



US 20130189243A1

(19) United States

(12) Patent Application Publication

Barr et al.

(10) Pub. No.: US 2013/0189243 A1

(43) Pub. Date: Jul. 25, 2013

## (54) BIOMARKERS FOR ACUTE ISCHEMIC STROKE

(76) Inventors: **Taura L. Barr**, Waynesburg, PA (US);  
**Maria Del Mar Matarin**, Kent (GB);  
**Steven Jay Warach**, West Lake Hills,  
TX (US); **Andrew Barry Singleton**,  
Poolesville, MD (US); **Yvette P. Conley**,  
North Huntingdon, PA (US)

(21) Appl. No.: 13/580,571

(22) PCT Filed: Feb. 22, 2011

(86) PCT No.: PCT/US11/25748

§ 371 (c)(1),  
(2), (4) Date: Mar. 11, 2013**Related U.S. Application Data**

(60) Provisional application No. 61/307,233, filed on Feb.  
23, 2010.

**Publication Classification**

## (51) Int. Cl.

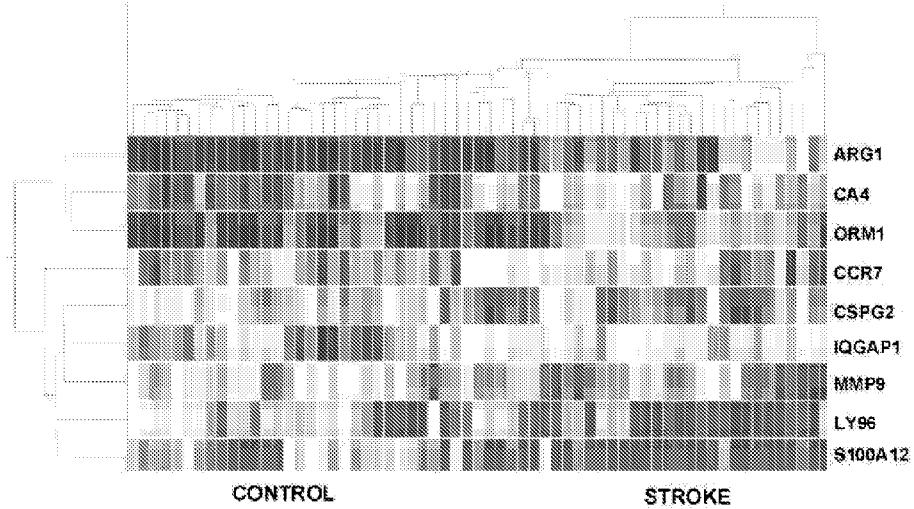
*C12Q 1/68* (2006.01)  
*A61B 6/00* (2006.01)  
*A61B 5/055* (2006.01)  
*G01N 33/68* (2006.01)  
*A61B 5/00* (2006.01)

## (52) U.S. Cl.

CPC ..... *C12Q 1/6876* (2013.01); *G01N 33/6893*  
(2013.01); *A61B 5/0042* (2013.01); *A61B  
5/055* (2013.01); *A61B 6/501* (2013.01)  
USPC ..... 424/94.64; 436/501; 435/7.92; 506/9;  
435/6.12; 435/6.11; 506/18; 506/16; 435/7.1;  
530/389.8; 536/24.31; 530/300; 600/407;  
600/410; 378/4

(57) **ABSTRACT**

The present invention provides methods and compositions for the diagnosis of acute ischemic stroke. The invention further provides methods and compositions for distinguishing acute ischemic stroke from other forms of stroke and TIAs and “stroke mimic” events. Moreover, methods and compositions are provided to facilitate the treatment of acute ischemic stroke patients.



<u>Gene</u>	<u>p-value</u>	<u>Fold change</u>	<u>Regulation</u>	<u>Description</u>
ARG1	2.84E-07	3.175	Up	Arginase-1
CA4	2.0E-04	2.122	Up	Carbonic anhydrase-4
CCR7	4.37E-05	2.094	Down	Chemokine CC Motif Receptor 7
CSPG2	3.45E-05	2.087	Up	Chondroitin sulfate proteoglycan 2
IQGAP1	7.97E-07	2.031	Up	IQ motif-containing GTPase-activating protein 1
LY96	0.001	2.159	Up	Lymphocyte antigen 96; MD2 protein
MMP9	1.11E-05	2.644	Up	Matrix metalloproteinase 9; gelatinase B
ORM1	0.006	2.246	Up	Orosomucoid 1; alpha 1 acid glycoprotein
S100A12	3.87E-04	2.354	Up	S100 calcium binding protein A12; calgranulin C

Figure 1

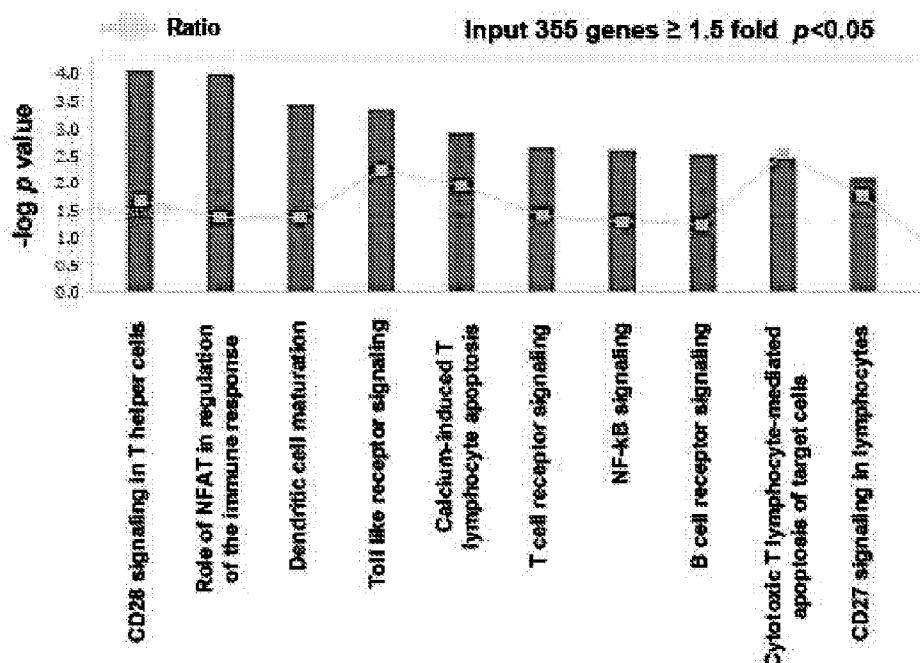


Figure 2

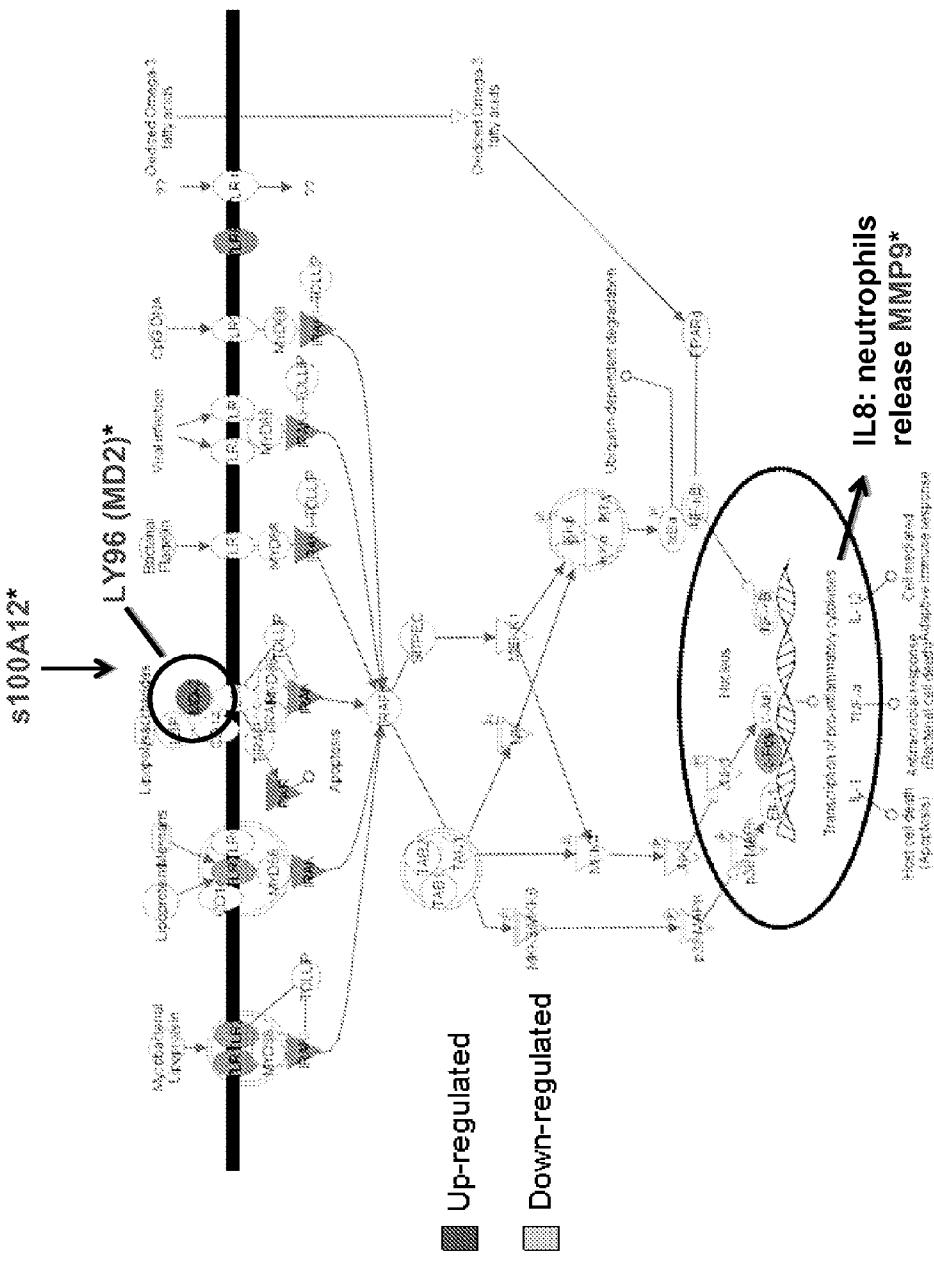


Figure 3

Homo sapiens chemokine (C-C motif) receptor 7 (CCR7), mRNA  
NCBI Reference Sequence: NM\_001838.2 (SEQ ID NO: 1)

Figure 4A

Chemokine (C-C motif) receptor 7 precursor [Homo sapiens]  
ACCESSION NP\_001829 (SEQ ID NO: 2)

```
1 MDLGKPMKSV IIVVALLVIFQ VCLCQDEVTD DYIGDNTTVD YTLFESLCSK KDVRNFKAWF
61 LPIMYSIICF VGLLGNGLVV LTYIYFKRLK TMTDTYLLNL AVADILFLLT LPFWAYSAAK
121 SWVFGVHFCK LIFAIYKMSF FSGMLLLICI SIDRYVAIVQ AVSAHRHRAR VLLISKLSCV
181 GIWILATVLS IPELLYSDLQ RSSSEQAMRC SLTEHVEAF ITIQVAQMVI GFLVPLLAMS
241 FCYLVIIRTL LQARNFERNK AIKVIIAVVV VFIVFQLPYN GVVLAQTVAN FNITSSTCEL
301 SKQLNIAYDV TYSLACVRCC VNPFLYAFIG VKFRNDLFKL FKDLGCLSQE QLRQWSSCRH
361 IRRSSSMSVEA ETTTFSP
```

Figure 4B

Homo sapiens versican (VCAN), mRNA; CSPG2  
NCBI Reference Sequence: NM\_004385.2 (SEQ ID NO: 3)

1 getccccga gccttctgg ggaagaactc caggcggtcg gacgcaacag ccgagaacat  
61 tagtgttgt ggacaggaggc tgggaccaag atcttcggcc agccccgcac cctcccgcat  
121 ctccagcac cgccccgcac cctccgcac cttcccccggg ccaccacgt tcctatgtga  
181 cccgcctggg caacgcggaa cccagtcgtcg cagcgctgc gtgaatttc cccccaact  
241 gcaataagcc gccttccaag gccaagatgt tcataaatat aaagagcata ttatggatgt  
301 gttcaaccct aatagtaacc catgcgtac ataaagtcaa agtggggaaaa agcccacccgg  
361 tgaggggctc cctctctggaa aaagtcaagc taccttgta ttttcaacg atgcctactt  
421 tgccacccag ttacaacacc agtgaatttc tccgcataa atggctcaag attgaagtgg  
481 acaaaaatgg aaaagatttgg aaagagacta ctgtccttgt ggcggaaaat ggaaatata  
541 agattggtca ggactacaaa gggagaggtgt ctgtccccac acatcccgag gctgtggcg  
601 atgcctccct cactgtggc aagctgtcg caagtgtgc gggctttac cgctgtgacg  
661 tcatgtacgg gattgaagac acacaagaca cgggtgtact gactgtggat ggggttgtgt  
721 ttcactacag ggccggcaacc agcaggta cactgaattt tgaggctgct cagaaggctt  
781 gtttggacgt tggggcagtc atagcaactc cagagcagct ctgtgtgcc tatgaagatg  
841 gatttgagca gtgtgacgc ggctggctgg ctgatcagac tgtcagatata cccatccggg  
901 ctccccagatg aggctgttat ggagataaga tgggaaaggc aggagtcagg acttatggat  
961 tccgttctcc ccaggaaact tacgatgtt attgttagt ggatcatctg gatgtgtatg  
1021 tggccacct cactgtcccc agtaaattca cttcgagga ggctgcaaaa gagtgtgaaa  
1081 accaggatgc caggctggca acagtggggg aactccaggc ggcattgggg aacggcttg  
1141 accagtgcga ttacgggtgg ctgtcgatg ccagcggtcg ccaccctgtg actgtggcca  
1201 gggcccagtg tggaggggtg ctacttgggg tgagaaccct gtatcggtt gagaaccaga  
1261 caggettcccc tccccctgtat agcagatttgc atgcctactg cttaaacctt aaagaggctt  
1321 caaccatcga tttagtatac ctcgcagaaa ctgcattcacc cagtttatcc aaagaaccac  
1381 aaatggtttgc ttagtataactt acaccaatca tcccttagt tgatgaatta cctgtcatc  
1441 caacagagtt ccctcccggtt gggaaatatttgc tcaatggggg acagaaagcc acagtccaa  
1501 ctcaaggctat cacagatgtt ttagccacca aattacccac acctactggc agtaccaaga  
1561 agccctggga tatggatgac tactcacccctt ctgttccagg acctttggaa aagcttagaca  
1621 tatcagaaat taaggaagaa gtgtccaga gtacaactgg cgctctcat tatgtctacgg  
1681 attcatggga tgggtcggtt gaaatggaaac aaacacaaga atcgggttaca cagattgaac  
1741 aaatagaagt gggccctttg gtaacatcta tggaaatctt aaaggcacatt cttcccaagg  
1801 aattccctgt aactgaaaca ccattggtaa ctgcaagaat gatccctggaa tccaaaactg  
1861 aaaagaaaat ggtaaagact gtttctgaat tggtaaccac aggtcactat ggattcacct  
1921 tgggagaaga ggtatgttggaa gacagaacac ttacagttgg atctgtatgag agcacccatg  
1981 tctttgacca aattccctgtt gtcattacgg tgcataagac ttcaagagac accatccaca  
2041 ctcatcttgc aagacttgggg tcaatgttcacatccacaac tggccatccat ttaattatgc  
2101 ctgataataa tggatcatcc atggatgact gggaaagagag acaaactgtt ggtggataaa  
2161 cggaaagagtt ctgtggcaaa tatctgtcta ctacaccctt tccatcacag catcgatc  
2221 aaatagaattt gtttcccttgc tctgggtataa aaatattatgtt agggaaatttccacatgtt  
2281 ttatcccttc tctacaaaca gaaatgacac atagaagaga aagaacagaa acactaata  
2341 cagagatgag aacagatact tatacagatg aaatacaaga agagatcact aaaagtccat  
2401 ttatgggaaa aacagaagaa gaagtcttctt ctggatgaa actctctaca tctctctcag

Figure 5A

2461 agccaaatca tggtacagag tcctctgtgg aaatgaccaa gtcctttgat ttcccaacat  
2521 tgataacaaa gttaaatgcg gagccaacag aagtaagaga tatggaggaa gactttacag  
2581 caactccagg tactacaaaa tatgtatgaaa atattacaac agtgctttg gcccatggta  
2641 cttaagtgt tgaagcagcc actgttatca aatggtcatg ggatgaagat aatacaacat  
2701 ccaagccctt agagtctaca gaacccctcg cctcttcaaa attggccct gccttactca  
2761 caactgtggg gatgaatgga aaggataaag acatccaaag ttcaactgaa gatggagcag  
2821 atgaatttac tcttaattcca gatagfactc aaaagcgatg agaggagggt actgtatgaaag  
2881 acatagcagc ccattggaaaa ttccacaatta gatttcagcc aactacatca actgttattg  
2941 cagaaaaagtgc aactttgaga gattctacaa ctgaaagaaaa agttcacct atcacaagca  
3001 ctgaaggcca agtttatgca accatggaag gaagtgcattt gggtgaagta gaagatgtgg  
3061 accctctaa gccagtatct actgttcccc aatttcacca cacttcagag gtggaaaggat  
3121 tagcatttgt tagttatagt agcaccctaa agcctactac ttatgttagac tcttcccata  
3181 ccattcccttc ttctgttaatt cccaaagacag actggggagt gtttagtacct tctgttccat  
3241 cagaagatga agttcttaggt gaaccctctc aagacataact tgcattgtat cagactcgcc  
3301 ttgaagcgcac tattttctca gaaactatga gaacaacaaa aatcacagag ggaacaactc  
3361 aggaagaatt ccattggaaa gaacagactg cagagaaacc agttctgtct ctcagttct  
3421 cagcttggac tcccaaggag gcagtaacac cactgatga acaagaggc gatggatcag  
3481 catatacagt ctctgaagat gaattttgtga cagggtctga gagggtccca gttttagaaaa  
3541 caactccagt tggaaaaatt gtcacatgt tgcattatcc accagggtgct gtaactgagc  
3601 acaaagtgaa aacagatgaa gtggtaacac taacaccac cattggccca aaagtatctt  
3661 taagtccagg gcctgaacaa aaatatgaaa cagaaggttag tagtacaaca ggatttacat  
3721 catctttagt cccttttagt accccacatta cccagctt ggaagaaacc actactgaga  
3781 aaacatccct agaggatatt gatttaggtc caggattatt tggaaagccc aaagccacag  
3841 aactcataga atttcaaca atcaaagtca cagttcaag tgatattacc actgccttca  
3901 gttcagtaga cagacttcac acaacttcag cattcaagcc atcttcccgatc atactaaga  
3961 aaccacccct catcgacagg gaacctgggt aagaaacaac cagtgcacatg gtaatcttgc  
4021 gagaatcaac atctcatgtt cctccacta cccttgaaga tatttttagcc aagggaaacag  
4081 aaaccgatata tgcatacggat tatttcacgca cttcaagttc tctgttacca cagccaaacaa  
4141 gaccacccac tggaaagac aaagaggcct tggacccatca ggcgccttct acgccacagc  
4201 ccccaagcaag cacaaaattt caccctgaca ttaatgttta tatttttagt gtcagagaaaa  
4261 ataagacagg tgcataatgtt gattttgttgc taatggcata tccaaatagat tcagaatcta  
4321 aagaagatga accttgtatgt gaagaaacag atccatgtca tgatctaatg gctgaaattt  
4381 tacctgaatt ccctgacata attgaaatag acctataccca cagtgaagaa aatgaagaag  
4441 aagaagaaga gtgtgcataat gctactgtc tgacaaaccac cccatctgtc cagtagat  
4501 atgggaagca ttcgttacc actgtgcacca aggaccaga agetgcagaa getaggcgtg  
4561 gccagtttgc aagtgttgc ctttcgttgc attttctcgaa cagctgttgc agetgtatc  
4621 atccattttgtt aatagccaaa acggaattttt ctactgttgc gcaacccataat gaatctacag  
4681 aaacaactgatgc tgcattttgc gttacatggc agcctgatgc ttacccttgc acatcagaac  
4741 atttttcagg tgggtttcc ccacagtccc attccatgttgc gaattttggaaa  
4801 gtggaaacgc caaaaaagggg gcagaatcag tcacagagag agatactgaa gttggcattc  
4861 aggcacatgatgc acatactgaa cctgttatc tgcattttgc agatgtttca ggagagattt  
4921 ccattttgc aagaatctcg aaaaatgcctt tggcaaggcc tacagaatgttgc acatttgggt  
4981 aagaggtaga aaaaagtact tctgttgc acatccac tataatgttca agttctgtat  
5041 cagcatatgttccatggaa gaagcaggta ccctaatagg aaatcccttgg ccagatgacc

5101 tgggtctac caaagaaagc tggtagaaag caactccat acaagttgt aagctctcag  
5161 ggaggcttc gattccaatt acagaaggct ctggagaagc agaagaagat gaagatacaa  
5221 tggtcaccat gttaactgat ttacacaga gaaataactc tgataacactt attacttag  
5281 acactagcag gataatcaca gaaagcttt ttgaggtcc tgcaaccacc atttatccag  
5341 ttctgaaca accttgc aaaaatggc ctaccaagtt tgaagtgaa acagacactt  
5401 ctgagtggtt ttccagtacc actgttggg aaaagaaaag gaaggaggag gagggaacta  
5461 caggtacggc ttctacattt gaggtatatt catctacaca gagatcggtt caattttttt  
5521 tacccttga attagaaagt ccaaatgttag ctacatctag tgattcagg accaggaaaa  
5581 gtttatgtc cttgacaaca ccaacacagt ctgaaaggaa aatgacagat tctactcctg  
5641 tcttacaga aacaaataca ttagaaaatt tggggcaca gaccactgag cacagcagta  
5701 tccatcaacc tgggttcag gaagggctga ccactctcc acgtgtcct gcctctgtct  
5761 ttatggagca gggctctgaa gaagctgtc cggaccaga aaccaccact gtttcttcat  
5821 ttccattttt ctagatgtt gcaattcaag ccgaaaagga agtagctggc acttttgtc  
5881 cgcattgttga aactacattc tccactgagc caacaggact ggtttttagt acagtaatgg  
5941 acagagtagt tgcgtttttt ataaaccctt catccaggaa aatagtgtt tcagagcgat  
6001 taggagaacc aaattatggg gcagaaataa ggggttttc cacagttttt cctttggagg  
6061 aagatttcag tggtagactt agagaataact caacagtgtc tcatccata gcaaaagaag  
6121 aaacggtaat gatggaaaggc tctggagatg cagcatttag ggacacccag acttcaccat  
6181 ctacagtacc tacttcgtt cacatcgtc acatatctga ctcagaagga cccagtagca  
6241 ccatggtcag cacttcagcc ttcccctggg aagagtttac atcctcagct gagggttcag  
6301 gtgagcaact ggtcacatgc agcagcttg ttgttcagg tgcctccagg gctgtgc  
6361 agttttctgg tacagttcc tccattatcg acgaaggatt gggagaagtg ggtactgtca  
6421 atgaaatttga tagaagatcc accattttac caacagcaga agtggaaaggat acgaaaggct  
6481 cagtagagaa ggaggaagta aaggtagtgc gcacagtcc aacaactt cccaaacta  
6541 tagagccagc caaattatgg tctaggcaag aagtcaaccc tgtaagacaa gaaattgaaa  
6601 gtgaaacaac atcagagggaa caaattcaag aagaaaagtc atttgaatcc cctcaaaact  
6661 ctccgtcaac agaacaaca atctttgatt cacagacatt tactgaaact gaactcaaaa  
6721 ccacagatta ttctgtacta acaacaaaga aaacttacag tgatgataaa gaaatgaagg  
6781 aggaagacac ttcttttagtt aacatgtcta ctccagatcc agatgcaaat ggettgaaat  
6841 ctacacaac tctccctgaa gtcactgaaa agtcacattt ttcttagt actgcattag  
6901 taactgaatc tataccatgtt gaacatgttag tcacagattt accaatcaaa aaggaagaaa  
6961 gtacaaaaca ttccggaaa ggcattgagac caacaattca agagtcagat actgagctt  
7021 tattctctgg actggatca ggagaagaag ttatccat tctaccaaca ggttcact  
7081 attttactga agtggaaacaa atcaataaca cattatatec ccacacttca caagtggaaa  
7141 gtacctcaag tgacaaaattt gaagacttta acagaatggaa aatgtggca aaagaagtt  
7201 gaccacttgt atctcaaca gacatcttgg aaggttagtgg gtcagtaacc agcacaacat  
7261 taatagaaat ttaagtgc actggagcag aaggaccac ggtggcacct ctcccttct  
7321 ccacggacat cggacatcc caaaatcaga ctgtcagggt ggcagaagaa atccagacta  
7381 gtagaccaca aaccataact gaacaagact ctaacaagaa ttctcaaca gcaaaaat  
7441 acgaaacaac aacccatct actgattttc tggctagac ttatggttt gaaatggca  
7501 aagaattttt tacatcagca cccaaaccat ctgactgttta ttatgaacct tctggagaag  
7561 gatctggaga agtggatattt gttgattcat ttcacacttc tgcaactact caggcaacca  
7621 gacaagaaag cagcaccaca ttgttctg atgggtccct ggaaaaacat cctgaggtgc  
7681 caagcgctaa agtggatgttact gtcgttactt tcccaacactt ttctgtatg ctgccttc

7741 attcagagca gaacaaaagc tcccgtatc caactagcac actgtcaat acagtgtcat  
7801 atgagaggc cacagacggt agttccaag accgttcag ggaattcgag gattccacct  
7861 taaaacctaa cagaaaaaaa cccactgaaa atattatcat agacctggac aaagaggaca  
7921 aggatttaat attgacaatt acagagaga ccatcctga aattctacct gagctgacat  
7981 cgataaaaa tactatcata gatattgatc atactaaacc tgtgtatgaa gacattctg  
8041 gaatgcaaac agatatacat acagaggtac catcagaacc acatgacagt aatgatgaaa  
8101 gtaatgatga cagcactcaa gtcaagaga tctatgaggc agctgtcaac ctltcttaa  
8161 ctgagggaaac atttgagggc tctgtgtatc ttctggctag ctacactcg gcaacacatg  
8221 atgaatcaat gacttatgaa gatagaagcc aactagatca catgggctt cacttcacaa  
8281 ctggatccc tgctcttagc acagaaacag aattagacgt ttacttccc acggcaacat  
8341 ccctgccaat tcctcgtaag tctgccacag ttattccaga gattgaagga ataaaagctg  
8401 aagcaaaagc cctggatgac atgttgaat caagcactt gtctgttgtt caagcttattg  
8461 cagaccaaag taaaataata ccaacattgg gccaatttga aaggactcg gaggagttatg  
8521 aagacaaaaa acatgctggt cttctttc agccagaatt ctcttcagga gctgaggagg  
8581 cattagtaga ccatactccc tatctaagta ttgctactac ccacctttagt gatcagatgt  
8641 taacagaggt gcctgtatgt atgaaaggat ccaatcccc atattacact gatacaacat  
8701 tagcagttc aacatttgcg aagtgtctt ctcagacacc atcatctccc ctcactatct  
8761 actcaggcag tgaagcctt ggacacacag agatccccca gcccagtgtctt ctggcaggaa  
8821 tagacgtcg ctcatctgtatgttccac aggattttt taagggaaatt catgtaaata  
8881 ttgaagcaac ttccaaacca tcaagtggg aataccctca cataactggg cccctctt  
8941 tatctcttgc cacaattaa gaacccctcg aagatgtatgg taaaccttagg ttattagaag  
9001 aaatggaaatc ttctccacaa gaacttattt ctgtggagg aactgagatt ctccaagatt  
9061 tccaaaacaa aaccgttgtt caagttctg gagaagcaat caagatgtt cccaccattt  
9121 aaacacctga ggctggaaact gttattacaa ctgcccgtatc aattttttttaaattttt  
9181 cacagtggcc acactctact tctgttgtt ccacctatgg ggtcgaggca ggtgtgggt  
9241 ctggctaaatccacatggact tctgttgttccacatggacttccatggacttccatggactt  
9301 ctgaaactca agcagcttta atcagaggc aggattccac gatagcagca tcagaacag  
9361 aagtggcage gagaattttt gattccatg atcaggcaac agtaaaccct gtggatttt  
9421 atactgaggt tgcaacacca ccattttcccttccatggacttccatggacttccatggactt  
9481 tgattggcat taatgaagag tcagtggaaat gcacggcaat ctatccatc ggacctgtatc  
9541 gctgcaaat gaaccctgtgc cttaaacggg gcacccgtt ccactgtatc aacttccatc  
9601 tatgcacccgt tttccatgg tacagcgggg accagtgtatc acttggattt gatgtatc  
9661 actctaatcccttgcgtatc ggagccactt gtgttgtatgg ttttttttttttttttttttt  
9721 tctgttgttccatggacttccatggacttccatggacttccatggacttccatggacttccatgg  
9781 gctggcacaat ttccaaaggc cagttgtatc aataacttgc ccacccgtatc acatgggat  
9841 cagctgttgtt ggttgtatgg ttt  
9901 aacaaatgtt tgtaatgtt gtggggccatg attatcgtatc gataggccctc aatgacaaga  
9961 tggttgatgt tgacttccatggacttccatggacttccatggacttccatggacttccatgg  
10021 ccaaccagcc agacagcttccatggacttccatggacttccatggacttccatggacttccatgg  
10081 agaatggcca gtggaaatgtt gttccctgtatc attaccatc caccatatacg tgcaagaaag  
10141 gaacagttgc ttgcggccatg cccctgtatc tagaaaaatgc caagaccctt ggaaagatgt  
10201 aacatcgatgttcaac tccctgtatc gataccactg caaagatgtt ttcattcaac  
10261 gtcacccatccatggacttccatggacttccatggacttccatggacttccatggacttccatgg

Figure 5A (con't)

10321 gcatgaaccc atctgcatac caaaggactt attctatgaa atacttaaa aattcctcat  
10381 cagcaagga caattcaata aatacatcca aacatgtca tcgtggagc cgagggtggc  
10441 aggagtctcgag gcgcgtatcc ctaaaatggc gaacatgtgt ttcatcatt tcagccaaag  
10501 tcctaacttc ctgtgcctt cctatcacct cgagaagttt atcagttt gtttggattt  
10561 ttggaccacc gttcagtcat ttgggttgc cgtgctccca aaacattta aatgaaagta  
10621 ttggcattca aaaagacagc agacaaaatg aaagaaaatg agagcagaaa gtaagcattt  
10681 ccagcctatc taatttctt agtttctat ttgcctccag tgcagtccat ttccataatgt  
10741 ataccagcct actgtactat taaaatgct caatttcagc accgatggcc atgtaaataa  
10801 gatgattaa tggatttt aatcctgtat ataaaataaa aagtcacaat gagtttggc  
10861 atatttaatg atgattatgg agccttagag gtcttaatc attgggtcgg ctgcctttat  
10921 gtagtttagg ctggaaatgg tttcaactgc tctttgactg tcagcaagac tgaagatggc  
10981 ttccctgga cagctagaaa acacaaaatc ttgtaggtca ttgcacctat ctcagccata  
11041 ggtgcagttt gcttctacat gatgctaaag gctgcgaatg ggatccgtat ggaactaagg  
11101 actccaatgt cgaacttcc ttgctgcat tcctttct tcacttacaa gaaaggcctg  
11161 aatggaggac ttctgtaa ccagg

