



US 20110053167A1

(19) **United States**

(12) **Patent Application Publication**  
**Holvoet et al.**

(10) **Pub. No.: US 2011/0053167 A1**  
(43) **Pub. Date: Mar. 3, 2011**

(54) **GENE SIGNATURES**

(75) Inventors: **Paul Holvoet**, Leuven (BE); **Nora Benhabiles**, Paris (FR); **Benjamine Geeraert**, Kapellen (BE)

(73) Assignee: **Katholieke Universiteit Leuven**, Leuven (BE)

(21) Appl. No.: **12/922,739**

(22) PCT Filed: **Apr. 2, 2009**

(86) PCT No.: **PCT/BE09/00022**

§ 371 (c)(1),  
(2), (4) Date: **Sep. 15, 2010**

**Related U.S. Application Data**

(60) Provisional application No. 61/072,971, filed on Apr. 3, 2008, provisional application No. 61/128,205, filed on May 19, 2008, provisional application No. 61/200,705, filed on Dec. 2, 2008.

(30) **Foreign Application Priority Data**

May 19, 2008 (GB) ..... 0809066.8

**Publication Classification**

(51) **Int. Cl.**  
**C12Q 1/68** (2006.01)  
**G01N 33/53** (2006.01)

(52) **U.S. Cl.** ..... **435/6; 436/501**

(57) **ABSTRACT**

The present invention relates generally to a new cluster of correlating molecules in a tissue or at least one cell of a tissue for instance a cell of a blood tissue, preferably such myeloid cells and of identifying the condition of the genes expression said correlating molecules or of the expression levels of said molecules in a method or system for identifying obesity and the risk at obesity-related metabolic diseases such as the obesity associated cardiovascular risk or obesity-related insulin resistance. This system of method provides information on how to modulate the correlating molecules to treat or prevent obesity and to prevent the obesity-related metabolic diseases.

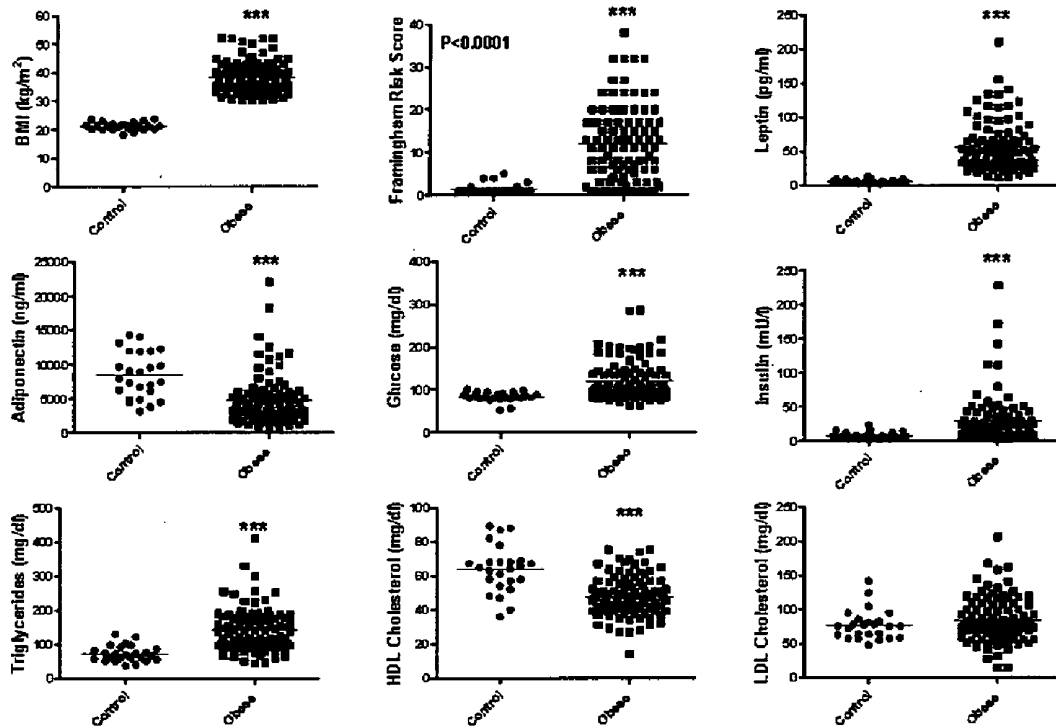


Figure 1

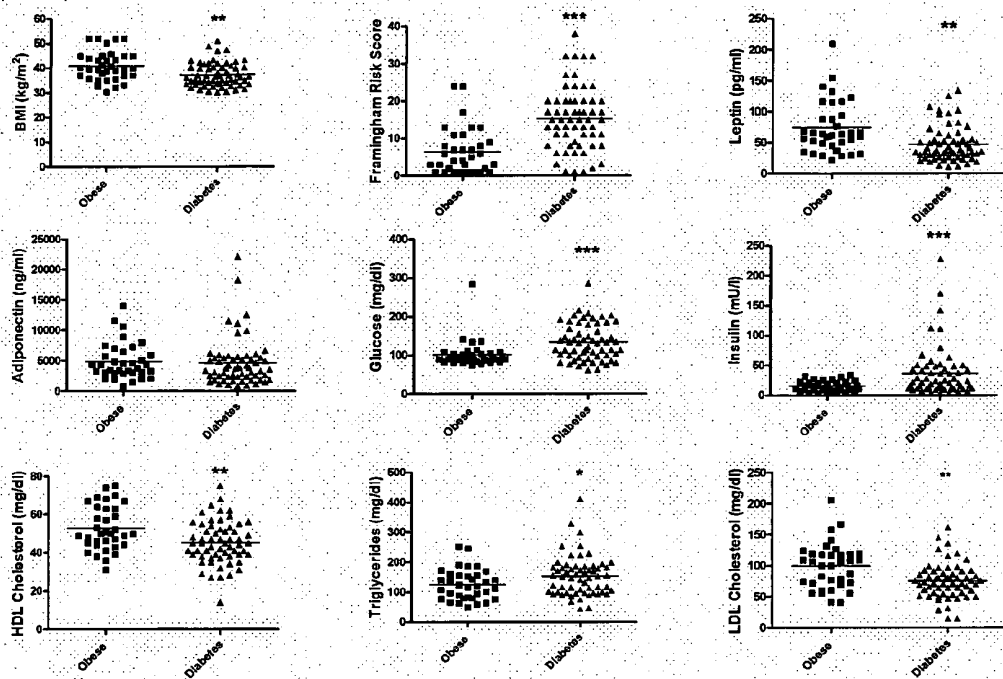


Figure 2

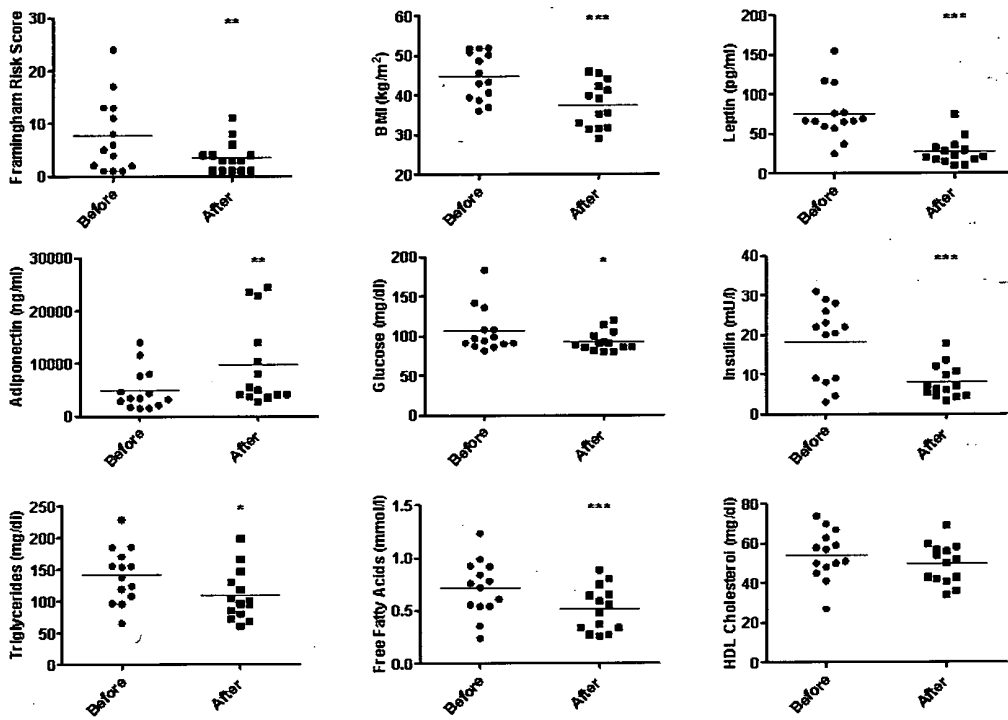


Figure 3

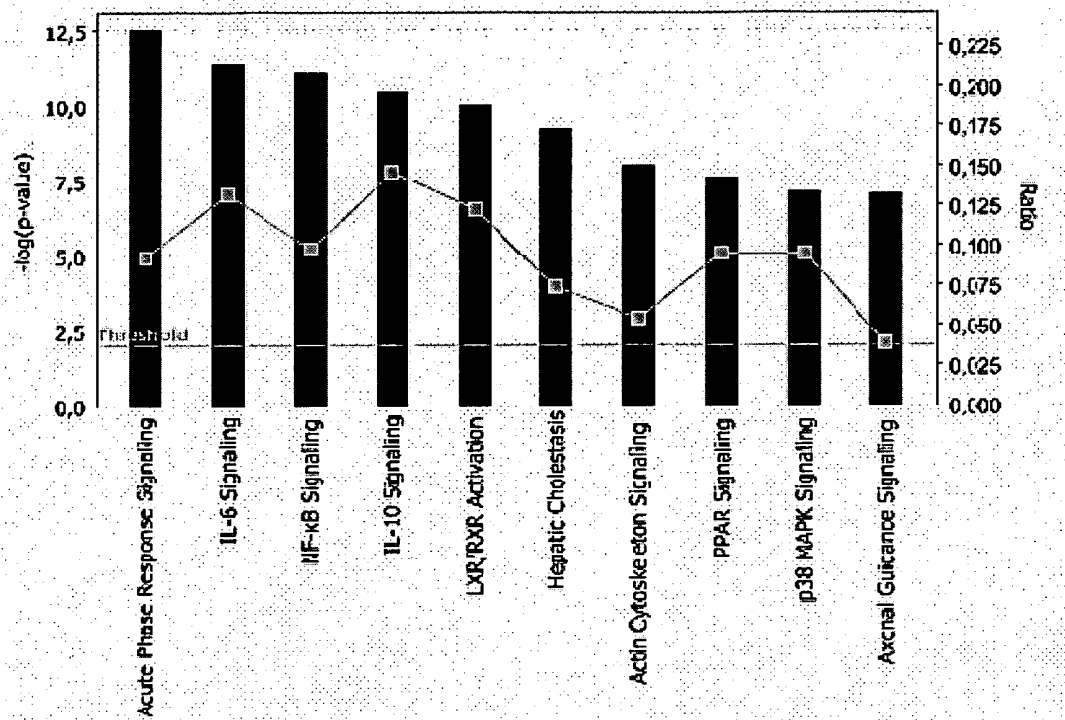


Figure 4

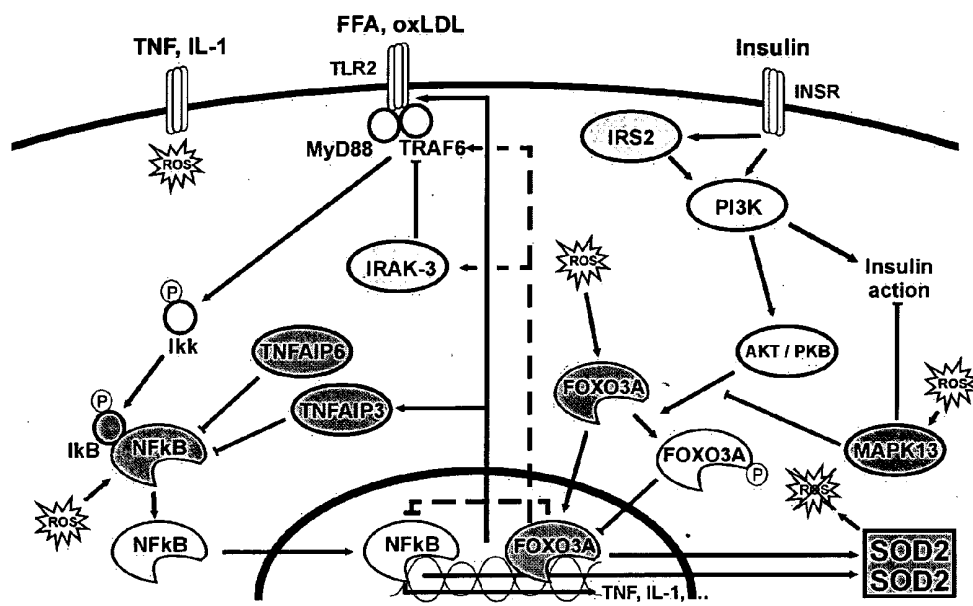


Figure 5

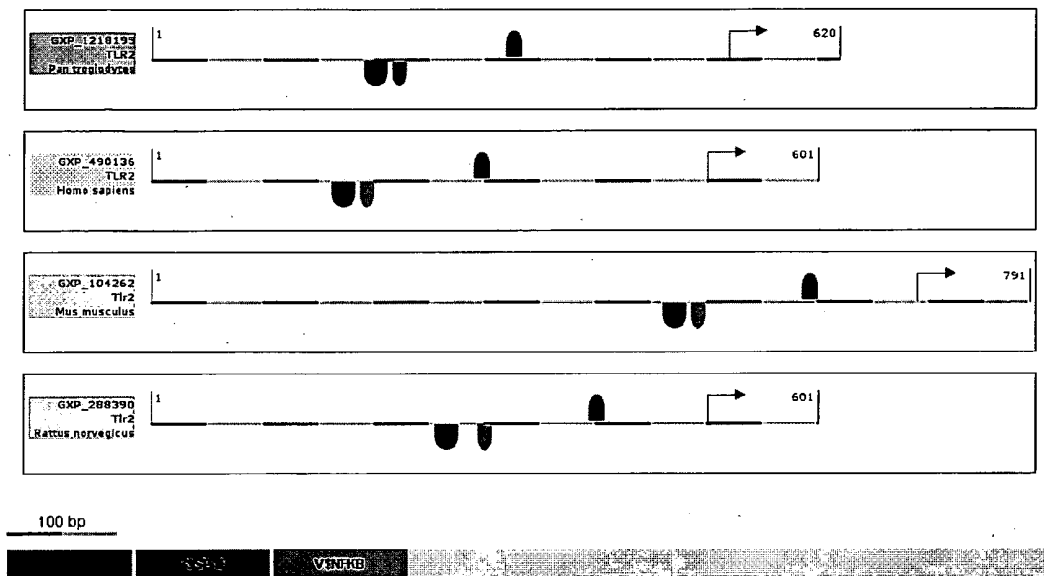


Figure 6

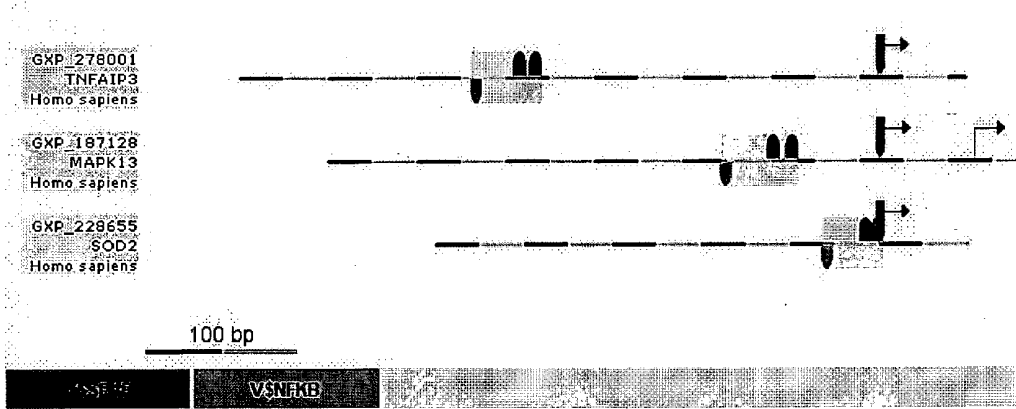
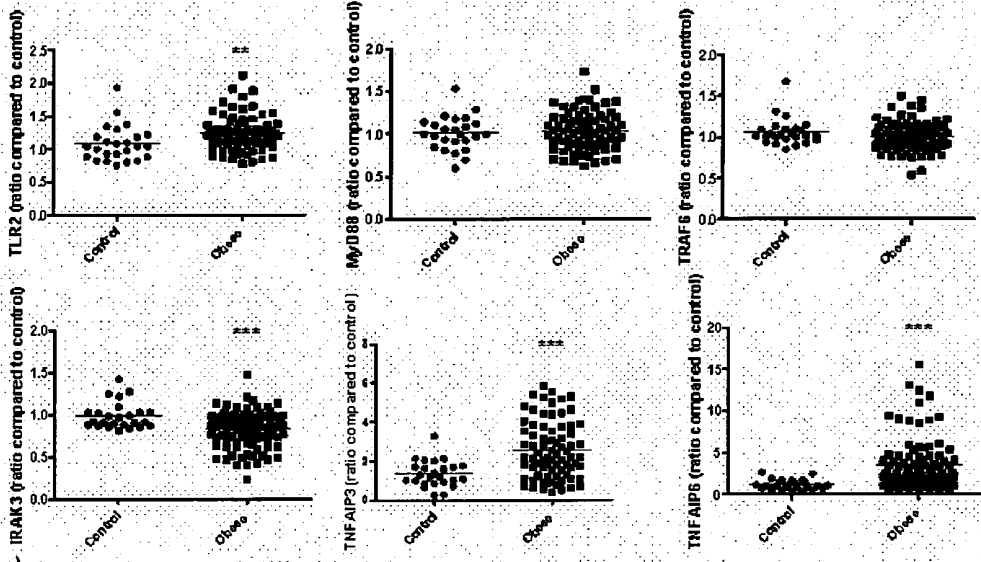
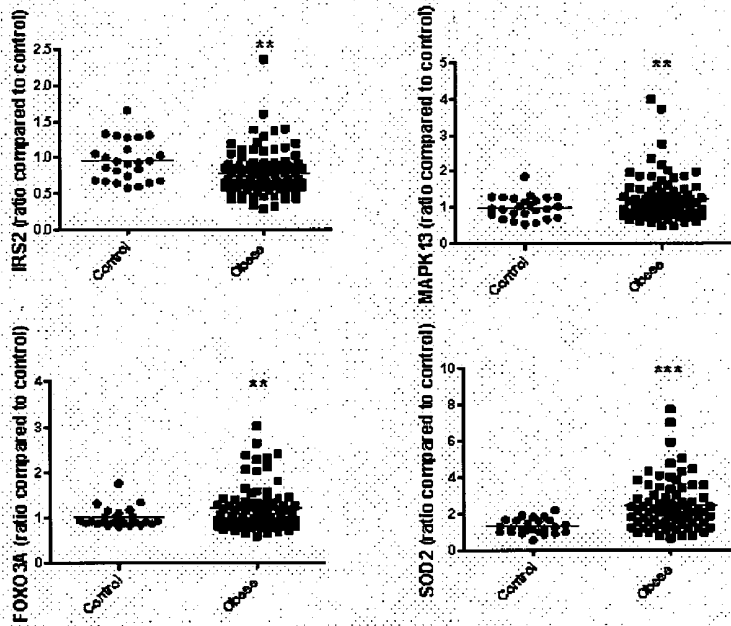


Figure 7

**TLR2 pathway**



**PI3K / AKT pathway**



**Figure 8**

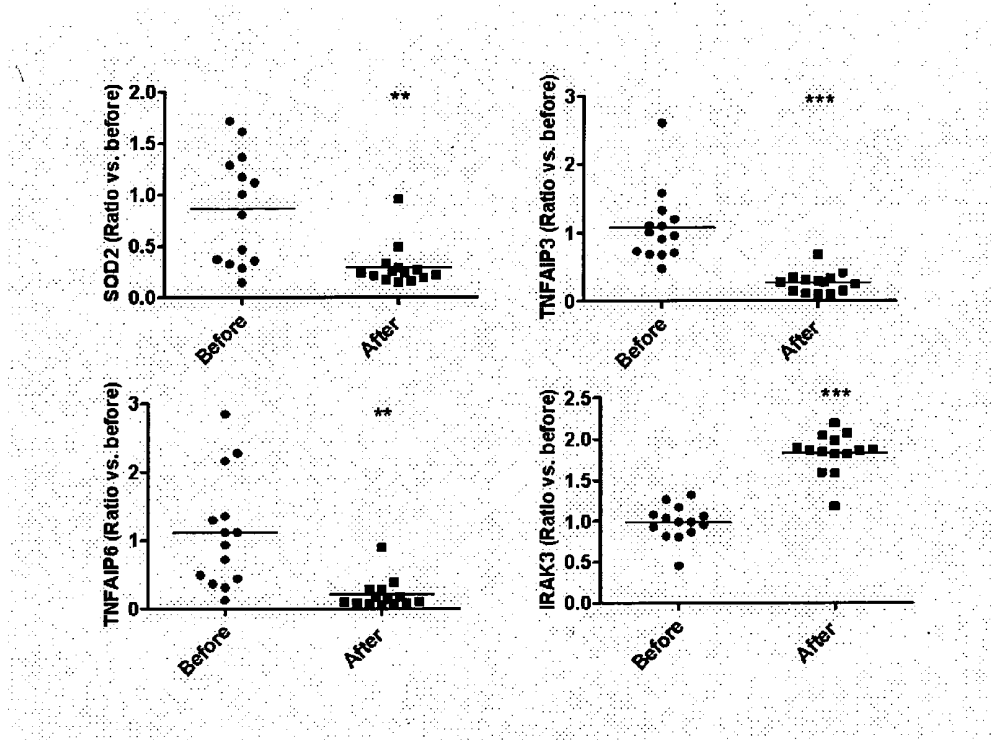
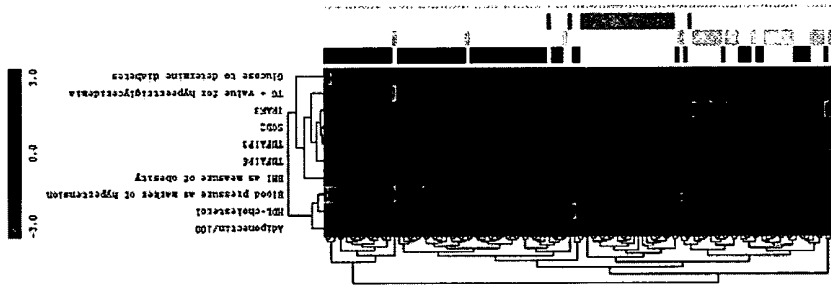


Figure 9

MetSyn + adiponectin + RNA expression



MetSyn + adiponectin

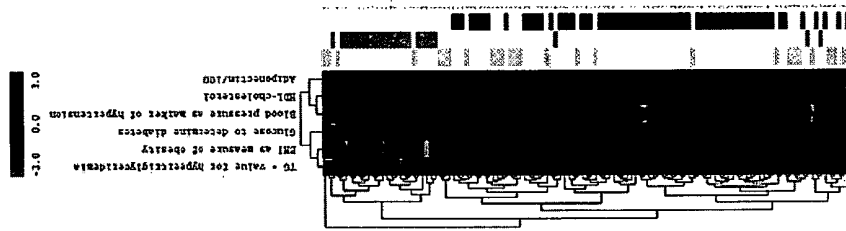
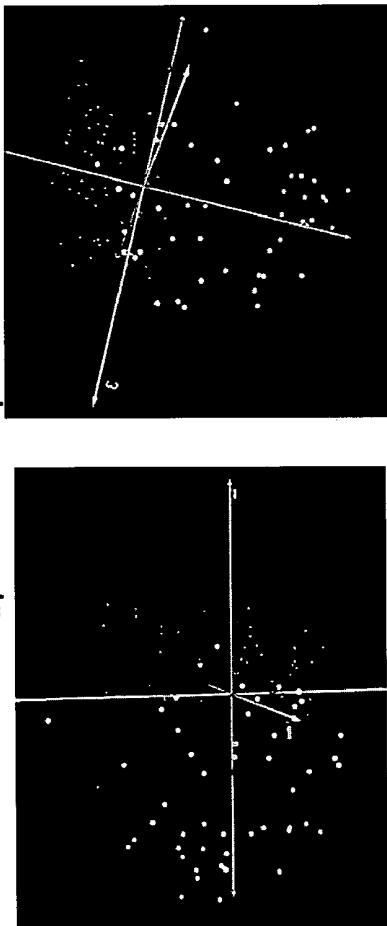
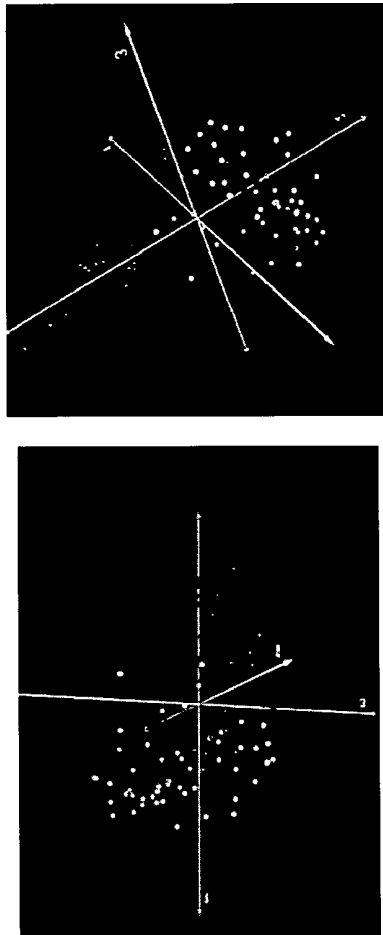


Figure 10

**PCA analysis 1: BMI, glucose, triglycerides, HDL-C, systolic blood pressure and adiponectin for 119 persons**



**PCA analysis 2: RNA expression levels of IRAK3, TNFAIP3, TNFAIP6, and SOD2 were added to variables included in model 1**



**Figure 11**

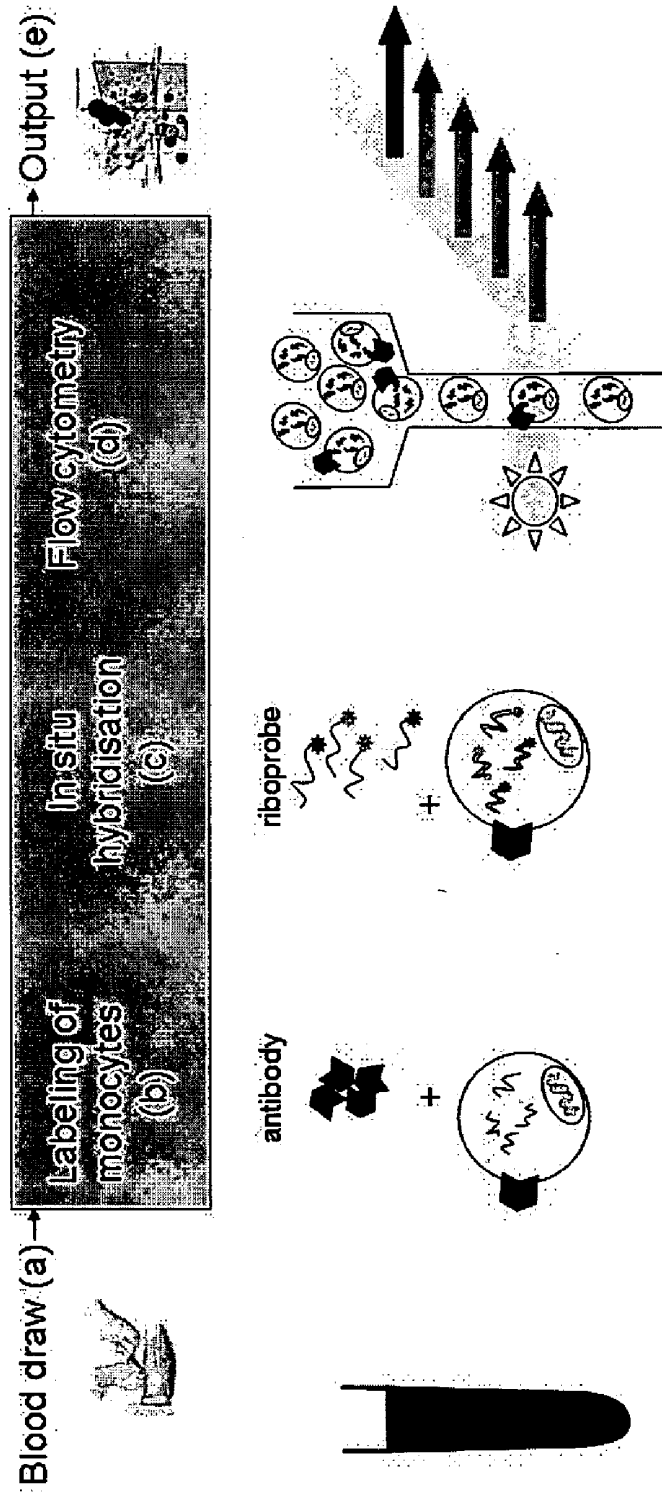


Figure 12

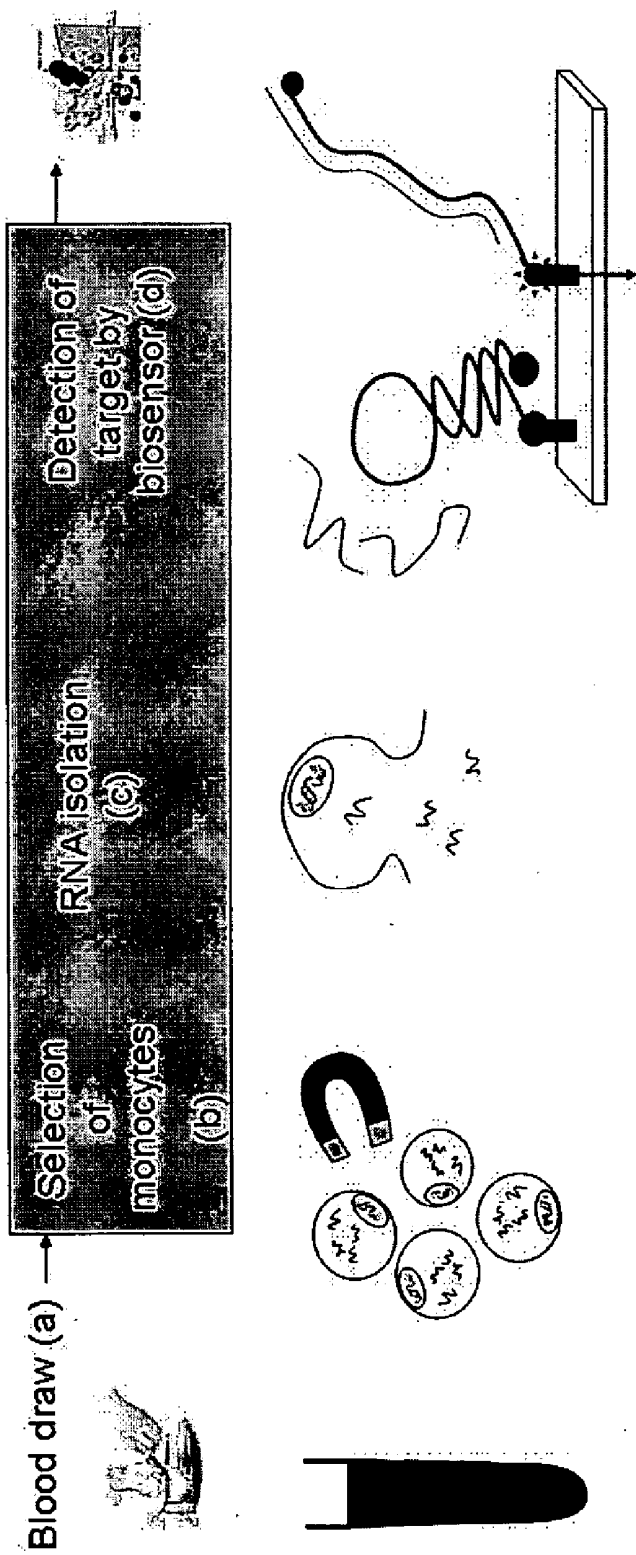


Figure 13

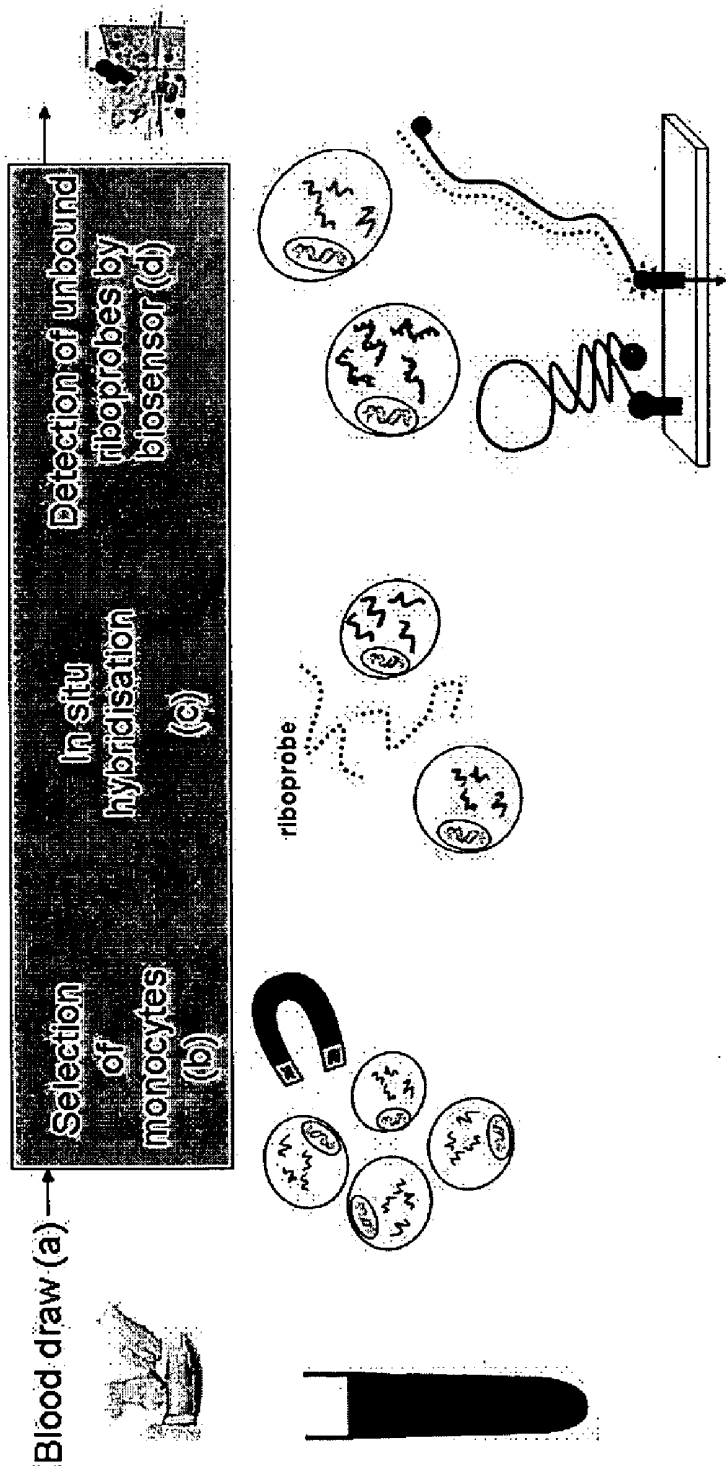


Figure 14

**sequence ID 1**

1 gctgtcgca ggcgtgcagg gacctggact ccgctcgtc cccggggctc gggcagccga  
61 gccatggcgg ggaactgtgg ggcccgcggc gcgctgtcgg cgcacacgct gctgttcgac  
121 ctgccgccc cgctgctcgg agagctctgc gctgttctgg acagctgcga cggcgcgctg  
181 ggctggcgcg gcctggcaga gagacttca agcagctggc tggatgttcg tcatattgaa  
241 aagtatgtag accaaggtaa aagtggaaca agagaattac tttggctctg ggcacagaaa  
301 aacaagacca tcggtgacct ttacaggct ctcaggaga tgggacatcg tcgagctatt  
361 cattaatta caaactatgg agcagtgtg agtccttcag agaagagta tcaggaaggt  
421 ggatttcaa atatattatt caaggaaca gccaatgtca ccgtggataa tgttcttatt  
481 cctgaacata atgaaaaagg agtactgctt aaatctcca tcagcttca aaatatcata  
541 gaaggaacta gaaatttcca caaagactc ctaattggag aaggagagat tttgaggta  
601 tacagagtgg agattcaaaa ctaacatat gctgtcaaat tatttaaca ggagaaaaaa  
661 atgcagtga agaagcattg gaagaggtt ttatctgagc tgaagttt actactgtt  
721 catcaccaa acatactaga gttggctgca tattttacag agactgagaa gttctgtctg  
781 atttatccat acatgagaaa tgaacactt ttgacagat tgcagtgtg aggtgacag  
841 gccccactcc cttggcacat tcgaatcgg atattaatag gaatatcaa agccattcac  
901 tacctgaca acgttcaacc atgctcggc atctgtggca gtatatcaag tgcaaacatc  
961 ctttggatg atcagttca acccaacta actgatttg ccatggcaca ctccggctc  
1021 cacctagaac atcagagtg taccataaat atgaccagca gcagcagtaa acatctgtg  
1081 tacatgccag aagagtacat cagacaggg aaacttcca ttaaacaga tgtctacag  
1141 tttggaattg taataatgga agttctaaca ggatgtagag tagtgttaga  
tgatcaaaa  
1201 catatccagc tgcgggatct ccttagagaa ttgatggaga agagaggcct ggattcatg  
1261 ctctcattc tagataagaa agtgctccc tgcctcgga atttctctgc caagctctc  
1321 tgttggcag gccggtgtc tgcaacgcg gcaaagtaa gaccatcaat ggatgaagtt  
1381 ttaaatactc ttgaaagtac tcaagccagc ttgtatttg ctgaagatcc tcccacatca  
1441 ctaaagtcct tcaggtgtcc ttctctcta ttctggaga atgtaccaag tattccagtg  
1501 gaagatgatg aaagccagaa taacaattta ctacctctg atgaaggcct gaggatagac  
1561 agaatgactc agaaaactcc tttgaaatgc agccagtctg aggttatgtt tctgagctg  
1621 gacaaaaagc cagagagcaa gagaaatgag gaagcttga acatgccag ttctctgt  
1681 gaagaaagtt ggttccaaa gtatatagtt ccatcccagg acttaaggcc ctataaggt  
1741 aatatagatc ctcttcaga agctccagg cattcttga ggagcaggcc agtggagagc

361 gmhllkcqeee kpvvihspgf ytgkpgyklc mrlhlqlpta qrcanyislf vhtmqqeyds  
421 hlpwfpqgti rltildqsea pvrqnheeim dakpellafq rptiprnpkg fgyvtfmhle  
481 alrqrtfikd dtllvrcevs trfdmgslrr egfqprstda gv //

**sequence ID 4**

1 gggcagtggc gtccgcagct ggggcttggc ctgcgggagg ccagcgaagg tggcgaaggc  
61 tccactgga tccagagttt gccgtccaag cagcctcgtc tcggcgcgca gtgtctgtgt  
121 ccgtcctcta ccggcgctt ggctgagcgg agtcgtgagg ttggtggggg agccctgccc  
181 tctggttcg gcctccccgc gactagaac gatcatgaac ttctgaaggg acccagctt  
241 cttgtgtgc tccaagtgt ttgcacaaat aataatata atattattg aaggagagaa  
301 tcagagcaag tgataatcaa gttactatga gtctgctaaa ctgtgaaaac agctgtggat  
361 tcagccagtc tgaaagtgc tgctgtgtgg ccatggccag ctctgtagc gctgtaacaa  
421 aagatgatag tgtgggtgga actgccagca cggggaacct ctccagctca tttatggagg  
481 agatccaggg atatgatga gagtttgacc caccctgga aagcaagtat gaatgcccc  
541 tctgcttgat ggcattacga gaagcagtc aaacgccatg cgccatagg tctgcaaag  
601 cctgcatcat aaaatcaata agggatgagc gtacaaaatg tccagttgac aatgaaatac  
661 tctggtgaaa tcaactattt ccagacaatt ttgcaaaacg tgagattctt tctctgatgg  
721 tgaatgtcc aatgaaggt tgtttgcaca agatggaact gagacatctt  
gaggatcatc  
781 aagcacattg tgagtttgc cttatggatt gtcccaatg ccagcgtccc ttcaaaaat  
841 tccatattaa tattcacatt ctgaaggatt gtccaaggag acaggtttct tgtgacaact  
901 gtgtgcatc aatggcattt gaagataaag agatccatga ccagaactgt cctttggcaa  
961 atgtcatctg tgaatactgc aatactatac tcatcagaga acagatgcct aatcattatg  
1021 atctagactg ccctacagcc ccaattccat gcacattcag tacttttggg tgccatgaaa  
1081 agatgcagag gaatcactg gcacgccacc tacaagagaa caccagtcac cacatgagaa  
1141 tgttggtcca ggctgttcat agtttgagcg ttataccga ctctgggtat atctcagagg  
1201 tccggaattt ccaggaaact attcaccagt tagagggtcg ccttgaaga caagaccatc  
1261 aaatccggga gctgactgct aaaatggaaa ctgagatgata gtatgtaagt gagctcaaac  
1321 gaaccattcg aaccctgag gacaaagttg ctgaaatcga agcacagcag tgcaatggaa

1381 ttatatttg gaagattggc aactttggaa tgcatttgaa atgtcaagaa gaggagaaac  
1441 ctgttgatg tcatagccct ggattctaca ctggcaaacc cgggtacaaa ctgtgcatgc  
1501 gcttgacact tcatgtaccg actgctcagc gctgtgcaaa ctatatatcc cttttgtcc  
1561 acacaatgca aggagaatat gacagccacc tcccttggcc ctccagggg acaatagcc  
1621 ttacaattct tgatcagtct gaagcacctg taaggcaaaa ccacgaagag ataatggatg  
1681 ccaaaccaga gctgctgct ttccagcagc ccacaatccc acggaacca aaaggtttg  
1741 gctatgtaac tttatgcat ctggaagccc taagacaaag aactttcatt aaggatgaca  
1801 cattattagt gcgctgtgag gtctccacc gcttgacat gggtagcctt cggagggagg  
1861 gtttcagcc acgaagtact gatgcagggg tatagcttgc cctcacttgc tcaaaaaaca  
1921 ctactggag aaaacagtgc cttccttgc cctgttctca ataacatgca acaaaacaag  
1981 ccacgggaaa tatgtaatat ctactagtga gtgtgttag agaggtcact tactattct  
2041 tctgttaca aatgatctga ggcagtttt tctgggaat ccacacgtc catgctttt  
2101 cagaaatgtt aggctgaag tgctgtggc atgttcagc agctatttg ccagttagta  
2161 tacctcttg ttgtacttc ttggctttt gctctggtg atttattgt cagaaagtc  
2221 agactcaaga gtactaaact ttaataata atggatttc cttaaactt cagtctttt  
2281 gtagtattat atgtaataa taaaagtga aatcactac cgccttgtc tagtgcctc  
2341 gagaagagt attgcttag aaagtgagt tctattttt ttaacctgt atagattca  
2401 gaggattga accataatcc ttgaaaact taagttctca tccaccag ttttctcc  
2461 aggtgttac taaggatatt caggatgag ttaaaccct aatataacc ttaattatt  
2521 agtgtaaaca tgtctgtga ataacttgg ttaagtgt aaaaaaaaaa aaaaaaagaa  
2581 aaaaaaaaaa aaaa //

**sequence ID 5**

1 gggcagtggc gtccgagct ggggcttggc ctgcggcgg ccagcgaagg tggcgaaggc  
61 tccactgga tccagagttt gccgtccaag cagcctcgtc tcggcgcga gtgtctgtg  
121 ccgtctcta ccggcgcctt ggctgagcgg agtcgtcgg ttggtggggg agccctgccc  
181 tctgtgttc gctccccgc gactagaac gatcatgaac ttctgaagg acccagctt  
241 cttgtgtgc tccaagtga ttgcacaaat aataatata atattattg aaggagagaa  
301 tcagagcaag tgataatcaa gttactatga gtctgctaaa ctgtgaaaac agctgtggat  
361 tcagccagtc tgaagtgc tgctgtgtgg ccatggccag ctctgtagc gctgtaacaa

421 aagatgatag tgtgggtgga actgccagca cggggaacct ctccagctca tttatggagg  
481 agatccaggg atatgatgta gagtttgacc caccctgga aagcaagtat gaatgcccc  
541 tctgcttgat ggcattacga gaagcagtgc aaacgccatg cggccatagg ttctgcaaag  
601 cctgcatcat aaaatcaata agggatgcag gtcacaaatg tccagttgac aatgaaatac  
661 tgctggaaaa tcaactattt ccagacaatt ttgcaaaacg tgagattctt tctctgatgg  
721 tgaatgtcc aatgaaggt tgttgcaca agatggaact gagacatctt gaggatcatc  
781 aagcacattg tgagtttgct cttatggatt gtcccaatg ccagcgtccc ttcaaaaat  
841 tccatattaa tattcacatt ctgaaggatt gtccaaggag acaggttctt tgtgacaact  
901 gtgctgcatc aatggcattt gaagataaag agatccatga ccagaactgt cctttggcaa  
961 atgtcatctg tgaatactgc aatactatac tcacagaga acagatgctt aatcattatg  
1021 atctagactg ccctacagcc ccaattccat gcacattcag tacttttggg tgccatgaaa  
1081 agatgcagag gaatcacttg gcacgccacc tacaagagaa caccagtcac cacatgagaa  
1141 tgttggccca ggctgttcat agtttgagcg ttatacccga ctctgggtat atctcagagg  
1201 tccggaattt ccaggaaact attcaccagt tagagggctg ccttgaaga caagaccatc  
1261 aaatccggga gctgactgct aaaatggaaa ctacagatg gtatgtaagt gagctcaaac  
1321 gaaccattcg aacccttgag gacaaagtg ctgaaatcga agcacagcag tgcaatggaa  
1381 tttatatttg gaagattggc aactttggaa tgcaattgaa atgtcaagaa gaggagaaac  
1441 ctgttgatg tcatagccct ggattctaca ctggcaaac cgggtacaaa ctgtgcatgc  
1501 gcttgacct tcagttaccg actgctcagc gctgtgcaaa ctatatatcc cttttgtcc  
1561 acacaatgca aggagaatat gacagccacc tccttggcc cttccagggt acaatagcc  
1621 ttacaattct tgatcagtct gaagcacctg taaggcaaaa ccacgaagag  
ataatgatg  
1681 ccaaaccaga gctgcttctt tccagcgac ccacaatccc acggaacca aaaggtttg  
1741 gctatgtaac tttatgcat ctggaagccc taagacaaag aactttcatt aaggatgaca  
1801 cattattagt gcgctgtgag gtctccacc gcttgacat gggtagcctt cggagggagg  
1861 gtttcagcc acgaagtact gatgcagggg tatagcttgc ctcacttgc tcaaaaacaa  
1921 ctacctggag aaaacagtgc cttccttgc cctgttctca ataacatgca acaaaacaag  
1981 ccacgggaaa tatgtaatat ctactagtga gtgttgtag agaggtcact tactattct  
2041 tctgttaca aatgatctga ggcagttttt tctggaat ccacagttc catgctttt  
2101 cagaaatgtt aggctgaag tgctgtggc atgttgagc agctattttg ccagttagta  
2161 tacctcttg ttgtacttc ttgggtttt gctctggtgt atttattgt cagaaagtc  
2221 agactcaaga gtactaaact ttaataata atggatttc ctaaaactt cagtctttt

2281 gtagtattat atgtaatata ttaaagtgga aaatcactac cgcttctgag tagtgcctc  
2341 gagaagagtt attgctctag aaagttgagt tctcattttt ttaacctgtt atagattca  
2401 gaggatttga accataatcc ttggaaaact taagttctca ttcacccag ttttctcc  
2461 aggtgttac taaggatatt cagggatgag ttaaaccct aaatataacc ttaattatt  
2521 agtgtaaaca tgtctgttga ataactctg ttaagtggt aaaaaaaaaa aaaaaagaa  
2581 aaaaaaaaaa aaaa //

**sequence ID 6**

1 atgcgaccg accgcgctga ggctccagga cgcccccca tggctgcagg aggtcccggc  
61 gcggggctg cgccccggt ctctccaca tctccctc cctggctgc tctcaacatg  
121 cgagtgcggc gccgctgtc tctgtcttg aacgtgcgga cacaggtggc ggccgactgg  
181 accgcgctgg cggaggagat ggactttgag tacttgagaga tccggcaact ggagacacaa  
241 gcggaccca ctggcaggct gctggacgcc tggcaggac gccctggcgc ctctgtaggc  
301 cgactgctcg agctgcttac caagctggg cgcgacgacg tgctgctgga gctgggaccc  
361 agcattgagg aggattgcca aaagtatctc ttgaagcagc agcaggagga ggctgagaag  
421 cctttacagg tggccgctgt agacagcagt gtcccacgga cagcagagct ggccggcatc  
481 accacactg atgacccct ggggcatatg cctgagcgtt tcgatgcctt  
catctgctat  
541 tgccccagcg acatccagtt tgtgcaggag atgatccggc aactggaaca gacaaactat  
601 cgactgaagt tgtgtgtgtc tgaccgcat gtctgctg gcacctgtgt ctggtctatt  
661 gctagtgagc tcatcgaaaa gaggtgccg cggatggtgg tggttgtctc tgatgattac  
721 ctgcagagca aggaatgtga ctccagacc aaattgcac tcagcctctc tccaggtgcc  
781 catcagaagc gactgatccc catcaagtac aaggcaatga agaaagagtt cccagcatc  
841 ctgaggttca tcaactgtct cgactacacc aaccctgca ccaaactttg gttctggact  
901 cgcttgcca aggcctgtc cctgcctga agactgttct gaggcctgg gtgtgtgtgt  
961 atctgtctg ctgtccatgt acttctgcc tgctctctc tttcgttga ggaggaatc  
1021 gtgtcttact tactctcaa ttctggaga tgccaactc acagacagct ctgcagcagc  
1081 tggacatcac attcatgtc ctgcatgga ccagtggtg tgagtggcat gtccactgc

1141 tggattatca gccaggacac tatagaacag gaccagctga gactaagaag gaccagcaga  
1201 gccagctcag ctctgagcca ttcacacatc ttcacctca gtttcctcac ttgaggagtg  
1261 ggatggggag aacagagagt agctgtgttt gaatccctgt aggaaatggg gaagcatagc  
1321 tctgggtctc ctgggggaga ccaggcttgg ctgcgggaga gctggctgtt gctggactac  
1381 atgctggcca ctgctgtgac cacgacactg ctggggcagc ttctccaca gtgatgccta  
1441 ctgatgcttc agtgcctctg cacaccgcc attccacttc ctcttcccc acagggcagg  
1501 tggggaagca gtttgccca gcccaaggag accccatctt gagccttatt tctaattggg  
1561 tccacctc atctgcatct ttcacacctc ccagcttctg cccaacctc agcagtgaca  
1621 agtcccaag agactgcct gagcagcttg ggctgcttt cattccacc tgcaggatg  
1681 cctgtgtca tgcttcagc tccacctggc atgagaagg atcctggcct ctggcatatt  
1741 catcaagtat gaggcttgg gatgagctac tgtaatgatg tgagcagga gccttctcc  
1801 ctggccacc tcagagagc tttccacca actttgtacc ttgattgcct tacaagtta  
1861 tttgttaca aacagcgacc atataaagc ctctgcccc aaagcttgg ggcacatggg  
1921 cacatacaga ctacataca gacacacaca tatatgtaca gacatgtact ctacacaca  
1981 caggcaccag catacacag ttttctagg tacagctccc aggaacagct aggtgggaaa  
2041 gtcccatcac tgaggagacc taaccatgct cctgaacaaa aattgggcac tcatctattc  
2101 ctttctctt gtgtccctac tcattgaaac caaactctgg aaaggacca atgtaccagt  
2161 attatacct ctaatgaagc acagagagag gaagagagct gcttaaactc acacaacaat  
2221 gaactgcaga cacagctgtt ctctcctct ctcttcca gagcaattta tactttacc  
2281 tcaggctgct ctctggggag aaggtgcat ggtcttaggt gtctgtgccc caggacagac  
2341 ctaggacc taaatcaat agaaaatgca tatcttctgct ccactttag  
ccaggctgga  
2401 gcaaggatc ttttctagg atctgggag ggaatggatg cccctctctg catgatctg  
2461 ttgaggcatt tagctgcat gcacctgtcc cctttaata ctgggcattt taaagccatc  
2521 tcaagaggca tcttctacat gtttgtacg cattaataa attcaaaga tatctgagaa  
2581 aagccgatat ttgccattct tctatatcc tggaaatata ctgcatcct gaggttataa  
2641 taataaataa tattctacct tggaaaaaaa aaaaaaaa //

**sequence ID 7**

"MAAGGPGAGSAAPVSSTSSLPLAALNMRVRRRLSLFLNVRTQVA  
ADWTALAEEMDFEYLEIRQLETQADPTGRLLDAWQGRPGASVGRLELLTKLGRDDVL

LELGPSIEEDCQKYILKQQQEEAEKPLQVAAVDSSVPRTAELAGITTLDDPLGHMPER  
FDAFICYCPSDIQFVQEMIRQLEQTNYRLKCVSDRDVLPGTCVWSIASIELIEKRCRR  
MVVVSDDYLSKECDFQTKFALSLSPGAHQKRLIPIKYKAMKKEFPSILRFITVCDY  
TNPCTKSWFWTRLAKALSLP"

**sequence ID 8**

1 tgccttgacc aggacttggg actttgcgaa aggatcgcg ggccccggaga ggtgttgag  
61 agcacaatgg ctgaacaagt ccttctcag gctttgtatt tgagcaatat gcggaaagct  
121 gtgaagatac gggagagAAC tccagaagac attttaaac ctactaatgg gatcattcat  
181 catttataaa ccatgcaccg atacacactg gaaatgttca gaacttgcca gttttgtcct  
241 cagtttcggg agatcatcca caaagccctc atcgacagaa acatccaggc caccctggaa  
301 agccagaaga aactcaactg gtgtcgagaa gtccggaagc ttgtggcgct gaaaacgaac  
361 ggtgacggca attgcctcat gcatgccact tctcagtaca tgtggggcgt tcaggacaca  
421 gacttggtac tgaggaaggc gctgttcagc acgctcaagg aacagacac acgcaactt  
481 aaattccgct ggcaactgga gtctctcaa tctcaggaat ttgttgaac ggggcttgc  
541 tatgatactc ggaactggaa tgatgaatgg gacaatctta tcaaatggc ttccacagac  
601 acaccatgg cccgaagtgg acttcagtac aactcactgg aagaaataca catattgtc  
661 cttgcaaca tctcagaag gccaatcatt gtcattcag acaaaatgct aagaagttg  
721 gaatcaggtt ccaatttcgc cctttgaaa gtgggtggaa ttacttgcc tctccactg  
781 ctgcccagg aatgctacag ataccatt gttctcgct atgacagcca tcatttga  
841 cccttggtga ccctgaagga cagtggcct gaaatccgag ctgttccact tgtaacaga  
901 gaccgggaa gattgaaga cttaaagtt cacttttga cagatcctga aaatgagatg  
961 aaggagaagc tcttaaaga gtacttaag                      gtgatagaa tccccgtcca  
aggctggac  
1021 catggcaca ctatctcat caatgccga aagttgatg aagtaactt accaaaagaa  
1081 atcaatctgg tagatgatta ctttgaactt gttcagcatg agtacaagaa atggcaggaa  
1141 aacagcgagc aggggaggag agaggggac gccagaatc ccatggaacc ttccgtgccc  
1201 cagctttctc tcatggatgt aaaatgtgaa acgcccact gcccttctt catgtctgtg  
1261 aacaccagc cttatgcca tgagtgtca gagaggcggc aaaagaatca aaacaaactc  
1321 ccaaagctga actcaagcc gggcctgag gggctcctg gcatggcgct cggggcctct  
1381 cggggagaag cctatgagcc ctggcggtg aaccctgagg agtccactgg gggcctcat  
1441 tcggccccac cgacagcacc cagcccttt ctgttcagtg agaccactgc catgaagtgc  
1501 aggagccccg gctgccctt cacactgaat gtgcagcaca acggattttg tgaacgttc  
1561 cacaacgccc ggcaactca cgccagccac gcccagacc acacaaggca cttggatccc  
1621 ggggaagtgc aagcctgct ccaggatgtt accaggacat ttaatgggat ctgcagtact

1681 tgcttcaaaa ggactacagc agaggcctcc tccagcctca gcaccagcct cctcctcc  
1741 tgtcaccagc gttccaagtc agatccctcg cggctcgtcc ggagcccctc cccgattct  
1801 tgccacagag ctggaacga cgcccctgct ggctgctgt ctcaagctgc acggactcct  
1861 ggggacagga cggggacgag caagtgcaga aaagccggct gcgtgtattt tgggactcca  
1921 gaaaacaagg gctttgcac actgtgttc atcgagtaca gagaaaaca acattttgct  
1981 gctgcctcag ggaaagtcag tcccacagcg tccaggtcc agaacacat tccgtgcctg  
2041 gggagggaat gcggcacctc tggaagcacc atgtttgaag gatactgcca gaagtgttc  
2101 attgaagctc agaatcagag atttcatgag gccaaaagga cagaagagca actgagatcg  
2161 agccagcgcga gagatgtgcc tcgaaccaca caaagcacct caaggcccaa gtgcgcccgg  
2221 gctcctgca agaacatcct ggctgcgcg agcgaggagc tctgcatgga gtgtcagcat  
2281 cccaaccaga ggatgggccc tggggccac cggggtgagc ctgccccga agaccccc  
2341 aagcagcgtt gccggcccc cgctgtgat cattttggca atgccaagtg caacggctac  
2401 tgcaacgaat gcttcagtt caagcagatg tatggctaac cggaaacagg tgggtcacct  
2461 cctgcaagaa gtggggcctc gagctgtcag tcatcatggt gctatcctct gaacccctca  
2521 gctgccactg caacagtggg ctttaaggtg tctgagcagg agaggaaaga taagctcttc  
2581 gtggtgcca cgatgctcag gtttgtaac ccgggagtgt tcccaggtgg ccttagaaag  
2641 caaagctgt aactggcaag ggatgatgtc agattcagcc caaggttct cctctctac  
2701 caagcaggag gccaggaact tctttggact tggaaaggtg gcggggactg gccgaggccc  
2761 ctgcacctg cgcatcagga ctgcttcac gtcttgctg agaaagggaa aagacacaca  
2821 agtcgctgg gttggagaag ccagagccat tccacctccc ctccccagc atctctcaga  
2881 gatgtgaagc cagatcctca tggcagcag gccctctgca agaagctcaa  
ggaagctcag  
2941 ggaaaatgga cgtattcaga gagtgtttgt agttcatggt tttccctac ctgcccggtt  
3001 cctttctgga ggaccggca gaaatgcaga accatccatg gactgtgatt ctgaggctgc  
3061 tgagactgaa catgttcaca ttgacagaaa aacaagctgc tcttataat atgcacctt  
3121 taaaaaatta gaatattta ctggaagac gtgtaactct ttgggttatt actgtctta  
3181 cttctaaaga agttagctg aactgaggag taaaagtgtg tacatatata atataccctt  
3241 acattatgta tgagggattt ttttaatta tattgaaatg ctgccctaga agtacaatag  
3301 gaaggctaaa taataatac ctgtttctg gttgtgttg gggcatgagc ttgtgtatac  
3361 actgcttga taaactcaac cagctgcctt tttaaagga gctctagtcc ttttgtga  
3421 attcacttta tttatttat tacaacttc aagattattt aagtgaagat atttctcag  
3481 ctctggggaa aatgccacag ttttctctg agagaacatc ctgctttga gtcaggctgt

3541 gggcaagttc ctgaccacag ggagtaaatt ggcctcttg atacactttt gcttgctcc  
3601 ccaggaaaga aggaattgca tccaaggat acatacatat tcatcgatgt ttcgtgctc  
3661 tccttatgaa actccagcta tgtaataaaa aactatactc tgtgttctgt taatgcctc  
3721 gagtgccta cctccttga gatgagatag ggaaggagca gggatgagac tggcaatggt  
3781 cacaggaaa gatgtggcct tttgtgatgg tttattttc tgtaacact gtgtcctggg  
3841 ggggctggga agtcccctgc atccatggt accctggat tgggacagca aaagccagta  
3901 accatgagta tgaggaaatc tctttctgtt gctggcttac agtttctctg tgtgctttgt  
3961 ggttgctgtc atatttgctc tagaagaaaa aaaaaaagg aggggaaatg cattttccc  
4021 agagataaag gctgccattt tgggggtctg tacttatggc ctgaaaatat ttgtgatcca  
4081 taactctaca cagccttac tcatactatt aggcacactt tcccctaga gccccctaa  
4141 ttttcccag acgaatctt ataatttctt ttcaaagat accaaataaa cttcagtgt  
4201 tcatctaata tctctaaag ttgatctt aatatttgt gttgatcatt attccattc  
4261 ttaatgtgaa aaaaagtaat tattatact tattataaaa agtatttgaa attgacat  
4321 ttaattgcc ctaatagaaa gccacatt cttgttga tttctcaag ttttctaaa  
4381 taaatgtaac tttcacaag agtcaacatt aaaaaataaa ttattt//

**sequence ID 9**

"MAEQVLPQALYLSNMRKAVKIRERTPEDIFKPTNGIIHHFKTMH

RYTLEMFRTCQFCPQFREIIHKALIDRNIQATLESQKKLNWCREVRKLVALKTNGDGN  
CLMHATSQYMWGVQDSDLVLRKALFSTLKETDRNFKFRWQLESLSQEFVETGLCY  
D  
TRNWNDEWDNLIKMASTDTPMARSGLQYNSLEEIHIFVLCNILRRPIIVISDKMLRSL  
ESGSNFAPLKVGGIYLPLHWPAQECYRYPVILGYDSHHFVPLVTLKDSGPEIRAVPLV  
NRDRGRFEDLVHFLTPENEMKEKLLKEYLMVIEIPVQGDHGTTHLINAAKLDEAN  
LPKEINLVDDYFELVQHEYKKWQENSEQGRREGHAQNPMEPSVPQLSLMDVKCETPN  
C  
PFFMSVNTQPLCHECSERRQKNQNKLPKLNKPGPEGLPGMALGASRGEAYEPLAWN  
P

EESTGGPHSAPPTAPSPFLFSETTAMKCRSPGCPFTLNVQHNGFCERCHNARQLHASH  
APDHTRHLDPGKCQACLQDVTRTFNGICSTCFKRTTAEASSLSTSLPPSCHQRSKSD  
PSRLVRSPSPHSCHRAGNDAPAGCLSQAARTPGDRTGTSKCRKAGCVYFGTPENKGF  
C  
TLCFIEYRENKHFAAASGKVSPTASRFQNTIPCLGRECGTLGSTMFEGYCQKCFIEAQ  
NQRFHEAKRTEEQLRSSQRRDVPRTTQSTS RPKCARASCKNILACRSEELCMECQHP  
N  
QRMGPGAHRGEPAPEDPPKQRCRAPACDHFGNAKCNNGYCNECFQFKQMYG"

**sequence ID 10**

1 cagtcacatt tcagccactg ctctgagaat ttgtgagcag ccctaacag gctgttactt  
61 cactacaact gacgatatga tcacttaat ttacttattt ctcttgctat gggaagacac  
121 tcaaggatgg ggattcaagg atggaatfff tcataactcc atatggcttg aacgagcagc  
181 cgggtgtgtac cacagagaag cacgggtctgg caaatacaag ctcacctacg cagaagctaa  
241 ggcgggtgtgt gaatttgaag gcggccatct cgcaacttac aagcagctag aggcagcccag  
301 aaaaattgga tttcatgtct gtgctgctgg atggatggct aagggcagag ttgatatccc  
361 cattgtgaag ccagggcca actgtggatt tggaaaaact ggcattattg attatggaat  
421 ccgtctcaat aggagtgaag gatgggatgc ctattgtctac aaccacacg caaaggatg  
481 tgggtggcgtc ttacagatc caaagcaaat ttttaaatct ccaggctcc caaatgagta  
541 cgaagataac caaatctgct actggcacat tagactcaag tatggtcagc  
gtattcacct  
601 gagtttttta gattttgacc ttgaagatga cccaggttgc ttgctgatt atgttgaat  
661 atatgacagt tacgatgatg tccatggctt tggggaaga tactgtggag atgagctcc  
721 agatgacatc atcagtacag gaaatgtcat gacctgaag ttttaagtg atgcttcagt  
781 gacagctgga ggtttcaaaa tcaaatatgt tgcaatggat cctgtatcca aatccagtca  
841 aggaaaaaat acaagtacta ctctactgg aaataaaaac ttttagctg gaagatttag  
901 ccattataa aaaaaaaaaa aaggatgatc aaaacacaca gtgttatgt tggatcttt  
961 tggaaactct ttgatctcac tgttattatt aacatttatt tattattttt ctaaagtga  
1021 aagcaataca taatttaggg aaaattggaa aatataggaa actttaaacg agaaaatgaa  
1081 accttcata atcccactgc atagaaataa caagcgttaa catttcata ttttttctt  
1141 tcagtcattt ttctattgtt ggtatatgta tatatgtacc tatatgtatt tgcattgaa

