. The method of claim **1**, wherein said polypeptide binds to said antigen to form a complex, and wherein said complex is immobilized on a solid support.

8. The method of claim **1**, wherein said polypeptide is linked to at least one entity other than an antibody.

9. The method of claim **8**, wherein said entity is selected from a group consisting of solid support, radioisotope, enzyme, detectable label, ligand, fluorophore, biotin, digoxegenin, avidin, streptavidin, Fc region of IgGs, a therapeutic agent, toxin, hormone, peptide, protein, vector, siRNA, micro-RNA and nucleic acid.

10. The method of claim **8**, wherein said solid support is selected from the group consisting of beads, biosensors, nanoparticles, microchannels, microarrays, and microfluidic devices, glass slides, glass chambers, and gold particles.

11. The method of claim **8**, wherein said enzyme is selected from the group consisting of alkaline phosphatase (AP), horse-raddish-peroxidase (HRP), Luciferase, and beta-galactosidase.

12. The method of claim **1**, wherein said polypeptide is selected from the group consisting of structures **5**, **5***a*, **7**, **7***a*, **14**, **14***a*, **15**, **15***a*, **19**, **20**, **31**, **32**, **33**.

13. The method of claim **1**, wherein said antigen is selected from the group consisting of Aβ42, Tau, ABAD (Abeta-binding alcohol dehydrogenase), a mitochondria regulating protein (MRP), Cyclophilin-D (Cyp-D: MRP), TOM (Translocase of Outer Mitochondria Membrane: MRP), hPReP (Human Presequence Protease: MRP), NMDAR (MRP), PtDS (MRP), mSOD1 (MRP), mHTT (MRP), ApoE4 (Demyelination Regulating Protein: dMRP), integrin-α4β1 (dMRP), integrin-α4β7 (dMRP), PPAR-gamma (dMRP), MAdCAM-1 (dMRP), α-synuclein, TDP-43 (TAR-DNA binding protein-43), ubiquitin, APP, ALZAS, gamma secre-

tase, BACE (β-secretase), Apo-A1, Apo-H, PV-1, PEDF, BDNF, Cystatin C, VGF nerve growth factor inducible, APO-E, GSK-3 binding protein, TEM1, PGD2, EGFR, EGFRT790M, Notch-4, ALDH-1, ESR-1, HER-2/neu, P53, RAS, KLKB1, SMAD4, Smad7, TNF-alfa, HPV, tPA, Mucin, Cadherin-2, FcRn alpha chain, TNF-alfa, Thrombin, cytokerratin 1-20, Celuloplasmin, Apo AII, VGF, Vif, LEDGF/p75, TS101, gp120, CCR5, CXCR4, HIV protease, HIV integrase, OST-577, H1N1, CD3, CD11a, CD20, CD33, CD25, CD52, Protein C5, VEGF, alfa-4-integrin, EPCA2, PSMA, PSA, TMPRSS2-ERG, PCA3, HAAH, AMACR, Glycoprotein IIb/IIIA, AP-1, VEGF-A, IgG-E, Bacillus anthracis protein, NadD (Nicotinate Mononucleotide Adenyltransferase, a enzyme implicated in drug-resistant bacteria), Plasmodium falciparum, STDs, TB, cGMP directed phosphodiestrase, chain B of Clostridium botulinum neurotroxin type E protein, *Borrelia* VlsE protein, ACE2 receptor, TTHY, AIAT, AFMN, APOE, SFRS4, SAMP, CD 71, GPA, epsilon- and gamma-glycophorins, TIMP-1, RGIA, EXTL3, biomarkers for: lung cancer, bladder cancer, gastric cancer, brain cancer, breast cancer, prostate cancer, cervical cancer, colorectal cancer, oral cancer, leukemia, childhood neuroblastoma, Non-Hodgkin lymphoma, Alzheimer's disease, Parkinson's disease, and AIDS.