Human CSPG2 amino acid sequence GenPept Accession No. NP\_004376 (SEQ ID NO: 4)

```

1 LHKVKGKSP PVRGSLSGKV SLPCHFSTMP TLPPSYNTSE FLRIKWSKIE VDKNGKDLKE
61 TTVLVAQNGN IKIGQDYKGR VSVPTHEPAV GDASLTVVKL LASDAGLYRC DVMYGIEDTQ
121 DTVSLTVGDV VFHYRAATSR YTLMFEAAQK ACLDVGAIVIA TPEQLFAAYE DGFEQCDAGW
181 LADQTVRYPI RAPRVCYGD KMKGAGVRTY GFRSPQETYD VYCYVDHLDG DVFHLTVPSK
241 FTFFEEAKAC ENQDARLATV GELQAAWRNG FDQCDYGWLS DASVRHPVTV ARAQCGGGLL
301 GVRTLYRFEN QTGFPPPDSDR FDAYCFKPKE ATTIDLSILA ETASPLSKE PQMVSDRITP
361 IIPLVDELPV IPTEFPPVGN IVSFEQKATV QPOAITDSLA TKLPTPTGST KKPWDMDDYS
421 PSASGPLGKL DISEIKEEVL QSTTGVSHYA TDSWDGVVED KQTQESVTQI EQIEVGPLVT
481 SMEILKHIPS KEFPVTETPL VTARMILESK TEKKMVSTVS ELVTTGHYGF TLGEEDDEDR
541 TLTVGSDEST LIFDQIPEVI TVSKTSEDTI HTHLEDLESV SASTTVSPLI MPDNNGSSMD
601 DWEERQTSGR ITEEFLGKYL STTPFPSQHR TEIELFPYSG DKILVEGIST VIYPSLQTEM
661 THRRERTETL IPEMERTDTYT DEIQQEITKS PFMGKTEEEV FSGMLSTSL SEPIHVTESS
721 VEMTKSFDFP TLITKLSAEP TEVRDMEEDF TATPGTTKYD ENITTULLAH GTLSVEAATV
781 SKWSWDDEDNT TSKPLESTEP SASSKLPPAL LTTWGMNGKD KDIPSFTEDG ADEFTLIEDS
841 TQKQLEEVTD EDIAAHGKFT IRFQPTSTG IAEKSTLRDS TTEEKVPPIT STEGQVYATM
901 EGGSALGEVED VDLSPKVSTV PQFAHTSEVE GLAFVSYSSST QEPPTTYVDSS HTIPLSVIPK
961 TDWGVLVPSV PSEDEVLPGE SQDILVIDQT RLEATISPET MRTIKITEGT TQEEFPWKEQ
1021 TAEKPVPALS STAWTPKEAV TPLDEQEGDG SAYTUSDEDEL LTGSERVPVL ETTPVGKIDH
1081 SVSYPPGAVT EHVKVKTDEVV TLTPRIGPKV SLSPGPEQKY ETEGSSTTGF TSSLSPFSTH
1141 ITQLMEETTT EKTSLEIDL GSGLFEKPKA TELIEFSTIK VTVPSDITTA FSSVDRLHTT
1201 SAFKPSSAIT KKPLLIDREP GEETTSMDVI IGEESTSHVPP TTLEDIVAKE TETDIDREYF
1261 TSSSPPATQP TRPPTVEDKE AFGPQALSTP QPPASTKFHP DINVYIIIEVR ENKTGRMSDL
1321 SVIGHPIDSE SKEDEPSEE TDPVHDLMAE ILPEFPDIIE IDLYHSEENE EEEEECANAT
1381 DVTTTPSVQY INGKHLTTV PKDPEAAEAR RGQFESVAPS QNFSDSSESQD THPFVIAKTE
1441 LSTAVQPNES TETTESLEVW WKPEPYTPEW EHFSGGEPEPV FPTVVFHEEF ESGTAKKAE
1501 SVTERDTEVG HQAHEHTEPV SLFPEESSGE IAIDQESQKI AFARATEVTF GEEVEKSTSV
1561 TYTPTIVPSS ASAYVSEEEA VTLLIGNWPW DLLSTKESWV EATPRQVVEL SGSSSIPIYE
1621 GSGEAEEDDED TMFTMVTDSL QRNTTDLIT LDTSRIITES FFEVPAATTIY PVSEQPSAKV
1681 VPTKFVSETD TSEWISSTV EKKRKREEEG TTGTASTFEV YSSTQRSDQL ILPFELESPN
1741 VATSSDSGTR KSFMSLTTPT QSEREMTDST PVFTETNTLE NLGAQTTTEHS SIHQPGVQEG
1801 LTTLPRSPAS VFMEQGSGEA AADPETTVS SFSLNVEYAI QAEKEVAGTL SPHVETTFST
1861 EPTGLVLSTV MDRVVAENIT QTSREIVISE RLGEPNYGAE IRGFSTGFP EEDFSGDFRE
1921 YSTVSHPIAK EETVMMEGSG DAAFRDQTQ TSPTVPTSVHI SHISDSEGPS STMVSTSAFP
1981 WEEFTSSAEG SGEQLVTVSS SVVPVLPSAV QKFSTGASSI IDEGLGEVGT VNEIDRRSTI
2041 LPTEVEVGTK APVEKEEVKV SGTNSTNFPQ TIEPAKLWSR QEVNPVRQEI ESETTSEBQI
2101 QEEKSFESPO NSPATEQTIF DSQFTETEL KTTDYSVLTT KKTYSDDKEM KEEDTSLVNM
2161 STPDPDANGL ESYTTLPEAT EKSHFFLATA LVTESIPAEH VVTDSPIKKE ESTKHFVKG
2221 RPTIQESDTE LLFSLGLSGE EVLPTLPTES VNFTEVEQIN NTLYPHTSQV ESTSSDKIED
2281 FNRMENVAKE VGPLVSQTDI FEQSGSVTST TLEIILSDTG AEGPTVAPLP FSTDIGHQON
2341 QTVRWAEEIQ TSRPQTTIEQ DSNKNNSTAE INETTSSTD FLARAYGFEM AKEFVTSAPK
2401 PSDLYYEPSG EGSGEVDIVD SFHTSATQD TRQESSTTFV SDGSLEKHPE VPSAKAVTAD
2461 GFPTVSVMLP LHSEQNKKSSP DPTSTLSNTV SYERSTDGSF QDRREFEDS TLKPNRKPT
2521 ENIIDLDKE DKDLILTITE STILEILPEL TSDKNTIIDI DHTKPVYEDI LGMQTDIDTE
2581 VPSEPHDSND ESNDDSTQVQ EIYEAAVNLS LTEETFEGSA DVLASYTQAT HDESMTYEDR
2641 SQLDHMGFHF TTGIPAPSTE TELDVLPTA TSLPIPRKSA TVIPIELEGK AEAALKDDMF
2701 ESSTLSDGQA IADQSEIIIPQ LGQFERTQEE YEDKKHAGPS FQPEFSSGAE EALVDHTPYL
2761 SIATTHLMDQ SVTEVPDVME GSNPPYYTDT TLAVSTFAKL SSQTPSSPLT IYSGSEASGH
2821 TEIPQPSALP GIDVGSSVMS PQDSFKEIHV NIEATFKPSS EYELHITEPP SLSPTDKLEP
2881 SEDDGKPELL EEMEASPTEL IAVEGTEILQ DFQNKTGDQV SGEAIKMFP TIKTPEAGTVI
2941 TTADEIELEG ATQWPHSTSA SATYGVEAGV VPWLSPQTSE RPTLSSSPEI NPETQAALIR

```

Figure 5B

3001 GQDSTIAASE QQVAARILDS NDQATVNPVE FNTEVATPPF SLLETSNETD FLIGINEESV  
3061 EGTAIYLPGP DRCKMNPCLN GGTCYPTETS YVCTCVPGYS GDQCeldorf CHSNPCRNGA  
3121 TCVDGFNTFR CLCLPSYVGA LCEQDTECD YGWHKFQGQC YKYFAHRTW DAAERECLQ  
3181 GAHLTSILSH EEQMFVNVRVG HDYQWIGLND KMFEHDFRWT DGSTLQYENW RPNQPDSFFS  
3241 AGEDCVIIW HENGQWNDVP CNYHLYTCK KGTVACGQPP VVENAKTFGK MKPRYEINSL  
3301 IRYHCKDGFI QRHLPTIRCL GNGRWAIPKI TCMNPSAYQR TYSMKYFKNS SSAKDONSINT  
3361 SKHDHRWSRR WQESRR

Homo sapiens IQ motif containing GTPase activating protein 1 (IQGAP1), mRNA  
NCBI Reference Sequence: NM\_003870.3 (SEQ ID NO: 5)

1 ggacccggc aagccgcgc acttggcagg agctgttaget accgcccgtcc ggcgcctcaa  
61 ggttcacgg cttcctcagc agagactcggt gctcgccgc catgtccgc gcagacgagg  
121 ttgacgggct gggcggtggcc cgccccact atggctctgt cctggataat gaaagactta  
181 ctgcagagga gatggatgaa aggagacgtc agaacgtggc ttatgagttac ctttgtcatt  
241 tggagaaggc gaagagggtgg atggaagcat gcctaggggaa agatctgcct cccaccacag  
301 aactggagga gggcggttagg aatggggct accttgccaa actggggaaac ttcttccttc  
361 cccaaagtgt gtccctgaaa aaaatctatg atcgagaaca gaccagatac aagggcgactg  
421 gcctccactt tagacacact gataatgtt aatcggtt gaatgccatg gatgagattt  
481 gattgcctaa gattttttac ccagaaacta cagatatcta tgatcgaaag aacatgcctaa  
541 gatgtatcta ctgtatccat gcactcgatgt tgcctgtt caagctaggc ctggcccttc  
601 agattcaaga cctatatggaa aagggttact tcacagaaga agaaatcaac aacatgaaga  
661 ctgagtttggaa gaagtatggc atccagatgc ctgccttttag caagatttggg ggcattttgg  
721 ctaatgaact gtcagttggal gaagccgtat tacatgtgc tttttttttt attaatgaag  
781 ctattgaccg tagaatcca ggcgacacat ttgcagttt gaaaaatccg aatgcctatgc  
841 ttgttaatct tgaagagccc ttggcatcca ttaccaggataatactttac caggctaaagc  
901 aggacaaaat gacaaatgtc aaaaacagggaa cagaaaactc agagagagaa agagatgtt  
961 atgaggagct gtcacgcgaa gtcgaaattc aaggcaatataa aacaaatgc aatacattt  
1021 ctgcatttgc aaatatcgac ctgcatttttag aacaaggaga tgcacigggc ttgttcaggg  
1081 ctctgcagtc accagccctg ggcgttcggag gactgcagca acagaatgc gactggact  
1141 tgaaggcgtt ctcgtgtat aacacaggcaga agagacagag tggcagact gaccccttc  
1201 agaaggagga gtcgcgttggatggatggatgc ctgcacaaacag tgctgcccag caatatcaga  
1261 gaagatttggc agcgtatgc ctgtttaatgc ctgcacatcca gaagggttgg gctgagaaga  
1321 ctgttttggaa actgtatgtatccaa ggcgacacat gtcgttatcca ttggccggcc  
1381 atctctatca gaaggagctg gtcaccccttc acgcacaaag tccgtacat aatctcacc  
1441 acccgagct ctctgtcgca gtggagatgt tgcacatggt ggcctgtatc aacaggccat  
1501 tggaatcaggat agatgtgaat acagtgtggaa agcaatttgc cagttcgtt actggctt  
1561 ccaatattgtt ggaagaaaac tgcacggatgtt atctcgatgtt gttgtatggaa ctgaaaggc  
1621 aggcacatgc agagaataat gaattcatta catggatgtatcaatgc tgcgtggacc  
1681 atgtgaacct ggtgggtggaa gaggaacatgc agaggattttt agccattttt gttttttt  
1741 aagccctggaa tgaagggtgtat gcccggccaa ctctgcaggc ctttacatgcattt ctttgc  
1801 aacttgagggg agtcccttgca gaagtggccc agcattacca agacacgcgtt attagagc  
1861 agagagagaa agcccaggaa atccaggatgtt agtcacgtt gttatgggtt gatgtttt  
1921 aagggtggaaat ctggcgttccaa aacaaagaca cccaaaggc acagaatgtt gctttaggg  
1981 tctttggccat taatgaggca gtagaaatgtt gttgtgtttt gttttttttt gttttttt  
2041 gtcacccctgtt tttttttttt gttttttttt tttttttttt tttttttttt tttttttttt  
2101 atcttgcgttgc acgcacaaatggt gttttttttt tttttttttt tttttttttt tttttttttt  
2161 agcactgggtt aaaaagggtggaa tttttttttt tttttttttt tttttttttt tttttttttt  
2221 gggatgttgcacc tccaaatttt gttttttttt tttttttttt tttttttttt tttttttttt  
2281 gttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt  
2341 gtcgttgcgttgc acgcacaaatggt gttttttttt tttttttttt tttttttttt tttttttttt  
2401 ccaggatgttgc accggatgttgc gttttttttt tttttttttt tttttttttt tttttttttt

Figure 6A

2461 gaggatacaa gcagaagaag gcatatcaag atcggttagc ttacctgcgc tcccacaaag  
2521 atgaagtgt aaagattcg tcacctgcaa ggtgcacca agtcgaaag cgctatcgag  
2581 atgcctgca gtactccgg gaccatataa atgacattat caaaatccag gctttatc  
2641 gggcaaacaag agctcgggat gactacaaga ctctcatca tgctgaggat cctctatgg  
2701 ttgtggccg aaaattgtc cacctgctgg accaaagtga ccaggattt caggaggagc  
2761 ttgaccitat gaagatcgga gaagaggta tcaccctcat tcgtltaac cagcagctgg  
2821 agaatgacct caatctcatg gatatcaaaa ttggactgct agtgaaaaat aagattacgt  
2881 tgcaggatgt ggccccac agtaaaaaac ttacaaaaaa aaataaggaa cagttgtctg  
2941 atatgatgat gataataaa cagaagggag gtctcaaggc tttgagcaag gagaagagag  
3001 agaagttgga agcttaccag cacctgtttt atttattgca aaccaatccc acctatctgg  
3061 ccaagctcat tttcagatg cccagaaca agtccaccaa gttcatggac tctgtaatct  
3121 tcacactcta caactacgcg tccaaccagc gagaggagta cctgctctg cggtcttta  
3181 agacagcact ccaagaggaa atcaagtcga aggtatca gattcaagag atttgacag  
3241 gaaatctac ggttattaaa atgggtgtaa gttcaaccg tggtccccgt gcccagaatg  
3301 ccctgagaca gatctggcc ccagtcgtga aggaaattat ggtgacaaa tctctaaca  
3361 tcaaaaactga ccctgtggat atttacaaat ctgggttaa tcagatggag tctcagacag  
3421 gagaggcaag caaactgccc tatgtgtga cccctgagca ggccgttagt catgaagaag  
3481 tgaagacacg gtagacagc tccatcagga acatgcgggc tgcacagac aagttctct  
3541 cagccattgt cagctctgtg gacaaaatcc ctatggat ggcgttcatt gccaaagtgc  
3601 tgaaggactc gttcatgag aagttccctg atgctggta ggtgagctg ctgaagatta  
3661 ttggtaactt gctttattat cgatacatga atccagccat ttggctctt gatgccttgc  
3721 acatcattga cctgtcagca ggaggccagc ttaccacaga ccaacgccga aatctggct  
3781 ccattgcaaa aatgttcag catgtctgtt ccaataagat gttctggga gataatgccc  
3841 acttaagcat cattaatgaa tatctttccc agtcctacca gaaattcaga cggttttcc  
3901 aaactgttg ttagtccca gagttcagg ataattaa tggatgatg tactctgatt  
3961 tagtaaccct caccaaaacca gtaatctaca ttccattgg tggaaatcatc aacacccaca  
4021 ctctcctgtt ggatcaccag gatgccattg ctccggagca caatgtatcca atccacgaac  
4081 tgctggacga cctcggcgag gtgcacca tcgagtcct gatagggaa agtctggca  
4141 atttaatga cccaaataag gaggactgtt ctaagacgga agtgtcttc accctgacca  
4201 acaagttcga cgtcctgga gatgagaatg cagaatgtt gtcgtcaacc atcttactga  
4261 atacaaaacg ttaattgtt gatgtcatcc gttccagcc aggagagacc ttgactgaaa  
4321 tcctagaaac accagccacc agtgaacagg aagcagaaca tcagagagcc atgcagagac  
4381 gtgttatccg ttagtccaaa acacctgaca agatgaaaaa gtcaaaatct gtaaaggaaag  
4441 acagcaaccc cactctcaa gagaagaaag agaagatcca gacaggatca aagaagctaa  
4501 cagagttgg aaccgtggac cccaaagaaca aataccagga actgtatcaac gacattgcca  
4561 gggatattcg gaatcagcg aggtaccgac agaggagaaa ggccgaacta gtgaaactgc  
4621 aacagacata cgctgctcg aactctaagg ccaccttta tggggagcag gtggattact  
4681 ataaaagcta tatcaaaaacc tgcttgata acttagccag caagggcaaa gtctccaaaa  
4741 agcctaggaa aatgaaagga aagaaaaagca aaaagatttctgaaatatac acagcagcaa  
4801 gactacatga aaaaggatgtt ctctggaaa ttgaggaccc gcaagtgaat cagttaaaa  
4861 atgttatatt tggaaatcatgtt ccaacagaag aagttggaga ctgcgttgcgaaat  
4921 tcatggagt tcaaattggag actttatgt tacattatca ggacctgtc cagctacagt  
4981 atgaaggatgt tgcgtcatg aaattttttagt agtagctaa agttaatgtc aacctctgaa  
5041 tcttccttcaacaaaatgatcggaa agttaattgtatcgttgc cagccagaa

5101 ggatgaagga aagaaggcacc tcacagctcc ttcttaggtc cttcttcct catttggaaagc  
5161 aaagacctag ccaacaacag cacctcaatc tgatacac tc ccgatgccac atttttaact  
5221 cctctcgctc tgatgggaca ttgttaccc tttttcata gtgaaattgt gtttcaggct  
5281 tagtctgacc ttctgggtt ctgcatttc ttccattact taggaaagag tggaaactcc  
5341 actaaaattt ctctgtgtt acatgttcc agagggtgc gtaactatatt gtaagcttg  
5401 gtgttgtt aatttgcata agggatggta ggattcaat gtgtgtcatt tagaagtgg  
5461 agctttagc accaatgaca taaatacata caagacacac aactaaaatg tcatgttatt  
5521 aacagttttt aggttgcatt taaaaataa agttcctta tatttctgtc ccacaggaa  
5581 aactgaagga tatgggaaat cattgggtt ctccattgt gttttctt atggacagga  
5641 gctaattggaa gtgacagtca tggtaaagg aagcatttc agaaaaaagg agataatgt  
5701 tttaaatttc attatcaaac ttggcaatt ctgttgtt aactccccga ctatggatg  
5761 ggagagtccc attgctaaaa ttcatgtact cagataaattt cagaatgggt caaggcac  
5821 gcctgtttt ttgggtgcac agagattgac ttgattcaga gagacaattc actccatccc  
5881 tatggcagag gaatgggta gccctaattt agaatgtcat tgttttaaa actgttttat  
5941 atcttaagag tgccttata aagtatagat gtatgttta aatgtgggt gataggaatt  
6001 ttaaagattt atataatgca tcaaaagcct tagaataaga aaagctttt tttaatttgc  
6061 ttatctgtt atctgaactc ttgaaacttta tagctaaaac actaggattt atctgcagtg  
6121 ttcagggaga taattctgcc tttaatttgc taaaacaaaaa acaaaaaccag ccaacctatg  
6181 ttacacgtga gatttttacc aatttttcc ccatttttc tcctttttc ttctgtgc  
6241 cacattgtgc cttttattta tgagccccag ttctgggc ttgtttaaa aaaaaaaatca  
6301 agtctaaaca ttgcatttag aagcttttgc ttcttgata aaaagtata cactttaaa  
6361 aaaaaaaaaa cttttccag gaaaatata tggaaatcatg ctgctgagcc tctattttct  
6421 ttcttgatg ttgttattca gtattttt atcataaattt tttagcattt aaaaattc  
6481 tggatgtacat taagccaata aactgttta atgaataaca aactatgttag tggatgttca  
6541 ttataatgc attggagaag tattttatg agactctta ctcaggtgc tggttacagc  
6601 ccacaggggag gcatggagtg ccatggaggatttgcact acccagaccc tggatgttca  
6661 tggatgttca aagacagggtt ttttttttttccat cagattttttt gatgtatgt  
6721 ttacaacttag cattgcctca aaaactggga ccaacccaaag tggatgttcaacc ctgttttcc  
6781 aaaagaggct atgaatccca aaggccacat ccaagacagg caataatgttag cagatgttac  
6841 agtccctta ataaaaatgtg tggatgttta taagggtt atgtccctca acacaatgt  
6901 taatgcagaa tagtgcataa tggatgttcaaa gatgttgc gatgtatgttag  
6961 gcttttagtag cacagaggat gccccaaactcatggcg ttgaaaccac acagttctca  
7021 ttactgttta ttattagctg tagcatttc tggatgttcaacc ttgaccc  
7081 cctcgaccag ccatcatgac atttaccatg aatttacttc ctcccaagag ttggactgc  
7141 ccgtcagatt gttgctgcac atagttgcct tggatgttcaacc acagttctca  
7201 tggatgttca aaaaaaaaaa

Figure 6A (con't)

NCBI Reference Sequence: NP\_003861.1

IQ motif containing GTPase activating protein 1 [Homo sapiens] (SEQ ID NO: 6)

1 MSAADEVDGL GVARPHYGSV LDNERLTAEE MDERRRQNVA YEYLCHLEEA KRWMEACLGE  
61 DLPPPTTELEE GLRNGVYIYAK LGNFFSPKVV SLKKIYDREQ TRYKATGLHF RHTDNVIQWL  
121 NAMDEIGLPK IFYPETTDIY DRKNMPRCIY CIHALSLYLF KLGLAPQIQL YGKVDFTEE  
181 EINNMKTELE KYGIQMPAFS KIGGILANEL SVDEAALHAA VIAINEAIDR RIPADTFAAL  
241 KNPNAMLVNL EEPPLASTYQD ILYQAKQDKM TNAKNRTENS ERERDVYEEL LTQAEIQGNI  
301 NKVNNTFSALA NIDLALEQGD ALALFRALQS PALGLRGLQQ QNSDWYLKQL LSDKQQKRQS  
361 GQTDPLQKEE IQLSGVDAANS AAQQYQRRLA AVALINAAIQ KGVAEKTVLE LMNPEAQLPQ  
421 VYPFAAADLYQ KELATLQRQS PEHNLTPEL SVALEMSSV ALINRALESQ DVNTVWKQLS  
481 SSVTGLTNIE EENCQRYLDE LMKLKAQAHQ ENNEFITWND IQACVDHVNL VVQEEHERIL  
541 AIGLINEALD EGDAQKTLOA LOPIPAAKLEG VLAEVQAHYQ DTLIRAKREK AQEIQDESAV  
601 LWLDEIQGGI WQSNKDTQEA QKFALGIFAI NEAVESGDVG KTLSALRSPD VGLYGVIPPEC  
661 GETYHSDILAE AKKKKLAVGD NNSKWVKHWV KGGYYYYYHNL ETQEGGWDEP PNFTVQNSMQL  
721 SREEIQSSIS GVTAAYNREQ LWLANEGLIT RLQARCRGYL VRQEFRSRMN FLKKQIPIAIT  
781 CIQSQWRGYK QKKAYQDRLA YLRSHKDEVV KIQSLARMHQ ARKRYRDRRLQ YFRDHINDII  
841 KIQAFIRANK ARDDYKTLIN AEDPPMVVVR KFVHLLDQSD QDFQEELDLM KMREEVITLI  
901 RSNQQLENDL NLMDIKIGLL VKNKITLQDV VSHSKKLTKK NKEQLSDMM INKQKGGLKA  
961 LSKEKREKLE AYQHLFYLLQ TNPTYLAKLI FQMPQNKSTK FMDSVIFTLY NYASNQREY  
1021 LLLRLFKTAL QEEIKSKVDQ IQEIVTGNPT VIKMVVSFNR GARGQNALRQ ILAPVVKEIM  
1081 DDKSLNIKTD PVDIYKSWN QMESQTGEAS KLPYDVTPEQ ALAHEEVKTR LDSSIRNMRA  
1141 VTDKFLSAIV SSVDKIPYGM RFIAKVLKDS LHEKFPDAGE DELLKIIGNL LYRYMNP  
1201 VAPDAFDIID LSAGGQLTTD QRRNLGSIAK MLQHAASNKM FLGDNAHLSI INEYLSQSYQ  
1261 KFRRFFQTAC DVPELQDKFN VDEYSDLVTL TKPVIYISIG EIINTHTLLL DHQDAIAPEH  
1321 NDPIHELLDD LGEVPPTIESL IGESSIONLND PNEAKALAKTE VSLTLNKFD VPGDENAE  
1381 ARTILLNTKR LIVDVIRFQP GETLTEILET PATSEQEAEH QRAMQRAIR DAKTPDKMKK  
1441 SKSVKEDSNL TLQEKKEKIQ TGLKKLTELQ TVDPKNKYQE LINDIARDIR NQRRYRQRRK  
1501 AELVKLQQTY AALNSKATFY GEQVDYYKSY IKTCLDNLAS KGKVSKPRE MKGKSKKIS  
1561 LKYTAARLHE KGVLLEIEDL QVNQFKNVIF EISPTEEVGQ FEVKAKFMGV QMETFMLHYQ  
1621 DLLQLQYEGV AVMKLFDRAK VNVNLLIFLL NKKFYGK

Figure 6B

NCBI Reference Sequence: NM\_000607.2  
Homo sapiens orosomucoid 1 (ORM1), mRNA (SEQ ID NO: 7)

```
1 acagagtaaa ctttgctgg gctccaagtg accgcccata gtttattata aagggtgactg
 61 caccctgcag ccaccagcac tgcctggctc caacgtgcctc ctgggttcag tatggcgctg
121 tcctgggttc ttacagtccct gagcctcccta cctctgctgg aagcccagat cccattgtgt
181 gccaaccttag taccgggtgcc catcaccaac gccaccctgg accggatcac tggcaagtgg
241 ttttatatcg catcggcctt tcgaaacgag gagtacaata agtcggttca ggagatccaa
301 gcaaccttct ttacttac ccccaacaag acagaggaca cgatcttct cagagagtac
361 cagacccgac aggaccagtg catctataaac accacacctacc tgaatgtcca gcgggaaaaat
421 gggaccatct ccagatacgt gggaggccaa gagcatttcg ctcacttgct gatccctcagg
481 gacaccaaga cctacatgct tgcttttgcgt gtaacatcg agaagaactg ggggctgtct
541 gtctatgctg acaagccaga gacgaccaag gagcaactgg gagagtctca cgaagctctc
601 gactgcttgc gcattcccaa gtcagatgtc gtgtacacccg attggaaaaaa ggataagtgt
661 gagccactgg agaagcagca cgagaaggag agggaaacagg aggagggggaa atcttagcag
721 gacacagcct tggatcgaga cagagacttg ggggccccatcc tgccctcca acccgacatg
781 tgtaccttag cttttccct cacttcgtcatc aataaaagctt ctgtgtttgg aacagctaaa
841 aaaaaaaaa
```

Figure 7A

NCBI Reference Sequence: NP\_000598.2  
orosomucoid 1 precursor [Homo sapiens] (SEQ ID NO: 8)

```
1 M ALSWVLTVL SLLPLLEAQI PLCANLVPVP ITNATLDRIT GKWFYIASAF RNEEYNKSVQ
61 EIQATFFYFT PNKTEDITFL REYQTRQDQC IYNTTYLNVQ RENGTLISRYV GGQEHLFAHLL
121 ILRDTKTYML AFDVNDEKNW GLSVYADKPE TTKEQLGEFY EALDCLRIPK SDVYYTDWKK
181 DKCEPLEKQH EKERKQEEGE S
```

Figure 7B

Homo sapiens arginase, liver (ARG1), mRNA  
NCBI Reference Sequence: NM\_000045.2 (SEQ ID NO: 9)

1 tgtcactgag ggttactgta ctggagatc caagtgcagc aaagagaagt gtcagagcat  
61 gagcgccaag tccagaacca tagggattat tggagctct ttctcaaagg gacagccacg  
121 aggaggggtg gaagaaggcc ctacagtatt gagaaggct ggtctgcttg agaaaacttaa  
181 agaacaagag tgtgtatgtga aggattatgg ggacctgccccc ttgtctgaca tccctaata  
241 cagtcctttt caaatgtga agaatccaag gtctgtggaa aaagcaagcg agcagctggc  
301 tggcaagggtg gcagaaggta agaagaacgg aagaatcgc ctggctgg gcggagacca  
361 cagttggca atttggca tctctggca tgccagggtc caccctgtate ttggagtc  
421 ctgggtggat gtcacactg atatcaacac tccactgaca accacaatgt gaaacttgca  
481 tggacaacct gtatcttc tcttgaagga actaaaaggaa aagattcccg atgtgccagg  
541 attctccctgg gtgactccct gtatctgc caaggatatt gtgtatattt gcttgagaga  
601 cgtggaccct ggggaacact acatttgaa aactcttaggc attaaataact ttcaatgac  
661 tgaagtggac agacttagaa ttggcaagggt gatggaaagaa acactcagct atctactagg  
721 aagaaagaaa agccaattt atctaagttt tgatgttgac ggactggacc catcttcac  
781 accagctact ggcacaccag tcgtgggagg tctgacatac agagaaggc tctacatcac  
841 agaagaaatc tacaacacag ggctactctc aggatttagat ataatggaa tgaacccatc  
901 cctggggaaag acaccagaag aagtaactcg aacagtgaac acagcagttg caataacctt  
961 ggcttgttgc ggacttgc tcacaaggctt attgactacc ttaacccacc  
1021 taagtaatggaaacatc cgatataat ctcatagtt atggcataat tagaaagcta  
1081 atcattttct taagcataga gttatccctc taaagacttg ttcttcaga aaaatgtttt  
1141 tccaatttagt ataaactcta caaattccctt ctgggtgaa aattcaagat gtggaaattc  
1201 taactttttt gaaattttaa agcttattt ttcataacttg gcaaaagact tatccttaga  
1261 aagagaagtg tacatttgatt tccaattttaa aatttgctgg cattaaaaat aagcacactt  
1321 acataagccc ccatacatag agtggactc ttggaaatcg gagacaaagc taccacatgt  
1381 ggaaaggtagtac tatgtgtcca tgcattcaa aaaatgttat tttttataat aaactcttta  
1441 taacaag

//

NCBI Reference Sequence: NP\_000036.2  
arginase 1 [Homo sapiens] (SEQ ID NO: 10)

```
1 MSAKSRTIGI IGAPFSKGQP RGGVEEGPTV LRKAGLLEKL KEQECDVKDY GDLPFADIPN
61 DSPFQIVKNP RSVGKASEQL AGKVAEVKKN GRISLVLGGD HSLAIGSISG HARVHPDLGV
121 IWVDAHTDIN TPLTTTSGNL HGQPVSFLLK ELKGKIPDVP GFSWVTPCIS AKDIVYIGLR
181 DVDPGEHYIL KTLGIKYFSM TEVDRLGIGK VMEETLSYLL GRKKRPIHLS FDVDGLDBSF
241 TPATGTPVVG GLTYREGLYI TEEIYKTGLL SGLDIMEVNP SLGKTPEEVT RTVNNTAVAIT
301 LACFGLAREG NHKPIDYLNP PK
```

Figure 8B

Homo sapiens lymphocyte antigen 96 (LY96), mRNA  
NCBI Reference Sequence: NM\_015364.3 (SEQ ID NO: 11)

```
1 agttaaatct tttctgctta ctgaaaagga agagtctgat gattagttac tgatcctctt
 61 tgcatttgta aagctttgga gatattgaat catgttacca tttctgtttt tttccaccct
121 gttttcttcc atatttactg aagtcagaa gcagtattgg gtctgcaact catccgatgc
181 aagtatttca tacacctact gtgataaaat gcaataccca atttcaatta atgttaaccc
241 ctgtatagaa ttgaaaagat ccaaaggatt attgcacatt ttctacattc caaggagaga
301 tttaaagcaa ttatattca atctctatat aactgtcaac accatgaatc ttccaaagcg
361 caaagaagtt atttgccag gatctgatga cgattactt ttttcagag ctctgaaggg
421 agagactgtg aataacaaca tatcattctc cttcaaggga ataaaatttt ctaagggaaa
481 atacaaatgt gtttgtgaag ctattctgg gagcccagaa gaaatgctct tttgcttgga
541 gtttgtcatc ctacaccaac ctaattcaaa tttagataaa ttgagtattt aaaaaaaaaaa
601 aaaaaaaaaa aaaaaaaaaa
```

//

Figure 9A

Homo sapiens matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase) (MMP9), mRNA (NCBI Reference Sequence: NM\_004994.2) (SEQ ID N0: 12)

1 agacacctct gcccacca tgagcctctg gcagccccctg gtccctgggtgc tcctgggtct  
61 gggcgtcgtc tttgtgtccc ccagacagcg ccagtccacc ctttgtcttc tcctggaga  
121 cctgagaacc aatctcaccc acaggcagct ggcagaggaa tacctgtacc gctatggta  
181 cactcggttg gcagagatgc gtggagagtc gaaatctctg gggcctgcgc tgctgcttct  
241 ccagaagcaa ctgtccctgc cegagaccgg tgagctggat agcgcacacg tgaaggccat  
301 gcgaaaccca cggtgcgggg tcccaagacct gggcagattc caaaccttgc agggcgaccc  
361 caagtggcac caccacaaca tcaccaattt gatccaaaac tactcggaag acttgcgcgc  
421 ggcgggttattt gacgacgcct ttgcgcgcgc ctgcgcactg tggagcgcgg tgacgcgcct  
481 cacccactact cgcgtgtaca gcccggacgc agacatcgatc atccagtttgc tggtgcgg  
541 gcacggagac gggtatccct tegacggaa ggacgggcct ctggcacacgc ccttcctcc  
601 tggccccggc attcaggag acgcccattt cgacgatgac gagttgtgtt ccctggca  
661 gggcgctgtt gttccaaactc ggtttggaaa cgcagatggc gcccgcctgcc acttccccctt  
721 catctcgag ggcgcgttactctgcgtt caccaccgc ggtcgctccg acggcttgcc  
781 ctggtgcaactt accacggcca actacgacac cgacgaccgg ttggcttctt gccccagcga  
841 gagactctac acccaggacg gcaatgttgc tggaaaccc tgccagtttcatcatctt  
901 ccaaggccaa tctactccg ctcgcaccac ggacgggtgc tccgacggctt accgctgg  
961 cgccaccacc gccaactacg accgggacaa ctgcgttccgc ttctgcgcga cccgagctga  
1021 ctgcacgggtt atggggggca actcggcggg ggagctgttgc gtctcccttc tcaacttct  
1081 ggtaaggag tactgacccctt gtaccagcga gggccgcggat gatggggccctc tctgg  
1141 taccacccg aacttttgcata gogacaagaa gtggggcttc tgccggacc aaggatacag  
1201 ttgttccttc gtggcgccgc atgagttcgcc acacgcgttgc ggcttagatc attcctca  
1261 gccggaggcg ctcatgtacc ctatgtaccctt ctcactgatgg gggccccctt tgcatag  
1321 cgacgtgaat ggcatecgcc acctctatgg tccctgcctt gaacctgagc caaggcc  
1381 aaccaccacc acaccgcage ccacggctcc cccgacggtc tgccccaccg gacecccc  
1441 tgtccacccccc teagagcgcctt ccacagctgg ccccacaggt ccccccctt cttgg  
1501 aggtcccccc actgtggccctt ctcactcgcc cactactgtt ctttgatgc cgggtgg  
1561 tgcctgaaac gtgaacatct tcgacgcctt cggggaggattt gggaaaccaggc ttttt  
1621 caaggatggg aagtactggc gattctctga gggcagggggg agccggccgc agggcc  
1681 ctttacgcgc gacaagtggc cccgcgttcc cccgacggatc gactcggtt tggagg  
1741 gctctcaag aagctttctt ttttctctgg ggcgcagggtt tgggtgtaca caggcg  
1801 ggtgtggc cccggggcgcc tggacaagat gggcctgggatc ggcgcacgtgg cccagg  
1861 cggggccctt cggaggatggc gggggaaagat gctgtgttcc agccggccgc gctctgg  
1921 ttgcacgtt aaggcgcaga tgggtggatcc cccgacggccgc agcgaggatgg accggatgtt  
1981 cccgggggtt ctttggaca cccgcacgtt cttccacgttcc cggaggatgg tggacc  
2041 ccaggaccgc ttctactggc cccgttgcgttcc cccgacggatgg tggacc  
2101 gggctacgtt acctatgaca tcctgcgttcc cccgttgcgttcc gctctgtt  
2161 ggcacgttccca tggaaatccc cactgggacc aaccctgggg aaggaggccag  
2221 acaaaactggt attctgttctt ggaggaaagg gaggaggatgg ggtggggctgg  
2281 ctcacccctt tttttgttgc gaggatggatgg ttttgcgttcc  
2341 aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa

Figure 9B

NCBI Reference Sequence: NP\_004985.2  
matrix metalloproteinase 9 preproprotein [Homo sapiens] (SEQ ID N0: 13)