1201 attttgaat cctgctctat gtacagtttt gtattatact ttttaaatct tgaactttat  
1261 aaacattttc tgaatcatt gattattcta caaaaacatg attttaaaca gctgtaaaat  
1321 attctatgat atgaatgttt tatgcattat ttaagcctgt ctctattggt ggaatttcag  
1381 gtcattttca taaatattgt tgcaataaat atccttgaac acaaaaaaaaa aaaaaaaaaa //

**sequence ID 11**

"MIILYLFLLWEDTQGWGFKDGIHNSIWLERAAGVYHREARS  
GKYKLYAEAKAVCEFEFGHLATYKQLEAARKIGFHVCAAGWMAKGRVGYPIVKPGPN  
CGFGKTGIIDYGIRLNRSERWDAYCYNPHAKECGGVFTDPKQIFKSPGFPNEYEDNQI  
CYWHIRLKYQRIHLSFLDFDLEDDPGCLADYVEIYDSYDDVHGFVGRYCGDELDDI  
ISTGNVMTLKFSLDASVTAGGFQIKYVAMDPVSKSSQGKNTSTTSTGNKNFLAGRFS  
HL"

**sequence ID 12**

1 cggcggcgcg gtcggagggg gccggcgcgc agagccagac gccgcccgtt gtttggtt  
61 gggctctcgg caactctcg aggaggagga ggaggagga ggaggggaga agtaactgca  
121 gcggcagcgc cctcccagg aacaggcgtc tccccgaac cttcccaaa cctcccct  
181 cccctctcgc cctgtcccc tcccctctc cccagccgc tggagcgagg ggcagggatg  
241 agtctgtccc tccggccgtt cccagctgc agtggctgcc cggatcgtt tgcagtgaa  
301 aagccactt cccaccgc cgagatggc ccggatggg ctgcagagga  
cgccccgcg  
361 ggcgccgca gcagcagcag cagcagcagc agcaacagca acagccgag cgccgcggtc  
421 tctgcgactg agctgtatt tggcggtg gtggcggtg ggacggttg ggggtgggag  
481 gagcggaagg aggagggaga acccgtgca acgttgggac ttggcaacc gcctcccct  
541 gcccaaggat attaattg cctcggaat cgtgcttc agaggggaac tcaggagga  
601 aggcgcgcgc gcgcgcgcgc tctggaggg gcaccgagg gaccccgac tgcgctcc  
661 ctgtgccgga ctccagccg ggcgacgaga gatgcatctt cgtccttc tggtgccg  
721 ggcggtgag aggagactg gctctcgag gatcggggt gcctcacc cgagcact  
781 gcctcccgc cggcgtgaag cccccgaaa ctccggtcg gctctctct gggctcagca  
841 gctgcgtct cctcagctg cccctcccgc gcgccccggg cggcgtgat tcagagtcg  
901 gggtttctgc tgctccagc cctgttgca tgtccgggc cgcgccgagg agcctccgc

961 cccacccgg ttgttttcg gagcctcct ctgctcagcg ttggtggtgg cggtggcagc  
1021 atggcgagcc ctccggagag cgatggcttc tcggacgtgc gcaaggtggg ctacctgcg  
1081 aaaccaaga gcatgcacaa acgcttctc gtactgcgcg cggccagcga ggctgggggc  
1141 ccggcgcgcc tcgagtacta cgagaacgag aagaagtggc ggcacaagtc gagcgcccc  
1201 aaacgctcga tccccctga gagctgctc aacatcaaca agcgggctga ctccaagaac  
1261 aagcacctgg tggctctta caccgggac gagcacttg ccatcgggc ggacagcgag  
1321 gccgagcaag acagctgga ccaggctctc ctacagctgc acaaccgtgc taagggccac  
1381 caccagggag ctgcggcct cggggcggga ggtggtggg gcagctgcag cggcagctcc  
1441 ggccttggt aggtgggga ggactgagc tacggtgacg tgccccagg acccgcttc  
1501 aaagaggtct ggcaagtgat cctgaagccc aaggcctgg gtcagacaaa gaactgatt  
1561 ggtatctacc gccttgct gaccagcaag accatcagct tcgtgaagct gaactcggag  
1621 gcagcggccg tggctgca gctgatgaac atcaggcgt gggccactc ggaaaactc  
1681 ttctcatcg aggtgggccc ttctgctg acggggccc gggagtctg gatcaggtg  
1741 gatgactctg tggggcca gaacatgac gagaccatc tggaggccat gcgggcatg  
1801 agtgatgagt tccgctcag cagcaagagc cagtctcgt ccaactgctc taacccatc  
1861 agcgtcccc tgcgcccga ccatctaac aatccccgc ccagccaggt ggggctgacc  
1921 cgccgatcac gactgagag catcaccgc acctccccg ccagcatggt gggcgggaag  
1981 ccaggctct tccgtgccc gcctccagt gacggcgaag gcaccatgct ccgcccagc  
2041 tcggtggacg gcagccctgt gactcccagc accaacagaa cccacgccc ccggcatcgg  
2101 ggcagcggc ggctgcccc ccgctcaac cacagccgt ccatccccat gccggcttc  
2161 cgctgctgc ctccggccac cagcccgtc agtctgctg ccagtagcac  
cagtggccat  
2221 ggctccact cggattgtct ctcccacgg cgatctagtg ctccggtgc tgggtcccc  
2281 agcgatggcg gttcatctc ctccgatgag tatggctcca gtcctgcga ttccggagt  
2341 tcctccgca gtgtactcc ggaltccctg ggccacacc caccagccc cggtgaggag  
2401 gagctaagca actatatctg catgggtggc aagggccct ccaccctgac ccccccaac  
2461 ggtcactaca tttgtctc gggggcaat ggccaccgt gcacccagc aacaggctg  
2521 ggcacgagtc cagcctggc tggggatgaa gcagccagtg ctgcagatct ggataatcg  
2581 ttccgaaaga gaactcactc ggcaggcaca tccctacca ttaccacca gaagacccc  
2641 tccagctct cagtggctc cattgaggag tacacagaga tgatgctgc ctaccacca  
2701 ggaggtggca gtggaggcc actgccgga cacaggcact ccgctctg gccacccgc  
2761 tcctaccag aggagggtct ggaaatgac ccctggagc gtcgggggg gcaccaccg

2821 ccagacagct ccaccctcca cacggatgat ggctacatgc ccatgtcccc aggggtggcc  
2881 ccagtgccca gtggccgaaa gggcagtgga gactatatgc ccatgagccc caagagcgta  
2941 tctgccccac agcagatcat caatcccatc agacgccatc cccagagagt ggacccaat  
3001 ggctacatga tgatgtcccc cagcggtagc tgctctctg acattggagg tggccccagc  
3061 agcagcagca gcagcagcaa cgccgtccct tccgggacca gctatggaaa gctgtggaca  
3121 aacgggtag ggggccacca ctctcatgtc ttgctcacc ccaaaccacc agtggagagc  
3181 agcggtagta agctcttacc ttgcacaggt gactacatga acatgtcacc agtgggggac  
3241 tccaacacca gcagcccctc cgactgtctac tacggccctg aggaccccca gcacaagcca  
3301 gtctctctct actactcatt gccaaatcc ttaagcaca cccagcgcgc cggggagccg  
3361 gaggagggg cccggcatca gcacctccgc ctttccacta gctctggtcg ctttctctat  
3421 gctgcaacag cagatgattc ttctcttcc accagcagcg acagcctggg tgggggatac  
3481 tgcgggggta ggctggagcc cagccttcca catccccacc atcaggttct gcagcccat  
3541 ctgctcgaa aggtggacac agctgctcag accaatagcc gcctggcccc gccacgagg  
3601 ctgtccctgg gggatcccaa ggccagcacc ttacctcggg cccgagagca gcagcagcag  
3661 cagcagccct tgctgcaccc tccagagccc aagagcccgg gggaatatgt caatattgaa  
3721 tttgggagtg atcagtctgg ctactgtct ggcccggtag ctttccacag ctcaccttct  
3781 gtcaggtgtc catccagct ccagccagct cccagagagg aagagactgg cactgaggag  
3841 tacatgaaga tggacctggg gccgggcccg agggcagcct ggcaggagag cactggggtc  
3901 gagatgggca gactgggccc tgacctccc ggggctgcta gcatttgag gcctaccgg  
3961 gcagtgccca gcagccgggg tgactacatg accatgcaga tgagttgtcc ccgtcagagc  
4021 tacgtggaca cctcgcagc tgcccctgta agctatgctg acatgcgaac  
aggcattgct  
4081 gcagaggagg tgagcctgcc cagggccacc atggctgctg cctcctcatc ctcagcagcc  
4141 tctgcttccc cgactgggcc tcaaggggca gcagagctgg ctgcccactc gtcctgctg  
4201 gggggccac aaggacctgg gggcatgagc gccttcaccc ggggtaacct cagtcctaac  
4261 cgcaaccaga gtgcaaaagt gatccgtgca gaccacaag ggtgccggcg gaggcatagc  
4321 tccgagactt tctctcaac acccagtgcc acccgggtgg gcaacacagt gccctttgga  
4381 gcgggggag cagtaggggg cggtaggggt agcagcagca gcagcgagga  
tgtgaaacgc  
4441 cacagctctg cttccttga gaatgtgtg ctgaggcctg gggagcttg gggagcccc  
4501 aaggagccag ccaaactgtg tggggctgt gggggttg agaatgtct taactacata  
4561 gacctggatt tggtaagga ctcaaacag tgccctcagg agtgcacccc tgaaccgag

4621 cctccccac cccacccc tcatcaacc ctgggcagcg gtgagagcag ctccaccgc  
4681 cgctcaagtg aggatttag cgctatgcc agcatcagtt tccagaagca gccagaggac  
4741 cgtcagtagc tcaactggac atcacagcag aatgaagacc taaatgacct cagcaaatcc  
4801 tcttctaact catgggtacc cagactctaa atatttcag attcacaact aggacctcat  
4861 atcttctca tcagtagatg gtacgatgca tccattcag tttgttact ttatccaatc  
4921 ctcaggattt cattgactga actgcagctt ctatattgtg ccaagcgaaa aaaaaaatg  
4981 cactgtgaca ccagaataat gagtctgcat aaacttcac ttcaacctta aggacttagc  
5041 tggccacagt gagctgatgt gccaccacc gtgcatgag agaatgggtt tactctcaat  
5101 gcattttcaa gatactttc atctgctgt gaaactgtgt acgacaaagc atcattgtaa  
5161 attatttcac acaaaactgt tcacgttggg tggagagagt attaatatt taacataggt  
5221 tttgattat atgtgtaatt tttaaataa aatgtaact tttcttacag cacatcttt  
5281 ttttgatgt gggatggagg tatacaatgt tctgtttaa agagtggagc aatgcttaa  
5341 aacaaggctt aaaagagtag aatagggat gatcctgtt ttaagattgt aattcagaaa  
5401 acataatata agaatcatag tgccatagat ggttctcaat tgtatagta tatttgctga  
5461 tactatctct tgtcatataa acctgatgt gagctgagtt ccttataaga attaatctta  
5521 atttgtatt tttctctga agacaatagg ccatgtaat taaactgaag aaggatatat  
5581 ttgctgggt gtttcaaat gtcagctaa aattgtaat tgaatggaag caaattata  
5641 agaagaggaa attaaagtct tccattgcat gtattgtaa cagaaggaga tgggtgattc  
5701 ctcaattca aaagctctt ttggaatgaa caatgtgggc gtttgtaa tctggaaatg  
5761 tcttctatt cataataaac tagatactgt tgatcttta aaaaaaaaaa aaaaaaaaaa  
5821 aaaaaaaa //

**sequence ID 13**

MASPPESDGFSDVRKVGYLKPKSMHKRFFVLRAASEAGGPARL  
EYENEKKWRHKSSAPKRSIPLESCFNINKRADSKNKHLVALYTRDEHFIAAADSEAE  
QDSWYQALLQLHNRAKGHHDGAAALGAGGGGGSCSGSSGLGEAGEDLSYGDVPPG  
PAF  
KEVWQVILKPKGLGQTKNLIGIYRLCLTSKTISFVKLNSEAAAVVLQLMNIRRCGHSE  
NFFFIEVGRSAVTGPGEFWMQVDDSVVAQNMHETILEAMRAMSDEFPRRSKSQSS  
NC  
SNPISVPLRRHHLNPPPSQVGLTRRSRTESITATSPASMVGGKPGSFRVRASSDGEG  
TMSRPASVDGSPVSPSTNRTHAHRHRGSARLHPPLNHSRSIPMPASRCSPSATSPVSL

SSSSTSGHGSTSDCLFPRRSSASVSGSPSDGGFISSDEYGSSPCDFRSSFVTPDSL  
GHTPPARGEELSNYICMGGKGPSTLTAPNGHYILSRGGNGHRCTPGTGLGTSPALAG  
DEAASAADLDNRFRKRTHSAGTSPITTHQKTPSQSSVASIEEYTEMMPAYPPGGGSG  
G  
RLPGHRHSAFVPTRSYPEEGLEMHPLERRGGHHRPDSSTLHTDDGYMPMSPGVAPVP  
S  
GRKGS GDYMPMSPKSVSAPQQIINPIRRHPQRVDPNGYMMMSPSGGCSPDIGGGPS  
SS  
SSSSNAVPSGTSYGKLWTNGVGGHSHVLPHPKPPVESSGGKLLPCTGDYMNMSPV  
GD  
SNTSSPDCYYGPEDPQH KPVLSYYS LPRSFKHTQRPGEPEEGARHQHLRLSTSSGRL  
LYAATADSSSSSTSSDSLGGGYCGARLEPSLPHPHHQLVLPHLPRKVDTAAQTNSRL  
A  
RPTRLSLGDPKASTLPRAREQQQQQPLLHPPEPKSPGEYVNIEFGSDQSGYLSGPVA  
FHSSPSVRCPSQLQPAPREEETGTEEYMKMDLGPGRRAAWQESTGVEMGRLGPAPP  
A  
ASICRPTRAVPSSRGDYMTMQMSCPRQSYVDTSPAAPVSYADMRTGIAAEEVSLPRAT  
MAAASSSSAASASPTGPQGAELAHSLLGGPQGPGGMSAFTRVNLSPNRNQSAK  
VI  
RADPQGCRRRHSSETFSSTPSATRVGNTVPFGAGAAVGGGGGSSSSSEDVKRHSSA  
SF  
  
ENVWLRPGELGGAPKEPAKLCGAAGGLENLNYIDLVLKDFKQCPQECTPEPQPPPP  
PPPHQPLGSGESSSTRSSEDLSAYASISFQKQPEDRQ"

**sequence ID 14**

1 gcgggcgggg gcgggcgag cgggggctcg ggcgctggga gccggtggg ccgcaacgc  
61 agccgccacg ccggggccgc cgagatcggg tgcccgggat gagcctcatc cggaaaaag  
121 gcttctacaa gcaggacgct aacaagaccg cctgggagct gcccaagacc tacgtgtcc  
181 cgagcacgct cggcagcggg gcctatggct ccgtgtgctc ggccatcgac aagcggtcag  
241 gggagaaggt ggccatcaag aagctgagcc gacccttca gtccgagatc ttcgccaagc  
301 ggcctaccg ggagctgctg ctgctgaagc acatgcagca tgagaacgctc attgggctcc

361 tggatgtctt caccagcc tctccctgc gcaacttcta tgacttctac ctggtgatgc  
421 ccttcatgca gacggatctg cagaagatca tggggatgga gttcagtgag gagaagatcc  
481 agtacctggg gttcagatg ctcaaaggcc ttaagtacat ccactctgct ggggtcgtgc  
541 acagggacct gaagccaggc aacctggctg tgaatgagga ctgtgaactg aagattctgg  
601 attttgggct ggcgcgacat gcagacgccg agatgactgg ctacgtggtg acccgtggg  
661 accgagcccc cgaggtgatc ctacgtgga tgactataa ccagacagtg gacatctggt  
721 ctgtgggctg tatcatggca gagatgctga cagggaaac tctgttcaag gggaaagatt  
781 acctggacca gctgaccag atcctgaaag tgaccggggg gcttggcagc gagtttctgc  
841 agaagctgaa cgacaaagcg gccaaatcct acatccagtc cctgccacag acccccagga  
901 aggatttac tcagctgttc ccacgggcca gccccaggc tgcggacctg ctggagaaga  
961 tgctggagct agacgtggac aagcgcctga cggccgcgca ggcctcacc catccctct  
1021 ttgaaccctt cggggacct gaggaagaga cggaggcca gcagccgtt gatgattct  
1081 tagaacacga gaaactaca gtggatgaat ggaagcagca catctacaag gagattgtga  
1141 actttagccc cattgcccgg aaggactcac ggcgccggag tggcatgaag ctgtaggac  
1201 tcatcttgca tggcaccacc ggccagacac tgccaagga ccagtattg tactaccaa  
1261 actcagccct tcttgaata cagccttca agcagaggac agaagggctc ttctcttat  
1321 gtgggaaatg ggcctagtag atgcagaatt caaagatgac ggttgggaga aactagctct  
1381 gatcctaaca ggccacgta aactgcccac ctggagaatc gctgcaggt ggggccctt  
1441 ccttcccgc agagtggggc tgagtgggag ctgagccagg ccggggcct atggcagtg  
1501 tgctgtgtg gtttctagg gatgctctaa cgaattacca caaacctggg ggattgaaac  
1561 agcagaactt gattccctta cagttctgga ggctggaaat ctgggatgga  
gggtttgga  
1621 gggctgtggt cctttgaag gctctgggga agaactctc cttggctctt ttagcttgt  
1681 ggcggcagtg ggcagtcctt ggcattcccc agcttattgc tgatcactc cagtctctg  
1741 ctctctgtt ctctctctt ttaacaacag tcattggatt tagggccac ctaactctg  
1801 tgtgatctta tctgatcct tattaattaa acctgcaaat actctagttc caataaagt  
1861 cacattctca gttccaggt ggacatga

**sequence ID 15**

MSLIRKKGFKYQDVNKTAWELPKTYVSPHTVGSAYGSVCSAID

KRSGEKVAIKKLSRPFQSEIFAKRAYRELLLLKHMQHENVIGLLDVFTPASSLRNFYD  
FYLVMPFMQTDLQKIMGMEFSEEKIQYLVYQMLKGLKYIHSAGVVHRDLKPGNLAVNE  
DCELKILDFGLARHADAEMTGYVVTRWYRAPEVILSWMHYNQTVDIWSVGCIMAEML  
T  
GKTLFKGKDYLQDQILKVTGVPGTEFVQKLNDKAAKSYIQSLPQTPRKDFTQLFPR  
ASPQAADLLEKMLELDVDKRLTAAQALHPFFEPFRDPEEETEAQQPFDDSLHEKLT  
VDEWKQHIYKEIVNFSPIARKDSRRRSGMKL

**sequence ID 16**

1 gcgcgaggcc gtcgattcgc tcgcggctcc atcgcggcct ggccgggggg cgtgtctgc  
61 tgcgccaggt tcgctggccg cacgtctca ggtcctcctg ttctgggag gcgggcgcg  
121 caggactggg aggtggcggc agcgggag gactcgccga ggacggggct ccgccccggg  
181 ataaccaact ctcttctct cttctttggt gcttcccag gcggcgcg cggcggcg  
241 gagccggagc cttcgcgcg tccacgtccc tccccgctg caccggccc cggcgcgaga  
301 ggagagcgcg agagccccag ccgcgggcgg gcggggcg aagatggcag aggcaccggc  
361 ttccccggcc ccgctctctc cgctcgaagt ggagctggac ccggagttcg agccccagag  
421 ccgtccgca tctgtacgt ggcccctgca aaggccggag ctccaagcga gcctgcca  
481 gcctcgggg gagacggccg ccgactccat gatccccgag gaggaggacg atgaagacga  
541 cgaggacggc gggggacggg ccggctcggc catggcgatc ggcgcgcg gcgggagcgg  
601 cacgctggg tccgggctgc tcttgagga ctcggcccgg gtgctggcac ccggagggca  
661 agacccccgg tctgggcccag ccaccggcg gggcgggctg agcgggggta cacaggcgt  
721 gctgcagcct cagcaaccgc tgccaccgcc gcagccgggg gcggctgggg gctccgggca  
781 gccgagaaa tgttcgtgc ggcggaacgc ctggggaaac ctgtctacg cggacctgat  
841 caccggccc atcgagagct cccggacaa acggctcact ctgtcccaga tctacgagtg  
901 gatggtgctg tgcgtgccct actcaagga taagggcgac agcaacagct ctgccggctg  
961 gaagaactcc atccggcaca acctgtcact gcatagtcga tcatgcggg  
tccagaatga  
1021 gggaactggc aagagctctt ggtggatcat caaccctgat ggggggaaga gcgaaaagc  
1081 cccccggcgg cgggctgtct ccattggaca tagcaacaag tataccaaga gccgtggccg  
1141 cgagccaag aagaaggcag cctgcagac agccccgaa tcagctgacg acagtcctc  
1201 ccagctctcc aagtggctg gcagccccac gtcacgcagc agtgatgagc tggatgctg  
1261 gacggacttc cgttcacgca ccaattctaa cgccagcaca gtcagtggcc gctgtgcc  
1321 catcatggca agcacagagt tggatgaagt ccaggacgat gatgcgctc tctgcccat  
1381 gctctacagc agctcagcca gcctgtcacc ttcagtaagc aagccgtgca cggtggaact  
1441 gccacggctg actgatatgg caggcaccat gaatctgaat gatgggctga ctgaaaact  
1501 catggacgac ctgctggata acatcacgct cccgccatcc cagccatgc cactggggg  
1561 actcatgcag cggagctcta gcttcccga taccaccaag ggctcgggccc tgggctccc

1621 aaccagctcc ttaacagca cgggttcgg acctcatct ctgaactccc tacgccagtc  
1681 tccatgcag accatccaag agaacaagcc agctacctc tcttccatgt cacactatgg  
1741 taaccagaca ctccaggacc tgctcactc ggactcactt agccacagcg atgtcatgat  
1801 gacacagtcg gacccttga tgtctcaggc cagcaccgct gtgtctgccc agaattcccg  
1861 ccggaacgtg atgcttcgca atgatccgat gatgtccttt gctgcccagc ctaaccaggg  
1921 aagtttggtc aatcagaact tgctccacca ccagcacca acccagggcg ctcttggtgg  
1981 cagccgtgcc ttgtcgaatt ctgtcagcaa catgggcttg agtgagtcca gcagccttg  
2041 gtcagccaaa caccagcagc agtctcctgt cagccagtct atgcaaacc tctcggactc  
2101 tctctcaggc tctccttgt actcaactag tgcaaactg cccgtcatgg gccatgagaa  
2161 gttccccagc gacttgacc tggacatgtt caatgggagc ttggaatgtg acatggagtc  
2221 cattatccgt agtgaactca tggatgctga tgggttgat ttaactttg attcctcat  
2281 ctccacacag aatgttggtg gttgaacgt ggggaactc actggtgcta agcaggcctc  
2341 atctcagagc tgggtgccag gctgaaggat cactgaggaa ggggaagtgg gcaaagcaga  
2401 ccctcaaact gacacaagac ctacagagaa aacccttgc caaatctgt ctcagcaagt  
2461 ggacagtgat accgtttaca gcttaacacc tttgtgaatc ccacgccatt ttctaaacc  
2521 agcagagact gtaaatggcc cttaccctg ggtgaagcac ttacccttg aacagaactc  
2581 taaaagtat gcaaaatctt cttgtacag ggtggtgagc cgctgcccag tggaggacag  
2641 caccctcag caccaccac cctcattcag agcacaccgt gagccccgt cggccattct  
2701 gtggtgtttt aatattgca tggttatgg gacgtttaa gtgtgttct tgtgtttgt  
2761 ttctttgac tttctgagtt ttcacatgc attaactgc ggtattttc tgttaaatg  
2821 ttaaccgtcc tcccctagc aaatttaaaa acagaaagaa aatgtgtac  
cagttaccat  
2881 tccgggttcg agcatcaca gctttgagc gcatggaact ccataaacta acaaattaca  
2941 taaactaaag ggggatttc tttctctt tgttggtag aaaattatcc ttttcaaaa  
3001 actgaacaat ggcacaattg tttgctatgt gcacccgtcc aggacagaac cgtgcatagg  
3061 caaaaggagt ggagcacagc gtccggcca gtgtgttcc ggtctgagt cagggtgatc  
3121 tgtggacggg acccagcac caagtctacg ggtgccagat cagtagggcc tgtgattcc  
3181 tgtcagtgct ctcagctaat gtgaacagtg ttggtctgct ggtagaaac tagaatattg  
3241 atatttcag gaaagaaatc agctcagctc tccactcatt gccaaatgc actaaaggt  
3301 ttagtttaa ggagaaagaa aaggaaaaa aaaaaaaca aaaaagtcct gtttgcttt  
3361 gcagaacaaa tgaactaca ggtgagcatt aagcttgag tgagaaatgt gcgaagagta  
3421 aaaaccaag tcaatgctga ggcagttcta acttactgt tttctaaat acacatcctt

3481 gattatttc agccttgcta tataatctga tctgctagaa gtgtatgagt gagaggcaat  
3541 agcatacaaa ctgattttt aatataagc ttaggttgta attgtacaag tgaactcaatg  
3601 gaagtacaaa atagggcagt ttaactttt tttctgctt ctatggattt cattttgtg  
3661 tgtttcaaaa aagttatggt gctgtatagg tgctttctgt ttaacctgga aagtgtgatt  
3721 atattcgta cttctttgg tagacggaat agttgggacc accttgga cataagaaat  
3781 tggataacg atgctctgat tagcacagta tatgcatact tctcaaagt gatatatgaa  
3841 gactctttc ttgcataaa aagcattagg catataaatg tataaatata tttatcatg  
3901 tacagtacaa aatggaacc ttatgcatgg gccttaggaa tacaggctag tatttcagca  
3961 cagacttccc tgcttgagtt ctgctgatg ctgaccctg gacagtgggc accaacacag  
4021 acgtgccacc caaccctg cacacaccac cggccaccag gggccccctt gtgcgcttg  
4081 gctttataac tctctgggg gtgatattgg tggatcac agctcctagc ataagagag  
4141 ttccattgg taltgtcaca cgtctctgc ctgctggg ttgcatggt tgagcgatgg  
4201 ccctgtgat ttaccctgc ctttactga atctgaaat tgggtgcaa ttgtggtat  
4261 agtagactgt agcacattgc ctttctaaa ctgctacatg ttataatct tcattttaa  
4321 agtatgtgta attttttaa gtatgtattc tattcatatg gtctgctgt cagtgagcca  
4381 gacttgcta ctatattcct ttataaat gctagccact tctggattc tttagtaatg  
4441 tgctgatgc aagaacttc cagtagcagt gaaggagggt tgctctcca agcttctaa  
4501 gggatgctgc cctgtgtgg gatgcattgc agaggcacta gtagcatgg ggctagagtg  
4561 gggagcgaga tgtaaaagg tggggggata ggagaattcc agagtgttc cagcattagg  
4621 gtcctgagaa ctctgagtt cagagaaaca tgcaaagtga ctacaaaat agctactac  
4681 cttgagtt ttacagacc tgggagctg tttgggagtg agaaaggcaa  
ccctcaatg  
4741 tgttcaact taaaatggt gaattcttt cagacatggt atctcattt ttctcttt  
4801 ctacgcttg ttgaattca ggcagaatgt cttacagaat gtctagaac cagattatca  
4861 ttaactctga aacagctgag gaagggacag agaaggta agggaaggc agcacaaaac  
4921 agatcaggag aatgaagagg gaatgcttg gttttgtt ttgtttgt tttctttt  
4981 caagtaacta aacagcctc tacatgtaga gtgtgtgga gagctgagac cagggtaaag  
5041 tcaagtgcag catcagtact gcgagacca ccagcccctg gagagggtca gccgagaatc  
5101 tggtagtgaa gctgtctag ggtcccggca cctcacctc cagccactg cagagaggcc  
5161 agggccccag agactagcct ggttctgaag tgggcagggg tgctgccaga gccctctgc  
5221 ccttatgtg agaccctgct tttaggacag gccagccgtt ggccaccatg tcacattctg  
5281 agtgagtgtc acaggtccct aacaataatt ttctgatctg gagcatatca gcagaatgct

5341 tagcctcaag gggcctggca gctgtaatgt ttgatttatg atgagaacta tccgaggcca  
5401 cccttgccct ctaaataagc tgctctaggg agccgcctac ttttgatga gaaattagaa  
5461 gagtacctaa tgttgaaaac atgacatgcg ctctgggat ctgctgttct ctccagggt  
5521 ccagaacctg atacctgta ccaagctag gaaagagctt tatcacaagc ctctactgtc  
5581 ctggcatgag aactggctgc caggctcagt gtacccatt aactgtgaat gaatctgagc  
5641 ttggtttct ttattgctc ctctgcaata tgattgctga aacacattt aaaaattcag  
5701 aagctgtca ctctgttaa tgggaggatc agtcacacat gtgtagtaca aggcggactt  
5761 tgtgtttgtt tttggtgta attttagca ttgtgtgtgt tgctcccca cctgaggag  
5821 aggacaccat ggcttactac tcaggacaag tatgccccgc tcagggtgtg atttcagggt  
5881 gcttcaaac ttgtacgag ttaaagatg gtggggacag acttgcctc tacctagtga  
5941 acccactta agaataagg agcattgaa tctctggaa aaggccatga agaataaagc  
6001 agtcaaaaag aagtcctcca tgttggtgcc aaggactgc gaggggaaat aaaaatgta  
6061 tccagcctga ccaacatgga gaaaccccg ctccattaaa aatacaaat tagcctggca  
6121 tgggtggcga tgctgtaat ccagctact ctggaggctg aggcaggaga atcgctgaa  
6181 ccaggaggc ggaggtcga gtgagccgag atcatgccag tgactccag cctgggtaac  
6241 aagagtgaaa ctccgtgca aaaaaaaaa aaaaatgta ctatcctct ctgaaagcaa  
6301 aaaggaaacc ctaacagctc tgaactctgg tttattttt ctgctgtat ttgggtgaac  
6361 attgtatgat taggcataat gttaaaaaa aaaattttt tttgtagaa atgcaatcac  
6421 cagtaaagag gtacgaaaaa gctagcctct ctacagagacc ggggaggcag agtactacta  
6481 gaggaagtga agttctgatg gaatcatgcc tgtcaaatga ggtcttgaag cggatgccca  
6541 aataaaagag tatattttat ctaaatctta agtgggtaac atttatgca  
gtttaaataga  
6601 atggaatatt ttctcttgt ttagtgtat ctgtttgtat tttctttga tgaatgattg  
6661 gtcatgaggc ctctgccac actccagaaa tacgtgtgcg gctgcttta agaactatgt  
6721 gtctggcac ttatttctct aaaattatct cattgcctgg caatcagtct tctctgtat  
6781 acttgtccta gcacattatg tacatgggaa atgtaaaca atgtgaagga ggaccagaaa  
6841 aattagttaa tatttaaaaa aatgtattgt gcattttggc ttcacatgtt taacttttt  
6901 taagaaaaa gttgatgaa tggaaaaaa aatctgtata cagtatctgt aaaaactatc  
6961 ttatctgtt caattcctg ctcatatccc atataatcta gaactaaata tgggtgtggtg  
7021 ccatatntaa acactgaga gtcaagcagt tgagactttg atttgaagca cctcatcctt  
7081 cttcaatgc gaacactatc atatggcatt ctactgagg atttgtcta accatagtt  
7141 gccatgaatt aactctgcc ctttcttaa ggatcaaac cagtttgatt tgggaatctt

7201 ccctttcca aatgaaatag agatgcagta ctaacttc cttggtgtt gtagatattg  
7261 ccttggtat tccacttaa accgtaatct agttgtaa agagatggtg acgcatgaa  
7321 ataaagcatc agtgacactc t //

**sequence ID 17**

"MAEAPASPAPLSPLEVELDPEFEPQSRPRRCTWPLQRPELQASP  
AKPSGETAADSMIPEEEDDEDDEDGGGRAGSAMAIGGGGGSGTLGSGLLLED SARV  
LA  
PGGQDPGSGPATAAGGLSGGTQALLQPQQPLPPPQGAAGGSGQPRKCSSRRNAWG  
NL  
SYADLITRAIESSPKRLTLSQIYEMVRCVPYFKDKGDSNSSAGWKNSIRHNLSLHS  
RFMRVQNEGTGKSSWWIINPDGGKSGKAPRRRAVSMDNSNKYTKSRGRAAKKKAAL  
QT  
APESADDSPSQLSKWPGSPTSRSSELDAWTDFRSRTNSNASTVSGRLSPIMASTEL  
D  
EVQDDDAPLSPMLYSSSASLSPSVSKPCTVELPRLTDMAGTMNLNDGLTENLMDDL  
D  
NITLPPSQPSPTGGLMQRSSFPYTTKSGSLGSPSTSSFNSTVFGPSSLNSLRQSPMQT  
IQENKPATFSSMSHYGNQTLQDLLTSDSLSHSDVMMTQSDPLMSQASTAVSAQNSR  
RN  
VMLRNDPMMSFAAQPNQGSLVNQNLLHHQHQTQGALGGSRALSNVSNMGLSESS  
SLG  
  
SAKHQQQSPVSQSMQTLSDSLSGSSLYSTSANLPVMGHEKFPSDLDLDMFNGLSLEC  
DM  
ESIIRSELMADGLDFNFDLISTQNVVGLNVGNFTGAKQASSQSWVPG"

**sequence ID 18**

1 gcggtgcct tgcggcgag ctggggtcgc ggcctgctc cccgcgttt ctaaggccc  
61 gcgggcgcg caggagcggc actcgtggct gtggtggctt cggcagcggc ttcagcagat  
121 cggcggcatc agcggtagca ccagcactag cagcatggtg agccgggcag tgtgcggcac

181 cagcaggcag ctggctccg tttggggta tctgggctcc aggcagaagc acagcctccc  
241 cgacctgcc tacgactacg gcgcctgga acctcacatc aacgcgcaga tcatgcagct  
301 gcaccacagc aagcaccacg cggcctacgt gaacaacctg aacgtcaccg aggagaagta  
361 ccaggaggcg ttggccaagg gagatgttac agcccagata gctcttcagc ctgcactgaa  
421 gttcaatggt ggtggtcata tcaatcatag cattttctgg acaaacctca gccctaacgg  
481 tgggtggagaa cccaaagggg agttgctgga agccatcaaa cgtgactttg gttccttga  
541 caagtttaag gagaagctga cggctgcac tgttggtgtc caaggctcag gttggggttg  
601 gcttggttc aataaggaac ggggacactt acaaattgct gcttgccaa atcaggatcc  
661 actgcaagga acaacaggcc ttattccact gctggggatt gatgtgtggg agcacgctta  
721 ctacctcag tataaaaatg tcaggcctga ttatctaaaa gctatttga atgtaatcaa  
781 ctgggagaat gtaactgaaa gatacatggc ttgcaaaaag taaaccacga tcgttatgct  
841 gagtatgta agctctttat gactgtttt gtagtggtat agagtactgc agaatacagt  
901 aagctgctct attgtagcat ttctgatgt tgcttagtca ctatttcat aaacaacta  
961 atgttctgaa taatttcta ctaaacattt tgttattggg caagtgattg aaaatagtaa  
1021 atgctttgtg tgattgaatc tgattggaca ttttctcag agagctaaat tacaattgtc  
1081 attataaaa ccatcaaaaa tattccatcc atatacttg gggacttga gggatgcctt  
1141 tctagtcta ttctattgca gttatagaaa atctagtctt tgccccagt tacttaaaaa  
1201 taaaatatta acactttccc aagggaaca ctcggcttc tatagaaaat tgcactttt  
1261 gtcgagtaat cctctgcagt gatacttctg gtagatgtca cccagtgggt tttgttaggt  
1321 caaatgttcc tgtatagttt ttgcaaatag agctgtatac tgtttaaag tagcagggtga  
1381 actgaactgg ggtttgctca cctgcacagt aaaggcaaac ttcaacagca  
aaactgcaaa  
1441 aagggtggtt ttgcagtagg agaaaggagg atgtttattt gcagggcgcc aagcaaggag  
1501 aattgggcag ctcatgcttg agaccaatc tccatgatga cctacaagct agagtattta  
1561 aaggcagtgg taaatttcag gaaagcagaa gtt //

**sequence ID 19**

MLSRVCGTSRQLAPVLGYLGSRQKHSLPDLPHYDYGALEPHINA  
QIMQLHHSKHHAAYVNNLNVTEEKYQEALAKGDVTAQIALQPALKFNGGGHINHSIF  
W

TNLSPNGGGEPKGELLEAIKRDFGSFDKFKELTAASVGVQGGSGWGWLGFNKERGHL  
Q  
IAACPNQDPLQGTGLIPLLIDVWEHAYYLQYKNVRPDYLKAIWNVINWENVTERYM  
ACKK

**sequence ID 20**

1 mgigkskins cplslwgkr hsvdtspyh esdskksedl slcnvaehsn ttegtgkqe  
61 gaqsveemfe eeaeeflk fvilhaeddt dealrvqnll qddfikpgi ifaempcgrq  
121 hlqnlldavn gsawtilllt enflrdtwcn fqfytslmns vnrqhkynsv ipmrplnpl  
181 prertpfaql tinaleesr gfptqverif qesvyktqt iwketmmvq rqlia //

**sequence ID 21**

1 gtgtatcttt ggggggtggg cgccaacagc agtcaggcct gacaagcggc gacctccaag  
61 ggtgaggcct ctgcgggccc ccgactcacg cgcgtccggg ctctgcaagg gcggtgggga  
121 gcaggctgct gtggtcgcgg ggactgggtt gcggcgcgcc gcgtacggga cggcccaaaa  
181 ctctcgacgc ccggggcaag acgcccaccc cctgggctct ctcgctgggc cagaaaggaa  
241 gacagaaaag ccgcgggctg actgtggtgg cgctcgctg cagattgaaa agaaatgctg  
301 agaatacat aaagtttcc tcttctgct tggatattta taatgggtat cgggaagtct  
361 aaaataaatt cctgccctct ttctctct tggggtaaaa ggcacagtgt ggatacaagt  
421 ccaggatattc atgagtcaga ttccaagaag tctgaagatc tatccttg  
taatgttgct  
481 gagcacagca atacaacaga ggggccaaca ggaaagcagg agggagctca gagcgtggaa  
541 gagatgttg aagaagaagc tgaagaagag gtgttctca aattgtgat attgcatgca  
601 gaagatgaca cagatgaagc cctcagagtc cagaatctgc tacaagatga cttggtatc  
661 aaacccggaa taatcttgc tgagatgcca tgtggcagac agcatttaca gaatttagat  
721 gatgctgtaa atgggtctgc atggacaatc ttattactga ctgaaaactt ttaagagat  
781 acttggtgta atttccagtt ctatacgtcc ctaatgaact ccgtaacag gcagcataaa  
841 tacaactctg ttataccat gcgggcccctg aacaatcccc tccccgaga aaggactccc  
901 tttgccctcc aaaccatcaa tgccttagag gaagaaagtc gtggatttcc tacacaagta  
961 gaaagaattt tcaggagtc tgtgtataag acacaacaaa ctatatggaa agagacaaga  
1021 aatatgttac aaagacaatt tattgctga gatgaacat ataacatgtg gctggctctt

1081 gttttgtaaa ccaaagatt aatcttact tgagaaagca gtttctagga aatgtttaa  
1141 taaaagagag tcttcacctt aaagaaacct atggagcaca agaaagataa atttctgcag  
1201 gacagtctat aaaattgtgg tacttttga tgtttcagta aacttgacat tgcagagtt  
1261 tcaaggactt tctttcaca attttctag tcatggata tgaaaaagga atttcaatc  
1321 catattcctt gattgaacc tgaacaaaa acttgatga cagacattt taaaaatgtg  
1381 acaacacttt tattctctga atttgatct caaaggacac agaaaaaaaa tggccccagg  
1441 agatctgatc acacttctc ctgaggcacc tctcatggat gttgcaataa gcattcgggt  
1501 actatcacc agaaataga attgccagaa tagaacattt agcatgtaa gcgtgatgc  
1561 atataaatc agaaatagat gtgagaatgg tgaactttt taaaagaacc cagtcaatg  
1621 tattttctgc tgaaatctgc atattggag gcatttccca ccaccgattc //

**sequence ID 22**

PAMDVLPTGGGRPGLRTELEFRGGGGEARLESQEEETIPAAPPAPRLRGAERPRRSR  
DTWDGDEDTEPGEACGGRTSRTASLVSGLLNELYSCTEEEEAAGGGRGAEGRRRRR  
DSLDSSTEASGSDVVLGGRSGAGDSRVLQELQERPSQRHQMLYLRQKDANELKTILR  
ELKYRIGIQSAKLLRHLKQKDRLLHKVQRNCDIVTACLQAVSQKRRVDTKLKFTLEPSL  
GQNGFQQWYDALKAVARLSTGIPKEWRRKVWLTADHYLHSAIDWDKTMRTFNER  
SNPDDDSMGIQIVKDLHRTGCSSYCGQEAEQDRVVLKRVLLAYARWNKTVGYCQGF

NILAALILEVMEGNEGDALKIMIYLIDKVLPESYFVNNLRALSVDMAVFRDLLRMKLPEL  
SQHLDTLQRTANKESGGGYEPPLTNVFTMQWFLTLFATCLPNQTVLKIWDSVFFEGSE  
IILRVSLAIWAKLGEQIECCETADEFYSTMGRLTQEMLENDLLQSHELMQTVYSMAFPF  
FPQLAELREKYTYNITPPATVKPTSVSGRHSKARDSDEENDPDEDAVVNAVGLGP  
FSGFLAPELQKYQKQIKEPNEEQSLRSNNIAELSPGAINSCRSEYHAAFNSMMMERMT  
TDINALKRQYSRIKKKQQQVHQVYIRADKGPVTSILPSQVNSSPVINHLLLGKKMKM  
TNRAAKNAVIHIPGHTGGKISPVYEDLKTCLNSPWRTHIRVHKKNMPRTKSHPGCGD  
TIGLIDEQNEASKTNGLGAAEFPSGCTATAGREGSSPEGSTRRTIEGQSPEPVFGDA  
DVDVSAVQAKLGALELNQRDAAAETELRVHPPCQRHCPEPPSAPEENKATSKAPQGS  
NSKTPIFSPFSPVKPLRKSATARNLGLYGPTERTPTVHFPQMSRSFSKPGGGNSGTKKR

**sequence ID 23**

MAGNCGARGALSAHTLLFDLPPALLGELCAVLDSCDGALGWRGLVSKNKLCEGVLLI  
QGFVHVVVCKYKDSKYGCVALF  
EKTPSETYLHMET

**Sequence ID 24**

*Zinc finger 217 (ZNF217)*

1..5653, organism="Homo sapiens" mRNA"

1 gacaaagaga actaatgctt tgtgctgatt catatttgaa tcgaggcatt ggaaccctg  
61 tatgccttgt ttgtggaaag aaccagtgc accatcactg agcttcctaa aagttcgaag  
121 aagttagagg actatacact ttctttgaa ctttataat aatatttgc tctggtttt  
181 ggaaccagg gctgttagag ggtgagtg caagtcttac aagttgcctt attccaactc  
241 cagaaattgc ccaacggaac ttgagatta tatgcaatcg aaagtgcag gaaacatgcc  
301 aactcaatcc ctctaatgt acatggatgg gccagaagtg attggcagct ctctggcag  
361 tccgatggag atggaggatg cctgtcaat gaaagggacc gctgtgttc cattccgagc  
421 tacacaagaa aaaaatgca tccaatcga ggggtatatg cccttgatt  
gcatgttctg  
481 cagccagacc ttcacacatt cagaagacct taataacat gtcttaatgc aacaccggcc  
541 taccctctgt gaaccagcag ttctcgggt tgaagcagag tatctcagtc cgcttgataa  
601 aagtcaagtg cgaacagaac ctccaagga aaagaattgc aaggaaaatg  
aatttagctg  
661 tgaggtatgt gggcagacat ttagagtcgc tttgatggt gagatccaca tgagaacaca  
721 caagattct tcacttacg ggtgtaacat gtgcggaaga agattcaagg agccttggtt  
781 tcttaaaaat cacatgcgga cacataatgg caaatcgggg gccagaagca  
aactgcagca  
841 aggcttgag agtagtccag caacgatcaa cgaggtcgtc caggtgcacg  
cggccgagag  
901 catctctct cttacaaaa tctgcatggt ttgtggctc ctatttcaa ataaagaaag

961 tctaattgag caccgcaagg tgcacaccaa aaaaactgct ttcggtacca gcagcgcgca  
1021 gacagactct ccacaaggag gaatgccgtc ctcgagggag gacttctctc agttgttcaa  
1081 cttgagacca aaatctcacc ctgaaacggg gaagaagcct gtcagatgca tccctcagct  
1141 cgatccgttc accaccttcc aggcttgga gctggctacc aaaggaaaag ttgccatttg  
1201 ccaagaagtg aaggaatcgg ggcaagaagg gagcaccgac aacgacgatt  
cgagttccga

1261 gaaggagctt ggagaaacaa ataagggcag ttgtgcaggc ctctcgaag  
agaagagaa

1321 gtgcaaacac tcccacggcg aagcgccctc cgtggacgcg gatcccaagt  
taccagtag

1381 caaggagaag cccactcact gctccgagtg cggcaaagct ttcagaacct accaccagct  
1441 ggtcttgac tccaggtcc acaagaagga ccggagggcc ggcgcggagt  
cgcccacat

1501 gtctgtggac gggaggcagc cggggacgtg ttctctgac ctgcccgc ctctggatga  
1561 aaatggagcc gtggatcgag gggaaagtg ttctgaagac ggatctgagg  
atgggcttcc

1621 cgaaggaatc catctggata aaaatgatga tggaggaaaa ataaaacatc ttacatctc  
1681 aagagagtgt agttattgtg gaaagtttt ccgttcaa atattcatct  
1741 cagaacgcat acaggtgaaa aaccatacaa atgtgaattt tgtgaatag ctgcagcca  
1801 gaagacatct ctgaggtatc acttgagag acatcacaag gaaaaacaaa  
ccgatgttc

1861 tgctgaagtc aagaacgatg gtaaaaatca ggacactgaa gatgcactat  
taaccgctga

1921 cagtgcgcaa accaaaaatt tgaaaagatt tttgatggt gccaaagatg ttacaggcag  
1981 tccacctgca aagcagctta aggagatgcc ttctgtttt cagaatgttc tggcagcgc  
2041 tgtcctctca ccagcacaca aagatactca ggatttccat aaaaatgcag ctgatgacag  
2101 tgctgataaa gtgaataaaa accctacccc tgcttacctg gacctgtta aaaagagatc  
2161 agcagttgaa actcaggcaa ataacctcat ctgtagaacc aaggcggatg ttactcctcc  
2221 tccggatggc agtaccaccc ataacctga agttagcccc aaagagaagc  
aaacggagac

2281 cgcagctgac tgcagataca ggccaagtgt ggattgtcac gaaaaacctt taaattatc  
2341 cgtgggggct ctcacaatt gcccgcaat ttcttgagt aaaagttga ttccaagtat

2401 cacctgtcca tttgtacct tcaagacatt ttatccagaa gtttaatga tgcaccagag  
2461 actggagcat aaatacaatc ctgacgttca taaaaactgt cgaaacaagt ccttgcttag  
2521 aagtcgacgt accggatgcc cgccagcgtt gctgggaaaa gatgtgcctc ccctctctag  
2581 tttctgtaaa cccaagccca agtctgcttt cccggcgcag tccaaatccc tgccatctgc  
2641 gaaggggaag cagagccctc ctgggccagg caaggccctc ctgacttcag  
ggatagactc  
2701 tagcacttta gcccgaagta acctgaagtc ccacagacca cagcagaatg tgggggtcca  
2761 aggggccgcc accaggcaac agcaatctga gatgtttcct aaaaccagtg tttccctgc  
2821 accggataag acaaaaagac ccgagacaaa attgaaacct ctccagtag ctctttctca  
2881 gccacccctc ggcagcagta acatcaatgg ttccatcgac taccggcca agaacgacag  
2941 cccgtgggca cctccgggaa gagactattt ctgtaatcgg agtgccagca atactgcagc  
3001 agaatttggg gagccccttc caaaaagact gaagtccagc gtggttggcc ttgacgttga  
3061 ccagcccggg gccaaattaca gaagaggcta tgaccttccc aagtaccata  
tggtcagagg  
3121 catcacatca ctgttaccgc aggactgtgt gtatccgtcg caggcgtgc ctccaaacc  
3181 aaggttctctg agctccagcg aggtcgattc tccaaatgtg ctgactgttc agaagcccta  
3241 tgggtgctcc gggccacttt acattgtgt gctgctggt agtccagcat ccagctcgac  
3301 gttagaagga aaaaggcctg tgtcatatca acattatct aacagcatgg  
cacaaaagag  
3361 aaactatgag aattttattg ggaatgcaca ttatcgacca aatgacaaaa aaacttgatt  
3421 cactaattag ggggaaaaaa ggtcttggtg gatgtcagtg ctactcccc  
atgaaattaa  
3481 atttacttc atccttgag aagcgaatgg tgaaagctac tgaaataagc tgtgattgta  
3541 ctgtacataa aacatatgag gaatctgcaa ggaacactac agttgtgtaa agttgtctg  
3601 ttaacttttg taccaaatag caatacaaac tagttggaac agttggaact tacatacatg  
3661 gggactggaa atctctattt tgtccctgaa taatatttt cttagaattg accaaataag  
3721 aagtggaatt tttgcatact tgagcgtctc tgaaaagaaa tcattgggt tgggtggggc  
3781 gggatggggg aaaagtatat aatgcttgca cctcaggtaa aaatctgtaa atatctaagt  
3841 tgtaaacctg ctgttcaaa tactgtgtgt attcctttc tctaatgcag ctcatcactt  
3901 ggagcagttt ctgctattgt gctcttcat ttaaaatgta tgtttttt ttttaact  
3961 gtcaatgatt ttgtattatg ttgaatcac ccaaatctat tgtgtctta aaattgtaa  
4021 tggaagtatt gaccctctat gatatgtgct gcagatatcg aggtcagcca ttcggaagct

4081 ggcagcattt tatcgcaact ttgagcatct cagatgggga aggcaccttc ttctcgct  
4141 ctccagattg tcctggaacc tccaggatcc ttgactgagg gcgttgggat ggctttagg  
4201 attttaaga gagtgtgtct acagacaagc attttctctg tagagcagcc acacgttgta  
4261 taaacataaa ctgtatgtgc agttatttaa atttgttct gcaaaataa tcattttgt  
4321 tggacgattc agtggcgggg gggtttgcc taaattagta tataaaaaca aaaatgctaa  
4381 attatatctg tgaattgcag gtattgggga acagtttaa gggaaattht ggggggaaca  
4441 ttttaggtht gtattgga gtctaatgt atctggcatt tgggtgaact gtggacatac  
4501 tagagttgat tatagacaca ttgatttga ataaggaact gctggccgag cccgctggga  
4561 gtctagaaag agaaaactg tttctagacc tcagttattt tcccatttht ggttgttht  
4621 aagcagtaac atthttctca gtgcacatgc aatttgggtt ttagagaaga tggccaccag  
4681 ctggcttct agatathta aactthtgtt ctthaatag ctgtccatgg ctgagtht  
4741 tagtacatgg gcttagtgac cacaaaatat thattaaga aactgthta aaaataaht  
4801 tgcactgttc atthttctgg cctcgctgtt cccatagag caagggtaat cctagaaaa  
4861 atthththt thaaatht gcaacgtaag atgtctctct tgatagaagt cttagctct  
4921 gtgttacaag ggagaactca tttgagatca gtctgttggc attgcaatga agtgctht  
4981 atcaggaaag tgtacactat tgaccttht tctgttcac aagctgagcc atatgtacat  
5041 aatctagatt ttgtthcat agthttgcac thttatagcc taththtga gattaacaca  
5101 ttgcaagat gattgactca atcttgcct aatccaatga gtgttacaga gagcttgctg  
5161 tgactagaac cataaatctt aaaggggga tgtgataata gagggctgga atthaaact  
5221 gtathtaaaa aaaagaatca ccaaatctat ttgaaaaca gtcgattgt attagctgg  
5281 aaththtgg gctthcagat thctththt aaccacatt ctgaatgtat  
aaaaatacca  
5341 attaththcc tacagccctt tgtacttcaa aatagththt tgtgtccatc agtathtaact  
5401 attgtatac tactggtht atathththt thctthtga caacagtaca tataatagag  
5461 gtacaattcg ttgaththt gthtatgtat thattcatt ccagthttag thattthaat  
5521 tgttgatact taagttgtcg aacagtagac attactgtt thattatga tataththcag  
5581 ctthaaagta tgtattata tgggaagtg thaatataga thtggtht thcaaaaaa  
5641 aaaaaaaaaaaa aaa

**sequence ID 25**

zinc finger protein 217 organism="Homo sapiens" protein" 1..1048

1 mqskvtgnmp tqslmymdg pevigsslgs pmemedalsm kgtavvpfra  
tqeknvigie

61 gympldcmfc sqftthsedl nkhlvmqhrp ticepavlrv eaeylspldk sqvrteppke

121 knckenefsc evcgqtrva fdveihmrth kdsftygcnm cgrrfkepwf

lknhrthng

181 ksgarsklqq glesspatin evvqvhaaes isspykicmv cgflfpnkes liehrkvhtk

241 ktafgtssaq tdspqggmps sredflqlfn lrpkshtpetg kkpvrqipql dpfttfqawq

301 latkgkvaic qevkesgqeg stdnddssse kelgetnkgs caglsqekek

ckhshgeaps

361 vdadplpss kekpthcsec gkafryhql vlhsvhkdd rragaesptm svdgrqpgtc

421 spdlaapide ngavdrgegg sedgsedglp egihldkndd ggkikhltss recsycgkff

481 rsnylnihl rthtgekpyk cefceyaaaq ktslryhler hhkekqtava aevkndgknq

541 dtedalltad saqtknkrf fdgakdvtgs ppakqlkemp svfqnvlgsa vispahkdtq

601 dfhknaadds advknkntp ayldllkkr avetqannli crtkadvtpg pdgstthnie

661 vspkekqtet aadcryrpsv dchekplns vgalhncpai slsklipsi tcpfctfktf

721 ypevlmmhqr lehkynpdvh kncrnksllr srrtgcpal lgkdvpplss fckpkpksaf

781 paqsklpsa kgkqsppggp kapltsgids stlpsnlks hrpqqnvqvq

gaatrqqse

841 mfpktsvspa pdtkrpetk lkplpvapsq ptlgssning sidypaknds

pwappgrdyf

901 cnrsasntaa efgeplpkr kssvvaldvd qpganyrrgy dlpyhmvrg

itsllpqdcv

961 ypsqalppkp rflssevds pnvltvqkpy ggsgplytcv pagspasst legkrpvsyq

1021 hlsnsmaqkr nyenfignah yrpdkkt

## GENE SIGNATURES

### BACKGROUND AND SUMMARY

#### Background of the Invention

##### [0001] A. Field of the Invention

[0002] The present invention relates generally to a new cluster of correlating molecules in a tissue or at least one cell of a tissue for instance a cell of a blood tissue, preferably such myeloid cells and of identifying the condition of the genes expression said correlating molecules or of the expression levels of said molecules in a method or system for identifying the risk of obesity-related metabolic syndrome disorder phenotype characterized by dyslipidemia, hypertension, glucose intolerance, insulin resistance and diabetes, lipid homeostasis disorders and/or cardiovascular diseases This system of method provides information on how to modulate the correlating molecules to treat or prevent obesity and to treat or prevent the obesity-related metabolic syndrome disorders.

[0003] Several documents are cited throughout the text of this specification. Each of the documents herein (including any manufacturer's specifications, instructions etc.) are hereby incorporated by reference. However, there is no admission that any document cited is indeed prior art of the present invention.

##### [0004] B. Description of the Related Art

[0005] Currently more than 1 billion adults are overweight—and at least 300 million of them are clinically obese. This suggests the role of relatively modern environmental or behavioral risk factors such as high caloric intake or sedentary lifestyle.

[0006] The epidemic of obesity is a global health issue across all age groups, especially in industrialized countries (American Obesity Association, 2006). According to WHO's estimate there are more than 300 million obese people (BMI>30) world-wide. Today, for example almost 65% of adult Americans (about 127 million) are categorized as being overweight or obese. There is also evidence that obesity is increasing problem among children, for example in the USA, the percentage of overweight children (aged 5-14 years) has doubled in the last 30 years, from 15% to 32%. The degree of health impairment of obesity is determined by three factors: 1) the amount of fat 2) the distribution of fat and 3) the presence of other risk factors. It is the second leading cause of preventable death in the Western society and an increasing cause on modernizing societies. Obesity affects all major bodily systems—heart, lung, muscle and bones—and is considered as a major risk factor for several chronic disease conditions, including coronary heart disease, type 2 diabetes mellitus, hypertension, stroke, and cancers of the breast, endometrium, prostate and colon (Burton & Foster 1985).

[0007] A large number of medical conditions have been associated with obesity. Health consequences are categorized as being the result of either increased fat mass (osteoarthritis, obstructive sleep apnea, social stigma) or increased number of fat cells (diabetes, cancer, cardiovascular disease, non-alcoholic fatty liver disease) (Bray G A (2004). *J. Clin. Endocrinol. Metab.* 89 (6): 2583-9). Mortality is increased in obesity, with a BMI of over 32 being associated with a doubled risk of death (Manson J E, et al (1995). *N. Engl. J. Med.* 333 (11): 677-85). There are alterations in the body's response to insulin (insulin resistance), a pro-inflammatory state and an increased tendency to thrombosis (pro-thrombotic state) (Bray G A (2004). *J. Clin. Endocrinol. Metab.* 89 (6): 2583-

9). Disease associations may be dependent or independent of the distribution of adipose tissue. Central obesity (male-type or waist-predominant obesity, characterized by a high waist-hip ratio), is an important risk factor for the metabolic syndrome, the clustering of a number of diseases and risk factors that heavily predispose for cardiovascular disease. These are diabetes mellitus type 2, high blood pressure, high blood cholesterol, and triglyceride levels (combined hyperlipidemia) (Grundy S M (2004). *J. Clin. Endocrinol. Metab.* 89 (6): 2595-600. doi:10.1210/jc.2004-0372. PMID 15181029). Apart from the metabolic syndrome, obesity is also correlated with a variety of other complications. For some of these complaints, it has not been clearly established to what extent they are caused directly by obesity itself, or have some other cause (such as limited exercise) that causes obesity as well. Cardiovascular: congestive heart failure, enlarged heart and its associated arrhythmias and dizziness, varicose veins, and pulmonary embolism. Endocrine: polycystic ovarian syndrome (PCOS), menstrual disorders, and infertility (van der Steeg J W, Steures P, Eijkemans M J, et al (2008). "Obesity affects spontaneous pregnancy chances in subfertile, ovulatory women". *Hum. Reprod.* 23 (2): 324-8.) Gastrointestinal: gastroesophageal reflux disease (GERD), fatty liver disease, cholelithiasis (gallstones), hernia, and colorectal cancer. Renal and genitourinary: erectile dysfunction (Esposito K, et al. (2004). *JAMA* 291 (24): 2978-84. doi:10.1001/jama.291.24.2978) urinary incontinence, chronic renal failure, (Ejerblad E, Fored C M, Lindblad P, Fryzek J, McLaughlin J K, Nyrén O (2006). *J. Am. Soc. Nephrol.* 17 (6): 1695-702) hypogonadism (male), breast cancer (female), uterine cancer (female), stillbirth. Integument (skin and appendages): stretch marks, acanthosis nigricans, lymphedema, cellulitis, carbuncles, intertrigo. Musculoskeletal: hyperuricemia (which predisposes to gout), immobility, osteoarthritis, low back pain. Neurologic: stroke, meralgia paresthetica, headache, carpal tunnel syndrome, dementia, (Whitmer R A, et al. (2005). *BMJ* 330 (7504): 1360) idiopathic intracranial hypertension. Respiratory: obstructive sleep apnea, obesity hypoventilation syndrome, asthma. Psychological: Depression, low self esteem, body dimorphic disorder, social stigmatization.

[0008] The economic cost attributable to obesity is substantial and is close to \$100 billion/yr (Wolf & Colditz 1998). Obesity accounts for 2-6% of total health care costs in several developed countries; some estimates put the figure as high as 7%. The true costs are undoubtedly much greater as not all obesity-related conditions are included in the calculations.

[0009] There is thus a clear need in the art to have accurate predictive tools for the risk of obesity associated disorders and there is a need as well for accurate diagnostic tools that are indicative for a proper treatment of obesity associated disorders such as obesity associated metabolic syndrome, obesity related cardiovascular disorders and obesity related insulin resistance for person in need thereto.

[0010] Present invention provides such solution to these problems in the art.

#### SUMMARY OF THE INVENTION

[0011] In accordance with the purpose of the invention, as embodied and broadly described herein, the invention is broadly drawn to a method of diagnosis or a diagnostic tool for assessing the condition of a subject or for testing the efficacy of his or her treatment.

**[0012]** The present invention solves the problems of the related art by providing an accurate diagnosis tool for the progression of a metabolic syndrome disorder and in particular for the progression of a lipid homeostasis disorder, such as an impaired glucose tolerance and/or insulin resistance condition seen with said lipid homeostasis disorder; and/or the progression of an adipocyte tissue disorder, such as an impaired adipose tissue accumulation or adipocyte function; hereinafter also referred to as an obesity-related metabolic syndrome, both with increased risk to related cardiovascular diseases.

**[0013]** In particular this invention demonstrates that IRAK3 is a causal biomarker for obesity-related metabolic syndrome disorders.

**[0014]** Accordingly, in a first embodiment the present invention provides a method of diagnosing a metabolic syndrome disorder phenotype in a subject, i.e. determining whether a sample is from tissue of a mammal having a metabolic syndrome disorder phenotype, said method comprising: (a) determine or analyze the level of IRAK3 expression or activity of an IRAK3 expression product in a biological sample isolated from said subject, and (b) compare said level of expression or activity of said expression product with the IRAK3 expression or activity in a control sample; whereby a decreased level of IRAK3 expression or activity is indicative for the progression to an impaired glucose tolerance and/or insulin resistance condition seen with said metabolic syndrome disorder both with increased risk to related cardiovascular diseases.

**[0015]** As provided in more details in the examples hereinafter, in addition to IRAK3, further genes of interest in diagnosing the progression of a metabolic syndrome disorder and in discriminating lean from obese subjects, include SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2, and ZNF217.

**[0016]** Thus in a further aspect, the present invention provides a method of diagnosing a metabolic syndrome disorder phenotype in a subject, including discriminating lean from obese subjects, said method comprising: (a) obtaining an expression profile in a biological sample isolated from said subject, wherein said expression profile consists of the analysis of the level of IRAK3 expression or activity of an IRAK3 expression product in combination with the gene expression level or activity of a gene product of at least one gene selected from the group consisting of SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 and ZNF217; and (b) comparing said obtained expression profile to a reference expression profile to determine whether said sample is from subject having a metabolic syndrome disorder phenotype. A lower than reference expression profile of IRAK3 gene is indicative for said metabolic syndrome disorder phenotype. In a particular embodiment of the aforementioned method, the genes analyzed in the expression profile consist of any one of the following combinations; IRAK3 and SOD2; IRAK3 and TNFAIP6; IRAK3 and TNFAIP3; IRAK3 and ZNF217; IRAK3 and TNFAIP6 and TNFAIP3; IRAK3 and SOD2 and TNFAIP3; IRAK3 and SOD2 and TNFAIP6; IRAK3 and SOD2 and ZNF217; IRAK3 and TNFAIP3 and ZNF217; IRAK3 and TNFAIP6 and ZNF217; IRAK3 and SOD2 and TNFAIP3 and TNFAIP6; IRAK3 and SOD2 and TNFAIP3 and ZNF217; IRAK3 and SOD2 and TNFAIP6 and ZNF217; IRAK3 and TNFAIP3, TNFAIP6, SOD2 and ZNF217. In said embodiment, a lower than reference expression profile of IRAK3 and a higher than reference expression profile of SOD2,

TNFAIP3, TNFAIP6 and/or ZNF217 gene is indicative for said metabolic syndrome disorder phenotype.

**[0017]** In a particular embodiment of present invention, the metabolic syndrome disorder in the aforementioned methods is a lipid homeostasis disorder; more in particular a glucose intolerance and/or insulin resistance phenotype.

**[0018]** Yet another aspect of present invention is a method of diagnosing the risk for cardiovascular diseases in a subject, i.e. determining whether a sample is from tissue of a mammal having a high risk for cardiovascular diseases, said method comprising: (a) determine or analyze the level of IRAK3 expression or activity of an IRAK3 expression product in a biological sample isolated from said subject, and (b) compare said level of expression or activity of said expression product with the IRAK3 expression or activity in a control sample; whereby a decreased level of IRAK3 expression or activity is indicative for said risk for cardiovascular diseases. A higher than reference expression profile of SOD2, TNFAIP3, TNFAIP6 and/or ZNF217 gene is indicative for said cardiovascular diseases.

**[0019]** The present invention also concerns the above described methods for testing or evaluating whether a medication, a diet or nutraceutical is efficient in curing or decreasing the disorders.

**[0020]** Another aspect of present invention is a method of treating a disease, condition or disorder selected from the group consisting of (1) non-insulin dependent Type 2 diabetes mellitus (NIDDM), (2) hyperglycemia, (3) low glucose tolerance, (4) insulin resistance, (6) a lipid disorder, (7) dyslipidemia, (8) hyperlipidemia, (9) hypertriglyceridemia, (10) hypercholesterolemia, (11) low HDL levels, (12) high LDL levels, (13) atherosclerosis, comprising administering to subject in need thereof an effective amount of a compound which increases the expression of IRAK3 in the monocytes or macrophages. Increased expression of in blood monocytes macrophages can for instance be induced by hyaluronan or fragments thereof, lipopolysaccharide, or peptidoglycan stimulation

**[0021]** Further scope of applicability of the present invention will become apparent from the detailed description given hereinafter. However, it should be understood that the detailed description and specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description. It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention, as claimed.