14. A method for diagnosing an individual with a disease, said method comprising:

a) obtaining a sample from said individual

b) detecting the presence or absence of one or more biomarkers associated with said disease, wherein said detection comprises utilizing a polypeptide, wherein said polypeptide comprises all or a portion of at least one variable antigen-binding (Vab) domain of camelid and/or shark single-domain heavy chain antibodies lacking light-chains, at least ten contiguous amino acids derived from a source other than camelid and/or shark single-domain heavy chain antibodies lacking light-chains, wherein said polypeptide binds specifically to at least one of said biomarkers; and said binding of said polypeptide to one or more of said biomarkers is indicative of the presence of said one or more biomarkers in said sample,

c) identifying said individual as having said disease when said one or more biomarkers are present in said individual's sample.

15. The method of claim **14** further comprising determining the amount of said biomarker in said sample and comparing said amount to a reference value, wherein an amount higher than said reference value is indicative of a disease.

16. The method of claim **14**, wherein, the said polypeptide is capable of binding specifically to a biomarker selected from the group consisting of:

biomarkers associated with Alzheimer's Disease, wherein said biomarkers for Alzheimer's disease is selected from the group consisting of Amyloid-beta (Aβ), ALZAS, Tau, Cyclophilinp-D, ABAD, TOM, hPReP, PtDS, PLSCR1, mSOD1, mHTT, integrin-α4β1, integrin-α4β7, PPAR-gamma, MAdCAM-1, NMDAR, integrin-DJ-1, Bax-1, PEDF, HPX, Cystatin-C, Beta-2-Microglobulin, BDNF, Tau-Kinase, gamma-Secretase, beta-Secretase, Apo-E4, and VGF-Peptide;

biomarkers associated with Parkinson's Disease, wherein said biomarkers for Parkinson's disease is selected from the group consisting of Apo-H, Cerulopasmin, Chromogranin-B, VDBP, Apo-E, Apo-AII, and alfa-Synuclein;

biomarkers associated with Brain Cancer, wherein said biomarkers for Brain cancer is selected from the group consisting of TEM1, Plasmalemmal Vesicle (PV-1), Prostaglandin D Synthetase, and (PGD-S);

biomarkers associated with HIV/AIDS, wherein said biomarkers for HIV/AIDS is selected from the group consisting of gp120, Vif, LEDGF/p75, TS101, HIV-Integrase, HIV-Reverse Transcriptase, HIV-Protease, CCR5, and CXCR4;

biomarkers associated with Lung Cancer, wherein said biomarkers for lung cancer is selected from the group consisting of KRAS, Ki67, EGFR, KLKB1, EpCAM, CYFRA21-1, tPA, ProGRP, Neuron-specific Enolase (NSE), and hnRNP;

biomarkers associated with Prostate Cancer, wherein said biomarkers for prostate cancer is selected from the group consisting of AMACR, PCA3, TMPRSS2-ERG, HEPSIN, B7-H3, SSeCKs, EPCA-2, PSMA, BAG-1, PSA, MUC6, hK2, PCA-1, PCNA, RKIP, and c-HGK;

biomarkers associated with Breast Cancer, wherein said biomarkers for breast cancer is selected from the group consisting of EGFR, EGFRT790M, HER-2, Notch-4, ALDH-1, ESR1, SBEM, HSP70, hK-10, MSA, p53, MMP-2, PTEN, Pepsinigen-C, Sigma-S, Topo-11-alfauKPA, BRCA-1, BRCA-2, SCGB2A1, and SCGB1D2;

biomarkers associated with Colorectal Cancer, wherein said biomarkers for colorectal cancer is selected from the group consisting of SMAD4, EGFR, KRAS, p53, TS, MSI-H, REGIA, EXTL3, p1K3CA, VEGF, HAAH, EpCAM, TEM8, TK1, STAT-3, SMAD-7, beta-Catenin, CK20, MMP-1, MMP-2, MMP-7,9,11, and VEGF-D;

biomarkers associated with Ovarian Cancer, wherein said biomarkers for ovarian cancer is selected from the group consisting of CD24, CD34, EpCAM, hK8, 10, 13, CKB, Cathesin B, M-CAM, c-ETS1, and EMMPRIN;