1 MSLWQPLVLV LLVLGCCFAA PRQRQSTLVL FPGDLRTNLT DRQLAEELY RGYTRVAEM  
61 RGESKSLGPA LLLLQKQLSL PETGELEDSAT LKAMRTPRCG VPDLGRFQTF EGDLKWHHHN  
121 ITYWIQNYSE DLPRAVIdda FARAFALWSA VTPLTFRVY SRDADIVIQF GVAEHGDGYP  
181 FDGKDGLLAH AFPPGPGIQC DAHFDDDELW SLGKGVVVPT RFGNADGAAC HFPFIFEGRS  
241 YSACTTDGRSDG DGLPWCSSTA NYDTDDRFGF CPSERLYTQD GNADGKPCQF PFIFQGQSYS  
301 ACTTDGRSDG YRWCATTANY DRDKLFGFCP TRADSTVMGG NSAGEILCVFP FTFLGKEYST  
361 CTSEGRRGDGR LWCATTSNFD SDKKWGFCPD QGYSLFLVAA HEFGHALGLD HSSVPEALMY  
421 PMYRFTEGPP LHKKDDVNGIR HLYGPRPEPE PRPPPTTTTPQ PTAPPTVCPT GPPTVHPSER  
481 PTAGPTGPPS AGPTGPPTAG PSTATTVPLS PVDDACNVNI FDAIAEIGNQ LYLFKDGYW  
541 RFSEGRGSRP QGPFLIADKW PALPRKLDSV FEERLSKKLF FFSGRQVWVY TGASVLGPRR  
601 LDKLGLGADV AQVTGALRSG RGKMLLFSGR RLWRFDVKAQ MVDPRSASEV DRMFPGVPLD  
661 THDVFQYREK AYFCQDRFYW RVSSRSELNQ VDQVGYVTYD ILQCPED

//

Figure 10A

NCBI Reference Sequence: NP\_056179.2  
LY96 protein precursor [Homo sapiens] (SEQ ID N0: 14)

1 MLPFLFFSTL FSSIFTEAQK QYWVCNSSDA SISYTYCDKM QYPISINVNP CIELKRSKGL  
61 LHIFYIPRRD LKQLYFNLYI TVNTMNLPKR KEVICRGSDD DYSFCRALKG ETVNNTTISFS  
121 FKGIKFSKGK YKCVVEAISG SPEEMLFCLE FVILHQPNNSN

//

Homo sapiens carbonic anhydrase IV (CA4), mRNA  
NCBI Reference Sequence: NM\_000717.3 (SEQ ID NO: 15)

ORIGIN

1 cgctataaaa cccaggccgg caggatcgct gcacccgcgg cggcctcctc ggtgcgcgac  
61 ccccggctca gaggactctt tgctgtcccg caagatgcgg atgctgctgg cgctccctggc  
121 cctctccgcg ggcggccat cggccagtgc agagtcacac tggtgctacg aggttcaagc  
181 cgagtccctcc aactaccctt gcttggtgcc agtcaagtgg ggtggaaact gccagaagga  
241 cccgcaggccccc catcaaca tcgtcaccac caaggcaaag gtggacaaaaa aactgggacg  
301 cttcttccttc tctggctacg ataagaagca aacgtggact gtccaaaata acgggcaactc  
361 agtgatgtat ttgctggaga acaaggccag catttcttggaa ggaggactgc ctgccccata  
421 ccaggccaaa cagttgcacc tgcaactggc cgacttgcca tataagggtc cggagcacag  
481 cctcgatggg gagcaacttttgc ccatggagat gcacatagta catgagaaag agaaggggac  
541 atcgaggaat gtgaaagagg cccaggaccc tgaagacgaa attgcgtgc tggccttct  
601 ggtggaggct ggaacccagg tgaacgaggg cttccagcca ctgtggagg cactgtctaa  
661 tatccccaaa cctgagatga gcactacgtat ggcagagacg agcctgttgg acctgctccc  
721 caaggaggag aaactgaggc actactcccg ctacctggc tcactcacca caccgacactg  
781 cgatgagaag gtcgtcttggat ctgtgttcccg ggagccatt cagcttcaca gagaacagat  
841 cctggcattc tctcagaagc tgtaactacga caaggaacag acagtggca tgaaggacaa  
901 tgtcaggccc ctgcagcagc tggggcagcg cacggtgata aagtccgggg ccccggtcg  
961 gccgctgccc tggggcctgc ctgcccctgt gggcccatg ctggcctgcc tgctggccgg  
1021 cttcctgcga tgatggctca ctgtgcacg cagcctctcgttgcctcag ctctccaagt  
1081 tccaggcttc cggtccttag ctttcccagg tgggacttta ggcattgatta aatatggac  
1141 atatttttgg agaaaaaaaaaaaaa aaaaaa

//

Figure 11A

NCBI Reference Sequence: NP\_000708.1

carbonic anhydrase IV precursor [Homo sapiens] (SEQ ID N0: 16)

```
1 MRMLLALLAL SAARPSASAE SHWCYEVQAE SSNYPCLVPV KWGGNCQKDR QSPINIVTTK
 61 AKVDKKLGRF FFSGYDKKQT WTVQNGHHSV MMLLENKASI SGGGLPAPYQ AKQLHLHWSD
121 LPYKGSEHSL DGEHFAMEMH IVHEKEKGTS RNVKEAQDPE DEIAVLAFLV EAGTQVNNEGF
181 QPLVEALSNI PKPEMSTTMA ESSLDDLLPK EEKLRHYFRY LGSLTTPTCD EKVWWTVFRE
241 PIQLHREQIL AFSQKLYYDK EQTVSMKDGV RPLQQLGQRT VIKSGAPGRP LPWALPALLG
301 PMLACLLAGF LR
//
```

Figure 11B

NCBI Reference Sequence: NM\_005621.1

Homo sapiens S100 calcium binding protein A12 (S100A12), mRNA (SEQ ID N0: 17)

```
1 accactgctg gcttttgct gtagctccac attcctgtgc attgaggggt taacattagg
 61 ctggaaat gacaaaacctt gaagagcatc tggagggaaat tgtcaataatc ttccaccaat
121 actcagttcg gaagggccat tttgacaccc tctctaagg tgagctgaag cagctgctta
181 caaaggagct tgcaaaccacc atcaagaata tcaaagataa agctgtcatt gatgaaatat
241 tccaaggccct ggatgctaattt caagatgaac aggtcgactt tcaagaattc atatccctgg
301 tagccattgc gctgaaggct gcccattacc acacccacaa agagtaggta gctctctgaa
361 ggcttttac ccagcaatgt cctcaatgag ggtttttctt ttccctcacc aaaacccagc
421 cttggccgtg gggagtaaga gttaataaac acactcacga aaagtt
//
```

Figure 12A

NCBI Reference Sequence: NP\_005612.1  
S100 calcium-binding protein A12 [Homo sapiens] (SEQ ID N0: 18)

1 MTKLEEHLEG IVNIFHQYSV RKGHFDTLSK GELKQLLTKE LANTIKNIKD KAVIDEIFQG  
61 LDANQDEQVD FQEFLISLVAI ALKAAYHHTH KE  
//

Figure 12B

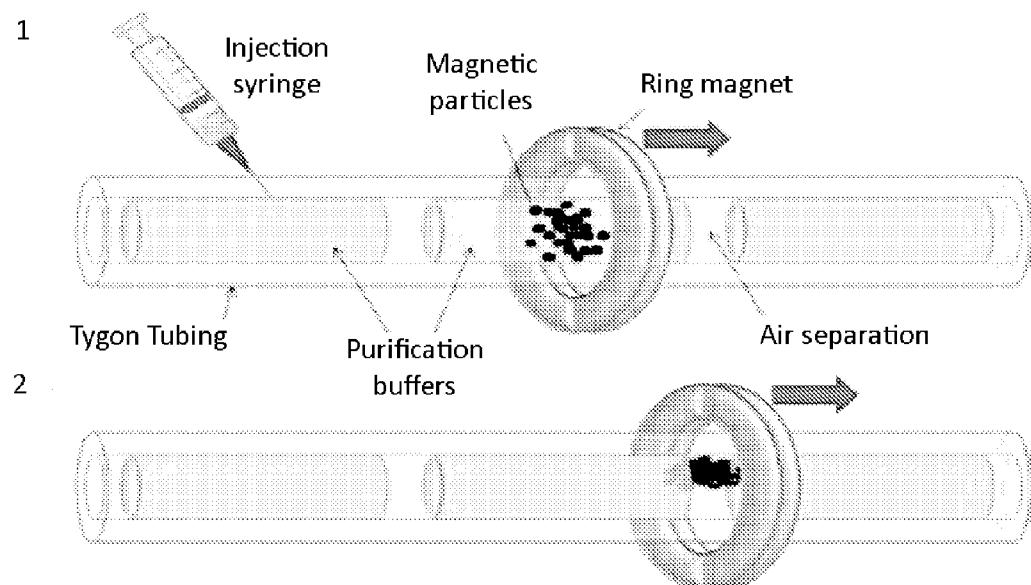


Figure 13

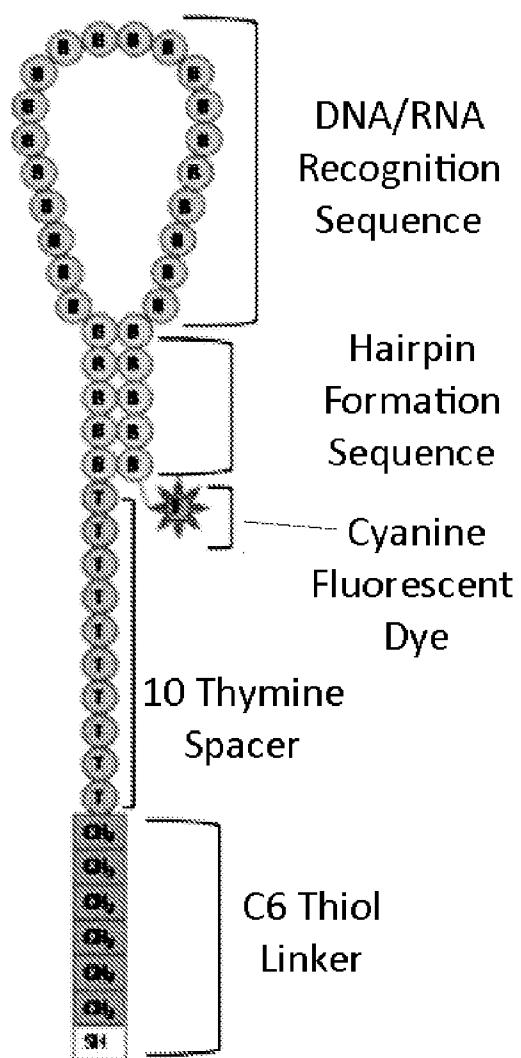


Figure 14

## Biomarkers for Acute Ischemic Stroke

### INCORPORATION BY REFERENCE

[0001] This application claims priority to U.S. Provisional Application Ser. No. 61/307,233, filed Feb. 23, 2010, the entire contents of which are incorporated herein by reference.

[0002] All documents cited or referenced herein and all documents cited or referenced in the herein cited documents, together with any manufacturer's instructions, descriptions, product specifications, and product sheets for any products mentioned herein or in any document incorporated by reference herein, are hereby incorporated by reference, and may be employed in the practice of the invention.

### GOVERNMENT SUPPORT

[0003] This application was supported by a pre-doctoral Intramural Research Training Award via the Graduate Partnerships Program through the National Institute of Nursing Research, National Institutes of Health. This research presented in this application was also supported by the Division of Intramural Research of the National Institute of Neurological Disorders and Stroke (Grant No. Z01 AG000957-05), and the Laboratory of Neurogenetics, National Institute of Aging National Institutes of Health. Accordingly, the Government may have certain rights in this invention.

### BACKGROUND OF THE INVENTION

[0004] 1. Field of the Invention

[0005] The present invention relates to the identification and use of diagnostic markers for acute ischemic stroke. In a various aspects, the invention relates to methods for rapid and early detection of acute ischemic stroke, as well as the identification of individuals at risk for acute ischemic stroke.

[0006] 2. Background

[0007] Stroke is clinically defined as a rapidly developing syndrome of vascular origin that manifests itself in focal loss of cerebral function. In more severe situations, the loss of cerebral function is global. Stroke can be categorized into two broad types, "ischemic stroke" (about 87%) and "hemorrhagic stroke" (about 10%). Ischemic stroke occurs when the blood supply to the brain is suddenly interrupted. Hemorrhagic stroke happens when a blood vessel located in our around the brain bursts leading to the leakage and accumulation of blood directly in the brain tissue. Additionally, a patient may experience transient ischemic attacks, which indicates a high risk for the future development of a more severe episode. Stroke also includes subarachnoid hemorrhage (about 3%).

[0008] The symptoms of stroke often include numbness or weakness, especially on one side of the body; sudden confusion or trouble speaking or understanding speech; sudden trouble seeing in one or both eyes; sudden trouble walking; dizziness; or loss of balance or coordination. Stroke is the most common devastating neurologic disease in the world and despite recent progress understanding stroke mechanisms, stroke management is still not optimal.

[0009] Stroke is the third leading cause of death in the world, after only heart disease and cancer. In the United States alone, approximately 780,000 people experience a stroke each year, which contributes to an overall financial burden of over \$65 billion per year. As noted, ischemic stroke accounts

for most instances of stroke in patients, and consequently, the category of stroke having the greatest financial burden.

[0010] Ischemic stroke encompasses thrombotic, embolic, lacunar and hypoperfusion types of strokes. Thrombi are occlusions of arteries created in situ within the brain, while emboli are occlusions caused by material from a distant source, such as the heart and major vessels, often dislodged due to myocardial infarct or atrial fibrillation. Less frequently, thrombi may also result from vascular inflammation due to disorders such as meningitis. Thrombi or emboli can result from atherosclerosis or other disorders, for example, arteritis, and lead to physical obstruction of arterial blood supply to the brain. Lacunar stroke refers to an infarct within non-cortical regions of the brain. Hypoperfusion embodies diffuse injury caused by non-localized cerebral ischemia, typically caused by myocardial infarction and arrhythmia.

[0011] Of the 88% of ischemic strokes that occur each year, 8-12% results in death within 30 days. The risk of ischemic stroke is associated with various familial and environmental factors, such as the presence of hypertension, obesity, tobacco use, and a positive family history. Determinants of outcome include non-modifiable risk factors such as age, race, gender and genetic variation along with clinical phenotypes of severity, such as stroke scale score, the presence of fever and serologic blood markers. Advances in neuroimaging and acute clinical management have resulted in greater numbers of patients surviving the initial insult. However, morbidity remains high secondary to complications following the primary event and initial misdiagnosis.

[0012] The onset of ischemic stroke is often abrupt, and can begin with a manifestation of neurologic deficits that worsen over a 24-48 hour period. Stroke-associated symptoms commonly include unilateral neurologic dysfunction that extends progressively, without producing headache or fever. Early manifestations may rapidly progress to more severe symptoms within a few minutes.

[0013] Hemorrhagic stroke is caused by intracerebral or subarachnoid hemorrhage, i.e., bleeding into brain tissue, following blood vessel rupture within the brain. Intracerebral and subarachnoid hemorrhage are subsets of a broader category of hemorrhage referred to as intracranial hemorrhage. Intracerebral hemorrhage is typically due to chronic hypertension, and a resulting rupture of an arteriosclerotic vessel. Stroke-associated symptom(s) of intracerebral hemorrhage are abrupt, with the onset of headache and steadily increasing neurological deficits. Nausea, vomiting, delirium, seizures and loss of consciousness are additional common stroke-associated symptoms.

[0014] In contrast, most subarachnoid hemorrhage is caused by head trauma or aneurysm rupture which is accompanied by high blood pressure release which also causes direct cellular trauma. Prior to rupture, aneurysms may be asymptomatic, or occasionally associated with tension or migraine headaches. However, headache typically becomes acute and severe upon rupture, and may be accompanied by varying degrees of neurological deficit, vomiting, dizziness, and altered pulse and respiratory rates.

[0015] Transient ischemic attacks (TIAs) have a sudden onset and brief duration, typically 2-30 minutes. Most TIAs are due to emboli from atherosclerotic plaques, often originating in the arteries of the neck, and can result from brief interruptions of blood flow. The symptoms of TIAs are identical to those of stroke, but are only transient. Concomitant

with underlying risk factors, patients experiencing TIAs are at a markedly increased risk for stroke.

[0016] There are few pharmaceutical therapies for treating stroke. In point of fact, the only Food and Drug Administration (FDA) approved treatment for ischemic stroke is recombinant tissue plasminogen activator (rtPA), alteplase. Multiple attempts to identify other pharmacologic agents have resulted in negative findings; therefore a redirection of the science is necessary to understand the human variable response to stroke, in particular, to ischemic stroke, to provide alternative avenues for therapeutic treatment.

[0017] Since its commencement into the clinical arena in 1996, recombinant tissue plasminogen activator (rtPA) has proven to be a promising therapeutic treatment for ischemic stroke and is safe and effective for use in routine clinical practice. However, its powerful effects are not seen without significant clinical complications. In addition, rtPA is only approved for use when patients present to the hospital within three hours from onset of symptoms. The downside is that the median time from stroke symptom onset to presentation to the emergency department is 3-6 hours. A recent study addressed the possibility of extending this limited therapeutic time window and it was shown that intravenous rtPA given between 3 and 4.5 hours after onset of symptoms significantly improved clinical outcomes following ischemic stroke compared to placebo. This is promising given that the time window limit of 3 hours and a large list of contraindicating factors for thrombolytic therapy currently results in only 3-8% of stroke patients receiving rtPA.

[0018] The advancements of rtPA therapy aside, there is still a demand for alternative acute ischemic stroke therapies in clinical practice. Unfortunately, the results of recent clinical trials have demonstrated that there is still a gap in the understanding of the variable human response to ischemic stroke. Numerous promising pre-clinical therapeutics display insignificant clinical utility in human patients, which speaks to the difficulty of translating what is learned at the bench to the patient at the bedside.

[0019] These negative findings may be due in part to the complexity of the human physiologic response to ischemic stroke, limited knowledge about the multiple pathways interacting in response to ischemic stroke and the implications of genomic variability on individual recovery from ischemic stroke. The difficulty may also be attributable to insufficient classification of ischemic stroke subtype. It is possible that gene expression profiling can help to identify subtypes of ischemic stroke, which has tremendous utility in designing therapeutic strategies for treatment. A better understanding of stroke pathophysiology in humans and more appropriate stroke subtyping may provide the foundation needed to design appropriate therapeutics for battling ischemic stroke and other stroke types.

[0020] Immediate diagnosis and care of a patient experiencing stroke can be critical. As noted, tissue plasminogen activator (rtPA) given within three hours of symptom onset in ischemic stroke is beneficial for selected acute ischemic stroke patients. Patients may also benefit from anticoagulants (e.g., heparin) if they are not candidates for rtPA therapy. In contrast, thrombolytics and anticoagulants are strongly contraindicated in hemorrhagic strokes. Thus, early and rapid differentiation of ischemic stroke from hemorrhagic-type stroke is imperative and often critical. Delays in the confir-

mation of stroke diagnosis and the identification of stroke type limit the number of patients that may benefit from early intervention therapy.

[0021] Another limitation in the diagnosis of ischemic stroke relates to the fact that, due to the rapidity of onset and progression of acute ischemic stroke, circumstances are such that ischemic stroke patients are often not seen by clinicians having the appropriate knowledge and training to be able to provide a correct, life-saving diagnosis. For instance, brain imaging technology is an integral and key aspect of the clinical stroke evaluations. However, such technology is often not available. Even if the technology is available, proper interpretation of brain imaging results concerning stroke diagnoses is best suited for those clinicians who are highly and specifically trained in the treatment and care of stroke patients. Indeed, due to the rapid onset of an acute ischemic stroke and other factors, such as the scarcity of trained stroke-clinicians and neurologists, clinical assessment of a potential victim is often carried out by a non-stroke specialist, e.g., a family practitioner, paramedic or triage nurse. Thus, achieving and accurate and rapid early diagnosis is often not possible under present clinical circumstances.

[0022] Patients, even those with mild symptoms, may be eligible for various ischemic stroke therapies if they can be started within a few hours of symptom onset. Patients who do not receive such early therapies are at an increased risk of recurrent stroke, often occurring even within a matter of days. Thus, prompt administration of a suitable therapy in a timely manner can substantially increase the efficacy of treatment while reducing the risk of recurrent stroke.

[0023] However, presently there are no rapid diagnostic procedures or methods that can be used to reliably determine whether a patient has suffered a stroke, in particular, an acute ischemic stroke, or whether a subject is at risk for ischemic stroke.

[0024] Accordingly, a rapid diagnostic test capable of making an accurate clinical diagnosis of ischemic stroke irrespective of the clinician's level of stroke expertise or lack thereof would be extremely useful. To date, the identification of appropriate biomarkers for ischemic stroke have proven to be extremely difficult. This may be tied, in part, to the observation that the proteins associated with stroke-affected brain and neurological tissues are slow to be released into the blood due to the blood-brain barrier. In addition, many potential stroke markers, including markers of cerebral ischemia and inflammation, located in the blood are associated with other conditions that may mimic stroke, e.g., severe myocardial infarction and brain infection.

[0025] This invention solves these deficiencies in the art.

## SUMMARY OF THE INVENTION

[0026] The present invention relates to the identification and use of diagnostic markers for acute ischemic stroke. The methods and compositions described herein can meet the need in the art for rapid, sensitive and specific diagnostic assay to be used in the diagnosis of acute ischemic stroke in patients or in the assessment of the risk of developing acute ischemic stroke. The methods and compositions can also be used to distinguish acute ischemic stroke from other various forms of stroke and TIAs and "stroke mimic" events. Moreover, the methods and compositions of the present invention can also be used to facilitate the treatment of stroke patients and the development of additional diagnostic and/or prognostic indicators.

[0027] Prior to this invention, the diagnosis of ischemic stroke has been difficult and inaccurate. These problems are due in part to limitations in the technology currently used to evaluate a patient for a stroke and limitations in the level of experience and/or proper training possessed by medical clinicians who engage the patients. These circumstances are detrimental to stroke victims because an accurate and quick diagnosis of acute ischemic stroke is extremely important to the health and outcome of the patients.

[0028] Brain imaging technology is an integral and key aspect of the clinical stroke evaluations but may suffer from a number of limitations, including lack of sensitivity and lack of availability. Moreover, proper interpretation of brain imaging results concerning stroke diagnoses is best suited for those clinicians who are highly and specifically trained in the treatment and care of stroke patients. Indeed, due to the rapid onset of an acute ischemic stroke and other factors, such as the scarcity of trained stroke-clinicians and neurologists, clinical assessment of a potential victim is often carried out by a non-stroke specialist, e.g., a family practitioner, paramedic or triage nurse. Thus, achieving an accurate and rapid early diagnosis is often not possible.

[0029] Patients, even those with mild symptoms, may be eligible for various ischemic stroke therapies if they can be started within a few hours of symptom onset. Patients who do not receive such early therapies are at an increased risk of recurrent stroke, often occurring even within a matter of days. Thus, prompt administration of a suitable therapy in a timely manner can substantially increase the risk of treatment while reduce the risk of recurrent stroke.

[0030] Accordingly, a rapid diagnostic test capable of making an accurate clinical diagnosis of ischemic stroke irrespective of the clinician's level of stroke expertise or lack thereof would be extremely useful. However, to date, the identification of appropriate blood biomarkers for ischemic stroke have proven to be extremely difficult. This may be tied, in part, to the observation that the proteins associated with stroke-affected brain and neurological tissues are slow to be released into the blood due to the blood-brain barrier. In addition, many potential stroke markers (including markers of cerebral ischemia and inflammation) located in the blood are associated with other conditions that may mimic stroke (e.g., severe myocardial infarction and brain infection). A recent comprehensive review of published blood-based stroke biomarkers revealed significant problems and weaknesses associated with currently published biomarkers, indicating that no single marker or panel of markers could be recommended for routine clinical practice.

[0031] Thus, at the time of the present invention, there remained a substantial need for a single biomarker or panel of biomarkers which were clinically effective in the diagnosis of acute ischemic stroke.

[0032] In one particular embodiment, using a gene expression profiling approach, this application describes obtaining peripheral blood samples from 39 stroke patients and 25 healthy controls at various times relative to the onset of acute ischemic stroke symptoms and then evaluated changes in gene expression profiles in both stroke and control groups over time. The gene expression profiles were analyzed to identify nine gene candidates having a likely role in acute ischemic stroke, thereby serving as biomarkers therefore.

[0033] Thus, in one aspect, the invention provides a set of biomarkers for use in methods for diagnosing acute ischemic stroke. In addition, the present invention is directed to com-

positions (e.g., arrays, probes, biomarker panels) that comprise the nine identified genes which can be used in the diagnosis of acute ischemic stroke and/or to distinguish acute ischemic stroke from TIAs and "stroke mimic" events. Further, since the biomarkers of the invention represent potential targets of intervention for the treatment of stroke, the biomarkers of the invention can be used in methods for screening compounds or agents that can treat acute ischemic stroke or a symptom thereof and which are detectable by the evaluation of the biomarkers of the invention.

[0034] In addition, the invention is directed to compositions that are useful in the detection of the biomarkers, including nucleic acid probes and antibodies that are specific for the biomarkers of the invention, as well as to compositions comprising purified biomarkers and their corresponding encoding nucleic acid molecules.

[0035] In a particular aspect, the present invention provides a method for the diagnosis of acute ischemic stroke in a subject, comprising detecting in a sample of whole peripheral blood obtained from the subject the presence of two or more biomarkers selected from the group consisting of: (a) chemokine receptor 7 (CCR7); (b) chondroitin sulfate proteoglycan 2 (CSPG2); (c) IQ motif-containing GTPase activation protein 1 (IQGAP1); (d) orosomucoid 1 (ORM1); (e) arginase 1 (ARG1); (f) lymphocyte antigen 96 (LY96); (g) matrix metalloproteinase 9 (MMP9); (h) carbonic anhydrase 4 (CA4); (i) s100 calcium binding protein A12 (s100A12); and wherein at least one of the biomarkers is (a), (b), (c) or (d), and wherein detection of the presence of the two or more biomarkers is indicative of acute ischemic stroke in the subject.

[0036] In one embodiment, the above method further comprises obtaining brain imaging data of the subject and evaluating the data to detect an acute ischemic stroke. The brain imaging data can be obtained through MRI or CT scan.

[0037] In another embodiment, the invention provides a method for identifying a candidate for acute ischemic stroke therapy using the biomarkers of the invention to diagnose acute ischemic stroke in a patient. The therapy can be the administration of a therapeutically effective amount of recombinant plasminogen activator (rtPA).

[0038] In yet another aspect, the invention provides a method for the diagnosis of acute ischemic stroke in a subject, comprising detecting in a biological sample obtained from the subject one or more biomarkers selected from the group consisting of: (a) chemokine receptor 7 (CCR7); (b) chondroitin sulfate proteoglycan 2 (CSPG2); (c) IQ motif-containing GTPase activation protein 1 (IQGAP1); and (d) orosomucoid 1 (ORM1).

[0039] In a further aspect, the invention provides a method for the diagnosis of acute ischemic stroke in a subject, comprising detecting in a biological sample obtained from the subject two or more biomarkers selected from the group consisting of: chemokine receptor 7 (CCR7); chondroitin sulfate proteoglycan 2 (CSPG2); IQ motif-containing GTPase activation protein 1 (IQGAP1); orosomucoid 1 (ORM1); arginase 1 (ARG1); lymphocyte antigen 96 (LY96); matrix metalloproteinase 9 (MMP9); carbonic anhydrase 4 (CA4); and s100 calcium binding protein A12 (s100A12), wherein at least one of the biomarkers is chemokine receptor 7 (CCR7); chondroitin sulfate proteoglycan 2 (CSPG2); IQ motif-containing GTPase activation protein 1 (IQGAP1); or orosomucoid 1 (ORM1).

[0040] In a still further aspect, the invention provides a method for differentiating an acute ischemic stroke from a transient ischemic attack (TIA), a hemorrhagic stroke and a stroke mimic in a subject presenting symptoms characteristic of a stroke or at risk of having a stroke, comprising:

(a) obtaining a biological sample from the patient;  
(b) contacting the biological sample with a detection means capable of detecting the presence of at least one biomarker selected from the group consisting of: chemokine receptor 7 (CCR7); chondroitin sulfate proteoglycan 2 (CSPG2); IQ motif-containing GTPase activation protein 1 (IQGAP1); and orosomucoid 1 (ORM1), wherein the presence of at least one of the biomarkers in the biological sample is indicative of an acute ischemic stroke but not indicative of a transient ischemic attack (TIAs), hemorrhagic stroke or stroke mimic.

[0041] In other aspects, the invention provides a kit comprising a means for detecting one or more biomarkers diagnostic of acute ischemic stroke, said biomarkers being selected from the group consisting of:

(a) chemokine receptor 7 (CCR7);  
(b) chondroitin sulfate proteoglycan 2 (CSPG2);  
(c) IQ motif-containing GTPase activation protein 1 (IQGAP1); and  
(d) orosomucoid 1 (ORM1).

[0042] In certain other aspects, the invention provides a filament-based diagnostic system comprising a panel of detectable polypeptides or functional polypeptide fragments thereof each corresponding to an acute ischemic stroke biomarker selected from the group consisting of:

(a) chemokine receptor 7 (CCR7);  
(b) chondroitin sulfate proteoglycan 2 (CSPG2);  
(c) IQ motif-containing GTPase activation protein 1 (IQGAP1); and  
(d) orosomucoid 1 (ORM1).

[0043] In still other aspects, the invention provides a filament-based diagnostic system comprising a panel of detectable oligonucleotides each corresponding to an acute ischemic stroke biomarker selected from the group consisting of:

(a) chemokine receptor 7 (CCR7);  
(b) chondroitin sulfate proteoglycan 2 (CSPG2);  
(c) IQ motif-containing GTPase activation protein 1 (IQGAP1); and  
(d) orosomucoid 1 (ORM1).

[0044] In still further aspects, the invention provides a filament-based diagnostic system comprising a panel of detectable antibodies each capable of specifically binding an acute ischemic stroke biomarker selected from the group consisting of:

(a) chemokine receptor 7 (CCR7);  
(b) chondroitin sulfate proteoglycan 2 (CSPG2);  
(c) IQ motif-containing GTPase activation protein 1 (IQGAP1); and  
(d) orosomucoid 1 (ORM1).

[0045] In certain embodiments, the sample is or is obtained from whole peripheral blood from the subject.

[0046] In other embodiments, the method of the invention is executed on the subject no more than 3 hours after onset of presenting acute ischemic stroke symptoms. In still further embodiments, the method is executed on the subject no more than 4.5 hours after onset of presenting acute ischemic stroke symptoms.

[0047] Other embodiments of the invention provide that the detecting step comprises contacting the biological sample

with a detection means capable of detecting the biomarker. The biomarker can be a nucleic acid molecule (e.g., mRNA) corresponding to or encoding one of (a) chemokine receptor 7 (CCR7); (b) chondroitin sulfate proteoglycan 2 (CSPG2); (c) IQ motif-containing GTPase activation protein 1 (IQGAP1); (d) orosomucoid 1 (ORM1); (e) arginase 1 (ARG1); (f) lymphocyte antigen 96 (LY96); (g) matrix metalloproteinase 9 (MMP9); (h) carbonic anhydrase 4 (CA4); or (i) s100 calcium binding protein A12 (s100A12).

[0048] In still other embodiments, the biomarker can be a polypeptide or active fragment thereof of (a) chemokine receptor 7 (CCR7); (b) chondroitin sulfate proteoglycan 2 (CSPG2); (c) IQ motif-containing GTPase activation protein 1 (IQGAP1); (d) orosomucoid 1 (ORM1); (e) arginase 1 (ARG1); (f) lymphocyte antigen 96 (LY96); (g) matrix metalloproteinase 9 (MMP9); (h) carbonic anhydrase 4 (CA4); or (i) s100 calcium binding protein A12 (s100A12).

[0049] Depending on the form of the biomarker of the invention (e.g., the mRNA or polypeptide of (a) chemokine receptor 7 (CCR7); (b) chondroitin sulfate proteoglycan 2 (CSPG2); (c) IQ motif-containing GTPase activation protein 1 (IQGAP1); (d) orosomucoid 1 (ORM1); (e) arginase 1 (ARG1); (f) lymphocyte antigen 96 (LY96); (g) matrix metalloproteinase 9 (MMP9); (h) carbonic anhydrase 4 (CA4); or (i) s100 calcium binding protein A12 (s100A12)), the detection means can be, but is not limited to, an antibody or oligonucleotide probe. The detection means can also be in the form of a kit or assay, such as a filament-based diagnostic system capable of detecting a polypeptide biomarker or a nucleic acid molecule biomarker of the invention.

[0050] In certain other embodiments, the inventive methods include treating the subject with a stroke therapy if the subject is diagnosed as having had an acute ischemic stroke or is at risk of therefore. The stroke therapy can include the administration of a therapeutically effective amount of recombinant plasminogen activator (rtPA).

[0051] In still other embodiments, the one or more detected biomarkers of the invention using the methods of the invention has at least a 1.5 fold increase or decrease in expression level as compared to the levels of the one or more biomarkers in a non-stroke subject.

[0052] In other embodiments, the one or more detectable biomarkers of the invention using the methods of the invention has at least a 2.0 fold increase or decrease in expression level as compared to the levels of the one or more biomarkers in a non-stroke subject.

[0053] In certain other embodiments of the invention, kits comprising a filament-based diagnostic system includes a surface on which is attached at known locations one or more oligonucleotides capable of hybridizing to the biomarkers.

[0054] In other embodiments of the invention, kits comprising a filament-based diagnostic system includes a surface on which is attached at known locations one or more antibodies capable of binding to the biomarkers. The surface can be a microarray, microtiter plate, membrane or the like. The kits may also include instructions for use.

[0055] These and other embodiments are disclosed or are obvious from and encompassed by, the following Detailed Description.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0056] The following detailed description, given by way of example, but not intended to limit the invention solely to the

specific embodiments described, may best be understood in conjunction with the accompanying drawings.

[0057] FIG. 1. Depicts the results of a nine-gene panel for ischemic stroke diagnosis. After comparison between both statistical packages there were 9 genes differentially expressed with at least a 2 fold difference in expression and Bonferroni corrected p<0.05 between stroke patients and control subjects. (ARG1, CA4, CCR7, CSPG2, IQGAP1, LY96, MMP9, ORM1, S100A12)

[0058] FIG. 2 provides a graph depicting the top pathways affected by acute ischemic stroke. A total of 355 genes with at least a 1.5 fold difference in expression between stroke patients and control subjects and a Bonferroni corrected p<0.05 were eligible for pathway analysis. A score, which was derived from a p value, was generated for each pathway; scores of 2 or higher are considered to have at least a 99% confidence of not being generated by chance alone. IPA analysis showed that the top 5 most significant canonical pathways in the peripheral blood of AICS patients were associated with CD28 signaling in T-helper cells (p=4.03E00), nuclear factor of activated T cells (NFAT) in regulation of the immune response (p=4.03E00) dendritic cell maturation (p=3.4E00), toll-like receptor signaling (p=3.33E00), and calcium-induced T-lymphocyte apoptosis (p=2.92E00).