## ILLUSTRATIVE EMBODIMENTS OF THE INVENTION

### Definitions

**[0022]** Myeloid refers to the nonlymphocytic groups of white blood cells, including the granulocytes, monocytes and platelets.

**[0023]** Dyslipidemia (From dys+lipid (fat)+emia (in the blood)=essentially, disordered lipids in the blood) is a disorder of lipoprotein metabolism. Dyslipidemias may be manifested by elevation of the triglyceride concentrations, and a decrease in the "good" high-density lipoprotein (HDL) cholesterol concentration in the blood. Dyslipidemia comes

under consideration in many situations including diabetes, a common cause of lipidemia. For adults with diabetes, it has been recommended that the levels HDL-cholesterol, and triglyceride be measured every year. Optimal HDL-cholesterol levels are equal to or greater than 40 mg/dL (1.02 mmol/L), and desirable triglyceride levels are less than 150 mg/dL (1.7 mmol/L).

**[0024]** Insulinemia concerns an abnormally large concentration of insulin in the blood.

**[0025]** Glycemia concerns the presence of glucose in the blood. It is a medical term meaning that the blood glucose is elevated, typically above 100 mg/dl.

**[0026]** Hypercholesterolemia is manifested by elevation of the total cholesterol due to elevation of the “bad” low-density lipoprotein (LDL) cholesterol in the blood. Optimal LDL-cholesterol levels for adults with diabetes are less than 100 mg/dL (2.60 mmol/L).

**[0027]** Triglycerides are the major form of fat. A triglyceride consists of three molecules of fatty acid combined with a molecule of the alcohol glycerol. Triglycerides serve as the backbone of many types of lipids (fats). Triglycerides come from the food we eat as well as from being produced by the body. Triglyceride levels are influenced by recent fat and alcohol intake, and should be measured after fasting for at least 12 hours. A period of abstinence from alcohol is advised before testing for triglycerides. Markedly high triglyceride levels (greater than 500 mg/dl) can cause inflammation of the pancreas (pancreatitis). Therefore, these high levels should be treated aggressively with low fat diets and medications, if needed. The word “triglyceride” reflects the fact that a triglyceride consists of three (“tri-”) molecules of fatty acid combined with a molecule of the alcohol glycerol (“-glyceride”) that serves as the backbone in many types of lipids (fats).

**[0028]** HDL-cholesterol concerns lipoproteins, which are combinations of lipids (fats) and proteins, are the form in which lipids are transported in the blood. The high-density lipoproteins transport cholesterol from the tissues of the body to the liver so it can be gotten rid of (in the bile). HDL-cholesterol is therefore considered the “good” cholesterol. The higher the HDL-cholesterol level, the lower the risk of coronary artery disease. Even small increases in HDL-cholesterol reduce the frequency of heart attacks. For each 1 mg/dl increase in HDL-cholesterol there is a 2 to 4% reduction in the risk of coronary heart disease. Although there are no formal guidelines, proposed treatment goals for patients with low HDL-cholesterol are to increase HDL-cholesterol to above 35 mg/dl in men and 45 mg/dl in women with a family history of coronary heart disease; and to increase HDL-cholesterol to approach 45 mg/dl in men and 55 mg/dl in women with known coronary heart disease. The first step in increasing HDL-cholesterol levels is life style modification. Regular aerobic exercise, loss of excess weight (fat), and cessation of cigarette smoking cigarettes will increase HDL-cholesterol levels. Moderate alcohol consumption (such as one drink a day) also raises HDL-cholesterol. When life style modifications are insufficient, medications are used. Medications that are effective in increasing HDL-cholesterol include nicotinic acid (niacin), gemfibrozil (Lopid), estrogen, and to a lesser extent, the statin drugs.

**[0029]** Hypertension or High blood pressure is defined as a repeatedly elevated blood pressure exceeding 140 over 90 mmHg—a systolic pressure above 140 with a diastolic pressure above 90. Chronic hypertension is a “silent” condition. Stealthy as a cat, it can cause blood vessel changes in the back

of the eye (retina), abnormal thickening of the heart muscle, kidney failure, and brain damage. For diagnosis, there is no substitute for measurement of blood pressure. Not having your blood pressure checked (or checking it yourself) is an invitation to hypertension. No specific cause for hypertension is found in 95% of cases. Hypertension is treated with regular aerobic exercise, weight reduction (if overweight), salt restriction, and medications.

**[0030]** Diabetes, type 2 is one of the two major types of diabetes, the type in which the beta cells of the pancreas produce insulin but the body is unable to use it effectively because the cells of the body are resistant to the action of insulin. Although this type of diabetes may not carry the same risk of death from ketoacidosis, it otherwise involves many of the same risks of complications as type 1 diabetes (in which there is a lack of insulin). The aim of treatment is to normalize the blood glucose in an attempt to prevent or minimize complications. People with type 2 diabetes may experience marked hyperglycemia, but most do not require insulin injections. In fact, 80% of all people with type 2 diabetes can be treated with diet, exercise, and, if needed be, oral hypoglycemic agents (drugs taken by mouth to lower the blood sugar). Type 2 diabetes requires good dietary control including the restriction of calories, lowered consumption of simple carbohydrates and fat with increased consumption of complex carbohydrates and fiber. Regular aerobic exercise is also an important method for treating both type 2 diabetes since it decreases insulin resistance and helps burn excessive glucose. Regular exercise also may help lower blood lipids and reduce some effects of stress, both important factors in treating diabetes and preventing complications. Type 2 diabetes is also known as insulin-resistant diabetes, non-insulin dependent diabetes, and adult-onset diabetes.

**[0031]** Systolic: The blood pressure when the heart is contracting. It is specifically the maximum arterial pressure during contraction of the left ventricle of the heart. The time at which ventricular contraction occurs is called systole. In a blood pressure reading, the systolic pressure is typically the first number recorded. For example, with a blood pressure of 120/80 (“120 over 80”), the systolic pressure is 120. By “120” is meant 120 mm Hg (millimeters of mercury). A systolic murmur is a heart murmur heard during systole, the time the heart contracts, between the normal first and second heart sounds. “Systolic” comes from the Greek systole meaning “a drawing together or a contraction.” The term has been in use since the 16th century to denote the contraction of the heart muscle.

**[0032]** Osteoarthritis is a type of arthritis caused by inflammation, breakdown, and eventual loss of cartilage in the joints. It is also known as degenerative arthritis.

**[0033]** An ischemic stroke is death of an area of brain tissue (cerebral infarction) resulting from an inadequate supply of blood and oxygen to the brain due to blockage of an artery. Ischemic stroke usually results when an artery to the brain is blocked, often by a blood clot or a fatty deposit due to atherosclerosis. Symptoms occur suddenly and may include muscle weakness, paralysis, lost or abnormal sensation on one side of the body, difficulty speaking, confusion, problems with vision, dizziness, and loss of balance and coordination. Diagnosis is usually based on symptoms and results of a physical examination, imaging tests, and blood tests. Treatment may include drugs to break up blood clots or to make blood less likely to clot and surgery, followed by rehabilitation. About one third of people recover all or most of normal

function after an ischemic stroke. Ischemic stroke occurs when local blood flow is suddenly limited by vessel occlusion. The rate of neuronal death varies with blood flow. If blood flow falls to less than 15 mL/100 g/min, energy failure and subsequent cell death occur within minutes. Even suboptimal flow for longer periods may cause the cells to die by an apoptotic mechanism over days to weeks. Rapid restoration of blood flow is essential to save brain tissue. The mechanism of stroke involving the PCA territory is variable. It is commonly due to embolization from the heart, the aortic arch, the vertebral artery, or the basilar artery. Other mechanisms include intrinsic atherosclerotic disease and vasospasm. Migrainous strokes tend to involve PCAs preferentially. Less commonly, the anterior circulation is to blame (e.g., internal carotid stenosis), when a fetal PCA is present. Rare causes of stroke may be considered when usual culprits such as coagulation abnormalities, vasculitis, sympathomimetic drugs, and metabolic disorders are not present.

**[0034]** Insulin resistance is the diminished ability of cells to respond to the action of insulin in transporting glucose (sugar) from the bloodstream into muscle and other tissues. Insulin resistance typically develops with obesity and heralds the onset of type 2 diabetes. It is as if insulin is “knocking” on the door of muscle. The muscle hears the knock, opens up, and lets glucose in. But with insulin resistance, the muscle cannot hear the knocking of the insulin (the muscle is “resistant”). The pancreas makes more insulin, which increases insulin levels in the blood and causes a louder “knock.” Eventually, the pancreas produces far more insulin than normal and the muscles continue to be resistant to the knock. As long as one can produce enough insulin to overcome this resistance, blood glucose levels remain normal. Once the pancreas is no longer able to keep up, blood glucose starts to rise, initially after meals, eventually even in the fasting state. Insulin resistance is an early feature and finding in the pathogenesis of type 2 diabetes associated with obesity is the development is insulin resistance, defined as impaired insulin-mediated glucose clearance in insulin-sensitive tissues (skeletal muscle, liver and adipose tissue) (Martin B C et al. *Lancet*. 1992 Oct. 17; 340(8825):925-9 and Warram J H et al, *Ann Intern Med*. 1990 Dec. 15; 113(12):909-15). Insulin resistance is the condition in which normal amounts of insulin are inadequate to produce a normal insulin response from fat, muscle and liver cells. Insulin resistance in fat cells reduces the effects of insulin and results in elevated hydrolysis of stored triglycerides in the absence of measures which either increase insulin sensitivity or which provide additional insulin. Increased mobilization of stored lipids in these cells elevates free fatty acids in the blood plasma. Insulin resistance in muscle cells reduces glucose uptake (and so local storage of glucose as (glycogen), whereas insulin resistance in liver cells reduces storage of glycogen, making it unavailable for release into the blood when blood insulin levels fall (normally only when blood glucose levels are at low storage: Both lead to elevated blood glucose levels. High plasma levels of insulin and glucose due to insulin resistance often lead to metabolic syndrome and type 2 diabetes, including its complications. In 2000, there were approximately 171 million people, worldwide, with diabetes. The numbers of diabetes patients will expectedly more than double over the next 25 years, to reach a total of 366 million by 2030 (WHO/IDF, 2004). The two main contributors to the worldwide increase in prevalence of diabetes are population ageing and urbaniza-

tion, especially in developing countries, with the consequent increase in the prevalence of obesity (WHO/IDF, 2004).

**[0035]** Cardiovascular diseases refer to the class of diseases that involve the heart or blood vessels (arteries and veins). While the term technically refers to any disease that affects the cardiovascular system, it is usually used to refer to those related to atherosclerosis (arterial disease). The circulatory system (or cardiovascular system) is an organ system that moves nutrients, gases, and wastes to and from cells, helps fight diseases and helps stabilize body temperature and pH to maintain homeostasis. While humans, as well as other vertebrates, have a closed circulatory system (meaning that the blood never leaves the network of arteries, veins and capillaries), some invertebrate groups have open circulatory system. The present diagnostic invention is particularly suitable for a cardiovascular disease of the group consisting of hypertension, coronary heart disease, heart failure, congestive heart failure, atherosclerosis, arteriosclerosis, stroke, cerebrovascular disease, myocardial infarction and peripheral vascular disease.

**[0036]** The following terms are similar, yet distinct, in both spelling and meaning, and can be easily confused: arteriosclerosis, arteriolosclerosis and atherosclerosis.

**[0037]** Arteriosclerosis also called hardening of the arteries chronic disease is characterized by abnormal thickening and hardening of the walls of arteries, with a resulting loss of elasticity. The major form of arteriosclerosis is atherosclerosis, in which plaques of consisting of macrophages, fatty deposits in foam cells, or atheromas, form on the inner walls of the arteries. These fatty acids are largely due to the uptake of oxidized LDL by macrophages. Arteriosclerosis is a general term describing any hardening (and loss of elasticity) of medium or large arteries (in Greek, “Arterio” meaning artery and “sclerosis” meaning hardening); arteriolosclerosis is arteriosclerosis mainly affecting the arterioles (small arteries); atherosclerosis is a hardening of an artery specifically due to an atheromatous plaque. Therefore, atherosclerosis is a form of arteriosclerosis. Arteriosclerosis (“hardening of the artery”) results from a deposition of tough, rigid collagen inside the vessel wall and around the atheroma. This increases the stiffness, decreases the elasticity of the artery wall.

**[0038]** Arteriolosclerosis (hardening of small arteries, the arterioles) is the result of collagen deposition, but also muscle wall thickening and deposition of protein (“hyaline”). Calcification, sometimes even ossification (formation of complete bone tissue) occurs within the deepest and oldest layers of the sclerosed vessel wall.

**[0039]** Atherosclerosis causes two main problems. First, the atheromatous plaques, though long compensated for by artery enlargement, eventually lead to plaque ruptures and stenosis (narrowing) of the artery and, therefore, an insufficient blood supply to the organ it feeds. If the compensating artery enlargement is excessive, a net aneurysm occurs. Atherosclerosis chronic disease is caused by the deposition of fats, cholesterol, calcium, and other substances in the innermost layer (endothelium) of the large and medium-sized arteries. Atherosclerosis is a disease affecting the arterial blood vessel. It is commonly referred to as a “hardening” or “furring” of the arteries. It is caused by the formation of multiple plaques within the arteries.

**[0040]** These complications are chronic, slowly progressing and cumulative. Most commonly, soft plaque suddenly ruptures (see vulnerable plaque), causing the formation of a thrombus that will rapidly slow or stop blood flow, e.g. 5

minutes, leading to death of the tissues fed by the artery. This catastrophic event is called an infarction. One of the most common recognized scenarios is called coronary thrombosis of a coronary artery causing myocardial infarction (a heart attack). Another common scenario in very advanced disease is claudication from insufficient blood supply to the legs, typically due to a combination of both stenosis and aneurismal segments narrowed with clots. Since atherosclerosis is a body wide process, similar events also occur in the arteries to the brain, intestines, kidneys, legs, etc.

**[0041]** Pathologically, the atheromatous plaque is divided into three distinct components: the nodular accumulation of a soft, flaky, yellowish material at the centre of large plaques composed of macrophages nearest the lumen of the artery; sometimes with underlying areas of cholesterol crystals; and possibly also calcification at the outer base of older/more advanced lesions.

**[0042]** Thrombogenicity refers to the tendency of a material in contact with the blood to produce a thrombus, or clot. It not only refers to fixed thrombi but also to emboli, thrombi which have become detached and travel through the bloodstream. Thrombogenicity can also encompass events such as the activation of immune pathways and the complement system. All materials are considered to be thrombogenic with the exception of the endothelial cells which line the vasculature. Certain medical implants appear non-thrombogenic due to high flow rates of blood past the implant, but in reality, all are thrombogenic to a degree. A thrombogenic implant will eventually be covered by a fibrous capsule, the thickness of this capsule can be considered one measure of thrombogenicity, and if extreme can lead to the failure of the implant.

**[0043]** Low-density lipoprotein (LDL) belongs to the lipoprotein particle family. Its size is approx. 22 nm and its mass is about 3 million Daltons; but, since LDL particles contain a changing number of fatty acids, they actually have a mass and size distribution. Each native LDL particle contains a single apolipoprotein B-100 molecule (Apo B-100, a protein with 4536 amino acid residues) that circles the fatty acids, keeping them soluble in the aqueous environment. In addition, LDL has a highly-hydrophobic core consisting of polyunsaturated fatty acid known as linoleate and about 1500 esterified cholesterol molecules. This core is surrounded by a shell of phospholipids and unesterified cholesterol as well as a single copy of B-100 large protein (514 kD)[Segrest, J. P. et al (September 2001 ture of apolipoprotein B-100 in low density lipoproteins). *Journal of Lipid Research* 42: 1346-1367]. Cholesterol is an animal sterol that is normally synthesized by the liver. The main types, low-density lipoprotein (LDL) and high-density lipoprotein (HDL) carry cholesterol from and to the liver, respectively. LDL-cholesterol concerns thus the cholesterol in low-density lipoproteins. Cholesterol is required in the membrane of mammalian cells for normal cellular function, and is either synthesized in the endoplasmic reticulum, or derived from the diet, in which case it is delivered by the bloodstream in low-density lipoproteins. These are taken into the cell by LDL receptor-mediated endocytosis in clathrin-coated pits, and then hydrolyzed in lysosomes. Oxidized LDL-cholesterol concerns a LDL-cholesterol that has been bombarded by free radicals; it is thought to cause atherosclerosis; the 'bad' cholesterol; a high level in the blood is thought to be related to various pathogenic conditions

**[0044]** Metabolic syndrome is a combination of medical disorders that increase the risk of developing cardiovascular disease and type 2 diabetes. It affects a large number of

people, and prevalence increases with age. Some studies estimate the prevalence in the USA to be up to 25% of the population. Metabolic syndrome is also known as metabolic syndrome X, syndrome X, insulin resistance syndrome, Reaven's syndrome or CHAOS. Metabolic syndrome components were defined as detailed in the Third Report of the National Cholesterol Education Program Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in adults (ATPIII) report: 1) waist circumference  $\geq 102$  cm in men and  $\geq 88$  cm in women; 2) fasting triglycerides  $\leq 150$  mg/dl (1.70 mmol/l); 3) HDL-cholesterol  $< 40$  mg/dl (1.03 mmol/l) in men and  $< 50$  mg/dl (1.29 mmol/l) in women; 4) blood pressure  $\geq 130/85$  mmHg or on anti-hypertensive medication; 5) fasting-glucose  $\geq 100$  mg/dl (5.55 mmol/l) or on anti-diabetic medication.

**[0045]** "Sample" or "biological sample" as used herein can be any organ, tissue, cell, or cell extract isolated from a subject, such as a sample isolated from a mammal having a metabolic syndrome disorder or at risk for a metabolic syndrome disorder (e.g., based on family history or personal history). For example, a sample can include, without limitation, cells or tissue (e.g., from a biopsy or autopsy), peripheral blood, whole blood, red cell concentrates, platelet concentrates, leukocyte concentrates, blood cell proteins, blood plasma, platelet-rich plasma, a plasma concentrate, a precipitate from any fractionation of the plasma, a supernatant from any fractionation of the plasma, blood plasma protein fractions, purified or partially purified blood proteins or other components, serum, tissue or fine needle biopsy samples, or any other specimen, or any extract thereof, obtained from a patient (human or animal), test subject, healthy volunteer, or experimental animal. A subject can be a human, rat, mouse, non-human primate, etc. A sample may also include sections of tissues such as frozen sections taken for histological purposes. A "sample" may also be a cell or cell line created under experimental conditions, that is not directly isolated from a subject.

**[0046]** In a particular embodiment the sample is selected from the group consisting of (a) a liquid containing cells; (b) a tissue-sample; (c) a cell-sample and (d) a cell biopsy; more in particular the sample comprises hematopoietic cells or blood cells; even more in particular the sample comprises at least one myeloid cell or debris thereof. In an even further embodiment the sample comprises at least one of monocytes or peripheral blood mononuclear cells or debris thereof.

**[0047]** A "control" or "reference" includes a sample obtained for use in determining base-line expression or activity. Accordingly, a control sample may be obtained by a number of means including from subjects not having a metabolic syndrome disorder; from subjects not suspected of being at risk for developing a metabolic syndrome disorder; or from cells or cell lines derived from such subjects. A control also includes a previously established standard, such as a previously characterized pool of RNA or protein extracts from monocytes of at least 20 subjects without any of the metabolic syndrome components as defined above. Accordingly, any test or assay conducted according to the invention may be compared with the established standard and it may not be necessary to obtain a control sample for comparison each time.

**[0048]** Unless defined otherwise, 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. Singleton et al., *Dictionary of Microbiol-*

ogy and Molecular Biology 2nd ed., J. Wiley & Sons (New York, N.Y. 1994), Microarrays in Clinical Diagnostics (© 2005 Humana Press Inc.) provide one skilled in the art with a general guide to many of the terms used in the present application.

**[0049]** For purposes of the present invention, the following terms are defined below.

**[0050]** The term “array” or “microarray” in general refers to an ordered arrangement of hybridizable array elements such as polynucleotide probes on a substrate. An “array” is typically a spatially or logically organized collection, e.g., of oligonucleotide sequences or nucleotide sequence products such as RNA or proteins encoded by an oligonucleotide sequence. In some embodiments, an array includes antibodies or other binding reagents specific for products of a candidate library. The array element may be an oligonucleotide, DNA fragment, polynucleotide, or the like, as defined below. The array element may include any element immobilized on a solid support that is capable of binding with specificity to a target sequence such that gene expression may be determined, either qualitatively or quantitatively.

**[0051]** When referring to a pattern of expression, a “qualitative” difference in gene expression refers to a difference that is not assigned a relative value. That is, such a difference is designated by an “all or nothing” valuation. Such an all or nothing variation can be, for example, expression above or below a threshold of detection (an on/off pattern of expression). Alternatively, a qualitative difference can refer to expression of different types of expression products, e.g., different alleles (e.g., a mutant or polymorphic allele), variants (including sequence variants as well as post-translationally modified variants), etc. In contrast, a “quantitative” difference, when referring to a pattern of gene expression, refers to a difference in expression that can be assigned a value on a graduated scale, (e.g., a 0-5 or 1-10 scale, a ++++ scale, a grade 1 grade 5 scale, or the like; it will be understood that the numbers selected for illustration are entirely arbitrary and in no-way are meant to be interpreted to limit the invention). Microarrays are useful in carrying out the methods disclosed herein because of the reproducibility between different experiments. DNA microarrays provide one method for the simultaneous measurement of the expression levels of large numbers of genes. Each array consists of a reproducible pattern of capture probes attached to a solid support. Labeled RNA or DNA is hybridized to complementary probes on the array and then detected for instance by laser scanning. Hybridization intensities for each probe on the array are determined and converted to a quantitative value representing relative gene expression levels. See the patent publications Nos. U.S. Pat. No. 6,040,138, U.S. Pat. No. 5,800,992 and U.S. Pat. No. 6,020,135, U.S. Pat. No. 6,033,860, U.S. Pat. No. 6,344,316, U.S. Pat. No. 7,439,346, U.S. Pat. No. 7,371,516, U.S. Pat. No. 7,353,116, U.S. Pat. No. 7,348,181, U.S. Pat. No. 7,347,921, U.S. Pat. No. 7,335,762, U.S. Pat. No. 7,335,470, U.S. Pat. No. 7,323,308, U.S. Pat. No. 7,321,829, U.S. Pat. No. 7,302,348, U.S. Pat. No. 7,276,592, U.S. Pat. No. 7,264,929, U.S. Pat. No. 7,244,559, U.S. Pat. No. 7,221,785, U.S. Pat. No. 7,211,390, U.S. Pat. No. 7,189,509, U.S. Pat. No. 7,138,506, U.S. Pat. No. 7,052,842, U.S. Pat. No. 7,047,141 and U.S. Pat. No. 7,031,845 which are incorporated herein by reference. High-density oligonucleotide arrays are particularly useful for determining the gene expression profile for a large number of RNA's in a sample.

**[0052]** A “DNA fragment” includes polynucleotides and/or oligonucleotides and refers to a plurality of joined nucleotide units formed from naturally-occurring bases and cyclofuranosyl groups joined by native phosphodiester bonds. This term effectively refers to naturally-occurring species or synthetic species formed from naturally-occurring subunits. “DNA fragment” also refers to purine and pyrimidine groups and moieties which function similarly but which have no naturally-occurring portions. Thus, DNA fragments may have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species. They may also contain altered base units or other modifications, provided that biological activity is retained. DNA fragments may also include species that include at least some modified base forms. Thus, purines and pyrimidines other than those normally found in nature may be so employed. Similarly, modifications on the cyclofuranose portions of the nucleotide subunits may also occur as long as biological function is not eliminated by such modifications.

**[0053]** The term “polynucleotide,” when used in singular or plural generally refers to any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. Thus, for instance, polynucleotides as defined herein include, without limitation, single- and double-stranded DNA, DNA including single- and double-stranded regions, single- and double-stranded RNA, and RNA including single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or include single- and double-stranded regions. In addition, the term “polynucleotide” as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are “polynucleotides” as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritiated bases, are included within the term “polynucleotides” as defined herein. In general, the term “polynucleotide” embraces all chemically, enzymatically and/or metabolically modified forms of unmodified polynucleotides, as well as the chemical forms of DNA and RNA characteristic of cells, including simple and complex cells.

**[0054]** The term “oligonucleotide” refers to a relatively short polynucleotide, including, without limitation, single-stranded deoxyribonucleotides, single- or double-stranded ribonucleotides, RNA: DNA hybrids and double-stranded DNAs. Oligonucleotides, such as single-stranded DNA oligonucleotides, are often synthesized by chemical methods, for example using automated oligonucleotide synthesizers that are commercially available. However, oligonucleotides can be made by a variety of other methods, including in vitro recombinant DNA-mediated techniques and by expression of DNAs in cells.

**[0055]** The terms “differentially expressed gene,” “differential gene expression” and their synonyms, which are used interchangeably, refer to a gene whose expression is activated to a higher or lower level in a subject, relative to its expression in a normal or control subject. A differentially expressed gene may be either activated or inhibited at the nucleic acid level or

protein level, or may be subject to alternative splicing to result in a different polypeptide product. Such differences may be evidenced by a change in mRNA levels, surface expression, secretion or other partitioning of a polypeptide, for example. Differential gene expression may include a comparison of expression between two or more genes, or a comparison of the ratios of the expression between two or more genes, or even a comparison of two differently processed products of the same gene, which differ between normal subjects and subjects suffering from a disease, or between various stages of the same disease. Differential expression includes both quantitative, as well as qualitative, differences in the temporal or cellular expression pattern in a gene or its expression products. As used herein, "differential gene expression" can be present when there is, for example, at least an about a one to about two-fold, or about two to about four-fold, or about four to about six-fold, or about six to about eight-fold, or about eight to about ten-fold, or greater than about 11-fold difference between the expression of a given gene in a patient of interest compared to a suitable control. However, folds change less than one is not intended to be excluded and to the extent such change can be accurately measured, a fold change less than one may be reasonably relied upon in carrying out the methods disclosed herein.

**[0056]** In some embodiments, the fold change may be greater than about five or about 10 or about 20 or about 30 or about 40.

**[0057]** The phrase "gene expression profile" as used herein, is intended to encompass the general usage of the term as used in the art, and generally means the collective data representing gene expression with respect to a selected group of two or more genes, wherein the gene expression may be upregulated, downregulated, or unchanged as compared to a reference standard. A gene expression profile is obtained via measurement of the expression level of many individual genes. The expression profiles can be prepared using different methods. Suitable methods for preparing a gene expression profile include, but are not limited to reverse transcription loop-mediated amplification (RT-LAMP), for instance one-step RT-LAMP, quantitative RT-PCR, Northern Blot, in situ hybridization, slot-blotting, nuclease protection assay, nucleic acid arrays, and immunoassays. The gene expression profile may also be determined indirectly via measurement of one or more gene products (whether a full or partial gene product) for a given gene sequence, where that gene product is known or determined to correlate with gene expression.

**[0058]** The phrase "gene product" is intended to have the meaning as generally understood in the art and is intended to generally encompass the product(s) of RNA translation resulting in a protein and/or a protein fragment. The gene products of the genes identified herein may also be used for the purposes of diagnosis or treatment in accordance with the methods described herein.

**[0059]** A "reference gene expression profile" as used herein, is intended to indicate the gene expression profile, as defined above, for a pre selected group which is useful for comparison to the gene expression profile of a subject of interest. For example, the reference gene expression profile may be the gene expression profile of a single individual known to not have an metabolic syndrome disorder phenotype or a propensity thereto (i.e. a "normal" subject) or the gene expression profile represented by a collection of RNA samples from "normal" individuals that has been processed as

a single sample. The "reference gene expression profile" may vary and such variance will be readily appreciated by one of ordinary skill in the art.

**[0060]** The phrase "reference standard" as used herein may refer to the phrase "reference gene expression profile" or may more broadly encompass any suitable reference standard which may be used as a basis of comparison with respect to the measured variable. For example, a reference standard may be an internal control, the gene expression or a gene product of a "healthy" or "normal" subject, a housekeeping gene, or any unregulated gene or gene product. The phrase is intended to be generally non-limiting in that the choice of a reference standard is well within the level of skill in the art and is understood to vary based on the assay conditions and reagents available to one using the methods disclosed herein.

**[0061]** "Gene expression profiling" as used herein, refers to any method that can analyze the expression of selected genes in selected samples.

**[0062]** The phrase "gene expression system" as used herein, refers to any system, device or means to detect gene expression and includes diagnostic agents, candidate libraries, oligonucleotide sets or probe sets.

**[0063]** The terms "diagnostic oligonucleotide" or "diagnostic oligonucleotide set" generally refers to an oligonucleotide or to a set of two or more oligonucleotides that, when evaluated for differential expression their corresponding diagnostic genes, collectively yields predictive data.

**[0064]** Such predictive data typically relates to diagnosis, prognosis, selection of therapeutic agents, monitoring of therapeutic outcomes, and the like. In general, the components of a diagnostic oligonucleotide or a diagnostic oligonucleotide set are distinguished from oligonucleotide sequences that are evaluated by analysis of the DNA to directly determine the genotype of an individual as it correlates with a specified trait or phenotype, such as a disease, in that it is the pattern of expression of the components of the diagnostic oligonucleotide set, rather than mutation or polymorphism of the DNA sequence that provides predictive value. It will be understood that a particular component (or member) of a diagnostic oligonucleotide set can, in some cases, also present one or more mutations, or polymorphisms that are amenable to direct genotyping by any of a variety of well known analysis methods, e.g., Southern blotting, RFLP, AFLP, SSCP, SNP, and the like.

**[0065]** The phrase "gene amplification" refers to a process by which multiple copies of a gene or gene fragment are formed in a particular cell or cell line. The duplicated region (a stretch of amplified DNA) is often referred to as "amplicon." Usually, the amount of the messenger RNA (mRNA) produced, i.e., the level of gene expression, also increases in the proportion of the number of copies made of the particular gene expressed.

**[0066]** A "gene expression system" refers to any system, device or means to detect gene expression and includes diagnostic agents, candidate libraries oligonucleotide, diagnostic gene sets, oligonucleotide sets, array sets, or probe sets.

**[0067]** As used herein, a "gene probe" refers to the gene sequence arrayed on a substrate.

**[0068]** As used herein, a "nucleotide probe" refers to the oligonucleotide, DNA fragment, polynucleotide sequence arrayed on a substrate.

**[0069]** The terms "splicing" and "RNA splicing" are used interchangeably and refer to RNA processing that removes

introns and joins exons to produce mature mRNA with continuous coding sequence that moves into the cytoplasm of a eukaryotic cell.

**[0070]** “Stringency” of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to re-anneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., *Current Protocols in Molecular Biology*, Wiley Interscience Publishers, (1995) and in *Current Protocols in Molecular Biology* Copyright© 2007 by John Wiley and Sons, Inc., 2008.

**[0071]** As used herein, a “gene target” refers to the sequence derived from a biological sample that is labeled and suitable for hybridization to a gene probe affixed on a substrate and a “nucleotide target” refers to the sequence derived from a biological sample that is labeled and suitable for hybridization to a nucleotide probe affixed on a substrate.

**[0072]** The term “treatment” refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented. The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology and biochemistry, which are within the skill of the art.

#### DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION

**[0073]** The following detailed description of the invention refers to the accompanying drawings. Also, the following detailed description does not limit the invention. Instead, the scope of the invention is defined by the appended claims and equivalents thereof.

**[0074]** The risk for developing heart disease is directly related to the concomitant burden of obesity-related cardiovascular risk factors clustered in the metabolic syndrome (MetSyn)<sup>1</sup>: dyslipidemia (i.e. high triglycerides and low HDL-cholesterol), hypertension, and type 2 diabetes.

**[0075]** Persons with the metabolic syndrome (MetSyn) are at increased risk of developing coronary heart diseases (CHD) as well as increased mortality from CHD and any other cause<sup>2, 3</sup>.

**[0076]** The Third Report of the National Cholesterol Education Program Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in adults (ATPIII), draws attention to the importance of MetSyn and provides a working definition of this syndrome<sup>4</sup>.

**[0077]** Findings from the Third National Health and Nutrition Examination Survey showed that MetSyn is highly prevalent in the United States. Its prevalence has increased

from 6.7% among participants aged 20 to 29 years, to 43.5% and 42.0% for participants aged 60 to 69 years and aged at least 70 years, respectively<sup>5</sup>.

**[0078]** Over 75% of hypertension cases are reported to be directly attributable to obesity, and the risk of developing hypertension is five to six times greater in obese adult Americans age 20 to 45 compared to non-obese individuals of the same age. Obesity and insulin resistance, and the interaction between these two components, are associated with a high cardiovascular risk<sup>6, 7</sup>. As many as 90% of individuals with type 2 diabetes are overweight or obese.<sup>8</sup> Obesity-related type 2 diabetes is a leading cause of morbidity and mortality in western societies, and is quickly approaching pandemic proportions<sup>9</sup>. In addition to heart disease, obesity is reported to increase the risk of ischemic stroke independent of other risk factors, including age and systolic blood pressure. The incidence of osteoarthritis increases with BMI and is associated with arthritis of the hand, hip, back and, in particular, the knee. Increased weight adds stress to bones and joints due to increased load. Lastly, there is evidence that some cancers (endometrial, breast and colon) are associated with obesity.

**[0079]** Although obesity and insulin resistance, and the interaction between these two components, are associated with a high cardiovascular risk<sup>6, 7</sup>, the severity of insulinemia and glycaemia during the diabetic phase can only to a minor extent explain this increased cardiovascular risk.

**[0080]** Therefore, the pathogenic mechanisms which link obesity with type-2 diabetes and with cardiovascular risk remain to be elucidated. One possible link is inflammation. Indeed, obesity is associated with increased infiltration in the adipose tissue of monocytes/macrophages that also produce inflammatory chemokines.<sup>10</sup>

**[0081]** Increased oxidative stress causes increased monocyte infiltration and is an early instigator of MetSyn. Several findings support this hypothesis. We demonstrated that MetSyn is associated with elevated levels of circulating oxidized LDL (oxLDL), a marker of oxidative stress. High triglycerides, low HDL-cholesterol, and high glucose and insulin predicted elevated levels of oxLDL independent of LDL-cholesterol levels. The association between MetSyn and elevated levels of oxLDL has been confirmed in European and Japanese cohorts<sup>11-13</sup>. Persons with high oxLDL levels showed a greater disposition to myocardial infarction, adjusting for all established cardiovascular risk factors.<sup>14</sup> Two other studies confirmed that elevated levels of circulating oxLDL predict future cardiovascular events even after adjustment for traditional cardiovascular risk factors and C-reactive protein<sup>15, 16</sup>. Recently, we have shown that persons with high oxLDL showed a 4.5-fold greater disposition to future MetSyn after 5 years follow-up, adjusted for age, gender, race, study centre, cigarette smoking, BMI, physical activity, and LDL-cholesterol, little changed by further adjustment for C-reactive protein, and adiponectin. In particular, oxLDL predicted the development of obesity, dyslipidemia and pre-diabetes.

**[0082]** We hypothesized that the identification of a cluster of genes and associated proteins which are associated with inflammation and oxidative stress and of which the expression pattern is improved by weight loss that significantly reduces cardiovascular risk could lead to a better estimate of the risk for cardiovascular disease for obese persons. We started from the observation in obese miniature pigs on an atherogenic diet that toll-like receptor 2 (TLR2) was over expressed in plaque macrophages isolated by laser capture

micro dissection and correlated with atherosclerotic plaque complexity<sup>17</sup>. Then, we performed micro array analysis of RNA extracted from monocytes of obese women. Because we found that TLR2 was over expressed, we searched for genes that correlated with TLR2. Structural modeling predicted a cluster of genes that besides TLR2 contains the following genes and associated proteins: IL1 receptor-associated kinase 3 (IRAK3), Tumor Necrosis Factor (TNF)-Associated Factor 6 (TRAF6), the myeloid differentiation marker MYD88, TNF-alpha-induced protein 3 and 6 (TNFAIP3; TNFAIP6), the Insulin Receptor Substrate 2 (IRS2), mitogen-activated protein kinase 13 (MAPK13), the Forkhead Box O3A (FOXO3A), and superoxide dismutase 2 (SOD2). These genes and associated proteins form the backbone of a pathway that links the toll-like receptor-mediated inflammation with the protection against oxidative stress by means of SOD2. Recently, we identified a protein correlating with that cluster that may be of importance for regulating the expression of that cluster: zinc finger 217 (ZNF217).

**[0083]** Present application first summarizes knowledge about the function of relevant genes and associated proteins. Thereafter examples which support our findings that some of these predicted molecules indeed are novel (bio)markers of cardiovascular risk in association with obesity, lipid homeostasis disorder related cardiovascular disease and/or an impaired glucose tolerance condition and that some are even causal biomarkers.

#### The Genes in the New Cluster of Correlating Molecules

##### **[0084]** Toll-Like Receptors

**[0085]** The Toll like/Interleukin 1 receptor family consists of a large number of transmembrane proteins which are involved in host defense and have conserved intracellular domains. This superfamily is divided into 2 subgroups, based on the components of the extracellular domains: the Toll like receptors (TLRs) with leucine-rich repeats, and the Interleukin-1 receptors (IL1Rs) with immunoglobulin-like motifs. Signal transduction pathways in these receptor families ultimately lead to activation of members of the Rel and AP1 family of transcription factors. An important mediator in this pathway is IL1 receptor-associated kinase (IRAK1). Using a murine EST sequence encoding a polypeptide with significant homology to IRAK1 to screen a human phytohemagglutinin-activated peripheral blood leukocyte (PBL) cDNA library, Wesche et al.<sup>18</sup> isolated a full-length cDNA clone encoding a 596-amino acid protein. Sequence analysis revealed an overall sequence similarity of 30 to 40% with IRAK1 and IRAK2 as well as structural similarity in an N-terminal death domain and a central kinase domain. Northern blot analysis revealed expression of 3 mRNA transcripts (an approximately 8-kb doublet and a 2.5-kb species that matched the isolated cDNA) predominantly in PBL and the monocytic cell lines U937 and THP-1, in contrast to the other IRAKs that are expressed in most cell types. Because of the restriction of expression of this IRAK to monocytic cells, the authors termed the protein IRAKM, now called IRAK3. The IRAK3 (or IRAKM or Interleukin-1 Receptor-Associated Kinase 3 or Interleukin-1 Receptor-Associated Kinase M) gene consists of 12 exons spanning a region of approximately 60 kb in chromosome 12q14.3<sup>19</sup>. Like IRAK2, the expression of IRAK3 in THP-1 cells is upregulated in the presence of phorbol ester and ionomycin, which also induce differentiation of these cells into more mature macrophages<sup>18</sup>. IRAK-3 (IRAK-M) is a member of the interleukine-1 receptor-asso-

ciated kinase (IRAK) family. The IRAK family is implicated in the Toll-like receptor (TLR) and IL-1R signaling pathway. IRAK3 interacts with the myeloid differentiation (MYD) marker MYD88 and TRAF6 signaling proteins in a manner similar to the other IRAKs. However, Kobayashi et al.<sup>20</sup> showed that IRAK3, in contrast to other IRAKs, is induced upon TLR stimulation but negatively regulates TLR signaling. IRAK3<sup>-/-</sup> cells exhibited increased cytokine production upon TLR/IL1 stimulation and bacterial challenge, and Irakm<sup>-/-</sup> mice showed increased inflammatory responses to bacterial infection. Endotoxin tolerance, a protection mechanism against endotoxin shock, was significantly reduced in IRAKM<sup>-/-</sup> cells. Thus, the authors concluded that IRAK3 regulates TLR signaling and innate immune homeostasis. IRAK-2 and IRAK-M. Data with IRAK-M knockout mice have revealed that IRAK-M serves as a negative regulator of IL-1R/TLR signaling. In contrast to other member of the interleukine-1 receptor-associated kinase (IRAK) family, IRAK-1 and IRAK-4 IRAK-M do not possess any detectable kinase activity. Moreover IRAK-M expression is mainly restricted to cells of a myeloid origin.

**[0086]** The *Homo sapiens* interleukin-1 receptor-associated kinase 3 (IRAK3) mRNA has been deposited in the NCBI database as under the accession number ACCESSION NM\_007199, VERSION NM\_007199.1 (LOCUS:NM\_007199 2288 bp mRNA linear PRI 11-FEB-2008) with the nucleotide sequence as in sequence ID 1.

**[0087]** The *Homo sapiens* interleukin-1 receptor-associated kinase 3 (IRAK3) protein has been deposited in the NCBI database as under the accession number ACCESSION NP\_009130 VERSION NP\_009130.1 GI:6005792 (LOCUS NP\_009130 596 aa linear PRI 11-FEB-2008) with the amino acid sequence as in sequence ID 2. The isoform hPA5312.2 [924 aa] Entry created: 2007, 4 Feb. and as published by the Last update of 2007, 12 Dec. has been depicted in Seq. ID 22 and the IRAK3 10 kDa protein—Length: 93 AS in sequence ID 23.

**[0088]** Tumor Necrosis Factor-Associated Factor (TRAF)  
**[0089]** The transcription factor NF-kappa-B (NFκB) is also activated by many cytokines which signal through different cell surface receptors. Members of the Tumor Necrosis Factor-Associated Factor (TRAF) protein family have been implicated in the activation of this transcription factor by the tumor necrosis factor (TNF) superfamily. TRAF2 is required for activation of this transcription factor by 2 TNF receptors, TNFR1 and TNFR2, as well as CD40, and TRAF5 may be responsible for NFκB activation signaled by the lymphotoxin B receptor. Cao et al.<sup>21</sup> identified a new member of the TRAF family, designated TRAF6. They showed that when over expressed in cultured human cells, TRAF6 activates NFκB. A dominant-negative mutant of TRAF6 inhibited this activation signaled by interleukin-1 (IL1A). IL1A treatment of the same cells induced the association of TRAF6 with IRAK. The findings were interpreted as indicating that TRAF proteins function as signal transducers for distinct receptor families in that TRAF6 participates in IL1A signaling. TRAF6 is a signal transducer in the NFκB pathway that activates I-kappa-B kinase (IKK) in response to pro-inflammatory cytokines. Cell-permeable peptides with the TRAF6-binding motif inhibited TRAF6 signaling, which indicated their potential as therapeutic modulators. It was concluded that their studies identified a universal mechanism by which TRAF6 regulates several signaling cascades in adaptive immunity, innate immunity, and bone homeostasis. TRAF6 is required for IL1

signaling. IL1B stimulation failed to induce NFκB or JNK/SAPK activation in cells from Traf6<sup>-/-</sup> mice. Inducible nitrous oxide synthase (iNOS) production in response to TNF plus IFNG, but not to IL1, was intact in Traf6-deficient mice. King et al.<sup>22</sup> generated healthy mice lacking Traf6 specifically in T lymphocytes. At 10 to 12 weeks of age, these mice developed splenomegaly and lymphadenopathy, with increased B-cell and Cd4-positive T-cell numbers, but fewer Cd8-positive T cells. Histopathologic analysis showed systemic inflammation in multiple organs of mutant mice. Cd4-positive/Cd25-positive regulatory T cells were present and appeared functional in mutant mice, but proliferation of Traf6<sup>-/-</sup> T cells could not be suppressed by wild-type or mutant regulatory T cells, suggesting the presence of a responder T-cell mechanism necessary to render T cells susceptible to regulation. The resistance to suppression was accompanied by hyperactivation of PI3K-dependent pathways. It was concluded that TRAF6 is important in maintaining the balance between immune activation and suppression.

**[0090]** Variant (1) contains an additional segment in the 5' UTR compared to variant 2. Both variants encode the same protein. Source sequence(s) BC031052, BI463192 There are two transcript Variant: This variant (2) lacks a segment in the 5' UTR compared to variant 1. Both variants encode the same protein. Source sequence(s) BC<sub>031052</sub>, BI463192, U78798 Consensus CDS CCDS7901.1 UniProtKB/TrEMBL A8KAB3 UniProtKB/Swiss-Prot Q9Y4K3 The *Homo sapiens* TNF receptor-associated factor 6 (TRAF6) protein has been deposited in the NCBI database as under the accession number ACCESSION NP\_004611, VERSION NP\_004611.1 GI:4759254, DBSOURCE REFSEQ: accession NM\_004620.2 with the amino acid sequence as in sequence ID 3.

**[0091]** mRNA of *Homo sapiens* TNF receptor-associated factor 6 (TRAF6), transcript variant 2 has been deposited in the NCBI database as under the accession number ACCESSION NM\_004620 REGION: 248.1816 VERSION NM\_004620.2 GI:22027628 (LOCUS NM\_004620 1569 bp mRNA linear PRI 10-FEB-2008) with the nucleotide sequence as in sequence ID 4.

**[0092]** mRNA of *Homo sapiens* TNF receptor-associated factor 6 (TRAF6), transcript variant 1 has been deposited in the NCBI database as under the accession number ACCESSION BC031052 VERSION BC031052.1 GI:21410268 (LOCUS BC031052 2594 bp mRNA linear PRI 17-JUL-2006) with the nucleotide sequence as in sequence ID 5.

**[0093]** Myeloid Differentiation Marker MYD88

**[0094]** The myeloid differentiation (MYD) marker MYD88 was first characterized during a study of the early genetic responses of murine myeloid cells to various differentiation and growth inhibitory stimuli<sup>23</sup>. Myeloid differentiation primary response genes are activated in M1 myeloleukemic cells in response to IL6, which induces both growth arrest and terminal differentiation. Northern blot analysis revealed widespread expression of the gene in many adult mouse tissues, and RT-PCR detected MYD88 mRNA in T- and B-cell lines and differentiating embryonic stem cells. The broad expression pattern demonstrated that mouse MYD88 expression is not restricted to cells of myeloid lineage as was originally believed. Medzhitov et al.<sup>24</sup> showed that signaling by the human TLR4 employs an adaptor protein, MYD88, and induces activation of NFκB via IRAK1 and TRAF6 protein. These findings implicate MYD88 as a general adaptor/regulator molecule for the Toll/IL1R family of receptors for

innate immunity. Adachi et al.<sup>25</sup> observed that mice with a targeted disruption of the MYD88 gene were unable to respond to iL.L. as determined by defective T-cell proliferation and the production of cytokines. Likewise, MYD88-deficient mice were unable to produce IFNG and mediate natural killer cell activity in response to IL18. NFκB activation in response to IL1 or IL18 was also impaired. These results indicated that MYD88 is a critical component in the IL1R and IL18R signaling cascades. Kawai et al.<sup>26</sup> extended these studies to show that responses to lipopolysaccharide, mediated by TLR4 and CD14, were lost or delayed in MYD88-deficient mice, establishing that MYD88 is part of the TLR signaling cascade as well, acting just upstream of IRAK. Bjorkbacka et al.<sup>27</sup> examined atherosclerotic lesion development in uninfected Apoe single-null mice and Apoe<sup>-/-</sup> MYD88<sup>-/-</sup> double-null mice, and found that the MYD88-deficient mice showed a marked reduction in early atherosclerosis. Inactivation of the MYD88 pathway led to a reduction in atherosclerosis through a decrease in macrophage recruitment to the artery wall that was associated with reduced chemokine levels. The findings linked elevated serum lipid levels to a pro-inflammatory signaling cascade that is also engaged by microbial pathogens.

**[0095]** mRNA of *Homo sapiens* Myeloid differentiation primary response gene (88) MYD88 has been deposited in the NCBI database as under the accession number ACCESSION BC013589 VERSION BC013589.1 GI:15488922 (LOCUS BC013589, 2678 bp mRNA linear PRI 19-JUN-2006) with the mRNA nucleotide sequence as in sequence ID 6.

**[0096]** The MYD88 translation product of the MYD88 gene CDS 40.930 of clone "MGC: 9601 IMAGE: 3900951" is depicted in sequence ID 7.

**[0097]** TNF-Alpha-Induced Protein 3 (TNFAIP3)

**[0098]** TNF-alpha-induced protein 3 (TNFAIP3; or A20) is a cytoplasmic zinc finger protein that inhibits NFκB activity and TNF-mediated programmed cell death. TNF dramatically increases TNFAIP3 mRNA expression in all tissues<sup>28, 29</sup>. Cytokines such as TNF profoundly affect endothelial cell function, promoting, for example, interaction with leukocytes and inducing a procoagulant phenotype. Changes of this nature are likely to be central to the pro-inflammatory effects of TNF. Dixit et al.<sup>28</sup> analyzed TNF-induced primary response genes in human umbilical vein endothelial cells. Of the 6 induced cDNAs identified, 2 encoded paracrine factors, neutrophil chemotactic factor and monocyte chemotactic factor; 1 encoded a membrane receptor for neutrophils, endothelial leukocyte adhesion molecule 1 (ELAM1); and 3 encoded hitherto undescribed TNF primary response genes. On exposure of endothelial cells to TNF, there was a rapid and substantial increase in the levels of mRNA encoding the 6 genes, which were further superinduced by cycloheximide. Thus these represent primary response genes, as their induction does not depend on protein synthesis. One of the 3 new proteins, designated A20 (new name TNFAIP3), was found on sequence analysis to code for a novel zinc finger protein<sup>30</sup>.<sup>31</sup> expressed TNFAIP3 and TLR4 in a human embryonic kidney cell line and observed inhibition of NFκB activation after LPS stimulation. Mutation analysis showed that the C-terminal zinc finger domain of A20 was sufficient for NFκB inhibition, whereas the full-length protein was required for inhibition of AP1 activation, and for induction of IL8. They concluded that TNFAIP3 (A20) modulates TLR4 signaling at or downstream of MEKK1 (MAP3K1). Wertz et al.<sup>32</sup> demonstrated that A20 down regulates NFκB signaling through the cooperative activity of its 2 ubiquitin-editing

domains. The N-terminal domain of TNFAIP3, which is a deubiquitinating enzyme of the OTU (ovarian tumor) family, removes lysine-63-linked ubiquitin chains from receptor-interacting protein (RIP), an essential mediator of the proximal TNF receptor-1 (TNFR1) signaling complex. The C-terminal domain of TNFAIP3 composed of 7 C2/C2 zinc fingers, then functions as a ubiquitin ligase by polyubiquitinating RIP with lysine-48-linked ubiquitin chains, thereby targeting RIP for proteasomal degradation. They defined a novel ubiquitin ligase domain and identified 2 sequential mechanisms by which A20 downregulates NF $\kappa$ B signaling. They also provided an example of a protein containing separate ubiquitin ligase and deubiquitinating domains, both of which participate in mediating a distinct regulatory effect. Lee et al.<sup>29</sup> generated A20-deficient mice by targeted disruption. A20 $^{-/-}$  mice appeared normal without evidence of pathology. A20 $^{-/-}$  mice, born from interbred A20 $^{+/-}$  mice in Mendelian ratios, developed runting as early as 1 week of age. Mice deficient for A20 developed severe inflammation and cachexia, were hypersensitive to both lipopolysaccharide and TNF, and died prematurely. A20-deficient cells failed to terminate TNF-induced NF $\kappa$ B responses. These cells were also more susceptible than control cells to undergo TNF-mediated program cell death. Thus, A20 is critical for limiting inflammation by terminating TNF-induced NF $\kappa$ B responses in vivo. Using mice doubly deficient in either A20 and Tnf or A20 and Tnfr1, Boone et al.<sup>33</sup> showed that, in addition to terminating TNF-induced signals, A20 is required for terminating TLR (e.g., TLR4)-induced activity of NF $\kappa$ B. Mutation and immunoblot analyses indicated that A20 acts, via its conserved OTU like domain, as a deubiquitinating enzyme on ubiquitinated TRAF6. In mice subjected to aortic banding, Cook et al.<sup>34</sup> detected greater than 4-fold A20 upregulation (p less than 0.05) at 3 hours, coinciding with peak NF $\kappa$ B activation. Cardiomyocytes infected with an adenoviral vector (Ad) encoding A20 inhibited TNF-stimulated NF $\kappa$ B signaling with an efficacy comparable to dominant-negative inhibitor of kappa-B kinase-beta (IKK $\beta$ ). Ad-IKK $\beta$ -infected cardiomyocytes exhibited increased apoptosis when serum-starved or subjected to hypoxia-reoxygenation, whereas Ad-A20-infected cardiomyocytes did not. Expression of Ad-A20 inhibited the hypertrophic response in cardiomyocytes stimulated with phenylephrine or endothelin-1. They concluded that A20 is dynamically regulated during acute biomechanical stress in the heart and functions to attenuate cardiac hypertrophy through the inhibition of NF $\kappa$ B signaling without sensitizing cardiomyocytes to apoptosis.

**[0099]** mRNA of *Homo sapiens* TNF-alpha-induced protein 3 (TNFAIP3; or A20) has been deposited in the NCBI database as under the accession number **ACCESSION M59465 305610 VERSION M59465.1 GI:177865 (LOCUS HUMA20 4426 bp mRNA linear PRI 30-OCT-1994)** with the mRNA nucleotide sequence as in sequence ID 8.

**[0100]** The human tumor necrosis factor alpha inducible protein A20 translation product of the CDS 67.2439 of the TNFAIP1 gene is depicted in sequence ID 9.

#### TNF-Alpha-Induced Protein 6 (TNFAIP6)

**[0101]** Lee et al.<sup>35</sup> described a gene, which they designated TSG6 (current name TNFAIP6), that is transcribed in normal fibroblasts and activated by binding of TNF-alpha and IL1 at AP-1 and NF-IL6 sites in its promoter. The cDNA was isolated from a library made from TNF-treated human fibroblasts. TNFAIP6 is a member of the hyaluronan-binding pro-

tein family, which includes cartilage link protein, proteoglycan core protein, and the adhesion receptor CD44. The predicted polypeptide is 277 amino acids long and includes a typical cleavage signal peptide. TNFAIP6 is highly homologous to CD44, particularly in the hyaluronic acid-binding domain. Western blots with antibodies made to a TNFAIP6 fusion protein detected a 39-kD glycoprotein in TNF-treated cells, and hyaluronate binding was shown by co-precipitation. TNFAIP6 expression is rapidly activated by TNF-alpha, IL1, and lipopolysaccharide in normal fibroblasts, peripheral blood mononuclear cells, synovial cells, and chondrocytes.

**[0102]** mRNA of *Homo sapiens* tumor necrosis factor, alpha-induced protein 6 (TNFAIP6) has been deposited in the NCBI database as under the accession number **ACCESSION NM\_007115 VERSION NM\_007115.2 GI: 26051242 (LOCUS NM\_007115 1440 bp mRNA linear PRI 11-FEB-2008)** with the mRNA nucleotide sequence as in sequence ID 10.

**[0103]** The translated product from the CDS 77.910 of the TNFAIP6 gene is depicted in Sequence ID 11.

**[0104]** Insulin Receptor Substrates

**[0105]** The Insulin Receptor Substrate 1 (IRS1) acts as an interface between signaling proteins with Src homology-2 domains (SH2 proteins) and the receptors for insulin, IGF2, growth hormone, several interleukins, and other cytokines. It regulates gene expression and stimulates mitogenesis and appears to mediate insulin/IGF1-stimulated glucose transport. Thus, the finding that the homozygous Irs1 KO mouse survives with only mild resistance to hypertension was surprising. This dilemma was provisionally resolved by the discovery by Sun et al.<sup>36</sup> of a second IRS signaling protein in mouse. They purified and cloned a likely candidate from mouse myeloid progenitor cells and, because of its resemblance to IRS1, they designated it IRS2. Withers et al.<sup>37</sup> demonstrated that homozygous absence of the Irs2 gene results in type II diabetes in mice. Heterozygous and wild type animals were unaffected. The authors demonstrated profound insulin resistance in both skeletal muscle and liver in the homozygous Irs2 $^{-/-}$  mice. Male mice lacking the Irs2 locus showed polydipsia and polyuria without ketosis and died from dehydration and hyperosmolar coma. A similar disease progression was observed in female mice, with the exception that the females rarely died. The authors concluded that dysfunction of IRS2 may contribute to the pathogenesis of human type II diabetes. To be et al.<sup>38</sup> observed that Irs2-deficient mice showed increased adiposity with increased serum leptin level, suggesting leptin resistance before the mice developed diabetes. Using oligonucleotide micro array and Northern blot analyses to analyze gene expression they detected increased expression of SREBP1, a downstream target of insulin, in Irs2-deficient mouse liver. Using high dose leptin administration, they provided evidence that leptin resistance in Irs2-deficient mice is causally related to SREBP1 gene induction. The authors concluded that Irs2 gene disruption results in leptin resistance, causing SREBP1 gene induction, obesity, fatty liver, and diabetes. Taguchi et al.<sup>39</sup> showed that, in mice, less Irs2 signaling throughout the body or only in brain extended life span up to 18%. At 22 months of age, brain-specific Irs2 knockout mice were overweight, hyperinsulinemic, and glucose intolerant; however, compared with control mice, they were more active and displayed greater glucose oxidation, and during meals they displayed stable SOD2 concentrations in the hypothalamus.

Thus, they concluded that less Irs2 signaling in aging brains can promote healthy metabolism, attenuate meal-induced oxidative stress, and extend the life span of overweight and insulin-resistant mice.

**[0106]** mRNA of *Homo sapiens* insulin receptor substrate 1 (IRS1) has been deposited in the NCBI database as under the accession number **ACCESSION** NM\_005544 **VERSION** NM\_005544.1 **GI**: 5031804 (LOCUS NM\_005544 5828 bp mRNA linear PRI 16-MAR-2008) with the mRNA nucleotide sequence as in sequence ID 12.

**[0107]** The translation product of the CDS 1021.4749 IRS1 gene is depicted in sequence ID 13.

**[0108]** Mitogen-Activated Protein Kinase 13

**[0109]** Mitogen-activated protein kinase (MAPK) cascades represent one of the major signal systems used by eukaryotic cells to transduce extracellular signals into cellular responses. Goedert et al.<sup>40</sup> isolated cDNAs encoding a protein that they designated SAPK4 (current name MAPK13). The sequence of the predicted 365-amino acid protein is approximately 60% identical to those of SAPK3 (or p38-gamma), SAPK2a (or p38), and SAPK2b. Like those SAPKs, SAPK4 has a TGY dual phosphorylation motif and is activated in response to cellular stresses and pro-inflammatory cytokines. Wang et al.<sup>41</sup> also isolated SAPK4 cDNAs. Using Northern blot analysis, they compared the expression patterns of SAPK4 (designated p38-delta by them), SAPK2a, SAPK2b, and SAPK3 in 50 human tissues. They detected 2 SAPK4 transcripts: a predominant 1.8-kb mRNA and a lower abundance 6-kb mRNA. The highest levels of expression were observed in exocrine/endocrine tissues including salivary gland, pituitary gland, adrenal gland and placenta. Recently, it has been shown that MAP kinase activation is contributing to atherosclerosis<sup>42, 43</sup>.

**[0110]** mRNA of *Homo sapiens* mitogen-activated protein kinase 13 (MAPK13), has been deposited in the NCBI database as under the accession number **ACCESSION** NM\_002754 **VERSION** NM\_002754.3 **GI**: 20986527 (LOCUS NM\_002754 1888 bp mRNA linear PRI 21-MAR-2008) with the mRNA nucleotide sequence as in sequence ID 14.

**[0111]** The translation product of the CDS 99.1196 of the MAPK13 gene is depicted in the sequence ID 15.

**[0112]** FOXO3A

**[0113]** Survival factors can suppress apoptosis in a transcription-independent manner by activating the serine/threonine kinase AKT1, which then phosphorylates and inactivates components of the apoptotic machinery, including BAD and caspase-9. Brunet et al.<sup>44</sup> demonstrated that AKT1 also regulates the activity of FKHL1 (current name FOXO3A). In the presence of survival factors, AKT1 phosphorylates FKHL1, leading to the association of FKHL1 with 14-3-3 proteins and its retention in the cytoplasm. Survival factor withdrawal leads to FKHL1 dephosphorylation, nuclear translocation, and target gene activation. Within the nucleus, FKHL1 most likely triggers apoptosis by inducing the expression of genes that are critical for cell death, such as the TNF ligand superfamily 6 (TNFSF6). Nemoto and Finkel<sup>45</sup> observed that exposure to intracellular ROS induced an increase in phosphorylated Fkhl1 and a shift from a nuclear to a cytosolic localization. They found that serum starvation, a stimulus that increases oxidative stress, resulted in lower levels of hydrogen peroxide in Shc1<sup>-/-</sup> cells or in cells expressing a ser36-to-ala (S36A) Shc1 mutant compared with wild type cells. Serum starvation also increased Fkhl1-dependent transcriptional activity,

which was further augmented in the Shc1-deficient cells. Increased ROS exposure failed to induce increased Fkhl1 phosphorylation in the mutant cells. Essers et al.<sup>46</sup> reported an evolutionarily conserved interaction of beta-catenin with FOXO transcription factors, which are regulated by insulin and oxidative stress signaling. In mammalian cells, beta-catenin binds directly to FOXO and enhances FOXO transcriptional activity. In *C. elegans*, loss of the beta-catenin BAR1 reduces the activity of the FOXO ortholog DAF16 in dauer formation and life span. Association of beta-catenin with FOXO was enhanced in cells exposed to oxidative stress. Furthermore, BAR1 was required for the oxidative stress-induced expression of the DAF16 target gene sod3 and for resistance to oxidative damage. They concluded that their results demonstrated a role for beta-catenin in regulating FOXO function that is particularly important under conditions of oxidative stress.

**[0114]** mRNA of *Homo sapiens* forkhead box O3 (FOXO3), transcript variant 1, has been deposited in the NCBI database as under the accession number **ACCESSION** NM\_001455 **VERSION** NM\_001455.3 **GI**: 146260266 with the mRNA nucleotide sequence as in sequence ID 16.

**[0115]** The translation product of the CDS 344.2365 of the human forkhead box O3 (FOXO3), transcript variant 1 is depicted in sequence ID 17.

**[0116]** Superoxide Dismutase 2

**[0117]** The superoxide dismutase 2 (SOD2; or manganese or mitochondrial superoxide dismutase) gene encodes an intramitochondrial free radical scavenging enzyme that is the first line of defense against superoxide produced as a by-product of oxidative phosphorylation. Li et al.<sup>47</sup> inactivated the Sod2 gene in transgenic mice by homologous recombination. Homozygous mutant mice died within the first 10 days of life with a dilated cardiomyopathy, accumulation of lipid in liver and skeletal muscle, and metabolic acidosis. The findings suggested SOD2 that is required for normal biologic function of tissues by maintaining the integrity of mitochondrial enzymes susceptible to direct inactivation by superoxide. Oxidative stress has been implicated in many diseases. The chief source of reactive oxygen species (ROS) within the cell is the mitochondrion. ROS have been implicated in a wide range of degenerative processes including amyotrophic lateral sclerosis, ischemic heart disease, Alzheimer disease, Parkinson disease, and aging. ROS are generated by mitochondria as the toxic by-products of oxidative phosphorylation, their energy generating pathway. Melov et al.<sup>48</sup> characterized a variety of biochemical and metabolic effects of inactivation of the mouse sod2 gen. The Sod2 mutant mice exhibited a tissue-specific inhibition of the respiratory chain enzymes NADH-dehydrogenase (complex I) and succinate dehydrogenase (complex II), inactivation of the tricarboxylic acid cycle enzyme aconitase, development of a urinary organic aciduria in conjunction with a partial defect in 3-hydroxy-3-methylglutaryl-CoA lyase, and accumulation of oxidative DNA damage. Treatment with an SOD mimetic, MnTBAP, rescued Sod2<sup>-/-</sup> mutant mice from this systemic pathology and dramatically prolonged their survival. Surviving animals developed a pronounced movement disorder progressing to total debilitation by 3 weeks of age. Neuropathologic evaluation showed a striking spongiform degeneration of the cortex and specific brainstem nuclei, associated with gliosis and intramyelinic vacuolization similar to that observed in cytotoxic edema and disorders associated with mitochondrial abnormalities such as Leigh disease and Canavan disease.

Their data suggested that because of the failure of MnTBAP to cross the blood-brain barrier progressive neuropathology is caused by excessive mitochondrial production of ROS<sup>48</sup>.

**[0118]** mRNA of *Homo sapiens* superoxide dismutase 2, mitochondrial (SOD2), nuclear gene encoding mitochondrial protein, transcript variant 1, has been deposited in the NCBI database as under the accession number ACCESSION NM\_000636 VERSION NM\_000636.2 GI:67782304 (LOCUS NM\_000636 1593 bp mRNA linear PRI 30-MAR-2008) with the mRNA nucleotide sequence as in sequence ID 18.

**[0119]** The translation product of the CDS 155.823 *Homo sapiens* superoxide dismutase 2, mitochondrial (SOD2), nuclear gene encoding mitochondrial protein, transcript variant 1 gene is depicted in sequence ID 19.

**[0120]** The biomarkers, i.e. the genes identified hereinbefore, according to the invention include substantially identical homologues and variants of the nucleic acid molecules and expression products thereof described herein, for example, a molecule that includes nucleotide sequences encoding polypeptides functionally equivalent to the biomarkers of the invention, e.g., sequences having one or more nucleotide substitutions, additions, or deletions, such as allelic variants or splice variants or species variants or molecules differing from the nucleic acid molecules and polypeptides referred to in the Tables herein due to the degeneracy of the genetic code. Species variants are nucleic acid sequences that vary from one species to another, although the resulting polypeptides generally will have significant amino acid identity and functional similarity relative to each other. A polymorphic variant (e.g., a single nucleotide polymorphism or SNP) is a variation in the nucleic acid sequence of a particular gene between individuals of a given species.