biomarkers associated with Cervical Cancer, wherein said biomarkers for cervical cancer is selected from the group consisting of HPV, CD34, ERCC1, Beta-CF, Id-1, UGF, SCC, p16, p21WAF1, PP-4, and TPS;

biomarkers associated with Bladder Cancer is selected from the group consisting of CK18, CK20, BLCa-1, BLCA-4, CYFRA21-1, TFT, BTA, Survivin, UCA1, UPII, FAS, and DD23;

biomarkers associated with a disease causing bacteria, wherein said bacteria or biomarker associated with disease causing bacteria is selected from the group consisting of *Clostridium Botulinum, Bacillus Anthracis, Salmonella Typhi, Treponema Pallidum, Plasmodinum, Chlamadyia, Borrelia* B, *Staphyloccus Aureus*, Tetanus, Meningococcal Meningitis, and *Mycobacterium tuberculosis*, and Nicitinate Mononuceleotide adenyltransferase (NadD);

biomarkers associated with a disease causing virus, wherein said virus or biomarker associated with a disease causing virus is selected from the group consisting of Pandemic Flu Virus H1N1 strain, Influenza virus H5N1 strain, Hepatitis B virus (HBV) antigen OSt-577, HBV core antigen HBcAg (HBV), HBV antigen Wnt-1, Hepatitis C Virus (HCV) antigen Wnt-1, and HCV RNA.

17. A method for detecting the presence or absence of circulating cells in a sample comprising

a) obtaining a sample suspected of having circulating cells,

b) detecting the level of one or more antigen associated with said circulating cell in said sample utilizing a polypeptide, wherein said polypeptide comprises all or a portion of at least one variable antigen-binding (Vab) domain of camelid and/or shark single-domain heavy chain antibodies lacking light-chains, at least ten contiguous amino acids derived from a source other than camelid and/or shark single-domain heavy chain antibodies lacking light-chains, wherein said polypeptide binds specifically to said one or more antigens, and wherein said binding of said polypeptide to said antigens is indicative of the presence of circulating cells in said sample.

18. The method of claim **17**, wherein said circulating cells are circulating tumor cells.

19. The method of claim **18**, wherein one or more antigens associated with tumor cell is selected from the group consisting of MUC-1, VCAM-1, EpCAm-1, CD44, CD133, E-Cadherin, VEGF, bFGF, sFASL, CD95, p53, Bcl-2 CyclinD1, Cyclin E, TNF-alfa, TGF-beta1, Her-2, EGFR, IGF-1 and IGF-1R, 1L-2R, Ras, and cMyc.

20. The method of claim **17**, wherein said circulating cells are circulating fetal cells.

21. The method of claim **20**, wherein one or more antigens associated with fetal cell is selected from the group consisting of GPA, CD71, CD133, CD34, CD44, ITCAM, ITGB1 (Integrin beta-1), Trop-1, Trop-2, HLA-G233, and 6B5.

* * * * *

专利名称(译)	使用抗体及其类似物的方法		
公开(公告)号	US20100136584A1	公开(公告)日	2010-06-03
申请号	US12/606818	申请日	2009-10-27
[标]申请(专利权)人(译)	ICB INT		
申请(专利权)人(译)	ICB INTERNATIONAL，INC.		
当前申请(专利权)人(译)	ICB INTERNATIONAL，INC.		
[标]发明人	BHATT RAM S BHATT RISHI		
发明人	BHATT, RAM S. BHATT, RISHI		
IPC分类号	G01N33/574 G01N33/53 G01N33/573		
CPC分类号	C07K2317/20 C07K2317/22 C07K2317/569 G01N2800/387 G01N33/574 G01N33/6857 G01N33/6893 G01N33/56966 Y02A50/57 Y02A50/58 Y02A50/59		
优先权	61/192732 2008-09-22 US 61/197601 2008-10-29 US		
外部链接	Espacenet USPTO		

摘要(译)

缺少轻链的骆驼科动物和鲨鱼单结构域重链抗体及其类似物被公开。公开了使用抗体及其类似物检测抗原的方法。描述了衍生这些抗体和类似物以开发诊断测定的方法。还提供了试剂盒，以及使用这种诊断试验的方法。