[0059] FIG. 3 provides a schematic depicting Toll-like Receptor Signaling. There were more genes involved in the Toll-like receptor signaling pathway in our dataset compared to the total number of genes in the Toll-like receptor pathway with a ratio >2 and p=3.33E00 (TLR2, TLR1, FOS, LY96, TLR8 (includes EG: 51311), IF2AK2, IRAK3). The activation of the TLR4/LY96(MD2) complex and TLR2 from endogenous alarmins such as s100A12 results in activation of IRAK1 through interaction with MyD88. This leads to engagement of TRAF6, which is a member of the TNF receptor family. Then through the IkappB kinase pathway, NF kappaB translocates from the cytoplasm to the nucleus where it stimulates transcription of both proinflammatory and anti-inflammatory cytokines and chemokines (e.g. pro inflammatory IL6 and IL8 (resulting in neutrophil release of MMP9); anti-inflammatory IL10 and TNF $\alpha$ ). Simultaneous with NfkappaB activation, engagement of TRAF6 also results in: 1) the stimulation of the JNK pathway and activation of immediate early response genes cJun and cFos which come together to form the API early response transcription factor which increases the activation of the MAPK signaling cascade with subsequent ELK1 activation with binding to the serum response factor (SRF) in the promoter of the cFos proto-oncogene. Therefore, within the TLR pathway there are both pro-inflammatory mechanisms of engagement (NF-kappaB) and anti-inflammatory and neurotrophic pathways of regeneration (BDNF and SRF).

[0060] FIGS. 4A and 4B provide the nucleotide and amino acid sequences, respectively, for human CCR7.

[0061] FIGS. 5A and 5B provides the nucleotide and amino acid sequences, respectively, for human CSPG2.

[0062] FIGS. 6A and 6B provides the nucleotide and amino acid sequences, respectively, for human IQGAP1.

[0063] FIGS. 7A and 7B provides the nucleotide and amino acid sequences, respectively, for human ORM1.

[0064] FIGS. 8A and 8B provides the nucleotide and amino acid sequences, respectively, for human ARG1.

[0065] FIGS. 9A and 9B provides the nucleotide and amino acid sequences, respectively, for human LY96.

[0066] FIGS. 10A and 10B provides the nucleotide and amino acid sequences, respectively, for human MMP9.

[0067] FIGS. 11A and 11B provides the nucleotide and amino acid sequences, respectively, for human CA4.

[0068] FIGS. 12A and 12B provides the nucleotide and amino acid sequences, respectively, for human s100A12.

[0069] FIG. 13 depicts a simple pull-through DNA/RNA extraction design, i.e., a filament based Point of Care (POC) test of the invention. In this system, 100  $\mu$ l of each consecutive treatment buffer is pre-loaded by injecting through the wall of the Tygon tubing. High surface tension forces between the liquid/air interface in this small diameter tubing maintain solution separation. A ring magnet is moved longitudinally along the length of the Tygon tubing thereby pulling magnetic particles containing nucleic acid molecule of interest bound thereto across air separations from one purification buffer to the next.

[0070] FIG. 14 is a schematic of a molecular beacon for use in the present invention. Near the 3' end, binding of DNA or RNA complementary to the loop region opens the stem region and the cyanine fluorescent dye moves away from the gold surface attached via a T spacer & C6 thiol linker at the 5' end.

#### DETAILED DESCRIPTION OF THE INVENTION

[0071] The present invention is based, at least in part, on the unexpected discovery of certain biomarkers for acute ischemic stroke in peripheral blood. Accordingly, the present invention provides methods and compositions for the identification and use of biomarkers that are associated with the diagnosis and prognosis of acute ischemic stroke in a subject. Such biomarkers can be used in diagnosing and treating a subject for acute ischemic stroke and/or to monitor the course of an ischemic stroke treatment regimen; for screening subjects for the occurrence or risk of acute ischemic stroke; and for screening compounds and pharmaceutical compositions that might provide a benefit in treating or preventing such conditions.

#### Acute Ischemic Stroke

[0072] Acute ischemic stroke occurs when there is a decrease or loss of blood flow to an area of the brain resulting in tissue damage or destruction. It is the largest subtype of stroke pathologies and therefore accounts for the majority of the death and disability associated with stroke.

[0073] There are numerous scenarios that contribute to compromised cerebral perfusion; combine this with the multifactorial effects of the environment and individual genomic responses secondary to DNA variation and epigenetic DNA modification and the result is variability of patient presentation and recovery. This complexity makes ischemic stroke very difficult to treat, both medically and pharmacologically. Most cerebral ischemic pathologic conditions involve alterations in cerebrovascular reactivity and clot formation.

[0074] Ischemia is the consequence of one or more of the following causes: thrombosis, embolism or decreased systemic circulation. Thrombosis is a localized obstruction of blood flow in one or more blood vessels, most commonly caused by atherosclerosis. Stroke refers to a blockage of the blood vessel not caused by a localized process, but rather material originating from outside of the cerebral circulation, most commonly the heart. Decreased systemic perfusion

results in a decrease in cerebral perfusion pressure, and ultimately cerebral blood flow secondary to cardiac pump failure or systemic hypotension.

[0075] Although each scenario has a different origin, the result is either a temporary or permanent decrease or loss of cerebral blood flow. Permanent loss of cerebral blood flow to an area of the brain resulting in cell death is termed infarction. The penumbra is the area of the brain receiving less than optimal blood flow and is damaged but "salvageable", not yet infarcted. In all cases of ischemic stroke, the intent of therapy is to rescue this penumbral tissue and if therapy or reperfusion occurs quickly, this tissue can be rescued. The extent of tissue damage depends on the location and duration of the infarction or lack of blood flow and the extent to which collateral vessels can supply oxygen and nutrients to compromised areas. The cerebral ischemic response is complex and involves a decrease in oxygen and glucose delivery but also an accumulation of detrimental metabolic waste products. Therefore reactive oxygen species (ROS) and inflammatory mediators play a critical role in the events following ischemia.

[0076] Immediately following ischemic brain injury a cascade of events occurs in response to loss of blood flow. Alterations at the cell membrane result in release of glutamate, activation of N-methyl-D-aspartic acid (NMDA) receptors and release of calcium ( $\text{Ca}^{+2}$ ) into the extracellular space. This process ultimately leads to the activation of Immediate Early Genes (IEG's), such as c-fos and c-jun. IEG's propagate the physiologic response by participating in transcription of neurotrophic factors (endogenous neuroprotection), heat shock proteins (general stress response), cytokines and immune mediated complexes (inflammatory and immune activation), and nitric oxide synthase (NOS) activation (neuronal stimulation). Pathway specific responses are mediated by non-modifiable factors (DNA variation, age, gender, and severity of injury), modifiable factors (diet, physical activity, temperature, and environmental stress) and the interaction of signaling molecules within the pathways themselves.

[0077] The collective response, secondary to human genetic variation, results in either propagation of injurious mechanisms and cell death or initiation of repair mechanisms and neuronal sustainment.

[0078] Since the brain does not store oxygen or glucose, cellular energy production fails to maintain normal metabolism within minutes following compromised cerebral blood flow. Within the mitochondria, the electron transport chain removes electrons from an electron donor and passes them onto oxygen to form water through a series of redox reactions. These reactions create a proton gradient across the mitochondrial membrane that drives production of adenosine triphosphate (ATP). ATP then enters the Krebs cycle (citric acid cycle) to become part of a metabolic pathway that converts carbohydrates, fats, and proteins into usable forms of energy (e.g. carbon dioxide and water). When oxygen is unavailable the electron transport chain can no longer accept electrons; a proton gradient is not produced, ATP production ceases, and pyruvate becomes the final acceptor of electrons in the chain. This switch from oxygen dependent aerobic metabolism to anaerobic energy production results in an accumulation of lactic acid and ionic pump failures.

[0079] Sodium potassium ( $\text{Na}^{+}/\text{K}^{+}$ ) pumps are highly dependent upon ATP energy production and begin to fail within minutes of anaerobic metabolism.  $\text{Na}^{+}$ , water and calcium ( $\text{Ca}^{+2}$ ) begin to pass from the extracellular space to the intracellular space and cerebral cells begin to swell,

resulting in cytotoxic edema. Capillary endothelial cells begin to function abnormally and the tight junctions between them lose their integrity, leading to blood brain barrier (BBB) disruption. Intravascular fluid leaks into the extravascular space and spreads easily throughout the white matter, resulting in vasogenic edema. Additionally, excess intracellular  $\text{Ca}^{+2}$  triggers free radical production along with free radicals produced during anaerobic metabolism contributing to protease and lipase activation. Superoxide and peroxynitrite (free radicals) production increases beyond the cells capacity to quench them, which in turn activates the production of other detrimental ROS. Eventually the swollen cells enter cell death pathways through necrotic or apoptotic processes dependent upon the cell type, severity of injury and the level of available ATP. Unfortunately, reoxygenation through reperfusion also acts as a substrate for enzymatic reactions that produce ROS. Cells attempt to minimize damage caused by ischemia by rebalancing energy supply and demand. This early neuroprotective response results in an overall suppression of non-essential energy consumption.

[0080] All of these events together result in blood-brain-barrier (BBB) permeability, loss of cell ion homeostasis, and excitotoxicity, resulting in a modulation of gene and protein expression. The molecular imprint of these processes is visible within all cells that migrate and circulate throughout the area of cerebral injury. These cells then circulate out of the central nervous system into the peripheral blood.

[0081] Clinical diagnosis of ischemic stroke is often difficult, complicated by its multiple etiologies and variable clinical presentation. In most hospitals, diagnosis is made when the patient presents with symptoms suggestive of acute cerebral ischemia in conjunction with pathologic findings on cerebral imaging that are most likely associated with the presenting symptoms. When possible, medical history is obtained via inquisition of the following: personal and family medical history; discussion of history of stroke or symptoms suggestive of stroke; time and activity at the onset of symptoms; temporal progression of symptoms; and whether or not they are accompanied by other factors, such as headache or nausea. Unfortunately more often than not, this information is unobtainable secondary to severity of the stroke and whether or not the patient has family available that can provide the history. At this point the physical examination and brain imaging findings are used to make the definitive diagnosis. Identifying the physical location of the stroke is made by assessment of neurologic status via the neurologic examination, which may include National Institutes of Health Stroke scale (NIHSS) score and the presence of pathologic findings on computed tomography (CT) or magnetic resonance imaging (MRI).

[0082] The majority of hospitals in the United States use CT to rule out stroke; however it has been found that CT is less than optimal for identifying acute ischemia. A recent study of emergency room (ER) neurology consults found that the initial diagnosis of the ER physician agrees with the final diagnosis ~60% of the time. There was a significant pattern of mis-diagnosis for stroke and seizure; other benign medical conditions (e.g. migraine) and psychiatric disorders were originally diagnosed and medically treated as stroke. Although over-diagnosis of stroke early may appear to err on the side of patient safety, it puts a percentage of patients into a category where they are being treated for an acute stroke that they do not have. Given the complications associated with bleeding following rtPA administration, this practice is quite

risky. For this reason, some ER physicians are reluctant to treat a patient with rtPA unless they are completely sure of a stroke diagnosis.

[0083] Where possible, hospitals are moving toward using MRI for acute diagnosis of stroke; however this is facilitated best by the presence of a dedicated stroke clinical team and only possible in facilities with 24 hour MRI availability. Additionally, even though rtPA is FDA approved, only a small number of stroke patients actually receive the drug. The Brain Attack Coalition has recommended a movement toward the creation of widespread primary stroke centers to increase the utilization of rtPA and creation of standards of care for ischemic stroke patients. In a small community hospital located in Bethesda, Md. the establishment of a primary stroke team resulted in a 7-fold increase in the proportion of stroke patients treated with rtPA within 24 months after the onset of the program. This paradigm can be applied to other small community hospitals and has the potential to increase the numbers of stroke patients treated with rtPA by an additional 30,000 patients per year. More and more hospitals are moving towards the establishment of primary stroke teams; however there are still quite a few hospitals that rely solely on the ER physician's expertise in diagnosing ischemic stroke.

[0084] Quick and definitive diagnosis in the acute care setting is essential to separate stroke from non-stroke, distinguish hemorrhage from ischemia, and identify the potential cause of the infarction, but most importantly to determine eligibility for thrombolytic therapy (e.g., rtPA-alteplase) to begin treatment within the three hour window of opportunity. An additional diagnostic measure, such as a serologic blood test or a screen of a panel of markers, would be extremely beneficial in obtaining a definitive diagnosis of acute stroke and increasing the utilization of rtPA, especially in hospitals where primary stroke centers are non-existent.

#### DEFINITIONS AND USE OF TERMS

[0085] The present invention may be understood more readily by reference to the following detailed description of preferred embodiments of the invention and the Examples included therein. Before the present methods and techniques are disclosed and described, it is to be understood that this invention is not limited to specific analytical or synthetic methods as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting. Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by one of ordinary skill in the art to which this invention belongs.

[0086] As used herein and in the appended claims, the singular forms "a," "and," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a gene" is a reference to one or more genes and includes equivalents thereof known to those skilled in the art, and so forth.

[0087] As used herein, the term "antibody" refers to immunoglobulin molecules (e.g., any type, including IgG, IgE, IgM, IgD, IgA and IgY, and/or any class, including, IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2) isolated from nature or prepared by recombinant means or chemically synthesized. The terms "antibody" and "immunoglobulin" can be used synonymously throughout the specification, unless indicated otherwise. Antibodies or immunoglobulins of the invention can be used for various purposes, including, for example, the

detection of the biomarkers of the invention through the use of any suitable detection assay, e.g., ELISA.

[0088] As used herein, the terms "biological sample" or "patient sample" or "test sample" or "sample" as used herein, refer to a sample obtained from an organism or from components (e.g., cells) of a subject or patient for the purpose of diagnosis, prognosis, or evaluation of a subject of interest, such as a patient. In certain embodiments, such a sample may be obtained for the purpose of determining the outcome of an ongoing condition or the effect of a treatment regimen on a condition. The sample may be of any biological tissue or fluid. The sample may be a clinical sample which is a sample derived from a patient. Such samples include, but are not limited to, brain cells or tissues, cerebrospinal fluid, nerve tissue, sputum, blood, serum, plasma, blood cells (e.g., white cells), tissue samples, biopsy samples, urine, peritoneal fluid, and pleural fluid, saliva, semen, breast exudate, tears, mucous, lymph, cytosols, ascites, amniotic fluid, bladder washes, and bronchioalveolar lavages or cells therefrom, among other body fluid samples. Preferably, the sample is peripheral blood. Preferable, the sample contains one or more of the biomarkers of the invention, or a nucleic acid encoding a biomarker of the invention (e.g., mRNA). The patient samples may be fresh or frozen, and may be treated, e.g. with heparin, citrate, or EDTA. Biological samples may also include sections of tissues such as frozen sections taken for histological purposes.

[0089] As used in this invention, the term "epitope" means any antigenic determinant on an antigen, e.g., a biomarker of the invention, to which an antibody binds through an antigenic binding site. Determinants or antigenic determinants on an antigen usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three dimensional structural characteristics, as well as specific charge characteristics.

[0090] As used herein, the term antibody that "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide is one that binds to that particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope. Alternatively, an antibody that specifically binds to an antigen, in accordance with this invention, refers to the binding of an antigen by an antibody or fragment thereof with a dissociation constant (IQ) of 104 or lower, as measured by a suitable detection instrument, e.g., surface plasmon resonance analysis using, for example, a BIACORE® surface plasmon resonance system and BIACORE® kinetic evaluation software (eg. version 2.1). The affinity or dissociation constant ( $K_d$ ) for a specific binding interaction is preferably about 500 nM or lower, more preferably about 300 nM or lower and preferably at least 300 nM to 50 pM, 200 nM to 50 pM, and more preferably at least 100 nM to 50 pM, 75 nM to 50 pM, 10 nM to 50 pM.

[0091] The term "treatment" includes any process, action, application, therapy, or the like, wherein a subject (or patient), including a human being, is provided medical aid with the object of improving the subject's condition, directly or indirectly, or slowing the progression of a condition or disorder in the subject, or ameliorating at least one symptom of the disease or disorder under treatment.

[0092] The term "combination therapy" or "co-therapy" means the administration of two or more therapeutic agents to treat a disease, condition, and/or disorder, e.g., acute ischemic stroke. Such administration encompasses "co-administra-

tion” of two or more therapeutic agents in a substantially simultaneous manner. One therapy can be based on the biomarkers of the invention. A second therapy can be based on a known therapy for a disorder, e.g., acute ischemic stroke, such as tissue plasminogen activator (rtPA). The order of administration of two or more sequentially co-administered therapeutic agents is not limited.

[0093] The phrase “therapeutically effective amount” means the amount of each agent (e.g., an agent that beneficially interacts with a biomarker of the invention to treat acute ischemic stroke) administered that will achieve the goal of improvement in a disease, condition, and/or disorder severity, and/or symptom thereof, while avoiding or minimizing adverse side effects associated with the given therapeutic treatment.

[0094] The term “pharmaceutically acceptable” means that the subject item is appropriate for use in a pharmaceutical product.

[0095] The term “biomarker” or “marker” as used herein refers to proteins or polypeptides (or active fragment thereof) that are associated with brain tissue or neural cells, and which can be correlated with acute ischemic stroke, but are not correlated with other types of injury. Such specific biomarkers of acute ischemic stroke identified by the methods of the invention include (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); and (9) s100 calcium binding protein A12 (s100A12), and the like. These specific biomarkers are described in detail hereinafter. Biomarkers can be detected or identified or measured and the like using any suitable methods or instrumentation for measuring, identifying or detecting polypeptides or proteins. In certain embodiments, nucleic acids encoding the biomarkers can be measured, identified or detected using any suitable means by which to analyze nucleic acid molecules (e.g., mRNA molecules in peripheral blood encoding a biomarker of the invention). Where the detection or diagnosis is made through measuring, detecting, or the like of the nucleic acid molecules (e.g., mRNA) corresponding to or encoding a biomarker polypeptide of the invention, the term “biomarker” can also be in reference to the nucleic acid molecule itself. For example, a “biomarker” of the invention can include mRNA (or DNA) encoding (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); and (9) s100 calcium binding protein A12 (s100A12). Other biomarkers can be identified using the screening methods of the invention.

[0096] The term “active fragment of a biomarker” refers to a fragment of a biomarker having sufficient sequence such that it still possesses the same or substantially the same function as the full-size biomarker. Preferably, an active fragment of a biomarker retains at least 100% of the activity of the full-size biomarker, or at least 99%, 95%, 90%, 85%, 80% 75%, 70%, 65% or 60% of its activity. In certain embodiments, an active fragment of a biomarker is one which is immunologically detectable (i.e., detectable using an antibody).

[0097] “Proteins or polypeptides” used as biomarkers in the present invention are contemplated to include any fragments thereof, in particular, immunologically detectable fragments. One of skill in the art would recognize that proteins which are released by cells of the central nervous system which become damaged during a cerebral attack (e.g., acute ischemic stroke) could become degraded or cleaved into such fragments. Additionally, certain markers are synthesized in an inactive form, which may be subsequently activated, e.g., by proteolysis. Examples of such markers are described hereinafter. The term “related marker” as used herein refers to one or more fragments of a particular marker that may be detected as a surrogate for the marker itself. These related markers may be, for example, “pre,” “pro,” or “prepro” forms of biomarkers, or the “pre,” “pro,” or “prepro” fragment removed to form the mature marker. In preferred embodiments, these “pre,” “pro,” or “prepro” forms or the removed “pre,” “pro,” or “prepro” fragments are used in an equivalent fashion to the mature markers in the methods described herein.

[0098] The phrase “diagnosis” as used herein refers to methods by which the skilled artisan can estimate and/or determine whether or not a patient is suffering from, or is at some level of risk of developing, a given disease or condition. The skilled artisan (e.g., stroke clinician or emergency room physician) often makes a diagnosis on the basis of one or more diagnostic indicators, i.e., a biomarker, the risk, presence, absence, or amount of which is indicative of the presence, severity, or absence of the condition, e.g., acute ischemic stroke.

[0099] The term “interaction” refers to direct or indirect binding or alteration of a biological activity of a biomolecule, e.g., a biomarker.

[0100] The term “sensitivity”, as used herein in the context of its application to diagnostic assays, refers to the proportion of all subjects with acute ischemic stroke or at risk for developing acute ischemic stroke that are correctly identified as such (that is, the number of true positives divided by the sum of the number of true positives and false negatives).

[0101] The term “specificity” of a diagnostic assay, as used herein in the context of its application to diagnostic assays, refers to the proportion of all subjects not having acute ischemic stroke or who are not at risk for developing acute ischemic stroke that are correctly identified as such (that is, the number of true negatives divided by the sum of the number of true negatives and false positives).

[0102] Within the context of the invention, the terms “detect”, “detection” or “detecting” refer to the identification of the presence, absence, or quantity of a given biomarker.

[0103] As used herein, the term “acute ischemic stroke” refers to those patients having or at risk for “definite acute ischemic cerebrovascular syndrome (AICS)” as defined by the diagnostic criteria of Kidwell et al. “Acute Ischemic Cerebrovascular Syndrome: Diagnostic Criteria,” Stroke, 2003, 34, pp. 2995-2998 (incorporated herein by reference). AICS diagnostic criteria are as follow:

Definite AICS: Acute onset of neurologic dysfunction of any severity consistent with focal brain ischemia AND imaging/laboratory CONFIRMATION of an acute vascular ischemic pathology.

Probable AICS: Acute onset of neurologic dysfunction of any severity suggestive of focal brain ischemic syndrome but WITHOUT imaging/laboratory CONFIRMATION of acute ischemic pathology\* (diagnostic studies were negative but INSENSITIVE for ischemic pathology of the given duration,

severity and location). Imaging, laboratory, and clinical data studies do not suggest nonischemic etiology: possible alternative etiologies ARE ruled out.

Possible AICS: Acute neurologic dysfunction of any duration or severity possibly consistent with focal brain ischemia WITHOUT imaging/laboratory CONFIRMATION of acute ischemic pathology\* (diagnostic studies were not performed or were negative and SENSITIVE for ischemic pathology of the given duration, severity and location). Possible alternative etiologies are NOT ruled out. Symptoms may be nonfocal or difficult to localize.

Not AICS: Acute onset of neurologic dysfunction with imaging/laboratory CONFIRMATION of NONISCHEMIC pathology/(including normal imaging/laboratory studies that are highly sensitive for ischemic pathology of the given duration, severity, and location) as the cause of the neurologic syndrome.

[0104] As used herein, reference to "stroke symptoms" or "symptoms characteristic of a stroke" can refer to those symptoms that may present at the onset of any type of stroke (including acute ischemic and hemorrhagic stroke and others), including those symptoms recognized by the National Stroke Association ([www.stroke.org](http://www.stroke.org)), which are as follows: (a) sudden numbness or weakness of the face, arm or leg—especially on one side of the body; (b) sudden confusion, trouble speaking or understanding; (c) sudden trouble seeing in one or both eyes; (d) sudden trouble walking, dizziness, loss of balance or coordination, and (e) sudden severe headache with no known cause.

#### Biomarkers

[0105] In one aspect, the present invention provides biomarkers for diagnosing and detecting acute ischemic stroke in a patient at risk for ischemic stroke or who has already had a stroke event.

[0106] In one embodiment, the biomarkers include a 9-biomarker panel comprising: (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); and (9) s100 calcium binding protein A12 (s100A12). The biomarkers include both the polypeptides or nucleic acid molecules corresponding to the polypeptides, e.g., mRNA encoding the polypeptide biomarkers.

[0107] In other embodiments, the invention provides additional suitable biomarkers for detecting and diagnosing stroke events utilizing the biomarker screening methods of the invention as exemplified in the Examples. These screening methods relate to the surprising finding that biomarkers for stroke may be found by analyzing gene expression profiling of whole peripheral blood. Such methods can be utilized to identify other biomarkers for acute ischemic stroke, or other types of stroke or brain trauma.

[0108] The biomarkers of the invention, including the 9-biomarker panel identified herein, have various utilities, including, for example, their use in rapid blood tests to evaluate risk of acute ischemic stroke or to provide a diagnosis of acute ischemic stroke in a patient or to diagnose other forms of stroke or brain trauma.

[0109] As noted, a rapid blood test for the diagnosis of acute ischemic stroke would transform stroke care in the United States and throughout the world. Most hospitals

across the U.S. are not large academic centers where stroke neurologists are available at all times and where an MRI can be used for acute assessment of cerebral ischemic changes prior to the administration of rtPA. The standard, however, is something quite different. More often than not, an emergency room (ER) physician would be given the task of assessing, diagnosing and treating acute ischemic stroke through clinical history assessment and CT without the assistance of a stroke-trained neurologist.

[0110] Recent studies have reinforced that although ER physicians are more than capable of treating stroke patients, they are often reluctant to administer stroke therapies (e.g., rtPA) unless the diagnosis is definitive. In addition, there is a shortage of trained emergency personnel and ER nursing staff capable of identifying stroke symptoms or conducting an appropriate stroke assessment. The small percentage of patients who actually receive rtPA (3-5%) and the large numbers of patients who leave the hospital (without treatment) with either a diagnosis of transient ischemic attack (TIA) or stroke of undetermined cause pays testament to the need to identify additional means of stroke diagnosis.

[0111] The skilled artisan will appreciate that peripheral blood markers specific for brain injury have proven virtually impossible to identify. Some groups have even begun to question the use of blood biomarkers in the study of acute brain injury. Numerous studies over the years have resulted in either insignificant findings or findings that could not be replicated. Traditional methods for the identification of these biomarkers have fallen short of the rigor and sensitivity necessary to identify such markers for brain injury, including ischemic stroke.

[0112] In one aspect, the present invention provides a "panel" of genes (or biomarkers) and their cognate encoded polypeptide products that can be used to detect or diagnose acute ischemic stroke. As shown in the Examples, the invention provides at least 9 genes identified by the methods of the invention that can predict acute ischemic stroke in a patient with a substantially high degree of accuracy as compared to MRI or CT based methods. Preferably, the predictive value of the 9 biomarker panel is at least 95%; or preferably at least 90%, 85%, 80%, 75%, 70%, 65%, or 60% accurate, as compared to the diagnostic capability of both MRI (85% accurate) and CT (54% accurate).

[0113] The 9-biomarker panel identified in this study comprises: (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); and (9) s100 calcium binding protein A12 (s100A12). Each of these biomarkers is described further as follows.

[0114] (1) Chemokine Receptor 7.

[0115] Chemokines are a family of small proteins that regulate leukocyte trafficking. Aside from their role in inflammatory and immune responses there is increasing evidence that they play a significant role in glial cell proliferation and migration as part of the neuro-immune response. Several chemokines have been identified in both the serum and CSF (cerebrospinal fluid) of stroke patients. CXCL5, CCL2, CCL3, and CXCL8 are significantly increased following stroke and play a modulatory role of inflammation during the acute phase of ischemia. In addition CCR8 is expressed in activated microglia on brain sections of ischemic stroke

patients. The down regulation of CCR7 in peripheral blood as shown by the methods of this invention (see Examples) suggests there is decreased glial cell proliferation and migration very early in the acute phase of ischemic stroke; which coincides with the literature that these cytokines become increasingly more active during recovery and repair. The human CCR7 mRNA sequence is publicly available as GenBank Accession No. NM\_001838, the complete sequence of which is shown in FIG. 4A. The human CCR7 amino acid sequence is publicly available as GenPept Accession No. NP\_001829, the complete sequence of which is shown in FIG. 4B.

[0116] (2) Chondroitin Sulfate Proteoglycan 2.

[0117] Chondroitin sulfate proteoglycan 2 (CSPG2) also known as versican, was first identified in hyaline cartilage where it provides mechanical support. Recent studies have identified CSPG2 as a primary component of the extracellular matrix in the CNS. A disaccharide degradation product of CSPG2 has been shown to stimulate microglia to possess increase phagocytic activity without cytotoxic effects. This suggests a role for CSPG2 in immune-related neurodegenerative disorders. In addition, increased CSPGs exhibit growth inhibiting properties and inhibit axonal sprouting within the glial scar. Within the infarct core CSPG2 expression is dramatically increased, resulting in increased cell death and reactive astrocytosis. Several enzymatic processes cleave CSPG2, including the matrix metalloproteinases. The up-regulation of CSPG2 as shown by the methods of this invention (see Examples) suggests there is inhibited axonal growth in the acute phase of ischemic stroke. The human CSPG2 mRNA sequence is publicly available as GenBank Accession No. NM\_004385, the complete sequence of which is shown in FIG. 5A. The human CSPG2 amino acid sequence is publicly available as GenPept Accession No. NP\_004376, the complete sequence of which is shown in FIG. 5B.

[0118] (3) IQ Motif-Containing GTPase Activation Protein 1 (IQGAP1).

[0119] IQ Motif-containing GTPase activating protein 1 (IQGAP1) is an evolutionarily conserved molecule that serves as a scaffold protein and plays a fundamental role in cell polarity. It modulates several cellular activities including cytoskeletal architecture, cell-cell adhesion, transcription and signaling (ERK signaling). Rho-family GTPases, including Cdc42 require IQGAP1 to regulate actin cytoskeleton and produce a gradient of signaling molecules. Cdc42 and IQGAP1 it co-localizes with actin filaments throughout the brain. In addition, increased Cdc42 activity has been implicated in the breakdown of the blood brain barrier (BBB). An up-regulation of IQGAP1 expression as shown by the methods of this invention (see Examples) suggests there is an increase in cellular signaling and transcription in the acute phase of ischemic stroke and IQGAP1 may mediate the disruption of the BBB as a means by which signals from the brain enter the periphery to augment cellular recruitment. The human IQGAP1 mRNA sequence is publicly available as GenBank Accession No. NM\_003870, the complete sequence of which is shown in FIG. 6A. The human IQGAP1 amino acid sequence is publicly available as GenPept Accession No. NP\_003861, the complete sequence of which is shown in FIG. 6B.

[0120] (4) Orosomucoid 1 (ORM1).

[0121] Oromucosid 1 (ORM1) also known as alpha-1 acid glycoprotein is an acute phase protein and increases 2-5 times

during an acute phase response. It has been shown to suppress lymphocyte response to LPS (thereby preventing ongoing tissue damage by neutrophil proteases), decrease platelet aggregation (and thus further platelet recruitment), and enhance cytokine secretion (as possibly part of a feedback mechanism). It exhibits both pro and anti-inflammatory effects and is therefore suggested to play a significant role in immunomodulation. An up-regulation of ORM1 as shown by the methods of this invention (see Examples) suggests a neuroimmune response in acute ischemic stroke mediated by a balance between pro and anti-inflammatory signaling molecules. The human ORM1 mRNA sequence is publicly available as GenBank Accession No. NM\_000607, the complete sequence of which is shown in FIG. 7A. The human ORM1 amino acid sequence is publicly available as GenPept Accession No. NP\_000598, the complete sequence of which is shown in FIG. 7B.

[0122] (5) Arginase 1 (ARG1).

[0123] Arginase-1 (ARG1) is an enzyme induced by T-helper 2 cytokines that metabolizes L-arginine to ornithine and urea and is a critical regulator of nitric oxide (NO) synthesis. Inflammatory stimuli (T-helper 1 cytokines) result in an increased expression of inducible NO synthetase (iNOS) through L-arginine metabolism. It is possible to determine the type of inflammatory response to injury depending on the relative amount of ARG1 and iNOS since both compete for L-arginine. Trauma is associated with an increase activity of ARG1 and a decrease in the level of arginine. In addition recent studies suggest activation of the JAK and STAT pathways induces ARG1 in smooth muscle. Since humoral anti-inflammatory cytokines induce ARG1, the up-regulation of ARG1 as shown by the methods of this invention (see Examples) suggests that the response to acute ischemic stroke favors an innate humoral immune response. The human ARG1 mRNA sequence is publicly available as GenBank Accession No. NM\_000045, the complete sequence of which is shown in FIG. 8A. The human ARG1 amino acid sequence is publicly available as GenPept Accession No. NP\_000036, the complete sequence of which is shown in FIG. 8B.

[0124] (6) Lymphocyte Antigen 96 (LY96).

[0125] Lymphocyte antigen 96 (LY96) also known as MD2 protein, is critical for toll-like receptor 4 (TLR4) activation as an innate response to lipopolysaccharide (LPS) which is the main constituent of gram-negative bacteria. TLR4 activation induces transduction pathways resulting in NF-kappaB expression and subsequent release of pro-inflammatory cytokines (e.g. IL6 and IL8). Interestingly, natural selection has shaped the sequence patterns of TLR genes in primate evolution. However, pathogens and LPS are not the only cause of tissue damage; ischemia is another mechanism. There is accumulating evidence that ischemic tissue damage is recognized at the cell level via receptor-mediated detection of proteins (alarmins) released by dead cells. Therefore there are exogenous pathogen-associated molecular patterns (PAMPs; such as LPS) and endogenous alarmins that elicit similar responses of the innate immune system known as damage associated molecular patterns (DAMPs). The upregulation of LY96 as shown by the methods of this invention (see Examples) suggests that the response to acute ischemic stroke is mediated by the innate immune system and TLR signaling. The human LY96 mRNA sequence is publicly available as GenBank Accession No. NM\_015364, the complete sequence of which is shown in FIG. 9A. The human

LY96 amino acid sequence is publicly available as GenPept Accession No. NP\_056179, the complete sequence of which is shown in FIG. 9B.

[0126] (7) Matrix Metalloproteinase 9 (MMP9).

[0127] Matrix Metalloproteinase 9 (MMP9) is a zinc and calcium dependent endopeptidase responsible for regulation of the extracellular matrix (ECM). Ischemia and reperfusion injury results in oxidative stress that mediates BBB disruption through metalloproteinase activation. MMP9 expression is the result of activated leukocytes (particularly neutrophils), and results in IL1beta activation and initiation of the inflammatory cascade, further contributing to BBB impairment. Up-regulation of MMP9 following acute ischemic stroke as shown by the methods of this invention (see Examples) suggests an increase in proteolytic activity early that may contribute to BBB disruption, which would allow cellular migration and signaling to and throughout the CNS. The human MMP9 mRNA sequence is publicly available as GenBank Accession No. NM\_004994, the complete sequence of which is shown in FIG. 10A. The human MMP9 amino acid sequence is publicly available as GenPept Accession No. NP\_004985, the complete sequence of which is shown in FIG. 10B.

[0128] (8) Carbonic Anhydrase 4 (CA4).