**[0121]** A "substantially identical" sequence is an amino acid or nucleotide sequence that differs from a reference sequence only by one or more conservative substitutions, as discussed herein, or by one or more non-conservative substitutions, deletions, or insertions located at positions of the sequence that do not destroy the biological function of the amino acid or nucleic acid molecule. Such a sequence can be any integer from 10% to 99%, or more generally at least 10%, 20%, 30%, 40%, 50, 55% or 60%, or at least 65%, 75%, 80%, 85%, 90%, or 95%, or as much as 96%, 97%, 98%, or 99% identical when optimally aligned at the amino acid or nucleotide level to the sequence used for comparison using, for example, the Align Program (Myers and Miller, CABIOS, 1989, 4:11-17) or FASTA. For polypeptides, the length of comparison sequences may be at least 2, 5, 10, or 15 amino acids, or at least 20, 25, or 30 amino acids. In alternate embodiments, the length of comparison sequences may be at least 35, 40, or 50 amino acids, or over 60, 80, or 100 amino acids. For nucleic acid molecules, the length of comparison sequences may be at least 5, 10, 15, 20, or nucleotides, or at least 30, 40, or 50 nucleotides. In alternate embodiments, the length of comparison sequences may be at least 60, 70, 80, or 90 nucleotides, or over 100, 200, or 500 nucleotides. Sequence identity can be readily measured using publicly available sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705, or BLAST software available from the National Library of Medicine, or as described herein). Examples of useful software include the programs Pile-up and Pretty Box. Such software matches similar

sequences by assigning degrees of homology to various substitutions, deletions, substitutions, and other modifications. Alternatively, or additionally, two nucleic acid sequences may be "substantially identical" if they hybridize under high stringency conditions. In some embodiments, high stringency conditions are, for example, conditions that allow hybridization comparable with the hybridization that occurs using a DNA probe of at least 500 nucleotides in length, in a buffer containing 0.5 M NaHPO<sub>4</sub>, pH 7.2, 7% SDS, 1 mM EDTA, and 1% BSA (fraction V), at a temperature of 65[deg.]C., or a buffer containing 48% formamide, 4.8xSSC, 0.2 M Tris-Cl, pH 7.6, 1xDenhardt's solution, 10% dextran sulfate, and 0.1% SDS, at a temperature of 42[deg.]C. (These are typical conditions for high stringency northern or Southern hybridizations.) Hybridizations may be carried out over a period of about 20 to 30 minutes, or about 2 to 6 hours, or about 10 to 15 hours, or over 24 hours or more. High stringency hybridization is also relied upon for the success of numerous techniques routinely performed by molecular biologists, such as high stringency PCR, DNA sequencing, single strand conformational polymorphism analysis, and in situ hybridization. In contrast to northern and Southern hybridizations, these techniques are usually performed with relatively short probes (e.g., usually about 16 nucleotides or longer for PCR or sequencing and about 40 nucleotides or longer for in situ hybridization). The high stringency conditions used in these techniques are well known to those skilled in the art of molecular biology, and examples of them can be found, for example, in Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y., 1998, which is hereby incorporated by reference.

**[0122]** Zinc Finger 217 (ZNF217)

**[0123]** Studies in which comparative genomic hybridization was used revealed approximately 20 regions of recurrent increased DNA sequence copy number in breast tumors<sup>49, 50</sup>. These regions are predicted to encode dominantly acting genes that may play a role in tumor progression or response to therapy. Three of these regions had been associated with established oncogenes: ERBB2 at 17q12, MYC at 8q24, and CCND1 and EMS1 at 11q13. In breast cancer, ERBB2 and CCND1/EMS1 amplification and overexpression are associated with decreased life expectancy, whereas MYC amplification has been associated with lymph node involvement, advanced stage, and an increased rate of relapse. Amplification at 20q13 occurs in a variety of tumor types and is associated with aggressive tumor behavior. Kallioniemi et al.<sup>50</sup> found increased copy number involving 20q13 in 40% of breast cancer cell lines and 18% of primary breast tumors. Other comparative genomic hybridization studies revealed copy number gains at 20q13 in greater than 25% of cancers of the ovary, colon, head and neck, brain, and pancreas. Collins et al.<sup>51</sup> reported the molecular cloning of an approximately 1-Mb region at 20q13.2 involved in recurrent amplification in breast cancer and other tumors, and the delineation of a 260-kb common region of amplification. Analysis of the 1-Mb region produced evidence of 5 genes, of which ZNF217 and NABC1 (novel amplified in breast cancer-1) emerged as strong candidate oncogenes and were characterized in detail. ZNF217 was found to be centrally located in the 260-kb common region of amplification, transcribed in multiple normal tissues, and overexpressed in all cell lines and tumors in which it was amplified and in 2 in which it was not. ZNF217 was predicted to encode alternatively spliced, Kruppel-like transcription factors of 1,062 and 1,108 amino acids, each

having a DNA-binding domain (8 C2H2 zinc fingers) and a proline-rich transcription activation domain. One of the previously known genes identified in the region of amplification was the CYP24 gene.

**[0124]** Over-expression of the zinc-finger protein 217 (ZNF217), a candidate oncogene on 20q13.2, in cultured human mammary and ovarian epithelial cells can lead to their immortalization, indicating that selection for ZNF217 expression may drive 20q13 amplification during critical early stages of cancer progression. ZNF217 can also attenuate apoptotic signals resulting from exposure to doxorubicin, suggesting that ZNF217 expression may also be involved in resistance to chemotherapy. Recent findings indicate that ZNF217 binds specific DNA sequences, recruits the co-repressor C-terminal binding protein (CtBP), and represses the transcription of a variety of genes. Inappropriate expression of ZNF217 may lead to aberrant down-regulation of genes involved in limiting the proliferation, survival, and/or invasiveness of cancer cells. Better understanding of ZNF217 and its associated pathways may provide new targets for therapeutic intervention in human cancers<sup>52</sup>.

**[0125]** Huang et al<sup>53</sup> presented evidence that ZNF217 can attenuate apoptotic signals resulting from telomere dysfunction as well as from doxorubicin-induced DNA damage and that silencing ZNF217 with siRNA restores sensitivity to doxorubicin. Moreover, elevated ZNF217 leads to increased phosphorylation of Akt, whereas inhibition of the phosphatidylinositol 3 kinase pathway and Akt phosphorylation decreases ZNF217 protein levels and increases sensitivity to doxorubicin. These results suggest that ZNF217 may promote neoplastic transformation by increasing cell survival during telomeric crisis and may promote later stages of malignancy by increasing cell survival during chemotherapy.

**[0126]** Transforming growth factor-beta (TGF-beta) is a tumor suppressor, the function of which is compromised in many types of human cancer, including breast cancer. The tumor suppressive effects of TGF-beta are caused by potent inhibition of cell proliferation due to cell cycle arrest in the G1 phase. Such antiproliferative responses are mediated by a signaling system that includes two types of cell surface receptors and intracellular signal transducers, the SMAD proteins. Different molecular mechanisms can lead to loss of antiproliferative TGF-beta responses in tumor cells, including mutations in components of the signaling system and inhibition of the SMAD signaling pathway by aberrant activities of various regulatory molecules<sup>54</sup>. Thillainadesan et al<sup>55</sup> demonstrated that stimulation of HaCaT cells with transforming growth factor beta (TGF-beta) resulted in a release of ZNF217 and a concomitant binding of SMAD2 to the proximal promoter, which preceded increases in ink4b protein expression. Furthermore, the changes in chromatin marks at the p15 (ink4b) promoter following TGF-beta stimulation were similar to those observed following ZNF217 down-regulation. Collectively, these results establish the ZNF217 complex as a novel negative regulator of the p15 (ink4b) gene and may constitute an important link between amplification of ZNF217 and the loss of TGF-beta responsiveness in breast cancer. The regulation of TGF-beta in human monocytes is also important in regulating the balance between macrophage deactivating and macrophage activating effects<sup>56</sup>.

**[0127]** mRNA and protein of *Homo sapiens* Zinc finger 217 (ZNF217), has been deposited in the NCBI database as under the accession number ACCSSION NM\_006526.2 (deposited 21-DEC-2008) and NP\_006517.1 (deposited 21-DEC-

2008) respectively with the mRNA nucleotide sequence as in sequence ID 24 (Zinc finger 217 (ZNF217) 1.5653, organism="Homo sapiens" mRNA") and the protein as in sequence ID 25 (zinc finger protein 217 organism="Homo sapiens" protein" 1.1048)

#### Methods to Determine the Gene Expression/Activity

**[0128]** Preparation of Reagents Using Biomarkers

**[0129]** The biomarkers described herein may be used to prepare oligonucleotide probes and antibodies that hybridize to or specifically bind the biomarkers mentioned herein, and homologues and variants thereof.

**[0130]** Antibodies or Other Binding Agents

**[0131]** The in vitro method of analyzing the level of IRAK3 expression or activity of expression product in a biological sample isolated from said subject, an of analyzing the level of IRAK3 expression or activity of an IRAK3 expression product in combination with the gene expression level or activity of a gene product of at least one gene selected from the group consisting of SOD2, TNFAIP6, TNFAIP3, TLR2 and IRS2 may for present invention also be carried of a binding assay. Methods, techniques and equipment for performing such binding assays will also be clear to the skilled person. For the binding assay, the binding agent may also be immobilized on a suitable support, as will be again clear to the skilled person.

**[0132]** An "antibody" includes molecules having antigen binding regions, such as whole antibodies of any isotype (IgG, IgA, IgM, IgE, etc.) and fragments thereof. Antibody fragments include Fab', Fab, F(ab')<sub>2</sub>, single domain antibodies, Fv, scFv, etc. Antibodies may be prepared using standard techniques of preparation as, for example, described in Harlow and Lane (Harlow and Lane Antibodies; A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1988), or known to those skilled in the art. For example, a coding sequence for a polypeptide biomarker of the invention may be purified to the degree necessary for immunization of rabbits. To attempt to minimize the potential problems of low affinity or specificity of antisera, two or three polypeptide constructs may be generated for each protein, and each construct may be injected into at least two rabbits. Antisera may be raised by injections in a series, preferably including at least three booster injections. Primary immunizations may be carried out with Freund's complete adjuvant and subsequent immunizations with Freund's incomplete adjuvant. Antibody titers may be monitored by Western blot and immunoprecipitation analyses using the purified protein. Immune sera may be affinity purified using CNBr-Sepharose-coupled protein. Antiserum specificity may be determined using a panel of unrelated proteins. Antibody fragments may be prepared recombinantly or by proteolytic cleavage. Peptides corresponding to relatively unique immunogenic regions of a polypeptide biomarker of the invention may be generated and coupled to keyhole limpet hemocyanin (KLH) through an introduced C-terminal lysine. Antiserum to each of these peptides may be affinity purified on peptides conjugated to BSA, and specificity tested in ELISA and Western blots using peptide conjugates and by Western blot and immunoprecipitation.

**[0133]** Monoclonal antibodies which specifically bind any one of the polypeptide biomarkers of the invention are prepared according to standard hybridoma technology (see, e.g., Kohler et al., Nature 256:495, 1975; Kohler et al., Eur. J. Immunol. 6:511, 1976; Kohler et al., Eur. J. Immunol. 6:292, 1976; Hammerling et al., In Monoclonal Antibodies and T

Cell Hybridomas, Elsevier, N.Y., 1981). Alternatively monoclonal antibodies may be prepared using the polypeptides of the invention and a phage display library (Vaughan et al., Nature Biotech 14:309-314, 1996). Once produced, monoclonal antibodies may also be tested for specific recognition by Western blot or immunoprecipitation.

**[0134]** In some embodiments, antibodies may be produced using polypeptide fragments that appear likely to be immunogenic, by criteria such as high frequency of charged residues. Antibodies can be tailored to minimize adverse host immune response by using chimerical antibodies contain an antigen binding domain from one species and the Fc portion from another species, or by using antibodies made from hybridomas of the appropriate species.

**[0135]** An antibody “specifically binds” an antigen when it recognizes and binds the antigen, for example, a biomarker as described herein, but does not substantially recognize and bind other molecules in a sample. Such an antibody has, for example, an affinity for the antigen which is at least 2, 5, 10, 100, 1000 or 10000 times greater than the affinity of the antibody for another reference molecule in a sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular biomarker. For example, a polyclonal antibody directed against a biomarker from a specific species such as rat, mouse, or human may be selected for only those polyclonal antibodies that are specifically immunoreactive with the biomarker and not with other proteins, except for polymorphic variants and alleles of the biomarker. In some embodiments, a polyclonal antibody raised to a biomarker from a specific species such as rat, mouse, or human may be selected for only those polyclonal antibodies that are specifically immunoreactive with the biomarker from that species and not with other proteins, including polymorphic variants and alleles of the biomarker.

**[0136]** In principle any suitable binding agent that can specifically binds the IRAK3 expression product and for the combined assay any suitable binding agent that can specifically binds SOD2, TNFAIP6, TNFAIP3, TLR2 and/or IRS2 can be used in the binding assay of present invention. For example, the binding agent can be a protein or polypeptide that is capable of specifically binding, such as an antibody that is capable of specifically binding; a part of fragment of an antibody, in which said part or fragment is capable of specifically binding; or a protein or polypeptide that contains and/or comprises one or more parts of fragments of an antibody, in which at least one of said parts or fragments is capable of specifically binding. In particular, said part or fragment may be a variable domain, such as a heavy chain variable domain and/or a light chain variable domain, or a ScFv comprising both a heavy chain variable domain and/or a light chain variable domain. Such antibodies and fragments, and methods for obtaining the same, will be clear to the skilled person; reference is for example made to Roitt et al., “Immunology” (6th Ed.), Mosby/Elsevier, Edinburgh (2001); and Janeway et al., “Immunobiology” (6th Ed.), Garland Science Publishing/Churchill Livingstone, N.Y. (2005).

**[0137]** According to one preferred, but non-limiting embodiment, the binding agent can be a so-called “heavy chain antibody” that is capable of specifically binding the expression products IRAK3, SOD2, TNFAIP6, TNFAIP3, TLR2 and/or IRS2; a part of fragment of a heavy chain antibody, in which said part or fragment is capable of specifically binding to these expression products; or a protein or

polypeptide that contains and/or comprises one or more parts of fragments of a heavy chain antibody, in which at least one of said parts or fragments is capable of specifically binding these expression products (for example, a multivalent protein containing two or more such fragments. Heavy chain antibodies and methods-for obtaining the same have been described in the art, see for example the following references, that are cited as general background art: WO 94/04678 (=EP 656 946), WO 96/34103 (=EP 0 822 985) and WO 97/49805 by Vrije Universiteit Brussel; WO 97/49805 by Vlaams Interuniversitair Instituut voor Biotechnologie; WO 94/25591 (=EP 0 698 097) and WO 00/43507 by Unilever N.V.; WO 01/90190 by the National Research Council of Canada; WO 03/025020 (=EP 1 433 793) by the Institute of Antibodies; WO 04/062551, WO 04/041863, WO 04/041865, WO 04/041862 by applicant; as well as for example Hamers-Casterman et al., Nature, Vol. 363, p. 446 (1993) and Riechmann and Muyldermans, Journal of Immunological Methods, 231 (1999), p. 25-38. For example, heavy chain antibodies against a desired antigen can be obtained from a species of Camelid or of Chondrostei immunized with said antigen, as described in the general prior art mentioned above. As also mentioned in the above references, naturally occurring heavy chain antibodies do not contain the light chains present in naturally occurring conventional 4-chain antibodies (that natively contain both heavy chains and light chains). Because of this, such naturally occurring heavy chain antibodies have also been referred to in the art as “single chain antibodies” (see for example WO 02/085945; and not to be confused with so-called “single chain Fv’s” or “scFv’s”, which are synthetic polypeptides comprising a  $V_H$  domain covalently linked to a  $V_L$  domain) and as “immunoglobulins devoid of light chains” (see for example EP 0 656 946 and some of the further general background art mentioned above), which terms for the purposes of the present description should be considered equivalent to the term “heavy chain antibody” as used herein. As also mentioned in these references, the heavy chains of naturally occurring heavy chain antibodies contain CH3 domains, CH2 domains and a variable domain, but—in addition to the light chains—lack the CH1 domains present in the heavy chains of naturally occurring conventional 4-chain antibodies. Herein, the variable domains from naturally occurring heavy chain antibodies will also be referred to as “ $V_{HH}$  domains”, in order to distinguish said variable domains from the variable domains from conventional 4-chain antibodies, which are commonly referred to as “ $V_H$  domains”. Generally,  $V_{HH}$  domains have a structure that retains the immunoglobulin fold of conventional  $V_H$  domains. However, compared to  $V_H$  domains,  $V_{HH}$  domains contain one or more substitutions in their amino acid sequence (and in particular in their framework regions) that make the region(s)/residues of the  $V_{HH}$  domain that in a  $V_H$  domain would form the  $V_H/V_L$  interphase more hydrophobic (see the general background art cited above). Also, as mentioned in the general background art cited above, heavy chain antibodies and  $V_{HH}$  domains have the major advantage that they are capable of binding an antigen without the presence of any light chains or light chain variable domains, respectively. This makes heavy chain antibodies and  $V_{HH}$  domains easier to obtain, to develop, to prepare (in particular) on a large scale, to use and/or to bind to a support than conventional 4-chain antibodies or light chain or heavy chain variable domains thereof. For example, the immobilization of  $V_{HH}$

domains on a solid support has been described in the International application WO 01/40310 by Hindustan Lever Limited.

**[0138]** According to a particularly preferred embodiment, the binding agent is a variable domain sequences of heavy chain antibodies, for which term reference is made to the prior art mentioned above, to US20080096223, US20070077249, US20060062784, US20050271663, WO2007063311, WO2007063308, WO2006122786 and WO2005033130. Generally, such ligands can be described as proteins that have some of the functional properties and structural features that are characteristic of naturally occurring  $V_{HH}$  domains. They may for example be a naturally occurring  $V_{HH}$  domain, a "humanized" or human  $V_{HH}$  domains or a "camelized"  $V_H$  domains, as well as a partially or fully synthetic protein, as long as the foregoing have (at least some of) the functional properties and structural features that are characteristic of naturally occurring  $V_{HH}$  domains.

**[0139]** Alternatively in the binding assays and diagnostic method of present invention aptamer sequences which specifically bind to RNA or DNA sequences or expression products of SOD2, TNFAIP6, TNFAIP3, TLR2 and/or IRS2 are used. As used herein, the terms "aptamer(s)" or "aptamer sequence(s)" are meant to refer to single stranded nucleic acids (RNA or DNA) whose distinct nucleotide sequence determines the folding of the molecule into a unique three dimensional structure. Aptamers comprising 15 to 120 nucleotides can be selected in vitro from a randomized pool of oligonucleotides (1014-1015 molecules). Any aptamers of the invention as described herein further contemplates the use of both native and modified DNA and RNA bases, such as beta-D-Glucosyl-Hydroxymethyluracil. Aptamer sequences can be isolated through methods such as those disclosed in co-pending U.S. patent application Ser. No. 10/934,856, entitled, "Aptamers and Methods for their In vitro Selection and Uses Thereof," which is hereby incorporated by reference. It is contemplated that the sequences described herein may be varied to result in substantially homologous sequences which retain the same function as the original. As used herein, a polynucleotide or fragment thereof is "substantially homologous" (or "substantially similar") to another if, when optimally aligned (with appropriate nucleotide insertions or deletions) with the other polynucleotide (or its complementary strand), using an alignment program such as BLASTN (Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) "Basic local alignment search tool." *J. Mol. Biol.* 215:403-410), and there is nucleotide sequence identity in at least about 80%, preferably at least about 90%, and more preferably at least about 95-98% of the nucleotide bases. Nucleic acids encoding sequences of SOD2, TNFAIP6, TNFAIP3, TLR2 and/or IRS2 or their expression products can also be isolated from expression libraries using antibodies as probes. Such polyclonal or monoclonal antibodies can be raised using methods known in the art (see, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual* (1988)).

**[0140]** Probes and Primers

**[0141]** A "probe" or "primer" is a single-stranded DNA or RNA molecule of defined sequence that can base pair to a second DNA or RNA molecule that contains a complementary sequence (the target). The stability of the resulting hybrid molecule depends upon the extent of the base pairing that occurs, and is affected by parameters such as the degree of complementarity between the probe and target molecule, and the degree of stringency of the hybridization conditions. The

degree of hybridization stringency is affected by parameters such as the temperature, salt concentration, and concentration of organic molecules, such as formamide, and is determined by methods that are known to those skilled in the art. Probes or primers specific for the nucleic acid biomarkers described herein, or portions thereof, may vary in length by any integer from at least 8 nucleotides to over 500 nucleotides, including any value in between, depending on the purpose for which, and conditions under which, the probe or primer is used. For example, a probe or primer may be 8, 10, 15, 20, or 25 nucleotides in length, or may be at least 30, 40, 50, or 60 nucleotides in length, or may be over 100, 200, 500, or 1000 nucleotides in length. Probes or primers specific for the nucleic acid biomarkers described herein may have greater than 20-30% sequence identity, or at least 55-75% sequence identity, or at least 75-85% sequence identity, or at least 85-99% sequence identity, or 100% sequence identity to the nucleic acid biomarkers described herein. Probes or primers may be derived from genomic DNA or cDNA, for example, by amplification, or from cloned DNA segments, and may contain either genomic DNA or cDNA sequences representing all or a portion of a single gene from a single individual. A probe may have a unique sequence (e.g., 100% identity to a nucleic acid biomarker) and/or have a known sequence. Probes or primers may be chemically synthesized. A probe or primer may hybridize to a nucleic acid biomarker under high stringency conditions as described herein.

**[0142]** Diagnostic Use

**[0143]** In a preferred embodiment, the invention involves methods to assess quantitative and qualitative aspects of the biomarker gene expression(s). In one example the decreased expression of an IRAK3 gene or gene product as provided by the present invention is indicative for the progression of a metabolic syndrome disorder in a subject or the increased risk to develop related cardiovascular diseases in said subject. Techniques well known in the art, e.g., quantitative or semi-quantitative RT PCR for instance real time RT PCR, for instance mRNA analysis by the fluorescence-based real-time reverse transcription polymerase chain reaction (qRT-PCR or RT-qPCR) or reverse transcription loop-mediated amplification (RT-LAMP), for instance one-step RT-LAMP, or real-time NASBA for detection, quantification and differentiation of the RNA and DNA targets (LOENS K. et al. *Journal of microbiological methods* 2006, vol. 67, no3, pp. 408-415, or Northern blot, can be used to measure expression levels of IRAK3.

**[0144]** The measurement of IRAK3 gene expression levels may include measuring naturally occurring IRAK3 transcripts and variants thereof as well as non-naturally occurring variants thereof. The diagnosis and/or prognosis of metabolic syndrome disorder in a subject, however is preferably directed to detecting a naturally occurring IRAK3 gene product or variant thereof. Thus, the invention relates to methods of diagnosing and/or predicting metabolic syndrome disorder in a subject by measuring the expression of an IRAK3 gene in a subject. For example a decreased level of mRNA encoded by an IRAK3 nucleic acid sequence (e.g., SEQ ID NO: 1).

**[0145]** Diagnostic methods for the detection of IRAK3 nucleic acid molecules, in patient samples or other appropriate cell sources, may involve the pre-amplification of specific gene sequences, e.g., by PCR (See Mullis, K. B., 1987, U.S. Pat. No. 4,683,202), followed by the analysis of the amplified molecules using techniques well known to those of skill in the art, such as, for example, those listed above. Utilizing analy-

sis techniques such as these, amplified sequences can be compared to the levels in control samples.

**[0146]** In a particular embodiment, the analyzing techniques include the application of detectably-labeled probes or primers. The probes or primers can be detectably-labeled, either radioactively or non-radioactively, by methods that are known to those skilled in the art, and their use in the methods according to the invention, involves nucleic acid hybridization, such as nucleic acid sequencing, nucleic acid amplification by the polymerase chain reaction (e.g., RT-PCR), single stranded conformational polymorphism (SSCP) analysis, restriction fragment polymorphism (RFLP), analysis, Southern hybridization, northern hybridization, in situ hybridization, electrophoretic mobility shift assay (EMSA), fluorescent in situ hybridization (FISH), and other methods that are known to those skilled in the art.

**[0147]** By "detectably labeled" is meant any means for marking and identifying the presence of a molecule, e.g., an oligonucleotide probe or primer, a gene or fragment thereof, or a cDNA molecule. Methods for detectably-labeling a molecule are well known in the art and include, without limitation, radioactive labeling (e.g., with an isotope such as <sup>32</sup>P or <sup>35</sup>S) and nonradioactive labeling such as, enzymatic labeling (for example, using horseradish peroxidase or alkaline phosphatase), chemiluminescent labeling, fluorescent labeling (for example, using fluorescein), bioluminescent labeling, or antibody detection of a ligand attached to the probe. Also included in this definition is a molecule that is detectably labeled by an indirect means, for example, a molecule that is bound with a first moiety (such as biotin) that is, in turn, bound to a second moiety that may be observed or assayed (such as fluorescein-labeled streptavidin). Labels also include digoxigenin, luciferases, and aequorin.

**[0148]** Immunoassays

**[0149]** Antibodies that specifically bind any of the biomarkers described herein may be employed in an immunoassay by contacting a sample with the antibody and detecting the presence of a complex of the antibody bound to the biomarker in the sample. The antibodies used in an immunoassay may be produced as described herein or known in the art, or may be commercially available from suppliers, such as Dako Canada, Inc., Mississauga, ON. The antibody may be fixed to a solid substrate (e.g., nylon, glass, ceramic, plastic, etc.) before being contacted with the sample, to facilitate subsequent assay procedures. The antibody-biomarker complex may be visualized or detected using a variety of standard procedures, such as detection of radioactivity, fluorescence, luminescence, chemiluminescence, absorbance, or by microscopy, imaging, etc. Immunoassays include immunohistochemistry, enzyme-linked immunosorbent assay (ELISA), western blotting, and other methods known to those of skill in the art. Immunoassays can be used to determine presence or absence of a biomarker in a sample as well as the amount of a biomarker in a sample. The amount of an antibody-biomarker complex can be determined by comparison to a reference or standard, such as a polypeptide known to be present in the sample. The amount of an antibody-biomarker complex can also be determined by comparison to a reference or standard, such as the amount of the biomarker in a reference or control sample. Accordingly, the amount of a biomarker in a sample need not be quantified in absolute terms, but may be measured in relative terms with respect to a reference or control.

Methods of Treatment

**[0150]** Detection of the biomarkers described herein may enable a medical practitioner to determine the appropriate

course of action for a subject (e.g., further testing, drug or dietary therapy, surgery, no action, etc.) based on the diagnosis. Detection of the biomarkers described herein may also help determine the presence or absence of a metabolic syndrome disorder, early diagnosis of a metabolic syndrome disorder, prognosis of a metabolic syndrome disorder, efficacy of a therapy for a metabolic syndrome disorder, or monitoring a metabolic syndrome disorder therapy in a subject. In alternative aspects, the biomarkers and reagents prepared using the biomarkers may be used to identify metabolic syndrome disorder therapeutics. The methods according to the invention allow a medical practitioner to monitor a metabolic syndrome disorder therapy in a subject, enabling the medical practitioner to modify the treatment based upon the results of the test. The methods can also be used to identify and validate breast cancer therapeutics, such as small molecules; peptides, etc.

**[0151]** In said aspect of the present invention, it has for example been found that a metabolic syndrome disorder can be treated by administering to subject in need thereof an effective amount of a therapeutic that increases the expression of IRAK3 in the monocytes or macrophages or any white blood cell.

**[0152]** A particular aspect of present invention is a method of treating a disease, condition or disorder selected from the group consisting of (1) non-insulin dependent Type 2 diabetes mellitus (NIDDM), (2) hyperglycemia, (3) low glucose tolerance, (4) insulin resistance, (6) a lipid disorder, (7) dyslipidemia, (8) hyperlipidemia, (9) hypertriglyceridemia, (10) hypercholesterolemia, (11) low HDL levels, (12) high LDL levels, or (13) atherosclerosis, comprising administering to subject in need thereof an effective amount of a compound which increases the expression of IRAK3 in the monocytes or macrophages.

**[0153]** Increased expression of IRAK3 or induction of elevation of both IRAK3 mRNA and IRAK3 protein in blood monocytes can for instance been achieved by teichoic acid, hyaluronan, lipoarabinomannan lipopolysaccharide, mannan, lipoarabinomannan, highly mannosylated polymers, peptidoglycan or fragments thereof.

**[0154]** Present invention involves a compound of the group consisting of teichoic acid, hyaluronan, lipoarabinomannan lipopolysaccharide, mannan, lipoarabinomannan, highly mannosylated polymers, peptidoglycan or fragments thereof for use in a treatment of metabolic syndrome.

**[0155]** Present invention involves a compound of the group consisting of teichoic acid, hyaluronan, lipoarabinomannan lipopolysaccharide, mannan, lipoarabinomannan, highly mannosylated polymers, peptidoglycan or fragments thereof for use in treating a disease, condition or disorder selected from the group consisting of (1) non-insulin dependent Type 2 diabetes mellitus (NIDDM), (2) hyperglycemia, (3) low glucose tolerance, (4) insulin resistance, (6) a lipid disorder, (7) dyslipidemia, (8) hyperlipidemia, (9) hypertriglyceridemia, (10) hypercholesterolemia, (11) low HDL levels, (12) high LDL levels, or (13) atherosclerosis is an embodiment of present invention.

**[0156]** The effective amount of a compound, which is required to achieve a therapeutic effect will be, of course, vary with the type of therapeutic component, such as small molecules, peptides, etc; the route of administration; the age and condition of the recipient; and the particular disorder or disease being treated. In all aspects of the invention, the daily maintenance dose can be given for a period clinically desir-

able in the patient, for example from 1 day up to several years (e.g. for the mammal's entire remaining life); for example from about (2 or 3 or 5 days, 1 or 2 weeks, or 1 month) upwards and/or for example up to about (5 years, 1 year, 6 months, 1 month, 1 week, or 3 or 5 days). Administration of the daily maintenance dose for about 3 to about 5 days or for about 1 week to about 1 year is typical. Nevertheless, unit doses should preferably be administered from twice daily to once every two weeks until a therapeutic effect is observed.

**[0157]** The saccharide mannan, polysaccharide or oligosaccharide, is a chain of mannose molecules polymer of the sugar mannose of plants that can be derived from the cell wall of yeast for instance from bakers' yeast (*Saccharomyces cerevisiae*). The mannan polysaccharide or mannan oligosaccharide may be derived from fungus or plants. Or it can be the phosphorylated glucomannan polysaccharide which is obtainable from the cell wall of *Candida utilis* or other *Candida* species such as *Candida albicans* for instance according to the methods described in patents P9900408 (Spain) and PCT/ES99/00338. A common source of mannan oligosaccharides are cell wall fragments obtained from *Saccharomyces cerevisiae*. Such bulk materials are obtainable by lysing yeast cells and centrifuging the resulting culture to isolate the cell wall components, which are subsequently washed and spray dried (Spring et al., 2000 *Poult. Sci.* 79:205-211). It is for instance also obtainable by citrate buffer extraction (Peat, S., W. J. Whelan, and T. E. Edwards. 1961. Polysaccharides of bakers' yeast. IV. Mannan. *J. Chem. Soc.* p. 129-34; Peat, S., Whelan, W. J., AND Edwards, T. E., *J. Chem. Soc.*, 3862 (1968)) or the methods of Yasuhito Okubo et al., *Journal of Bacteriology*, October 1978, p. 63-68 or Okubo Y., Suzuki S., *Carbohydr. Res.* 62, 135-141 (1978). Two fractions of neutral and acid mannan may be obtainable by firstly applying one of the previous methods (Peat, S., W. J. Whelan, and T. E. Edwards. 1961. Polysaccharides of bakers' yeast. I V. Mannan. *J. Chem. Soc.* p. 129-34; Peat, S., Whelan, W. J., AND Edwards, T. E., *J. Chem. Soc.*, 3862 (1968), Yasuhito Okubo et al., *Journal of Bacteriology*, October 1978, p. 63-68 or Okubo Y., Suzuki S., *Carbohydr. Res.*, 62, 135-141 (1978) to obtain bulk mannan fraction for instance extracted from bakers' yeast and consequently was further fractionation by DEAE-Sephadex A-50 column chromatography whereby the neutral and the strongly acidic fractions can be eluted with water and 0.25 M NaCl respectively. Both the neutral mannan and the acidic mannan fractions from for instance baker's yeast (*Saccharomyces cerevisiae*) can be used for the metabolic syndrome treatment or the treatment of disorder selected from the group consisting of (1) non-insulin dependent Type 2 diabetes mellitus (NIDDM), (2) hyperglycemia, (3) low glucose tolerance, (4) insulin resistance, (6) a lipid disorder, (7) dyslipidemia, (8) hyperlipidemia, (9) hypertriglyceridemia, (10) hypercholesterolemia, (11) low HDL levels, (12) high LDL levels or (13) atherosclerosis of present invention. The acidic fractions are the preferred fractions for such treatment. Mannan oligosaccharide preparations are also available in the art such as the concentrated MOS-500 Extract M.O.S.500 is which is a naturally derived extract from the cell wall of *Saccharomyces cerevisiae* (mannan oligosaccharide content is approximately 50% of the carbohydrate fraction) and a food grade ingredient and fermentation additive (Ultra Bio-Logics Inc.), the Red Star Company, Oakland, Calif. or the mannan oligosaccharides (MOS), (BioMos®) from Alltech Inc., Nicholasville, Ky.

**[0158]** Peptidoglycan, also known as murein, is a polymer consisting of sugars and amino acids that forms a mesh-like layer outside the plasma membrane of eubacteria. The sugar component consists of alternating residues of  $\beta$ -(1,4) linked N-acetylglucosamine and N-acetylmuramic acid residues. Peptidoglycans are for instance also obtainable from whale cartilage according to a method described by Kazuyuki Sugahara et al *Eur. J. Biochem.* 202, 805-811 (1991).

**[0159]** Examples of hyaluronic acid, recently renamed hyaluronan are the naturally occurring hyaluronan (Hyalgan), the synthetic hylan G-F 20 (Synvisc) and HMW-HA (HEALON™) for clinical application with an endotoxin content <0.1 ng/mg of Amersham Pharmacia Biotech. Hylans are cross-linked hyaluronic acids, which gives them a higher molecular weight and increased elastoviscous properties. High molecular weight hyaluronans (for instance with a MW between 600 to 1500 kDa) for the treatment of present invention are preferred above low molecular weight LMW-HA) hyaluronans (for instance <600 kDa). LMWA can be prepared by enzymatic digestion as described previously (Termeer, C., et al. 2000. *J. Immunol.* 165:1863-1870). These compounds can be used to increase IRAK3 expression in myeloid cells such as monocytes in order to treat metabolic syndrome or a disease, condition or disorder selected from the group consisting of (1) non-insulin dependent Type 2 diabetes mellitus (NIDDM), (2) hyperglycemia, (3) low glucose tolerance, (4) insulin resistance, (6) a lipid disorder, (7) dyslipidemia, (8) hyperlipidemia, (9) hypertriglyceridemia, (10) hypercholesterolemia, (11) low HDL levels, (12) high LDL levels, or (13) atherosclerosis

**[0160]** Peptidoglycan, also known as murein, is a polymer consisting of sugars and amino acids that forms a mesh-like layer outside the plasma membrane of eubacteria. The sugar component consists of alternating residues of  $\beta$ -(1,4) linked N-acetylglucosamine and N-acetylmuramic acid residues. Attached to the N-acetylmuramic acid is a peptide chain of three to five amino acids. The peptide chain can be cross-linked to the peptide chain of another strand forming the 3D mesh-like layer. Some Archaea have a similar layer of pseudopeptidoglycan. These compounds can be used to increase IRAK3 expression in myeloid cells such as monocytes in order to treat metabolic syndrome or a disease, condition or disorder selected from the group consisting of (1) non-insulin dependent Type 2 diabetes mellitus (NIDDM), (2) hyperglycemia, (3) low glucose tolerance, (4) insulin resistance, (6) a lipid disorder, (7) dyslipidemia, (8) hyperlipidemia, (9) hypertriglyceridemia, (10) hypercholesterolemia, (11) low HDL levels, (12) high LDL levels, or (13) atherosclerosis.

**[0161]** A particular aspect of present invention is a method of treating a metabolic syndrome or a disease, condition or disorder selected from the group consisting of (1) non-insulin dependent Type 2 diabetes mellitus (NIDDM), (2) hyperglycemia, (3) low glucose tolerance, (4) insulin resistance, (6) a lipid disorder, (7) dyslipidemia, (8) hyperlipidemia, (9) hypertriglyceridemia, (10) hypercholesterolemia, (11) low HDL levels, (12) high LDL levels, (13) atherosclerosis, comprising administering to subject in need thereof an effective amount of at least one bacterial or fungal microorganism of which the cell wall is broken for releasing components for instance by the cell wall is broken by lytic enzymes such as glucosaminidases, amidases, and endopeptidases for instance lysozyme. The microorganisms suspended in a buffer containing lytic enzyme may be incubated at between about 20

Celsius to about 50 Celsius until at least the probiotic is digested. And separating the suspension by centrifugation and consequently inactivating the lysozyme by boiling. Eventually the supernatant can be processed to remove endotoxins for instance by chromatography with Detoxi-Gel (Pierce, Cat#20344) to remove any endotoxin. Such suitable microorganisms are of the Bifidobacteria (such as *Bifidobacterium bifidum*, *Bifidobacterium breve*, *Bifidobacterium infantis*, and *Bifidobacterium longum*), Lactobacilli (such as *Lactobacillus acidophilus*, *Lactobacillus bulgaricus*, *Lactobacillus casei*, *Lactobacillus plantarum*, *Lactobacillus rhamnosus*, *Lactobacillus GG*, and *Lactobacillus reuteri*), Streptococci (such as *Streptococcus thermophilus*); or yeast (such as *Saccharomyces boulardii*).

**[0162]** A particular embodiment of present invention is muramyl dipeptide or Acetylmuramyl-Alanyl-Isoglutamine (MDP) inducer of expression of IRAK3 for use in a treatment of curing or preventing metabolic syndrome disorder. Muramyl dipeptide is a peptidoglycan constituent of both Gram positive and Gram negative bacteria that is composed of N-acetylmuramic acid linked by its lactic acid moiety to the N-terminus of an L-alanine D-isoglutamine dipeptide and that is particularly known to induce expression of IRAK3 (PNAS Dec. 4, 2007 vol. 104 no. 49 19440-19445).

**[0163]** Furthermore the present invention involves muramyl dipeptide or Acetylmuramyl-Alanyl-Isoglutamine (MDP) (inducer of expression of IRAK3) for use in treating a disease, condition or disorder selected from the group consisting of (1) non-insulin dependent Type 2 diabetes mellitus (NIDDM), (2) hyperglycemia, (3) low glucose tolerance, (4) insulin resistance, (6) a lipid disorder, (7) dyslipidemia, (8) hyperlipidemia, (9) hypertriglyceridemia, (10) hypercholesterolemia, (11) low HDL levels, (12) high LDL levels, or (13) atherosclerosis is an embodiment of present invention.

#### Compositions

**[0164]** It is also an object of the present invention to provide a composition comprising the above mentioned components. In particular, suitable for use in treating and/or preventing a metabolic syndrome or a disease, condition or disorder selected from the group consisting of (1) non-insulin dependent Type 2 diabetes mellitus (NIDDM), (2) hyperglycemia, (3) low glucose tolerance, (4) insulin resistance, (6) a lipid disorder, (7) dyslipidemia, (8) hyperlipidemia, (9) hypertriglyceridemia, (10) hypercholesterolemia, (11) low HDL levels, (12) high LDL levels, (13) atherosclerosis, in a subject in need thereof.

**[0165]** The compositions of the present invention, for use in the methods of the present invention, can be prepared in any known or otherwise effective dosage or product form suitable for use in providing topical or systemic delivery of the therapeutic compounds, which would include both pharmaceutical dosage forms as well as nutritional product forms suitable for use in the methods described herein.

**[0166]** The above mentioned components may be administered to induce an increase in IRAK3 mRNA and IRAK3 protein in myeloid cells in particular in blood monocytes. Such administration can be in any form by any effective route, including, for example, oral, parenteral, enteral, intraperitoneal, topical, transdermal (e.g., using any standard patch), ophthalmic, nasally, local, non-oral, such as aerosol, spray, inhalation, subcutaneous, intravenous, intramuscular, buccal, sublingual, rectal, vaginal, intra-arterial, and intrathecal, etc. Oral administration is preferred. Such dosage forms can be

prepared by conventional methods well known in the art, and would include both pharmaceutical dosage forms as well as nutritional products.

#### Pharmaceutical Compositions

**[0167]** The pharmaceutical compositions of the present invention can be prepared by any known or otherwise effective method for formulating or manufacturing the selected product form. For example, the above mentioned components can be formulated along with common excipients, diluents, or carriers, and formed into oral tablets, capsules, sprays, mouth washes, lozenges, treated substrates (e.g. oral or topical swabs, pads, or disposable, non-digestible substrate treated with the compositions of the present invention); oral liquids (e.g. suspensions, solutions, emulsions), powders, or any other suitable dosage form.

**[0168]** Non-limiting examples of suitable excipients, diluents, and carriers include: fillers and extenders such as starch, sugars, mannitol, and silicic derivatives; binding agents such as carboxymethyl cellulose and other cellulose derivatives, alginates, gelatin, and polyvinyl pyrrolidone; moisturizing agents such as glycerol; disintegrating agents such as calcium carbonate and sodium bicarbonate; agents for retarding dissolution such as paraffin; resorption accelerators such as quaternary ammonium compounds; surface active agents such as acetyl alcohol, glycerol monostearate; adsorptive carriers such as kaolin and bentonite; carriers such as propylene glycol and ethyl alcohol, and lubricants such as talc, calcium and magnesium stearate, and solid polyethyl glycols.

**[0169]** A particular embodiment of present invention is also an in vitro method of diagnosing a metabolic syndrome disorder phenotype in a subject, said method comprising: (a) analyzing the level of ZNF217 expression or activity of expression product in a biological sample isolated from said subject, and (b) compare said level of expression or activity with the ZNF217 expression or activity in a control sample; whereby a increased level of ZNF217 expression or activity relative to a control sample is an indication of such metabolic syndrome disorder phenotype or a propensity thereto.

#### Measuring ZNF217 Expression

**[0170]** Another particular embodiment of present invention is an in vitro method of diagnosing a metabolic syndrome disorder, or a propensity thereto in a subject, said method (a) obtaining an expression profile in a biological sample isolated from said subject, wherein said expression profile consists of the analysis of the level of ZNF217 expression or activity of an ZNF217 expression product in combination with the gene expression level or activity of a gene product of at least one gene selected from the group consisting of SOD2, TNFAIP6, TNFAIP3, TLR2, IRAK3 and IRS2; and (b) comparing said obtained expression profile to a reference expression profile to determine whether said sample is from subject having a metabolic syndrome disorder phenotype or a propensity thereto.

**[0171]** The present invention can comprise the above described in vitro methods wherein the metabolic syndrome disorder is an impaired adipose tissue accumulation or adipocyte function, an impaired glucose tolerance condition, an insulin resistance or type II diabetes disorder, a lipid homeostasis disorder or a cardiovascular disease related thereto. Such cardiovascular disease can hereby be of the group consisting of hypertension, coronary heart disease, heart failure,

congestive heart failure, atherosclerosis, arteriosclerosis, stroke, cerebrovascular disease, myocardial infarction and peripheral vascular disease.

[0172] Such in vitro methods as described hereabove can comprise analyzing the level of ZNF217 expression or activity of expression product in a sample isolated from said subject, whereby a decreased level of ZNF217 expression or activity relative to a control sample is an indication of such adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto. The expression product can be a nucleic acid molecule selected from the group consisting of mRNA and cDNA mRNA or derived polypeptides. Moreover the in vitro method can be carried out on a sample isolated from said subject is selected from a group consisting of (a) a liquid containing cells; (b) a tissue-sample; (c) a cell-sample and (a) a cell biopsy. Such sample can comprise haematopoietic cells or blood cells. For instance the sample can comprise at least one myeloid cell or debris thereof or the sample can comprise at least one of monocytes or peripheral blood mononuclear cells or debris thereof. Moreover the in-vitro method as described above can be by carrying out the detection of the level of the nucleic acids or polypeptides is carried using at least one binding agent specifically binding to the nucleic acids or polypeptides to be detected. The binding agent can be detectably labelled and the label can be selected from the group consisting of a radioisotope, a bioluminescent compound, a chemiluminescent compound, a fluorescent compound, a metal chelate, biotin, digoxigenin and an enzyme. The in vitro methods described above can be wherein at least one binding agent is an aptamer or an antibody selected from a group comprising

- [0173] (a) a monoclonal antibody;
- [0174] (b) a polyclonal antibody;
- [0175] (c) a Fab-Fragment; and
- [0176] (d) a single chain antibody.
- [0177] (e) an antibody variable domain sequence.

[0178] The detection can comprise an immuno-cytochemical detection procedure or a procedure wherein at least one binding agent being a nucleic acid hybridising to a nucleic acid for the detection of the marker molecules, in particular for the detection of ZNF217 alone or in combination with a gene of the group consisting of IRAK3, SOD2, TNFAIP6, TNFAIP3, TLR2 and IRS2. Furthermore the detection reaction can comprise a nucleic acid amplification reaction. Finally the above described in vitro method can be for in-situ detection.

[0179] The present invention also comprises an in vitro method for identifying or monitoring a metabolic syndrome disorder therapy in a subject said method comprising analysing the level of ZNF217 expression or activity of expression product in a sample isolated from said subject before and after treatment with said therapy, whereby an increased level of ZNF217 compared to a sample of said subject before the therapy is indicative for the efficacy of said therapy.

[0180] An in vitro method for identifying or monitoring a metabolic syndrome disorder therapy in a subject said method comprising analyzing the expression profile in a biological sample isolated from said subject before and after treatment with said therapy, wherein said expression profile consists of the analysis of the level of ZNF217 expression or activity of an ZNF217 expression product in combination with the gene expression level or activity of a gene product of

at least one gene selected from the group consisting of SOD2, TNFAIP6, TNFAIP3, TLR2, IRAK3 and IRS2; and whereby an increased level of ZNF217, a decreased level of SOD2, TNFAIP3 and/or TNFAIP3 and an increased level of IRAK3 compared to a sample of said subject before the therapy is indicative for the efficacy of said therapy. Such method can be used in identifying or monitoring a metabolic syndrome disorder therapy. Such therapy can be a treatment by a medication or a nutraceutical or it can concern the administration to said subject of a medicament of the group consisting of pogli-tazone, rosiglitazone, netoglitazone, rivoglitazone (CS-011), FK-614, tesaglitazar (AZ-242), ragaglitazar (N,N-622), muraglitazar (BMS-298585), edaglitazone (BM-13-1258), metaglidasen (MBX-102), naveglitazar (LY-519818), MX-6054, LY-510929, AMG-131(T-131), THR-0921), voglibose, acarbose, miglitol, emiglitate, phenformin, metformin, buformin, Vidagliptin (LAF237), P32/98, Sitagliptin (MK-431), P93/01, PT-100, saxagliptin (BMS-477118), T-6666, TS-021), AJ-9677, GLP-1, GLP-1MR agent, N,N-2211, AC-2993 (exendin-4), BIM-51077, Aib (8,35)hGLP-1 (7,37)NH<sub>2</sub>, CJC-[131], pramlintide, sodium vanadate, BVT-3498, AS-2868 Ro-28-1675, and GIP (Glucose-dependent insulinotropic peptide). Moreover the treatment can concern administration to said subject of a medicament of the group consisting of pravastatin, simvastatin, lovastatin, atorvastatin, fluvastatin, rosuvastatin, pitavastatin N—[[[(3R,5S)-1-(3-acetoxy-2,2-dimethylpropyl)-7-chloro-5-(2,3-dimethoxyphenyl)-2-oxo-1,2,3,5-tetrahydro-4,1-benzoxazepin-3-yl] acetyl] piperidine-4-acetic acid, ciprofibrate, fenofibrate, bezafibrate, clofibrate, simfibrate, clinofibrate, Avasimibe, Eflucimibe, colestyramine, probucol, nicomol, niceritrol, ethyl icosapentate, soysterol and  $\gamma$ -oryzanol. The method can also concern administration to said subject of a medicament of the group consisting of dexfenfluramine, fenfluramine, phentermine, sibutramine, amfepramone, dexamphetamine, mazindol, phenylpropanolamine, clobenzorex, SB-568849; SNAP-7941, CP-422935, SR-141716, SR-147778, BVT-3498, orlistat, cetilistat (ATL-962)), A3-9677, leptin, CNTF (Ciliary Neurotropic Factor), linitript, FPL-15849.

[0181] A particular aspect of present invention is a method of treating a disease, condition or disorder selected from the group consisting of (1) non-insulin dependent Type 2 diabetes mellitus (NIDDM), (2) hyperglycemia, (3) low glucose tolerance, (4) insulin resistance, (6) a lipid disorder, (7) dyslipidemia, (8) hyperlipidemia, (9) hypertriglyceridemia, (10) hypercholesterolemia, (11) low HDL levels, (12) high LDL levels, or (13) atherosclerosis, comprising administering to subject in need thereof an effective amount of a compound which increases the expression of ZNF217 in the monocytes or macrophages.

[0182] Increased expression of ZNF217 or induction of elevation of both ZNF217 mRNA and ZNF217 protein in blood monocytes can for instance been achieved by an organic fluid extract, for instance an alcohol extract, for instance a C<sub>1</sub>-C<sub>4</sub> alcohol extract or an primary, secondary and tertiary alcohol, or for instance an ethanol extract of *Fagopyrum dibotrys* or by procyanidin B-2 (US20060165825).

[0183] An organic fluid extract or a CO<sub>2</sub> supercritical fluid extract of the rhizomes of *Fagopyrum dibotrys* for use in a treatment of curing or preventing metabolic syndrome is an embodiment of present invention.

[0184] An organic fluid extract or a CO<sub>2</sub> supercritical fluid extract of the rhizomes of *Fagopyrum dibotrys* for use in treating a disease, condition or disorder selected from the

group consisting of (1) non-insulin dependent Type 2 diabetes mellitus (NIDDM), (2) hyperglycemia, (3) low glucose tolerance, (4) insulin resistance, (6) a lipid disorder, (7) dyslipidemia, (8) hyperlipidemia, (9) hypertriglyceridemia, (10) hypercholesterolemia, (11) low HDL levels, (12) high LDL levels, or (13) atherosclerosis is an embodiment of present invention.

**[0185]** Other embodiments of the invention will be apparent to those skilled in the art from consideration of the specification and practice of the invention disclosed herein. It is intended that the specification and examples be considered as exemplary only, with a true scope and spirit of the invention being indicated by the following claims.

#### DRAWING DESCRIPTION

##### Brief Description of the Drawings

**[0186]** The present invention will become more fully understood from the detailed description given herein below and the accompanying drawings which are given by way of illustration only, and thus are not limitative of the present invention, and wherein:

**[0187]** FIG. 1 demonstrates the predicted cardiovascular risk (by Framingham risk scoring or FRS) and metabolic syndrome-related characteristics of controls (n=25), and obese women (n=95). \*\*\*P<0.001 compared to controls.

**[0188]** FIG. 2: displays the predicted cardiovascular risk (by Framingham risk scoring or FRS) and metabolic syndrome-related characteristics of obese women without (n=35) and with (n=60) diabetes. \*, \*\*, \*\*\*P<0.05, P<0.01, and P<0.001 compared to obese women without diabetes.

**[0189]** FIG. 3: demonstrates the predicted cardiovascular risk (by Framingham risk scoring or FRS) and metabolic syndrome-related characteristics of obese women (N=14) without diabetes before and 3 months after bariatric surgery. \*, \*\*, \*\*\*P<0.05, P<0.01, and P<0.001 compared to values before weight loss.

**[0190]** FIG. 4: provides the top ten canonical signaling pathways which were identified by means of Ingenuity Pathways Analysis (Ingenuity Systems® Ingenuity IPA 5.5-802, www.ingenuity.com). The significance of the association between the dataset and the canonical pathway was measured in 2 ways. First, a ratio (yellow squares) of the number of genes which were differentially expressed in monocytes of obese women and were assigned to a particular signaling pathway to the total number of genes that belong to this signaling canonical pathway and were present on the microchip. Second, the Fischer's exact test was used to calculate a p-value determining the probability that the association between the differentially expressed genes in the dataset and the assigned genes in the canonical pathway is not explained by chance alone. The threshold was P=0.01.

**[0191]** FIG. 5: displays the proposed structural model that predicts the interactions between the TLR2 and the PI3K/AKT pathway. Flow of the pathways at the protein interaction level is indicated by black arrows. Blunted arrows indicate inhibition. Phosphorylation is indicated by a circled P. Note that both Foxo3a and NFκB are regulated at the level of nuclear transport. NFκB is translocated to the nucleus after dissociation from the phosphorylated IκB. Foxo3a is a direct target of AKT/PKB and the phosphorylated transcription is unable to enter the nucleus and become active. Red arrows indicate downstream transcriptional events. Dashed arrows indicate the hypothetical involvement of Foxo3a in the tran-

scription of TRAF6 and IRAK3. Dark grey: unregulated genes; light grey; down regulated genes; white: genes of which the expression is not different in monocytes of lean controls and obese women.

**[0192]** FIG. 6 is a schematic view of the conserved NFκB-binding sites in the TLR2 promoter.

**[0193]** FIG. 7: displays common transcription factor binding sites in the promoters of TNFAIP3, MAPK13, and SOD2.

**[0194]** FIG. 8: displays the results of RT-PCR analysis of indicated genes in monocytes from control and obese women; obese women without and with diabetes are combined. \*\*, \*\*\*P<0.01, and P<0.001 compared to controls

**[0195]** FIG. 9: provides a comparison of RNA expression before and after weight loss. Individual data before weight loss were expressed as a ratio compared to a pool of 29 samples of non-obese women. Individual data after weight loss were expressed as a ratio compared to values before weight loss. \*\*, \*\*\*P<0.01, and P<0.001 compared to before.

**[0196]** FIG. 10: An example of hierarchical clustering analysis based on Euclidean distance after normalization. Group attribution: lean controls (green), obese persons without cardiovascular risk equivalents (yellow), and obese persons with cardiovascular risk equivalents (red). Cardiovascular risk equivalents are: an at least 10% risk of developing a cardiovascular disease within the next 10 years based on Framingham scoring and/or diabetes and or metabolic syndrome (at least 3 metabolic syndrome or MetSyn components as defined above).

**[0197]** FIG. 11: An example of Principal component analysis of 6-10 variables (indicated on FIG. 10) for 119 participants. Color attributions are as in FIG. 10. We are showing two views for the same PCA along two different factorial axes.

**[0198]** FIG. 12: Determination of RNA expression of specific cells in suspension combines flow cytometry with in situ hybridization. After blood draw (a), monocytes in whole blood are labeled with a specific monocyte specific antibody carrying a marker (b). Next, riboprobes that carry a marker for detection are added for in situ hybridization (c). Multi-color flow cytometry detects markers simultaneously (d). Data is analyzed to generate a risk profile (e).

**[0199]** FIG. 13: Example 1 of biosensor approach. After blood draw (a), monocytes in whole blood are labeled with a monocyte specific antibody carrying a magnet to remove them from whole blood (b). Next, RNA is isolated from the monocytes (c). Target RNAs hybridize with their complementary molecular beacons immobilized onto the recognition layer of the biosensor, generating a fluorescent signal that is transduced by the transducer layer of the biosensor (d).

**[0200]** FIG. 14: Example 2 of biosensor approach. After blood draw (a), monocytes in whole blood are labeled with a monocyte specific antibody carrying a magnet to remove them from whole blood (b). Next, unlabelled rib probes are added to the cell suspension for in situ hybridization with target mRNAs (c). Unbound rib probes can hybridize with their complementary molecular beacons immobilized onto the recognition layer of the biosensor, generating a fluorescent signal that is inversely related to the expression of the target mRNAs (d).

#### GENE EXPRESSION PROFILING

**[0201]** The present invention relates to an in vitro method of diagnosing a metabolic syndrome disorder phenotype in a subject, said method comprising: (a) analyzing the level of

IRAK3 expression or activity of expression product in a biological sample isolated from said subject, and (b) compare said level of expression or activity with the IRAK3 expression or activity in a control sample; whereby a decreased level of IRAK3 expression or activity relative to a control sample is an indication of such metabolic syndrome disorder phenotype or a propensity thereto or to an in vitro method of diagnosing a metabolic syndrome disorder, or a propensity thereto in a subject, said method (a) obtaining an expression profile in a biological sample isolated from said subject, wherein said expression profile consists of the analysis of the level of IRAK3 expression in combination with the gene expression level of at least one gene selected from the group consisting of SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 and ZNF217; and (b) comparing said obtained expression profile to a reference expression profile to determine whether said sample is from subject having a metabolic syndrome disorder phenotype or a propensity thereto. Furthermore the present invention relates to an in vitro method of diagnosing adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto in a subject, said method comprising: (a) analyzing the level of IRAK3 expression in a biological sample isolated from said subject, and (b) compare said level of expression with the IRAK3 expression in a control sample; whereby a decreased level of IRAK3 expression relative to a control sample is an indication of such adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto or to an in vitro method of diagnosing an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto, or a propensity thereto in a subject, said method (a) obtaining an expression profile in a biological sample isolated from said subject, wherein said expression profile consists of the analysis of the level of IRAK3 expression in combination with the gene expression level or activity of a gene product of at least one gene selected from the group consisting of SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 and ZNF217; and (b) comparing said obtained expression profile to a reference expression profile to determine whether said sample is from subject having an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto

**[0202]** Furthermore the present invention concerns a method for classifying individuals having or suspected of having a metabolic syndrome disorder phenotype as a responder or a non-responder to first line treatment, comprising the steps of measuring the gene expression of IRAK3 in a biological sample obtained from the individual to obtain a gene expression profile, and comparing the gene expression profile to that of a suitable control.

**[0203]** Another embodiment of present invention is method for classifying individuals having or suspected of having a metabolic syndrome disorder phenotype as a responder or a non-responder to first line treatment, comprising the steps of measuring the gene expression of IRAK3 expression or activity of an IRAK3 expression product in combination with the gene expression level or activity of a gene product of at least one gene selected from the group consisting of SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 and ZNF217; in a biologi-

cal sample obtained from the individual to obtain a gene expression profile, and comparing the gene expression profile to that of a suitable control.

**[0204]** This method is used for predicting the optimal course of therapy for patients having an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto using a diagnostic oligonucleotide set or gene expression profile as described herein, via classification of a individual having or suspected of having a an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto as being either a “responder” or “non-responder” to first-line therapy. In one embodiment, the methods described herein may be used to predict the optimal course of therapy, or identify the efficacy of a given treatment in an individual having, or suspected of having an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto. In other embodiments, the methods described herein may be used to predict the optimal course of therapy post-diagnosis, for example, after treatment of an individual having an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto has begun, such that the therapy may be changed or adjusted, in accordance with the outcome of the diagnostic methods.

**[0205]** The present invention also relates to diagnostic oligonucleotides and diagnostic oligonucleotide sets and methods of using the diagnostic oligonucleotides and oligonucleotide sets to diagnose or monitor disease, assess severity of disease, predict future occurrence of disease, predict future complications of disease, determine disease prognosis, evaluate the patient's risk, “stratify” or classify a group of patients, assess response to current drug therapy, assess response to current non-pharmacological therapy, identify novel therapeutic compounds, determine the most appropriate medication or treatment for the patient, predict whether a patient is likely to respond to a particular drug, and determine most appropriate additional diagnostic testing for the patient, as well as other clinically and epidemiologically relevant applications. As set forth above, the term “diagnostic oligonucleotide set” generally refers to a set of two or more oligonucleotides that, when evaluated for differential expression of their products, collectively yields predictive data. Such predictive data typically relates to diagnosis, prognosis, monitoring of therapeutic outcomes, and the like. In general, the components of a diagnostic oligonucleotide set are distinguished from nucleotide sequences that are evaluated by analysis of the DNA to directly determine the genotype of an individual as it correlates with a specified trait or phenotype, such as a disease, in that it is the pattern of expression of the components of the diagnostic nucleotide set, rather than mutation or polymorphism of the DNA sequence that provides predictive value. It will be understood that a particular component (or member) of a diagnostic nucleotide set can, in some cases, also present one or more mutations, or polymorphisms that are amenable to direct genotyping by any of a variety of well known analysis methods, e.g., Southern blotting, RFLP, AFLP, SSCP, SNP, and the like.

**[0206]** In another embodiment of the present invention, a gene expression system useful for carrying out the described

methods is also provided. This gene expression system can be conveniently used for determining a diagnosis, prognosis, or selecting a treatment for patients having or suspected of having an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto. In one embodiment, the methods disclosed herein allow one to classify an individual of interest as either a "responder" or a "non-responder" to first-line treatment using a gene expression profile. For purposes of the methods disclosed herein, the term "responder" refers to a patient that responds to first line therapy and does not require a second induction of remission during the year following the induction of remission. In contrast, the term "non-responder" refers to a patient having an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto that will require a second induction of remission using any therapy. For example, treatment non-responders may require more than one course of a medicament of the group consisting of pogliptazone, rosiglitazone, netoglitazone, rivoglitazone (CS-011), FK-614, tesaglitazar (AZ-242), ragaglitazar (N,N-622), muraglitazar (BMS-298585), edaglitazone (BM-13-1258), metaglidase (MBX-102), naveglitazar (LY-519818), MX-6054, LY-510929, AMG-131(T-131), THR-0921), voglibose, acarbose, miglitol, emiglitate, phenformin, metformin, buformin, Vidagliptin (LAF237), P32/98, Sitagliptin (MK-431), P93/01, PT-100, saxagliptin (BMS-477118), T-6666, TS-021), AJ-9677, GLP-1, GLP-1MR agent, N,N-2211, AC-2993 (exendin-4), BIM-51077, Aib (8,35)hGLP-1 (7,37)NH<sub>2</sub>, CJC-[131], pramlintide, sodium vanadate, BVT-3498, AS-2868 Ro-28-1675, and GIP (Glucose-dependent insulinotropic peptide) or a medicament of the group consisting of pravastatin, simvastatin, lovastatin, atorvastatin, fluvastatin, rosuvastatin, pitavastatin N—[[[(3R,5S)-1-(3-acetoxy-2,2-dimethylpropyl)-7-chloro-5-(2,3-dimethoxyphenyl)-2-oxo-1,2,3,5-tetrahydro-4,1-benzoxazepin-3-yl]acetyl]piperidine-4-acetic acid, ciprofibrate, fenofibrate, bezafibrate, clofibrate, simfibrate, clinofibrate, Avasimibe, Eflucimibe, colestyramine, probucol, nicomol, niceritrol, ethyl icosapentate, soysterol and  $\gamma$ -oryzanol or a medicament of the group consisting of dexfenfluramine, fenfluramine, phentermine, sibutramine, amfepramone, dexamphetamine, mazindol, phenylpropanolamine, clobenzorex, SB-568849; SNAP-7941, CP-422935, SR-141716, SR-147778, BVT-3498, orlistat, cetilistat (ATL-962)), AJ-9677, leptin, CNTF (Ciliary Neurotropic Factor), linitript, FPL-15849 or a nutraceutical.

**[0207]** Thus, in accordance with the methods, a classification of an individual as a "responder" indicates that first line treatment is likely to be successful in treating the adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto, and as such, may be the treatment of choice, while an individual identified as being a non-responder would generally not be an ideal candidate for traditional first-line therapies.

**[0208]** Rather, an individual identified as a non-responder would likely benefit from more aggressive, or second-line therapies typically reserved for individuals that have not responded to first-line treatment.

**[0209]** Classifying patients as either a "responder" or a "non-responder" is advantageous, in that it allows one to predict the optimal course of therapy for the patient. This

classification may be useful at the outset of therapy (at the time of diagnosis) or later, when first-line therapy has already been initiated, such that treatment may be altered to the benefit of the patient.

**[0210]** In general, the method of using a gene expression profile or gene expression system for diagnosing an individual as a responder or a non-responder comprises measuring the gene expression of IRAK3 and a gene identified in any of tables 3-12 or the sequence listing. Gene expression, as used herein, may be determined using any method known in the art reasonably calculated to determine whether the expression of a gene is upregulated, down-regulated, or unchanged, and may include measurement of RNA or the gene product itself.

**[0211]** In one embodiment, an individual is characterized as a responder or nonresponder to first line therapy via measurement of the expression of IRAK3 one or more genes of tables 3-12 in the individual as compared to the expression of one or more genes of tables 3-12 in a suitable control (such as an individual previously determined to be a responder or non responder). In another embodiment the one or more genes are selected from Table 3. In another embodiment the one or more genes are selected from Table 4. In another embodiment the one or more genes are selected from Table 5. In another embodiment the one or more genes are selected from Table 6. In another embodiment the one or more genes are selected from Table 7. In another embodiment the one or more genes are selected from Table 8. In another embodiment the one or more genes are selected from Table 9. In another embodiment the one or more genes are selected from Table 10. In another embodiment the one or more genes are selected from Table 11. In another embodiment the one or more genes are selected from Table 12. The genes selected for measurement of expression may be selected on the basis of fold difference.

**[0212]** In yet another embodiment, the method of identifying an individual having or suspected of having an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto such as comprises the steps of: 1) providing an array set immobilized on a substrate, wherein the array set comprises one or more oligonucleotides derived from IRAK3 or derived from SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 or ZNF217, or the Sequence Listing in this application, 2) providing a labeled target obtained from mRNA isolated from a biological sample from a patient having an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto, 3) hybridizing the labeled target to the array set under suitable hybridization conditions such that the labeled target hybridizes to the array elements, 4) determining the relative amounts of gene expression in the patient's biological sample as compared to a reference sample by detecting labeled target that is hybridized to the array set; 5) using the gene expression profile to classify the patient as a responder or a non-responder; and 6) predicting the optimal course of therapy based on said classification.

**[0213]** The one or more sequences that comprise the array elements may be selected from any of the sequences listed in Tables 3-12 or the Sequence Listing in this application. In one embodiment, the gene expression system comprises one or more array elements wherein the one or more array elements correspond to sequences selected from those sequences listed in Tables 3-12, or the Sequence Listing. In one embodiment,

the array set comprises the sequences listed on IRAK3. In another embodiment, the array set comprises the sequences listed on of SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 and ZNF217.

**[0214]** The present invention also relates to an apparatus for predicting the optimal course of therapy in a patient having an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto. The apparatus comprises a solid support having an array set immobilized thereon, wherein labeled target derived from mRNA from a patient of interest is hybridized to the one or more sequences of the array set on the solid support, such that a change in gene expression for each sequence compared to a reference sample or other suitable control may be determined, permitting a determination of the optimal course of therapy for the patient. The array set comprises one or more sequences selected from those sequences IRAK3, SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 or ZNF217 listed in the Sequence Listing described herein.