[0129] Carbonic anhydrase IV (CA4) is a zinc enzyme that catalyzes the conversion between carbon dioxide and the bicarbonate ion, thus making it crucial for all physiologic processes involved in cellular respiration and transport. CA4 is a membrane-bound protein found in tissues throughout the body and is found in the brain within the luminal surface of capillary endothelial cells suggesting a role for CA4 in the blood brain barrier as a regulator of CO<sub>2</sub> and bicarbonate homeostasis in the brain. The up-regulation of CA4 as shown by the methods of this invention (see Examples) suggests there is an increase in cellular respiration following acute ischemic stroke that requires an increase in CA4 to convert CO<sub>2</sub> to HCO<sub>3</sub> to maintain pH. The human CA4 mRNA sequence is publicly available as GenBank Accession No. NM\_000717, the complete sequence of which is shown in FIG. 11A. The human CA4 amino acid sequence is publicly available as GenPept Accession No. NP\_000708, the complete sequence of which is shown in FIG. 11B.

[0130] (9) s100 Calcium Binding Protein A12 (s100A12).

[0131] s100 calcium binding protein A12 (s100A12) also known as calgranulin C and EN-RAGE (extracellular newly identified RAGE binding protein) is specifically related to innate immune function. S100A12 is expressed by phagocytes and released at the site of tissue inflammation. It is an endogenous DAMP that turns pro-inflammatory after a release into the extracellular space following brain injury. The Receptor for Advanced Glycation End Products (RAGE) is a member of the immunoglobulin superfamily and is a specific cell surface reaction site for advanced glycation endproducts (AGEs) which increase with advancing age. Interaction between AGEs and RAGE has been linked to chronic inflammation. Once engaged RAGE interaction in inflammatory and vascular cells results in the increased expression of MMPs. The up-regulation of s100A12 as shown by the methods of this invention (see Examples) supports the claim that the response to acute ischemic stroke is largely driven by innate immunity. The human s100A12 mRNA sequence is publicly available as GenBank Accession No. NM\_005621, the complete sequence of which is shown in FIG. 12A. The human s100A12 amino acid sequence is publicly available as

GenPept Accession No. NP\_005612, the complete sequence of which is shown in FIG. 12B.

[0132] The biomarkers described herein, including the 9-biomarkers above, may be used individually, or as part of one or more panels as described hereinafter, and such panels may comprise 2, 3, 4, 5, 6, 7, 8, or 9 or more of individual biomarkers or related markers.

[0133] Particularly preferred markers for the diagnosis and/or prognosis of acute ischemic stroke include (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); and (9) s100 calcium binding protein A12 (s100A12) or markers related (i.e., related markers of these biomarkers as defined herein) thereto.

[0134] Certain preferred marker panels include at least one biomarker, and preferably 2, 3, or 4 biomarkers selected from the group consisting of (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1). In particularly preferred embodiments, one or more of these markers, and preferably 2, 3, or 4 of these biomarkers, may be combined with one or more different markers and preferably 2, 3, 4, 5, 6, 7, 8, or 9 or more different markers or related markers. In preferred embodiments, the invention provide biomarker panels that includes at least one biomarker selected from the group consisting of (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1) and any number of other biomarkers identified herein or which could be identified using the methods of the invention or other previously known biomarkers for acute ischemic stroke or other type of stroke.

[0135] The biomarker panels of the invention can include any suitable biomarker for acute ischemic stroke, or if useful, any other type of stroke or brain injury if such inclusion is deemed suitable by the user of the panel. Biomarkers for ischemic stroke previously disclosed in the art can include those described in Tang et al., "Gene expression in blood changes rapidly in neutrophils and monocytes after ischemic stroke in humans: a microarray study," Journal of Cerebral Blood Flow & Metabolism (2006) 26, pp. 1089-1102; Whiteley et al., "Blood Biomarkers in the Diagnosis of Ischemic Stroke: A Systematic Review," Stroke (2008) 39, pp. 2902-2909; Flex et al. "Proinflammatory Genetic Profiles in Subjects With History of Ischemic Stroke," Stroke (2004) 35, pp. 2270-2275; and Moore et al., "Using Peripheral Blood Mononuclear Cells to Determine a Gene Expression Profile of Acute Ischemic Stroke: A Pilot Investigation," Circulation (2005) 111, pp. 212-221, each of which are incorporated herein in their entireties by reference.

[0136] In other embodiments, the biomarker panels of the invention can comprise biomarkers that are diagnostic to different types of stroke, including acute ischemic stroke, hemorrhagic stroke, transient ischemic attacks, and subarachnoid hemorrhage and other forms of cerebral injury. Examples of other such markers can be found in the art, including, for example, in U.S. Pat. Nos. 7,608,406, 7,622,114, 6,896,872, 7,361,473, 7,358,055, and 6,897,030, each of which is incorporated herein by reference in their entireties.

[0137] In one embodiment, the present invention provides a biomarker panel for detecting or diagnosis from a test sample (e.g., peripheral blood) evidence of risk for or an occurrence of acute ischemic stroke in a patient comprising at least one, or preferably 2, 3 or 4 biomarker(s) selected from the group consisting of chemokine receptor 7 (CCR7) or chondroitin sulfate proteoglycan 2 (CSPG2) or IQ motif-containing GTPase activation protein 1 (IQGAP1) or orosomucoid 1 (ORM1).

[0138] As noted above, the amino acid and corresponding nucleic acid sequences of the biomarkers of the invention are known in the art and can be found in publicly available publications and databases. Exemplary sequences are set forth below in the form of GenBank accession numbers. The nucleic acid and polypeptide accession numbers, respectively, are set forth in parenthesis after each biomarker; (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); and (9) s100 calcium binding protein A12 (s100A12). One of skill in the art will understand that although accession numbers are provided, each biomarker may exist in multiple forms, each of which are encompassed by the invention. For example, variants may exist in which a small number, e.g., 1, 2, 3, 4, 5, 10 or more, nucleotides or amino acid residues are different in relation to the exemplary accession numbers set forth above. However, these variants are intended to be used in the methods of the invention. In addition, "derivatives" of the biomarkers are contemplated.

[0139] As used herein, a "derivative" of a biomarker (or of if encoding nucleic acid molecule) to a modified form of a biomarker of the invention. A modified form of a given biomarker may include at least one amino acid substitution, deletion, or insertion, wherein said modified form retains a biological activity of an unmodified form. An amino acid substitution may be considered "conservative" when the substitution results in similar structural or chemical properties (e.g., replacement of leucine with isoleucine). An amino acid substitution may be "non-conservative" in nature wherein the structure and chemical properties vary (e.g., replacement of arginine with alanine). A modified form of a given biomarker may include chemical modifications, wherein a modified form retains a biological activity of a given biomarker. Such modifications include, but are not limited to, glycosylation, phosphorylation, acetylation, alkylation, methylation, biotinylation, glutamylated glycation, isoprenylation, lipoylation, pegylation, phosphopantetheinylation, sulfation, selenation, and C-terminal amidation. Other modifications include those involving other proteins such as ISGylation, SUMOylation, and ubiquitination. In addition, modifications may also include those involved in changing the chemical nature of an amino acid such as deimination and deamidation.

[0140] The term "biomarker panel" refers to a collection, e.g., 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, preferably 2, 3, 4, 5, 6, 7, 8 or 9 or more biomarkers (e.g., in the form of polypeptides or nucleic acid molecules or the like), that may be analyzed, tested, assayed, probed, measured, quantified, evaluated, an the like, in a generally simultaneous manner. This includes where the individual biomarkers or means for detecting the biomarkers (e.g., oligonucleotide probe or antibody) are situated on a single surface or support medium, e.g., fixed to an

array or a multi-well plate or a filament-based diagnostic system (described in further detail herein), at the time they are analyzed, tested, assayed, probed, measured, quantified, evaluated, an the like. This also includes where the biomarkers are separately analyzed, tested, assayed, probed, measured, quantified, evaluated, and the like, i.e., in separate reaction vessels or reaction environments, such that their assay results are obtained at substantially the same time. The biomarker panel can refer to the constitution of polypeptides or nucleic acid molecules in a biological sample against which are analyzed, tested, assayed, probed, measured, quantified, evaluated, and the like. Alternatively, the biomarker panel can refer a testing device on which isolated or purified biomarkers of the invention (or nucleic acid molecules encoding the biomarkers of the invention or parts thereof or antibodies specific for the biomarkers) are placed to enable the interaction of the biomarker panel with a test biological sample.

[0141] For example, a biomarker panel can include an array of antibodies specific for the biomarkers of the invention, which can be used to detect the presence of the biomarkers in a test biological sample (e.g., peripheral blood). In another embodiment, the biomarker panel can include an array of nucleic acid molecules (e.g., oligonucleotide probes) which are complimentary to mRNA encoding the biomarkers of the invention which may be present in a test sample (e.g., peripheral blood). In yet another embodiment, the biomarker panel can include purified biomarker polypeptides of the invention for the screening of compounds or agents in a test sample or agent library that may interact or bind to the biomarker polypeptides of the invention.

[0142] It will be appreciated that the biomarker panels preferably include at least one biomarker for diagnosis or predicting acute ischemic stroke, including (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); and (9) s100 calcium binding protein A12 (s100A12) or markers related (i.e., related markers of these biomarkers as defined herein) thereto. Preferably, the biomarker panels include at least one of (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1). As noted, the biomarker panels of the invention can be comprised of the biomarker polypeptides themselves, antibodies that are specific to the biomarker polypeptides or even nucleic acid molecules that are complimentary or that recognize corresponding nucleic acid molecules in a sample which encode a biomarker of the invention.

[0143] In addition, the biomarker panels of the invention can include other biomarkers that pertain to other diseases or conditions other than acute ischemic stroke, including any other type of stroke, or other non-stroke condition, in the event a user wishes to test or detect not only acute ischemic strokes, but also other conditions at the same time or using the same panel or set of biomarkers. Examples of other such biomarkers include those related to blood pressure (e.g., A-type natriuretic peptide, C-type antriretic peptide, urokinase II, vasopressin, calcitonin, angiotensin II, adrenomedullin, and endothenlins), coagulation and hemostasis (D-dimer, plasmin, b-thromboglobulin, platelet factor 4, fibrinopeptide A, platelet-derived growth factor, prothrom-

bin, P-selectin and thrombin), acute phase response (C-reactive protein, mannose-binding protein, human neutrophil elastase, inducible nitric oxide synthase, lysophosphatidic acid, malondialdehyde LDL, lipopolysaccharide binding protein) and markers related to inflammation (interleukins, tumor necrosis factor, myeloperoxidase, soluble intercellular adhesion molecule, vascular cell adhesion molecule, monocyte chemotactic protein-1). Such other biomarkers may assist in gaining a better overall clinical picture of the health of the patient and the potential causes of stroke. Such markers can be selected on the basis of the knowledge of one or ordinary skill in the art. Additional examples of such markers can be found in the art, for example, in U.S. Pat. No. 7,608,406, which is incorporated herein by reference.

#### Biomarker Forms

[0144] One of ordinary skill in the art will appreciate that proteins frequently exist in a biological sample in a plurality of different forms. These forms can result from either or both of pre- and post-translational modification. Pre-translational modified forms include allelic variants, splice variants and RNA editing forms. Post-translationally modified forms include forms resulting from proteolytic cleavage (e.g., cleavage of a signal sequence or fragments of a parent protein), glycosylation, phosphorylation, lipidation, oxidation, methylation, cysteinylation, sulphonation and acetylation.

[0145] When detecting or measuring a biomarker of the invention in a sample, the ability to differentiate between different forms of a protein biomarker depends upon the nature of the difference and the method used to detect or measure. For example, an immunoassay using a monoclonal antibody will detect all forms of a protein containing the epitope and will not distinguish between them. However, a sandwich immunoassay that uses two antibodies directed against different epitopes on a protein will detect all forms of the protein that contain both epitopes and will not detect those forms that contain only one of the epitopes.

[0146] In diagnostic assays, the inability to distinguish different forms of a biomarker protein has little impact when the forms detected by the particular method used are equally good biomarkers as any other particular form. However, when a particular form (or a subset of particular forms) of a protein is a better biomarker than the collection of different forms detected together by a particular method, the power of the assay may suffer. In this case, it may be useful to employ an assay method that distinguishes between forms of a protein and that specifically detects and measures a desired form or forms of the protein. Distinguishing different forms of an analyte (e.g., a biomarker) or specifically detecting a particular form of an analyte is referred to as "resolving" the analyte.

[0147] Mass spectrometry is a particularly powerful methodology to resolve different forms of a protein because the different forms typically have different masses that can be resolved by mass spectrometry. Accordingly, if one form of a protein is a superior biomarker for a disease than another form of the biomarker, mass spectrometry may be able to specifically detect and measure the useful form where traditional immunoassay fails to distinguish the forms and fails to specifically detect to useful biomarker.

[0148] One useful methodology combines mass spectrometry with immunoassay. First, a biospecific capture reagent (e.g., an antibody that recognizes the biomarker and other forms of it) is used to capture the biomarker of interest. Preferably, the biospecific capture reagent is bound to a solid

phase, such as a bead, a plate, a membrane or an array. After unbound materials are washed away, the captured analytes are detected and/or measured by mass spectrometry. (This method also will also result in the capture of protein interactors that are bound to the proteins or that are otherwise recognized by antibodies and that, themselves, can be biomarkers.) Various forms of mass spectrometry are useful for detecting the protein forms, including laser desorption approaches, such as traditional MALDI or SELDI, and electrospray ionization.

#### Detection Methods

[0149] Numerous methods and devices are well known to the skilled artisan for the detection and analysis of the biomarkers of the instant invention.

[0150] With regard to polypeptides or proteins in patient test samples, immunoassay devices and methods can be used. See, e.g., U.S. Pat. Nos. 6,143,576; 6,113,855; 6,019,944; 5,985,579; 5,947,124; 5,939,272; 5,922,615; 5,885,527; 5,851,776; 5,824,799; 5,679,526; 5,525,524; and 5,480,792, each of which is hereby incorporated by reference in its entirety, including all tables, figures and claims. These devices and methods can utilize labeled molecules in various sandwich, competitive, or non-competitive assay formats, to generate a signal that is related to the presence or amount of an analyte of interest.

[0151] Additionally, certain methods and devices, such as biosensors and optical immunoassays, may be employed to determine the presence or amount of analytes without the need for a labeled molecule. See, e.g., U.S. Pat. Nos. 5,631,171; and 5,955,377, each of which is hereby incorporated by reference in its entirety, including all tables, figures and claims. One skilled in the art also recognizes that robotic instrumentation including but not limited to Beckman ACCESS®, Abbott AXSYM®, Roche ELECSYS®, Dade Behring STRATUS® systems are among the immunoassay analyzers that are capable of performing the immunoassays taught herein.

[0152] Preferably the biomarkers are analyzed using an immunoassay, although other methods are well known to those skilled in the art (for example, the measurement of marker RNA levels). The presence or amount of a marker is generally determined using antibodies specific for each marker and detecting specific binding. Any suitable immunoassay may be utilized, for example, enzyme-linked immunoassays (ELISA), radioimmunoassays (RIAs), competitive binding assays, and the like. Specific immunological binding of the antibody to the marker can be detected directly or indirectly. Direct labels include fluorescent or luminescent tags, metals, dyes, radionuclides, and the like, attached to the antibody. Indirect labels include various enzymes well known in the art, such as alkaline phosphatase, horseradish peroxidase and the like.

[0153] The use of immobilized antibodies specific for the markers is also contemplated by the present invention. The antibodies could be immobilized onto a variety of solid supports, such as magnetic or chromatographic matrix particles, the surface of an assay place (such as microtiter wells), pieces of a solid substrate material or membrane (such as plastic, nylon, paper), and the like. An assay strip could be prepared by coating the antibody or a plurality of antibodies in an array on solid support. This strip could then be dipped into the test

sample and then processed quickly through washes and detection steps to generate a measurable signal, such as a colored spot.

[0154] The analysis of a plurality of biomarkers may be carried out separately or simultaneously with one test sample. For separate or sequential assay of markers, suitable apparatuses include clinical laboratory analyzers such as the ELECSYS® (Roche), the AXSYM® (Abbott), the ACCESS® (Beckman), the ADVIA® CENTAUR® (Bayer) immunoassay systems, the NICHOLS ADVANTAGE® (Nichols Institute) immunoassay system, etc. Preferred apparatuses or protein chips perform simultaneous assays of a plurality of markers on a single surface. Particularly useful physical formats comprise surfaces having a plurality of discrete, addressable locations for the detection of a plurality of different analytes. Such formats include protein microarrays, or "protein chips" (see, e.g., Ng and Ilag, *J. Cell Mol. Med.* 6: 329-340 (2002)) and certain capillary devices (see e.g., U.S. Pat. No. 6,019,944). In these embodiments each discrete surface location may comprise antibodies to immobilize one or more analyte(s) (e.g., a marker) for detection at each location. Surfaces may alternatively comprise one or more discrete particles (e.g., microparticles or nanoparticles) immobilized at discrete locations of a surface, where the microparticles comprise antibodies to immobilize one analyte (e.g., a marker) for detection. As noted, many protein biochips are described in the art. These further include, for example, protein biochips produced by Ciphergen Biosystems, Inc. (Fremont, Calif.), Packard BioScience Company (Meriden Conn.), Zyomyx (Hayward, Calif.), Phylos (Lexington, Mass.) and Biacore (Uppsala, Sweden). Examples of such protein bio chips are described in the following patents or published patent applications: U.S. Pat. No. 6,225,047; PCT International Publication No. WO 99/51773; U.S. Pat. No. 6,329,209, PCT International Publication No. WO 00/56934 and U.S. Pat. No. 5,242,828, each of which are incorporated by reference.

[0155] Several markers may be combined into one test for efficient processing of a multiple of samples. In addition, one skilled in the art would recognize the value of testing multiple samples (for example, at successive time points) from the same individual. Such testing of serial samples will allow the identification of changes in marker levels over time. Increases or decreases in marker levels, as well as the absence of change in marker levels, would provide useful information about the disease status that includes, but is not limited to identifying the approximate time from onset of the event, the presence and amount of salvageable tissue, the appropriateness of drug therapies, the effectiveness of various therapies as indicated by reperfusion or resolution of symptoms, differentiation of the various types of stroke, identification of the severity of the event, identification of the disease severity, and identification of the patient's outcome, including risk of future events.

[0156] As noted, a biomarker panel consisting of the biomarkers referenced above may be constructed to provide relevant information related to diagnosis of acute ischemic stroke. Such a panel may be constructed using 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, or more or individual markers, but preferably includes at least one of (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); and (9) s100 cal-

cium binding protein A12 (s100A12) or related markers, and more preferably at least one of (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); or (4) orosomucoid 1 (ORM1) or related markers.

[0157] The analysis of a single marker or subsets of markers comprising a larger panel of markers could be carried out by one skilled in the art to optimize clinical sensitivity or specificity in various clinical settings. These include, but are not limited to ambulatory, urgent care, emergency care, critical care, intensive care, monitoring unit, inpatient, outpatient, physician office, medical clinic, and health screening settings. Furthermore, one skilled in the art can use a single marker or a subset of markers comprising a larger panel of markers in combination with an adjustment of the diagnostic threshold in each of the aforementioned settings to optimize clinical sensitivity and specificity.

[0158] In addition, the analysis of the biomarkers of the invention can be carried out by a person of skill who may not necessarily have an expertise with stroke-specific medicine and care, e.g., emergency room or urgent care clinicians, ambulatory clinicians, or any physician not having an expertise in stroke-specific medicine.

[0159] The analysis of markers could be carried out in a variety of physical formats as well. For example, the use of microtiter plates or automation could be used to facilitate the processing of large numbers of test samples. Alternatively, single sample formats could be developed to facilitate immediate treatment and diagnosis in a timely fashion, for example, in ambulatory transport or emergency room settings.

[0160] The present invention also contemplates the use of filament-based detection systems for rapidly detecting the biomarkers of the invention. Filament-based detection systems are known in the art and can be found, for example, in US Published Application No. 2006/012148 A1, and which are described in further detail below.

#### Diagnosis Methods

[0161] The invention provides methods and systems for the identification of one or more biomarkers for the diagnosis of disease, including preferably acute ischemic stroke. One skilled in the art will also recognize that analysis of markers can be performed and the data from the analyses of multiple markers can be combined to form panels of markers to increase the sensitivity and reliability of a diagnosis.

[0162] In developing a panel of markers useful in diagnosis of a particular disease or condition, e.g., acute ischemic stroke, data for a number of potential markers may be obtained from a group of subjects by testing for the presence or level of certain markers. The group of subjects can be divided into two sets, and preferably the first set and the second set each have an approximately equal number of subjects. The first set includes subjects who have been confirmed as having a disease or, more generally, being in a first condition state (e.g., acute ischemic stroke). For example, this first set of patients may be those that have recently had an acute ischemic stroke, or may be those having a specific type of stroke (e.g., thrombotic, embolic, lacunar, hypoperfusion, intracerebral hemorrhage, and subarachnoid hemorrhage types of strokes). The confirmation of this condition state may be made through a more rigorous and/or expensive testing

such as MRI or CT or other instrumentation-based confirmatory test. Hereinafter, subjects in this first set will be referred to as "diseased".

[0163] The second set of subjects are simply those who do not fall within the first set. Subjects in this second set may be "non-diseased;" that is, normal subjects. Alternatively, subjects in this second set may be selected to exhibit one symptom or a constellation of symptoms that mimic those symptoms exhibited by the "diseased" subjects. In the case of neurological disorders, for example, the skilled artisan will understand that neurologic dysfunction is a common symptom in various systemic disorders (e.g., alcoholism, vascular disease, stroke, a specific type of stroke (e.g., thrombotic, embolic, lacunar, hypoperfusion, intracerebral hemorrhage, and subarachnoid hemorrhage types of strokes) autoimmunity, metabolic disorders, aging, etc.).

[0164] Specific neurologic dysfunctions or "stroke-associated symptoms" or "stroke-mimicking symptoms" may include, but are not limited to, pain, headache, aphasia, apraxia, agnosia, amnesia, stupor, confusion, vertigo, coma, delirium, dementia, seizure, migraine insomnia, hypersomnia, sleep apnea, tremor, dyskinesia, paralysis, visual disturbances, diplopia, paresthesias, dysarthria, hemiplegia, hemianesthesia, hemianopia, etc. Patients exhibiting one or more of these symptoms but that have not suffered from a stroke are referred to herein as "stroke mimics." Conditions within the differential diagnosis of stroke include brain tumor (including primary and metastatic disease), aneurysm, electrocution, burns, infections (e.g., meningitis), cerebral hypoxia, head injury (including concussion), stress, dehydration, nerve palsy (cranial or peripheral), hypoglycemia, migraine, multiple sclerosis, peripheral vascular disease, peripheral neuropathy, seizure (including grand mal seizure), subdural hematoma, syncope, and transient unilateral weakness. Preferred markers and marker panels are those that can distinguish acute ischemic stroke from these stroke mimicking conditions.

[0165] The data obtained from subjects in these sets includes levels of a plurality of markers. Preferably, data for the same set of markers is available for each patient. This set of markers may include all candidate markers which may be suspected as being relevant to the detection of a particular disease or condition, e.g., those identified 9-biomarkers of the invention. Embodiments of the methods and systems described herein may be used to determine which of the candidate markers are most relevant to the diagnosis of the disease or condition. The levels of each marker in the two sets of subjects may be distributed across a broad range, e.g., as a Gaussian distribution. However, no distribution fit is required.

[0166] As noted above, a marker often is incapable of definitively identifying a patient as either diseased or non-diseased. For example, if a patient is measured as having a marker level that falls within the overlapping region, the results of the test may not be helpful in diagnosing the patient. An artificial cutoff may be used to distinguish between a positive and a negative test result for the detection of the disease or condition. Regardless of where the cutoff is selected, the effectiveness of the single marker as a diagnosis tool is unaffected. Changing the cutoff merely trades off between the number of false positives and the number of false negatives resulting from the use of the single marker. The effectiveness of a test having such an overlap is often

expressed using a ROC (Receiver Operating Characteristic) curve. ROC curves are well known to those skilled in the art.

[0167] The horizontal axis of the ROC curve represents (1-specificity), which increases with the rate of false positives. The vertical axis of the curve represents sensitivity, which increases with the rate of true positives. Thus, for a particular cutoff selected, the value of (1-specificity) may be determined, and a corresponding sensitivity may be obtained. The area under the ROC curve is a measure of the probability that the measured marker level will allow correct identification of a disease or condition. Thus, the area under the ROC curve can be used to determine the effectiveness of the test.

[0168] As discussed above, the measurement of the level of a single marker may have limited usefulness. The measurement of additional markers provides additional information, but the difficulty lies in properly combining the levels of two potentially unrelated measurements. In the methods and systems according to embodiments of the present invention, data relating to levels of various markers for the sets of diseased and non-diseased patients may be used to develop a panel of markers to provide a useful panel response. The data may be provided in a database such as Microsoft Access, Oracle, other SQL databases or simply in a data file. The database or data file may contain, for example, a patient identifier such as a name or number, the levels of the various markers present, and whether the patient is diseased or non-diseased.

[0169] Next, an artificial cutoff region may be initially selected for each marker. The location of the cutoff region may initially be selected at any point, but the selection may affect the optimization process described below. In this regard, selection near a suspected optimal location may facilitate faster convergence of the optimizer. In a preferred method, the cutoff region is initially centered about the center of the overlap region of the two sets of patients. In one embodiment, the cutoff region may simply be a cutoff point. In other embodiments, the cutoff region may have a length of greater than zero. In this regard, the cutoff region may be defined by a center value and a magnitude of length. In practice, the initial selection of the limits of the cutoff region may be determined according to a pre-selected percentile of each set of subjects. For example, a point above which a pre-selected percentile of diseased patients are measured may be used as the right (upper) end of the cutoff range.

[0170] Each marker value for each patient may then be mapped to an indicator. The indicator is assigned one value below the cutoff region and another value above the cutoff region. For example, if a marker generally has a lower value for non-diseased patients and a higher value for diseased patients, a zero indicator will be assigned to a low value for a particular marker, indicating a potentially low likelihood of a positive diagnosis. In other embodiments, the indicator may be calculated based on a polynomial. The coefficients of the polynomial may be determined based on the distributions of the marker values among the diseased and non-diseased subjects.

[0171] The relative importance of the various markers may be indicated by a weighting factor. The weighting factor may initially be assigned as a coefficient for each marker. As with the cutoff region, the initial selection of the weighting factor may be selected at any acceptable value, but the selection may affect the optimization process. In this regard, selection near a suspected optimal location may facilitate faster convergence of the optimizer. In a preferred method, acceptable weighting coefficients may range between zero and one, and

an initial weighting coefficient for each marker may be assigned as 0.5. In a preferred embodiment, the initial weighting coefficient for each marker may be associated with the effectiveness of that marker by itself. For example, a ROC curve may be generated for the single marker, and the area under the ROC curve may be used as the initial weighting coefficient for that marker.

[0172] Next, a panel response may be calculated for each subject in each of the two sets. The panel response is a function of the indicators to which each marker level is mapped and the weighting coefficients for each marker. One advantage of using an indicator value rather than the marker value is that an extraordinarily high or low marker levels do not change the probability of a diagnosis of diseased or non-diseased for that particular marker. Typically, a marker value above a certain level generally indicates a certain condition state. Marker values above that level indicate the condition state with the same certainty. Thus, an extraordinarily high marker value may not indicate an extraordinarily high probability of that condition state. The use of an indicator which is constant on one side of the cutoff region eliminates this concern.

[0173] The panel response may also be a general function of several parameters including the marker levels and other factors including, for example, race and gender of the patient. Other factors contributing to the panel response may include the slope of the value of a particular marker over time. For example, a patient may be measured when first arriving at the hospital for a particular marker. The same marker may be measured again an hour later or some other time increment later, and the level of change may be reflected in the panel response. Further, additional markers may be derived from other markers and may contribute to the value of the panel response. For example, the ratio of values of two markers may be a factor in calculating the panel response.

[0174] Having obtained panel responses for each subject in each set of subjects, the distribution of the panel responses for each set may now be analyzed. An objective function may be defined to facilitate the selection of an effective panel. The objective function should generally be indicative of the effectiveness of the panel, as may be expressed by, for example, overlap of the panel responses of the diseased set of subjects and the panel responses of the non-diseased set of subjects. In this manner, the objective function may be optimized to maximize the effectiveness of the panel by, for example, minimizing the overlap.

[0175] In a preferred embodiment, the ROC curve representing the panel responses of the two sets of subjects may be used to define the objective function. For example, the objective function may reflect the area under the ROC curve. By maximizing the area under the curve, one may maximize the effectiveness of the panel of markers. In other embodiments, other features of the ROC curve may be used to define the objective function. For example, the point at which the slope of the ROC curve is equal to one may be a useful feature. In other embodiments, the point at which the product of sensitivity and specificity is a maximum, sometimes referred to as the "knee," may be used. In an embodiment, the sensitivity at the knee may be maximized. In further embodiments, the sensitivity at a predetermined specificity level may be used to define the objective function. Other embodiments may use the specificity at a predetermined sensitivity level may be used. In still other embodiments, combinations of two or more of these ROC-curve features may be used.

[0176] It is possible that one of the markers in the panel is specific to the disease or condition being diagnosed. When such markers are present at above or below a certain threshold, the panel response may be set to return a "positive" test result. When the threshold is not satisfied, however, the levels of the marker may nevertheless be used as possible contributors to the objective function.

[0177] An optimization algorithm may be used to maximize or minimize the objective function. Optimization algorithms are well-known to those skilled in the art and include several commonly available minimizing or maximizing functions including the Simplex method and other constrained optimization techniques. It is understood by those skilled in the art that some minimization functions are better than others at searching for global minimums, rather than local minimums. In the optimization process, the location and size of the cutoff region for each marker may be allowed to vary to provide at least two degrees of freedom per marker. Such variable parameters are referred to herein as independent variables. In a preferred embodiment, the weighting coefficient for each marker is also allowed to vary across iterations of the optimization algorithm. In various embodiments, any permutation of these parameters may be used as independent variables.

[0178] In addition to the above-described parameters, the sense of each marker may also be used as an independent variable. For example, in many cases, it may not be known whether a higher level for a certain marker is generally indicative of a diseased state or a non-diseased state. In such a case, it may be useful to allow the optimization process to search on both sides. In practice, this may be implemented in several ways. For example, in one embodiment, the sense may be a truly separate independent variable which may be flipped between positive and negative by the optimization process. Alternatively, the sense may be implemented by allowing the weighting coefficient to be negative.

[0179] The optimization algorithm may be provided with certain constraints as well. For example, the resulting ROC curve may be constrained to provide an area-under-curve of greater than a particular value. ROC curves having an area under the curve of 0.5 indicate complete randomness, while an area under the curve of 1.0 reflects perfect separation of the two sets. Thus, a minimum acceptable value, such as 0.75, may be used as a constraint, particularly if the objective function does not incorporate the area under the curve. Other constraints may include limitations on the weighting coefficients of particular markers. Additional constraints may limit the sum of all the weighting coefficients to a particular value, such as 1.0.

[0180] The iterations of the optimization algorithm generally vary the independent parameters to satisfy the constraints while minimizing or maximizing the objective function. The number of iterations may be limited in the optimization process. Further, the optimization process may be terminated when the difference in the objective function between two consecutive iterations is below a predetermined threshold, thereby indicating that the optimization algorithm has reached a region of a local minimum or a maximum.

[0181] Thus, the optimization process may provide a panel of markers including weighting coefficients for each marker and cutoff regions for the mapping of marker values to indicators. In order to develop lower-cost panels which require the measurement of fewer marker levels, certain markers may be eliminated from the panel. In this regard, the effective

contribution of each marker in the panel may be determined to identify the relative importance of the markers. In one embodiment, the weighting coefficients resulting from the optimization process may be used to determine the relative importance of each marker. The markers with the lowest coefficients may be eliminated.

[0182] In certain cases, the lower weighting coefficients may not be indicative of a low importance. Similarly, a higher weighting coefficient may not be indicative of a high importance. For example, the optimization process may result in a high coefficient if the associated marker is irrelevant to the diagnosis. In this instance, there may not be any advantage that will drive the coefficient lower. Varying this coefficient may not affect the value of the objective function.

[0183] Individual panel response values may also be used as markers in the methods described herein. For example, a panel may be constructed from a plurality of markers, and each marker of the panel may be described by a function and a weighting factor to be applied to that marker (as determined by the methods described above). Each individual marker level is determined for a sample to be tested, and that level is applied to the predetermined function and weighting factor for that particular marker to arrive at a sample value for that marker. The sample values for each marker are added together to arrive at the panel response for that particular sample to be tested. For a "diseased" and "non-diseased" group of patients, the resulting panel responses may be treated as if they were just levels of another disease marker.

[0184] One could use such a method to define new biomarkers. For example, one may divide stroke subjects and non-stroke subjects as follows: (1) ischemic stroke; (2) hemorrhagic stroke; (3) normals; (4) TIAs; (5) other stroke mimics. One would define a first panel constructed from a plurality of markers as described above, and obtain the panel responses from this first panel for all the subjects. Then, the members of any one of these 5 groups may be compared to the panel responses of the members of any other of these groups to define a function and weighting factor that best differentiates these two groups based on the panel responses. This can be repeated as all 5 groups are compared pairwise. The "markers" used to define a second panel might include any or all of the following as a new "marker": ischemic stroke versus normals as marker 1; hemorrhagic stroke versus normals as marker 2; ischemic stroke versus TIAs as marker 3; hemorrhagic stroke versus TIAs as marker 4; ischemic stroke versus other mimics as marker 5; and hemorrhagic stroke versus other mimics as marker 6.

[0185] Measures of test accuracy may be obtained as described in Fischer et al., *Intensive Care Med.* 29: 1043-51, 2003, and used to determine the effectiveness of a given marker or panel of markers. These measures include sensitivity and specificity, predictive values, likelihood ratios, diagnostic odds ratios, and ROC curve areas. As discussed above, suitable tests may exhibit one or more of the following results on these various measures: at least 75% sensitivity, combined with at least 75% specificity; ROC curve area of at least 0.7, more preferably at least 0.8, even more preferably at least 0.9, and most preferably at least 0.95; and/or a positive likelihood ratio (calculated as sensitivity/(1-specificity)) of at least 5, more preferably at least 10, and most preferably at least 20, and a negative likelihood ratio (calculated as (1-sensitivity)/specificity) of less than or equal to 0.3, more preferably less than or equal to 0.2, and most preferably less than or equal to 0.1.

[0186] As noted, a number of immunoassays or nucleic acid based tests can be used to rapidly detect the presence of the biomarkers of the invention in a biological sample, in particular, when done in the context of the urgent clinical setting. Examples include radioimmunoassays, enzyme immunoassays (e.g. ELISA), immunofluorescence, immunoprecipitation, latex agglutination, hemagglutination, and histochemical tests. A particularly preferred method, however, because of its speed and ease of use, is latex agglutination.

[0187] Latex agglutination assays have been described in Beltz, G. A. et al., in *Molecular Probes: Techniques and Medical Applications*, A. Albertini et al., eds., Raven Press, New York, 1989, incorporated herein by reference. In the latex agglutination assay, antibody raised against a particular biomarker is immobilized on latex particles. A drop of the latex particles is added to an appropriate dilution of the serum to be tested and mixed by gentle rocking of the card. With samples lacking sufficient levels of the biomarkers, the latex particles remain in suspension and retain a smooth, milky appearance. However, if biomarkers reactive with the antibody are present, the latex particles clump into visibly detectable aggregates.

[0188] An agglutination assay can also be used to detect biomarkers wherein the corresponding antibody is immobilized on a suitable particle other than latex beads, for example, on gelatin, red blood cells, nylon, liposomes, gold particles, etc. The presence of antibodies in the assay causes agglutination, similar to that of a precipitation reaction, which can then be detected by such techniques as nephelometry, turbidity, infrared spectrometry, visual inspection, colorimetry, and the like.

[0189] The term latex agglutination is employed generically herein to refer to any method based upon the formation of detectable agglutination, and is not limited to the use of latex as the immunosorbent substrate. While preferred substrates for the agglutination are latex based, such as polystyrene and polypropylene, particularly polystyrene, other well-known substrates include beads formed from glass, paper, dextran, and nylon. The immobilized antibodies may be covalently, ionically, or physically bound to the solid-phase immunoabsorbent, by techniques such as covalent bonding via an amide or ester linkage, ionic attraction, or by adsorption. Those skilled in the art will know many other suitable carriers for binding antibodies, or will be able to ascertain such, using routine experimentation.

[0190] Conventional methods can be used to prepare the antibodies. For example, by using a peptide of a biomarker of the invention, polyclonal antisera or monoclonal antibodies can be made using standard methods. A mammal, (e.g., a mouse, hamster, or rabbit) can be immunized with an immunogenic form of the peptide which elicits an antibody response in the mammal. Techniques for conferring immunogenicity on a peptide include conjugation to carriers or other techniques well known in the art. For example, the peptide can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassay procedures can be used with the immunogen as antigen to assess the levels of antibodies. Following immunization, antisera can be administered and, if desired, polyclonal antibodies isolated from the sera.

[0191] To produce monoclonal antibodies, antibody producing cells (lymphocytes) can be harvested from an immunized animal and fused with myeloma cells by standard

somatic cell fusion procedures thus immortalizing these cells and yielding hybridoma cells. Such techniques are well known in the art, (e.g., the hybridoma technique originally developed by Kohler and Milstein (Nature 256, 495-497 (1975)) as well as other techniques such as the human B-cell hybridoma technique (Kozbor et al., Immunol. Today 4, 72 (1983)), the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al. Monoclonal Antibodies in Cancer Therapy (1985) Allen R. Bliss, Inc., pages 77-96), and screening of combinatorial antibody libraries (Huse et al., Science 246, 1275 (1989)]. Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with the peptide and the monoclonal antibodies can be isolated. Therefore, the invention also contemplates hybridoma cells secreting monoclonal antibodies with specificity for the biomarkers of the invention as described herein.

[0192] In addition, the biomarkers of the invention may be measured by detection and quantification of nucleic acids encoding the biomarkers, e.g., cDNAs corresponding to mRNAs present in the peripheral blood. Such detection methods may be carried out by any suitable means for analyzing nucleic acids, including traditional PCR assays such as cDNA hybridization, Northern blots, or Southern blots. These methods can be carried out using oligonucleotides that hybridize to nucleic acid molecules encoding the polypeptide biomarkers of the invention. One of ordinary skill in the art is fully capable of designing and selecting appropriate oligonucleotide molecules based on the known sequences of the biomarkers as noted above.

#### Compositions of Matter

[0193] In another aspect, this invention provides compositions of matter based on the biomarkers of the present invention.

[0194] In one embodiment, this invention provides biomarkers of this invention in purified form. Purified biomarkers have utility as antigens to raise antibodies. Purified biomarkers also have utility as standards in assay procedures. As used herein, a "purified biomarker" is a biomarker that has been isolated from other proteins and peptides, and/or other material from the biological sample in which the biomarker is found. Biomarkers may be purified using any method known in the art, including, but not limited to, mechanical separation (e.g., centrifugation), ammonium sulphate precipitation, dialysis (including size-exclusion dialysis), size-exclusion chromatography, affinity chromatography, anion-exchange chromatography, cation-exchange chromatography, and metal-chelate chromatography. Such methods may be performed at any appropriate scale, for example, in a chromatography column, or on a biochip.

[0195] Thus, in one embodiment, the present invention provides purified biomarkers of the invention, including (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); and (9) s100 calcium binding protein A12 (s100A12).

[0196] In another embodiment, this invention provides biospecific capture reagents that specifically bind a biomarker of this invention, optionally in purified form. Preferably, a biospecific capture reagent is an antibody. In one embodiment, a biospecific capture reagent is an antibody that binds (1)

chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); or (9) s100 calcium binding protein A12 (s100A12).

[0197] In another embodiment, this invention provides a complex between a biomarker of this invention and biospecific capture reagent that specifically binds the biomarker. In other embodiments, the biospecific capture reagent is bound to a solid phase. For example, this invention contemplates a device comprising bead or chip derivatized with a biospecific capture reagent that binds to a biomarker of this invention and, also, the device in which a biomarker of this invention is bound to the biospecific capture reagent.

[0198] In another embodiment, this invention provides a device comprising a solid substrate to which is attached an adsorbent, e.g., a chromatographic adsorbent, to which is further bound a biomarker of this invention.

#### Kits and Detection Systems (e.g., Filament-Based Detection Systems)

[0199] In another embodiment the invention provides kits for diagnosing acute ischemic stroke in a patient. Depending on how the kit is to be operated, the kit may include one or more biomarker polypeptides of the invention, including preferably (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); or (9) s100 calcium binding protein A12 (s100A12). In addition, the kit may include antibodies that specifically bind to any of the biomarker polypeptides of the invention. If the kit is to be used to detect nucleic acid molecules that correspond to the biomarkers of the invention, the kit may include oligonucleotide molecules or other nucleic acid molecules for use in the detection of the biomarker DNA or RNA in a sample. Both antibody and antigen preparations should preferably be provided in a suitable titrated form, with antigen concentrations and/or antibody titers given for easy reference in quantitative applications.

[0200] In certain embodiments, the kits may also include an immunodetection reagent or label for the detection of specific immunoreaction between the provided biomarkers and/or antibody, as the case may be, and the diagnostic sample. Suitable detection reagents are well known in the art as exemplified by radioactive, enzymatic or otherwise chromogenic ligands, which are typically employed in association with the antigen and/or antibody, or in association with a second antibody having specificity for first antibody. Thus, the reaction is detected or quantified by means of detecting or quantifying the label. Immunodetection reagents and processes suitable for application in connection with the novel methods of the present invention are generally well known in the art.

[0201] The reagents may also include ancillary agents such as buffering agents and protein stabilizing agents, e.g., polysaccharides and the like. The diagnostic kit may further include where necessary agents for reducing background interference in a test, agents for increasing signal, apparatus for conducting a test, calibration curves and charts, standardization curves and charts, and the like.

[0202] The kit can also comprise a washing solution or instructions for making a washing solution, in which the

combination of the capture reagent and the washing solution allows capture of the biomarker or biomarkers on the solid support for subsequent detection by, e.g., mass spectrometry. The kit may include more than type of adsorbent, each present on a different solid support.

[0203] In a further embodiment, such a kit can comprise instructions for suitable operational parameters in the form of a label or separate insert. For example, the instructions may inform a consumer about how to collect the sample, how to wash the probe or the particular biomarkers to be detected.

[0204] In yet another embodiment, the kit can comprise one or more containers with biomarker samples, to be used as standard(s) for calibration.

[0205] In a more particular aspect, the kit of the invention relates to a rapid biomarker panel for detecting acute ischemic stroke in a patient comprising antibodies to one or more biomarkers of the invention, including preferably (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); or (9) s100 calcium binding protein A12 (s100A12), and including more preferably (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); or (4) orosomucoid 1 (ORM1). Such kits may include other components, as needed and as described above.

[0206] In another particular aspect, the kit of the invention relates to a rapid biomarker panel for detecting acute ischemic stroke in a patient comprising one or more biomarkers of the invention, including preferably (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); or (9) s100 calcium binding protein A12 (s100A12), and including more preferably (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); or (4) orosomucoid 1 (ORM1). Such kits may include other components, as needed and as described above.

[0207] In yet another particular aspect the invention relates to a rapid biomarker panel for detecting acute ischemic stroke in a patient comprising a nucleic acid molecule (e.g., an oligonucleotide) that hybridizes with a nucleic acid molecule encoding one or more biomarkers of the invention (e.g., the mRNA corresponding to or encoding the biomarkers of the invention), including preferably (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); or (9) s100 calcium binding protein A12 (s100A12), and including more preferably (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); or (4) orosomucoid 1 (ORM1). Such kits may include other components, as needed and as described above.

[0208] In a particular embodiment, the present invention contemplates a filament-based rapid diagnostic kit or test system that can be used as a "point-of-care" (POC) diagnostic

system for rapid diagnosis of acute ischemic stroke or risk thereof utilizing the biomarkers of the invention, including, the nine gene panel of the invention (or some subgroup thereof, e.g., at least 2 biomarkers, or 3, or 4, or 5, or 6, or 7, or 8 of the biomarkers of the group including (1) chemokine receptor 7 (CCR7); (2) chondroitin sulfate proteoglycan 2 (CSPG2); (3) IQ motif-containing GTPase activation protein 1 (IQGAP1); (4) orosomucoid 1 (ORM1); (5) arginase 1 (ARG1); (6) lymphocyte antigen 96 (LY96); (7) matrix metalloproteinase 9 (MMP9); (8) carbonic anhydrase 4 (CA4); and (9) s100 calcium binding protein A12 (s100A12)), to rapidly and easily detect or diagnose acute ischemic stroke and/or to distinguish from diagnoses other forms of stroke, TIAs and stroke mimic events. Such a POC test advantageously be operated by anyone irrespective of a person's level of expertise in clinical stroke care and/or testing.

[0209] As used herein, a "filament-based test or diagnostic system" takes the meaning as contemplated in the art, and in particular, in U.S. Published Application No. US 2006/012148 A1, which is incorporated herein by reference in its entirety. In general, filament-based tests utilize either capture antibodies on a polyester filament, or DNA (or other nucleic acid) probe on a gold wire, each of which function as molecular hooks to troll for polypeptides or nucleic acid molecules of interest (e.g., the biomarker polypeptides of the invention, or their corresponding mRNA molecules) in a biological sample, e.g., peripheral blood. For antibody detection of "target" polypeptides (e.g., the biomarker polypeptides of the invention), a filament material immobilized with antibodies specific for the target polypeptides that has been exposed to a test sample (e.g., peripheral blood) is threaded through an array of chambers that carry out the washing and ultimate reporting of the result. For nucleic acid detection (e.g., mRNA encoding the biomarkers of the invention), a filament containing DNA or nucleotide probes bound to the filament (e.g., gold filament) that are specific or hybridize to target nucleic acid molecules in the a biological sample (e.g., mRNA of each biomarker in a sample of peripheral blood), that is passed through various chambers that carry out the washing and reporting of any probe/target interactions that have occurred on the filament surface.

[0210] In one aspect, the filament-based system includes a filament support which provides the opportunity to rapidly and efficiently move probes between different zones (e.g., chambers, such as the washing chamber or a reporting chamber) of an apparatus and still retain information about their location. It also permits the use of very small volumes of various samples—as little as nanoliter volume reactions. The filament may be constructed so that the probes are arranged in an annular fashion, forming a probe band around the circumference of the filament. This also permits bands to be deposited so as to achieve high linear density of probes on the filament.

[0211] The filament may be made of any of a number of different materials. Suitable materials include polystyrene, glass (e.g., fiber optic cores), nylon or other substrate derivatized with chemical moieties to impart desired surface structure (3-dimensional) and chemical activity. The filament may also be constructed to contain surface features such as pores, abrasions, invaginations, protrusions, or any other physical or chemical structures that increase effective surface area. These surface features may, in one aspect, provide for enhanced mixing of solutions as the filament passes through a solution-containing chamber, or increase the number and availability

of probe molecules. The filament may also contain a probe identifier which allows the user to track large numbers of different probes on a single filament. The probe identifiers may be dyes, magnetic, radioactive, fluorescent, or chemiluminescent molecules. Alternatively, they may comprise various digital or analog tags.

[0212] The probes that are attached to the filaments can be any of a variety of biomolecules, including, in particular with respect to this invention, nucleic acid molecules (e.g., oligonucleotides) and antibodies or antibodies fragments. The probes should be capable of binding to or interacting with a target substance of interest (e.g., the polypeptide biomarkers of the invention or their encoding mRNA molecules) in a sample to be tested (e.g., peripheral blood), such that the binding to or interaction is capable of being detected.

[0213] The term "nucleic acid" is well known in the art. A "nucleic acid" as used herein will generally refer to a molecule of DNA, RNA or a derivative or analog thereof, including synthetic molecules. Nucleic acids are also defined as molecules containing a series of naturally-occurring purine or pyrimidine bases. The term "nucleic acid" encompasses the terms "oligonucleotide" and "polynucleotide," each as a sub-genus of the term "nucleic acid." The term "oligonucleotide" refers to a molecule of between about 3 and about 100 nucleobases in length. The term "polynucleotide" refers to at least one molecule of greater than about 100 nucleobases in length.

[0214] These definitions generally refer to both single-stranded and double-stranded molecules, the latter being substantially or fully complementary to each other. A nucleic acid may even encompass a triple-stranded molecule. As used herein, a single stranded nucleic acid may be denoted by the prefix "ss," a double stranded nucleic acid by the prefix "ds," and a triple stranded nucleic acid by the prefix "ts."

[0215] The probes can also be proteinaceous materials, e.g., polypeptides or polypeptide fragments of the biomarkers of the invention. In another embodiment, the probe may be a proteinaceous compound. There are wide variety of protein-protein interactions; however, proteins also bind nucleic acids, metals and other non-proteinaceous compounds (e.g., lipids, hormones, transmitters). Some other examples of protein that may be used as either targets or probes include, but are not limited to, antibodies, enzymes, receptors, and DNA- or RNA-binding proteins.

[0216] In various embodiments, it may desirable to label probe or target molecules. Examples of labels include paramagnetic ions, radioactive isotopes; fluorochromes, NMR-detectable substances, and X-ray imaging compounds.

[0217] Paramagnetic ions include chromium (III), manganese (II), iron (III), iron (II), cobalt (II), nickel (II), copper (II), neodymium (II), samarium (III), ytterbium (III), gadolinium (III), vanadium (II), terbium (III), dysprosium (III), holmium (III) and/or erbium (III), with gadolinium being particularly preferred. Ions useful in other contexts, such as X-ray imaging, include but are not limited to lanthanum (III), gold (III), lead (II), and especially bismuth (III).

[0218] Radioactive isotopes include <sup>14</sup>-carbon, <sup>15</sup>chromium, <sup>36</sup>-chlorine, <sup>57</sup>cobalt, and the like may be utilized. Among the fluorescent labels contemplated for use include Alexa 350, Alexa 430, AMCA, BODIPY 630/650, BODIPY 650/665, BODIPY-FL, BODIPY-R6G, BODIPY-TMR, BODIPY-TRX, Cascade Blue, Cy3, Cy5,6-FAM, Fluorescein Isothiocyanate, HEX, 6-JOE, Oregon Green 488, Oregon Green 500, Oregon Green 514, Pacific Blue, REG, Rhodamine Green, Rhodamine Red, Renographin, ROX,

TAMRA, TET, Tetramethylrhodamine, and/or Texas Red. Enzymes (an enzyme tag) that will generate a colored product upon contact with a chromogenic substrate may also be used. Examples of suitable enzymes include urease, alkaline phosphatase, (horseradish) hydrogen peroxidase or glucose oxidase. Preferred secondary binding ligands are biotin and/or avidin and streptavidin compounds. The use of such labels is well known to those of skill in the art and are described, for example, in U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149 and 4,366,241; each incorporated herein by reference.

[0219] The chambers of the filament-based system can be of any suitable design or function, including processing chambers, pretreatment chambers, wash chambers and amplification chambers.

[0220] A variety of different types of chambers may be used in accordance with the present invention. The invention contemplates a processing chamber containing putative target molecules for binding. It also is possible, where convenient, to have a series of processing chambers that are connected by means other than the filament. For example, one may wish to "recycle" target solution by moving it from one chamber to another for reuse. A processing chamber may also be reused in the sense that the filament may be passed through a given the chamber more than once.

[0221] The present invention may utilize multiple processing in such chambers where different target solutions included therein. Thus, a single filament can be utilized for multiple reactions in a single "run." If a large number of reactions are to be run, a series of processing chambers may be utilized that can quickly be emptied, rinsed, and filled with new target solutions. Thus, one can image an apparatus with three processing chambers A, B and C, where after a filament passes through each chamber, the chambers can be emptied and refilled with new target solutions, and the movement of the filament is reversed. By repeating this process two more times, a series of four filament passes permit exposure to twelve different target solutions.

[0222] The probes or filaments can also be "pretreated" in such a way as to ensure that the ensuing reaction with the target has a high degree of fidelity, i.e., minimize non-specific attachment. A classic example is of a pretreatment is a "blocking" reaction. Non-specific protein-protein interactions by inhibited by pretreating a substrate with a non-specific protein such as BSA. Similarly, non-specific DNA reactions can be reduced by incubating the probe with a "random" DNA known to lack homology with the probe.

[0223] The filament-based system also contemplates washing chambers to remove non-specifically bound molecules from the probe. Though achieving the same goal as pretreatment, washing takes place after the exposure of probe to target. Typically, wash solutions comprise a buffer similar to that used in the target solution, but lacking the target itself. Occasionally, it will be desirable to alter the chemical properties of the wash solution by, for example, changing the salt concentration or pH.

[0224] The system can also include a signal-amplification chamber capable of recursively amplifying signals relating to binding of targets to probe. There are a variety of mechanisms for accomplishing this. However, a common feature will be the need for one or more chambers which effect the necessary steps to achieve the amplification.

[0225] The above disclosure generally describes the present invention. A more complete understanding can be

obtained by reference to the following specific examples, which are provided for purposes of illustration only and are not intended to limit the scope of the invention.

#### EXAMPLES

[0226] The structures, materials, compositions, and methods described herein are intended to be representative examples of the invention, and it will be understood that the scope of the invention is not limited by the scope of the examples. Those skilled in the art will recognize that the invention may be practiced with variations on the disclosed structures, materials, compositions and methods, and such variations are regarded as within the ambit of the invention.

##### Example 1

###### Identification of Blood Biomarkers for Acute Ischemic Stroke

[0227] Overview.

[0228] The objective of this study was to identify peripheral blood biomarkers for the differential diagnosis of acute ischemic cerebrovascular syndrome (AICS) through gene expression profiling. Peripheral whole blood samples were collected from n=39 patients who were 18 years of age with MRI diagnosed AICS and n=25 control subjects who were Non-stroke neurologically healthy. Total RNA was extracted from whole blood stabilized in Paxgene RNA tubes, amplified, and hybridized to Illumina humanRef-8v2 bead chips. Gene expression was compared in a univariate manner between stroke patients and control subjects using t-test in GeneSpring. The significant genes were tested in a logistic regression model controlling for age, hypertension and dyslipidemia. Inflation of type one error was corrected by Bonferroni and False Discovery Rate. Validation was performed by QRT-PCR using Taqman gene expression assays. A nine gene profile has been identified in the whole blood of AICS patients using gene expression profiling. Five of these nine genes were identified in a previously published expression profiling study of stroke and are therefore likely candidates for AICS diagnosis. Pathway analysis revealed toll like receptor (TLR) signaling as a highly significant canonical pathway present in the peripheral whole blood of AICS patients. This study replicates the findings of a previous expression profiling study of ischemic stroke and therefore supports the claim that gene expression profiling of peripheral whole blood can be used to identify biomarkers of AICS.

[0229] Introduction.

[0230] Stroke is the third leading cause of death in the United States(1) and one of the most common causes of death and disability in industrialized countries (2). Despite significant advances in neuroimaging and acute clinical management that have resulted in greater numbers of patients surviving the initial insult (3), the rate of false-positive diagnoses of ischemic stroke can be as high as 25% (4). The small percentage of patients who actually receive rtPA (3-5%) and the large numbers of patients who leave the hospital with either a diagnosis of transient ischemic attack (TIA) or stroke of undetermined cause reflects the need to identify additional definitive means of stroke diagnosis.

[0231] A novel approach to the study of ischemic stroke is the use of gene expression profiling to discover biomarkers that improve acute stroke diagnosis and classification, identify secondary complications such as blood brain barrier dis-

ruption or aid in the development of novel stroke therapeutics. A secondary advantage of gene expression profiling is its ability to unveil the molecular pathways involved in brain recovery and health and elucidate complex genomic interactions that may play a role in outcome.

[0232] Gene expression profiling has been utilized for the characterization of several neurological and immune disorders (5-7). In the case of cancer tissue samples the method has facilitated the identification and refinement of tumor subtypes (8), distinction between good-prognosis and poor-prognosis tumors (9) and the prediction of response to treatment (10). To date there are two reports of peripheral blood mononuclear cells (PBMCs) (11, 12) and one of peripheral whole blood (13) examining the changes of gene expression in patients following ischemic stroke. These studies provide a novel and sophisticated approach to identifying candidates for stroke diagnosis (14). However, the study designs are not complimentary; nor are the findings. Gene expression profiling of PBMCs captures a significantly smaller proportion of the differential gene expression following stroke compared to that which can be found in whole blood RNA (13). Thus, the purpose of our study was to determine the gene expression profile of peripheral whole blood following acute ischemic stroke in a larger cohort of stroke patients and control subjects, adjusting for common stroke risk factors in an attempt to replicate the findings of previous studies. This data can be used to examine the diagnostic capability of the candidate genes for ischemic stroke and to explore innate immune responses to ischemic stroke.

[0233] Methods.

[0234] This was a prospective case-control gene expression profiling study of peripheral whole blood in ischemic stroke patients. Recruitment was conducted from June 2007 through September 2008. Stroke patients were recruited from an IRB approved NINDS/NIH study at Suburban Hospital, Bethesda Md. after written informed consent was obtained when the following inclusion criteria were met: age 18 years; MRI diagnosed definite Acute Ischemic Cerebrovascular Syndrome (AICS15); and research blood draw within 24 hours from onset of symptoms. Patients with probable/possible AICS and hemorrhage (ruled out by MRI) were excluded from this study. The time of stroke symptom onset was determined as the time the patient was last known to be free of the acute stroke symptoms. Patient evaluations and management were standardized. Thrombolytic therapy with tissue plasminogen activator (rtPA) was given to patients with disabling symptoms within 3 hours from presentation to the clinical stroke team. Pre-morbid deficits were determined by the Modified Rankin Scale (MRS) during the acute clinical assessment for status prior to stroke and 30 days post-stroke and severity of injury was determined by the National Institutes of Health Stroke Scale (NIHSS) at the time of blood draw after stroke. Peripheral whole blood samples were collected from stroke patients into Paxgene blood RNA tubes (PreAnalytiX, Qiagen) within 24 hours from onset of stroke symptoms. Adult control subjects were recruited under a separate NIA/NIH protocol during support group sessions for patients with movement disorders if they were significant others of patients affected by movement disorders and they were neurologically normal per neurologist assessment at the time of enrollment. Clinical demographic data was collected from the stroke patient or significant other and all control subjects by trained neurologists.

[0235] RNA Extraction and Amplification.

[0236] Peripheral whole blood samples were collected from control subjects and stroke patients into Paxgene blood RNA tubes following consent. Paxgene RNA tubes were inverted 8-10 times to ensure RBC lysis, and immediately placed in a -80° C. freezer until RNA extraction. All frozen whole blood specimens were allowed to thaw at room temperature for 24 hours on a rotating bed prior to RNA isolation procedures to ensure complete red blood cell lysis. RNA was extracted from whole blood stabilized in Paxgene tubes in one batch per manufacturer's protocol using the Paxgene Blood RNA extraction Kit (PreAnalytiX, Qiagen). A recent study demonstrated that globin reduction does not increase the number of differentially expressed transcripts when hybridizing to HumanRef-8 v2 beadchips and therefore has little impact on probe detection when using the Illumina platform (16). Therefore, globin reduction was not conducted on any sample in this study. Biotinylated, amplified RNA was generated from the Illumina TotalPrep RNA amplification kit (Applied Biosystems). RNA quantity was determined by the Nanodrop and RNA quality was determined by A260/A280 ratio and the presence of two distinct ribosomal bands on gel electrophoresis.

[0237] Array Hybridization.

[0238] RNA was hybridized to Illumina HumanRef-8 v2 expression bead chips. The HumanRef-8 v2 bead chips have the capability to analyze >22,000 probes targeting genes and known alternative splice variants. Stroke patients and control subjects were randomly hybridized to each array for a total of 8 arrays (8 samples per array). Beadarrays were scanned by the Illumina BeadStation 500x and raw intensity values were saved in Illumina's Bead Studio program manager. Sample labeling, hybridization, and scanning were conducted using standard Illumina protocols.

[0239] Statistical Analysis.

[0240] Baseline descriptive statistics for the sample were computed using SPSS (version 15, SPSS, Inc., Chicago, Ill.). Descriptive and frequency analysis was conducted for all demographic and clinical data. Baseline demographic and clinical characteristics were compared between stroke patients and control subjects using chi-square analysis for the following categorical variables: gender, race, presence of comorbidities (hypertension, diabetes and hyperlipidemia), and medication history. Student's t-test was used to analyze the difference between stroke patients and control subjects by age. The level of significance for these descriptive comparisons was established at 0.05 for two-sided hypothesis testing.

[0241] Probe Level Analysis.

[0242] After scanning the beadchip the raw probe expression values were saved into Illumina BeadStudio Gene Expression (GX) Module (version 1, IlluminaR, San Diego Calif.) and GeneSpring GX v10 (Agilent technologies). Probes were filtered in GeneSpring based on signal intensity resulting in a final probe set of 24,424 to be used in analysis. The probe level data were collated using robust multi-array analysis (RMA) normalization with data processing occurring in the following order: 1) Background correction—using perfect match probe information only; 2) Quantile normalization—probe level normalization; and 3) Summarization—expression measure summary done in log base 2 scale and median was used to fit a linear model. Unsupervised clustering was performed without knowledge of class to determine phylogenetic distances between samples to detect potential outliers.

[0243] Gene Expression Level Analysis.

[0244] Data analysis for gene expression was conducted in Illumina BeadStudio Gene Expression Module and verified in GeneSpring. Genes with at least a 2 fold difference in expression were compared in a univariate manner between stroke patients and control subjects through the use of Illumina's custom model (modified t-test) in BeadStudio and t-test comparisons in GeneSpring. The influence of multiple testing was evaluated using the false discovery rate (FDR) and the Bonferroni Family wise error (FWER).

[0245] Logistic Regression for Identification of Off-Target Effects.

[0246] To assess the specificity of the 9 gene profile for ischemic stroke diagnosis, all 9 genes were tested independently in a logistic regression model controlling for age, hypertension and dyslipidemia. The normalized signal intensities for each gene were entered into separate models with age and then hypertension and dyslipidemia as the covariates of interest. A bonferroni corrected p of <0.005 (0.05/9) was considered to be statistically significant.

[0247] Pathway Analysis.

[0248] Data were interpreted through the use of INGENUITY® Systems Pathway analysis (IPA®) (INGENUITY®Systems, www.ingenuity.com). To increase the number of genes included in the analysis, genes with a 1.5 fold difference in expression between stroke patients and control subjects were chosen. The data set that contained gene identifiers and their corresponding expression signal intensities was uploaded into the INGENUITY® systems program. The gene list was compared one-by-one to the Canonical Pathways stored in the INGENUITY® systems knowledge base; pathways with significant p-values ( $p<0.05$ ) were identified. The p-value measures how likely genes from the gene list participate in the function described in the specific pathway. The INGENUITY® systems software queried the INGENUITY® knowledge base and generated a set of networks with a network size of 35 genes/gene products. A score, which was derived from a p-value, was generated for each network according to the fit of the set of significant genes. Scores of 2 or higher were considered to have at least a 99% confidence of not being generated by chance alone. Biological functions were then calculated and assigned to each network. The significance of the association between the data set and the canonical pathway was measured in two ways: 1) a ratio of the number of significant genes that mapped to the canonical pathway (the number of molecules in a given pathway that meet the 1.5 fold cut off, divided by the total number of molecules that make up that pathway); and 2) A right tailed Fisher's exact test to calculate a p-value determining the probability that the association between the genes in the dataset and the canonical pathway is not explained by chance alone.

[0249] Polymerase Chain Reaction Validation.

[0250] cDNA was generated from total RNA per manufacturer's protocol. (Invitrogen, SuperScript III first strand synthesis kit). Quantitative real-time polymerase chain reaction (QRT-PCR) using Taqman gene expression probes was used to validate the significant transcripts identified in this study that overlapped with previous findings and one novel gene identified in this study. An endogenous control gene with a constant expression level between stroke patients and control subjects (Beta-actin) based on the microarray data was used to normalize the relative expression of chosen genes. When

using Taqman gene expression assays the comparative CT method for determining relative fold change correlates well with expected fold change values (17). Therefore relative fold change differences between stroke patients and control subjects were calculated using the delta CT method (18). Validation was determined positive if the relative fold change in expression was in the same direction as what was identified with the microarray results, t-test analysis revealed significance and there was a positive correlation between QRT-PCR and microarray results.

bid neurological deficits. Severity of stroke was mild with a median baseline National Institutes of Health Stroke Scale score (NIHSS) of 3 with a range from 0-23 and a hospital discharge NIHSS median of 0 with a range from 0-10. There was no difference by race or gender between the stroke patient and control subject groups. However, stroke patients were significantly older than control subjects ( $t=-4.03$ ;  $p=0.000$ ) and stroke patients were more likely to have the presence of comorbidities for which they were receiving medication. See Table 1, below.

TABLE 1

	Univariate associations between stroke patients and control subjects				
	Total Sample	Stroke n = 39 (61.9%)	Control n = 24 (38.1%)	Statistic/df	p value
Gender (% female)	36 (57.1%)	22 (56.4%)	14 (58.3%)	$\chi^2$ 0.02/1	0.883
Mean age, years	68.1 ± 14.02	73.1 ± 14	59.9 ± 9.73	$t$ 4.4/61	0.000
Hypertension	32 (50.8%)	25 (64.1%)	7 (29.2%)	$\chi^2$ 6.6/1	0.010
Diabetes	13 (20.6%)	11 (28.2%)	2 (8.3%)	$\chi^2$ 3.3/1	0.068
Dyslipiderma	18 (28.6%)	18 (46.2%)	0	$\chi^2$ 14.9/1	0.000
Atrial Fibrillation	6 (9.5%)	6 (15.4%)	0	$\chi^2$ 3.9/1	0.048
Myocardial Infarction	6 (9.5%)	6 (15.4%)	0	$\chi^2$ 3.9/1	0.048
Previous Ischemic Stroke	8 (12.7%)	6 (15.4%)	2 (8.3%)	$\chi^2$ 0.7/1	0.414
Previous or Current smoker	30 (47.6%)	15 (38.5%)	15 (62.5%)	$\chi^2$ 7.8/2	0.020
Hypertension Medication	37 (47.6%)	29 (74.4%)	8 (33.3%)	$\chi^2$ 10.3/1	0.001
Diabetes Medication	8 (12.7%)	7 (17.9%)	1 (4.2%)	$\chi^2$ 2.66/1	0.103
Cholesterol Medication	22 (34.9%)	17 (43.6%)	5 (20.8%)	$\chi^2$ 3.4/1	0.066
Anticongulant or antiplatelet	21 (33.3%)	20 (51.3%)	1 (4.2%)	$\chi^2$ 14.8/1	0.000
Family history of Stroke	19 (30.2%)	15 (38.5%)	4 (16.7%)	$\chi^2$ 3.6/1	0.169

[0251] Results.

[0252] Clinical Characteristics.

[0253] A total of 92 subjects (67 stroke patients and 25 control subjects) were recruited to address the aims of the study. Of the 67 stroke patients enrolled, 39 stroke patients received a diagnosis of definite AICS15 with an acute blood draw within 24 hours from onset of symptoms. Using the TOAST (Trial of ORG 10172 in Acute Stroke Treatment) subtype criterion 43.6% (n=17) of the causes were classified as cardioembolic stroke; 28.2% (n=11) were of undetermined cause; 12.8% (n=5) were large artery embolus/thrombosis and the remaining 15.4% (n=6) were small vessel or other cause.

[0254] The presence of comorbidity was prevalent in the stroke patient group, with 64% having a history of hypertension, 28% with a history of diabetes, 15.4% with a history of prior stroke, and 38% with a smoking history (previous or current). The mean time from symptom onset to acute blood draw (baseline RNA profile) was 10:06 hours±6:31. Nine (23.1%) of the patients received rtPA, of which only one patient had their blood drawn before rtPA administration. Stroke patients had a median pre stroke Modified Rankin Scale (MRS) score of zero, implying the absence of premor-

[0255] Array Quality Control.

[0256] Total RNA purified using the Paxgene system was highly pure, with A260/A280 values between 1.9 and 2.2 and RNA yields >1-2 µg from 2.5 ml of peripheral whole blood. Hybridization controls were appropriate for low, medium and high. Negative control, background, and noise signals were low (<200) across all bead arrays and housekeeping and biotin signals were consistently high (>20,000). The average signal for internal controls across the arrays was similar. The control plots were consistent with high quality data.

[0257] Nine Gene Profile for Stroke.

[0258] All analyses were conducted first in Illumina BeadStudio Gene Expression (GX) Module (version 1, Illumina®, San Diego Calif.) and then in GeneSpring GX v10 (Agilent technologies) to verify the findings. Unsupervised clustering of samples revealed two outliers in the dataset (one stroke patient and one control subject). The outlying control subject was removed. It could not be determined if the patient sample was an outlier because of technical or biological variability and to keep credibility of the dataset, the patient outlier was kept in the analysis resulting in a final total of 39 stroke patients and 24 control subjects. BeadStudio identified 344 genes with a 1.5 fold difference in expression with a Diff score >13 (corrected p<0.05) between stroke patients and control subjects. There were 19 genes with a 2 fold difference in expression with a Diff score >13 (corrected p<0.05). See Supplemental Table 1, below.