**[0215]** In yet another embodiment, the method of classifying an individual having or suspected of having an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto as a responder or non-responder comprises the steps of: 1) obtaining mRNA isolated from a biological sample from a patient having or suspected of having an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto, 2) reverse transcribing mRNA to obtain the corresponding DNA; 3) selecting suitable oligonucleotide primers corresponding to IRAK3 and one or more genes selected from SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 and ZNF217 which sequence listed herein, 4) combining the DNA and oligonucleotide primers in a suitable hybridization solution; 5) incubating the solution under conditions that permit amplification of the sequences corresponding to the primers; and 6) determining the relative amounts of gene expression in the patient's biological sample as compared to a reference sample or other suitable control; wherein the resulting gene expression profile can be used to classify the patient as a responder or a non-responder. In other embodiments, real time PCR methods or any other method useful in measuring mRNA levels as known in the art may also be used. Alternatively, measurement of one or more gene products using any standard method of measuring protein (such as radioimmunoassay methods or Western blot analysis) may be used to determine a gene expression profile.

**[0216]** The methods of gene expression profiling that may be used with the methods and apparatus described herein are well-known in the art. In general, methods of gene expression profiling can be divided into methods based on hybridization analysis of polynucleotides, and methods based on sequencing of polynucleotides. Commonly used methods known in the art for the quantification of mRNA expression in a sample include northern blotting and in situ hybridization (Parker & Barnes, *Methods in Molecular Biology* 106:247-283 (1999)), RNase protection assays (Hod, *Biotechniques* 13:852-854 (1992)), and reverse transcription polymerase chain reaction (RT-PCR) (Weis et al., *Trends in Genetics* 8:263-264 (1992)), or modified RT-PCR methods, such as that described in U.S. Pat. No. 6,618,679 or in *Methods in Molecular Biology* #193: RT-PCR Protocols by Joe O'Connell Publisher: Humana Press

or Real-Time PCR: Current Technology and Applications Publisher Caister Academic Press January 2009. Alternatively, antibodies, antibody fragments, domain antibodies or nanobodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. Representative methods for sequencing-based gene expression analysis include Serial Analysis of Gene Expression (SAGE) and gene expression analysis by massively parallel signature sequencing (MPSS). In one embodiment described herein, gene array technology such as microarray technology is used to profile gene expression.

#### Arrays and Microarray Technologies

**[0217]** Array and microarray techniques known in the art to determine gene expression may be employed with the invention described herein.

**[0218]** Where used herein, array refers to either an array or microarray. An array is commonly a solid-state grid containing sequences of polynucleotides or oligonucleotides (array elements) of known sequences are immobilized at a particular position (also referred to as an "address") on the grid. Microarrays are a type of array termed as such due to the small size of the grid and the small amounts of nucleotide (such as nanogram, nanomolar or nanoliter quantities) that are usually present at each address. The immobilized array elements (collectively, the "array set") serve as hybridization probes for cDNA or cRNA derived from messenger RNA (mRNA) isolated from a biological sample. An array set is defined herein as one or more DNA fragments or oligonucleotides, as defined above, that are immobilized on a solid support to form an array.

**[0219]** In one embodiment, for example, the array is a "chip" composed, e.g., of one of the above specified materials. Polynucleotide probes, e.g., RNA or DNA, such as cDNA, synthetic oligonucleotides, and the like, or binding proteins such as antibodies, that specifically interact with expression products of individual components of the candidate library are affixed to the chip in a logically ordered manner, i.e., in an array. In addition, any molecule with a specific affinity for either the sense or anti-sense sequence of the marker nucleotide sequence (depending on the design of the sample labeling), can be fixed to the array surface without loss of specific affinity for the marker and can be obtained and produced for array production, for example, proteins that specifically recognize the specific nucleic acid sequence of the marker, ribozymes, peptide nucleic acids (PNA), or other chemicals or molecules with specific affinity.

**[0220]** The techniques described herein, including array and microarray techniques, may be used to compare the gene expression profile of a biological sample from a patient of interest to the gene expression profile of a reference sample or other suitable control. The gene expression profile is determined by first extracting RNA from a biological sample of interest, such as from a patient diagnosed with an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or a propensity thereto. The RNA is then reverse transcribed into cDNA and labeled. In another embodiment, the cDNA may be transcribed into cRNA and labeled. The labeled cDNA or cRNA forms the target that may be hybridized to the array set comprising probes selected according to the methods described herein. The reference sample obtained from a control patient is prepared

in the same way. In one embodiment, both a test sample and reference sample may be used, the targets from each sample being differentially labeled (for example, with fluorophores having different excitation properties), and then combined and hybridized to the array under controlled conditions. In general, the labeled target and immobilized array sets are permitted, under appropriate conditions known to one of ordinary skill in the art, to hybridize such that the targets hybridize to complementary sequences on the arrays. After the array is washed with solutions of appropriately determined stringency to remove or reduce non-specific binding of labeled target, gene expression may be determined. The ratio of gene expression between the test sample and reference sample for a given gene determines the color and/or intensity of each spot, which can then be measured using standard techniques as known in the art. Analysis of the differential gene expression of a given array set provides an "expression profile" or "gene signature" for that array set.

**[0221]** The expression profile is the pattern of gene expression produced by the experimental sample, wherein transcription of some genes are increased or decreased compared to the reference sample.

**[0222]** Amplification methods using in vitro transcription may also be used to yield increased quantities of material to array where sample quantities are limited. In one embodiment, the Nugen Ovation amplification system may be incorporated into the protocol, as described below. Commercially-produced, high-density arrays such as those manufactured by Affymetrix GeneChip (available from Affymetrix, Santa Clara, Calif.) containing synthesized oligonucleotides may be used with the methods disclosed herein. In one embodiment, the HGU133 Plus Version 2 Affymetrix GeneChip may be used to determine gene expression of an array sets comprising sequences listed in Tables 4-8 or the Sequence Listing. In another embodiment, customized cDNA or oligonucleotide arrays may be manufactured by first selecting one or more array elements to be deposited on the array, selected from one or more sequences of IRAK3, SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 or ZNF217 listed in this application or of natural sequence variations on these human genes or homologs and variants of the disclosed nucleic acid molecules. Purified PCR products or other suitably derived oligonucleotides having the selected sequence may then be spotted or otherwise deposited onto a suitable matrix. The support may be selected from any suitable support known in the art, for example, microscope slides, glass, plastic or silicon chips, membranes such as nitrocellulose or paper, fibrous mesh arrangement, nylon filter arrays, glass-based arrays or the like. The array may be a chip array, a plate array, a bead array, a pin array, a membrane array, a solid surface array, a liquid array, an oligonucleotide array, a polynucleotide array, a cDNA array, a microfilter plate, a membrane or a chip. Where transparent surfaces such as microscope slides are used, the support provides the additional advantage of two-color fluorescent labeling with low inherent background fluorescence. The gene expression systems described above, such as arrays or microarrays, may be manufactured using any techniques known in the art, including, for example, printing with fine-pointed pins onto glass slides, photolithography using dynamic micromirror devices, ink-jet printing, or electrochemistry on microelectrode arrays.

**[0223]** Oligonucleotide adherence to the slide may be enhanced, for example, by treatment with polylysine or other cross-linking chemical coating or by any other method known

in the art. The DNA or oligonucleotide may then be cross-linked by ultraviolet irradiation and denatured by exposure to either heat or alkali. The microarray may then be hybridized with labeled target derived from mRNA from one or more samples to be analyzed. For example, in one embodiment, cDNA or cRNA obtained from mRNA from colon samples derived from both a patient diagnosed with metabolic syndrome and a healthy control sample is used. The samples may be labeled with different detectable labels such as, for example, fluorophores that exhibit different excitation properties. The samples may then be mixed and hybridized to a single microarray that is then scanned, allowing the visualization of up-regulated or down-regulated genes. The Dual-Chip™ platform available from Eppendorf is an example of this type of array.

**[0224]** The probes affixed to the solid support in the gene expression system comprising the array elements may be a candidate library, a diagnostic agent, a diagnostic oligonucleotide set or a diagnostic probe set. In one embodiment of the present invention, the one or more array elements comprising the array set are selected from those sequences of IRAK3, SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 or ZNF217 listed in this application or of natural sequence variations on these human genes or homologs and variants of the disclosed nucleic acid molecules.

#### Determination of Array Sets

**[0225]** A global pattern of gene expression in biological samples such as blood cells, blood monocyte preferable a myeloid cell from patients with adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease at diagnosis (CDD), treated CD patients refractory to first line therapy (chronic refractory, CDT), and healthy controls can be determined. cRNA can be prepared from biopsies obtained from endoscopically affected segments, predominantly the ascending colon, with control biopsies obtained from matched segments in healthy patients. cRNA can be labeled and then hybridized. RNA obtained from a pool of RNA from one normal specimen can be labeled and hybridized with each batch of new samples to serve as an internal control for batch to batch variability in signal intensity. Results can be interpreted utilizing a proper software e.g. GeneSpring™ 7.3 Software (Silicon Genetics). Differentially expressed genes are then identified by filtering levels of gene-specific signal intensity for statistically significant differences when grouped by clinical forms (e.g. healthy control versus diagnosis for adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease and healthy control versus chronic refractory to first line therapy).

#### Determination of a Gene Expression Profile

**[0226]** The present invention is related to methods of detecting gene expression using a gene expression system having one or more array elements wherein the array elements comprise one or more sequence that corresponds to sequence selected from IRAK3, SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 or ZNF217 listed in this application or of natural sequence variations on these human genes or homologs and variants of the disclosed nucleic acid molecules, forming an array set. From IRAK3, SOD2, TNFAIP6, TNFAIP3, TLR2,

IRS2 or ZNF217 listed in this application or of natural sequence variations on these human genes or homologs and variants of the disclosed nucleic acid molecules, it should be understood by one of ordinary skill in the art that standard methods of data analysis or using the disclosed methods (such as cluster analysis, K-nearest neighbors class prediction algorithms, or class prediction analysis using appropriately selected parameters) can be used to identify a smaller number of array elements, while still retaining the predictive characteristics of the array sets disclosed herein. Non-limiting examples of data analysis that may be used are listed below. In one embodiment, an array may be used to determine gene expression as described above. For example, PCR amplified inserts of cDNA clones may be applied to a substrate in a dense array. These cDNA may be selected from one or more of those sequences of IRAK3, SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 or ZNF217 listed in this application or of natural sequence variations on these human genes or homologs and variants of the disclosed nucleic acid molecules.

**[0227]** The probes, immobilized on the selected substrate, are suitable for hybridization under conditions with appropriately determined stringency, such that targets binding non-specifically to the substrate or array elements are substantially removed. Appropriately labeled targets generated from mRNA are generated using any standard method as known in the art. For example, the targets may be cDNA targets generated through incorporation of fluorescent nucleotides by reverse transcription of RNA extracted from tissues of interest.

**[0228]** Alternatively, biotin labeled targets may be used, such as using the method described herein. It should be clear that any suitable oligonucleotide-based target may be used. In another embodiment, suitably labeled cRNA targets may be used. Regardless of the type of target, the targets are such that the labeled targets applied to the chip hybridize to complementary probes on the array. After washing to minimize non-specific binding, the chip may be scanned by confocal laser microscopy or by any other suitable detection method known in the art, for example, a CCD camera. Quantification of hybridization at each spot in the array allows a determination of corresponding mRNA expression. With dual color fluorescence, separately labeled cDNA targets generated from two sources of RNA are hybridized pairwise to the array. The relative abundance of the transcripts from the two sources corresponding to each specified gene can then be determined simultaneously. (See Schena et al., Proc. Natl. Acad. Sci. USA 93(2): 106 149 (1996)). Microarray analysis can be performed by commercially available equipment, following manufacturer's protocols, such as by using the Affymetrix GenChip technology, or Incyte's microarray technology, or using any other methods as known in the art.

**[0229]** Homologs and variants of these nucleic acid molecules such of IRAK3, SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 or ZNF217 listed in this application typically possess a relatively high degree of sequence identity when aligned using standard methods. Sequences suitable for use in the methods described herein have at least about 40-50, about 50-60, about 70-80, about 80-85, about 85-90, about 90-95 or about 95-100% sequence identity to the sequences disclosed herein. It is understood that for determination of a gene expression profile, variations in the disclosed sequences will still permit detection of gene expression. The degree of sequence identity required to detect gene expression varies depending on the length of the oligomer. For example, in a

60-mer, (an oligonucleotide with about 60 nucleotides), about 6 to about 8 random mutations or about 6 to about 8 random deletions in a 60-mer do not affect gene expression detection. Hughes, T R, et al. "Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. Nature Biotechnology, 19:343-347 (2001). As the length of the DNA sequence is increased, the number of mutations or deletions permitted while still allowing gene expression detection is increased. As will be appreciated by those skilled in the art, the sequences of the present invention may contain sequencing errors. That is, there may be incorrect nucleotides, frame shifts, unknown nucleotides, or other types of sequencing errors in any of the sequences; however, the correct sequences will fall within the homology and stringency definitions herein.

#### Additional Methods of Determining Gene Expression

**[0230]** The array sets disclosed herein may also be used to determine a gene expression profile such that a patient may be classified as a responder or a nonresponder any other techniques that measure gene expression. For example, the expression of genes disclosed in the array sets herein may be detected using RT-PCR methods or modified RT-PCR methods. In this embodiment, RT-PCR is used to detect gene expression of genes selected from one or more genes selected from of IRAK3, SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2 or ZNF217 listed in this application or of natural sequence variations on these human genes or homologs and variants of the disclosed nucleic acid molecules. Various methods using RT-PCR may be employed. For example, standard RT-PCR methods may be used. Using this method, well-known in the art, isolated RNA may be reverse transcribed using into cDNA using standard methods as known in the art. This cDNA is then exponentially amplified in a PCR reaction using standard PCR techniques. The reverse transcription step is typically primed using specific primers, random hexamers, or oligo-dT primers, depending on the circumstances and the goal of expression profiling. For example, extracted RNA can be reverse-transcribed using a GeneAmp RNA PCR kit (Perkin Elmer, Calif., USA), following the manufacturer's instructions. The derived cDNA can then be used as a template in the subsequent PCR reaction. Although the PCR step can use a variety of thermostable DNA-dependent DNA polymerases, it typically employs the Taq DNA polymerase, which has a 5'-3' nuclease activity but lacks a 3'-5' proofreading endonuclease activity. Thus, TaqMan® PCR typically utilizes the 5'-nuclease activity of Taq or Tth polymerase to hydrolyze a hybridization probe bound to its target amplicon, but any enzyme with equivalent 5' nuclease activity can be used. Two oligonucleotide primers are used to generate an amplicon typical of a PCR reaction. A third oligonucleotide is designed to detect nucleotide sequence located between the two PCR primers. The third oligonucleotide is non-extendible by Taq DNA polymerase enzyme, and is labeled with a reporter fluorescent dye and a quencher fluorescent dye. Any laser-induced emission from the reporter dye is quenched by the quenching dye when the two dyes are located close together as they are on the probe. During the amplification reaction, the Taq DNA polymerase enzyme cleaves the third oligonucleotide in a template-dependent manner. The resultant fragments dissociate in solution, and signal from the released reporter dye is free from the quenching effect of the second fluorophore. One molecule of reporter dye is liberated for each new molecule synthesized, and detection of the

unquenched reporter dye provides the basis for quantitative interpretation of the data. TaqMan® RT-PCR can be performed using commercially available equipment, such as, for example, ABI PRISM 7700™ Sequence Detection System™ (Perkin-Elmer-Applied Biosystems, Foster City, Calif., USA), or Lightcycler (Roche Molecular Biochemicals, Mannheim, Germany). In a preferred embodiment, the 5' nuclease procedure is run on a real-time quantitative PCR device such as the ABI PRISM 7700™ Sequence Detection System™. The system consists of a thermocycler, laser, charge-coupled device (CCD), camera and computer. The system amplifies samples in a 96-well format on a thermocycler. During amplification, laser-induced fluorescent signal is collected in real-time through fiber optics cables for all 96 wells, and detected at the CCD. The system includes software for running the instrument and for analyzing the data. To minimize errors and the effect of sample-to-sample variation, RT-PCR is usually performed using an internal standard. The ideal internal standard is expressed at a constant level among different tissues, and is unaffected by the experimental treatment. RNAs most frequently used to normalize patterns of gene expression are mRNAs for the housekeeping genes glyceraldehyde-3-phosphate-dehydrogenase (GAPDH) and  $\beta$ -actin, although any other housekeeping gene or other gene established to be expressed at constant levels between comparison groups can be used. Real time quantitative PCR techniques, which measure PCR product accumulation through a dual-labeled fluorogenic target (i.e., TaqMan® probe), may also be used with the methods disclosed herein to determine a gene expression profile. The Stratagene Brilliant SYBR Green QPCR reagent, available from 11011 N. Torrey Pines Road, La Jolla, Calif. 92037, may also be used. The SYBR® Green dye binds specifically to double-stranded PCR products, without the need for sequence-specific targets. Real time PCR is compatible both with quantitative competitive PCR, where internal competitor for each target sequence is used for normalization, and with quantitative comparative PCR using a normalization gene contained within the sample, or a housekeeping gene for RT-PCR. For further details see, e.g. Held et al., *Genome Research* 6:986-994 (1996). Alternatively, a modified RT-PCR method such as express Profiling™ (XP) technology for high-throughput gene expression analysis, available from Althea Technologies, Inc. 11040 Roselle Street, San Diego, Calif. 92121 U.S.A. may be used to determine a gene expression profiles of a patient diagnosed with metabolic syndrome. The gene expression analysis may be limited to one or more array sets as disclosed herein. This technology is described in U.S. Pat. No. 6,618,679, incorporated herein by reference. This technology uses a modified RT-PCR process that permits simultaneous, quantitative detection of expression levels of about 20 genes. This method may be complementary to or used in place of array technology or PCR and RT-PCR methods to determine or confirm a gene expression profile, for example, when classifying the status of a patient as a responder or non-responder. Multiplex mRNA assays may also be used, for example, that described in Tian, et al., "Multiplex mRNA assay using Electrophoretic tags for high-throughput gene expression analysis," *Nucleic Acids Research* 2004, Vol. 32, No. 16, published online Sep. 8, 2004 and Elnifro, et al. "Multiplex PCR: Optimization and Application in Diagnostic Virology," *Clinical Microbiology Reviews*, October 2000, p. 559-570, both incorporated herein by reference. In multiplex CR, more than one target sequence can be amplified by including more than one pair of primers

in the reaction. Alternatively, Real-Time NASBA can be used. This is an isothermal nucleic acid amplification method which is particularly suited to detection and quantification of genomic, ribosomal or messenger RNA and which has been proved to be the basis of sensitive and specific assays for detection, quantification and differentiation of RNA and DNA targets. The product of NASBA is single-stranded RNA of opposite sense to the original target. The first developed NASBA methods relied on liquid or gel-based probe-hybridization for post-amplification detection of products. But more recently, real-time procedures incorporating amplification and detection in a single step have been reported and applied to a wide range of RNA and some DNA targets (Loens K. et al. *Journal of microbiological methods*. 2006, vol. 67, no3, pp. 408-415, Schneider et al. *Real-time nucleic acid sequence-based amplification is more convenient than real-time PCR for quantification of Plasmodium falciparum*. *J Clin Microbiol*. 2005; 43(1):402-5 and Keightley et al. *Real-time NASBA detection of SARS-associated coronavirus and comparison with real-time reverse transcription-PCR*. *J Med. Virol*. 2005; 77(4):602-8.) Another alternative is that can be used is the one-step, reverse transcription loop-mediated amplification (RT-LAMP) assay. RT-LAMP is sensitive, rapid, specific, and cost effective, with the potential for field deployment and use in first line diagnosis. (Horibe, D., et al. Rapid detection of metastasis of gastric cancer using reverse transcription loop-mediated isothermal amplification. *Int. J. Cancer* (2007)

**[0231]** The methods disclosed herein employ a biological sample derived from patients diagnosed with metabolic syndrome. The samples may include, for example, tissue samples obtained by biopsy of endoscopically affected colonic segments including the cecum/ascending, transverse/descending or sigmoid/rectum; small intestine; ileum; intestine; cell lysates; serum; or blood samples.

**[0232]** Colon epithelia cells and lamina propria cells may be used for mRNA isolation. Control biopsies are obtained from the same source.

**[0233]** Sample collection will depend on the target tissue or sample to be assayed.

**[0234]** Immediately after collection of a biological sample, for instance the blood cell blood cells, blood monocyte preferable a myeloid cell from patients with adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease or an healthy patient the sample may be placed in a medium appropriate for storage of the sample such that degradation of mRNA is minimized and stored on ice. For example, a suitable medium for storage of sample until processing is RNALater®, available from Applied Biosystems, 850 Lincoln Centre Drive, Foster City Calif. 94404, U.S.A. Total RNA may then be prepared from a target sample using standard methods for RNA extraction known in the art and disclosed in standard textbooks of molecular biology, including Ausubel et al., *Current Protocols of Molecular Biology*, John Wiley and Sons (1997). For example, RNA isolation can be performed using purification kit, buffer set and protease from commercial manufacturers, such as Qiagen, according to the manufacturer's instructions. In one embodiment, total RNA is prepared utilizing the Qiagen RNeasy mini-column, available from QIAGEN Inc., 27220 Turnberry Lane Suite 200, Valencia, Calif. 91355. Other commercially available RNA isolation kits include MasterPure™ Complete DNA and RNA Purification Kit (EPICENTRE®),

Madison, Wis.), or Paraffin Block RNA Isolation Kit (Ambion, Inc.). Total RNA from tissue samples may also be isolated using RNA Stat-60 (Tel-Test). RNA may also be prepared, for example, by cesium chloride density gradient centrifugation. RNA quality may then be assessed. RNA quality may be determined using, for example, the Agilent 2100 Bioanalyzer. Acceptable RNA samples have distinctive 18S and 28S Ribosomal RNA Bands and a 28S/18S ribosomal RNA ratio of about 1.5 to about 2.0. In one embodiment, about 400 to about 500 nanograms of total RNA per sample is used to prepare labeled mRNA as targets. The RNA may be labeled using any methods known in the art, including for example, the TargetAmp 1-Round Aminoallyl-aRNA Amplification Kit available from Epicentre to prepare cRNA, following the manufacturer's instructions. The TargetAmp 1-Round Aminoallyl-aRNA Amplification Kit (Epicentre) is used to make double-stranded cDNA from total RNA. An in vitro transcription reaction creates cRNA target. Biotin-X-X-NHS (Epicentre) is used to label the aminoallyl-aRNA with biotin following the manufacturer's instructions. In one embodiment, the biotin-labeled cRNA target is then chemically fragmented and a hybridization cocktail is prepared and hybridized to a suitable array set immobilized on a suitable substrate. In another embodiment, the total RNA may be used to prepare cDNA targets. The targets may be labeled using any suitable labels known in the art. The labeled cDNA targets may then be hybridized under suitable conditions to any array set or subset of an array set described herein, such that a gene expression profile may be obtained.

#### Normalization

**[0235]** Normalization is an adjustment made to microarray gene expression values to correct for potential bias or error introduced into an experiment. With respect to array-type analyses, such errors may be the result of unequal amounts of cDNA probe, differences in dye properties, differences in dye incorporation etc. Where appropriate, the present methods include the step of normalizing data to minimize the effects of bias or error. The type of normalization used will depend on the experimental design and the type of array being used. The type of normalization used will be understood by one of ordinary skill in the art.

#### Levels of Normalization

**[0236]** There may be two types of normalization levels used with the methods disclosed herein: "within slide" (this compensates, for example, for variation introduced by using different printing pins, unevenness in hybridization or, in the case of two channel arrays, differences in dye incorporation between the two samples) or "between slides," which is sometimes referred to as "scaling" and permits comparison of results of different slides in an experiment, replicates, or different experiments.

#### Normalization Methods

**[0237]** Within slide normalization can be accomplished using local or global methods as known in the art. Local normalization methods include the use of "housekeeping genes" and "spikes" or "internal controls". "Housekeeping" genes are genes which are known, or expected, not to change in expression level despite changes in disease state or phenotype or between groups of interest (such as between known non-responders and responders). For example, common

housekeeping genes used to normalize data are those that encode for ubiquitin, actin and elongation factors. Where housekeeping genes are used, expression intensities on a slide are adjusted such that the housekeeping genes have the same intensity in all sample assays. Normalization may also be achieved using spikes or internal controls that rely on RNA corresponding to particular probes on the microarray slide being added to each sample. These probes may be from a different species than the sample RNAs and optimally should not cross-hybridize to sample RNAs. For two channel arrays, the same amount of spike RNA is added to each sample prior to labeling and normalization is determined via measurement of the spiked features. Spikes can also be used to normalize spatially across a slide if the controls have been printed by each pin—the same controls on different parts of the slide should hybridize equally. Spikes may also be used to normalize between slides.

**[0238]** Reference samples may be any suitable reference sample or control as will be readily understood by one of skill in the art. For example, the reference sample may be selected from normal patients, "responder" patients, "non-responder" patients, or "chronic-refractory patients." Normal patients are those not diagnosed with an adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease. "Responder" patients and "non-responder" patients are described above. "Chronic refractory" patients are patients with moderate to severe disease that require a second induction of remission using any drug. In one embodiment of the present invention, the control sample comprises cDNA from one or more patients that do not have adipose tissue disorder, lipid homeostasis disorder and/or an impaired glucose tolerance or insulin resistance condition and/or related cardiovascular disease. In this embodiment, the cDNA of multiple normal samples are combined prior to labeling, and used as a control when determining gene expression of experimental samples. The data obtained from the gene expression analysis may then be normalized to the control cDNA.

**[0239]** A variety of global normalization methods may be used including, for example, linear regression. This method is suitable for two channel arrays and involves plotting the intensity values of one sample against the intensity values of the other sample. A regression line is then fitted to the data and the slope and intercept calculated. Intensity values in one channel are then adjusted so that the slope=1 and the intercept is 0. Linear regression can also be carried out using MA plots. These are plots of the log ratio between the Cy5 and Cy3 channel values against the average intensity of the two channels. Again regression lines are plotted and the normalized log ratios are calculated by subtracting the fitted value from the raw log ratio. In the alternative, Lowess regression (locally weighted polynomial regression) may be used. This regression method again uses MA plots but is a non-linear regression method. This normalization method is suitable if the MA plots show that the intensity of gene expression is influencing the log ratio between the channels. Lowess essentially applies a large number of linear regressions using a sliding window of the data. Yet another alternative method of normalization is "print tip normalization." This is a form of spatial normalization that relies on the assumption that the majority of genes printed with individual print tips do not show differential expression. Either linear or non-linear regression can be used to normalize the data. Data from features printed by different print tips are normalized inde-

pendently. This type of normalization is especially important when using single channel arrays. Yet another method of normalization is "2D Lowess normalization." This form of spatial normalization uses a 2d polynomial Lowess regression that is fitted to the data using a false color plot of log ratio or intensity as a function of the position of the feature on the array. Values are adjusted according to this polynomial. "Between slide normalization" enables you to compare results from different slides, whether they are two channel or single channel arrays.

**[0240]** Centering and scaling may also be used. This adjusts the distributions of the data (either of log ratios or signal intensity) on different slides such that the data is more similar. These adjustments ensure that the mean of the data distribution on each slide is zero and the standard deviation is 1. For each value on a slide, the mean of that slide is subtracted and the resulting value divided by the standard deviation of the slide. This ensures that the "spread" of the data is the same in each slide you are comparing. Quantile normalization is yet another method that is particularly useful for comparing single channel arrays. Using this method, the data points in each slide are ranked from highest to lowest and the average computed for the highest values, second highest values and so on. The average value for that position is then assigned to each slide, i.e. the top ranked data point in each slide becomes the average of the original highest values and so on. This adjustment ensures that the data distributions on the different slides are identical. Various tools for normalizing data are known in the art, and include GenePix, Excel, GEPAS, TMeV/MIDAS and R.

#### Hybridization Techniques

**[0241]** [Where array techniques are used to determine a gene expression profile, the targets must be hybridized to the array sets under suitable hybridization conditions using hybridization and wash solutions having appropriate stringency, such that labeled targets may hybridize to complementary probe sequences on the array. Washes of appropriate stringency are then used to remove non-specific binding of target to the array elements or substrate. Determination of appropriate stringency is within the ordinary skill of one skilled in the art. In one embodiment of the present invention, the array set is that of the Affymetrix Genechip Array (HGU133 Plus Version 2 Affymetrix GeneChip, available from Affymetrix, 3420 Central Expressway, Santa Clara, Calif. 95051). In this embodiment, suitably labeled cRNA and hybridization cocktail are first prepared. In this embodiment, the hybridization cocktail contains about 0.034 ug/uL fragmented cRNA, about 50 pM Control Oligonucleotide B2 (available from Affymetrix), 20x Eukaryotic Hybridization Controls (1.5 pM bioB, 5 pM blot, 25 pM bioD, 100 pM ere) (available from Affymetrix), about 0.1 mg/mL Herring Sperm DNA (Promega), about 0.5 mg/mL Acetylated BSA (Invitrogen). The hybridization cocktail is heated to 99° C. for 5 minutes, to 45° C. for 5 minutes, and spun at maximum speed in a microcentrifuge for 5 minutes. The probe array is then filled with 200 uL of 1x Hybridization Buffer (available from Affymetrix) and incubated at 45° C. for 10 minutes while rotating at 60 rpm. The 1x Hybridization Buffer is removed and the probe array filled with 200 uL of the hybridization cocktail. The probe array is then incubated at 45° C. for about 16 hours in a hybridization oven rotating at 60 rpm. The array is then washed and stained using any method as known in the art. In one embodiment, the Fluidics Station 450

(Affymetrix) and the fluidics protocol EukGE-WS2v4\_450 is used. This protocol comprises the steps of a first post-hybridization wash (10 cycles of 2 mixes/cycle with Affymetrix Wash Buffer A at 25° C.), a second post-hybridization wash (4 cycles of 15 mixes/cycle with Affymetrix Wash Buffer B at 50° C.), a first stain (staining the probe array for 10 minutes with Affymetrix Stain Cocktail 1 at 25° C.), a post-stain wash (10 cycles of 4 mixes/cycle with Affymetrix Wash Buffer A at 25° C.), a second stain (stain the probe array for 10 minutes with Stain Cocktail 2 at 25° C.), a third stain (stain the probe array for 10 minutes with Stain Cocktail 3 at 25° C.) and a final wash (15 cycles of 4 mixes/cycle with Wash Buffer A at 30° C. The holding temperature is 25° C.). All Wash Buffers and Stain Cocktails are those provided in the GeneChip® Hybridization, Wash and Stain Kit, Manufactured for Affymetrix, Inc., by Ambion, Inc., available from Affymetrix. In one embodiment, the stain used is R-Phycoerythrin Streptavidin, available from Molecular Probes. The antibody used is anti-streptavidin antibody (goat) biotinylated, available from Vector Laboratories.

#### Data Collection and Processing

**[0242]** When using an array to determine a gene expression profile, the data from the array must be obtained and processed. The data may then be used for any of the purposes set forth herein, such as to predict the outcome of a therapeutic treatment or to classify a patient as a responder or non-responder. Following appropriate hybridization and wash steps, the substrate containing the array set and hybridized target is scanned. Data is then collected and may be saved as both an image and a text file. Precise databases and tracking of files should be maintained regarding the location of the array elements on the substrates. Information on the location and names of genes should also be maintained. The files may then be imported to software programs that perform image analysis and statistical analysis functions. The gene expression profile of a patient of interest is then determined from the collected data. This may be done using any standard method that permits qualitative or quantitative measurements as described herein. Appropriate statistical methods may then be used to predict the significance of the variation in the gene expression profile, and the probability that the patient's gene expression profile is within the category of non-responder or responder. For example, in one embodiment, the data may be collected, then analyzed such that a class determination may be made (i.e., categorizing a patient as a responder or non-responder) using a class prediction algorithm and GeneSpring™ software as described below. Expression patterns can be evaluated by qualitative and/or quantitative measures. Qualitative methods detect differences in expression that classify expression into distinct modes without providing significant information regarding quantitative aspects of expression. For example, a technique can be described as a qualitative technique if it detects the presence or absence of expression of a candidate nucleotide sequence, i.e., an on/off pattern of expression. Alternatively, a qualitative technique measures the presence (and/or absence) of different alleles, or variants, of a gene product. In contrast, some methods provide data that characterize expression in a quantitative manner. That is, the methods relate expression on a numerical scale, e.g., a scale of 0-5, a scale of 1-10, a scale of +++++, from grade 1 to grade 5, a grade from a to z, or the like. It will be understood that the numerical, and symbolic examples provided are arbitrary, and that any graduated scale (or any

symbolic representation of a graduated scale) can be employed in the context of the present invention to describe quantitative differences in nucleotide sequence expression. Typically, such methods yield information corresponding to a relative increase or decrease in expression. Any method that yields either quantitative or qualitative expression data is suitable for evaluating expression. In some cases, e.g., when multiple methods are employed to determine expression patterns for a plurality of candidate nucleotide sequences, the recovered data, e.g., the expression profile, for the nucleotide sequences is a combination of quantitative and qualitative data. In some applications, expression of the plurality of candidate nucleotide sequences is evaluated sequentially. This is typically the case for methods that can be characterized as low- to moderate-throughput. In contrast, as the throughput of the elected assay increases, expression for the plurality of candidate nucleotide sequences in a sample, or multiple samples is assayed simultaneously.

**[0243]** Again, the methods (and throughput) are largely determined by the individual practitioner, although, typically, it is preferable to employ methods that permit rapid, e.g. automated or partially automated, preparation and detection, on a scale that is time-efficient and cost-effective.

**[0244]** It is understood that the preceding discussion is directed at both the assessment of expression of the members of candidate libraries and to the assessment of the expression of members of diagnostic nucleotide sets. Many techniques have been applied to the problem of making sense of large amounts of gene expression data. Cluster analysis techniques (e.g., K-Means), self-organizing maps (SOM), principal components analysis (PCA), and other analysis techniques are all widely available in packaged software used in correlating this type of gene expression data.

#### Class Prediction

**[0245]** In one embodiment, the data obtained may be analyzed using a class prediction algorithm to predict whether a subject is a non-responder or a responder, as defined above. Class prediction is a supervised learning method in which the algorithm learns from samples with known class membership (the training set) and establishes a prediction rule to classify new samples (the test set). Class prediction consists of several steps. The first is a feature selection i.e. a process by which genes within a defined gene set are scored for their ability to distinguish between classes (responders and non-responders) in the training set. Genes may be selected for uses as predictors, by individual examination and ranking based on the power of the gene to discriminate responders from non-responders. Genes may then be scored on the basis of the best prediction point for responders or non-responders. The score function is the negative natural logarithm of the p-value for a hyper geometric test of predicted versus actual group membership for responder versus non-responder. A combined list for responders and non-responders for the most discriminating genes may then be produced, up to the number of predictor genes specified by the user. The Golub method may then be used to test each gene considered for the predictor gene set for its ability to discriminate responders from non-responders using a signal-to-noise ratio. Genes with the highest scores may then be kept for subsequent calculations. A subset of genes with high predictive strength may then used in class prediction, with cross validation performed using the known groups from the training set. The K-nearest neighbors approach may be used to classify training set samples during

cross validation, and to classify test set samples once the predictive rule had been established. In this system, each sample is classified by finding the K-nearest neighboring training set samples (where K is the number of neighbors defined by the user) plotted based in Euclidean space over normalized expression intensity for each of the genes in the predictor set. For example, a predictive gene set of twenty members may be selected using four nearest neighbors. Depending on the number of samples available, the k value may vary. The class membership of the selected number of nearest neighbors to each sample is enumerated and p-values computed to determine the likelihood of seeing at least the observed number of neighbors from each class relative to the whole training set by chance in a K-sized neighborhood. With this method, the confidence in class prediction is best determined by the ratio of the smallest p-value and the second smallest p-value, termed the decision cut-off p-value. If it is lower, the test sample is classified as the class corresponding to the smallest p-value. If it is higher, a prediction is not made. In one embodiment, a decision cut-off p-value ratio of about 0.5 may be used. Cross validation in GeneSpring may then be then done by a drop-one-out algorithm, in which the accuracy of the prediction rule is tested. This approach removes one sample from the training set and uses it as a test sample. By predicting the class of a given sample only after it is removed from the training set, the rule makes unbiased prediction of the sample class. Once performance of the predictive rule has been optimized in this fashion, it may be tested using additional samples.

#### Cluster Analysis

**[0246]** Cluster analysis is a loose term covering many different algorithms for grouping data. Clustering can be divided into two main types: top-down and bottom-up. Top-down clustering starts with a given number of clusters or classes and proceeds to partition the data into these classes. Bottom-up clustering starts by grouping data at the lowest level and builds larger groups by bringing the smaller groups together at the next highest level's-Means is an example of top-down clustering. K-means groups data into K number of best-fit clusters. Before using the algorithm, the user defines the number of clusters that are to be used to classify the data (K clusters). The algorithm randomly assigns centers to each cluster and then partitions the nearest data into clusters with those centers. The algorithm then iteratively finds new centers by averaging over the data in the cluster and reassigning data to new clusters as the centers change. The analysis iteratively continues until the centers no longer move (Sherlock, G., *Current Opinion in Immunology*, 12:201, 2000). Tree clustering is an example of bottom-up clustering. Tree clustering joins data together by assigning nearest pairs as leaves on the tree. When all pairs have been assigned (often according to either information-theoretical criteria or regression methods), the algorithm progresses up to the next level joining the two nearest groups from the prior level as one group. Thus, the number and size of the clusters depends on the level. Often, the fewer clusters, the larger each cluster will be. The stoppage criteria for such algorithms varies, but often is determined by an analysis of the similarity of the members inside the cluster compared to the difference across the clusters. Self-organizing maps (SOMs) are competitive neural networks that group input data into nearest neighbors (Torkkola, K., et al., *Information Sciences*, 139:79, 2001; Toronen, P., et al., *FEBS Letters*, 451:142 146, 1999). As data is presented to

the neural network, neurons whose weights currently are capable of capturing that data (the winner neuron) are updated toward the input. Updating the weights, or training the neural net, shifts the recognition space of each neuron toward a center of similar data. SOMs are similar to K-means with the added constraint that all centers are on a 1 or 2 dimensional manifold (i.e., the feature space is mapped into a 1 or 2 dimensional array, where new neighborhoods are formed). In SOM, the number of neurons is chosen to be much larger than the possible number of the clusters. It is hoped that the clusters of trained neurons will provide a good estimation of the number of the neurons. In many cases, however, a number of small clusters are formed around the larger clusters, and there is no practical way of distinguishing such smaller clusters from, or of merging them into, the larger clusters. In addition, there is no guarantee that the resulting clusters of genes actually exhibit statistically independent expression profiles. Thus, the members of two different clusters may exhibit similar patterns of gene expression. Principal component analysis (PCA), although not a clustering technique in its nature (Jolliffe, I. T., *Principal Component Analysis*, New York: Springer-Verlag, 1986) can also be used for clustering (Yeung, K. Y., et al., *Bioinformatics*, 17:763, 2001). PCA is a stepwise analysis that attempts to create a new component axis at each step that contains most of the variation seen for the data. Thus, the first component explains the first most important basis for the variation in the data, the second component explains the second most important basis for the variation in the data, the third component the third most important basis, and so on. PCA projects the data into a new space spanned by the principal components. Each successive principal component is selected to be orthogonal to the previous ones, and to capture the maximum information that is not already present in the previous components. The principal components are therefore linear combinations (or eigenarrays) of the original data. These principal components are the classes of data in the new coordinate generated by PCA. If the data is highly non-correlated, then the number of significant principal components can be as high as the number of original data values. If, as in the case of DNA microarray experiments, the data is expected to correlate among groups, then the data should be described by a set of components which is fewer than the full complement of data points. A variety of systems known in the art may be used for image analysis and compiling the data. For example, where the mRNA is labeled with a fluorescent tag, and fluorescence imaging system (such as the microarray processor commercially available from AFFYMETRIX®, Santa Clara, Calif.) may be used to capture, and quantify the extent of hybridization at each address. Or, in the case where the mRNA is radioactive, the array may be exposed to X-ray film and a photographic image made. Once the data is collected, it may be compiled to quantify the extent of hybridization at each address as for example, using software to convert the measured signal to a numerical value. Any publicly available imaging software may be used. Examples include BioDiscovery (ImaGene), Axon Instruments (GenePix Pro 6.0), Eisen-Lab—Stanford University (ScanAlyze), Spotfinder (TIGR), Imaxia (ArrayFox), F-Scan (Analytical Biostatistics Section—NIH), MicroDiscovery (GeneSpotter), CLONDIAG (IconoClust), Koda Technology (Koadarray), Vigena Tech (MicroVigena), Nonlinear Dynamics (Phoretix), CSIRO Mathematical and Information Sciences (SPOT) Niles Scientific (SpotReader). Any commercially available data analysis

software may also be used. Examples include, BRB Array Tools (Biometric Research Branch—NCI), caGEDA (University of Pittsburgh), Cleaver 1.0 (Stanford Biomedical Informatics), ChipSC2C (Peterson Lab—Baylor College of Medicine), Cluster (Eisen Lab—Stanford/UC Berkeley), DNA-Chip Analyzer (dChip) (Wong Laboratory—Harvard University), Expression Profiler (European Bioinformatics Institute), FuzzyK (Eisen Lab—Stanford/UC Berkeley), GeneCluster 2.0 (Broad Institute), GenePattern (Broad Institute), GeneXPress (Stanford University), Genesis (Alexander Stum—Graz University of Technology), GEPAS (Spanish National Cancer Center), GLR (University of Utah), GQL (Max Planck Institute for Molecular Genetics), INCLUSive (Katholieke Universiteit Leuven), Maple Tree (Eisen Lab—Stanford/UC Berkeley) MeV (TIGR) MIDAS (TIGR), OntoTools (Sorin Draghici—Wayne State University), Short Time-series Expression Miner (Carnegie Mellon University), Significance Analysis of Microarrays (Rob Tibshirani—Stanford University), SNOMAD (Johns Hopkins Schools of Medicine and Public Health), SparseLOGREG (Shevade & Keerthi—National University of Singapore), SuperPC Microarrays (Rob Tibshirani—Stanford University), TableView (University of Minnesota), TreeView (Eisen Lab—Stanford/UC Berkeley), Venn Mapper (Universitais Medisch Centrum Rotterdam), Applied Maths (GeneMaths XT), Array Genetics (AffyMate), Axon Instruments (Acuity 4.0) BioDiscovery (GeneSight), BioSieve (ExpressionSieve), CytoGenomics (SilicoCyte), Microarray Data Analysis (GeneSifter), MediaCybernetics (ArrayPro Analyzer), Microarray Fuzzy Clustering (BioRainbow), Molmine (J-Express Pro), Optimal Design (ArrayMiner), Partek (Partek Pro) Predictive Patterns Software (GeneLinker), Promoter Extractor (BioRainbow) SAS Microarray Silicon Genetics (GeneSpring), Spotfire (Spotfire), Strand Genomics (Avadis) Vialogy Corp. It should also be understood that confounding factors may exist in individual subjects that may affect the ability of a given gene set to predict responders versus non-responders. These confounding variables include variation in medications, such as cases in which concurrent 6-MP with infliximab overcomes the adverse effects of an unfavorable FasL polymorphism on response, the CARD15 genotype status, or the location of the biopsy, due to variation of gene expression along the colon. To account for this variation, outliers may be identified, and subsequently determined whether the outliers may be accounted for by variations in medication use, CARD15 genotype, or the location of the colon biopsy.

#### Kits

**[0247]** In an additional aspect, the present invention provides kits embodying the methods, compositions, and systems for analysis of gene expression as described herein. Kits of the present invention may comprise one or more of the following: a) at least one pair of universal primers; b) at least one pair of target-specific primers, wherein the primers are specific to one or more sequences listed in Tables 4-8 or the sequence listing; c) at least one pair of reference gene-specific primers; and d) one or more amplification reaction enzymes, reagents, or buffers. The universal primers provided in the kit may include labeled primers. The target-specific primers may vary from kit to kit, depending upon the specified target gene(s) to be investigated, and may also be labeled. Exemplary reference gene-specific primers (e.g., target-specific primers for directing transcription of one or more reference

genes) include, but are not limited to, primers for  $\beta$ -actin, cyclophilin, GAPDH, and various rRNA molecules. The kits of the invention optionally include one or more preselected primer sets that are specific for the genes to be amplified. The preselected primer sets optionally comprise one or more labeled nucleic acid primers, contained in suitable receptacles or containers.

**[0248]** Exemplary labels include, but are not limited to, a fluorophore, a dye, a radiolabel, an enzyme tag, etc., that is linked to a nucleic acid primer itself. In addition, one or more materials and/or reagents required for preparing a biological sample for gene expression analysis are optionally included in the kit. Furthermore, optionally included in the kits are one or more enzymes suitable for amplifying nucleic acids, including various polymerases (RT, Taq, etc.), one or more deoxynucleotides, and buffers to provide the necessary reaction mixture for amplification. In one embodiment of the invention, the kits are employed for analyzing gene expression patterns using mRNA as the starting template. The mRNA template may be presented as either total cellular RNA or isolated mRNA. In other embodiments, the methods and kits described in the present invention allow quantification of other products of gene expression, including tRNA, rRNA, or other transcription products. In still further embodiments, other types of nucleic acids may serve as template in the assay, including genomic or extragenomic DNA, viral RNA or DNA, or nucleic acid polymers generated by non-replicative or artificial mechanism, including PNA or RNA/DNA copolymers. Optionally, the kits of the present invention further include software to expedite the generation, analysis and/or storage of data, and to facilitate access to databases. The software includes logical instructions, instructions sets, or suitable computer programs that can be used in the collection, storage and/or analysis of the data. Comparative and relational analysis of the data is possible using the software provided.

## EXAMPLES

### Studies in Obese Women

#### Study Design, Materials and Methods

**[0249]** We studied 25 healthy controls, and 35 successive obese ( $\text{BMI} > 35 \text{ kg/m}^2 + \text{co-morbidities}$ ) women without diabetes, and 60 obese women with diabetes. Fourteen obese women without diabetes were referred to our hospital for bariatric surgery. Before they were included, all women were evaluated by an endocrinologist, an abdominal surgeon, a psychologist and a dietician. Only after multidisciplinary deliberation the selected patients received a gastric bypass. This is a procedure in which the stomach is divided into a small, proximal pouch and a separate, large, distal remnant. The upper pouch is joined to the proximal jejunum through a narrow gastrojejunal anastomosis (Roux-en-Y configuration). The storage capacity of the stomach is reduced to approximately 5% of its normal value, and ingested food bypasses approximately 95% of the stomach, the entire duodenum, and a small portion (15-20 cm) of the proximal jejunum. Patients typically lose 35-40% of total body weight and most of this is maintained for at least 10-15 years<sup>57-59</sup>.

**[0250]** All participants did not have symptoms of clinical cardiovascular disease. However, their predicted cardiovascular risk based on their Framingham Risk Score (FRS) was elevated. This score was calculated according to Wilson<sup>60</sup> using the MedCalc (<http://www.medcalc.com/heartrisk.html>) program.

Blood samples were collected, and peripheral blood mononuclear cells (PBMC) were prepared from the anti-coagulated blood using gradient separation on Histo-paque-1077 performed directly in the blood collection tubes. Cells were washed three times in  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ -free Hanks's balanced salt solution. PBMC were incubated for 20 min at 4° C. with CD14 microbeads at 20  $\mu\text{l}/1 \times 10^7$  cells. The cells were washed once, re-suspended in 500  $\mu\text{l}$   $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ -free PBS containing 5% FBS/ $1 \times 10^8$  cells. The suspension was then applied to MidiMACS Separator (Miltenyi, Bergisch Gladbach, Germany)<sup>61, 62</sup>. Blood samples were also centrifuged to prepare plasma samples for analysis. Total and HDL-cholesterol and triglyceride levels were determined with enzymatic methods (Boehringer Mannheim, Vilvoorde, Belgium). LDL-cholesterol levels were calculated with the Friedewald formula. Free fatty acids (FFA) were measured with a FFA-Half Microtest kit (Roche Applied Science). Plasma glucose was measured with the glucose oxidase method (on Vitros 750XRC, Johnson & Johnson), and insulin with an immunoradiometric assay (Biosource Technologies, Inc., Fleurus, Belgium). Oxidized LDL (oxLDL) and interleukin-6 (IL6) were measured with ELISA. Blood pressure was taken three times with the participant in a seated position after 5 minutes quiet rest. The average of the last two measurements was used for systolic and diastolic blood pressure.

**[0251]** Total RNA was extracted from  $5.3 \times 10^6 \pm 1.7 \times 10^6$  (mean  $\pm$  s.d.) monocytes per person with the Trizol reagent (InVitrogen, Merelbeke, Belgium) and purified on an RNeasy Mini Kit column (Qiagen, Antwerpen, Belgium). Extraction yielded  $6.5 \pm 2.8 \mu\text{g}$  of RNA. After phenol-chloroform extraction and RNeasy Mini Kit cleanup of the RNA, its quality was assessed with the RNA 6000 Nano assay kit using the Agilent 2100 Bioanalyzer. The adjusted S28/S18 ratios of extracted RNA were  $1.60 \pm 0.07$ . First-strand cDNA was generated from total RNA by RT using random primers from Takara and Superscript III reverse transcriptase (InVitrogen). RNA expression levels were expressed as percentage of controls as previously described<sup>63</sup>.

**[0252]** RNA isolated from monocytes of obese women and non obese controls was analyzed on Illumina's Sentrix Human-6 v2 Expression BeadChip Kit containing around 46,713 probes/array targeting genes and known alternative splice variants from the RefSeq database release 17 and UniGene build 188. RNA was labeled, hybridized and scan according to Illumina GLP standards by Aros AB laboratory (Aarhus, Denmark). The raw data were normalized with the rank invariant method (Illumina BeadStudio V2). This method uses a linear scaling of the populations being compared. The scaling factor is determined by rank-invariant genes. 'Rank-invariant' genes are those genes whose expression values show a consistent order relative to other genes in the population. Of the 46,713 transcripts, 1,725 transcripts which were mapped in the Ingenuity Pathway Analysis (IPA) program 5.5-802 were differentially expressed in monocytes of obese women compared to controls at the  $P\text{-value} \leq 0.05$  level. Out of these genes, we selected a subpopulation of the 592 genes with a ratio threshold of +1.2. The Ingenuity V5 Pathway analysis program (IPA 5.0) was used to model the canonical and metabolic function pathways that may be activated in the monocytes of obese women compared to controls. Networks were built by means of the "Connect" and the "Path explorer" tool in IPA 5.0. Canonical pathways were derived using the "Analysis" tool in IPA 5.0.

**[0253]** Quantitative real-time PCR was performed on a 7500 fast Real-Time PCR system using the power SYBR Green Master mix (Applied Biosystems, Lennik, Belgium). Table 1 summarizes forward and reverse primers used in RT-PCR analysis. The RNA expression level was calculated with the threshold cycle ( $C_t$ ) value, i.e., the number of PCR cycles at which the fluorescent signal reaches a fixed threshold. For each sample, both the  $C_t$  for the gene of interest and for the housekeeping gene  $\beta$ -actin were determined to calculate  $\Delta C_{t, sample}(C_{t, target gene} - C_{t, housekeeping gene})$ , thus normalizing the data and correcting for differences in amount and/or quality among the different RNA samples. The expression levels were related to an external calibrator. Subsequently,  $\Delta\Delta C_t(\Delta C_{t, sample} - \Delta C_{t, calibrator})$  was determined, and the relative expression levels were calculated from  $2^{-\Delta\Delta C_t}$ , according to the manufacturer's instructions (Applied Biosystems). RNA expression levels were expressed as percentage of controls as previously described<sup>64</sup>.

**[0254]** MM6 and THP-1 cells were cultured in RPMI-1640 medium containing 10% FBS (low LPS), 0.2 mM Streptomycin, 100 U/ml Penicillin, 2 mM L-glutamine, 1xMEM NEAA, 1 mM MEM sodium Pyruvate Solution. At day 1, cells were seeded in a 24 well plate at a cell density of  $1 \times 10^6$  cells/ml. At day 2, we replaced growth medium with complete growth medium (-FBS)+HSA (1 mg/ml). At day 3 we added ox-LDL (50  $\mu$ g/ml) or FFA (710.9  $\mu$ M final concentration) or placebo. At day 4, cells were collected, washed and RNA was extracted as described above. IRAK3 RNA expression was assessed as described above.

**[0255]** Groups were compared with unpaired t-test with Welch's correction (two groups) or with paired t-test (GraphPad 5.0). Pearson correlation coefficients were calculated to determine the relations between gene expressions and of gene expressions with risk factors. Receiver operating characteristic (ROC) curve analysis was performed using the MedCalc program. Fisher exact test analysis was performed to determine relative risks, sensitivity and specificity, and positive and negative predictive values using cut-offs which were determined by ROC curve analysis. A P-value of less than 0.05 was considered statistically significant. Principal component analysis (PCA) was performed using the TMEV 4-2 program<sup>65</sup>. Data are normalized in TMEV4 and the 119 patients are clustered according to the values of the 6 to 10 variables selected with the Hierarchical Clustering (HCL) with the average linkage method on Euclidean distances<sup>66, 67</sup>. Tree leaves ordering were optimized using the "Gene Leaf Order" option. The hierarchical trees obtained are shown in FIG. 10. A principal Component Analysis (PCA) is then used to attribute the overall variability in the data to a reduced set of variables termed principal components<sup>68</sup>. The centering mode is the median and the number for neighbors for KNN imputation is 10. (Available at [http://smi-web.stanford.edu/pubs/SMI\\_Abstracts/SMI-1999-0804.html](http://smi-web.stanford.edu/pubs/SMI_Abstracts/SMI-1999-0804.html)).

#### Characteristics of Study Cohort and Effect of Weight Loss

**[0256]** We studied 25 healthy lean control women (age: 39.17 years), and 95 obese women (age:  $51 \pm 11$ ,  $P < 0.001$ ). FIG. 1 shows that obese women had a higher BMI associated with higher FRS. Leptin concentrations were higher; adiponectin was lower. Glucose, insulin, and triglyceride concentrations were higher; HDL-cholesterol was lower. LDL-cholesterol levels were not different. Plasma concentrations of oxLDL ( $50 \pm 20$  vs.  $32 \pm 10$  U/I;  $P < 0.001$ ) and interleukin-6 ( $5.4 \pm 2.9$  vs.  $2.7 \pm 0.94$  pg/ml) were also higher.

**[0257]** Among obese women 35 were without diabetes (age:  $45 \pm 12$ ,  $P < 0.05$ ; BMI:  $41.0.3 \pm 5.9$ ,  $P < 0.0001$  vs. controls); 60 women had diabetes (age:  $55 \pm 9$ ,  $P < 0.001$  vs. controls and vs. obese women without diabetes; BMI:  $37.2 \pm 4.8$ ;  $P < 0.0001$  vs. controls, and  $P < 0.05$  vs. obese without diabetes). FIG. 2 shows that diabetics had lower BMI but higher FRS. Their leptin concentrations were higher; adiponectin was not different. Diabetics had higher glucose, insulin, and triglyceride concentrations; their HDL-cholesterol was lower. Their LDL-cholesterol was lower due to more frequent use of statins (73% vs. 24% of women without diabetes).

**[0258]** Controls had lower systolic and diastolic blood pressure ( $121 \pm 11$ ,  $71 \pm 9.3$  mm<sub>Hg</sub>) than obese women without ( $135 \pm 18$ ,  $88 \pm 12$  mm<sub>Hg</sub>;  $P < 0.01$  and  $< 0.001$  vs. controls) and obese women with diabetes ( $139 \pm 16$ ,  $82 \pm 10$  mm<sub>Hg</sub>;  $P < 0.001$  vs. controls). Diabetics were more often treated for arterial hypertension (57%) than obese women without diabetes (42%). OxLDL concentrations in diabetics were lower ( $46 \pm 17$  vs.  $57 \pm 23$  U/I;  $P < 0.05$ ) likely due to more frequent statin use and associated lower LDL-cholesterol. Interleukin-6 concentrations were not different.

**[0259]** We collected samples from 14 out of 35 obese women without diabetes after bariatric surgery. Their BMI was 17% lower 3 months after bariatric surgery (FIG. 3). Their FRS predicted a decrease of the cardiovascular risk to  $4 \pm 3\%$  (compared to  $8 \pm 7\%$  before weight loss;  $P < 0.01$ ).

**[0260]** Weight loss was associated with a significant decrease of leptin, glucose, insulin, and triglyceride concentrations, and of systolic and diastolic blood pressure ( $118 \pm 4$ ,  $61 \pm 3$  mm<sub>Hg</sub>;  $P < 0.001$  vs. before). Adiponectin levels were higher. Weight loss did not affect LDL- and HDL-cholesterol, oxLDL and interleukin-6 concentrations.

#### Microanalysis of RNA Extracts of Blood Monocytes of Obese Women

**[0261]** We performed microarray analysis of RNA extracts from monocytes of 14 obese women. Table 2 summarizes the most significant gene functions of the 592 genes which were overexpressed compared to age-matched controls ( $N=10$ ). Especially genes which mediate cell-to-cell signaling and immune response were overexpressed. They are known to be involved in hematological, immunological, neurological diseases, and in cancer. FIG. 4 summarizes the top ten canonical signaling pathways which were identified by means of Ingenuity Pathways Analysis (IPA 5.0) program. The acute phase response, the interleukin, the NF $\kappa$ B, the LXRL/RXR, and the PPAR signaling pathways were among them.

#### Proposal of a Structural Network

**[0262]** In order to identify the most representative structural network the following software tools were used together with the IPA program:

Software	Description	Input Data	Output Data
GEMS Launcher task:	Search a weight matrix in DNA	DNA sequences, DNA databases.	A list of weight matrix matches.
"MatInspector: Search for transcription factor binding sites." (Quandt et al., 1995, Cartharius et al., 2005)			

-continued

Software	Description	Input Data	Output Data
GEMS Launcher task: "FrameWorker: Definition of common framework"	Search for a common framework of two or more DNA patterns in a set of sequences.	A set of two or more co-regulated promoter sequences and a library of matrices describing transcription factor binding sites.	A framework of at least two conserved DNA elements in at least a subset of the sequences.
Genomatix Portal: "Gene2Promoter"	Retrieve & Analyze Promoters	DNA sequences, accession numbers, GeneIDs, or gene symbols	A set of promoters with the ability to extract sequences and start basic analysis tasks
Genomatix Portal: "ElDorado"	Extended Genome Annotation	DNA sequences, accession numbers, or GeneIDs, or gene symbols	Retrieve annotation and genomic context of an input gene

**[0263]** FIG. 5 shows a predicted structural model in which two pathways interact:

**[0264]** Activation of the TLR2 receptor pathway by free fatty acids (FFA) and oxLDL, which are elevated in blood of obese persons, occurs through activation of MYD88 and TRAF6 resulting in the activation of NFκB and associated expression of the inflammatory TNF and IL-1 which contribute to oxidative stress and release of ROS<sup>69</sup>. TRAF6 induces the expression and the activation of NFκB that then further induces the expression of TLR2. In support of this model, several NFκB-binding sites have been identified of the promoter of TLR2 (FIG. 6). This feedback may be used to couple the TLR2-pathway of innate immune response to other NFκB activating response mechanisms thereby generally amplifying the immune response. ROS activate directly NFκB. This pathway can be inhibited at two sites. IRAK3 is inhibiting TRAF6<sup>70</sup>; TNFAIP3<sup>71</sup> and TNFAIP6<sup>72</sup> are inhibiting NFκB.

**[0265]** Insulin through interaction with its receptor (INSR) and IRS2 activates the PI3K/AKT pathway. This activation is required for proper insulin action. ROS induces the expression of MAPK13 that blocks proper insulin action<sup>73</sup>. ROS induces the expression of the forkhead transcription factor FOXO3A that induces the expression of SOD2. FOXO3A, like NFκB, is regulated at the level of nuclear transport. In the nucleus FOXO3A blocks NFκB<sup>74</sup>. Promoter analysis suggested that FOXO3A is also involved in the regulation of TRAF6 and IRAK3. Indeed, a forkhead binding site was identified within a strongly conserved TFBS model shared by two promoters for these genes. Activation of the PI3K/AKT pathway results in the phosphorylation of FOXO3A, thereby blocking its transport into the nucleus. This can result in activation of NFκB, and the repression of the expression of SOD2 (directly), and of IRAK3, TRAF6, and TNFAIP3 (indirectly). This inactivation of FOXO3A can be reverted by phosphorylation

of FOXO3A at another site by MAPK13<sup>75</sup> thereby restoring SOD2 expression. As discussed above, SOD2 is essential for neutralizing ROS. We postulated that MAPK13 is an important link between the two pathways, other than that it affects FOXO3A activity, because it contains common transcription factor binding sites (TFBS) with a gene upstream of NFκB, i.e. TNFAIP3, and two genes downstream of NFκB, i.e. MAPK13 and SOD2 (FIG. 7).

**[0266]** Although the indicated references suggested separate interactions proposed in the model, the complete model has not been presented in its present form and its relationship with obesity and related disorders has never been validated.

**[0267]** We also used the IPA program for searching other correlating proteins which may be important regulators of the expression of the identified cluster. We identified ZNF217 as a possible candidate.

Validation of the Structural Model:

**[0268]** We measured the RNA expression of TLR2, MYD88, TRAF6, IRAK3, TNFAIP3, TNFAIP6, and IRS2, MAPK13, FOXO3A, SOD2, and ZNF217 in extracts of monocytes from healthy controls, and obese women without and with diabetes. FIG. 8 shows that the expressions of TLR2, TNFAIP3, TNFAIP6, MAPK13, FOXO3A, and SOD2 were higher in monocytes of obese women; the expressions of IRAK3, IRS2 and ZNF217 were lower. Those of MYD88 and TRAF6 were not different. Lack of overexpression of MYD88 and TRAF6 does not exclude their activation. In support of the latter, we observed an 18±4.2% (P<0.01) increase in NFκB in monocytes from obese women compared to controls. This increase was observed in spite of the increase in TNFAIP3, TNFAIP6, and of FOXO3A, which are inhibitors of NFκB. The increase in NFκB was observed despite the unexpected increase of the NFκB inhibitor (39.4±6.0% higher in monocytes from obese women; P<0.001). Although we did not find an upregulation of MYD88, we observed an indirect evidence of the action of MYD88. Recent insights<sup>76</sup> have revealed additional functions for MYD88 apart from NFκB activation, including activation of the transcription factors IRF1, IRF5 and IRF7, and also a role outside the TLRs in interferon-γ signaling. The expression of IRF1 was 94±6.3% higher in monocytes from obese women.

**[0269]** According to the proposed model, the increase in MAPK13 can explain why the increase in FOXO3A, despite the increase in NFκB, was associated with an increase in SOD2.

**[0270]** According to the proposed model IRAK3 was the only inhibitor of the TLR2 pathway of which the expression was significantly lowered suggesting a functional role of IRAK3 in the regulation of this pathway. The expressions of the other predicted inhibitors TNFAIP3 and TNFAIP6 were elevated. In spite of their increases we observed an activation of the TLR2 pathway. These data suggest that TNFAIP3 and TNFAIP6 are markers of stress which cannot prevent stress. Despite the increase of the antioxidant SOD2 we observed an increase in oxLDL suggesting that SOD2 is also a marker of stress that by itself can not prevent oxidative stress.

**[0271]** The expression of ZNF217 was lower in obese women than in lean controls: 0.73±0.092 vs. 0.93±0.03 (P<0.001).

Determination of Diagnostic Values

**[0272]** To determine the value of the above mentioned genes to discriminate between lean control and obese women,

we performed ROC curve analysis. Table 3 indicates that SOD2, TNFAIP6, TNFAIP3, IRAK3, TLR2, ZNF217, IRS2, and MAPK13 (ranked according to their areas under the curve) significantly discriminated between control and obese women; MYD88, FOXO3A, and TRAF6 did not.

**[0273]** In aggregate, we identified a cluster of genes containing IRAK3, SOD2, TNFAIP6, TNFAIP3, TLR2, ZNF217, IRS2, and MAPK13 with diagnostic value. Importantly, the expressions of these genes were not different between obese women with and with no diabetes, indicating that they are differentially regulated early in the onset of the above mentioned disorders.

**[0274]** As an alternative, we determined the relative risks, sensitivities and specificities, and positive and negative predictive values by Fisher exact test analysis. Low IRAK3 and ZNF217, and high SOD2, TNFAIP6, TNFAIP3 discriminated between lean controls and obese women (Table 4). The odds ratios of TLR2, MAPK13 and IRS2 were lower. We then investigated whether IRAK3, SOD2, TNFAIP6, TNFAIP3, and ZNF217 had an additive diagnostic value. Table 4 shows that the sensitivity of the combination of low IRAK3 and high SOD2 was 76%; specificity was 86%. Similar values were obtained when combining IRAK3 with TNFAIP3, TNFAIP6 or ZNF217. This was expected because of the high correlation between SOD2, TNFAIP3, and TNFAIP6 (correlation coefficients 0.88 and 0.83, respectively;  $P < 0.001$ ). Although ZNF217 correlated with TNFAIP3, TNFAIP6 and SOD2, the correlation coefficients were lower (Table 5). In contrast IRAK3 did not correlate with any of the 3 genes. Adding TNFAIP6 and TNFAIP3 to the IRAK3-SOD2 pair increased sensitivity to 95% or above and increased specificity 76%. Similar values were obtained when replacing SOD2 with ZNF217. Further additions of genes did not significantly increase sensitivities and specificities. Table 6 shows the correlation between gene expressions and risk factors.

**[0275]** We also investigated if statin treatment had an effect on the discrimination between controls and obese persons. Table 7 shows the outcome of the ROC curve analysis for obese persons who were not treated with a cholesterol lowering statin compared to lean controls; table 8 shows the outcome for obese persons who were treated with a statin. SOD2, TNFAIP3, TNFAIP6, IRAK3, TLR2, and IRS2 discriminated between lean controls and obese persons whether they were treated or not with a statin; MAPK13 did not when treated with statin.

**[0276]** Table 9 shows the outcome of the Fisher exact test analysis comparing lean controls with obese persons who were not on statin therapy. Table 10 shows the analysis for obese persons on statin therapy. Overall, there were no significant differences.

**[0277]** To further assess the relation between gene expression and obesity, we measured the expressions of those genes in monocytes from obese women before and after weight loss. We demonstrated above that weight loss was associated with a significant decrease of their predicted risk as indicated by their lower FRS. FIG. 9 shows that the expressions of SOD2, TNFAIP3, and TNFAIP6 decreased after weight loss; the expressions of IRAK3 increased. The expression of ZNF217 increased from  $0.68 \pm 0.04$  before to  $0.98 \pm 0.09$  after weight loss.

**[0278]** Table 11 shows the association between gene signatures in circulating monocytes and the occurrence of the metabolic syndrome. It shows that prediction of the metabolic syndrome improves by combining genes from the cluster

particularly IRAK3, SOD2, TNFAIP3, TNFAIP6, and ZNF217. The relative risks of having the metabolic syndrome increase from 1.2 (when the RNA expression level of 1 gene is different from control; cut-point determined by ROC analysis) to 6.8 (when the RNA expression level of at least three genes are different from control). The clinical sensitivities (the percentages of positive tests when the clinical disorder is present) increase from 52% (when the RNA expression level of 1 gene is positive) to 97% (when the RNA expression level of at least three genes are positive). In the current cohort, ROC analysis did not show a significant relationship between ZNF217 expression and cardiovascular risk. Therefore, ZNF217 was as yet not included in further analyses.

**[0279]** Table 12 shows the association between gene signatures in circulating monocytes and the present cardiovascular risk equivalents. It shows that the risk prediction improves by combining genes from the cluster particularly IRAK3, SOD2, TNFAIP3, and TNFAIP6. The relative risks increase from 1.4 (when the RNA expression level of 1 gene is different from control; cut-point determined by ROC analysis) to 7.2 (when the RNA expression level of at least three genes are different from control); the sensitivities increase from 52% to 95%. The positive predictive values can be as high as 80%. The negative predictive values can be as high as 92%.

**[0280]** In addition to ROC and Fisher exact analysis, we performed PCA analysis to reduce the number of dimensions to 3 per 119 (number of patients) to map the distance between lean controls, obese persons without cardiovascular risk equivalents, and obese persons with cardiovascular risk equivalents. These are: an at least 10% risk of developing a cardiovascular disease within the next 10 years based on Framingham scoring and/or diabetes and/or metabolic syndrome (at least 3 metabolic syndrome or MetSyn components as defined above). This type of analysis allows determining the additive value of emerging biomarkers on top of established risk factors: high BMI (obesity), high glucose (diabetes), high blood pressure (hypertension), and low HDL cholesterol and/or high triglycerides (dyslipidemia), and low adiponectin. Addition of RNA expression data for genes from the cluster particularly IRAK3, SOD2, TNFAIP3, and TNFAIP6 to the metabolic syndrome components together with adiponectin improved clearly the hierarchical clustering and the PCA representation of the 3 groups. Not only the lean controls (green) were separated better from the obese persons at low risk (yellow), but also the obese persons at low risk were separated better from the obese persons at high risk (red).

Assessment of FFA and oxLDL on the Expression of IRAK3 in a Human Monocyte Cell Line

**[0281]** As an example of a relevant human monocytic cell line we have chosen the MM6 cell. If we put the expression of IRAK3 at baseline at 100%, then the expression in the presence of FFA at a concentration equal to the mean FFA concentration in the plasma of obese persons before weight loss was 80%. The expression in the presence of in vitro oxidized LDL (50  $\mu\text{g/ml}$ ) was 49%. As another example of a human monocytic cell line we have used THP-1 cells. Incubation of THP-1 cells with oxidized LDL reduced IRAK3 expression with 58%.

Examples of Tests for Diagnosis and Prognosis of Metabolic Syndrome Disorders and Cardiovascular Diseases and for Assessing the Effect of Interventions

**[0282]** A first approach for determining mRNA expression of specific cells in suspension combines flow cytometry with

in situ hybridization<sup>77</sup>. This method maintains cellular integrity since it does not require the extraction and amplification of mRNA. Furthermore, it allows for the detection of the cellular source of the message in a heterogeneous population of cells thus avoiding monocyte isolation from whole blood. This way, the method provides the advantage to avoid 2 laborious and time-consuming steps. Indeed, the combination of flow cytometry with in situ hybridization makes a rapid analysis of large numbers of cells at a rate of approx 2000 cells/s feasible. Further, multicolor flow cytometry allows the resolution of very complex analytical mixtures. A schematic overview of the method is shown in FIG. 12. A small volume of blood is drawn from the patient. In a first step, monocytes in whole blood are stained with an antibody specific for monocyte surface markers e.g. CD14. This antibody is labeled with a marker for detection. In a second step, in situ hybridization is performed. Riboprobes that carry a marker for detection are added and hybridize with the corresponding mRNA of interest in a concentration-dependent way, which allows for quantitative detection of gene expression. The use of different color markers for each riboprobe set that hybridizes to one gene, allows for simultaneous detection of different messages.

**[0283]** Finally, cells are guided through a flow cytometer and lasers excite the markers associated with the riboprobes and the monocyte specific antibody. The marker specific fluorescent signals are measured and allow simultaneous message detection in the same cell.

**[0284]** Two other approaches make use of biosensors. The first approach requires the extraction of mRNA and depending on the sensitivity of the bioreceptor, amplification of messages. A possible work flow for this approach is depicted in FIG. 13. To detect mRNA messages specifically in mono-

cytes, we first need to select the monocytes out of whole blood. This can easily be done by adding magnetic bead coupled antibody specific for monocytes, e.g. CD14. Through magnetic forces, the monocytes can be moved out whole blood suspension. Next, the monocytes have to be lysed and RNA is extracted. Depending on the sensitivity of the bioreceptor, RNA can be amplified isothermally using NASBA<sup>78</sup>.

**[0285]** The bioreceptor for the detection of nucleic acids is a complementary nucleic acid<sup>79</sup>. Hybridization of the target is then detected through the generation of signal such as a current, fluorescence, . . . There are several possibilities. In this figure/In our proposal, the bioreceptor is a molecular beacon immobilized to the recognition layer of the biosensor<sup>80</sup>.

**[0286]** A molecular beacon is a single-stranded oligonucleotide hybridization probe that forms a stem-and-loop structure. A fluorophore is covalently linked to one end and a quencher is covalently linked to the other end. Molecular beacons do not fluoresce when they are unbound because the stem places the fluorophore close to the quencher that inhibits the fluorophore to fluoresce. When the probe encounters a target molecule, it forms a probe-target hybrid. Consequently, the molecular beacon undergoes a spontaneous conformational reorganization that forces the stem hybrid to dissociate and the fluorophore and the quencher to move away from each other, enabling them to fluoresce brightly.

**[0287]** Taking advantage of the properties of molecular beacons, we propose an alternative competitive assay that circumvents the need for nucleic acid extraction in FIG. 14. After monocyte selection as described above, unlabeled riboprobes are added to the monocyte suspension and can bind to the bioreceptor or to the intracellular target mRNA. The higher the expression of our target mRNA, the lower the signal detected by the biosensor.

TABLE 1

Primers for RT-PCR analysis

Gene Symbol	FORWARD	REVERSE
BCL2L11	GGGCCACCGTGTCATTA	CCTGTGCCAATTCATGA
FOXO3A	CAACAAAATGAAATCCATAGAGCAAGTGTATGAGTGAGAGGCAATAGCA	
IRAK3	TGCAACGCGGGCAA	TTTAGTGATGTGGAGGATCTTCA
IRS2	GCTTCCCAGTGCCTATCTTC	AAACCAACAACCTACATCTCCAATGA
MAPK13	AAAGCGGCCAAATCCTACATC	GGGAACAGCTGAGTGAAATCCT
MST1	TCTCAGCAACACGGGACTGT	AATCCCCTCTACCCCTGTGTGTA
MYD88	TGCATATCTTTGCTCCACTTTCA	ATTCCTCCCAAGATCCTAAGAA
SOD2	TGGAAGCCATCAAACGTGACT	TTTGTAAAGTGICCCCGTTCTTT
TLR2	TGCAAGTACGAGCTGGACTTCTC	GTGTTTATTATCTTCCGAGCTT
TNFAIP3	TCCTTGCTCCTTCCCTATCTC	ATGTTTCGTCTTCTCCTTATGAA
TNFAIP6	GGCCATCTCGCAACTTACAAG	GCAGACAGACATGAAATCCA
TRAF6	CATGAAAAGATGCAGAGGAATCAC	GAACAGCCTGGCCAACAT
ZNF217	CGCTGCGACGGATACAC	TGGTCAGAGGCATCACATCAC
BACTIN	GGACCTGACCGACTACCTCATG	CGACGTAGCAGAGCTTCTCCTT

TABLE 2

Top ten of gene functions which are overexpressed in monocytes of obese women				
Functions of 592 overexpressed genes	Rank	Number of overexpressed genes	Total number of genes with this function present on microchip	Percentage of overexpressed genes
Gene, cellular development, immune and lymphatic system development	1	29	35	83%
Cell-to-cell signalling and interaction, haematological system development and function, Immune	2	29	35	83%
Cancer, immunological disease, tumour	3	28	35	80%
Cell signalling, lipid, small molecule biochemistry	4	27	35	77%
Cancer, dermatological diseases and conditions, cellular development	5	23	35	66%
Cancer, cell-to-cell signalling and interaction, neurological disease	6	23	35	66%
Cancer, cell cycle, immunological disease	7	21	35	60%
Cell death, immunological disease, cellular growth	8	21	35	60%
Cellular assembly and cellular compromise, cellular function	9	20	35	57%
Cell-to-cell signalling and interaction, haematological system development and function, immune and lymphatic system development	10	19	35	54%

TABLE 3

Receiver Operating Characteristic Curve analysis of relation with obesity					
Gene	AUC	Sensitivity %	Specificity %	PPV %	NPV %
SOD2	0.81 (0.73-0.88)***	72	84	94	44
TNFAIP6	0.81 (0.73-0.88)***	70	88	96	44
TNFAIP3	0.76 (0.67-0.83)***	51	96	98	34
IRAK3	0.72 (0.63-0.80)***	39	92	96	33
TLR2	0.72 (0.63-0.80)***	76	64	88	41
ZNF217	0.74 (0.65-0.82)***	54	96	98	36
IRS2	0.70 (0.61-0.78)**	78	56	87	40
MAPK13	0.64 (0.55-0.73)**	37	96	97	29
TRAF6	0.60 (0.51-0.69)	44	92	92	29

TABLE 3-continued

Receiver Operating Characteristic Curve analysis of relation with obesity					
Gene	AUC	Sensitivity %	Specificity %	PPV %	NPV %
FOXO3A	0.60 (0.50-0.70)	55	76	87	40
MYD88	0.52 (0.42-0.62)	24	88	86	27

Data present area under the curve and 95% confidence intervals;  
 \*P < 0.05,  
 \*\*\*P < 0.01,  
 \*\*\*\*P < 0.001.  
 AUC: area under the curve; PPV: positive predictive value; NPV: negative predictive value

TABLE 4

Fisher exact test analysis of relation with obesity						
Gene	Cut point	OR	Sensitivity %	Specificity %	NPV %	PPV %
IRAK3	≤0.85	11	92 (74-99)	49 (39-60)	32 (22-45)	96 (86-100)
SOD2	≥1.67	10	80 (59-93)	72 (61-80)	43 (28-58)	93 (85-98)
TNFAIP6	≥1.72	9.6	80 (59-93)	70 (60-79)	42 (28-57)	93 (85-98)
TNFAIP3	≥2.14	7.8	88 (69-98)	51 (41-62)	32 (22-45)	94 (84-99)
ZNF217	≤0.70	28	54 (44-65)	96 (80-99)	36 (24-48)	98 (90-99)
IRAK3 + SOD2	Same as above	20	76 (55-91)	86 (78-93)	59 (41-76)	93 (86-97)
IRAK3 + TNFAIP3	Same as above	17	80 (59-93)	82 (72-88)	53 (36-69)	94 (86-98)
IRAK3 + TNFAIP6	Same as above	22	76 (55-91)	87 (79-93)	61 (42-78)	93 (86-97)
IRAK3 + ZNF217	Same as above	23	76 (66-84)	88 (69-97)	49 (34-64)	96 (89-99)
SOD2 + ZNF217	Same as above	19	79 (69-86)	84 (64-95)	51 (35-67)	95 (87-99)
TNFAIP3 + ZNF217	Same as above	22	81 (71-88)	84 (64-95)	54 (37-70)	95 (88-99)
TNFAIP6 + ZNF217	Same as above	22	87 (79-93)	76 (55-91)	61 (42-78)	93 (86-97)
IRAK3 + SOD2 + TNFAIP3	Same as above	56	95 (88-98)	76 (55-91)	79 (58-93)	94 (87-98)
IRAK3 + SOD2 + TNFAIP6	Same as above	96	97 (91-99)	76 (55-91)	86 (65-97)	94 (87-98)
IRAK3 + TNFAIP3 + TNFAIP6	Same as above	24	88 (80-96)	76 (55-91)	63 (44-80)	93 (86-97)

TABLE 4-continued

Fisher exact test analysis of relation with obesity						
Gene	Cut point	OR	Sensitivity %	Specificity %	NPV %	PPV %
IRAK3 + SOD2 + ZNF217	Same as above	18	85 (77-92)	76 (55-91)	58 (39-75)	93 (86-97)
IRAK3 + TNFAIP3 + ZNF217	Same as above	21	84 (75-91)	80 (59-93)	57 (39-74)	94 (87-98)
IRAK3 + TNFAIP6 + ZNF217	Same as above	96	97 (91-99)	76 (55-91)	86 (65-97)	94 (87-98)
SOD2 + TNFAIP3 + ZNF217	Same as above	50	93 (85-97)	80 (59-93)	74 (54-89)	95 (88-98)
SOD2 + TNFAIP6 + ZNF217	Same as above	56	95 (88-98)	76 (55-91)	79 (58-93)	94 (87-98)
TNFAIP3 + TNFAIP6 + ZNF217	Same as above	30	90 (83-96)	76 (55-91)	68 (86-98)	93 (86-98)
IRAK3 + SOD2 + TNFAIP3 + TNFAIP6	Same as above	47	76 (55-91)	94 (87-98)	76 (55-91)	94 (87-98)
IRAK3 + SOD2 + TNFAIP3 + ZNF217	Same as above	294	99 (94-99)	76 (55-91)	95 (75-99)	94 (87-98)
IRAK3 + SOD2 + TNFAIP6 + ZNF217	Same as above	466	100 (96-100)	72 (51-88)	100 (81-100)	93 (86-97)
IRAK3 + TNFAIP3 + TNFAIP6 + ZNF217	Same as above	118	98 (93-99)	72 (51-88)	90 (68-99)	93 (86-97)
IRAK3 + SOD2 + TNFAIP3 + TNFAIP6 + ZNF217	Same as above	466	100 (96-100)	72 (51-88)	100 (81-100)	93 (86-97)

Cut-off values were determined by ROC curve analysis (see table 3). Sensitivities, specificities, and positive and negative predictive values were presented as means and 95% confidence intervals. OR: chance of being obese when having high or low (for IRAK3 and ZNF217) expressions compared with those having low or high expressions.

TABLE 5

Correlation between expressions of genes within the selected cluster				
Genes	IRAK3	TNFAIP3	TNFAIP6	SOD2
IRAK3	—	NS	NS	NS
TNFAIP3	NS	—	0.76****	0.83****
TNFAIP6	NS	0.76****	—	0.88****
SOD2	NS	0.83****	0.88****	—
ZNF217	NS	-0.39***	-0.47***	-0.48***

\*\*\*P < 0.001;  
\*\*\*\*P < 0.0001

TABLE 6

Correlations between gene expressions of genes within the selected cluster and risk factors									
Genes	BMI	Leptin	Adiponectin	Glucose	Insulin	HDL-C	TG	OxLDL	IL6
IRAK3	-0.40***	-0.34***	0.19*	NS	-0.21*	NS	-0.23*	-0.33***	NS
SOD2	0.23*	0.25**	NS	NS	NS	0.22*	0.19	0.26**	0.19*
TNFAIP6	0.30***	0.29**	NS	NS	NS	NS	NS	0.27	NS
TNFAIP3	NS	NS	NS	NS	NS	NS	NS	NS	NS
ZNF217	NS	-0.20*	NS	NS	-0.21*	NS	NS	NS	-0.22*

\*P < 0.05;  
\*\*P < 0.01;  
\*\*\*P < 0.001;  
NS: not significant

TABLE 7

Receiver Operating Characteristic Curve analysis of obese persons not treated with a statin					
Gene	AUC	Sensitivity %	Specificity %	PPV %	NPV %
SOD2	0.82 (0.71-0.90)***	63	96	97	58
TNFAIP6	0.85 (0.75-0.92)***	76	88	92	66
TNFAIP3	0.77 (0.65-0.86)***	50	96	96	51

TABLE 7-continued

Receiver Operating Characteristic Curve analysis of obese persons not treated with a statin					
Gene	AUC	Sensitivity %	Specificity %	PPV %	NPV %
IRAK3	0.73 (0.62-0.83)***	46	100	100	55
TLR2	0.78 (0.66-0.87)***	76	76	85	63
IRS2	0.67 (0.54-0.77)*	39	92	90	45

TABLE 7-continued

Receiver Operating Characteristic Curve analysis of obese persons not treated with a statin					
Gene	AUC	Sensitivity %	Specificity %	PPV %	NPV %
MAPK13	0.65 (0.57-0.80)**	46	96	96	49
ZNF217	0.75 (0.65-0.83)***	55	96	98	36

Data present area under the curve and 95% confidence intervals;

\*P < 0.05,

\*\*P < 0.01,

\*\*\*P < 0.001.

AUC: area under the curve; PPV: positive predictive value; NPV: negative predictive value

TABLE 8

Receiver Operating Characteristic Curve analysis of obese persons treated with a statin					
Gene	AUC	Sensitivity %	Specificity %	PPV %	NPV %
SOD2	0.80 (0.69-0.88)***	69	84	89	58
TNFAIP6	0.78 (0.67-0.87)***	63	88	91	55
TNFAIP3	0.75 (0.63-0.84)***	51	96	96	58
IRAK3	0.70 (0.59-0.81)**	47	92	92	47
TLR2	0.68 (0.56-0.78)**	80	56	78	59
IRS2	0.73 (0.61-0.82)***	84	56	79	64
MAPK13	0.60 (0.48-0.71)	33	92	89	41
ZNF217	0.72 (0.63-0.81)***	53	94	96	36

Data present area under the curve and 95% confidence intervals;

\*P < 0.05,

\*\*P < 0.01,

\*\*\*P < 0.001.

AUC: area under the curve; PPV: positive predictive value; NPV: negative predictive value

TABLE 9

Fisher exact test analysis of obese persons not treated with a statin						
Gene	Cut point	OR	Sensitivity %	Specificity %	PPV %	NPV %
IRAK3	≤0.77	39.5	100 (86-100)	43 (29-59)	49 (35-63)	100 (83-100)
SOD2	≥1.82	19.6	92 (74-99)	63 (48-77)	58 (41-73)	94 (79-99)
TNFAIP6	≥1.72	12.7	80 (59-93)	76 (61-87)	65 (45-81)	88 (73-96)
TNFAIP3	≥2.14	6.7	88 (69-97)	48 (33-63)	48 (33-63)	88 (69-98)
IRAK3 + SOD2	Same as above	41.4	92 (74-99)	78 (64-89)	70 (51-84)	95 (82-99)
IRAK3 + TNFAIP3	Same as above	41.4	92 (74-99)	78 (64-89)	70 (51-84)	95 (82-99)
IRAK3 + TNFAIP6	Same as above	26.7	80 (59-93)	87 (74-95)	77 (56-91)	89 (76-96)
IRAK3 + SOD2 + TNFAIP6 + TNFAIP3	Same as above	26.7	80 (59-93)	87 (74-95)	77 (56-91)	89 (76-96)

Cut-off values were determined by ROC curve analysis (see table 3). Sensitivities, specificities, and positive and negative predictive values were presented as means and 95% confidence intervals. OR: chance of being obese when having high or low (for IRAK3) expressions.

TABLE 10

Fisher exact test analysis of obese persons treated with a statin						
Gene	Cut point	OR	Sensitivity %	Specificity %	PPV %	NPV %
IRAK3	≤0.85	10.0	92 (74-99)	47 (33-62)	47 (33-62)	92 (74-99)
SOD2	≥1.67	9.1	80 (59-93)	69 (55-82)	57 (39-74)	87 (73-96)
TNFAIP6	≥1.72	7.5	80 (59-93)	65 (50-78)	54 (37-71)	86 (71-95)
TNFAIP3	≥2.14	7.6	88 (69-97)	51 (36-66)	48 (33-63)	89 (72-98)
IRAK3 + SOD2	Same as above	16	76 (55-91)	84 (70-93)	70 (50-86)	87 (74-95)
IRAK3 + TNFAIP3	Same as above	15.6	80 (59-93)	80 (66-90)	67 (47-83)	89 (75-96)
IRAK3 + TNFAIP6	Same as above	16	76 (55-91)	84 (70-93)	70 (58-86)	87 (74-95)
IRAK3 + SOD2 + TNFAIP6 + TNFAIP3	Same as above	18	72 (51-88)	88 (75-95)	75 (53-90)	86 (73-94)

Cut-off values were determined by ROC curve analysis (see table 3). Sensitivities, specificities, and positive and negative predictive values were presented as means and 95% confidence intervals. OR: chance of being obese when having high or low (for IRAK3) expressions.

TABLE 11

Association between gene expression in monocytes and the occurrence of metabolic syndrome determined by Fisher exact test						
Gene	Cut point	RR	Sensitivity %	Specificity %	PPV %	NPV %
SOD2	≥1.67	1.2 (0.9-1.8)	60 (46-72)	53 (39-65)	58 (45-70)	54 (40-68)
IRAK3	≥0.77	1.6 (1.1-2.2)	52 (39-65)	72 (58-83)	67 (52-80)	58 (45-69)
TNFAIP6	≥1.72	1.7 (1.1-2.5)	71 (58-82)	53 (39-66)	62 (50-73)	63 (47-76)
TNFAIP3	≥2.14	1.7 (1.2-2.4)	53 (40-66)	74 (60-84)	69 (54-81)	59 (47-71)

TABLE 11-continued

Association between gene expression in monocytes and the occurrence of metabolic syndrome determined by Fisher exact test						
Gene	Cut point	RR	Sensitivity %	Specificity %	PPV %	NPV %
ZNF217	≤0.77	1.4 (0.98-1.9)	52 (39-65)	65 (51-77)	62 (47-75)	55 (43-67)
IRAK3 + SOD2	Same as above	1.7 (1.0-2.8)	81 (69-90)	40 (28-54)	60 (48-70)	66 (48-80)
IRAK3 + TNFAIP3	Same as above	2.1 (1.3-3.3)	79 (67-88)	50 (37-64)	64 (52-74)	69 (53-82)
IRAK3 + TNFAIP6	Same as above	2.0 (1.1-3.6)	85 (74-93)	38 (25-51)	60 (49-71)	70 (51-85)
IRAK3 + ZNF217	Same as above	1.7 (1.1-2.7)	74 (62-84)	51 (37-64)	62 (50-73)	64 (49-78)
SOD2 + TNFAIP6	Same as above	2.1 (1.1-3.8)	87 (76-94)	35 (23-49)	59 (48-70)	71 (51-87)
SOD2 + TNFAIP3	Same as above	2.7 (1.4-5.1)	87 (76-94)	46 (32-59)	64 (52-74)	76 (59-89)
SOD2 + ZNF217	Same as above	1.7 (1.1-2.6)	76 (63-86)	46 (32-59)	60 (49-71)	63 (47-78)
TNFAIP3 + TNFAIP6	Same as above	1.8 (1.2-2.8)	74 (62-84)	53 (39-66)	63 (51-74)	65 (50-79)
TNFAIP3 + ZNF217	Same as above	2.4 (1.4-4.1)	82 (70-91)	53 (39-66)	65 (54-76)	73 (57-86)
TNFAIP6 + ZNF217	Same as above	1.8 (1.0-3.1)	85 (74-93)	33 (21-47)	60 (49-70)	67 (46-83)
IRAK3 + SOD2 + TNFAIP3	Same as above	3.9 (1.5-9.6)	94 (84-98)	37 (24-51)	62 (51-72)	84 (64-95)
IRAK3 + SOD2 + TNFAIP6	Same as above	4.5 (1.5-13)	95 (87-99)	33 (21-47)	61 (51-71)	86 (65-97)
IRAK3 + SOD2 + ZNF217	Same as above	1.9 (1.1-3.3)	88 (72-92)	37 (26-52)	60 (49-70)	69 (50-84)
IRAK3 + TNFAIP3 + TNFAIP6	Same as above	2.3 (1.2-4.2)	87 (76-94)	39 (26-52)	61 (50-71)	73 (54-88)
IRAK3 + TNFAIP3 + ZNF217	Same as above	3.8 (1.7-8.7)	92 (82-97)	44 (31-58)	64 (53-74)	83 (65-94)
IRAK3 + TNFAIP6 + ZNF217	Same as above	4.5 (1.5-13)	95 (86-99)	33 (21-47)	61 (50-71)	86 (65-97)
SOD2 + TNFAIP3 + TNFAIP6	Same as above	1.6 (0.99-2.7)	82 (70-91)	35 (23-49)	58 (47-68)	65 (45-81)
SOD2 + TNFAIP3 + ZNF217	Same as above	4.2 (1.7-11)	94 (84-98)	41 (28-54)	63 (52-73)	85 (66-96)
SOD2 + TNFAIP6 + ZNF217	Same as above	2.9 (1.3-6.4)	92 (82-97)	33 (21-47)	60 (49-70)	79 (58-93)
TNFAIP3 + TNFAIP6 + ZNF217	Same as above	2.4 (1.2-4.7)	89 (78-95)	37 (24-51)	60 (50-71)	75 (55-89)
IRAK3 + SOD2 + TNFAIP3 + TNFAIP6	Same as above	4.5 (1.5-13)	95 (87-99)	33 (21-47)	61 (50-71)	86 (65-97)
IRAK3 + SOD2 + TNFAIP3 + ZNF217	Same as above	13 (1.9-89)	98 (91-99)	35 (23-49)	62 (52-72)	95 (76-100)
IRAK3 + SOD2 + TNFAIP6 + ZNF217	Same as above	12 (1.7-79)	98 (91-99)	32 (20-45)	61 (51-71)	95 (74-100)
IRAK3 + TNFAIP3 + TNFAIP6 + ZNF217	Same as above	4.5 (1.5-13)	95 (86-99)	33 (21-47)	61 (50-71)	86 (65-97)
IRAK3 + SOD2 + TNFAIP3 + TNFAIP6 + ZNF217	Same as above	12 (1.9-89)	98 (91-99)	32 (20-45)	61 (51-71)	95 (74-100)

Cut points were determined by ROC curve analysis. Sensitivities, specificities, and positive and negative predictive values were presented as mean and 95% confidence intervals. RR: relative risk. NPV: negative predictive value; PPV: positive predictive value. Persons treated and not treated with a statin were included in analysis.

TABLE 12

Association between gene expression in monocytes and cardiovascular risk equivalents						
Gene	Cut point	RR	Sensitivity %	Specificity %	PPV %	NPV %
SOD2	≥1.85	1.4 (1.1-1.9)	61 (49-72)	61 (49-72)	73 (60-83)	48 (35-62)
TNFAIP3	≥2.14	1.7 (1.3-2.3)	54 (42-66)	82 (67-92)	83 (70-93)	51 (39-64)
IRAK3	≤0.77	1.7 (1.3-2.2)	51 (39-62)	84 (70-93)	84 (70-94)	50 (38-62)
TNFAIP6	≥1.77	1.9 (1.4-2.7)	72 (60-82)	68 (52-81)	79 (68-88)	59 (44-72)
TNFAIP3 + TNFAIP6	Same as above	1.8 (1.2-2.6)	75 (63-84)	59 (43-74)	76 (64-85)	58 (42-72)
IRAK3 + TNFAIP3	Same as above	2.2 (1.5-3.4)	80 (69-88)	62 (47-67)	78 (67-87)	65 (49-79)
IRAK3 + SOD2	Same as above	2.2 (1.4-3.4)	83 (72-90)	55 (39-70)	76 (65-84)	65 (47-80)
TNFAIP6 + SOD2	Same as above	2.8 (1.5-5.2)	89 (80-95)	50 (35-65)	75 (65-84)	73 (54-88)
TNFAIP3 + SOD2	Same as above	2.9 (1.7-5.2)	88 (78-94)	57 (41-72)	78 (67-86)	74 (56-87)
IRAK3 + TNFAIP6	Same as above	3.2 (1.8-5.7)	88 (78-94)	61 (45-76)	80 (69-88)	75 (58-88)
IRAK3 + TNFAIP3 + TNFAIP6	Same as above	3.2 (1.7-5.9)	89 (80-95)	56 (40-70)	77 (67-85)	76 (58-89)
IRAK3 + TNFAIP3 + SOD2	Same as above	5.0 (2.0-1.2)	95 (87-99)	50 (35-65)	76 (66-85)	85 (65-96)
IRAK3 + TNFAIP6 + SOD2	Same as above	6.4 (2.2-18)	96 (89-99)	50 (35-65)	77 (67-85)	88 (69-97)
TNFAIP3 + TNFAIP6 + SOD2	Same as above	9.2 (2.4-3.5)	97 (91-99)	50 (35-65)	77 (67-85)	92 (73-99)
IRAK3 + TNFAIP3 + TNFAIP6 + SOD2	Same as above	8.7 (2.3-33)	97 (91-99)	48 (32-65)	76 (66-84)	91 (72-99)

Cut points were determined by ROC curve analysis. Sensitivities, specificities, and positive and negative predictive values were presented as mean and 95% confidence intervals. RR: relative risk. NPV: negative predictive value; PPV: positive predictive value.

## REFERENCE LIST

[0288] (1) Grundy S M, Brewer H B, Jr., Cleeman J I, Smith S C, Jr., Lenfant C. Definition of metabolic syndrome: Report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues related to definition. *Circulation* 2004 Jan. 27; 109 (3):433-438.

[0289] (2) Isomaa B, Almgren P, Tuomi T, Forsen B, Lahti K, Nissen M, Taskinen M R, Groop L. Cardiovascular morbidity and mortality associated with the metabolic syndrome. *Diabetes Care* 2001 April; 24(4):683-689.

[0290] (3) Trevisan M, Liu J, Bahsas F B, Menotti A. Syndrome X and mortality: a population-based study. Risk Factor and Life Expectancy Research Group. *Am J Epidemiol* 1998 Nov. 15; 148(10):958-966.

- [0291] (4) National Institutes of Health. Third Report of the National Cholesterol Education Program Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Bethesda, Md.: National Institutes of Health; 2002. Report No.: NIH Publication 01-3670.
- [0292] (5) Ford E S, Giles W H, Dietz W H. Prevalence of the metabolic syndrome among US adults: findings from the third National Health and Nutrition Examination Survey. *JAMA* 2002 Jan. 16; 287(3):356-359.
- [0293] (6) Fagerberg B, Bondjers L, Nilsson P. Low birth weight in combination with catch-up growth predicts the occurrence of the metabolic syndrome in men at late middle age: the Atherosclerosis and Insulin Resistance study. *J Intern Med* 2004 September; 256(3):254-259.
- [0294] (7) Sinaiko A R, Steinberger J, Moran A, Prineas R J, Vessby B, Basu S, Tracy R, Jacobs D R, Jr. Relation of body mass index and insulin resistance to cardiovascular risk factors, inflammatory factors, and oxidative stress during adolescence. *Circulation* 2005 Apr. 19; 111(15):1985-1991.
- [0295] (8) Sjöholm A, Nystrom T. Inflammation and the etiology of type 2 diabetes. *Diabetes Metab Res Rev* 2005 Jul. 1.
- [0296] (9) Sjöholm A, Nystrom T. Inflammation and the etiology of type 2 diabetes. *Diabetes Metab Res Rev* 2005 Jul. 1.
- [0297] (10) Reilly M P, Lehrke M, Wolfe M L, Rohatgi A, Lazar M A, Rader D J. Resistin is an inflammatory marker of atherosclerosis in humans. *Circulation* 2005 Feb. 22; 111(7):932-939.
- [0298] (11) Lapointe A, Couillard C, Piche M E, Weisnagel S J, Bergeron J, Nadeau A, Lemieux S. Circulating oxidized LDL is associated with parameters of the metabolic syndrome in postmenopausal women. *Atherosclerosis* 2006 May 3.
- [0299] (12) Weinbrenner T, Schroder H, Escurriol V, Fito M, Elosua R, Vila J, Marrugat J, Covas M I. Circulating oxidized LDL is associated with increased waist circumference independent of body mass index in men and women. *Am J Clin Nutr* 2006 January; 83(1):30-35.
- [0300] (13) Yamagishi S I, Matsuo H, Kitano S, Hibi N, Jinnouchi Y, Umei H, Iida S, Takenaka K, Matsui T, Nakamura K, Imaizumi T. Elevated circulating oxidized LDL levels in Japanese subjects with the metabolic syndrome. *Int J Cardiol* 2006 Oct. 5.
- [0301] (14) Holvoet P, Kritchevsky S B, Tracy R P, Mertens A, Rubin S M, Butler J, Goodpaster B, Harris T B. The metabolic syndrome, circulating oxidized LDL, and risk of myocardial infarction in well-functioning elderly people in the health, aging, and body composition cohort. *Diabetes* 2004 April; 53(4):1068-1073.
- [0302] (15) Johnston N, Jernberg T, Lagerqvist B, Siegbahn A, Wallentin L. Improved identification of patients with coronary artery disease by the use of new lipid and lipoprotein biomarkers. *Am J Cardiol* 2006 Mar. 1; 97(5):640-645.
- [0303] (16) Meisinger C, Baumert J, Khuseynova N, Loewel H, Koenig W. Plasma oxidized low-density lipoprotein, a strong predictor for acute coronary heart disease events in apparently healthy, middle-aged men from the general population. *Circulation* 2005 Aug. 2; 112(5):651-657.
- [0304] (17) Holvoet P, Davey P C, De K D, Doukoure M, Derudder E, Bochaton-Piallat M L, Gabbiani G, Beaufort E, Bishay K, Andrieux N, Benhabiles N, Marguerie G. Oxidized low-density lipoprotein correlates positively with toll-like receptor 2 and interferon regulatory factor-1 and inversely with superoxide dismutase-1 expression: studies in hypercholesterolemic swine and THP-1 cells. *Arterioscler Thromb Vasc Biol* 2006 July; 26(7):1558-1565.
- [0305] (18) Wesche H, Gao X, Li X, Kirschning C J, Stark G R, Cao Z. IRAK-M is a novel member of the Pelle/interleukin-1 receptor-associated kinase (IRAK) family. *J Biol Chem* 1999 Jul. 2; 274(27):19403-19410.
- [0306] (19) Balaci L, Spada M C, Olla N, Sole G, Loddo L, Anedda F, Naitza S, Zuncheddu M A, Maschio A, Altea D, Uda M, Pilia S, Sanna S, Masala M, Crisponi L, Fattori M, Devoto M, Doratiotto S, Rassu S, Mereu S, Giua E, Cadeddu N G, Atzeni R, Pelosi U, Corrias A, Perra R, Torrazza P L, Pirina P, Ginesu F, Marcias S, Schintu M G, Del Giacco G S, Manconi P E, Malerba G, Bisognin A, Trabetti E, Boner A, Pescolliderung L, Pignatti P F, Schlessinger D, Cao A, Pilia G. IRAK-M is involved in the pathogenesis of early-onset persistent asthma. *Am J Hum Genet* 2007 June; 80(6):1103-1114.
- [0307] (20) Kobayashi K, Hernandez L D, Galan J E, Janeway C A, Jr., Medzhitov R, Flavell R A. IRAK-M is a negative regulator of Toll-like receptor signaling. *Cell* 2002 Jul. 26; 110(2):191-202.
- [0308] (21) Cao Z, Xiong J, Takeuchi M, Kurama T, Goeddel D V. TRAF6 is a signal transducer for interleukin-1. *Nature* 1996 Oct. 3; 383(6599):443-446.
- [0309] (22) King C G, Kobayashi T, Cejas P J, Kim T, Yoon K, Kim G K, Chiffolleau E, Hickman S P, Walsh P T, Turka L A, Choi Y. TRAF6 is a T cell-intrinsic negative regulator required for the maintenance of immune homeostasis. *Nat Med* 2006 September; 12(9):1088-1092.
- [0310] (23) Lord K A, Hoffman-Liebermann B, Liebermann D A. Sequence of MyD116 cDNA: a novel myeloid differentiation primary response gene induced by IL6. *Nucleic Acids Res* 1990 May 11; 18(9):2823.
- [0311] (24) Medzhitov R, Preston-Hurlburt P, Kopp E, Stauden A, Chen C, Ghosh S, Janeway C A, Jr. MyD88 is an adaptor protein in the hToll/IL-1 receptor family signaling pathways. *Mol Cell* 1998 August; 2(2):253-258.
- [0312] (25) Adachi O, Kawai T, Takeda K, Matsumoto M, Tsutsui H, Sakagami M, Nakanishi K, Akira S. Targeted disruption of the MyD88 gene results in loss of IL-1- and IL-18-mediated function. *Immunity* 1998 July; 9(1):143-150.
- [0313] (26) Kawai T, Adachi O, Ogawa T, Takeda K, Akira S. Unresponsiveness of MyD88-deficient mice to endotoxin. *Immunity* 1999 July; 11(1):115-122.
- [0314] (27) Bjorkbacka H, Kunjathoor V V, Moore K J, Koehn S, Ordijao C M, Lee M A, Means T, Halmen K, Luster A D, Golenbock D T, Freeman M W. Reduced atherosclerosis in MyD88-null mice links elevated serum cholesterol levels to activation of innate immunity signaling pathways. *Nat Med* 2004 April; 10(4):416-421.
- [0315] (28) Dixit V M, Green S, Sarma V, Holzman L B, Wolff W, O'Rourke K, Ward P A, Prochownik E V, Marks R M. Tumor necrosis factor-alpha induction of novel gene products in human endothelial cells including a macrophage-specific chemotaxin. *J Biol Chem* 1990 Feb. 15; 265(5):2973-2978.

- [0316] (29) Lee E G, Boone D L, Chai S, Libby S L, Chien M, Lodolce J P, Ma A. Failure to regulate TNF-induced NF-kappaB and cell death responses in A20-deficient mice. *Science* 2000 Sep. 29; 289(5488):2350-2354.
- [0317] (30) Opipari A W, Jr., Boguski M S, Dixit V M. The A20 cDNA induced by tumor necrosis factor alpha encodes a novel type of zinc finger protein. *J Biol Chem* 1990 Sep. 5; 265(25):14705-14708.
- [0318] (31) O'Reilly S M, Moynagh P N. Regulation of Toll-like receptor 4 signalling by A20 zinc finger protein. *Biochem Biophys Res Commun* 2003 Apr. 4; 303(2):586-593.
- [0319] (32) Wertz I E, O'Rourke K M, Zhou H, Eby M, Aravind L, Seshagiri S, Wu P, Wiesmann C, Baker R, Boone D L, Ma A, Koonin E V, Dixit V M. De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-kappaB signalling. *Nature* 2004 Aug. 5; 430(7000):694-699.
- [0320] (33) Boone D L, Turer E E, Lee E G, Ahmad R C, Wheeler M T, Tsui C, Hurley P, Chien M, Chai S, Hitotsumatsu O, McNally E, Pickart C, Ma A. The ubiquitin-modifying enzyme A20 is required for termination of Toll-like receptor responses. *Nat Immunol* 2004 October; 5(10):1052-1060.
- [0321] (34) Cook S A, Novikov M S, Ahn Y, Matsui T, Rosenzweig A. A20 is dynamically regulated in the heart and inhibits the hypertrophic response. *Circulation* 2003 Aug. 12; 108(6):664-667.
- [0322] (35) Lee T H, Klampfer L, Shows T B, Vilcek J. Transcriptional regulation of TSG6, a tumor necrosis factor- and interleukin-1-inducible primary response gene coding for a secreted hyaluronan-binding protein. *J Biol Chem* 1993 Mar. 25; 268(9):6154-6160.
- [0323] (36) Sun X J, Wang L M, Zhang Y, Yenush L, Myers M G, Jr., Glasheen E, Lane W S, Pierce J H, White M F. Role of IRS-2 in insulin and cytokine signalling. *Nature* 1995 Sep. 14; 377(6545):173-177.
- [0324] (37) Withers D J, Burks D J, Towery H H, Altamuro S L, Flint C L, White M F. Irs-2 coordinates Igf-1 receptor-mediated beta-cell development and peripheral insulin signalling. *Nat Genet* 1999 September; 23(1):32-40.
- [0325] (38) Tobe K, Suzuki R, Aoyama M, Yamauchi T, Kamon J, Kubota N, Terauchi Y, Matsui J, Akanuma Y, Kimura S, Tanaka J, Abe M, Ohsumi J, Nagai R, Kadowaki T. Increased expression of the sterol regulatory element-binding protein-1 gene in insulin receptor substrate-2(-/-) mouse liver. *J Biol Chem* 2001 Oct. 19; 276(42):38337-38340.
- [0326] (39) Taguchi A, Wartschow L M, White M F. Brain IRS2 signaling coordinates life span and nutrient homeostasis. *Science* 2007 Jul. 20; 317(5836):369-372.
- [0327] (40) Goedert M, Cuenda A, Craxton M, Jakes R, Cohen P. Activation of the novel stress-activated protein kinase SAPK4 by cytokines and cellular stresses is mediated by SKK3 (MKK6); comparison of its substrate specificity with that of other SAP kinases. *EMBO J* 1997 Jun. 16; 16(12):3563-3571.
- [0328] (41) Wang X S, Diener K, Manthey C L, Wang S, Rosenzweig B, Bray J, Delaney J, Cole C N, Chan-Hui P Y, Mantlo N, Lichenstein H S, Zukowski M, Yao Z. Molecular cloning and characterization of a novel p38 mitogen-activated protein kinase. *J Biol Chem* 1997 Sep. 19; 272(38):23668-23674.
- [0329] (42) Jagavelu K, Tietge U J, Gaestel M, Drexler H, Schieffer B, Bavendiek U. Systemic deficiency of the MAP kinase-activated protein kinase 2 reduces atherosclerosis in hypercholesterolemic mice. *Circ Res* 2007 Nov. 26; 101(11):1104-1112.
- [0330] (43) Morris J B, Olzinski A R, Bernard R E, Aravindhnan K, Mirabile R C, Boyce R, Willette R N, Jucker B M. p38 MAPK inhibition reduces aortic ultra-small superparamagnetic iron oxide uptake in a mouse model of atherosclerosis: MRI assessment. *Arterioscler Thromb Vasc Biol* 2008 February; 28(2):265-271.
- [0331] (44) Brunet A, Bonni A, Zigmond M J, Lin M Z, Juo P, Hu L S, Anderson M J, Arden K C, Blenis J, Greenberg M E. Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell* 1999 Mar. 19; 96(6):857-868.
- [0332] (45) Nemoto S, Finkel T. Redox regulation of forkhead proteins through a p66shc-dependent signaling pathway. *Science* 2002 Mar. 29; 295(5564):2450-2452.
- [0333] (46) Essers M A, de Vries-Smits L M, Barker N, Polderman P E, Burgering B M, Korswagen H C. Functional interaction between beta-catenin and FOXO in oxidative stress signaling. *Science* 2005 May 20; 308(5725):1181-1184.
- [0334] (47) Li Y, Huang T T, Carlson E J, Melov S, Ursell P C, Olson J L, Noble L J, Yoshimura M P, Berger C, Chan P H, Wallace D C, Epstein C J. Dilated cardiomyopathy and neonatal lethality in mutant mice lacking manganese superoxide dismutase. *Nat Genet* 1995 December; 11(4):376-381.
- [0335] (48) Melov S, Coskun P, Patel M, Tuinstra R, Cottrell B, Jun A S, Zastawny T H, Dizdaroglu M, Goodman S I, Huang T T, Miziorko H, Epstein C J, Wallace D C. Mitochondrial disease in superoxide dismutase 2 mutant mice. *Proc Natl Acad Sci USA* 1999 Feb. 2; 96(3):846-851.
- [0336] (49) Isola J J, Kallioniemi O P, Chu L W, Fuqua S A, Hilsenbeck S G, Osborne C K, Waldman F M. Genetic aberrations detected by comparative genomic hybridization predict outcome in node-negative breast cancer. *Am J Pathol* 1995 October; 147(4):905-911.
- [0337] (50) Kallioniemi A, Kallioniemi O P, Piper J, Tanner M, Stokke T, Chen L, Smith H S, Pinkel D, Gray J W, Waldman F M. Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization. *Proc Natl Acad Sci USA* 1994 Mar. 15; 91(6):2156-2160.
- [0338] (51) Collins C, Rommens J M, Kowbel D, Godfrey T, Tanner M, Hwang S I, Polikoff D, Nonet G, Cochran J, Myambo K, Jay K E, Froula J, Cloutier T, Kuo W L, Yaswen P, Dairkee S, Giovanola J, Hutchinson G B, Isola J, Kallioniemi O P, Palazzolo M, Martin C, Ericsson C, Pinkel D, Albertson D, Li W B, Gray J W. Positional cloning of ZNF217 and NABC1: genes amplified at 20q13.2 and overexpressed in breast carcinoma. *Proc Natl Acad Sci USA* 1998 July 21; 95(15):8703-8708.
- [0339] (52) Quinlan K G, Verger A, Yaswen P, Crossley M. Amplification of zinc finger gene 217 (ZNF217) and cancer: when good fingers go bad. *Biochim Biophys Acta* 2007 June; 1775(2):333-340.
- [0340] (53) Huang G, Krig S, Kowbel D, Xu H, Hyun B, Volik S, Feuerstein B, Mills G B, Stokoe D, Yaswen P, Collins C. ZNF217 suppresses cell death associated with chemotherapy and telomere dysfunction. *Hum Mol Genet* 2005 Nov. 1; 14(21):3219-3225.

- [0341] (54) Kretzschmar M. Transforming growth factor-beta and breast cancer: Transforming growth factor-beta/SMAD signaling defects and cancer. *Breast Cancer Res* 2000; 2(2):107-115.
- [0342] (55) Thillainadesan G, Iovic M, Loney E, Andrews J, Tini M, Torchia J. Genome analysis identifies the p15ink4b tumor suppressor as a direct target of the ZNF217/CoREST complex. *Mol Cell Biol* 2008 October; 28(19):6066-6077.
- [0343] (56) Bogdan C, Nathan C. Modulation of macrophage function by transforming growth factor beta, interleukin-4, and interleukin-10. *Ann NY Acad Sci* 1993 Jun. 23; 685:713-739.
- [0344] (57) Cummings DE, Overduin J, Foster-Schubert K E. Gastric bypass for obesity: mechanisms of weight loss and diabetes resolution. *J Clin Endocrinol Metab* 2004 June; 89(6):2608-2615.
- [0345] (58) Klein S, Burke L E, Bray G A, Blair S, Allison D B, Pi-Sunyer X, Hong Y, Eckel R H. Clinical implications of obesity with specific focus on cardiovascular disease: a statement for professionals from the American Heart Association Council on Nutrition, Physical Activity, and Metabolism: endorsed by the American College of Cardiology Foundation. *Circulation* 2004 Nov. 2; 110(18): 2952-2967.
- [0346] (59) Sjostrom L, Lindroos A K, Peltonen M, Torgerson J, Bouchard C, Carlsson B, Dahlgren S, Larsson B, Narbro K, Sjostrom C D, Sullivan M, Wedel H. Lifestyle, diabetes, and cardiovascular risk factors 10 years after bariatric surgery. *N Engl J Med* 2004 Dec. 23; 351(26): 2683-2693.
- [0347] (60) Wilson P W, D'Agostino R B, Levy D, Belanger A M, Silbershatz H, Kannel W B. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998 May 12; 97(18):1837-1847.
- [0348] (61) Pickl W F, Majdic O, Kohl P, Stockl J, Riedl E, Scheinecker C, Bello-Fernandez C, Knapp W. Molecular and functional characteristics of dendritic cells generated from highly purified CD14+ peripheral blood monocytes. *J Immunol* 1996 Nov. 1; 157(9):3850-3859.
- [0349] (62) Salio M, Cerundolo V, Lanzavecchia A. Dendritic cell maturation is induced by mycoplasma infection but not by necrotic cells. *Eur J Immunol* 2000 February; 30(2):705-708.
- [0350] (63) Holvoet P, Davey P C, De Keyzer D., Doukoure M, Deridder E, Bochaton-Piallat M L, Gabbiani G, Beaufort E, Bishay K, Andrieux N, Benhabiles N, Marguerie G. Oxidized low-density lipoprotein correlates positively with toll-like receptor 2 and interferon regulatory factor-1 and inversely with superoxide dismutase-1 expression: studies in hypercholesterolemic swine and THP-1 cells. *Arterioscler Thromb Vasc Biol* 2006 July; 26(7):1558-1565.
- [0351] (64) Holvoet P, Davey P C, De Keyzer D., Doukoure M, Deridder E, Bochaton-Piallat M L, Gabbiani G, Beaufort E, Bishay K, Andrieux N, Benhabiles N, Marguerie G. Oxidized low-density lipoprotein correlates positively with toll-like receptor 2 and interferon regulatory factor-1 and inversely with superoxide dismutase-1 expression: studies in hypercholesterolemic swine and THP-1 cells. *Arterioscler Thromb Vasc Biol* 2006 July; 26(7):1558-1565.
- [0352] (65) Saeed A I, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* 2003 February; 34(2):374-378.
- [0353] (66) Bar-Joseph Z, Gifford D K, Jaakkola T S. Fast optimal leaf ordering for hierarchical clustering. *Bioinformatics* 2001; 17 Suppl 1:S22-S29.
- [0354] (67) Eisen M B, Spellman P T, Brown P O, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998 Dec. 8; 95(25): 14863-14868.
- [0355] (68) Raychaudhuri S, Stuart J M, Altman R B. Principal components analysis to summarize microarray experiments: application to sporulation time series. *Pac Symp Biocomput* 2000; 455-466.
- [0356] (69) Bowie A, O'Neill L A. Oxidative stress and nuclear factor-kappaB activation: a reassessment of the evidence in the light of recent discoveries. *Biochem Pharmacol* 2000 Jan. 1; 59(1):13-23.
- [0357] (70) Pathak S K, Basu S, Bhattacharyya A, Pathak S, Kundu M, Basu J. *Mycobacterium tuberculosis* lipoarabinomannan-mediated IRAK-M induction negatively regulates Toll-like receptor-dependent interleukin-12 p40 production in macrophages. *J Biol Chem* 2005 December 30; 280(52):42794-42800.
- [0358] (71) Heyninck K, Kreike M M, Beyaert R. Structure-function analysis of the A20-binding inhibitor of NF-kappa B activation, ABIN-1. *FEBS Lett* 2003 Feb. 11; 536(1-3):135-140.
- [0359] (72) Wisniewski H G, Hua J C, Poppers D M, Naime D, Vilcek J, Cronstein B N. TNF/IL-1-inducible protein TSG-6 potentiates plasmin inhibition by inter-alpha-inhibitor and exerts a strong anti-inflammatory effect in vivo. *J Immunol* 1996 Feb. 15; 156(4):1609-1615.
- [0360] (73) Evans J L, Goldfine I D, Maddux B A, Grodsky G M. Are oxidative stress-activated signaling pathways mediators of insulin resistance and beta-cell dysfunction? *Diabetes* 2003 January; 52(1):1-8.
- [0361] (74) Nunn A V, Bell J, Barter P. The integration of lipid-sensing and anti-inflammatory effects: how the PPARs play a role in metabolic balance. *Nucl Recept* 2007; 5(1):1.
- [0362] (75) Asada S, Daitoku H, Matsuzaki H, Saito T, Sudo T, Mukai H, Iwashita S, Kako K, Kishi T, Kasuya Y, Fukamizu A. Mitogen-activated protein kinases, Erk and p38, phosphorylate and regulate Foxo1. *Cell Signal* 2007 March; 19(3):519-527.
- [0363] (76) O'Neill L A, Bowie A G. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nat Rev Immunol* 2007 May; 7(5):353-364.
- [0364] (77) Pennline K J. FISHing for cytokines. Methodology combining flow cytometry and in situ hybridization. *Methods Mol Biol* 1998; 91:197-215.
- [0365] (78) Moore C, Hibbitts S, Owen N, Corden S A, Harrison G, Fox J, Gelder C, Westmoreland D. Development and evaluation of a real-time nucleic acid sequence based amplification assay for rapid detection of influenza A. *J Med Virol* 2004 December; 74(4):619-628.
- [0366] (79) Belluzo M S, Ribone M E, Lagier C M. Assembling amperometric biosensors for clinical diagnostics. *Sensors* 2008 March; 8(3):1366-1399.
- [0367] (80) Kim Y, Sohn D, Tan W. Molecular beacons in biomedical detection and clinical diagnosis. *Int J Clin Exp Pathol* 2008; 1(2):105-116.

## SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 25

<210> SEQ ID NO 1

<211> LENGTH: 2288

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

```
gcctgtcgca ggcgtgcagg gacctggact ccgcctcgtc cccggggctc gggcagccga    60
gccatggcgg ggaactgtgg ggcccgcggc gcgctgtcgg cgcacacgct gctgttcgac    120
ctgcgcccgcg cgtgctcgg agagctctgc gctgttctgg acagctgcga cggcgcgctg    180
ggctggcgcg gctggcaga gagactttca agcagctggc tggatgttcg tcatattgaa    240
aagtatgtag accaaggtaa aagtggaaca agagaattac tttggctctg ggcacagaaa    300
aacaagacca tcggtgacct tttcacagtc ctccaggaga tgggacatcg tcgagctatt    360
catttaatta caaactatgg agcagtggtg agtccttcag agaagagtta tcaggaaggt    420
ggattttcaa atatattatt caaggaaaca gccaatgtca ccgtggataa tgttcttatt    480
cctgaacata atgaaaaagg agtactgctt aaatcttcca tcagctttca aaatatcata    540
gaaggaacta gaaattttca caaagacttc ctaattggag aaggagagat ttttgaggta    600
tacagagtgg agattcaaaa cctaacatat gctgtcaaat tatttaaaaa ggagaaaaaa    660
atgcagtgta agaagcattg gaagaggttt ttatctgagc ttgaagtttt actactgttt    720
catcacccaa acatactaga gttggctgca tattttacag agactgagaa gttctgtctg    780
atztatccat acatgagaaa tggaaacctt tttgacagat tgcagtggtg aggtgacacg    840
gccccactcc cttggcacat tcgaatcggg atattaatag gaatatccaa agccattcac    900
tacctgcaca acgttcaacc atgctcgggc atctgtggca gtatatcaag tgcaaacatc    960
cttttgatg atcagtttca acccaacta actgattttg ccatggcaca cttccgggtc   1020
cacctagaac atcagagtgg taccataaat atgaccagca gcagcagtaa acatctgtgg   1080
tacatgccag aagagtacat cagacagggg aaactttcca ttaaacaga tgtctacagc   1140
tttggaattg taataatgga agttctaaca ggatgtagag tagtgtaga tgatccaaaa   1200
catatccagc tgcgggatct ccttagagaa ttgatggaga agagaggcct ggattcatgt   1260
ctctcatttc tagataagaa agtgccctcc tgcctcggga atttctctgc caagctcttc   1320
tgtttgccag gccggtgtgc tgcaacgcgg gcaaagttaa gaccatcaat ggatgaagtt   1380
ttaaatactc ttgaaagtac tcaagccagc ttgtattttg ctgaagatcc tcccacatca   1440
ctaaagtcc tcaagtgctc ttctcctcta ttctgggaga atgtaccaag tattccagtg   1500
gaagatgatg aaagccagaa taacaattta ctacctctg atgaaggcct gaggatagac   1560
agaatgactc agaaaactcc ttttgaatgc agccagtctg aggttatggt tctgagcttg   1620
gacaaaaagc cagagagcaa gagaaatgag gaagcttgca acatgccccag ttcttcttgt   1680
gaagaaagtg ggttcccaaa gtatatagtt ccatcccagg acttaaggcc ctataaggta   1740
aatatagatc cttcttcaga agctccaggg cattcttgca ggagcaggcc agtggagagc   1800
agctgttctc ccaaattttc ctgggatgaa tatgaacagt acaaaaaaga ataaattcta   1860
ccagaagata aagaaaaaag caagtattgc ataggcacct gagcataggt atgaccttgg   1920
```

-continued

---

```

gaagacattg gctccataag caatgccaaag agaatgatca atagtgagtt tgggtgatgc 1980
agataaaciaa tctggataat tccatttttt ttttccaaa cctcaaaaca gaggcctta 2040
aaaaattggtt ttatcaggat aattgtctca tgaccaaatac cacgctcaat tagagccatt 2100
caaaattcct taagatcatg ggttctgact tcagccaaac aaaacaatca aaacctacca 2160
aaaagggact ggattgtaat gtcctctcca tcatcctcag tgtgagtcct cagagcctcc 2220
atctgccaaag aacattcagt tggattccat cgtttgggtt agcttgcttg cacgggttgt 2280
agggaaatg 2288

```

```

<210> SEQ ID NO 2
<211> LENGTH: 596
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

```

```

<400> SEQUENCE: 2

```

```

Met Ala Gly Asn Cys Gly Ala Arg Gly Ala Leu Ser Ala His Thr Leu
1          5          10          15
Leu Phe Asp Leu Pro Pro Ala Leu Leu Gly Glu Leu Cys Ala Val Leu
20         25         30
Asp Ser Cys Asp Gly Ala Leu Gly Trp Arg Gly Leu Ala Glu Arg Leu
35         40         45
Ser Ser Ser Trp Leu Asp Val Arg His Ile Glu Lys Tyr Val Asp Gln
50         55         60
Gly Lys Ser Gly Thr Arg Glu Leu Leu Trp Ser Trp Ala Gln Lys Asn
65         70         75         80
Lys Thr Ile Gly Asp Leu Leu Gln Val Leu Gln Glu Met Gly His Arg
85         90         95
Arg Ala Ile His Leu Ile Thr Asn Tyr Gly Ala Val Leu Ser Pro Ser
100        105        110
Glu Lys Ser Tyr Gln Glu Gly Gly Phe Pro Asn Ile Leu Phe Lys Glu
115        120        125
Thr Ala Asn Val Thr Val Asp Asn Val Leu Ile Pro Glu His Asn Glu
130        135        140
Lys Gly Val Leu Leu Lys Ser Ser Ile Ser Phe Gln Asn Ile Ile Glu
145        150        155        160
Gly Thr Arg Asn Phe His Lys Asp Phe Leu Ile Gly Glu Gly Glu Ile
165        170        175
Phe Glu Val Tyr Arg Val Glu Ile Gln Asn Leu Thr Tyr Ala Val Lys
180        185        190
Leu Phe Lys Gln Glu Lys Lys Met Gln Cys Lys Lys His Trp Lys Arg
195        200        205
Phe Leu Ser Glu Leu Glu Val Leu Leu Leu Phe His His Pro Asn Ile
210        215        220
Leu Glu Leu Ala Ala Tyr Phe Thr Glu Thr Glu Lys Phe Cys Leu Ile
225        230        235        240
Tyr Pro Tyr Met Arg Asn Gly Thr Leu Phe Asp Arg Leu Gln Cys Val
245        250        255
Gly Asp Thr Ala Pro Leu Pro Trp His Ile Arg Ile Gly Ile Leu Ile
260        265        270
Gly Ile Ser Lys Ala Ile His Tyr Leu His Asn Val Gln Pro Cys Ser
275        280        285

```

-continued

---

Val Ile Cys Gly Ser Ile Ser Ser Ala Asn Ile Leu Leu Asp Asp Gln  
 290 295 300  
 Phe Gln Pro Lys Leu Thr Asp Phe Ala Met Ala His Phe Arg Ser His  
 305 310 315 320  
 Leu Glu His Gln Ser Cys Thr Ile Asn Met Thr Ser Ser Ser Ser Lys  
 325 330 335  
 His Leu Trp Tyr Met Pro Glu Glu Tyr Ile Arg Gln Gly Lys Leu Ser  
 340 345 350  
 Ile Lys Thr Asp Val Tyr Ser Phe Gly Ile Val Ile Met Glu Val Leu  
 355 360 365  
 Thr Gly Cys Arg Val Val Leu Asp Asp Pro Lys His Ile Gln Leu Arg  
 370 375 380  
 Asp Leu Leu Arg Glu Leu Met Glu Lys Arg Gly Leu Asp Ser Cys Leu  
 385 390 395 400  
 Ser Phe Leu Asp Lys Lys Val Pro Pro Cys Pro Arg Asn Phe Ser Ala  
 405 410 415  
 Lys Leu Phe Cys Leu Ala Gly Arg Cys Ala Ala Thr Arg Ala Lys Leu  
 420 425 430  
 Arg Pro Ser Met Asp Glu Val Leu Asn Thr Leu Glu Ser Thr Gln Ala  
 435 440 445  
 Ser Leu Tyr Phe Ala Glu Asp Pro Pro Thr Ser Leu Lys Ser Phe Arg  
 450 455 460  
 Cys Pro Ser Pro Leu Phe Leu Glu Asn Val Pro Ser Ile Pro Val Glu  
 465 470 475 480  
 Asp Asp Glu Ser Gln Asn Asn Asn Leu Leu Pro Ser Asp Glu Gly Leu  
 485 490 495  
 Arg Ile Asp Arg Met Thr Gln Lys Thr Pro Phe Glu Cys Ser Gln Ser  
 500 505 510  
 Glu Val Met Phe Leu Ser Leu Asp Lys Lys Pro Glu Ser Lys Arg Asn  
 515 520 525  
 Glu Glu Ala Cys Asn Met Pro Ser Ser Ser Cys Glu Glu Ser Trp Phe  
 530 535 540  
 Pro Lys Tyr Ile Val Pro Ser Gln Asp Leu Arg Pro Tyr Lys Val Asn  
 545 550 555 560  
 Ile Asp Pro Ser Ser Glu Ala Pro Gly His Ser Cys Arg Ser Arg Pro  
 565 570 575  
 Val Glu Ser Ser Cys Ser Ser Lys Phe Ser Trp Asp Glu Tyr Glu Gln  
 580 585 590  
 Tyr Lys Lys Glu  
 595

<210> SEQ ID NO 3  
 <211> LENGTH: 522  
 <212> TYPE: PRT  
 <213> ORGANISM: Homo sapiens

<400> SEQUENCE: 3

Met Ser Leu Leu Asn Cys Glu Asn Ser Cys Gly Ser Ser Gln Ser Glu  
 1 5 10 15  
 Ser Asp Cys Cys Val Ala Met Ala Ser Ser Cys Ser Ala Val Thr Lys  
 20 25 30  
 Asp Asp Ser Val Gly Gly Thr Ala Ser Thr Gly Asn Leu Ser Ser Ser  
 35 40 45

-continued

---

Phe Met Glu Glu Ile Gln Gly Tyr Asp Val Glu Phe Asp Pro Pro Leu  
 50 55 60

Glu Ser Lys Tyr Glu Cys Pro Ile Cys Leu Met Ala Leu Arg Glu Ala  
 65 70 75 80

Val Gln Thr Pro Cys Gly His Arg Phe Cys Lys Ala Cys Ile Ile Lys  
 85 90 95

Ser Ile Arg Asp Ala Gly His Lys Cys Pro Val Asp Asn Glu Ile Leu  
 100 105 110

Leu Glu Asn Gln Leu Phe Pro Asp Asn Phe Ala Lys Arg Glu Ile Leu  
 115 120 125

Ser Leu Met Val Lys Cys Pro Asn Glu Gly Cys Leu His Lys Met Glu  
 130 135 140

Leu Arg His Leu Glu Asp His Gln Ala His Cys Glu Phe Ala Leu Met  
 145 150 155 160

Asp Cys Pro Gln Cys Gln Arg Pro Phe Gln Lys Phe His Ile Asn Ile  
 165 170 175

His Ile Leu Lys Asp Cys Pro Arg Arg Gln Val Ser Cys Asp Asn Cys  
 180 185 190

Ala Ala Ser Met Ala Phe Glu Asp Lys Glu Ile His Asp Gln Asn Cys  
 195 200 205

Pro Leu Ala Asn Val Ile Cys Glu Tyr Cys Asn Thr Ile Leu Ile Arg  
 210 215 220

Glu Gln Met Pro Asn His Tyr Asp Leu Asp Cys Pro Thr Ala Pro Ile  
 225 230 235 240

Pro Cys Thr Phe Ser Thr Phe Gly Cys His Glu Lys Met Gln Arg Asn  
 245 250 255

His Leu Ala Arg His Leu Gln Glu Asn Thr Gln Ser His Met Arg Met  
 260 265 270

Leu Ala Gln Ala Val His Ser Leu Ser Val Ile Pro Asp Ser Gly Tyr  
 275 280 285

Ile Ser Glu Val Arg Asn Phe Gln Glu Thr Ile His Gln Leu Glu Gly  
 290 295 300

Arg Leu Val Arg Gln Asp His Gln Ile Arg Glu Leu Thr Ala Lys Met  
 305 310 315 320

Glu Thr Gln Ser Met Tyr Val Ser Glu Leu Lys Arg Thr Ile Arg Thr  
 325 330 335

Leu Glu Asp Lys Val Ala Glu Ile Glu Ala Gln Gln Cys Asn Gly Ile  
 340 345 350

Tyr Ile Trp Lys Ile Gly Asn Phe Gly Met His Leu Lys Cys Gln Glu  
 355 360 365

Glu Glu Lys Pro Val Val Ile His Ser Pro Gly Phe Tyr Thr Gly Lys  
 370 375 380

Pro Gly Tyr Lys Leu Cys Met Arg Leu His Leu Gln Leu Pro Thr Ala  
 385 390 395 400