SUPPLEMENTAL TABLE 1

Beadstudio 2 fold, p < 0.05 Gene List, 19 Genes		
Gene	Diff Score*	Regulation
ARG1	44.13	up
BNIP3L	25.68	up
CA4	40.43	up
CCR7	-71.37	down
CEACAM6	16.8	up
CEACAM8	15.99	up
CSPG2	53.25	up
ECHDC3	17.85	up
FKBP5	15.1	up
IQGAP1	70.95	up
KCTD12	14.99	up
LY96	49.69	up
MMP9	30.01	up
OLFM4	18.97	up
ORM1	17.00	up
PDK4	102.71	up
S100A12	39.53	up
SDPR	41.33	up
TPST1	26.83	up

\*Diff score: Using the Illumina Custom Algorithm a Diff score is calculated from the p-value of significance. {DiffScore = (10sgn(|cond - ref|)log10(p))} For each gene, Diff scores of corresponding probes are averaged and concordance between the probes is reported. For a p-value of 0.05, Diff score = ±13

[0259] Results were comparable in GeneSpring with 355 genes having a 1.5 fold difference in expression (corrected p<0.05) and 16 genes with 2 fold difference in expression between stroke patients and control subjects (corrected p<0.05). See Supplemental Table 2, below.

SUPPLEMENTAL TABLE 2

GeneSpring 2 fold, p < 0.05 Gene List, 16 Genes		
Gene	p-value*	Regulation
ACSL1	4.03e-04	up
AKAP7	0.001	up
APOBEC3A	0.03	up
ARG1	2.84E-07	up
CA4	2.00E-04	up
CCR7	4.37E-05	down
CRISPLD2	4.74E-06	up
CSPG2	3.45E-05	up
FCGR3B	0.024	up
FOLR3	9.23E-04	up
IQGAP1	7.97E-07	up
LY96	0.001	up
MMP9	1.11E-05	up
ORM1	0.006	up
PADI4	4.70E-06	up
S100A12	3.87E-04	up

\*Bonferroni Family wise error (FWER) corrected p-value

Legend:

ACSL1 = acyl-CoA synthetase long-chain family member 1;

AKAP7 = A-kinase anchor protein 7;

APOBEC3A = apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3A;

ARG1 = Arginase 1;

CA4 = Carbonic anhydrase 4;

CCR7 = Chemokine receptor 7;

CRISPLD2 = cysteine-rich secretory protein LCCL domain containing 2;

CSPG2 = Chondroitin sulfate proteoglycan 2;

FCGR3B = Fc fragment of IgG, low affinity IIIb, receptor (CD16b);

FOLR3 = folate receptor 3;

IQGAP1 = IQ motif containing GTPase activating protein 1;

LY96 = Lymphocyte antigen 96;

MMP9 = Matrix metalloproteinase 9;

ORM1 = orosomucoid 1;

PADI4 = peptidyl arginine deiminase, type IV;

S100A12 = S100 calcium binding protein A12

[0260] After comparison between the findings of the two statistical programs, there were 9 genes significantly different between stroke patients and control subjects with at least a 2 fold difference in expression and corrected p<0.05 the same across the two statistical packages. See FIG. 1.

[0261] It is important to note that five of these 9 genes were also found to be significant in the first whole blood gene expression profiling study of stroke (ARG1; CA4; LY96; MMP9; S100A12) (13). See Table 2, below.

TABLE 2

Comparison of Gene List to Previous Study				
Gene	p value	Fold Change	p value	Fold Change
ARG1	2.84E-07	3.2	5.03E-04	3.8
CA4	2.0E-04	2.1	3.54E-05	2.2
LY96	0.001	2.2	3.67E-03	2.1
MMP9	1.11E-05	2.6	3.54E-05	3.2
s100A12	3.87E-04	2.4	2.59E-04	2.2

This Study  
Tang et al 2006

[0262] Regression for Identification of Off-Target Effects.

[0263] Given the significant difference between stroke patients and control subjects by age, a logistic regression was performed to determine if the 9 genes in the identified profile were preferentially regulated by age. After controlling for the effects of age and correcting for multiple testing (p=9/0.05), s100A12 fell out of the model (p=0.014). See Table 3, below.

TABLE 3

Logistic Regression Controlling for Age			
Predictor	Odds Ratio	95% CI for OR	p value
ARG1	0.061	(0.012-0.312)	0.001
CA4	0.167	(0.058-0.482)	0.001
CCR7	7.986	(2.229-28.61)	0.001
CSPG2	0.223	(0.086-0.628)	0.004
IQGAP1	0.061	(0.011-0.339)	0.001
LY96	0.361	(0.186-0.701)	0.003
MMP9	0.131	(0.041-0.422)	0.001
ORM1	0.033	(0.156-0.704)	0.004
s100A12	0.444	(0.233-0.846)	0.014

Note:

p < 0.005 statistically significant after Bonferroni correction

[0264] An additional logistic regression analysis revealed that ARG1 (p=0.002), CA4 (p=0.002), CCR7 (p=0.005), CSPG2 (p=0.003), IQGAP1 (p=0.003), and MMP9 (p=0.002) remained significantly associated with stroke diagnosis after controlling for history of hypertension and dyslipidemia.

[0265] Pathway Analysis.

[0266] There were 355 genes eligible for pathway analysis (at least a 1.5 fold difference in expression between stroke patients and control subjects and corrected p<0.05). IPA revealed the five most significant canonical pathways present in the peripheral whole blood RNA of stroke patients were CD28 signaling in T-helper cells (p=4.03E00), nuclear factor of activated T cells (NFAT) in regulation of the immune response (p=4.03E00), dendritic cell maturation (p=3.4E00), toll-like receptor (TLR) signaling (p=3.33E00), and calcium-induced T-lymphocyte apoptosis (p=2.92E00). Supplemental Data. However, there were more genes from our dataset dif-

ferentially expressed in the TLR signaling pathway between stroke patients and control subjects than in any other identified pathway, (with a ratio >2 and p=3.33E00) implying the TLR pathway is the most significant for this dataset. See FIG. 2.

[0267] Taqman Gene Expression Assay Validation.

[0268] Taqman gene expression assays were used to confirm the beadarray results. QRT-PCR reactions were performed using Taqman® gene expression probes (Applied Biosystems) for the genes ARG1, CCR7, LY96, and MMP9 by the 7900HT QRT-PCR system based on availability of RNA. ARG1, LY96, and MMP9 was chosen because they were also significant in the first gene expression profiling study of stroke. CCR7 was the only down-regulated gene for stroke found in our study. QRT-PCR validated significant changes in mRNA levels in all 4 genes. It is important to note that the first gene expression profiling study conducted by Tang et al was not able to validate microarray results by limitations in the availability of RNA; for this same reason we could not validate the entire gene profile.

[0269] Discussion

[0270] A rapid blood test to confirm the diagnosis of ischemic stroke would transform stroke care in the US and across the world. We found that 9 genes were differentially expressed with at least a 2 fold difference between stroke patients and control subjects. Although the major limitation of our study was a younger control group, post-hoc analyses controlling for age and other stroke risk factors supported the primary analysis findings. In addition, the fact that 5 of the 9 genes identified in this study were also found to be significant in the first whole blood gene expression profiling study of ischemic stroke suggests that confounding factors as age, time after stroke and other factors probably did not radically affect our results and confirms the validity of this method for the identification of diagnostic biomarkers in this population.

[0271] Peripheral whole blood gene expression analysis in AICS patients identified 9 genes; 8 of which are up-regulated. These are Arginase 1 (ARG1); carbonic anhydrase 4 (CA4); chondroitin sulfate proteoglycan 2 (CSPG2); IG motif-containing GTPase activation protein 1 (IQGAP1); lymphocyte antigen 96 (LY96); matrix metalloproteinase 9 (MMP9); orosomucoid 1 (ORM1) and s100 calcium binding protein A12 (s100A12) and one down-regulated, the chemokine receptor 7 (CCR7) gene (see FIGS. 4-12 and the Detailed Description for Accession numbers, nucleotide sequences and amino acid sequences for each marker).

#### Comparison to Previous Studies

[0272] The inconsistency between the first three peripheral blood gene expression profiling studies of AICS is most attributable to the different sources of RNA under study. The first human study compared the gene expression levels in peripheral blood mononuclear cells (PBMC) between 20 ischemic stroke patients and 20 control subjects. Using prediction analysis for microarrays (PAM), the group identified a 22 gene panel that classified stroke in a validation cohort with a sensitivity of 78% and a specificity of 80% (12). The subsequent study performed by Tang and colleagues examined the gene expression profiles of whole blood in 15 stroke patients at 3 hours, 5 hours, and at 24 hours after ischemic stroke in comparison to 8 control subjects (13). The study revealed that the majority of the genes induced between 2 and 24 hours following stroke symptoms were induced in neutrophils and monocytes. Nearly all of the genes regulated at 3

hours were also regulated at 5 and 24 hours with greater numbers of genes expressed as time passed. These findings of this study suggested the use of peripheral whole blood gene expression is the most useful for making early diagnosis of ischemic stroke in humans. The third study examined the changes of peripheral PBMC's in patients 24 hours following stroke, and identified one gene with measurable differences between stroke patients and control subjects, phosphodiesterase 4 D (PDE4D), but no genes in common with the first two studies. (11)

[0273] There were 156 genes (11%) with at least a 1.5 fold change in expression coincident between the first two studies (14). However, when comparing the significant gene lists, only 2 genes were identified in both studies: N-acetylneuraminate pyruvate lyase (NPL) and v-ets Erythroblastosis virus E26 oncogene homolog 2 (avian). Tang and colleagues interpreted this lack of replication secondary to the fact that the majority of patients in the Moore et al study were treated with rtPA, the symptom onset time from blood draw (RNA profile assessment) was variable, and different cell populations were examined. Considering this, the group suggested that the changes of peripheral blood gene expression identified in the first 3 hours after ischemic stroke symptom onset prior to rtPA administration would be the most useful for discovering biomarkers for early stroke diagnosis. Thus they reported 18 significant genes with differential expression between stroke patients and control subjects at 3 hours after ischemic stroke symptom onset.

[0274] In this study, with regard to significant genes with at least a 2 fold difference in expression there was one gene coincident with the Moore et al study (CSPG2) and 5 genes overlapping with the Tang et al study (ARG1; CA4; LY96; MMP9; S100A12). Nine patients (23%) in our study received rtPA, a similar proportion to that of Moore's study, and the mean time from symptom onset to blood draw was 10:06 hours; more than the 3 hours suggested by Tang et al (13). It is remarkable that besides these differences 50% of the genes identified by Tang et al have been replicated in our study. This indicates first that the differences between the two first expression studies were due primarily to differences in cell populations under study; and second that expression in the genes found to be coincident with the Tang et al study are not reflecting changes secondary to administration of rtPA as previously considered, but rather are changes associated with ischemic stroke. These findings suggesting a larger time window beyond 3 hours for the identification of biomarkers of stroke diagnosis has tremendous clinical applications as the majority of stroke patients do not come to the hospital within 3 hours from onset of stroke symptoms. Previous proteomic blood biomarker studies of stroke performing poorly beyond 3 hours from stroke symptom onset (19), reinforces gene expression profiling as an alternative method of biomarker identification for stroke diagnosis.

[0275] The control subjects in this present study were matched for race and there was no significant difference in gender between stroke patients and control subjects. Although the stroke patients in our study were older than the control subjects, post-hoc analyses controlling for age and other stroke risk factors (e.g. hypertension and dyslipidemia) supported the primary analysis findings with the exception of s100A12.

[0276] The results of this study and the previous studies, suggest that the relative expression of ARG1, CA4, LY96,

MMP9, and S100A12 taken together have strong evidence for diagnostic capability in acute ischemic stroke.

**[0277]** Aside from the comparability of these findings to the previous gene expression profiling study, MMP9 and various isoforms of S100 at the protein level have been implicated as biomarkers of ischemic stroke. One of our recent publications suggests baseline serum MMP9 may help to predict the occurrence of blood brain barrier disruption following ischemic stroke (20) and high levels of S100 serum protein have been associated with poor outcome following stroke (21), implying they may be useful as prognostic markers.

#### Novel Candidates for Biomarkers of Stroke

**[0278]** In addition to the 5 genes identified by this study and the previous study by Tang et al, we have identified other novel candidates for stroke diagnosis: CCR7, CSPG2, IQGAP1, and ORM1.

#### Chemokine Receptor 7 (CCR7)

**[0279]** CCR7 is a member of the G-coupled chemokine receptor family. Inflammatory chemokines function mainly as chemoattractants for leukocytes, recruiting monocytes, neutrophils and other effector cells from the blood to sites of infection or tissue damage (22). That cytokines could be the driving force in the neuroinflammatory immune response following stroke, is suggested as specific cytokines have shown to exacerbate brain damage (23) and pharmacological chemokine receptor antagonists have been reported to reduce infarct volume in mice (24, 25). In addition chemokines directly affect blood brain barrier (BBB) permeability via alterations in tight junction (TJ) proteins in an in vitro BBB model (co-cultures of endothelial cells and astrocytes) (26). Interestingly chemokines also play an important role in vasculogenesis by augmenting endothelial progenitor cell (EPC) recruitment in ischemic tissues (27, 28).

**[0280]** Most of the knowledge on the involvement of CCR7 in the development of immunity and tolerance in stroke is derived from mouse models, whereas the data on CCR7 function in humans is rather sparse. Increased levels of CCR7 have been found in experimental stroke rat brain at 22 hours following injury, in contrast with a significant reduction of CCR7 mRNA in spinal cord, suggesting a compensatory response to intracranial events (29). In humans, an increase of CCR7+ T cells has been reported in peripheral blood leukocytes (PBL) of ischemic stroke patients with mild to moderate severity 1 week following stroke (30). In this study, down regulation of CCR7 was found in the peripheral blood in the acute phase of ischemic stroke. Differences in the direction of the regulation could be explained as differences in tissue/cell-specific immune response or different immune responses following stroke progression. Another possibility is that our findings are reflecting immune responses after stroke of mild severity, as most of our patients had mild stroke symptoms. One third possibility is the down regulation of CCR7 is a consequence of the increased age of our stroke patients, as age-associated reductions in CCR7 expression has been reported in animal models (30). However, the logistic regression performed correcting for age, suggests that CCR7 is associated with stroke independent of age. If the response of CCR7 is related to age, tissue, cell population, severity of AICS or stage of AICS progression, further studies of CCR7 and stroke should address this issue.

#### Chondroitin Sulfate Proteoglycan 2 (CSPG2)

**[0281]** CSPG2, also known as versican, is a primary component of the extracellular matrix in the central nervous system (CNS). CSPG2 is involved in cell adhesion, proliferation, migration and angiogenesis (31). Several experimental studies have demonstrated elevated expression of CSPGs in response to brain injury (32, 33). Diminished versican expression in brain has been associated with deep cerebral white matter injury in neonatal hypoxic-ischemic rat injury (34). Furthermore, it has been reported that CSPG2 expression is dramatically increased within the infarct core following ischemic stroke, resulting in increased cell death and reactive astrogliosis (35). In a rat model of myocardial infarction, versican expression is upregulated in monocytes and macrophages within the infarcted myocardium (36). To our knowledge, studies of CSPG2 expression in humans are absent with the exception of the first blood gene expression profiling study of stroke (12), which identified CSPG2 as one of the significant genes that classified stroke with a specificity of 80%. Up-regulation of CSPG2 in our dataset is consistent with these findings.

#### IQ Motif-Containing GTPase Activating Protein 1 (IQGAP1)

**[0282]** IQGAP1 is an evolutionarily conserved molecule that serves as a scaffold protein and plays a fundamental role in cell polarity. It modulates several cellular activities including cytoskeletal architecture, cell-cell adhesion, transcription and signaling (ERK signaling) (37). Rho-family GTPases, are a family of small signaling G proteins that require IQGAP1 to regulate actin cytoskeleton to produce a gradient of signaling molecules (38). Experimental evidence suggests that the expression of RhoA increases in aortic and basilar arteries with age (39), therefore RhoA, and indirectly IQGAP1, may play a role in altered vascular responses associated with aging (39). Furthermore, studies in vitro have suggested that leukocyte transmigration and changes in endothelial permeability can be facilitated by RhoA (40). It has been shown that down-regulation of Rho improves endothelial barrier function (41), however long-term inactivation of Rho can lead to loss of intercellular junctions and an increase in endothelial permeability through disrupting VE-cadherin and cell-cell adhesion (42). The up-regulation of IQGAP1 expression in the context of ischemic stroke could suggest an increase in cellular signaling and transcription in the acute phase of ischemic stroke leading to increased permeability of the BBB. IQGAP1 may mediate the disruption of the BBB as a means by which signals from the brain enter the periphery to augment cellular recruitment.

#### Oromucosid 1 (ORM1)

**[0283]** Finally, ORM1 also known as alpha-1 acid glycoprotein is an acute phase protein and increases 2-5 times during an acute phase response. ORM1 has been shown to suppress lymphocyte response to lipopolysaccharides (LPS) (thereby preventing ongoing tissue damage by neutrophil proteases), decrease platelet aggregation (and thus further platelet recruitment), and enhance cytokine secretion as possibly part of a feedback mechanism (43). It exhibits anti-inflammatory effects by inhibiting polymorphonuclear neutrophil activation and increasing the secretion of IL-1 and is therefore suggested to play a significant role in immunomodulation of the acute phase response (44). The up-regulation of ORM1 in the context of ischemic stroke suggests an

acute phase response in ischemic stroke that is similar to trauma, infection and systemic tissue injury.

#### Toll-Like Receptor Signaling

[0284] IPA analysis identified TLR signaling as the most significant canonical pathway present in the peripheral blood of ischemic stroke patients in this dataset. The genes differentially expressed between stroke patients and control subjects identified in the TLR pathway are TLR2, TLR1, FOS, LY96, TLR8 (includes EG:51311), IF2AK2, and IRAK3. See FIG. 3.

[0285] The TLR pathway is a necessary component of the innate immune system. Recent evidence suggests the TLR system as a key player in ischemic preconditioning and therefore may be a novel target for stroke therapeutics (45). Activation of the innate immune response, through TLRs, is a primary component of pro-inflammatory cytokine generation following ischemic brain injury (46). TLRs recognize pathogen associated molecular patterns (PAMPs), a diverse set of stress and injury-induced molecules with highly conserved structures, to initiate the innate immune response to infection. Traditionally the PAMP/TLR activation pathway was primarily implicated in tolerance to endotoxins that alert an organism to intruding pathogens. However, it is becoming increasingly clear that microbial invasion is not the only mechanism by which the TLR pathway becomes activated (47, 48). Products of protein degradation, damaged DNA, fibrinogen and heat shock proteins have emerged as activators of the TLR pathway through a mechanism known as damage associated molecular pattern (DAMPs) recognition (48).

[0286] Several studies have shown that the responsiveness of the adaptive immune system dramatically decreases with age (49) and a pro-inflammatory shift in gene expression occurs with increasing age (50, 49). TLRs are upregulated in the aged mouse brain (51) and specifically, TLR4 is upregulated in the cardiovascular system of aged rodents (50). However, the expression of TLR2 and TLR4 on neutrophils is not affected by age, (52) which has implications for this present study given the RNA profile under study was extracted from peripheral whole blood. On the other hand, TLR2 and TLR4 have been shown to play a significant role in ischemic brain damage, regardless of age (53). Activation of the innate immune response, through TLRs, is a primary component of inflammatory activation following ischemic brain injury, (46) and TLR4 activation has recently been implicated as a negative effector of the innate immune response (54). Down regulation of TLR4 results in a decrease in final infarct volume and better outcome in a mouse middle cerebral artery occlusion (MCAO) model. TLR4 deficient mice are protected by ischemic injury through a down-regulation of inducible nitric oxide synthase (iNOS) production (55, 56). In addition, mice who are TLR4 deficient also have better behavior following MCAO that is preceded by significant psychological stress (57). Human studies have added to this theory of TLR mediated negative outcome and have identified that increased activation of TLR4 following ischemic stroke corresponds to worse clinical outcome (58, 59).

[0287] Strong evidence suggests that stroke and aging are associated with an increase in inflammatory mediators; however, the relationship between stroke and aging on innate immunity, toll like receptor signaling and inflammatory gene expression warrants further investigation. The predominance of these innate and inflammatory immune pathways in our study reinforces their importance in ischemic stroke. At

present, although progress has been made, a better understanding of how these pathways react and respond to ischemic stroke may lead to the emergence of new avenues for therapeutic intervention. The TLR pathway is a promising and worthy target to be considered and studied with more detail.

[0288] The results of this study provide insight into the molecular mechanisms involved in cerebral ischemia. Nine genes were differently expressed with at least a 2 fold difference between stroke patients and control subjects. The study replicates the findings of a previous gene expression study of ischemic stroke (13), and therefore supports the use of gene expression profiling of peripheral whole blood to identify biomarkers of AICS.

#### Example 2

##### Gene Expression Profile for the Diagnosis of Ischemic Stroke

###### Specific Aims Introduction.

[0289] Stroke is the third leading cause of death in the United States (60) and accounts for 10% of deaths worldwide (61). Clinical diagnosis of ischemic stroke is often difficult, complicated by its multiple etiologies and variable clinical presentation. Most hospitals in the United States use CT to initially evaluate patients suspected of having an acute stroke; however CT is less than optimal for identifying acute ischemia (62). The only Food and Drug Administration (FDA) approved treatment for ischemic stroke is recombinant tissue plasminogen activator (rtPA), and rtPA is only approved for use when patients present to the hospital within three hours after onset of symptoms. The downside is that the median time from stroke symptom onset to presentation to the emergency department is 3-6 hours (63). (It is noted that this Example 2 includes a separate listing of references to which the text makes reference to).

[0290] Currently only 3-8% of stroke patients receive rtPA (64). The recent extension of the time window to up to about 4.5 hours is likely to change this proportion only modestly. Where possible, hospitals are moving toward using MRI for acute diagnosis of stroke; however this often requires a dedicated stroke clinical team and is only possible in facilities with 24 hour MRI availability. Quick and definitive diagnosis in the acute care setting is essential to separate stroke from non-stroke, distinguish hemorrhage from ischemia, and identify the potential cause of the infarction, but most importantly to determine eligibility for thrombolytic therapy to begin treatment within the about 3 to 4.5 hour window of opportunity.

[0291] An additional diagnostic measure, such as a serologic blood test or a screen of a panel of markers, would be extremely beneficial in obtaining a definitive diagnosis of acute stroke to help increase the utilization of rtPA, especially in hospitals that are not stroke centers.

[0292] In Example 1, a nine gene panel of biomarkers was identified in the peripheral whole blood of ischemic stroke patients which can be used for acute diagnosis of ischemic stroke (65). The nine genes identified in Example 1 and discussed elsewhere in this application predicted stroke with an accuracy of 95%; which is higher than the diagnostic capability of either MRI (85%) or CT (54%) (62).

[0293] An additional experimental goal is to validate these findings in a larger age matched cohort of stroke patients,

disease control subjects and stroke-free control subjects to further ensure the differential diagnostic capability of the panel.

[0294] This Example outlines a project to validate gene expression profiles for the acute diagnosis of ischemic stroke, which will lead to improved assessment and treatment of patients who experience a stroke. Gene expression profiling is an effective approach to identify genes and pathways that predict a phenotype and clinical outcome.

[0295] The inventors have determined that expression profiling of peripheral whole blood could be used to differentiate stroke from stroke mimic, as well as to predict the clinical trajectory following a neurologic insult.

[0296] The Specific Aims of this Example:

[0297] Specific Aim [1] To validate and replicate a gene expression profile for the diagnosis of ischemic stroke in a larger cohort of ischemic stroke patients, using neurologic disease control patients. The neurologic disease groups include: 1) patients with CT or MRI confirmed acute ischemic stroke (IS) of all severities (n=75); 2) acute (<24 hrs) transient ischemic attack (TIA) patients (n=75); and 3) intracerebral hemorrhage (ICH) patients (n=50).

[0298] Specific Aim [2] To determine if the gene profile is specific for ischemic stroke by comparing the gene profile in neurologic disease patients (IA, TIA and ICH) with patients following acute inflammatory/ischemic stress (myocardial infarction (MI) or age matched normal controls.

[0299] Specific Aim [3] To develop a filament based point of care (POC) test that can be used at the bedside for differential diagnosis of ischemic stroke, which will be based on gene expression profiles validated in Specific Aims [1] and [2].

[0300] A secondary aim of this study is to identify the peripheral blood gene expression profile associated with clinical outcomes at 30 and 90 days in order to determine whether the identified gene activity can be used to predict recovery from ischemic stroke.

#### Background and Significance.

[0301] Gene Expression Profiling in Neurological Disease:

[0302] A novel approach to the study of ischemic stroke is the use of gene expression profiling to discover biomarkers that improve acute stroke diagnosis, identify secondary complications or aid in the development of novel stroke therapeutics. Gene expression profiling has the potential to identify biomarkers for many acute neurological diseases, as well as stratifying risk for patients with common asymptomatic neurological diseases, such as asymptomatic aneurysm and carotid stenosis. A stratification of risk based on a blood gene profile would aid in the decision to treat or not to treat, dramatically improving current treatments.

[0303] Blood Biomarkers for Stroke Diagnosis:

[0304] There has been a substantial attempt to identify blood biomarkers for ischemic stroke in the past; however the task has proven difficult and has not provided successful results. Many potential blood markers of ischemia and inflammation are also found in other conditions that may mimic stroke, which complicates the ability to identify a specific biomarker of stroke. A recent meta-analysis of published blood biomarker studies for stroke revealed significant methodological and design weaknesses in the studies including small sample size, poor references, and poor choice of control subjects and lack of validation. An ideal biomarker panel should distinguish ischemic stroke from stroke mimic

and hemorrhage, be available in small centers without the need for interpretation outside of the facility, and easily accessible.

[0305] Study:

[0306] This Example outlines a validation study to determine a gene expression profile that can be used in the acute care setting for ischemic stroke diagnosis. Example 1 identified nine genes that were differentially expressed with at least a 2 fold difference between stroke patients and control subjects. To further validate the results of Example 1, this Example proposes to recruit age-matched stroke-free controls and a separate group of neurologic disease and inflammatory stress disease controls to determine the differential diagnostic capability of the nine gene panel of the invention (markers identified as Arginase 1 (ARG1); carbonic anhydrase 4 (CA4); chondroitin sulfate proteoglycan 2 (CSPG2); IG motif-containing GTPase activation protein 1 (IQGAP1); lymphocyte antigen 96 (LY96); matrix metalloproteinase 9 (MMP9); orosomucoid 1 (ORM1) and s100 calcium binding protein A12 (s100A12) and one down-regulated, the chemokine receptor 7 (CCR7) gene).

[0307] Filament Based Point of Care (POC) Test:

[0308] Dr. Haselton and his group at Vanderbilt University have developed a filament-based point of care (POC) test for bedside diagnosis of infectious diseases (See U.S. Patent Application No. 20060121481, which is incorporated herein by reference; also see Stone et al., Ann Biomed Eng., 2006 November; 34(70):1778-85; Stone et al., J Biomed Opt., 2006, May-June; 11(62):34012; and Stone et al., Biotechnol. Bioeng., 2005, Sept. 20; 91(65):699-706, each of which are incorporated herein by reference).

[0309] The technology uses DNA beacons on a gold wire to detect RNA in a small biological solution (73). In the prototype design, extraction is performed as illustrated in FIG. 13. Silica-coated magnetic particles are mixed with a patient sample and pulled through processing solutions in the device by an external magnet. Nucleic acid, in this case RNA, is freed from the particle surface in elution buffer. At this point, capture beacons on a gold wire are exposed to the isolated RNA. As the wire is pulled upwards through the device by a small machine, the gold wire passes through further processing chambers. The molecular beacons are then positioned within the top reporting chamber. It is contemplated that this platform could be used in an over-the-counter capacity. As part of this study, a prototype of the POC will be developed and evaluated for the gene profile of the nine biomarkers (or a subset of same) identified and validated in this study. This prototype will then be tested in the field for acute stroke definitive diagnosis in a multi-site clinical trial.

#### Research Design and Methods.

[0310] Research Design:

[0311] A matched case-control, repeated measures design is proposed to validate the nine gene profile of the invention in its capacity to detect ischemic stroke in patients as compared to neurologic disease control patients, acute inflammatory/ischemic stress control patients and stroke-free age-matched control subjects. The diagnostic capability of the panel will be tested to ensure that the profile withstands differential diagnosis of TIA and other neurologic diseases that can mimic stroke in the acute care setting. This study will be crucial in identifying the utility of an RNA based POC test for differential diagnostics of neurological disease.

[0312] Subjects:

[0313] Male and female subjects will be recruited from Mayo Clinic, Jacksonville Fla. if they are over age 18 years, present to the emergency room within 24 hours from onset of symptoms and:

[0314] Ischemic Stroke (IS): Symptoms of acute onset of neurologic dysfunction of any severity consistent with focal brain ischemia and imaging (MRI or CT) or laboratory confirmation of an acute vascular ischemic pathology (74).

[0315] Transient Ischemic Attack (TIA): Symptoms of acute onset of neurologic dysfunction of any severity, that is transient and caused by focal brain, spinal cord, or retinal ischemia, without acute infarction (75).

[0316] Acute myocardial infarction (AMI): Symptoms associated with ST segment elevation MI (STEMI). ECG elevation of ST-segments by >0.1 mV in contiguous leads in patients with ischemic symptoms. Increased cardiac biomarkers (>99th percentile of the upper reference limit for cardiac troponin 1 (TnI), cutoff 0.04 µg/L).

[0317] Intracerebral hemorrhage (ICH): Have symptoms suggestive of intracerebral hemorrhage (ICH) with a diagnostic CT scan

[0318] Stroke-free age-matched control subjects: No major CNS disease or history of acute ischemic stroke or TIA. They answer positively to any item on the Questionnaire for Verifying Stroke-Free Status (QVSS) (17). The sedimentation rate is >19 mm for men or >22 mm for women (77).

[0319] Exclusion Criteria:

[0320] Diagnosis of pre-existing condition, including a chronic inflammatory disorder, such as celiac disease, vasculitis, lupus, chronic obstructive pulmonary disease, irritable bowel disease, arthritis, and psoriasis or pregnancy.

[0321] Study Material and Methods:

[0322] IS, TIA, ICH and AMI patients will be assessed as soon as possible following presentation to the ED (day 0), and at 24 hours after admission. Long term follow-up evaluations will be performed at 30 days following onset of symptoms for the 4 patient groups.

[0323] Stroke-free age-matched control subjects will undergo an evaluation at baseline and will not be followed over time.

[0324] Collection of Venous Blood Samples:

[0325] Peripheral blood samples will be collected on all patients and controls at each evaluation. Samples to be collected include: 5 ml blood in Paxgene blood tubes (for RNA), 8 ml blood in ACD tubes (DNA); and 8 ml blood in EDTA tubes (WBC). All samples will be processed appropriately and frozen at -80 until analysis.

[0326] Human Ref-8 v2 Expression Bead Chips:

[0327] RNA will be extracted from venous whole blood samples according to Paxgene blood RNA protocol. Illumina human Ref-8 v2 expression bead chips will be used for the study and have the capability to analyze 18,631 unique curated genes at once. The multi-sample format allows for up to eight samples to be arrayed in parallel, increasing throughput and decreasing experimental variability. Sample labeling, hybridization, and scanning will be conducted using standard Illumina protocols. QRT-PCR will be used to validate microarray data.

[0328] Clinical Severity and Outcome Measures:

[0329] The National Institutes of Health Stroke Scale Score (NIHSS) is a 15-item assessment tool that provides a quanti-

tative measure of neurologic deficit. It assesses level of consciousness, gaze, visual fields, facial weakness, motor performance of the extremities, sensory deficit, coordination (ataxia), language (aphasia), speech (dysarthria), and hemi-inattention (neglect). For all parameters, a value of 0 is normal; so, the higher the score, the worse the neurological deficit (the highest possible score is 42). The NIHSS will be performed during the baseline assessment and at subsequent evaluations (24 hours and day 30).

[0330] Functional outcome will be determined using the Modified Rankin Scale (MRS) at 30 days post onset of symptoms in all patients. The MRS (78) measures degree of disability following stroke on a 0-6 scale, from no symptoms to death.

[0331] The Barthel Index (BI) (79) is a test of independence and scores the ability of a patient to care for himself and will be determined at 30 days. The values assigned to each item are based on time and amount of physical assistance required. A patient scoring 100 BI is continent, feeds himself, dresses himself, gets up out of bed and chairs, bathes himself, walks at least one block, and can ascend and descend stairs. The total score however, is not as significant or meaningful as the breakdown into individual items, since these indicate where deficiencies are.

[0332] The Extended Glasgow Outcome Scale Score (EGOS) is an 8 level scale ranging from Death to Upper Good Recovery and will be used to determine functional outcome at 30 days (80).

[0333] The Seattle Angina Questionnaire (SAQ) is a reliable instrument that measure five clinically important dimensions of health and will be used to address health-related quality of life outcomes in AMI patients at 30 days post onset of symptoms (81).

[0334] Covariates:

[0335] Potential covariates include severity of injury (determined by the NIHSS and clinical assessment), history of comorbidities (e.g. presence of hypertension), and medication history. A detailed history will be obtained from the patient and/or patient representative during the acute clinical work up. A modified k-prototypes algorithm for clustering biological samples based on simultaneously considering gene expression data and classes of known phenotypic variables will be used during statistical analysis to attempt to control for these covariates (82).

[0336] Power Analysis:

[0337] Sample size calculations have been conducted using PASS: Power Analysis and Sample Size System and JMP software. A 2.0 fold change is generally expected in genes of significance; however a smaller fold change of 1.5 may be used to identify genes not that differentially expressed between similar groups.

[0338] For Specific Aim 1 and Aim 2 group sample sizes of 50 achieve 95% power for each gene to detect a true difference in expression of at least 1.5 with estimated group standard deviations of 1.5 and 1.5 and with a false discovery rate (FDR) of 0.0500 using a two-sided two-sample T-Test. For a single test, the individual test alpha is 0.0002092. The probability of detecting 100 genes with true mean difference in expression >1.5, is 0.00668. Since patients will be matched to controls by age over-recruitment of patients will be conducted to ensure appropriate power to address Specific Aims 1, 2 and 3.

[0339] Therefore a total N of 325 will be recruited. N=75 Acute Ischemic Stroke; N=75 Transient Ischemic Attack;

N=50 Myocardial Infarction; N=75 Stroke-free Volunteers; and N=50 Intracerebral hemorrhage.

[0340] Development of POC Device:

[0341] The POC design includes nucleic acid isolation, target binding, and result reporting. Silica-coated magnetic particles are mixed with a patient sample and injected through the wall of Tygon tubing. In the prototype design the Tygon tubing has been pre-loaded with 100 ul of each of the processing buffers arranged along the length of tubing held in place by surface tension forces. Magnetic beads are entrained by an external magnet and pulled through each of the processing solutions. When the beads reach the liquid-air interface they pass through without entraining the solution. Entry into the next solution proceeds similarly until the entire cloud of particles has passed through all of the process steps (about 2 minutes). The final elution buffer removes the RNA from the bead surface and will be used in downstream processing to detect markers of interest.