Gln Arg Cys Ala Asn Tyr Ile Ser Leu Phe Val His Thr Met Gln Gly  
 405 410 415

Glu Tyr Asp Ser His Leu Pro Trp Pro Phe Gln Gly Thr Ile Arg Leu  
 420 425 430

Thr Ile Leu Asp Gln Ser Glu Ala Pro Val Arg Gln Asn His Glu Glu  
 435 440 445

-continued

---

Ile Met Asp Ala Lys Pro Glu Leu Leu Ala Phe Gln Arg Pro Thr Ile  
 450 455 460

Pro Arg Asn Pro Lys Gly Phe Gly Tyr Val Thr Phe Met His Leu Glu  
 465 470 475 480

Ala Leu Arg Gln Arg Thr Phe Ile Lys Asp Asp Thr Leu Leu Val Arg  
 485 490 495

Cys Glu Val Ser Thr Arg Phe Asp Met Gly Ser Leu Arg Arg Glu Gly  
 500 505 510

Phe Gln Pro Arg Ser Thr Asp Ala Gly Val  
 515 520

<210> SEQ ID NO 4  
 <211> LENGTH: 2594  
 <212> TYPE: DNA  
 <213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4

```

gggcagtggc gtccgcagct ggggcttggc ctgcgggcgg ccagcgaagg tggcgaaggc   60
tcccactgga tccagatgtt cccgtccaag cagcctcgtc tcggcgcgca gtgtctgtgt   120
cgcctctcta ccgggcctt ggctgagcgg agtcgtgcgg ttggtggggg agccctgccc   180
tcctggttcg gctcctccgc gcactagaac gatcatgaac ttctgaaggg acccagcttt   240
ctttgtgtgc tccaagtgat ttgcacaaat aataatata atatttattg aaggagagaa   300
tcagagcaag tgataatcaa gttactatga gtctgctaaa ctgtgaaaac agctgtggat   360
tcagccagtc tgaagtgc tgctgtgtgg ccatggccag ctctctgtagc gctgtaacaa   420
aagatgatag tgtgggtgga actgccagca cggggaacct ctccagctca tttatggagg   480
agatccaggg atatgatgta gagtttgacc caccctgga aagcaagtat gaatgcccc   540
tctgcttgat ggcattacga gaagcagtc aaacgccatg cggccatagg ttctgcaaa   600
cctgcatcat aaaatcaata agggatgcag gtcacaaatg tccagttgac aatgaaatac   660
tgctggaaaa tcaactatct ccagacaatt ttgcaaacg tgagattctt tctctgatgg   720
tgaatgtcc aaatgaaggt gtgttgaca agatggaact gagacatctt gaggatcatc   780
aagcacattg tgagtttgc tttatggatt gtecccaatg ccagcgtccc ttccaaaaat   840
tccatattaa tattcacatt ctgaaggatt gtccaaggag acaggtttct tgtgacaact   900
gtgctgcatc aatggcattt gaagataaag agatccatga ccagaactgt cctttggcaa   960
atgtcatctg tgaatactgc aatactatac tcatcagaga acagatgcct aatcattatg  1020
atctagactg ccctacagcc ccaattccat gcacattcag tacttttggg tgccatgaaa  1080
agatgcagag gaatcacttg gcacgccacc tacaagagaa caccagtcac cacatgagaa  1140
tgttggccca ggctgttcat agtttgagcg ttataccgca ctctgggtat atctcagagg  1200
tccggaatct ccaggaaact attcaccagt tagagggctg ccttgtaaga caagaccatc  1260
aaatccggga gctgactgct aaaatggaaa ctacagatg gtatgtaagt gagctcaaac  1320
gaaccattcg aacccttgag gacaaagtgt ctgaaatcga agcacagcag tgcaatggaa  1380
tttatatttg gaagattggc aactttggaa tgcatttgaa atgtcaagaa gaggagaaac  1440
ctgttgatgc tcatagccct ggattctaca ctggcaaac ccgggtacaaa ctgtgcatgc  1500
gcttgacact tcagttaccg actgctcagc gctgtgcaaa ctatatatcc ctttttgtcc  1560
acacaatgca aggagaatat gacagccacc tcccttgccc cttccagggt acaatacgcc  1620

```

-continued

---

```

ttacaattct tgatcagtct gaagcacctg taaggcaaaa ccacgaagag ataatggatg 1680
ccaaaccaga gctgcttgct ttccagcgcac ccacaatccc acggaaccca aaagggttttg 1740
gctatgtaac ttttatgcat ctggaagccc taagacaaaag aactttcatt aaggatgaca 1800
cattattagt gcgctgtgag gtctocaccc gctttgacat gggtagcctt cggaggggagg 1860
gttttcagcc acgaagtact gatgcagggg tatagcttgc cctcacttgc tcaaaaacaa 1920
ctacctggag aaaacagtgc ctttccttgc cctgttctca ataacatgca aacaaacaag 1980
ccacgggaaa tatgtaatat ctactagtga gtgttggtag agaggtcact tactatttct 2040
tcctgttaca aatgatctga ggcagttttt tcctgggaat ccacacgttc catgcttttt 2100
cagaaatgtt aggcctgaag tgctgtggc atgttgacgc agctattttg ccagttagta 2160
tacctctttg ttgacttttc ttgggctttt gctctgggtg attttatgt cagaaagtcc 2220
agactcaaga gtactaaact ttttaataata atggattttc cttaaaactt cagtcttttt 2280
gtagtattat atgtaatata ttaaaagtga aaatcactac cgccttgtgc tagtgccttc 2340
gagaagagtt attgctctag aaagttagt tctcattttt ttaacctggt atagatttca 2400
gaggatttga accataatcc ttggaaaact taagttctca ttcaccccag tttttcctcc 2460
aggttgttac taaggatatt cagggatgag tttaaaccct aaataataacc ttaattattt 2520
agtgtaaaca tgtctgttga ataactcttg ttaagtgtt aaaaaaaaaa aaaaaaagaa 2580
aaaaaaaaaa aaaa 2594

```

&lt;210&gt; SEQ ID NO 5

&lt;211&gt; LENGTH: 2594

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Homo sapiens

&lt;400&gt; SEQUENCE: 5

```

gggcagtggc gtccgcagct ggggcttggc ctgcgggcgg ccagcgaagg tggcgaaggc 60
tcccactgga tccagagttt gccgtccaag cagcctcgtc tcggcgcgca gtgtctgtgt 120
ccgtccteta ccggcgcctt ggctgagcgg agtcgtgcgg ttggtggggg agccctgccc 180
tcctgttctg gctcctccgc gcaactagaac gatcatgaac tttcgaaggg acccagcttt 240
ctttgtgtgc tccaagtgat ttgcacaaat aataatata atatttattg aaggagagaa 300
tcagagcaag tgataatcaa gttactatga gtctgctaaa ctgtgaaaac agctgtggat 360
tcagccagtc tgaagtgtac tgctgtgtgg ccatggccag ctctgttagc gctgtaacaa 420
aagatgatag tgtgggtgga actgccagca cggggaacct ctccagctca tttatggagg 480
agatccaggg atatgatgta gagtttgacc caccctgga aagcaagtat gaatgcccc 540
tctgcttgat ggcattacga gaagcagtgc aaacgccatg cggccatagg ttctgcaaa 600
cctgcatcat aaaatcaata agggatgcag gtcacaaatg tccagttgac aatgaaatac 660
tgctgaaaaa tcaactatth ccagacaatt ttgcaaaacg tgagattctt tctctgatgg 720
tgaaatgtcc aaatgaaggt tgtttgaca agatggaact gagacatctt gaggatcatc 780
aagcactatg tgagtttgc tttatggatt gtcccaatg ccagcgtccc ttccaaaaat 840
tccatattaa tattcacatt ctgaaggatt gtccaaggag acaggtttct tgtgacaact 900
gtgctgcatc aatggcattt gaagataaag agatccatga ccagaactgt cctttggcaa 960
atgcatctg tgaactctgc aatactatac tcatcagaga acagatgcct aatcattatg 1020

```

-continued

---

```

atctagactg cctacagcc ccaattccat gcacattcag tacttttggg tgccatgaaa 1080
agatgcagag gaatcacttg gcacgccacc tacaagagaa caccocagtca cacatgagaa 1140
tggtggccca ggtgttcat agtttgagcg ttatacccg cctcgggtat atctcagagg 1200
tccggaatth ccaggaactc attcaccagt tagagggtcg ccttgtaaga caagaccatc 1260
aaatccggga gctgactgct aaaatggaaa ctacagatg gtatgtaagt gagctcaaac 1320
gaaccattcg aacccttgag gacaaagtgg ctgaaatcga agcacagcag tgcaatggaa 1380
tttatatttg gaagattggc aactttggaa tgcatttgaa atgcaagaa gaggagaaac 1440
ctgttgatg tcatagccc ggattctaca ctggcaaac cgggtacaaa ctgtgcatgc 1500
gcttgacact tcagttaccg actgctcagc gctgtgcaaa ctatatatcc ctttttgtcc 1560
acacaatgca aggagaatat gacagccacc tcccttggcc ctccagggtt acaatacggc 1620
ttacaattct tgatcagtgc gaagcacctg taaggcaaaa ccacgaagag ataatggatg 1680
ccaaaccaga gctgcttgct ttccagcagc ccacaatccc acggaacca aaaggttttg 1740
gctatgtaac ttttatgcat ctggaagccc taagacaaa aactttcatt aaggatgaca 1800
cattattagt gcgctgtgag gtctccacc gctttgacat gggtagcctt cggaggggag 1860
gttttcagcc acgaagtact gatgcagggg tatagcttgc cctcactgct tcaaaaaaaa 1920
ctacctggag aaaacagtgc ctttccctgc cctgttctca ataacatgca acaaaaacag 1980
ccacgggaaa tatgtaatat ctactagtga gtgttggtag agaggteact tactatttct 2040
tcctgttaca aatgatctga ggcagttttt tcctgggaat ccacacgttc catgcttttt 2100
cagaaatggt aggcctgaag tgctgtggc atgttcagc agctattttg ccagttagta 2160
tacctctttg ttgtactttc ttgggctttt gctctgggtg attttattgt cagaaagtcc 2220
agactcaaga gtactaaact ttaataata atggatttcc cttaaaactt cagtcttttt 2280
gtagtattat atgtaataa ttaaaagtga aaatcactac cgccttgtgc tagtgccctc 2340
gagaagagtt attgctctag aaagttgagt tctcattttt ttaacctggt atagatttca 2400
gaggatttga accataatcc ttggaaaact taagtctca ttcaccccag ttttctctcc 2460
aggttgttac taaggatatt cagggatgag tttaaacctt aaatataacc ttaattattt 2520
agtgtaaaca tgtctgttga ataatacttg ttaagtgtt aaaaaaaaaa aaaaaaagaa 2580
aaaaaaaaaa aaaa 2594

```

&lt;210&gt; SEQ ID NO 6

&lt;211&gt; LENGTH: 2678

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Homo sapiens

&lt;400&gt; SEQUENCE: 6

```

atgcgacccg accgcgctga ggctccagga ccgcccgcca tggtgcaggg aggtcccggc 60
gccccggtct cgcgcccggt ctctccaca tctcccttc cctggctgc tctcaacatg 120
cgagtgcggc gccgcctgtc tctgttcttg aacgtgaggc cacagggtgg gcccgactgg 180
accgcgctgg cggaggagat ggactttgag tacttggaga tccggcaact ggagacacaa 240
gccccccca ctggcaggct gctggacgcc tggcaggggc gccctggcgc ctctgtaggc 300
cgactgctcg agctgcttac caagctgggc cgcgacgagc tgctgctgga gctgggacct 360
agcattgagg aggattgcca aaagtatatc ttgaagcagc agcaggagga ggctgagaag 420

```

-continued

---

cctttacagg tggccgctgt agacagcagt gtcccacgga cagcagagct ggcgggcatc	480
accacacttg atgacccccct ggggcatatg cctgagcgtt tccatgcctt catctgctat	540
tgccccagcg acatccagtt tgtgcaggag atgatccggc aactggaaca gacaaactat	600
cgactgaagt tgtgtgtgtc tgaccgcgat gtectgcctg gcacctgtgt ctgggtctatt	660
gctagtgagc tcatcgaaaa gaggtgccgc cggatgggtg tggttgtctc tgatgattac	720
ctgcagagca aggaatgtga cttccagacc aaatttgcac tcagcctctc tccaggtgcc	780
catcagaagc gactgatccc catcaagtac aaggcaatga agaaagagtt cccagcatc	840
ctgaggttca tcaactgtctg cgactacacc aaccctgca ccaaatcttg gttctggact	900
cgcttggcca aggccttgtc cctgccctga agactgttct gaggccctgg gtgtgtgtgt	960
atctgtctgc ctgtccatgt acttotgccc tgccctctcc tttctgtgta ggaggaatct	1020
gtgtcttact tacctctcaa ttctgggaga tgccaacttc acagacacgt ctgcagcagc	1080
tggacatcac atttcatgtc ctgcatggaa ccagtggctg tgagtggcat gtccacttgc	1140
tggattatca gccaggacac tatagaacag gaccagctga gactaagaag gaccagcaga	1200
gccagctcag ctctgagcca ttcacacatc ttcaccctca gtttctcac ttgaggagtg	1260
ggatggggag aacagagagt agctgtgttt gaatccctgt aggaaatggt gaagcatagc	1320
tctgggtctc ctgggggaga ccaggcttgg ctgcccggaga gctggctgtt gctggactac	1380
atgctggcca ctgctgtgac cagcacaactg ctggggcagc ttcttccaca gtgatgccta	1440
ctgatgcttc agtgcctctg cacaccgccc attccacttc ctcttcccc acagggcagg	1500
tggggaagca gtttggccca gcccaaggag accccatctt gagccttatt tcctaaggg	1560
tccacctctc atctgcatct ttcacacctc ccagcttctg cccaaccttc agcagtgaca	1620
agtccccaa agactgcctc gacagcttg ggctgtcttt catttccacc tgtcaggatg	1680
cctgtgttca tgctctcagc tccacctggc atgagaaggg atcctggcct ctggcatatt	1740
catcaagtat gagttctggg gatgagtcac tgtaatgatg tgagcaggga gccttctctc	1800
ctgggccacc tgcagagagc tttcccacca actttgtacc ttgattgcct tacaaagtta	1860
tttgtttaca aacagcgacc atataaaagc ctctgcccc aaagcttctg ggcacatggg	1920
cacatacaga ctcacataca gacacacaca tatatgtaca gacatgtact ctcacacaca	1980
caggcaccag catacacacg tttttctagg tacagctccc aggaacagct agtggggaaa	2040
gtcccacac tgaggagacc taacctatgc cctgaacaaa aattgggcac tcatctattc	2100
cttttctctt gtgtccctac tcattgaaac caaactctgg aaaggaccca atgtaccagt	2160
atttatacct ctaatgaagc acagagagag gaagagagct gcttaaacct acacaacaat	2220
gaactgcaga cacagctgtt ctctccctct ctcttccca gagcaattta tactttacc	2280
tcaggctgtc ctctggggag aagggtgcat ggtcttaggt gtctgtgccc caggacagac	2340
cctaggacc taaatccaat agaaaatgca tatctttgct ccaacttctcag ccaggctgga	2400
gcaaggtacc ttttcttagg atcttgggag ggaatggatg cccctctctg catgatcttg	2460
ttgaggcatt tagctgcat gcacctgtcc ccctttaaata ctgggcattt taaagccatc	2520
tcaagaggca tcttctacat gttttgtacg cattaaaata atttcaaaga tatctgagaa	2580
aagccgatat ttgccattct tcttatatcc tggaatatat cttgcatcct gagtttataa	2640
taataaataa tattctacct tggaaaaaaa aaaaaaaa	2678

-continued

```

<210> SEQ ID NO 7
<211> LENGTH: 296
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 7
Met Ala Ala Gly Gly Pro Gly Ala Gly Ser Ala Ala Pro Val Ser Ser
1          5          10
Thr Ser Ser Leu Pro Leu Ala Ala Leu Asn Met Arg Val Arg Arg Arg
20         25         30
Leu Ser Leu Phe Leu Asn Val Arg Thr Gln Val Ala Ala Asp Trp Thr
35         40         45
Ala Leu Ala Glu Glu Met Asp Phe Glu Tyr Leu Glu Ile Arg Gln Leu
50         55         60
Glu Thr Gln Ala Asp Pro Thr Gly Arg Leu Leu Asp Ala Trp Gln Gly
65         70         75         80
Arg Pro Gly Ala Ser Val Gly Arg Leu Leu Glu Leu Leu Thr Lys Leu
85         90         95
Gly Arg Asp Asp Val Leu Leu Glu Leu Gly Pro Ser Ile Glu Glu Asp
100        105        110
Cys Gln Lys Tyr Ile Leu Lys Gln Gln Gln Glu Glu Ala Glu Lys Pro
115        120        125
Leu Gln Val Ala Ala Val Asp Ser Ser Val Pro Arg Thr Ala Glu Leu
130        135        140
Ala Gly Ile Thr Thr Leu Asp Asp Pro Leu Gly His Met Pro Glu Arg
145        150        155        160
Phe Asp Ala Phe Ile Cys Tyr Cys Pro Ser Asp Ile Gln Phe Val Gln
165        170        175
Glu Met Ile Arg Gln Leu Glu Gln Thr Asn Tyr Arg Leu Lys Leu Cys
180        185        190
Val Ser Asp Arg Asp Val Leu Pro Gly Thr Cys Val Trp Ser Ile Ala
195        200        205
Ser Glu Leu Ile Glu Lys Arg Cys Arg Arg Met Val Val Val Val Ser
210        215        220
Asp Asp Tyr Leu Gln Ser Lys Glu Cys Asp Phe Gln Thr Lys Phe Ala
225        230        235        240
Leu Ser Leu Ser Pro Gly Ala His Gln Lys Arg Leu Ile Pro Ile Lys
245        250        255
Tyr Lys Ala Met Lys Lys Glu Phe Pro Ser Ile Leu Arg Phe Ile Thr
260        265        270
Val Cys Asp Tyr Thr Asn Pro Cys Thr Lys Ser Trp Phe Trp Thr Arg
275        280        285
Leu Ala Lys Ala Leu Ser Leu Pro
290        295

```

```

<210> SEQ ID NO 8
<211> LENGTH: 4426
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

```

```

<400> SEQUENCE: 8

```

```

tgccttgacc aggacttggg accttgcgaa aggatcgagg ggcccggaga ggtgttgagg 60

```

-continued

agcacaatgg ctgaacaagt ccttctcag gctttgtatt tgagcaatat gcggaaaget	120
gtgaagatac gggagagaac tccagaagac atttttaaac ctactaatgg gatcattcat	180
cattttaaaa ccatgcaccg atacacactg gaaatgttca gaacttgcca gttttgtcct	240
cagtttcggg agatcatcca caaagccctc atcgacagaa acatccaggc caccctggaa	300
agccagaaga aactcaactg gtgtcgagaa gtccggaagc ttgtggcgct gaaaacgaac	360
ggtagcggca attgcctcat gcatgccact tctcagtaca tgtggggcgt tcaggacaca	420
gacttggtag tgaggaaggc gctgttcagc acgctcaagg aacagacac acgcaacttt	480
aaattccgct ggcaactgga gtctctcaaa tctcaggaat ttgttgaaac ggggctttgc	540
tatgatactc ggaactggaa tgatgaatgg gacaatctta tcaaatggc ttccacagac	600
acaccatgg ccggaagtgg acttcagtag aactcactgg aagaaataca catattttgc	660
ctttgcaaca tctcagaag gccaatcatt gtcatttcag acaaatgct aagaagtgtg	720
gaatcagggt ccaatttcgc ccctttgaaa gtgggtggaa tttacttgcc tctccactgg	780
cctgcccagg aatgctacag ataccccatt gttctcggtc atgacagcca tcattttgta	840
cccttggtag ccctgaagga cagtgggctc gaaatccgag ctgttccact tgtaacaga	900
gaccggggaa gatttgaaga cttaaaagtt cactttttga cagatcctga aaatgagatg	960
aaggagaagc tcttaaaaga gtacttaatg gtgatagaaa tcccctcca aggctgggac	1020
catggcacia ctcatctcat caatgccgca aagttggatg aagctaactt accaaaagaa	1080
atcaatctgg tagatgatta ctttgaactt gttcagcatg agtacaagaa atggcaggaa	1140
aacagcagc agggggagg agaggggac gccagaatc ccattggaacc ttccgtgccc	1200
cagctttctc tcattgatgt aaaatgtgaa acgcccact gccctctctt catgtctgtg	1260
aacaccagc ctttatgcca tgagtgtcca gagaggcggc aaaagaatca aaacaaactc	1320
ccaaagtga actccaagcc gggccctgag gggctccctg gcattggcgt cggggcctct	1380
cggggagaag cctatgagcc cttggcgtgg aacctgagg agtccactgg ggggcctcat	1440
tcggccccc acgacagacc cagccctttt ctgttcagtg agaccactgc catgaagtgc	1500
aggagcccc gctgcccctt cacactgaat gtgcagcaca acggattttg tgaacgttgc	1560
cacaacgccc ggcaacttca cgccagccac gccccagacc acacaaggca cttggatccc	1620
gggaagtgcc aagcctgctt ccaggatgtt accaggacat ttaatgggat ctgcagtact	1680
tgcttcaaaa ggactacagc agaggcctcc tccagcctca gcaaccagcct cctccttcc	1740
tgccaccagc gttccaagtc agatccctcg cggctcgtec ggagcccctc ccgcattct	1800
tgccacagag ctggaacga cgccctgctt ggctgctgt ctcaagctgc acggactcct	1860
ggggacagga cggggacgag caagtgcaga aaagccggct gcgtgtatct tgggactcca	1920
gaaaaaagg gcttttgcac actgtgtttc atcgagtaca gagaaaaaa acattttgct	1980
gctgcctcag ggaagttag tcccacagc tccaggttcc agaaccat tccgtgcctg	2040
gggagggat gcggcaccct tgggaagcacc atgtttgaag gatactgcca gaagtgtttc	2100
attgaagctc agaatcagag atttcatgag gccaaaagga cagaagagca actgagatcg	2160
agccagcga gagatgtgct tcgaaccaca caaagcact caaggcccaa gtgcgcccgg	2220
gcctcctgca agaactcctt ggcctgcgc agcagaggagc totgatgga gtgtcagcat	2280
cccaaccaga ggaatggccc tggggcccac cggggtgagc ctgccccga agaccccc	2340

-continued

---

```

aagcagcggt gccgggcccc egcctgtgat ctttttgga atgccaaagt caacggctac 2400
tgcaaacgaat gctttcagtt caagcagatg tatggctaac cggaaacagg tgggtcacct 2460
cctgcaagaa gtggggcctc gagctgtcag tcatcatggt gctatectct gaaccctca 2520
gctgccactg caacagtggg ctttaagggg tctgagcagg agaggaaaga taagctcttc 2580
gtggtgcccc cgatgctcag gtttggtaac ccgggagtggt toccaggtgg ccttagaaaag 2640
caaagcttgt aactggcaag ggatgatgtc agattcagcc caaggttcct cctctcctac 2700
caagcaggag gccaggaact tctttggact tggaaagtggt gcggggactg gccgaggccc 2760
ctgcacctg cgcacagga ctgcttcctc gtcttggtg agaagggaa aagacacaca 2820
agtcgcgtgg gttggagaag ccagagccat tccacctccc ctccccagc atctctcaga 2880
gatgtgaagc cagatcctca tggcagcgag gccctctgca agaagctcaa ggaagctcag 2940
ggaaaatgga cgtattcaga gagtgttgt agttcatggt tttccctac ctgcccggt 3000
cctttcctga ggaccgggca gaaatgcaga accatccatg gactgtgatt ctgaggctgc 3060
tgagactgaa catgttcaca ttgacagaaa aacaagctgc tctttataat atgcacctt 3120
taaaaaatta gaatatttta ctgggaagac gtgtaactct ttgggttatt actgtcttta 3180
cttctaaaga agttagcttg aactgaggag taaaagtgtg tacatatata atataccctt 3240
acattatgta tgagggattt ttttaatta tattgaaatg ctgccctaga agtacaatag 3300
gaaggctaaa taataataac ctgtttctcg gttgtgttg gggcatgagc ttgtgtatac 3360
actgcttga taaactcaac cagctgcctt tttaaagga gctctagtc ttttgtgta 3420
attcaactta tttattttat tacaacttc aagattattt aagtgaagat atttcttcag 3480
ctctggggaa aatgccacag tgtctctctg agagaacatc cttgcttga gtcaggctgt 3540
gggcaagtcc ctgaccacag ggagtaaat ggctctttg atacactttt gcttgccctc 3600
ccaggaaaaga aggaattgca tccaaggtat acatacatat tcatogatgt ttcgtgcttc 3660
tccttatgaa actccagcta tgtaataaaa aactatactc tgtgttctgt taatgcctct 3720
gagtgctcta cctccttgga gatgagatag ggaaggagca gggatgagac tggcaatggt 3780
cacagggaaa gatgtggcct tttgtgatgg ttttatttc tgttaacact gtgtcctggg 3840
ggggtcggga agtccccctc atccccatggt accctgggat tgggacagca aaagccagta 3900
accatgagta tgaggaaatc tctttctggt gctggcttac agtttctctg tgtgctttgt 3960
ggttgctgtc atatttgctc tagaagaaaa aaaaaaagg aggggaaatg cttttcccc 4020
agagataaag gctgccattt tgggggtctg tacttatgac ctgaaaatat ttgtgatcca 4080
taactctaca cagcctttac tcatactatt aggcacactt tccccttaga gccccctaag 4140
ttttcccgag acgaatcttt ataatttctt ttccaagat accaaaataa cttcagtggt 4200
ttcatctaat tctcttaaag ttgatctctt aatattttgt gttgatcatt atttccatc 4260
ttaatgtgaa aaaaagtaat tatttatact tattataaaa agtatttgaa atttgacat 4320
ttaatgttcc ctaatagaaa gccacctatt ctttgttggg tttcttcaag tttttctaaa 4380
taaatgtaac ttttcacaag agtcaacatt aaaaaataa ttattt 4426

```

&lt;210&gt; SEQ ID NO 9

&lt;211&gt; LENGTH: 790

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Homo sapiens

-continued

&lt;400&gt; SEQUENCE: 9

Met Ala Glu Gln Val Leu Pro Gln Ala Leu Tyr Leu Ser Asn Met Arg  
1 5 10 15  
Lys Ala Val Lys Ile Arg Glu Arg Thr Pro Glu Asp Ile Phe Lys Pro  
20 25 30  
Thr Asn Gly Ile Ile His His Phe Lys Thr Met His Arg Tyr Thr Leu  
35 40 45  
Glu Met Phe Arg Thr Cys Gln Phe Cys Pro Gln Phe Arg Glu Ile Ile  
50 55 60  
His Lys Ala Leu Ile Asp Arg Asn Ile Gln Ala Thr Leu Glu Ser Gln  
65 70 75 80  
Lys Lys Leu Asn Trp Cys Arg Glu Val Arg Lys Leu Val Ala Leu Lys  
85 90 95  
Thr Asn Gly Asp Gly Asn Cys Leu Met His Ala Thr Ser Gln Tyr Met  
100 105 110  
Trp Gly Val Gln Asp Thr Asp Leu Val Leu Arg Lys Ala Leu Phe Ser  
115 120 125  
Thr Leu Lys Glu Thr Asp Thr Arg Asn Phe Lys Phe Arg Trp Gln Leu  
130 135 140  
Glu Ser Leu Lys Ser Gln Glu Phe Val Glu Thr Gly Leu Cys Tyr Asp  
145 150 155 160  
Thr Arg Asn Trp Asn Asp Glu Trp Asp Asn Leu Ile Lys Met Ala Ser  
165 170 175  
Thr Asp Thr Pro Met Ala Arg Ser Gly Leu Gln Tyr Asn Ser Leu Glu  
180 185 190  
Glu Ile His Ile Phe Val Leu Cys Asn Ile Leu Arg Arg Pro Ile Ile  
195 200 205  
Val Ile Ser Asp Lys Met Leu Arg Ser Leu Glu Ser Gly Ser Asn Phe  
210 215 220  
Ala Pro Leu Lys Val Gly Gly Ile Tyr Leu Pro Leu His Trp Pro Ala  
225 230 235 240  
Gln Glu Cys Tyr Arg Tyr Pro Ile Val Leu Gly Tyr Asp Ser His His  
245 250 255  
Phe Val Pro Leu Val Thr Leu Lys Asp Ser Gly Pro Glu Ile Arg Ala  
260 265 270  
Val Pro Leu Val Asn Arg Asp Arg Gly Arg Phe Glu Asp Leu Lys Val  
275 280 285  
His Phe Leu Thr Asp Pro Glu Asn Glu Met Lys Glu Lys Leu Leu Lys  
290 295 300  
Glu Tyr Leu Met Val Ile Glu Ile Pro Val Gln Gly Trp Asp His Gly  
305 310 315 320  
Thr Thr His Leu Ile Asn Ala Ala Lys Leu Asp Glu Ala Asn Leu Pro  
325 330 335  
Lys Glu Ile Asn Leu Val Asp Asp Tyr Phe Glu Leu Val Gln His Glu  
340 345 350  
Tyr Lys Lys Trp Gln Glu Asn Ser Glu Gln Gly Arg Arg Glu Gly His  
355 360 365  
Ala Gln Asn Pro Met Glu Pro Ser Val Pro Gln Leu Ser Leu Met Asp  
370 375 380  
Val Lys Cys Glu Thr Pro Asn Cys Pro Phe Phe Met Ser Val Asn Thr  
385 390 395 400

-continued

---

Gln Pro Leu Cys His Glu Cys Ser Glu Arg Arg Gln Lys Asn Gln Asn  
 405 410 415  
 Lys Leu Pro Lys Leu Asn Ser Lys Pro Gly Pro Glu Gly Leu Pro Gly  
 420 425 430  
 Met Ala Leu Gly Ala Ser Arg Gly Glu Ala Tyr Glu Pro Leu Ala Trp  
 435 440 445  
 Asn Pro Glu Glu Ser Thr Gly Gly Pro His Ser Ala Pro Pro Thr Ala  
 450 455 460  
 Pro Ser Pro Phe Leu Phe Ser Glu Thr Thr Ala Met Lys Cys Arg Ser  
 465 470 475 480  
 Pro Gly Cys Pro Phe Thr Leu Asn Val Gln His Asn Gly Phe Cys Glu  
 485 490 495  
 Arg Cys His Asn Ala Arg Gln Leu His Ala Ser His Ala Pro Asp His  
 500 505 510  
 Thr Arg His Leu Asp Pro Gly Lys Cys Gln Ala Cys Leu Gln Asp Val  
 515 520 525  
 Thr Arg Thr Phe Asn Gly Ile Cys Ser Thr Cys Phe Lys Arg Thr Thr  
 530 535 540  
 Ala Glu Ala Ser Ser Ser Leu Ser Thr Ser Leu Pro Pro Ser Cys His  
 545 550 555 560  
 Gln Arg Ser Lys Ser Asp Pro Ser Arg Leu Val Arg Ser Pro Ser Pro  
 565 570 575  
 His Ser Cys His Arg Ala Gly Asn Asp Ala Pro Ala Gly Cys Leu Ser  
 580 585 590  
 Gln Ala Ala Arg Thr Pro Gly Asp Arg Thr Gly Thr Ser Lys Cys Arg  
 595 600 605  
 Lys Ala Gly Cys Val Tyr Phe Gly Thr Pro Glu Asn Lys Gly Phe Cys  
 610 615 620  
 Thr Leu Cys Phe Ile Glu Tyr Arg Glu Asn Lys His Phe Ala Ala Ala  
 625 630 635 640  
 Ser Gly Lys Val Ser Pro Thr Ala Ser Arg Phe Gln Asn Thr Ile Pro  
 645 650 655  
 Cys Leu Gly Arg Glu Cys Gly Thr Leu Gly Ser Thr Met Phe Glu Gly  
 660 665 670  
 Tyr Cys Gln Lys Cys Phe Ile Glu Ala Gln Asn Gln Arg Phe His Glu  
 675 680 685  
 Ala Lys Arg Thr Glu Glu Gln Leu Arg Ser Ser Gln Arg Arg Asp Val  
 690 695 700  
 Pro Arg Thr Thr Gln Ser Thr Ser Arg Pro Lys Cys Ala Arg Ala Ser  
 705 710 715 720  
 Cys Lys Asn Ile Leu Ala Cys Arg Ser Glu Glu Leu Cys Met Glu Cys  
 725 730 735  
 Gln His Pro Asn Gln Arg Met Gly Pro Gly Ala His Arg Gly Glu Pro  
 740 745 750  
 Ala Pro Glu Asp Pro Pro Lys Gln Arg Cys Arg Ala Pro Ala Cys Asp  
 755 760 765  
 His Phe Gly Asn Ala Lys Cys Asn Gly Tyr Cys Asn Glu Cys Phe Gln  
 770 775 780  
 Phe Lys Gln Met Tyr Gly  
 785 790

-continued

```

<210> SEQ ID NO 10
<211> LENGTH: 1440
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 10
cagtcacatt tcagccactg ctctgagaat ttgtgagcag cccctaacag gctgttactt    60
cactacaact gacgatatga tcattttaat ttactttatt ctcttgctat ggaagacac    120
tcaaggatgg ggattcaagg atggaatfff tcataactcc atatggcttg aacgagcagc    180
cggtgtgtac cacagagaag cacggctctg caaatacaag ctcaoctacg cagaagctaa    240
ggcgggtgtg gaatttgaag gggccatctc cgcaacttac aagcagctag aggcagccag    300
aaaaattgga tttcatgtct gtgctgctgg atggatggct aagggcagag ttggataccc    360
cattgtgaag ccagggccca actgtggatt tggaaaaact ggcatattg attatggaat    420
ccgtctcaat aggagtgaag gatgggatgc ctattgctac aaccacacag caaaggagtg    480
tgggtggctc tttacagatc caaagcaaat ttttaaatct ccaggcttcc caaatgagta    540
cgaagataac caaatctgct actggcacat tagactcaag tatggtcagc gtattcacct    600
gagtttttta gattttgacc ttgaagatga cccaggttgc ttggctgatt atgttgaaat    660
atatgacagt tacgatgatg tccatggctt tgtgggaaga tactgtggag atgagcttcc    720
agatgacatc atcagtacag gaaatgtcat gaccttgaag tttctaagtg atgcttcagt    780
gacagctgga ggtttccaaa tcaaatatgt tgcaatggat cctgtatcca aatccagtca    840
aggaaaaaat acaagtacta cttctactgg aaataaaaac tttttagctg gaagatttag    900
ccacttataa aaaaaaaaaa aaggatgacg aaaacacaca gtgtttatgt tggaaatctt    960
tggaaactct ttgatctcac tgttattatt aacatttatt tattattttt ctaaatgtga   1020
aagcaataca taatttaggg aaaattggaa aatataaggaa actttaaacg agaaaatgaa   1080
acctctcata atcccactgc atagaaataa caagcgttaa cattttcata ttttttctt    1140
tcagtcattt ttctatttgg ggtatatgta tatatgtacc tatatgtatt tgcatttgaa   1200
athttggaa cctgctctat gtacagtttt gtattatact ttttaaatct tgaactttat   1260
aaacattttc tgaatcatt gattattcta caaaaacatg attttaaca gctgtaaaat   1320
attctatgat atgaatggtt tatgcattat ttaagcctgt ctctattggt ggaatttcag   1380
gtcattttca taaatattgt tgcaataaat atccttgaac acaaaaaaaaa aaaaaaaaaa   1440

```

```

<210> SEQ ID NO 11
<211> LENGTH: 277
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 11
Met Ile Ile Leu Ile Tyr Leu Phe Leu Leu Leu Trp Glu Asp Thr Gln
1         5         10        15
Gly Trp Gly Phe Lys Asp Gly Ile Phe His Asn Ser Ile Trp Leu Glu
20        25        30
Arg Ala Ala Gly Val Tyr His Arg Glu Ala Arg Ser Gly Lys Tyr Lys
35        40        45
Leu Thr Tyr Ala Glu Ala Lys Ala Val Cys Glu Phe Glu Gly Gly His
50        55        60

```

-continued

---

Leu Ala Thr Tyr Lys Gln Leu Glu Ala Ala Arg Lys Ile Gly Phe His  
 65 70 75 80

Val Cys Ala Ala Gly Trp Met Ala Lys Gly Arg Val Gly Tyr Pro Ile  
 85 90 95

Val Lys Pro Gly Pro Asn Cys Gly Phe Gly Lys Thr Gly Ile Ile Asp  
 100 105 110

Tyr Gly Ile Arg Leu Asn Arg Ser Glu Arg Trp Asp Ala Tyr Cys Tyr  
 115 120 125

Asn Pro His Ala Lys Glu Cys Gly Gly Val Phe Thr Asp Pro Lys Gln  
 130 135 140

Ile Phe Lys Ser Pro Gly Phe Pro Asn Glu Tyr Glu Asp Asn Gln Ile  
 145 150 155 160

Cys Tyr Trp His Ile Arg Leu Lys Tyr Gly Gln Arg Ile His Leu Ser  
 165 170 175

Phe Leu Asp Phe Asp Leu Glu Asp Asp Pro Gly Cys Leu Ala Asp Tyr  
 180 185 190

Val Glu Ile Tyr Asp Ser Tyr Asp Asp Val His Gly Phe Val Gly Arg  
 195 200 205

Tyr Cys Gly Asp Glu Leu Pro Asp Asp Ile Ile Ser Thr Gly Asn Val  
 210 215 220

Met Thr Leu Lys Phe Leu Ser Asp Ala Ser Val Thr Ala Gly Gly Phe  
 225 230 235 240

Gln Ile Lys Tyr Val Ala Met Asp Pro Val Ser Lys Ser Ser Gln Gly  
 245 250 255

Lys Asn Thr Ser Thr Thr Ser Thr Gly Asn Lys Asn Phe Leu Ala Gly  
 260 265 270

Arg Phe Ser His Leu  
 275

<210> SEQ ID NO 12  
 <211> LENGTH: 5828  
 <212> TYPE: DNA  
 <213> ORGANISM: Homo sapiens

<400> SEQUENCE: 12

```

cgggcggcgcg gtcggagggg gcccggcgcg agagccagac gccgccgctt gttttgggtg      60
gggctctcgg caactctccg aggaggagga ggaggagga ggaggggaga agtaactgca      120
gcggcagcgc cctcccagag aacaggcgtc ttcccgaac cttoccaaa cctcccccat      180
cccctctcgc cttgtcccc tcccctcctc cccagccgcc tggagcgagg ggcagggatg      240
agttctgtcc tccggccggt cccagctgc agtggctgcc cggtatcgtt tcgcatggaa      300
aagccacttt ctccaccgcg cgagatgggc cgggatgggg ctgcagagga cgcgcccgcg      360
ggcggcggca gcagcagcag cagcagcagc agcaacagca acagccgcag cgcgcggctc      420
tctgcgactg agctgttatt tggcggctg gtggcggctg ggaagggttg ggggtgggag      480
gaggcgaagg aggagggaga acccctgca acgttgggac ttggcaaccc gctccccct      540
gccaaggat atttaatttg cctcgggaat cgctgcttcc agaggggaac tcaggaggga      600
aggcgcgcgc gcgcgcgcgc tcctggaggg gcaccgcagg gacccccgac tgtcgcctcc      660
ctgtgccgga ctccagccgg ggcgacgaga gatgcatctt cgctccttcc tggtgccggc      720
ggcggctgag aggagacttg gctctcggag gatcggggct gccctcacc cggacgcact      780
    
```

-continued

gcctccccgc cggcgtgaag egcccgaaaa ctccggtegg gctctctcct gggctcagea	840
gotgogtctc ccttcagctg cccctccccg gcgcgggggg cggcgtggat ttcagagctg	900
gggtttctgc tgccctcagc cctgtttgca tgtgcccggc cgcggcgagg agcctccgcc	960
ccccaccceg ttgtttttcg gagcctccct ctgctcagcg ttggtggtgg cggtggcage	1020
atggcgagcc ctccggagag cgatggcttc tcggacgtgc gcaaggtggg ctacctgcgc	1080
aaacccaaga gcatgcacaa acgcttcttc gtactgcgcg cggccagcga ggcgggggc	1140
ccggcgcgcc tcgagtacta cgagaacgag aagaagtggc ggcacaagtc gagcgcccc	1200
aaacgctcga tcccccttga gagctgcttc aacatcaaca agcgggctga ctccaagaac	1260
aagcacctgg tggctctcta caccgggac gagcactttg ccatgcggc ggacagcgag	1320
gccgagcaag acagctggta ccaggctctc ctacagctgc acaaccgtgc taagggccac	1380
cacgacggag ctgcggccct cggggcggga ggtggtgggg gcagctgcag cggcagctcc	1440
ggccttggtg aggctgggga ggaattgagc tacggtgacg tgccccagg acccgcatc	1500
aaagaggtct ggaagtgat cctgaagccc aagggcctgg gtcagacaaa gaacctgatt	1560
ggtatctacc gcctttgcct gaccagcaag accatcagct tcgtgaagct gaactcggag	1620
gcagcggccg tggctgtgca gctgatgaac atcaggcctg gtggccactc ggaaaacttc	1680
ttcttcacgc aggtggggcc ttctgccgtg acggggcccc gggagtctct gatgcaggtg	1740
gatgactctg tggggccca gaacatgcac gagaccatcc tggaggccat gcgggcccag	1800
agtgatgagt tccgccctcg cagcaagagc cagtctctct ccaactgctc taaccccatc	1860
agcgtcccc tgccgcggga ccatctcaac aatccccgc ccagccaggg ggggctgacc	1920
cgccgatcac gcactgagag catcacccgc acctccccgg ccagcatggt gggcgggaag	1980
ccaggctcct tccgtgtccg cgcctccagt gacggcgaag gcaaccatgc ccgccagcc	2040
tcggtgagc gcagccctgt gagtcccagc accaacagaa cccacgccca ccggcatcgg	2100
ggcagcgcgc ggtgcaccc cccgctcaac cacagccgct ccatcccat gccggcttc	2160
cgtgctcgc ctccggccac cagcccggtc agtctgtctg ccagtagcac cagtggccat	2220
ggctccacct cggattgtct ctcccccagg cgatctagt ctccgggtgc tggttcccc	2280
agcgatggcg gtttcatctc ctccgatgag tatggctcca gtcctgcga tttccggagt	2340
tccttccgca gtgtcactcc ggattccctg ggccacaccc caccagcccg cggtgaggag	2400
gagctaagca actatatctg catgggtggc aaggggcctt ccaccctgac cgcgcccaac	2460
ggteactaca tttgtctcgg ggggtggcaat ggccaccgct gcaccccagg aacaggcttg	2520
ggcagagtc cagccttggc tggggatgaa gcagccagtg ctgcagatct ggataatcgg	2580
ttccgaaaga gaactcactc ggcaggcaca tcccctacca ttaccacca gaagaccccg	2640
tcccagtcct cagtggcttc cattgaggag tacacagaga tgatgcctgc ctaccacca	2700
ggagggtggca gtggaggccg actgccggga cacaggcact ccgctctctg gccacccgc	2760
tcctaccagc aggagggtct ggaatgcac ccttggagc gtcggggggg gcaccaccgc	2820
ccagacagct ccaccctcca cacggatgat ggctacatgc ccatgtcccc aggggtggcc	2880
ccagtgccca gtggccgaaa gggcagtgga gactatatgc ccatgagccc caagagcgtg	2940
totgccccac agcagatcat caatccatc agacgccatc cccagagagt ggacccaat	3000
ggctacatga tgatgtcccc cagcgggtggc tgctctcctg acattggagg tggccccagc	3060

-continued

---

agcagcagca gcagcagcaa cgccgtccct tccgggacca gctatggaaa gctgtggaca	3120
aacgggtag ggggccacca ctctcatgtc ttgcctcacc ccaaaccccc agtggagagc	3180
agcggtggtg agctcttacc ttgcacaggt gactacatga acatgtcacc agtgggggac	3240
tccaacacca gcagcccctc cgactgctac tacggccctg aggaccccca gcacaagcca	3300
gtcctctcct actactcatt gccaaagatcc ttttaagcaca cccagcgcgc cgggggagccg	3360
gaggaggggtg cccggcatca gcacctccgc ctttccacta gctctgggtcg ccttctctat	3420
gctgcaacag cagatgatcc ttctcttcc accagcagcg acagcctggg tgggggatac	3480
tgcgggggta ggtgggagcc cagccttcca catccccacc atcaggttct gcagcccac	3540
ctgcctcga aggtggacac agctgctcag accaatagcc gcttggcccg gccacagagg	3600
ctgtccctgg gggatcccaa ggcacgacc ttacctcggg cccgagagca gcagcagcag	3660
cagcagccct tgctgcacc tccagagccc aagagcccgg gggaaatgt caatattgaa	3720
tttgggagtg atcagctctg taacttctct ggcccgggtg ctttccacag ctcaccttct	3780
gtcaggtgtc catcccagct ccagccagct cccagagagg aagagactgg cactgaggag	3840
tacatgaaga tggacctggg gccgggcccg agggcagcct ggcaggagag cactggggtc	3900
gagatgggca gactgggccc tgcacctccc ggggtgctga gcatttgcag gctaccggg	3960
gcagtgccca gcagccgggg tgactacatg accatgcaga tgagttgtcc ccgtcagagc	4020
tacgtggaca cctcgcacc tgcccctgta agctatgctg acatgcgaac aggcattgct	4080
gcagaggagg tgagcctgcc cagggccacc atggctgctg cctcctcacc ctcagcagcc	4140
tctgcttccc cgactggggc tcaaggggca gcagagctgg ctgcccactc gtcctgctg	4200
ggggggcccac aaggacctgg gggcatgagc gccttcaacc ggggtaacct cagtctaac	4260
cgcaaccaga gtgccaaagt gatccgtgca gaccacaag ggtgccggcg gaggcatagc	4320
tccgagactt tctcctcaac acccagtgcc acccgggtgg gcaacacagt gccctttgga	4380
gccccgggag cagtaggggg cggtggcggg agcagcagca gcagcgagga tgtgaaacgc	4440
cacagctctg ctctctttga gaatgtgtgg ctgaggcctg gggagcttgg gggagcccc	4500
aaggagccag ccaaactgtg tggggctgct gggggtttgg agaatggtct taactacata	4560
gacctggatt tggtaagga cttcaaacag tgcctcagg agtgaccccc tgaaccgag	4620
cctccccacc ccccccccc tcatcaacc ctgggcagcg gtgagagcag ctccaccgc	4680
cgctcaagtg aggatttaag cgcctatgcc agcatcagtt tccagaagca gccagaggac	4740
cgtcagtagc tcaactggac atcacagcag aatgaagacc taaatgacct cagcaaatcc	4800
tcttctaact catgggtacc cagactctaa atatttcag attcacaact aggacctcat	4860
atcttctca tcagtagatg gtacgatgca tccatttcag tttgtttact ttaatcaatc	4920
ctcaggattt cattgactga actgcacggt ctatattgtg ccaagcgaaa aaaaaaatg	4980
cactgtgaca ccagaataat gagtctgcat aaacttcac ttcaacctta aggacttagc	5040
tggccacagt gagctgatgt gccaccacc gtgtcatgag agaatgggtt tactctcaat	5100
gcattttcaa gatacatttc atctgctgct gaaactgtgt acgacaaagc atcattgtaa	5160
attatttcat acaaaaactg tcacgttggg tggagagagt attaaatatt taacataggt	5220
tttgatttat atgtgtaatt ttttaaatga aaatgtaact tttcttacag cacatctttt	5280
ttttggatgt gggatggagg tatacaatgt tctgtttaa agagtggagc aaatgcttaa	5340

-continued

---

```

aacaaggctt aaaagagtag aatagggtat gatccttggt ttaagattgt aattcagaaa 5400
acataatata agaatcatag tgccatagat ggttctcaat tgtatagtta tatttgctga 5460
tactatctct tgcatataa acctgatggt gagctgagtt ccttataaga attaactctta 5520
atattgtatt ttttctgta agacaatagg ccatgttaat taaactgaag aaggatatat 5580
ttggctgggt gttttcaaat gtcagcttaa aattggtaat tgaatggaag caaaattata 5640
agaagaggaa attaaagtct tccattgcat gtattgtaaa cagaaggaga tgggtgattc 5700
cttcaattca aaagctctct ttggaatgaa caatgtgggc gtttgtaaat tctggaatg 5760
tctttctatt cataataaac tagatactgt tgatctttta aaaaaaaaaa aaaaaaaaaa 5820
aaaaaaaaa 5828

```

```

<210> SEQ ID NO 13
<211> LENGTH: 1242
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

```

```

<400> SEQUENCE: 13

```

```

Met Ala Ser Pro Pro Glu Ser Asp Gly Phe Ser Asp Val Arg Lys Val
1          5          10          15
Gly Tyr Leu Arg Lys Pro Lys Ser Met His Lys Arg Phe Phe Val Leu
20         25         30
Arg Ala Ala Ser Glu Ala Gly Gly Pro Ala Arg Leu Glu Tyr Tyr Glu
35         40         45
Asn Glu Lys Lys Trp Arg His Lys Ser Ser Ala Pro Lys Arg Ser Ile
50         55         60
Pro Leu Glu Ser Cys Phe Asn Ile Asn Lys Arg Ala Asp Ser Lys Asn
65         70         75         80
Lys His Leu Val Ala Leu Tyr Thr Arg Asp Glu His Phe Ala Ile Ala
85         90         95
Ala Asp Ser Glu Ala Glu Gln Asp Ser Trp Tyr Gln Ala Leu Leu Gln
100        105        110
Leu His Asn Arg Ala Lys Gly His His Asp Gly Ala Ala Ala Leu Gly
115        120        125
Ala Gly Gly Gly Gly Gly Ser Cys Ser Gly Ser Ser Gly Leu Gly Glu
130        135        140
Ala Gly Glu Asp Leu Ser Tyr Gly Asp Val Pro Pro Gly Pro Ala Phe
145        150        155        160
Lys Glu Val Trp Gln Val Ile Leu Lys Pro Lys Gly Leu Gly Gln Thr
165        170        175
Lys Asn Leu Ile Gly Ile Tyr Arg Leu Cys Leu Thr Ser Lys Thr Ile
180        185        190
Ser Phe Val Lys Leu Asn Ser Glu Ala Ala Ala Val Val Leu Gln Leu
195        200        205
Met Asn Ile Arg Arg Cys Gly His Ser Glu Asn Phe Phe Phe Ile Glu
210        215        220
Val Gly Arg Ser Ala Val Thr Gly Pro Gly Glu Phe Trp Met Gln Val
225        230        235        240
Asp Asp Ser Val Val Ala Gln Asn Met His Glu Thr Ile Leu Glu Ala
245        250        255
Met Arg Ala Met Ser Asp Glu Phe Arg Pro Arg Ser Lys Ser Gln Ser
260        265        270

```

-continued

---

Ser Ser Asn Cys Ser Asn Pro Ile Ser Val Pro Leu Arg Arg His His  
 275 280 285  
 Leu Asn Asn Pro Pro Pro Ser Gln Val Gly Leu Thr Arg Arg Ser Arg  
 290 295 300  
 Thr Glu Ser Ile Thr Ala Thr Ser Pro Ala Ser Met Val Gly Gly Lys  
 305 310 315 320  
 Pro Gly Ser Phe Arg Val Arg Ala Ser Ser Asp Gly Glu Gly Thr Met  
 325 330 335  
 Ser Arg Pro Ala Ser Val Asp Gly Ser Pro Val Ser Pro Ser Thr Asn  
 340 345 350  
 Arg Thr His Ala His Arg His Arg Gly Ser Ala Arg Leu His Pro Pro  
 355 360 365  
 Leu Asn His Ser Arg Ser Ile Pro Met Pro Ala Ser Arg Cys Ser Pro  
 370 375 380  
 Ser Ala Thr Ser Pro Val Ser Leu Ser Ser Ser Ser Thr Ser Gly His  
 385 390 395 400  
 Gly Ser Thr Ser Asp Cys Leu Phe Pro Arg Arg Ser Ser Ala Ser Val  
 405 410 415  
 Ser Gly Ser Pro Ser Asp Gly Gly Phe Ile Ser Ser Asp Glu Tyr Gly  
 420 425 430  
 Ser Ser Pro Cys Asp Phe Arg Ser Ser Phe Arg Ser Val Thr Pro Asp  
 435 440 445  
 Ser Leu Gly His Thr Pro Pro Ala Arg Gly Glu Glu Glu Leu Ser Asn  
 450 455 460  
 Tyr Ile Cys Met Gly Gly Lys Gly Pro Ser Thr Leu Thr Ala Pro Asn  
 465 470 475 480  
 Gly His Tyr Ile Leu Ser Arg Gly Gly Asn Gly His Arg Cys Thr Pro  
 485 490 495  
 Gly Thr Gly Leu Gly Thr Ser Pro Ala Leu Ala Gly Asp Glu Ala Ala  
 500 505 510  
 Ser Ala Ala Asp Leu Asp Asn Arg Phe Arg Lys Arg Thr His Ser Ala  
 515 520 525  
 Gly Thr Ser Pro Thr Ile Thr His Gln Lys Thr Pro Ser Gln Ser Ser  
 530 535 540  
 Val Ala Ser Ile Glu Glu Tyr Thr Glu Met Met Pro Ala Tyr Pro Pro  
 545 550 555 560  
 Gly Gly Gly Ser Gly Gly Arg Leu Pro Gly His Arg His Ser Ala Phe  
 565 570 575  
 Val Pro Thr Arg Ser Tyr Pro Glu Glu Gly Leu Glu Met His Pro Leu  
 580 585 590  
 Glu Arg Arg Gly Gly His His Arg Pro Asp Ser Ser Thr Leu His Thr  
 595 600 605  
 Asp Asp Gly Tyr Met Pro Met Ser Pro Gly Val Ala Pro Val Pro Ser  
 610 615 620  
 Gly Arg Lys Gly Ser Gly Asp Tyr Met Pro Met Ser Pro Lys Ser Val  
 625 630 635 640  
 Ser Ala Pro Gln Gln Ile Ile Asn Pro Ile Arg Arg His Pro Gln Arg  
 645 650 655  
 Val Asp Pro Asn Gly Tyr Met Met Met Ser Pro Ser Gly Gly Cys Ser  
 660 665 670

-continued

---

Pro Asp Ile Gly Gly Gly Pro Ser Ser Ser Ser Ser Ser Ser Asn Ala  
 675 680 685

Val Pro Ser Gly Thr Ser Tyr Gly Lys Leu Trp Thr Asn Gly Val Gly  
 690 695 700

Gly His His Ser His Val Leu Pro His Pro Lys Pro Pro Val Glu Ser  
 705 710 715 720

Ser Gly Gly Lys Leu Leu Pro Cys Thr Gly Asp Tyr Met Asn Met Ser  
 725 730 735

Pro Val Gly Asp Ser Asn Thr Ser Ser Pro Ser Asp Cys Tyr Tyr Gly  
 740 745 750

Pro Glu Asp Pro Gln His Lys Pro Val Leu Ser Tyr Tyr Ser Leu Pro  
 755 760 765

Arg Ser Phe Lys His Thr Gln Arg Pro Gly Glu Pro Glu Glu Gly Ala  
 770 775 780

Arg His Gln His Leu Arg Leu Ser Thr Ser Ser Gly Arg Leu Leu Tyr  
 785 790 795 800

Ala Ala Thr Ala Asp Asp Ser Ser Ser Ser Thr Ser Ser Asp Ser Leu  
 805 810 815

Gly Gly Gly Tyr Cys Gly Ala Arg Leu Glu Pro Ser Leu Pro His Pro  
 820 825 830

His His Gln Val Leu Gln Pro His Leu Pro Arg Lys Val Asp Thr Ala  
 835 840 845

Ala Gln Thr Asn Ser Arg Leu Ala Arg Pro Thr Arg Leu Ser Leu Gly  
 850 855 860

Asp Pro Lys Ala Ser Thr Leu Pro Arg Ala Arg Glu Gln Gln Gln Gln  
 865 870 875 880

Gln Gln Pro Leu Leu His Pro Pro Glu Pro Lys Ser Pro Gly Glu Tyr  
 885 890 895

Val Asn Ile Glu Phe Gly Ser Asp Gln Ser Gly Tyr Leu Ser Gly Pro  
 900 905 910

Val Ala Phe His Ser Ser Pro Ser Val Arg Cys Pro Ser Gln Leu Gln  
 915 920 925

Pro Ala Pro Arg Glu Glu Glu Thr Gly Thr Glu Glu Tyr Met Lys Met  
 930 935 940

Asp Leu Gly Pro Gly Arg Arg Ala Ala Trp Gln Glu Ser Thr Gly Val  
 945 950 955 960

Glu Met Gly Arg Leu Gly Pro Ala Pro Pro Gly Ala Ala Ser Ile Cys  
 965 970 975

Arg Pro Thr Arg Ala Val Pro Ser Ser Arg Gly Asp Tyr Met Thr Met  
 980 985 990

Gln Met Ser Cys Pro Arg Gln Ser Tyr Val Asp Thr Ser Pro Ala Ala  
 995 1000 1005

Pro Val Ser Tyr Ala Asp Met Arg Thr Gly Ile Ala Ala Glu Glu  
 1010 1015 1020

Val Ser Leu Pro Arg Ala Thr Met Ala Ala Ala Ser Ser Ser Ser  
 1025 1030 1035

Ala Ala Ser Ala Ser Pro Thr Gly Pro Gln Gly Ala Ala Glu Leu  
 1040 1045 1050

Ala Ala His Ser Ser Leu Leu Gly Gly Pro Gln Gly Pro Gly Gly  
 1055 1060 1065

Met Ser Ala Phe Thr Arg Val Asn Leu Ser Pro Asn Arg Asn Gln

-continued

1070	1075	1080
Ser Ala Lys Val Ile Arg 1085	Ala Asp Pro Gln Gly Cys 1090	Arg Arg Arg 1095
His Ser Ser Glu Thr Phe 1100	Ser Ser Thr Pro Ser 1105	Ala Thr Arg Val 1110
Gly Asn Thr Val Pro Phe 1115	Gly Ala Gly Ala Ala Val 1120	Gly Gly Gly 1125
Gly Gly Ser Ser Ser Ser 1130	Ser Glu Asp Val Lys Arg 1135	His Ser Ser 1140
Ala Ser Phe Glu Asn Val Trp 1145	Leu Arg Pro Gly Glu 1150	Leu Gly Gly 1155
Ala Pro Lys Glu Pro Ala Lys 1160	Leu Cys Gly Ala Ala 1165	Gly Gly Leu 1170
Glu Asn Gly Leu Asn Tyr Ile 1175	Asp Leu Asp Leu Val 1180	Lys Asp Phe 1185
Lys Gln Cys Pro Gln Glu Cys 1190	Thr Pro Glu Pro Gln 1195	Pro Pro Pro 1200
Pro Pro Pro Pro His Gln Pro 1205	Leu Gly Ser Gly Glu 1210	Ser Ser Ser 1215
Thr Arg Arg Ser Ser Glu Asp 1220	Leu Ser Ala Tyr Ala 1225	Ser Ile Ser 1230
Phe Gln Lys Gln Pro Glu Asp 1235	Arg Gln 1240	

<210> SEQ ID NO 14  
 <211> LENGTH: 1888  
 <212> TYPE: DNA  
 <213> ORGANISM: Homo sapiens

<400> SEQUENCE: 14

```

gcggcgcggg gcgggcgcag cgggggtcgg ggcgctggga gcccgttggg ccgcaaacgc    60
agccgccacg ccggggcccg cgagatcggg tgccccggat gagcctcacc cggaaaaagg    120
gctctacaaa gcaggcagtc aacaagaccg cctgggagct gcccaagacc tacgtgtccc    180
cgacgcacgt cggcagcggg gcctatggct cegtgtgctc ggccatcgac aagcggtcag    240
gggagaaggt ggccatcaag aagctgagcc gaccctttca gtccgagatc ttcgccaagc    300
gcgctaccg ggagctgctg ctgctgaagc acatgcagca tgagaacgtc attgggctcc    360
tggatgtctt caccaccagc tcctccctgc gcaactteta tgacttctac ctggtgatgc    420
cctcatgca gacggatctg cagaagatca tggggatgga gttcagtgag gagaagatcc    480
agtacctggt gtatcagatg ctcaaaggcc ttaagtacat ccactctgct ggggtcgtgc    540
acagggaacct gaagccaggc aacctggctg tgaatgagga ctgtgaaactg aagattctgg    600
atthtgggct ggcgcgacat gcagacgccc agatgactgg ctactgtgtg acccctggt    660
accgagcccc cgaggtgatc ctacagtgga tgcactacaa ccagacagtg gacatctggt    720
ctgtgggctg tatcatggca gagatgctga cagggaaaac tctgttcaag gggaaagatt    780
acctggacca gctgaccacg atcctgaaag tgaccggggt gcctggcacg gagtttgtgc    840
agaagctgaa cgacaaagcg gccaaatcct acatccagtc cctgccacag acccccagga    900
aggatttca ctagctgttc ccacgggcca gccccaggc tgccggacctg ctggagaaga    960
tgctggagct agacgtggac aagcgcctga cggccgcgca ggccctcacc catcccttct    1020
    
```

-continued

```

ttgaaccctt cggggaccct gaggaagaga cggaggccca gcagccgttt gatgattcct 1080
tagaacacga gaaactcaca gtggatgaat ggaagcagca catctacaag gagattgtga 1140
acttcagccc cattgcccgg aaggactcac ggcgcccggag tggcatgaag ctgtagggac 1200
tcctcttgca tggcaccacc gcccagacac tgcccaagga ccagtatttg tcaactacaa 1260
actcagccct tcttggaaata cagcctttca agcagaggac agaaggggtcc ttctccttat 1320
gtgggaaatg ggccctagtag atgcagaatt caaagatgtc ggttgggaga aactagctct 1380
gatcctaaca ggccacgtta aactgcccat ctggagaatc gcttcgaggt ggggcccttt 1440
ccttcccgcc agagtggggc tgagtgggcg ctgagccagg ccgggggect atggcagtga 1500
tgctgtgttg gtttcttagg gatgctctaa cgaattacca caaacctggt ggattgaaac 1560
agcagaactt gattccctta cagttctgga ggctggaaat ctgggatgga ggtgttgga 1620
gggctgtggt cctttgaaag gctctgggga agaatecttc cttggctctt ttagcttgt 1680
ggcggcagtg ggcagtcctg ggcattcccc agcttattgc tgcactcactc cagtctctgt 1740
ctcttctggt ctctctctt ttaacaacag tcattggatt tagggccac cctaactctg 1800
tgtgatctta tcttgatcct tattaattaa acctgcaaat actctagttc caataaaagt 1860
cacattctca ggttccaggt ggacatga 1888

```

```

<210> SEQ ID NO 15
<211> LENGTH: 365
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

```

```

<400> SEQUENCE: 15

```

```

Met Ser Leu Ile Arg Lys Lys Gly Phe Tyr Lys Gln Asp Val Asn Lys
1          5          10          15
Thr Ala Trp Glu Leu Pro Lys Thr Tyr Val Ser Pro Thr His Val Gly
20          25          30
Ser Gly Ala Tyr Gly Ser Val Cys Ser Ala Ile Asp Lys Arg Ser Gly
35          40          45
Glu Lys Val Ala Ile Lys Lys Leu Ser Arg Pro Phe Gln Ser Glu Ile
50          55          60
Phe Ala Lys Arg Ala Tyr Arg Glu Leu Leu Leu Lys His Met Gln
65          70          75          80
His Glu Asn Val Ile Gly Leu Leu Asp Val Phe Thr Pro Ala Ser Ser
85          90          95
Leu Arg Asn Phe Tyr Asp Phe Tyr Leu Val Met Pro Phe Met Gln Thr
100         105         110
Asp Leu Gln Lys Ile Met Gly Met Glu Phe Ser Glu Glu Lys Ile Gln
115         120         125
Tyr Leu Val Tyr Gln Met Leu Lys Gly Leu Lys Tyr Ile His Ser Ala
130         135         140
Gly Val Val His Arg Asp Leu Lys Pro Gly Asn Leu Ala Val Asn Glu
145         150         155         160
Asp Cys Glu Leu Lys Ile Leu Asp Phe Gly Leu Ala Arg His Ala Asp
165         170         175
Ala Glu Met Thr Gly Tyr Val Val Thr Arg Trp Tyr Arg Ala Pro Glu
180         185         190
Val Ile Leu Ser Trp Met His Tyr Asn Gln Thr Val Asp Ile Trp Ser

```

-continued

195	200	205	
Val Gly Cys Ile Met Ala Glu Met Leu Thr Gly Lys Thr Leu Phe Lys			
210	215	220	
Gly Lys Asp Tyr Leu Asp Gln Leu Thr Gln Ile Leu Lys Val Thr Gly			
225	230	235	240
Val Pro Gly Thr Glu Phe Val Gln Lys Leu Asn Asp Lys Ala Ala Lys			
245	250	255	
Ser Tyr Ile Gln Ser Leu Pro Gln Thr Pro Arg Lys Asp Phe Thr Gln			
260	265	270	
Leu Phe Pro Arg Ala Ser Pro Gln Ala Ala Asp Leu Leu Glu Lys Met			
275	280	285	
Leu Glu Leu Asp Val Asp Lys Arg Leu Thr Ala Ala Gln Ala Leu Thr			
290	295	300	
His Pro Phe Phe Glu Pro Phe Arg Asp Pro Glu Glu Glu Thr Glu Ala			
305	310	315	320
Gln Gln Pro Phe Asp Asp Ser Leu Glu His Glu Lys Leu Thr Val Asp			
325	330	335	
Glu Trp Lys Gln His Ile Tyr Lys Glu Ile Val Asn Phe Ser Pro Ile			
340	345	350	
Ala Arg Lys Asp Ser Arg Arg Arg Ser Gly Met Lys Leu			
355	360	365	

<210> SEQ ID NO 16  
 <211> LENGTH: 7341  
 <212> TYPE: DNA  
 <213> ORGANISM: Homo sapiens

<400> SEQUENCE: 16

```

gcgcgaggcc gtcgattcgc tcgcggtccc atcgcggcct ggccgggggg cgggtgtctgc      60
tgcgccaggt tcgctggcgc cacgtcttca ggtcctcctg ttctctgggag gcgggcgccc      120
caggactggg aggtggcggc agcggggcgg gactcgcgca ggacggggct ccggcccggg      180
ataaccaact ctctctctct ettctttggt gcttccccag gcggcggcgg cggegcgccg      240
gagccggagc cttecgggcg tccacgtccc tcccccgctg caccocgccc cggegcgaga      300
ggagagcgcg agagccccag ccgcgggcgg gcgggcggcg aagatggcag aggcaccggc      360
ttccccggcc ccgctctctc cgctcgaagt ggagctggac ccggagttcg agccccagag      420
ccgtccgcca tctgtacgt ggcccctgca aaggccggag ctccaagcga gcctgccaa      480
gccctcgggg gagacggcgc ccgactccat gatccccgag gaggaggacg atgaagacga      540
cgaggacggc gggggacggg ccggctcggc catggcgatc ggccggcgcg gcgggagcgg      600
cacgctgggc tcgggctgc tccttgagga ctggccccgg gtgctggcac ccggagggca      660
agacccccgg tctgggccag ccaccgcggc gggcgggctg agcgggggta cacagcgct      720
gctgcagcct cagcaaccgc tgccaccgcc gcagccgggg gcggctgggg gctccgggca      780
gccgaggaaa tgttcgtcgc ggcggaacgc ctggggaaac ctgtcctacg cggacctgat      840
caccocgccc atcgagagct ccccggaaca acggctcact ctgtcccaga tctacgagtg      900
gatggtcgtg tgcgtgcctt acttcaagga taaggcgac agcaacagct ctgccggctg      960
gaagaactcc atccgcaca acctgtcact gcatagtcga ttcatgcggg tccagaatga     1020
gggaactggc aagagctctt ggtggatcat caaccctgat ggggggaaga gcggaaaagc     1080
    
```

-continued

---

ccccggcgg	cgggctgtct	ccatggacaa	tagcaacaag	tataccaaga	gccgtggcgg	1140
cgcagccaag	aagaaggcag	cctgcagac	agccccgaa	tcagctgacg	acagtccttc	1200
ccagctctcc	aagtggcctg	gcagccccac	gtcacgcagc	agtgatgagc	tggatgcgtg	1260
gacggacttc	cgttcacgca	ccaattctaa	cgccagcaca	gtcagtgggc	gcctgtcgcc	1320
catcatggca	agcacagagt	tggatgaagt	ccaggacgat	gatgcgcctc	tctcgcccat	1380
gctctacagc	agctcagcca	gcctgtcacc	ttcagtaagc	aagccgtgca	cggtggaact	1440
gccacggctg	actgatatgg	caggcaccat	gaatctgaat	gatgggctga	ctgaaaacct	1500
catggacgac	ctgctggata	acatcacgct	cccgccatcc	cagccatcgc	ccactggggg	1560
actcatgcag	cggagctcta	gcttcccgtg	taccaccaag	ggctggggcc	tgggctcccc	1620
aaccagctcc	tttaacagca	cgggtgttcg	accttcactc	ctgaactccc	tacgccagtc	1680
tcccctgcag	accatccaag	agaacaagcc	agctaccttc	tcttccatgt	cacactatgg	1740
taaccagaca	ctccaggacc	tgtctcactc	ggactcactt	agccacagcg	atgtcatgat	1800
gacacagtcg	gaccccttga	tgtctcaggc	cagcaccgct	gtgtctgccc	agaattcccc	1860
ccggaacgtg	atgcttcgca	atgatccgat	gatgtccttt	gctgcccagc	ctaaccaggg	1920
aagtttggtc	aatcagaact	tgtctccacca	ccagcaccaa	accagggcgg	ctcttggtgg	1980
cagccgtgcc	ttgtcgaatt	ctgtcagcaa	catgggcttg	agtgagtcca	gcagccttgg	2040
gtcagccaaa	caccagcagc	agtctcctgt	cagccagtct	atgcaaaccc	tctcgacttc	2100
tctctcaggc	tctccttgt	actcaactag	tgcaaacctg	cccgtcatgg	gccatgagaa	2160
gttccccagc	gacttggacc	tggacatggt	caatgggagc	ttggaatgtg	acatggagtc	2220
cattatccgt	agtgaactca	tggatgctga	tgggttggat	tttaactttg	attccctcat	2280
ctccacacag	aatggttgtg	gtttgaaact	ggggaacttc	actggtgcta	agcaggcctc	2340
atctcagagc	tgggtgccag	gctgaaggat	caactgaggaa	ggggaagtgg	gcaaaacaga	2400
ccctcaaact	gacacaagac	ctacagagaa	aaccttttgc	caaatctgct	ctcagcaagt	2460
ggacagtgat	accgtttaca	gcttaacacc	tttgtgaatc	ccacgccatt	ttcctaacc	2520
agcagagact	gttaaatggc	ccttaccctg	ggtgaagcac	ttacccttgg	aacagaactc	2580
taaaaagtat	gcaaaatctt	ccttgtagag	ggtggtgagc	cgccctgcag	tggaggacag	2640
caccctcag	caccacccac	cctcattcag	agcacaccgt	gagccccctg	cggccattct	2700
gtggtgtttt	aatattgcga	tggtttatgg	gacgttttaa	gtgtgttctc	tgtgtttgtt	2760
ttcctttgac	tttctgagtt	tttcacatgc	attaacttgc	ggtattttcc	tgttaaatgt	2820
ttaaccgtcc	ttcccctagc	aaatttaaaa	acagaagaa	aatgttgtac	cagttaccat	2880
tccgggttcg	agcatcacia	gcttttgagc	gcatggaact	ccataaacta	acaaattaca	2940
taaactaaag	ggggattttc	tttcttcttt	tgtttggtag	aaaattatcc	ttttctaaaa	3000
actgaacaat	ggcacaattg	tttctatgt	gcacccgtcc	aggacagaac	cgtgcatagg	3060
caaaaggagt	ggagcacagc	gtccggccca	gtgtgtttcc	ggttctgagt	cagggtgatc	3120
tgtggacggg	accccagcac	caagtctacg	ggtgccagat	cagtagggcc	tgtgatttcc	3180
tgtcagtgtc	ctcagtaaat	gtgaacagtg	ttggtctgct	ggttagaaac	tagaatattg	3240
atattttcag	gaaagaaatc	agctcagctc	tccactcatt	gccaaatgtc	actaaagggt	3300
ttagttttaa	ggagaaagaa	aaggaaaaaa	aaaaaaaaaa	aaaaagtcct	gttttgcttt	3360

-continued

---

gcagaacaaa tgaacttaca ggtgagcatt aagcttgcag tgagaaatgt gcgaagagta	3420
aaaacccaag tcaatgctga ggcagttcta acttctactgt tttcctaata acacatcctt	3480
gattattttc agccttgcta tataatctga tctgctagaa gtgtatgagt gagaggcaat	3540
agcatacaaa ctgatttttt aaatataagc ttaggttcta attgtacaag tgactcaatg	3600
gaagtacaaa atagggcagt ttttaactttt ttttctgctt ctatggattt cattttggtg	3660
tgttttcaaa aagttatggt gctgtatagg tgctttctgt ttaacctgga aagtgtgatt	3720
atattcgtta ccttctttgg tagacggaat agttgggacc acctttggta cataagaaat	3780
tggtataacg atgctctgat tagcacagta tatgcatact tctccaaagt gatatatgaa	3840
gactcttttc tttgcataaa aagcattagg catataaatg tataaatata ttttatcatg	3900
tacagtacaa aaatggaacc ttatgcatgg gccttaggaa tacaggctag tatttcagca	3960
cagacttccc tgcttgagtt cttgctgatg cttgcaccgt gacagtgggc accaacacag	4020
acgtgccacc caacccctcg cacacaccac cggccaccag gggccccctt gtgcgccttg	4080
gctttataac tctctggtgg gtgatattgg tggatgatcac agctcctagc ataagagag	4140
ttccatttgg tattgtcaca cgtctcctgc ctgcttggg ttgccatggt tgagcgatgg	4200
ccctgttgat ttcacctgc cttttactga atctgtaaat tgtgtgcaa tttgtgttat	4260
agtagactgt agcacattgc cttttctaaa ctgctacatg tttataatct tcatttttaa	4320
agtatgtgta atttttttaa gtatgtatc tattcatatg gtctgcttgt cagttagcca	4380
gacttgctta ctatattcct ttataataat gctagccact tcctggattc tttagtaatg	4440
tgctgtatgc aagaacttcc cagtagcagt gaaggagggt tgcctctcca agcttctcaa	4500
gggatgctgc cctgtgtggg gatgcatcgc agaggcacta gtagcatggg ggctagagtg	4560
gggagcgaga tgtaaaaggg tggggggata ggagaattcc agagtgttc cagcattagg	4620
gtcctgagaa cttctgagtt cagagaaaca tgcaaagtga ctaacaaaat agctacttac	4680
ctttgcagtt ttacagacc tgggagctgc tttgggagtg agaaaggcaa ccctccaatg	4740
tgtttcaact ttaaaatggt gaattctttt cagacatggt atctcattta ttctcctttt	4800
ctagcgtttg ttgaatttca ggcagaatgt cttacagaat gtcttagaac cagattatca	4860
tttaacttga aacagctgag gaagggacag agaaggtaca agggcaaggc agcacaaaac	4920
agatcaggag aatgaagagg gaatgctttg gttttttggt ttgttttgtt ttttctttt	4980
caagtaacta aaacagcatc tacatgtaga gtgttggtga gagctgagac cagggtaaaag	5040
tcaagtgcag catcagtaact gcgagaccca ccagcccctg gagagggtca gccgagaatc	5100
tggtagttaa gcctgtctag ggtcccggca ccctcaccct cagccacctg cagagaggcc	5160
agggccccag agactagcct ggttctgaag tgggcagggg tgctgccaga gccctctgcc	5220
ccttatgttg agaccctgct ttcaggacag gccagccgtt ggccaccatg tcacattctg	5280
agtgagtgtc acaggtccc tacaataatt ttctgatctg gagcatatca gcagaatgct	5340
tagcctcaag gggcctggca cctgtaatgt ttgatttatg atgagaacta tccgaggcca	5400
cccttgccct ctaataaagc tgctctaggg agccgcctac tttttgatga gaaattagaa	5460
gagtacctaa tgttgaanaac atgacatgcg ctcttgggat ctgctgttct ctccagggt	5520
ccagaacctg atacctgtta ccaaagctag gaaagagctt tatcacaagc cttcactgtc	5580
ctggcatgag aactggctgc caggctcagt gtacccatt aactgtgaat gaatctgagc	5640

-continued

```

ttggtttcct ttattgcttc ctctgcaata tgattgctga aacacatttt aaaaattcag 5700
aagcttgta ctcctgttaa tgggaggatc agtcacacat gtgtagtaca aggcggactt 5760
tgtgtttgtt tttggtgta atttttagca ttgtgtgtgt tgcttcccca cectgaggag 5820
aggacaccat ggcttactac tcaggacaag tatgccccgc tcagggtgtg atttcagggtg 5880
gcttccaaac ttgtacgcag tttaaagatg gtggggacag actttgcctc tacctagtga 5940
accccactta aagaataagg agcatttgaa tctcttgtaa aaggccatga agaataaagc 6000
agtcaaaaag aagtcctcca tggttggtgcc aaggacttgc gaggggaaat aaaaatgtta 6060
tccagcctga ccaacatgga gaaaccccgct ctccattaaa aatacaaaat tagcctggca 6120
tggtggcgca tgctgtaat ccagctact ctggaggctg aggcaggaga atcgcttga 6180
cccaggaggc ggaggtcgca gtgagccgag atcatgccag tgcactccag cctgggtaac 6240
aagagtgaaa ctccgtgtca aaaaaaaaa aaaaatgtta ctcatcctct ctgaaagcaa 6300
aaaggaaacc ctaacagctc tgaactctgg ttttattttt cttgctgtat ttgggtgaac 6360
attgtatgat taggcataat gtaaaaaaaaa aaaatttttt tttggtagaa atgcaatcac 6420
cagtaaaagag gtacgaaaaa gctagcctct ctacagagacc ggggaggcag agtactacta 6480
gaggaagtga agttctgatg gaatcatgcc tgtcaaatga ggtcttgaag cggatgceca 6540
aataaaagag tatattttat ctaaacttta agtgggtaac attttatgca gtttaaatga 6600
atggaatatt ttctcttctg ttagtgtgat ctgtttggat tttctttga tgaatgattg 6660
gtcatgaggc ctcttgccac actccagaaa tacgtgtgcg gctgctttta agaactatgt 6720
gtctggctac ttatttctct aaaattatct cattgcctgg caatcagtct tctcttgtat 6780
acttgcctca gcacattatg tacatgggaa atgtaaaaa atgtaagga ggaccagaaa 6840
aattagttaa tatttaaaaa aatgtattgt gcattttggc ttcacatggt taactttttt 6900
taagaaaaaa gttgcatgaa tggaaaaaaa aatctgtata cagtatctgt aaaaactatc 6960
ttatctgttt caattccttg ctcatatccc atataatcta gaactaata tgggtgtgtg 7020
ccatatttaa acacctgaga gtcaagcagt tgagactttg atttgaagca cctcctcctt 7080
ctttcaatgc gaacactatc atatggcatt cttactgagg attttgtcta accatattgt 7140
gccatgaatt aactctgccc cctttcttaa ggatcaaaac cagtttgatt tgggaatcct 7200
cccctttcca aatgaaatag agatgcagta cttaactttc cttgggtgtt gtagatattg 7260
ccttgtgtat tccacttaaa accgtaatct agtttgtaaa agagatgggtg acgcatgtaa 7320
ataaagcctc agtgacactc t 7341

```

```

<210> SEQ ID NO 17
<211> LENGTH: 673
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 17

```

```

Met Ala Glu Ala Pro Ala Ser Pro Ala Pro Leu Ser Pro Leu Glu Val
1          5          10         15
Glu Leu Asp Pro Glu Phe Glu Pro Gln Ser Arg Pro Arg Ser Cys Thr
20         25         30
Trp Pro Leu Gln Arg Pro Glu Leu Gln Ala Ser Pro Ala Lys Pro Ser
35         40         45
Gly Glu Thr Ala Ala Asp Ser Met Ile Pro Glu Glu Glu Asp Asp Glu

```

-continued

50			55			60									
Asp 65	Asp 65	Glu 65	Asp 65	Gly 70	Gly 70	Gly 70	Arg 70	Ala 70	Gly 75	Ser 75	Ala 75	Met 75	Ala 75	Ile 75	Gly 80
Gly 85	Gly 85	Gly 85	Gly 85	Ser 85	Gly 85	Thr 85	Leu 85	Gly 90	Ser 90	Gly 90	Leu 90	Leu 90	Leu 90	Glu 95	Asp 95
Ser 100	Ala 100	Arg 100	Val 100	Leu 100	Ala 100	Pro 100	Gly 100	Gly 105	Gln 105	Asp 105	Pro 105	Gly 110	Ser 110	Gly 110	Pro 110
Ala 115	Thr 115	Ala 115	Ala 115	Gly 115	Gly 115	Leu 115	Ser 120	Gly 120	Gly 120	Thr 120	Gln 125	Ala 125	Leu 125	Leu 125	Gln 125
Pro 130	Gln 130	Gln 130	Pro 130	Leu 130	Pro 130	Pro 130	Pro 135	Gln 135	Pro 135	Gly 140	Ala 140	Ala 140	Gly 140	Gly 140	Ser 140
Gly 145	Gln 145	Pro 145	Arg 145	Lys 150	Cys 150	Ser 150	Ser 150	Arg 150	Arg 155	Asn 155	Ala 155	Trp 155	Gly 160	Asn 160	Leu 160
Ser 165	Tyr 165	Ala 165	Asp 165	Leu 165	Ile 165	Thr 165	Arg 165	Ala 170	Ile 170	Glu 170	Ser 170	Ser 170	Pro 175	Asp 175	Lys 175
Arg 180	Leu 180	Thr 180	Leu 180	Ser 180	Gln 180	Ile 180	Tyr 180	Glu 185	Trp 185	Met 185	Val 185	Arg 190	Cys 190	Val 190	Pro 190
Tyr 195	Phe 195	Lys 195	Asp 195	Lys 195	Gly 195	Asp 195	Ser 200	Asn 200	Ser 200	Ser 200	Ala 205	Gly 205	Trp 205	Lys 205	Asn 205
Ser 210	Ile 210	Arg 210	His 210	Asn 210	Leu 210	Ser 210	Leu 215	His 215	Ser 215	Arg 215	Phe 220	Met 220	Arg 220	Val 220	Gln 220
Asn 225	Glu 225	Gly 225	Thr 225	Gly 225	Lys 230	Ser 230	Ser 230	Trp 230	Trp 235	Ile 235	Ile 235	Asn 235	Pro 235	Asp 240	Gly 240
Gly 245	Lys 245	Ser 245	Gly 245	Lys 245	Ala 245	Pro 245	Arg 245	Arg 250	Arg 250	Ala 250	Val 250	Ser 250	Met 255	Asp 255	Asn 255
Ser 260	Asn 260	Lys 260	Tyr 260	Thr 260	Lys 260	Ser 260	Arg 260	Gly 265	Arg 265	Ala 265	Ala 265	Lys 265	Lys 270	Lys 270	Ala 270
Ala 275	Leu 275	Gln 275	Thr 275	Ala 275	Pro 275	Glu 275	Ser 275	Ala 280	Asp 280	Asp 280	Ser 280	Pro 285	Ser 285	Gln 285	Leu 285
Ser 290	Lys 290	Trp 290	Pro 290	Gly 290	Ser 290	Pro 290	Thr 295	Ser 295	Arg 295	Ser 295	Ser 300	Asp 300	Glu 300	Leu 300	Asp 300
Ala 305	Trp 305	Thr 305	Asp 305	Phe 305	Arg 310	Ser 310	Arg 310	Thr 310	Asn 315	Ser 315	Asn 315	Ala 315	Ser 315	Thr 320	Val 320
Ser 325	Gly 325	Arg 325	Leu 325	Ser 325	Pro 325	Ile 325	Met 325	Ala 325	Ser 330	Thr 330	Glu 330	Leu 330	Asp 335	Glu 335	Val 335
Gln 340	Asp 340	Asp 340	Asp 340	Ala 340	Pro 340	Leu 340	Ser 340	Pro 345	Met 345	Leu 345	Tyr 345	Ser 345	Ser 350	Ser 350	Ala 350
Ser 355	Leu 355	Ser 355	Pro 355	Ser 355	Val 355	Ser 355	Lys 355	Pro 360	Cys 360	Thr 360	Val 360	Glu 365	Leu 365	Pro 365	Arg 365
Leu 370	Thr 370	Asp 370	Met 370	Ala 370	Gly 370	Thr 375	Met 375	Asn 375	Leu 375	Asn 375	Asp 380	Gly 380	Leu 380	Thr 380	Glu 380
Asn 385	Leu 385	Met 385	Asp 385	Asp 385	Leu 390	Leu 390	Asp 390	Asn 390	Ile 395	Thr 395	Leu 395	Pro 395	Pro 395	Ser 400	Gln 400
Pro 405	Ser 405	Pro 405	Thr 405	Gly 405	Gly 405	Leu 405	Met 405	Gln 410	Arg 410	Ser 410	Ser 410	Ser 410	Phe 415	Pro 415	Tyr 415
Thr 420	Thr 420	Lys 420	Gly 420	Ser 420	Gly 420	Leu 420	Gly 420	Ser 425	Pro 425	Thr 425	Ser 425	Ser 425	Phe 430	Asn 430	Ser 430
Thr 435	Val 435	Phe 435	Gly 435	Pro 435	Ser 435	Ser 435	Leu 435	Asn 440	Ser 440	Leu 440	Arg 440	Gln 445	Ser 445	Pro 445	Met 445
Gln 450	Thr 450	Ile 450	Gln 450	Glu 450	Asn 450	Lys 450	Pro 450	Ala 450	Thr 450	Phe 450	Ser 450	Ser 450	Met 450	Ser 450	His 450

-continued

Tyr Gly Asn Gln Thr Leu Gln Asp Leu Leu Thr Ser Asp Ser Leu Ser  
 465 470 475 480  
 His Ser Asp Val Met Met Thr Gln Ser Asp Pro Leu Met Ser Gln Ala  
 485 490 495  
 Ser Thr Ala Val Ser Ala Gln Asn Ser Arg Arg Asn Val Met Leu Arg  
 500 505 510  
 Asn Asp Pro Met Met Ser Phe Ala Ala Gln Pro Asn Gln Gly Ser Leu  
 515 520 525  
 Val Asn Gln Asn Leu Leu His His Gln His Gln Thr Gln Gly Ala Leu  
 530 535 540  
 Gly Gly Ser Arg Ala Leu Ser Asn Ser Val Ser Asn Met Gly Leu Ser  
 545 550 555 560  
 Glu Ser Ser Ser Leu Gly Ser Ala Lys His Gln Gln Gln Ser Pro Val  
 565 570 575  
 Ser Gln Ser Met Gln Thr Leu Ser Asp Ser Leu Ser Gly Ser Ser Leu  
 580 585 590  
 Tyr Ser Thr Ser Ala Asn Leu Pro Val Met Gly His Glu Lys Phe Pro  
 595 600 605  
 Ser Asp Leu Asp Leu Asp Met Phe Asn Gly Ser Leu Glu Cys Asp Met  
 610 615 620  
 Glu Ser Ile Ile Arg Ser Glu Leu Met Asp Ala Asp Gly Leu Asp Phe  
 625 630 635 640  
 Asn Phe Asp Ser Leu Ile Ser Thr Gln Asn Val Val Gly Leu Asn Val  
 645 650 655  
 Gly Asn Phe Thr Gly Ala Lys Gln Ala Ser Ser Gln Ser Trp Val Pro  
 660 665 670

Gly

<210> SEQ ID NO 18  
 <211> LENGTH: 1593  
 <212> TYPE: DNA  
 <213> ORGANISM: Homo sapiens

<400> SEQUENCE: 18

gcggtgccct tgcggcgcag ctgggggtgc gccctgctc cccgcgcttt cttaaggccc 60  
 gcggggcggc caggagcggc actcgtggct gtggtggctt cggcagcggc ttcagcagat 120  
 cggcggcatc agcggtagca ccagcactag cagcatggtg agccgggcag tgtgcggcac 180  
 cagcaggcag ctggctccgg ttttggggta tctgggctcc aggcagaagc acagcctccc 240  
 cgacctgcc tacgactacg ggcacctgga acctcacatc aacgcgcaga tcatgcagct 300  
 gcaccacagc aagcaccacg cggcctacgt gaacaacctg aacgtcaccg aggagaagta 360  
 ccaggaggcg ttggccaagg gagatgttac agcccagata gctcttcagc ctgcaactgaa 420  
 gttcaatggt ggtggtcata tcaatcatag cattttctgg acaaacctca gcctaacg 480  
 tggtgagaaa cccaaagggg agttgctgga agccatcaaa cgtgactttg gttcctttga 540  
 caagttaag gagaagctga cggctgcacg tgttggtgtc caaggctcag gttggggttg 600  
 gcttggtttc aataaggaac ggggacactt acaaattgct gcttgccaa atcaggatcc 660  
 actgcaagga acaacaggcc ttattccact gctggggatt gatgtgtggg agcacgctta 720  
 ctaccttcag tataaaaatg tcaggcctga ttatctaaaa gctatttggg atgtaataca 780

-continued

```

ctgggagaat gtaactgaaa gatacatggc ttgcaaaaag taaaccaacga tcgttatgct 840
gagtatgtta agctctttat gactgttttt gtagtggtat agagtactgc agaatacagt 900
aagctgctct attgtagcat ttcttgatgt tgcttagtca cttatttcat aaacaactta 960
atgttctgaa taatttctta ctaaacattt tgttattggg caagtgattg aaaatagtaa 1020
atgctttgtg tgattgaato tgattggaca ttttcttcag agagctaaat tacaattgtc 1080
atttataaaa ccatcaaaaa tattocattc atatactttg gggacttcta gggatgcctt 1140
tctagtccca ttctattgca gttatagaaa atctagtctt ttgccccagt tacttaaaaa 1200
taaaaatatta acactttccc aagggaacaa ctcggcttcc tatagaaaat tgcacttttt 1260
gtcagagtaat cctctgcagt gatactctcg gtagatgtca cccagtggtt tttgttaggt 1320
caaatgttcc tgtatagttt ttgcaaatag agctgtatac tgtttaaatg tagcaggtga 1380
actgaactgg ggtttgctca cctgcacagt aaaggcaaac ttcaacagca aaactgcaaa 1440
aagggtggtt ttgcagtagg agaaaggagg atgtttattt gcagggcgcc aagcaaggag 1500
aattgggcag ctcagctctg agacccaatc tccatgatga cctacaagct agagtattta 1560
aaggcagtgg taaatttcag gaaagcagaa gtt 1593

```

&lt;210&gt; SEQ ID NO 19

&lt;211&gt; LENGTH: 222

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Homo sapiens

&lt;400&gt; SEQUENCE: 19

```

Met Leu Ser Arg Ala Val Cys Gly Thr Ser Arg Gln Leu Ala Pro Val
1          5          10          15
Leu Gly Tyr Leu Gly Ser Arg Gln Lys His Ser Leu Pro Asp Leu Pro
20         25         30
Tyr Asp Tyr Gly Ala Leu Glu Pro His Ile Asn Ala Gln Ile Met Gln
35         40         45
Leu His His Ser Lys His His Ala Ala Tyr Val Asn Asn Leu Asn Val
50         55         60
Thr Glu Glu Lys Tyr Gln Glu Ala Leu Ala Lys Gly Asp Val Thr Ala
65         70         75         80
Gln Ile Ala Leu Gln Pro Ala Leu Lys Phe Asn Gly Gly Gly His Ile
85         90         95
Asn His Ser Ile Phe Trp Thr Asn Leu Ser Pro Asn Gly Gly Gly Glu
100        105        110
Pro Lys Gly Glu Leu Leu Glu Ala Ile Lys Arg Asp Phe Gly Ser Phe
115        120        125
Asp Lys Phe Lys Glu Lys Leu Thr Ala Ala Ser Val Gly Val Gln Gly
130        135        140
Ser Gly Trp Gly Trp Leu Gly Phe Asn Lys Glu Arg Gly His Leu Gln
145        150        155        160
Ile Ala Ala Cys Pro Asn Gln Asp Pro Leu Gln Gly Thr Thr Gly Leu
165        170        175
Ile Pro Leu Leu Gly Ile Asp Val Trp Glu His Ala Tyr Tyr Leu Gln
180        185        190
Tyr Lys Asn Val Arg Pro Asp Tyr Leu Lys Ala Ile Trp Asn Val Ile
195        200        205
Asn Trp Glu Asn Val Thr Glu Arg Tyr Met Ala Cys Lys Lys

```

-continued

210	215	220	
<210> SEQ ID NO 20 <211> LENGTH: 235 <212> TYPE: PRT <213> ORGANISM: Homo sapiens <400> SEQUENCE: 20			
Met Gly Ile Gly Lys Ser Lys Ile Asn Ser Cys Pro Leu Ser Leu Ser			
1	5	10	15
Trp Gly Lys Arg His Ser Val Asp Thr Ser Pro Gly Tyr His Glu Ser			
	20	25	30
Asp Ser Lys Lys Ser Glu Asp Leu Ser Leu Cys Asn Val Ala Glu His			
	35	40	45
Ser Asn Thr Thr Glu Gly Pro Thr Gly Lys Gln Glu Gly Ala Gln Ser			
	50	55	60
Val Glu Glu Met Phe Glu Glu Glu Ala Glu Glu Glu Val Phe Leu Lys			
	65	70	75
Phe Val Ile Leu His Ala Glu Asp Asp Thr Asp Glu Ala Leu Arg Val			
	85	90	95
Gln Asn Leu Leu Gln Asp Asp Phe Gly Ile Lys Pro Gly Ile Ile Phe			
	100	105	110
Ala Glu Met Pro Cys Gly Arg Gln His Leu Gln Asn Leu Asp Asp Ala			
	115	120	125
Val Asn Gly Ser Ala Trp Thr Ile Leu Leu Leu Thr Glu Asn Phe Leu			
	130	135	140
Arg Asp Thr Trp Cys Asn Phe Gln Phe Tyr Thr Ser Leu Met Asn Ser			
	145	150	155
Val Asn Arg Gln His Lys Tyr Asn Ser Val Ile Pro Met Arg Pro Leu			
	165	170	175
Asn Asn Pro Leu Pro Arg Glu Arg Thr Pro Phe Ala Leu Gln Thr Ile			
	180	185	190
Asn Ala Leu Glu Glu Glu Ser Arg Gly Phe Pro Thr Gln Val Glu Arg			
	195	200	205
Ile Phe Gln Glu Ser Val Tyr Lys Thr Gln Gln Thr Ile Trp Lys Glu			
	210	215	220
Thr Arg Asn Met Val Gln Arg Gln Phe Ile Ala			
	225	230	235
<210> SEQ ID NO 21 <211> LENGTH: 1670 <212> TYPE: DNA <213> ORGANISM: Homo sapiens <400> SEQUENCE: 21			
gtgtatctttt gcggggtggg cgccaacagc agtcaggcct gacaagcggc gacctccaag			60
ggtgaggcct ctgcgggccc ccgactcacg cgcgtccggg ctctgcaagg gcggtgggga			120
gcaggctgct gtggctcggg ggactgggtt gcggcgcgcc gcgtacggga cggcccaaaa			180
ctctcgacgc ccggggcaag acgcccacc cctgggcgct ctgctgggc cagaaaggaa			240
gacagaaaag ccgcgggctg actgtgtgg cgctcgcctg cagattgaaa agaattgctg			300
agaaatacat aaagttttcc tctctgct tggatattta taatgggtat cgggaagtct			360
aaaataaatt cctgcctct tctctctct tggggtaaaa ggcacagtgt ggatacaagt			420

-continued

```

ccaggatatac atgagtcaga ttccaagaag tctgaagatc tatccttggtg taatgttgct 480
gagcacagca atacaacaga ggggccaaaca ggaaagcagg agggagctca gagcgtggaa 540
gagatgtttg aagaagaagc tgaagaagag gtgttcctca aatttgatgat attgcatgca 600
gaagatgaca cagatgaagc cctcagagtc cagaatctgc tacaagatga ctttggtatc 660
aaacccggaa taatctttgc tgagatgcca tgtggcagac agcatttaca gaatttagat 720
gatgctgtaa atgggtctgc atggacaatc ttattactga ctgaaaactt ttaagagat 780
acttgggtgta atttccagtt ctatacgtcc ctaatgaact cggttaacag gcagcataaa 840
tacaactctg ttataccatc gcggccctg aacaatcccc tccccgaga aaggactccc 900
tttgccctcc aaaccatcaa tgccttagag gaagaaagtc gtggatttcc tacacaagta 960
gaaagaatth ttcaggagtc tgtgtataag acacaacaaa ctatatggaa agagacaaga 1020
aatatggtac aaagacaatt tattgcctga gatgaaacat ataacatggtg gctggctctt 1080
gttttgtaaa ccaaatgatt aatcttctact tgagaaagca gtttctagga aatgtttaa 1140
taaaagagag tcttcacctt aaagaaacct atggagcaca agaaagataa atttctgcag 1200
gacagtctat aaaattgtgg tactttttga tgtttcagta aacttgacat tgtcagagtt 1260
tcaaggactt ttctttcaca attttctag ttcattggata tgaaaaagga attctcaatc 1320
catattcctt gtattgaacc ttgaacaaaa acttgtatga cagacatttt taaaaatgtg 1380
acaacacttt tattctctga attttgatct caaaggacac agaaaaaaaa tggccccagg 1440
agatctgac acacttctc ctgaggcacc tctcatggat gttgcaataa gcattcgggt 1500
actatcacc agaaatatga attgccagaa tagaacattt agcatgttaa gcgttgatgc 1560
ataaaaaatc agaaatagat gtgagaatgg tggaactttt taaaagaacc cagtcaaatg 1620
tattttctgc tgaatctgc atatttggag gcatttccca ccaccgatcc 1670

```

```

<210> SEQ ID NO 22
<211> LENGTH: 926
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

```

<400> SEQUENCE: 22

```

Pro Ala Met Asp Val Leu Pro Thr Gly Gly Gly Arg Pro Gly Leu Arg
1          5          10          15
Thr Glu Leu Glu Phe Arg Gly Gly Gly Gly Glu Ala Arg Leu Glu Ser
20          25          30
Gln Glu Glu Glu Thr Ile Pro Ala Ala Pro Pro Ala Pro Arg Leu Arg
35          40          45
Gly Ala Ala Glu Arg Pro Arg Arg Ser Arg Asp Thr Trp Asp Gly Asp
50          55          60
Glu Asp Thr Glu Pro Gly Glu Ala Cys Gly Gly Arg Thr Ser Arg Thr
65          70          75          80
Ala Ser Leu Val Ser Gly Leu Leu Asn Glu Leu Tyr Ser Cys Thr Glu
85          90          95
Glu Glu Glu Ala Ala Gly Gly Gly Arg Gly Ala Glu Gly Arg Arg Arg
100         105         110
Arg Arg Asp Ser Leu Asp Ser Ser Thr Glu Ala Ser Gly Ser Asp Val
115         120         125
Val Leu Gly Gly Arg Ser Gly Ala Gly Asp Ser Arg Val Leu Gln Glu

```

-continued

130			135			140									
Leu	Gln	Glu	Arg	Pro	Ser	Gln	Arg	His	Gln	Met	Leu	Tyr	Leu	Arg	Gln
145					150					155					160
Lys	Asp	Ala	Asn	Glu	Leu	Lys	Thr	Ile	Leu	Arg	Glu	Leu	Lys	Tyr	Arg
			165						170					175	
Ile	Gly	Ile	Gln	Ser	Ala	Lys	Leu	Leu	Arg	His	Leu	Lys	Gln	Lys	Asp
			180					185						190	
Arg	Leu	Leu	His	Lys	Val	Gln	Arg	Asn	Cys	Asp	Ile	Val	Thr	Ala	Cys
	195							200					205		
Leu	Gln	Ala	Val	Ser	Gln	Lys	Arg	Arg	Val	Asp	Thr	Lys	Leu	Lys	Phe
	210					215						220			
Thr	Leu	Glu	Pro	Ser	Leu	Gly	Gln	Asn	Gly	Phe	Gln	Gln	Trp	Tyr	Asp
	225				230					235					240
Ala	Leu	Lys	Ala	Val	Ala	Arg	Leu	Ser	Thr	Gly	Ile	Pro	Lys	Glu	Trp
			245						250					255	
Arg	Arg	Lys	Val	Trp	Leu	Thr	Leu	Ala	Asp	His	Tyr	Leu	His	Ser	Ile
			260					265						270	
Ala	Ile	Asp	Trp	Asp	Lys	Thr	Met	Arg	Phe	Thr	Phe	Asn	Glu	Arg	Ser
		275					280					285			
Asn	Pro	Asp	Asp	Asp	Ser	Met	Gly	Ile	Gln	Ile	Val	Lys	Asp	Leu	His
	290					295						300			
Arg	Thr	Gly	Cys	Ser	Ser	Tyr	Cys	Gly	Gln	Glu	Ala	Glu	Gln	Asp	Arg
	305				310					315					320
Val	Val	Leu	Lys	Arg	Val	Leu	Leu	Ala	Tyr	Ala	Arg	Trp	Asn	Lys	Thr
				325					330					335	
Val	Gly	Tyr	Cys	Gln	Gly	Phe	Asn	Ile	Leu	Ala	Ala	Leu	Ile	Leu	Glu
			340				345						350		
Val	Met	Glu	Gly	Asn	Glu	Gly	Asp	Ala	Leu	Lys	Ile	Met	Ile	Tyr	Leu
		355					360					365			
Ile	Asp	Lys	Val	Leu	Pro	Glu	Ser	Tyr	Phe	Val	Asn	Asn	Leu	Arg	Ala
	370					375						380			
Leu	Ser	Val	Asp	Met	Ala	Val	Phe	Arg	Asp	Leu	Leu	Arg	Met	Lys	Leu
	385				390					395					400
Pro	Glu	Leu	Ser	Gln	His	Leu	Asp	Thr	Leu	Gln	Arg	Thr	Ala	Asn	Lys
			405						410					415	
Glu	Ser	Gly	Gly	Gly	Tyr	Glu	Pro	Pro	Leu	Thr	Asn	Val	Phe	Thr	Met
			420					425					430		
Gln	Trp	Phe	Leu	Thr	Leu	Phe	Ala	Thr	Cys	Leu	Pro	Asn	Gln	Thr	Val
		435					440					445			
Leu	Lys	Ile	Trp	Asp	Ser	Val	Phe	Phe	Glu	Gly	Ser	Glu	Ile	Ile	Leu
	450					455					460				
Arg	Val	Ser	Leu	Ala	Ile	Trp	Ala	Lys	Leu	Gly	Glu	Gln	Ile	Glu	Cys
	465				470					475					480
Cys	Glu	Thr	Ala	Asp	Glu	Phe	Tyr	Ser	Thr	Met	Gly	Arg	Leu	Thr	Gln
			485						490					495	
Glu	Met	Leu	Glu	Asn	Asp	Leu	Leu	Gln	Ser	His	Glu	Leu	Met	Gln	Thr
			500					505					510		
Val	Tyr	Ser	Met	Ala	Pro	Phe	Pro	Phe	Pro	Gln	Leu	Ala	Glu	Leu	Arg
		515					520					525			
Glu	Lys	Tyr	Thr	Tyr	Asn	Ile	Thr	Pro	Phe	Pro	Ala	Thr	Val	Lys	Pro
	530					535						540			

-continued

Thr Ser Val Ser Gly Arg His Ser Lys Ala Arg Asp Ser Asp Glu Glu  
 545 550 555 560  
 Asn Asp Pro Asp Asp Glu Asp Ala Val Val Asn Ala Val Gly Cys Leu  
 565 570 575  
 Gly Pro Phe Ser Gly Phe Leu Ala Pro Glu Leu Gln Lys Tyr Gln Lys  
 580 585 590  
 Gln Ile Lys Glu Pro Asn Glu Glu Gln Ser Leu Arg Ser Asn Asn Ile  
 595 600 605  
 Ala Glu Leu Ser Pro Gly Ala Ile Asn Ser Cys Arg Ser Glu Tyr His  
 610 615 620  
 Ala Ala Phe Asn Ser Met Met Met Glu Arg Met Thr Thr Asp Ile Asn  
 625 630 635 640  
 Ala Leu Lys Arg Gln Tyr Ser Arg Ile Lys Lys Lys Gln Gln Gln Gln  
 645 650 655  
 Val His Gln Val Tyr Ile Arg Ala Asp Lys Gly Pro Val Thr Ser Ile  
 660 665 670  
 Leu Pro Ser Gln Val Asn Ser Ser Pro Val Ile Asn His Leu Leu Leu  
 675 680 685  
 Gly Lys Lys Met Lys Met Thr Asn Arg Ala Ala Lys Asn Ala Val Ile  
 690 695 700  
 His Ile Pro Gly His Thr Gly Gly Lys Ile Ser Pro Val Pro Tyr Glu  
 705 710 715 720  
 Asp Leu Lys Thr Lys Leu Asn Ser Pro Trp Arg Thr His Ile Arg Val  
 725 730 735  
 His Lys Lys Asn Met Pro Arg Thr Lys Ser His Pro Gly Cys Gly Asp  
 740 745 750  
 Thr Ile Gly Leu Ile Asp Glu Gln Asn Glu Ala Ser Lys Thr Asn Gly  
 755 760 765  
 Leu Gly Ala Ala Glu Ala Phe Pro Ser Gly Cys Thr Ala Thr Ala Gly  
 770 775 780  
 Arg Glu Gly Ser Ser Pro Glu Gly Ser Thr Arg Arg Thr Ile Glu Gly  
 785 790 795 800  
 Gln Ser Pro Glu Pro Val Phe Gly Asp Ala Asp Val Asp Val Ser Ala  
 805 810 815  
 Val Gln Ala Lys Leu Gly Ala Leu Glu Leu Asn Gln Arg Asp Ala Ala  
 820 825 830  
 Ala Glu Thr Glu Leu Arg Val His Pro Pro Cys Gln Arg His Cys Pro  
 835 840 845  
 Glu Pro Pro Ser Ala Pro Glu Glu Asn Lys Ala Thr Ser Lys Ala Pro  
 850 855 860  
 Gln Gly Ser Asn Ser Lys Thr Pro Ile Phe Ser Pro Phe Pro Ser Val  
 865 870 875 880  
 Lys Pro Leu Arg Lys Ser Ala Thr Ala Arg Asn Leu Gly Leu Tyr Gly  
 885 890 895  
 Pro Thr Glu Arg Thr Pro Thr Val His Phe Pro Gln Met Ser Arg Ser  
 900 905 910  
 Phe Ser Lys Pro Gly Gly Gly Asn Ser Gly Thr Lys Lys Arg  
 915 920 925

&lt;210&gt; SEQ ID NO 23

&lt;211&gt; LENGTH: 93

-continued

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Homo sapiens

&lt;400&gt; SEQUENCE: 23

```

Met Ala Gly Asn Cys Gly Ala Arg Gly Ala Leu Ser Ala His Thr Leu
1           5           10           15
Leu Phe Asp Leu Pro Pro Ala Leu Leu Gly Glu Leu Cys Ala Val Leu
                20           25           30
Asp Ser Cys Asp Gly Ala Leu Gly Trp Arg Gly Leu Val Ser Lys Asn
        35           40           45
Lys Leu Cys Glu Asn Gly Val Leu Ile Leu Gln Gly Phe Trp His Val
        50           55           60
Val Val Cys Lys Tyr Lys Asp Ser Lys Tyr Gly Cys Val Ala Leu Phe
65           70           75           80
Glu Lys Thr Pro Ser Glu Thr Tyr Leu His Met Glu Thr
                85           90

```

&lt;210&gt; SEQ ID NO 24

&lt;211&gt; LENGTH: 5653

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Homo sapiens

&lt;400&gt; SEQUENCE: 24

```

gacaaagaga actaatgctt tgtgctgatt catatttgaa tcgaggcatt ggaaccctg      60
tatgccttgt ttgtggaaag aaccagtgac accatcactg agcttcctaa aagttcgaag      120
aagtttagagg actatacact ttcttttgaa cttttataat aaatatttgc tctggttttt      180
ggaaccagg gctgtagag gggtagtga caagtcttac aagtggcctt attccaactc      240
cagaaattgc ccaacggaac ttgagatta tatgcaatcg aaagtgcag gaaacatgcc      300
aactcaatcc ctcttaatgt acatggatgg gccagaagtg attggcagct ctcttggcag      360
tccgatggag atggaggatg ccttgtcaat gaaagggacc gctgtgttcc cattccgagc      420
tacacaagaa aaaaatgtca tccaatcga ggggtatag cccttggatt gcatgttctg      480
cagccagacc ttcacacatt cagaagacct taataaacat gtcttaatgc aacaccggcc      540
taccctctgt gaaccagcag ttcttcgggt tgaagcagag tatctcagtc cgttgataa      600
aagtcaagtg cgaacagaac ctccaagga aaagaattgc aaggaaaatg aatttagctg      660
tgaggatgt gggcagacat ttagagtcgc ttttgatgtt gagatccaca tgagaacaca      720
caaagattct ttcacttacg ggtgtaacat gtgcggaaga agattcaagg agccttggtt      780
tcttaaaaat cacatgctga cacataatgg caaatcgggg gccagaagca aactgcagca      840
aggcttgagg agtagtccag caacgatcaa cgaggtcgtc caggtgcacg cggccgagag      900
catctcctct ccttacaanaa tctgcatggt ttgtggcttc ctatttccaa ataagaaaag      960
tctaattgag caccgcaagg tgcacaccaa aaaaactgct ttcggtacca gcagcgcgca     1020
gacagactct ccacaaggag gaatgccgtc ctcgaggag gacttctctg agttgttcaa     1080
cttgagacca aaatctcacc ctgaaacggg gaagaagcct gtcagatgca tcctcagct     1140
cgatccgttc accaccttcc aggcttggca gctggctacc aaaggaaaag ttgccatttg     1200
ccaagaagtg aaggaatcgg ggcaagaagg gagcaccgac aacgacgatt cgagttccga     1260
gaaggagctt ggagaacaaa ataagggcag ttgtgcaggc ctctcgcaag agaaagagaa     1320
gtgcaaacac tcccacggcg aagcgccctc cgtggacgcg gatcccaagt taccagtag     1380

```

-continued

---

caaggagaag cccactcact gctccgagtg cggcaaagct ttcagaaect accaccagct	1440
ggtcttgac tccagggctc acaagaagga cgggagggcc ggcgcggagt cccccaccat	1500
gtctgtggac gggaggcagc gggggaectg ttctcctgac ctcgcccctc ctctggatga	1560
aaatggagcc gtggatcgag gggaaagtgg ttctgaagac ggatctgagg atgggcttcc	1620
cgaaggaatc catctggata aaaatgatga tggaggaaaa ataaaacatc ttacatcttc	1680
aagagagtgt agttattgtg gaaagttttt cegttcaaat tattacctca atattcatct	1740
cagaacgcat acaggtgaaa aaccatacaa atgtgaattt tgtgaatag ctgcagccca	1800
gaagacatct ctgaggtatc acttgagag acatcacaag gaaaaacaaa ccgatgttgc	1860
tgctgaagtc aagaacgatg gtaaaaatca ggacactgaa gatgcactat taaccgctga	1920
cagtgcgcaa accaaaaatt tgaaaagatt ttttgatggt gccaaagatg ttacaggcag	1980
tccacctgca aagcagctta aggagatgcc ttctgttttt cagaatgttc tgggcagcgc	2040
tgctcctca ccagcacaca aagatactca ggatttccat aaaaatgcag ctgatgacag	2100
tgctgataaa gtgaataaaa accctacccc tgcttacctg gacctgtaa aaaagagatc	2160
agcagttgaa actcaggcaa ataacctcat ctgtagaacc aaggcggatg ttactcctcc	2220
tccggatggc agtaccaccc ataacctga agttagcccc aaagagaagc aaacggagac	2280
cgcagctgac tgcagataca ggccaagtgt ggattgtcac gaaaaacctt taaatttatc	2340
cgtgggggct cttcacaatt gccccgcaat ttctttgagt aaaagttga ttccaagtat	2400
cacctgtoca ttttgaect tcaagacatt ttatccagaa gttttaatga tgcaccagag	2460
actggagcat aaatacaatc ctgacgttca taaaaactgt cgaaacaagt ccttgcttag	2520
aagtcgacgt accggatgcc gcgccagcgt gctgggaaaa gatgtgcctc cctctctag	2580
ttctgtaaa cccaagccca agtctgcttt cccggcgcag tccaaatccc tgccatctgc	2640
gaaggggaag cagagccctc ctgggccagg caaggcccct ctgacttcag ggatagactc	2700
tagcacttta gccccaaagta acctgaagtc ccacagacca cagcagaatg tgggggtcca	2760
aggggcccgc accaggcaac agcaatctga gatgttctct aaaaccagtg tttcccctgc	2820
accggataag acaaaaagac ccgagacaaa attgaaacct cttccagtag ctccttctca	2880
gcccaccctc ggcagcagta acatcaatgg ttccatcgac taccocgcca agaacgacag	2940
cccgtgggca cctccgggaa gagactatct ctgtaatcgg agtgccagca atactgcagc	3000
agaatttggg gagcccctc caaaaagact gaagtccagc gtggttgccc ttgacgttga	3060
ccagcccggg gcccaattaca gaagaggcta tgaccttccc aagtaccata tggtcagagg	3120
catcacatca ctgttaccgc aggactgtgt gtatccgtcg caggcgctgc ctcccaaac	3180
aaggttcctg agctccagcg aggtcgatc tccaaatgtg ctgactgttc agaagcccta	3240
tgggtgctcc gggccacttt acacttgtgt gctgctggt agtccagcat ccagctcgac	3300
gttagaagga aaaaggcctg tgtcatatca acacttatct aacagcatgg cacaaaagag	3360
aaactatgag aattttattg ggaatgcaca ttatcgacca aatgacaaaa aaacttgatt	3420
cactaattag ggggaaaaaa ggtcttggtg gatgtcagtg cttactcccc atgaaattaa	3480
atttacttc atcctttgag aagcgaatgg tgaagctac tgaataaagc tgtgattgta	3540
ctgtacataa aacatatgag gaatctgcaa ggaacactac agttgtgtaa agttgttctg	3600
ttaacttttg taccaaatag caatacaaac tagttggaac agttggaact tacatacatg	3660

-continued

---

```

gggactggaa atctctatatt tgtcocctgaa taatatTTTT cttagaattg accaaataag 3720
aagtgggaatt tttgcatact tgagcgcctgc tgaaaagaaa tcatttgggt tgggtggggc 3780
gggatggggg aaaagtatat aatgcttgca cctcaggtaa aaatctgtaa atatctaagt 3840
tgtaaacctg cttgttcaaa tactgtgtgt attccttttc tctaatacag ctcatactt 3900
ggagcagttt ctgctattgt gctctttcat ttaaatgta tgtttttttt tttttaact 3960
gtcaatgatt ttgtattatg ttgaatccac ccaaatctat tgttgcctta aaattgttaa 4020
tggaagtatt gacctctat gatatgtgct gcagatctcg aggtcagcca ttcggaagct 4080
ggcagcattt tatcgcaact ttgagcatct cagatgggga aggcaccttc ttcctcgcct 4140
ctccagattg tcttggaaac tccaggatcc ttgactgagg gcgttgggat ggctttagg 4200
atttttaaga gagtgtgtct acagacaagc attttctctg tagagcagcc acacgttgta 4260
taaacataaa ctgtatgtgc agttatttaa atttgtttct gtcaaattaa tcatttttgt 4320
tggacgattc agtggcgggg gggttttgcc taaattagta tataaaaaa aaaatgctaa 4380
attatatctg tgaattgcag gtattgggga acagttttaa gggaaatTTT ggggggaaca 4440
ttttaggttt gtatttggta gtcttaatgt atctggcatt tgggtgaact gtggacatac 4500
tagagttgat tatagacaca ttgattttga ataaggaact gctggccgag cccgctggga 4560
gtctagaaag agaaaatctg tttctagacc tcagttatTT tccattttt ggttgtttt 4620
aagcagtaac atttttctca gtgcacatgc aatttgggtt ttagagaaga tggccaccag 4680
ctggcttctc agatatttta aacttttgtt ctttaatatg ctgtocatgg ctgagtttat 4740
tagtacatgg gcttagtgac caaaaatat tttattaaga aactgtttca aaaataaatt 4800
tgcactgttc atttttctgg cctcgcctgt ctccatagag caagggtaat cctagaaaaa 4860
atttttttt tttaaattat gcaacgtaag atgtcctcct tgatagaagt cttagctcct 4920
gtgttacaag ggagaactca tttgagatca gtctgttggc attgcaatga agtgctttgt 4980
atcaggaaag tgtacactat tgacctttt tctgtttcac aagctgagcc atatgtacat 5040
aatctagatt ttgttttcat agttttgcac ttttatagcc tattttttaa gattaacaca 5100
tttgcaagat gattgactca atctttgctt aatccaatga gtgttacaga gagcttgcctg 5160
tgactagaac cataaatctt aaagggggtg tgtgataata gagggctgga atttaaacct 5220
gtatttaaaa aaaagaatca ccaaatctat ttgaaaaaa gtcgatttgt attatgctgg 5280
aattttttg gctttcagat ttctctttt aaccacattt ctgaatgtat aaaaatacca 5340
attattttcc tacagcctt tgtaactcaa aatatgtttt tgtgtccatc agtattaact 5400
attggtatac tactggtttt atattttttt ttctttgaga caacagtaca tataatagag 5460
gtacaattcg ttgattttt gtttatgtat ttatttcatt ccagtttgat ttattttaa 5520
tgttgatact taagttgtcg aacagtagac attacttgtt ttatttatga tatatttcag 5580
cttaaagtta tgttattata tgtggaagtg taaatataga tttgggtttt tgcaaaaaa 5640
aaaaaaaaa aaa 5653

```

&lt;210&gt; SEQ ID NO 25

&lt;211&gt; LENGTH: 1048

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Homo sapiens

&lt;400&gt; SEQUENCE: 25

-continued

---

Met Gln Ser Lys Val Thr Gly Asn Met Pro Thr Gln Ser Leu Leu Met  
 1 5 10 15  
 Tyr Met Asp Gly Pro Glu Val Ile Gly Ser Ser Leu Gly Ser Pro Met  
 20 25 30  
 Glu Met Glu Asp Ala Leu Ser Met Lys Gly Thr Ala Val Val Pro Phe  
 35 40 45  
 Arg Ala Thr Gln Glu Lys Asn Val Ile Gln Ile Glu Gly Tyr Met Pro  
 50 55 60  
 Leu Asp Cys Met Phe Cys Ser Gln Thr Phe Thr His Ser Glu Asp Leu  
 65 70 75 80  
 Asn Lys His Val Leu Met Gln His Arg Pro Thr Leu Cys Glu Pro Ala  
 85 90 95  
 Val Leu Arg Val Glu Ala Glu Tyr Leu Ser Pro Leu Asp Lys Ser Gln  
 100 105 110  
 Val Arg Thr Glu Pro Pro Lys Glu Lys Asn Cys Lys Glu Asn Glu Phe  
 115 120 125  
 Ser Cys Glu Val Cys Gly Gln Thr Phe Arg Val Ala Phe Asp Val Glu  
 130 135 140  
 Ile His Met Arg Thr His Lys Asp Ser Phe Thr Tyr Gly Cys Asn Met  
 145 150 155 160  
 Cys Gly Arg Arg Phe Lys Glu Pro Trp Phe Leu Lys Asn His Met Arg  
 165 170 175  
 Thr His Asn Gly Lys Ser Gly Ala Arg Ser Lys Leu Gln Gln Gly Leu  
 180 185 190  
 Glu Ser Ser Pro Ala Thr Ile Asn Glu Val Val Gln Val His Ala Ala  
 195 200 205  
 Glu Ser Ile Ser Ser Pro Tyr Lys Ile Cys Met Val Cys Gly Phe Leu  
 210 215 220  
 Phe Pro Asn Lys Glu Ser Leu Ile Glu His Arg Lys Val His Thr Lys  
 225 230 235 240  
 Lys Thr Ala Phe Gly Thr Ser Ser Ala Gln Thr Asp Ser Pro Gln Gly  
 245 250 255  
 Gly Met Pro Ser Ser Arg Glu Asp Phe Leu Gln Leu Phe Asn Leu Arg  
 260 265 270  
 Pro Lys Ser His Pro Glu Thr Gly Lys Lys Pro Val Arg Cys Ile Pro  
 275 280 285  
 Gln Leu Asp Pro Phe Thr Thr Phe Gln Ala Trp Gln Leu Ala Thr Lys  
 290 295 300  
 Gly Lys Val Ala Ile Cys Gln Glu Val Lys Glu Ser Gly Gln Glu Gly  
 305 310 315 320  
 Ser Thr Asp Asn Asp Asp Ser Ser Ser Glu Lys Glu Leu Gly Glu Thr  
 325 330 335  
 Asn Lys Gly Ser Cys Ala Gly Leu Ser Gln Glu Lys Glu Lys Cys Lys  
 340 345 350  
 His Ser His Gly Glu Ala Pro Ser Val Asp Ala Asp Pro Lys Leu Pro  
 355 360 365  
 Ser Ser Lys Glu Lys Pro Thr His Cys Ser Glu Cys Gly Lys Ala Phe  
 370 375 380  
 Arg Thr Tyr His Gln Leu Val Leu His Ser Arg Val His Lys Lys Asp  
 385 390 395 400

-continued

---

Arg	Arg	Ala	Gly	Ala	Glu	Ser	Pro	Thr	Met	Ser	Val	Asp	Gly	Arg	Gln
				405							410				415
Pro	Gly	Thr	Cys	Ser	Pro	Asp	Leu	Ala	Ala	Pro	Leu	Asp	Glu	Asn	Gly
			420					425					430		
Ala	Val	Asp	Arg	Gly	Glu	Gly	Gly	Ser	Glu	Asp	Gly	Ser	Glu	Asp	Gly
		435					440					445			
Leu	Pro	Glu	Gly	Ile	His	Leu	Asp	Lys	Asn	Asp	Asp	Gly	Gly	Lys	Ile
	450					455						460			
Lys	His	Leu	Thr	Ser	Ser	Arg	Glu	Cys	Ser	Tyr	Cys	Gly	Lys	Phe	Phe
465					470					475					480
Arg	Ser	Asn	Tyr	Tyr	Leu	Asn	Ile	His	Leu	Arg	Thr	His	Thr	Gly	Glu
			485						490					495	
Lys	Pro	Tyr	Lys	Cys	Glu	Phe	Cys	Glu	Tyr	Ala	Ala	Ala	Gln	Lys	Thr
			500					505					510		
Ser	Leu	Arg	Tyr	His	Leu	Glu	Arg	His	His	Lys	Glu	Lys	Gln	Thr	Asp
	515						520					525			
Val	Ala	Ala	Glu	Val	Lys	Asn	Asp	Gly	Lys	Asn	Gln	Asp	Thr	Glu	Asp
	530					535					540				
Ala	Leu	Leu	Thr	Ala	Asp	Ser	Ala	Gln	Thr	Lys	Asn	Leu	Lys	Arg	Phe
545					550					555					560
Phe	Asp	Gly	Ala	Lys	Asp	Val	Thr	Gly	Ser	Pro	Pro	Ala	Lys	Gln	Leu
			565					570					575		
Lys	Glu	Met	Pro	Ser	Val	Phe	Gln	Asn	Val	Leu	Gly	Ser	Ala	Val	Leu
		580						585					590		
Ser	Pro	Ala	His	Lys	Asp	Thr	Gln	Asp	Phe	His	Lys	Asn	Ala	Ala	Asp
	595						600					605			
Asp	Ser	Ala	Asp	Lys	Val	Asn	Lys	Asn	Pro	Thr	Pro	Ala	Tyr	Leu	Asp
610						615					620				
Leu	Leu	Lys	Lys	Arg	Ser	Ala	Val	Glu	Thr	Gln	Ala	Asn	Asn	Leu	Ile
625					630					635					640
Cys	Arg	Thr	Lys	Ala	Asp	Val	Thr	Pro	Pro	Pro	Asp	Gly	Ser	Thr	Thr
			645						650					655	
His	Asn	Leu	Glu	Val	Ser	Pro	Lys	Glu	Lys	Gln	Thr	Glu	Thr	Ala	Ala
		660						665					670		
Asp	Cys	Arg	Tyr	Arg	Pro	Ser	Val	Asp	Cys	His	Glu	Lys	Pro	Leu	Asn
		675					680					685			
Leu	Ser	Val	Gly	Ala	Leu	His	Asn	Cys	Pro	Ala	Ile	Ser	Leu	Ser	Lys
	690					695					700				
Ser	Leu	Ile	Pro	Ser	Ile	Thr	Cys	Pro	Phe	Cys	Thr	Phe	Lys	Thr	Phe
705					710					715					720
Tyr	Pro	Glu	Val	Leu	Met	Met	His	Gln	Arg	Leu	Glu	His	Lys	Tyr	Asn
			725						730					735	
Pro	Asp	Val	His	Lys	Asn	Cys	Arg	Asn	Lys	Ser	Leu	Leu	Arg	Ser	Arg
			740					745					750		
Arg	Thr	Gly	Cys	Pro	Pro	Ala	Leu	Leu	Gly	Lys	Asp	Val	Pro	Pro	Leu
		755					760					765			
Ser	Ser	Phe	Cys	Lys	Pro	Lys	Pro	Lys	Ser	Ala	Phe	Pro	Ala	Gln	Ser
	770					775					780				
Lys	Ser	Leu	Pro	Ser	Ala	Lys	Gly	Lys	Gln	Ser	Pro	Pro	Gly	Pro	Gly
785					790					795					800
Lys	Ala	Pro	Leu	Thr	Ser	Gly	Ile	Asp	Ser	Ser	Thr	Leu	Ala	Pro	Ser



**48.** An in vitro method of diagnosing a metabolic syndrome disorder, or a propensity thereto in a subject, said method comprising:

- a) obtaining an expression profile in a biological sample isolated from said subject, wherein said expression profile consists of the analysis of the level of IRAK3 expression or activity of an IRAK3 expression product in combination with the gene expression level or activity of a gene product of at least one gene selected from the group consisting of SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2, and ZNF217; and
- b) comparing said obtained expression profile to a reference expression profile to determine whether said sample is from a subject having a metabolic syndrome disorder phenotype or a propensity thereto.

**49.** An in vitro method for identifying or monitoring a metabolic syndrome disorder therapy in a subject, said method comprising analyzing the level of IRAK3 expression or activity of expression product in a sample isolated from said subject before and after treatment with said therapy, wherein an increased level of IRAK3 expression or activity compared to a sample of said subject before the therapy is indicative of the efficacy of said therapy.

**50.** An in vitro method for identifying or monitoring a metabolic syndrome disorder therapy in a subject, said method comprising analyzing the expression profile in a biological sample isolated from said subject before and after treatment with said therapy, wherein said expression profile consists of the analysis of the level of IRAK3 expression or activity of an IRAK3 expression product in combination with the gene expression level or activity of a gene product of at least one gene selected from the group consisting of SOD2, TNFAIP6, TNFAIP3, TLR2, IRS2, and ZNF217; whereby an increased level of ZNF217 or IRS2, a decreased level of SOD2, TNFAIP3, TNFAIP3, or TLR2 and in increased level of IRAK3 compared to a sample of said subject before the therapy is indicative for the efficacy of said therapy.

**51.** A diagnostic test kit for diagnosing a metabolic syndrome disorder phenotype in a subject, monitoring the effectiveness of therapy of a metabolic syndrome disorder in patients receiving such therapy, or predicting whether a subject is a candidate for metabolic syndrome disorder comprising:

- a) a predetermined amount of an antibody specific for IRAK3;
- b) a predetermined amount of a specific binding partner to said antibody;
- c) buffers and other reagents necessary for monitoring detection of antibody bound to IRAK3; wherein either said antibody or said specific binding partner are detectably labeled.

**52.** A diagnostic test kit for diagnosing a metabolic syndrome disorder phenotype in a subject, for monitoring the effectiveness of therapy of metabolic syndrome disorder in patients receiving such therapy, or predicting whether a subject is a candidate for metabolic syndrome disorder comprising:

- a) a predetermined amount of an antibody specific for ZNF217;
- b) a predetermined amount of a specific binding partner to said antibody;
- c) buffers and other reagents necessary for monitoring detection of antibody bound to ZNF217; wherein either said antibody or said specific binding partner are detectably labeled.

**53.** A diagnostic test kit for diagnosing a metabolic syndrome disorder phenotype in a subject, for monitoring the effectiveness of therapy of metabolic syndrome disorder in patients receiving such therapy, or for predicting whether a subject is a candidate for metabolic syndrome disorder comprising:

- a) one or more nucleic acids encoding one or more of the proteins selected from the group consisting of ZNF217 and IRAK3;
- b) reagents useful for monitoring the expression level of said one or more nucleic acids or proteins encoded by said nucleic acids; and
- c) instructions for use of the kit.

**54.** The test kit of claim 51 wherein the kit further comprises at least one nucleic acid encoding ZNF217 or IRAK3 and at least one nucleic acid encoding a gene selected from the group consisting of SOD2, TNFAIP6, TNFAIP3, TLR2, and IRS2.

\* \* \* \* \*

专利名称(译)	基因签名		
公开(公告)号	<a href="#">US20110053167A1</a>	公开(公告)日	2011-03-03
申请号	US12/922739	申请日	2009-04-02
[标]申请(专利权)人(译)	天主教鲁汶大学		
申请(专利权)人(译)	鲁汶大学		
当前申请(专利权)人(译)	鲁汶大学		
[标]发明人	HOLVOET PAUL BENHABILES NORA GEERAERT BENJAMINE		
发明人	HOLVOET, PAUL BENHABILES, NORA GEERAERT, BENJAMINE		
IPC分类号	C12Q1/68 G01N33/53		
CPC分类号	C12Q1/6883 G01N33/6893 G01N2333/912 C12Q2600/158 G01N2800/044 G01N2800/32 G01N2800/52 G01N2800/042		
优先权	2008009066 2008-05-19 GB 61/200705 2008-12-02 US 61/128205 2008-05-19 US 61/072971 2008-04-03 US		
外部链接	<a href="#">Espacenet</a> <a href="#">USPTO</a>		

摘要(译)

本发明一般涉及组织或组织的至少一个细胞中的新的相关分子簇，例如血液组织的细胞，优选这种骨髓细胞，并且识别基因表达的条件，所述相关分子或所述分子在用于鉴定肥胖的方法或系统中的表达水平以及与肥胖相关的代谢疾病（例如肥胖相关的心血管风险或肥胖相关的胰岛素抗性）的风险。该方法系统提供了关于如何调节相关分子以治疗或预防肥胖以及预防肥胖相关代谢疾病的信息。

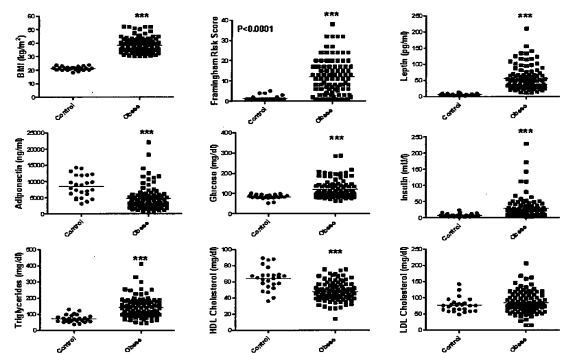


Figure 1