[0342] Candidate MB structures will be evaluated using synthetic targets using methods reported in our recent Analyst publication (83). The experimental studies of Aim 3 will be performed using real-time PCR and modifications of a single MB design coupled to the surface of gold nanoparticles. Performance as measured by fluorescence of candidate MB/nanoparticle construct will be compared to a matched control beacon using increasing concentrations of synthetic target alone and in the presence of a non-specific control target and by adding known concentrations of the synthetic target to patient samples before and after the performance of a RNA extraction. Our goal is to identify those features of the molecular beacon structure and its interaction with the gold surface that are the most robust in the presence of any residual patient sample constituents contained in solution with extracted RNA.

[0343] Optimization of the Stem and Loop Structures of the Molecular Beacon ("MB"):

[0344] One molecular beacon contemplated by the invention is presented in FIG. 14. One additional aim of this Example is focus on improving both the capture and reporting functions of MB in the likely patient sample matrix by modifying stem and loop region lengths (84).

[0345] Hydrolytic Amplification:

[0346] To increase the limit of detection, a mixture of MB will be coupled to the gold surface with loop structures complementary to sequences unique to the DNA. To quantify hydrolysis, the extracted samples will be run on an Agilent 2100 Bioanalyzer and the resulting electropherograms will be examined. The signal produced by homogenous MB on a gold surface will be compared to the signal produced by heterogeneous MB coupling on a gold surface, with and without hydrolysis.

[0347] Amplification of Fluorescence Signal Using Multivalent QDs:

[0348] To achieve the necessary level of detection, the incorporation of amplification features into the device design will be examined. The fluorescence signal associated with MB opening is increased by assembling large fluorescent structures through alternating exposure of the capture region of the gold containing these MB to multivalent streptavidin quantum dots and then multivalent biotin quantum dots. The upper limit of fluorescence signal will be determined by an alternating quantum dot coupling method by plotting the average fluorescence as a function of the number of amplification steps.

[0349] Thermal Specificity Studies of MB on Gold.

[0350] A patient extract sample may result in high non-specific MB opening and therefore loss of specificity which could negatively impact the limit of detection. The temperature control features of a real time PCR machine will be used to study molecular beacon fluorescence response to patient sample extracts as a function of temperature to establish the importance of controlling the temperature for maximum signal for obtaining reproducible results; and secondly, identify the optimum temperature for differentiating between targets which exactly match the beacon loop sequence and all other bound species.

[0351] Data Analysis.

[0352] Descriptive Statistics:

[0353] Baseline descriptive statistics for the sample will be computed using SPSS (version 15, SPSS, Inc., Chicago, Ill.). Baseline demographic and clinical characteristics will be compared between IS patients of differing severity and IS and control subjects using chi-square analysis for the following categorical variables: gender, race, presence of comorbidities (HTN, DM, etc.), and medication history. The level of significance for these descriptive comparisons will be established at 0.05 for two-sided hypothesis testing. Hierarchical linear modeling will be used to compare the trajectory of the gene expression response across time after injury by severity of IS.

[0354] Differential Expression Analysis:

[0355] After scanning the beadchip the raw probe expression values will be saved into Illumina BeadStudio Gene Expression (GX) Module (version 1, Illumina®, San Diego Calif.). Specific Aim One: Data analysis for gene expression will be conducted in GeneSpring GX v10 (Agilent technologies) and R© (GNU). Genes with at least a 2 fold difference in expression will be compared in a univariate manner between stroke patients and neurologic disease control patients through the use of ANOVA in GeneSpring. Specific Aim Two: Genes with at least a 2 fold difference in expression will be compared in a univariate manner between stroke patients, acute inflammatory/ischemic stress control patients and stroke-free age-matched control subjects through the use of the ANOVA in GeneSpring. For Aims 1 and 2, the uncorrected probability values will be assigned a cutoff threshold value of significance of <0.05. Inflation of type one error from multiple hypothesis testing will be corrected by the Bonferroni Family wise error (FWER) and a false detection rate of <0.05 after correction will be considered statistically significant.

[0356] Logistic Regression for Identification of Off-Target Effects:

[0357] To assess the specificity of the identified gene profile for ischemic stroke diagnosis, the profile will be tested independently in a logistic regression model controlling for hypertension and dyslipidemia. The normalized signal intensities for each gene will be entered into separate models with age and then hypertension and dyslipidemia as the covariates of interest. A bonferroni corrected p of <0.005 will be considered to be statistically significant.

[0358] Pathway Analysis:

[0359] Data will be interpreted through the use of INGENUITY® Systems Pathway analysis (IPA) (INGENUITY® Systems, www.ingenuity.com). Genes with a 1.5 fold difference in expression between stroke patients and control subjects will be identified. The data set that contains gene iden-

tifiers and their corresponding expression signal intensities will be uploaded into the INGENUITY® systems program.

[0360] Secondary Aim:

[0361] The change in the gene expression profile over time (baseline to 24 hours) will be analyzed in comparison to MRS. The profile obtained in stroke patients from the baseline time point in aim one will be assessed at 24 hours and the change in the profile will be explained as change in the regulation of the gene set from baseline to 24 hours. MRS at 30 days will be divided into a binary variable as good outcome (MRS 0-1) and bad outcome (MRS 3-6). Genes with at least a 2 fold change in expression over time will be compared in a univariate manner between stroke patients with good or bad outcomes through the use of t-tests in GeneSpring.

[0362] Relevance and Innovation:

[0363] This project has the capability to identify biomarkers associated with response to ischemic stroke (IS) and elucidate complex genomic interactions that may play a role in outcome following IS. This study is innovative in that we are proposing one of the first bedside RNA diagnostic tools for neurologic disease. Interdisciplinary collaborations across multiple sectors, such as that proposed in this project, are required to push bench concepts of RNA based diagnostics to patients at the bedside.

### Example 3

#### Validation of 9-Gene Biomarker Panel by Further Bioinformatic and Biostatistical Analysis Using INGENUITY® Systems, Inc.'s IPA® Software Package

[0364] The data of Example 1 was further analyzed using INGENUITY® Systems, Inc.'s IPA® software package to determine the most relevant biological pathways likely to involve the group of biomarkers of the invention, namely, chemokine receptor 7 (CCR7); chondroitin sulfate proteoglycan 2 (CSPG2); IQ motif-containing GTPase activation protein 1 (IQGAP1); orosomucoid 1 (ORM1); arginase 1 (ARG1); lymphocyte antigen 96 (LY96); matrix metalloproteinase 9 (MMP9); carbonic anhydrase 4 (CA4); and s100 calcium binding protein A12 (s100A12), wherein at least one of the biomarkers is chemokine receptor 7 (CCR7); chondroitin sulfate proteoglycan 2 (CSPG2); IQ motif-containing GTPase activation protein 1 (IQGAP1); or orosomucoid 1 (ORM1).

Result A. Analysis of stroke versus control data validated the 9 gene panel of the invention. INGENUITY® pathway analysis identified the following canonical pathways as most likely to involve the 9-gene biomarker panel of the invention:

[0365] PI3K signaling in B lymphocytes; p=4.72E-05; 11/147 (0.075)

[0366] Role of macrophages, fibroblasts and endothelial cells in rheumatoid arthritis; p=1.57E-04; 17/359 (0.047)

[0367] Altered T cell and B cell signaling in rheumatoid arthritis; p=1.9E-4; 8/91 (0.088)

[0368] Toll-like receptor signaling; p=2.6E-04; 6/54 (0.111)

[0369] Primary immunodeficiency signaling; p=4.5-04; 6/63 (0.095)

The pathways identified above, including in particular, the TLR pathway, suggest that the innate immunity response are involved in responding to acute ischemic stroke at an early stage.

Result B. Analysis of stroke over time data verified the change in the TLR pathway and the identification of the CTLA pathway. INGENUITY® pathway analysis validated the following canonical pathways as likely to involve the 9-gene biomarker panel of the invention:

[0370] Gene expression over time: genes 2 fold and p<0.05 different between baseline and follow up: IL8, LY96, SDPR

[0371] Ingenuity pathway analysis data:

[0372] Ephrin receptor signaling 8.5E-04; 14/198 (0.071)

[0373] Dopamine receptor signaling 1.95E-03; 8/93 (0.086)

[0374] CTLA4 signaling in cytotoxic T lymphocytes 2.44E-03; 9/100 (0.09)

[0375] P70S6K signaling 1.39E-02; 9/133 (0.068)

[0376] Regulation of actin-based motility by Rho 1.59E-02; 7/92 (0.076)

The pathways identified above, including in particular, the CTLA4 pathway, suggest that the cell-mediated immune response is involved in responding to acute ischemic stroke at a later stage.

[0377] The use of INGENUITY®'s IPA software is well-known in the art. Reference can be made, for example, to the following publications which cite to the use of IPA software:

[0378] "Mining knowledge and data to discover intelligent molecular biomarkers: Prostate cancer i-Biomarkers." Soft Computing Applications (SOFA), 2010 4th International Workshop on. Pages: 113-118. 15-17 Jul. 2010. Floares, A.; Balacescu, O.; Floares, C.; Balacescu, L.; Popa, T.; Vermesan, O.; Dept. of Artificial Intell., SAIA & OncoPredict, Cluj-Napoca, Romania.;

[0379] "Exploration of a genomic expression and pathway analysis approach to neurocognitive performance: preliminary findings." Neurobehavioral HIV Medicine. July 2010, Volume 2010:2 Pages 23-32. Chad A Bousman, Gursharan Chana, Stephen J Glatt, Sharon D Chandler, Todd May, James Lohr, Ian P Everall, William S Kremen, Ming T Tsuang;

[0380] "Comparison of the performance of two affinity depletion spin filters for quantitative proteomics of CSF: Evaluation of sensitivity and reproducibility of CSF analysis using GeLC-MS/MS and spectral counting." PROTEOMICS—Clinical Applications. Volume 4, Issue 6-7, pages 613-617. July 2010. Silvina A. Fratantoni, Sander R. Piersma, Connie R. Jimenez;

[0381] "Transcriptome profiling and network analysis of genetically hypertensive mice identifies potential pharmacological targets of hypertension." Physiol Genomics. 2010 Jun. 29. [Epub ahead of print] Puig O, Wang I M, Cheng P, Zhou P, Roy S, Cully D, Peters M A, Benita Y, Thompson J, Cai T Q;

[0382] "Osteosarcoma is characterised by reduced expression of markers of osteoclastogenesis and antigen presentation compared with normal bone." Br J. Cancer. 2010 Jun. 29; 103(1):73-81. Epub 2010 Jun. 15. Endo-Munoz L, Cumming A, Sommerville S, Dickinson I, Saunders N A; and

[0383] "P-225 Specific gene expression in human cumulus cells according to oocyte nuclear maturation stages under in vivo maturation: clinical applications." Human Reproduction. 2010 25(Supplement 1):1170-i210. G. Ouandago, S. Assou, D. Haouzi, A. Ferrieres, T. Anahory, J. De Vos and S. Hamamah; and

[0384] "Transcript abundance patterns in Kawasaki disease patients with intravenous immunoglobulin resistance." *Hum Immunol.* 2010 Jun; 20: Fury W, Tremoulet A H, Watson V E, Best B A, Shimizu C, Hamilton J, Kanegaye J T, Wei Y, Kao C, Mellis S, Lin C, Burns JC; each of which are incorporated herein by reference.

## REFERENCES

[0385] The following references are incorporated herein by reference and may be cited to as further background support of this specification.

- [0386] 1. Thom T, Haase N, Rosamond W, Howard V J, Rumsfeld J, Manolio T, Zheng Z J, Flegel K, O'Donnell C, Kittner S, Lloyd-Jones D, Goff D C, Jr, Hong Y, Adams R, Friday G, Furie K, Gorelick P, Kissela B, Marler J, Meigs J, Roger V, Sidney S, Sorlie P, Steinberger J, Wasserthiel-Smoller S, Wilson M, Wolf P. Heart disease and stroke statistics—2006 update: A report from the american heart association statistics committee and stroke statistics subcommittee. *Circulation.* 2006; 113:e85-151
- [0387] 2. McKay J, Mensah G A, ebrary Inc. The atlas of heart disease and stroke. Geneva: World Health Organization; 2005.
- [0388] 3. Kim D, Liebeskind D S. Neuroimaging advances and the transformation of acute stroke care. *Semin Neurol.* 2005; 25:345-361
- [0389] 4. Hemmen T M, Meyer B C, McClean T L, Lyden P D. Identification of nonischemic stroke mimics among 411 code strokes at the University of California, San Diego, stroke center. *J Stroke Cerebrovasc Dis.* 2008; 17:23-25
- [0390] 5. Tajouri L, Fernandez F, Griffiths L R. Gene expression studies in multiple sclerosis. *Curr Genomics.* 2007; 8:181-189
- [0391] 6. Scherzer C R, Eklund A C, Morse L J, Liao Z, Locascio J J, Fefer D, Schwarzschild M A, Schlossmacher M G, Hauser M A, Vance J M, Sudarsky L R, Standaert D G, Growdon J H, Jensen R V, Gullans S R. Molecular markers of early parkinson's disease based on gene expression in blood. *Proc Natl Acad Sci USA.* 2007; 104:955-960
- [0392] 7. Maes O C, Xu S, Yu B, Chertkow H M, Wang E, Schipper H M. Transcriptional profiling of alzheimer blood mononuclear cells by microarray. *Neurobiol Aging.* 2007; 28:1795-1809
- [0393] 8. Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, Hendrix M, Radmacher M, Simon R, Yakhini Z, Ben-Dor A, Sampas N, Dougherty E, Wang E, Marincola F, Gooden C, Lueders J, Glatfelter A, Pollock P, Carpten J, Gillanders E, Leja D, Dietrich K, Beaudry C, Berens M, Alberts D, Sondak V. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature.* 2000; 406:536-540
- [0394] 9. Jones M H, Virtanen C, Honjoh D, Miyoshi T, Satoh Y, Okumura S, Nakagawa K, Nomura H, Ishikawa Y. Two prognostically significant subtypes of high-grade lung neuroendocrine tumours independent of small-cell and large-cell neuroendocrine carcinomas identified by gene expression profiles. *Lancet.* 2004; 363:775-781
- [0395] 10. Chang J C, Wooten E C, Tsimelzon A, Hilsenbeck S G, Gutierrez M C, Elledge R, Mohsin S, Osborne C K, Chamness G C, Allred D C, O'Connell P. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet.* 2003; 362:362-369
- [0396] 11. Grond-Ginsbach C, Hummel M, Wiest T, Horstmann S, Pfleger K, Hergenhahn M, Hollstein M, Mansmann U, Grau A J, Wagner S. Gene expression in human peripheral blood mononuclear cells upon acute ischemic stroke. *J. Neurol.* 2008; 255:723-731
- [0397] 12. Moore D F, Li H, Jeffries N, Wright V, Cooper R A, Jr, Elkahloun A, Gelderman M P, Zudaire E, Blevis G, Yu H, Goldin E, Baird A E. Using peripheral blood mononuclear cells to determine a gene expression profile of acute ischemic stroke: A pilot investigation. *Circulation.* 2005; 111:212-221
- [0398] 13. Tang Y, Xu H, Du X, Lit L, Walker W, Lu A, Ran R, Gregg J P, Reilly M, Pancioli A, Khoury J C, Sauerbeck L R, Carrozzella J A, Spilker J, Clark J, Wagner K R, Jauch E C, Chang D J, Verro P, Broderick J P, Sharp F R. Gene expression in blood changes rapidly in neutrophils and monocytes after ischemic stroke in humans: A microarray study. *J Cereb Blood Flow Metab.* 2006; 26:1089-1102
- [0399] 14. Sharp F R, Xu H, Lit L, Walker W, Pinter J, Apperson M, Verro P. Genomic profiles of stroke in blood. *Stroke.* 2007; 38:691-693
- [0400] 15. Kidwell C S, Warach S. Acute ischemic cerebrovascular syndrome: Diagnostic criteria. *Stroke.* 2003; 34:2995-2998
- [0401] 16. D., N C. Expression profiling of whole blood specimens on illumina beadchips. Expression analysis tech note; www.Expressionanalysis.Com. October 2007
- [0402] 17. Amplification efficiency of taqman gene expression assays, www3.Appliedbiosystems.Com, stock#127ap05-03
- [0403] 18. Livak K J, Schmittgen T D. Analysis of relative gene expression data using real-time quantitative per and the 2(-delta delta c(t)) method. *Methods.* 2001; 25:402-408
- [0404] 19. Laskowitz D T, Kasner S E, Saver J, Remmel K S, Jauch E C. Clinical usefulness of a biomarkerbased diagnostic test for acute stroke: The biomarker rapid assessment in ischemic injury (brain) study. *Stroke.* 2009; 40:77-85
- [0405] 20. Barr T L, Latour L L, Lee K Y, Schaewe T J, Luby M, Chang G S, El-Zammar Z, Alam S, Hallenbeck J M, Kidwell C S, Warach S. Blood-brain barrier disruption in humans is independently associated with increased matrix metalloproteinase-9. *Stroke.* 2009
- [0406] 21. Brea D, Sobrino T, Blanco M, Cristobo I, Rodriguez-Gonzalez R, Rodriguez-Yanez M, Moldes O, Agulla J, Leira R, Castillo J. Temporal profile and clinical significance of serum neuron-specific enolase and s100 in ischemic and hemorrhagic stroke. *Clin Chem Lab Med.* 2009; 47:1513-1518
- [0407] 22. Fernandez E J, Lolis E. Structure, function, and inhibition of chemokines *Annu Rev Pharmacol Toxicol.* 2002; 42:469-499
- [0408] 23. Mines M, Ding Y, Fan G H. The many roles of chemokine receptors in neurodegenerative disorders: Emerging new therapeutical strategies. *Curr Med. Chem.* 2007; 14:2456-2470
- [0409] 24. Takami S, Minami M, Katayama T, Nagata I, Namura S, Satoh M. Tak-779, a nonpeptide cc chemokine receptor antagonist, protects the brain against focal cerebral ischemia in mice. *J Cereb Blood Flow Metab.* 2002; 22:780-784
- [0410] 25. Takami S, Minami M, Nagata I, Namura S, Satoh M. Chemokine receptor antagonist peptide, viral

- mip-ii, protects the brain against focal cerebral ischemia in mice. *J Cereb Blood Flow Metab.* 2001; 21:1430-1435
- [0411] 26. Stamatovic S M, Shakui P, Keep R F, Moore B B, Kunkel S L, Van Rooijen N, Andjelkovic A V. Monocyte chemoattractant protein-1 regulation of blood-brain barrier permeability. *J Cereb Blood Flow Metab.* 2005; 25:593-606
- [0412] 27. Ceradini D J, Kulkarni A R, Callaghan M J, Tepper O M, Bastidas N, Kleinman M E, Capla J M, Galiano R D, Levine J P, Gurtner G C. Progenitor cell trafficking is regulated by hypoxic gradients through hif-1 induction of sdf-1. *Nat. Med.* 2004; 10:858-864
- [0413] 28. Yamaguchi J, Kusano K F, Masuo O, Kawamoto A, Silver M, Murasawa S, Bosch-Marce M, Masuda H, Losordo D W, Isner J M, Asahara T. Stromal cell-derived factor-1 effects on ex vivo expanded endothelial progenitor cell recruitment for ischemic neovascularization. *Circulation.* 2003; 107:1322-1328
- [0414] 29. Offner H, Subramanian S, Parker S M, Afentoulis M E, Vandenbergk A A, Hurn P D. Experimental stroke induces massive, rapid activation of the peripheral immune system. *J Cereb Blood Flow Metab.* 2006; 26:654-665
- [0415] 30. Yan J, Greer J M, Etherington K, Cadigan G P, Cavanagh H, Henderson R D, O'Sullivan J D, Pandian J D, Read S J, McCombe P A. Immune activation in the peripheral blood of patients with acute ischemic stroke. *J. Neuropathol Immunol.* 2009; 206:112-117
- [0416] 31. Zheng P S, Wen J, Ang L C, Sheng W, Viloria-Petit A, Wang Y, Wu Y, Kerbel R S, Yang B B. Versican/pg-m g3 domain promotes tumor growth and angiogenesis. *FASEB J.* 2004; 18:754-756
- [0417] 32. Vorisek I, Hajek M, Tintera J, Nicolay K, Sykova E. Water adc, extracellular space volume, and tortuosity in the rat cortex after traumatic injury. *Magn Reson Med.* 2002; 48:994-1003
- [0418] 33. Beggah A T, Dours-Zimmermann M T, Barras F M, Brosius A, Zimmermann D R, Zurn A D. Lesion-induced differential expression and cell association of neurocan, brevican, versican v1 and v2 in the mouse dorsal root entry zone. *Neuroscience.* 2005; 133:749-762
- [0419] 34. Leonardo C C, Eakin A K, Ajmo J M, Gottschall P E. Versican and brevican are expressed with distinct pathology in neonatal hypoxic-ischemic injury. *J Neurosci Res.* 2008; 86:1106-1114
- [0420] 35. Carmichael S T, Archibeque I, Luke L, Nolan T, Momiy J, Li S. Growth-associated gene expression after stroke: Evidence for a growth-promoting region in peri-infarct cortex. *Exp Neurol.* 2005; 193:291-311
- [0421] 36. Toeda K, Nakamura K, Hirohata S, Hatipoglu O F, Demircan K, Yamawaki H, Ogawa H, Kusachi S, Shiratori Y, Ninomiya Y. Versican is induced in infiltrating monocytes in myocardial infarction. *Mol Cell Biochem.* 2005; 280:47-56
- [0422] 37. Brown M D, Sacks D B. Iggap1 in cellular signaling: Bridging the gap. *Trends Cell Biol.* 2006; 16:242-249
- [0423] 38. Fukata M, Nakagawa M, Kaibuchi K. Roles of rho-family gtpases in cell polarisation and directional migration. *Curr Opin Cell Biol.* 2003; 15:590-597
- [0424] 39. Miao L, Calvert J W, Tang J, Parent A D, Zhang J H. Age-related rhoa expression in blood vessels of rats. *Mech Ageing Dev.* 2001; 122:1757-1770
- [0425] 40. Wojciak-Stothard B, Ridley A J. Rho gtpases and the regulation of endothelial permeability. *Vascul Pharmacol.* 2002; 39:187-199
- [0426] 41. Carabajal J M, Schaeffer R C, Jr. Rhoa inactivation enhances endothelial barrier function. *Am J. Physiol.* 1999; 277:C955-964
- [0427] 42. Hordijk P L, Anthony E, Mul F P, Rentsma R, Oomen L C, Roos D. Vascular-endothelialcadherin modulates endothelial monolayer permeability. *J Cell Sci.* 1999; 112 (Pt 12):1915-1923
- [0428] 43. Hochepied T, Berger F G, Baumann H, Libert C. Alpha(1)-acid glycoprotein: An acute phase protein with inflammatory and immunomodulating properties. *Cytokine Growth Factor Rev.* 2003; 14:25-34
- [0429] 44. Fournier T, Medjoubi N N, Porquet D. Alpha-1-acid glycoprotein. *Biochim Biophys Acta.* 2000; 1482: 157-171
- [0430] 45. Marsh B J, Stevens S L, Hunter B, Stenzel-Poore M P. Inflammation and the emerging role of the toll-like receptor system in acute brain ischemia. *Stroke.* 2009; 40:S34-37
- [0431] 46. Medzhitov R. Toll-like receptors and innate immunity. *Nat Rev Immunol.* 2001; 1:135-145
- [0432] 47. Rubartelli A, Lotze M T. Inside, outside, upside down: Damage-associated molecular-pattern molecules (damps) and redox. *Trends Immunol.* 2007; 28:429-436
- [0433] 48. Bianchi M E. Damps, pamps and alarmins: All we need to know about danger. *J Leukoc Biol.* 2007; 81:1-5
- [0434] 49. Salminen A, Huusko J, Ojala J, Kauppinen A, Kaarniranta K, Suuronen T. Activation of innate immunity system during aging: Nf-kb signalling is the molecular culprit of inflamm-aging. *Ageing Res Rev.* 2008; 7:83-105
- [0435] 50. Csiszar A, Wang M, Lakatta E G, Ungvari Z. Inflammation and endothelial dysfunction during aging: Role of nf-kappab. *J Appl Physiol.* 2008; 105:1333-1341
- [0436] 51. Letiembre M, Hao W, Liu Y, Walter S, Mihajevic I, Rivest S, Hartmann T, Fassbender K. Innate immune receptor expression in normal brain aging. *Neuroscience.* 2007; 146:248-254
- [0437] 52. Fulop T, Larbi A, Douziech N, Fortin C, Guerard K P, Lesur O, Khalil A, Dupuis G. Signal transduction and functional changes in neutrophils with aging. *Aging Cell.* 2004; 3:217-226
- [0438] 53. Marsh B J, Williams-Karnesky R L, Stenzel-Poore M P. Toll-like receptor signaling in endogenous neuroprotection and stroke. *Neuroscience.* 2009; 158:1007-1020
- [0439] 54. Kilic U, Kilic E, Matter C M, Bassetti C L, Hermann D M. Tlr-4 deficiency protects against focal cerebral ischemia and axotomy-induced neurodegeneration. *Neurobiol Dis.* 2008; 31:33-40
- [0440] 55. Ziegler G, Harhausen D, Schepers C, Hoffmann O, Rohr C, Prinz V, Konig J, Lehrach H, Nietfeld W, Tredelenburg G. Tlr2 has a detrimental role in mouse transient focal cerebral ischemia. *Biochem Biophys Res Commun.* 2007; 359:574-579
- [0441] 56. Caso J R, Pradillo J M, Hurtado O, Lorenzo P, Moro M A, Lizasoain I. Toll-like receptor 4 is involved in brain damage and inflammation after experimental stroke. *Circulation.* 2007; 115:1599-1608
- [0442] 57. Caso J R, Pradillo J M, Hurtado O, Leza J C, Moro M A, Lizasoain I. Toll-like receptor 4 is involved in

- subacute stress-induced neuroinflammation and in the worsening of experimental stroke. *Stroke*. 2008; 39:1314-1320
- [0443] 58. Yang Q W, Li J C, Lu F L, Wen A Q, Xiang J, Zhang L L, Huang Z Y, Wang J Z. Upregulated expression of toll-like receptor 4 in monocytes correlates with severity of acute cerebral infarction. *J Cereb Blood Flow Metab*. 2008; 28:1588-1596
- [0444] 59. Urra X, Cervera A, Obach V, Climent N, Planas A M, Chamorro A. Monocytes are major players in the prognosis and risk of infection after acute stroke. *Stroke*. 2009; 40:1262-1268
- [0445] 60. Thom T, Haase N, Rosamond W, Howard V J, Rumsfeld J, Manolio T, Zheng Z J, Flegal K, O'Donnell C, Kittner S, Lloyd-Jones D, Goff D C, Jr, Hong Y, Adams R, Friday G, Furie K, Gorelick P, Kissela B, Marler J, Meigs J, Roger V, Sidney S, Sorlie P, Steinberger J, Wasserthiel-Smoller S, Wilson M, Wolf P. Heart disease and stroke statistics—2006 update: A report from the american heart association statistics committee and stroke statistics subcommittee. *Circulation*. 2006; 113:e85-151
- [0446] 61. McKay J, Mensah G A, ebrary Inc. *The atlas of heart disease and stroke*. Geneva: World Health Organization; 2005.
- [0447] 62. Chalela J A, Kidwell C S, Nentwich L M, Luby M, Butman J A, Demchuk A M, Hill M D, Patronas N, Latour L, Warach S. Magnetic resonance imaging and computed tomography in emergency assessment of patients with suspected acute stroke: A prospective comparison. *Lancet*. 2007; 369:293-298
- [0448] 63. Rosamond W, Flegal K, Furie K, Go A, Greenland K, Haase N, Hailpern S M, Ho M, Howard V, Kissela B, Kittner S, Lloyd-Jones D, McDermott M, Meigs J, Moy C, Nichol G, O'Donnell C, Roger V, Sorlie P, Steinberger J, Thom T, Wilson M, Hong Y. Heart disease and stroke statistics—2008 update: A report from the american heart association statistics committee and stroke statistics subcommittee. *Circulation*. 2008; 117:e25-146
- [0449] 64. Reeves M J, Arora S, Broderick J P, Frankel M, Heinrich J P, Hickenbottom S, Karp H, LaBresh K A, Malarcher A, Mensah G, Moonaw C J, Schwamm L, Weiss P. Acute stroke care in the us: Results from 4 pilot prototypes of the paul coverdell national acute stroke registry. *Stroke*. 2005; 36:1232-1240
- [0450] 65. Barr T L, Conley Y P, Dillman A, Ding J, Warach S, Singleton A, Matarin M. Gene expression profiling in human acute ischemic stroke. *University of Pittsburgh School of Nursing*. 2009; PhD:270
- [0451] 66. Tang Y, Xu H, Du X, Lit L, Walker W, Lu A, Ran R, Gregg J P, Reilly M, Pancioli A, Khouri J C, Sauerbeck L R, Carrozzella J A, Spilker J, Clark J, Wagner K R, Jauch E C, Chang D J, Verro P, Broderick J P, Sharp F R. Gene expression in blood changes rapidly in neutrophils and monocytes after ischemic stroke in humans: A microarray study. *J Cereb Blood Flow Metab*. 2006; 26:1089-1102
- [0452] 67. Basso G, Case C, Dell'Orto M. Diagnosis and genetic subtypes of leukemia combining gene expression and flow cytometry. *Blood Cells, Molecules, and Diseases*. 2007; 39:164-168
- [0453] 68. Andersson A, Ritz C, Lindgren D, Eden P, Lassen C, Heldrup J, Olofsson T, Rade J, Fontes M, Porwit-MacDonald A, Behrendtz M, Hoglund M, Johansson B, Fioretos T. Microarray-based classification of a consecutive series of 121 childhood acute leukemias: Prediction of leukemic and genetic subtype as well as of minimal residual disease status. *Leukemia* 2007; 21:1198-1203
- [0454] 69. Cardoso F, Van't Verr L, Rutgers E, Loi S, Mook S, Piccart-Gebhart M. Clinical application of the 70-gene profile: The mindact trial. *Journal of Clinical Oncology*. 2008; 26:729-735
- [0455] 70. Sharp F R, Xu H, Lit L, Walker W, Apperson M, Gilbert D L, Glauser T A, Wong B, Hershey A, Liu D Z, Pinter J, Zhan X, Liu X, Ran R. The future of genomic profiling of neurological diseases using blood. *Arch Neurol*. 2006; 63:1529-1536
- [0456] 71. Whiteley W, Tseng M C, Sandercock P. Blood biomarkers in the diagnosis of ischemic stroke. A systematic review. *Stroke*. 2008
- [0457] 72. Montaner J, Perea-Gainza M, Delgado P, Ribo M, Chacon P, Rosell A, Quintana M, Palacios M E, Molina C A, Alvarez-Sabin J. Etiologic diagnosis of ischemic stroke subtypes with plasma biomarkers. *Stroke*. 2008; 39:2280-2287
- [0458] 73. Stone G P, Mernaugh R, Haselton F R. Virus detection using filament-coupled antibodies. *Biotechnol Bioeng*. 2005; 91:699-706
- [0459] 74. Kidwell C S, Warach S. Acute ischemic cerebrovascular syndrome: Diagnostic criteria. *Stroke*. 2003; 34:2995-2998
- [0460] 75. Easton J D, Saver J L, Albers G W, Alberts M J, Chaturvedi S, Feldmann E, Hatsukami T S, Higashida R T, Johnston S C, Kidwell C S, Lutsep H L, Miller E, Sacco R L. Definition and evaluation of transient ischemic attack: A scientific statement for healthcare professionals from the american heart association/american stroke association stroke council; council on cardiovascular surgery and anesthesia; council on cardiovascular radiology and intervention; council on cardiovascular nursing; and the interdisciplinary council on peripheral vascular disease. The american academy of neurology affirms the value of this statement as an educational tool for neurologists. *Stroke*. 2009; 40:2276-2293
- [0461] 76. Meschia J F, Lojacono M A, Miller M J, Brott T G, Atkinson E J, O'Brien P C. Reliability of the questionnaire for verifying stroke-free status. *Cerebrovasc Dis*. 2004; 17:218-223
- [0462] 77. Griffiths R A, Good W R, Watson N P, O'Donnell H F, Fell P J, Shakespeare J M. Normal erythrocyte sedimentation rate in the elderly. *Br Med J (Clin Res Ed)*. 1984; 289:724-725
- [0463] 78. Bonita R, Beaglehole R. Recovery of motor function after stroke. *Stroke*. 1988; 19:1497-1500
- [0464] 79. Mahoney F I, Barthel D W. Functional evaluation: The barthel index. *Md State Med J*. 1965; 14:61-65
- [0465] 80. Wilson J T, Pettigrew L E, Teasdale G M. Structured interviews for the glasgow outcome scale and the extended glasgow outcome scale: Guidelines for their use. *J Neurotrauma*. 1998; 15:573-585
- [0466] 81. Spertus J A, Winder J A, Dewhurst T A, Deyo R A, Prodzienski J, McDonell M, Fihn S D. Development and evaluation of the seattle angina questionnaire: A new functional status measure for coronary artery disease. *J Am Coll Cardiol*. 1995; 25:333-341

[0467] 82. Bushel P R, Wolfinger R D, Gibson G. Simultaneous clustering of gene expression data with clinical chemistry and pathological evaluations reveals phenotypic prototypes. *BMC Syst Biol.* 2007; 1:15

[0468] 83. Perez J M, Hasleton, F. R., Wright, Dr. W. Viral detection using DNA functionalized gold filaments. *The Analyst.* 2009; 134, 1548-1553.

[0469] 84. Tsourkas A, Behlke M A, Rose S D, Bao G. Hybridization kinetics and thermodynamics of molecular beacons. *Nucleic Acids Res.* 2003; 31:1319-1330

[0470] 85. Bonnet G, Tyagi S, Libchaber A, Kramer F R. Thermodynamic basis of the enhanced specificity of structured DNA probes. *Proc Natl Acad Sci USA.* 1999; 96:6171-6176

[0471] Having thus described in detail preferred embodiments of the present invention, it is to be understood that the invention defined by the above paragraphs is not to be limited to particular details set forth in the above description as many apparent variations thereof are possible without departing from the spirit or scope of the present invention.

## SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 19

<210> SEQ ID NO 1

<211> LENGTH: 2188

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

agacagggggt agtgcgaggg cggcacagc cttctgtgt ggtttaccg cccagagagc	60
gtcatggacc tggggaaacc aatgaaaagc gtgtggtgg tggctctct tgcattttc	120
caggtatgcc tgtgtcaaga tgaggtcagc gacgattaca tcggagacaa caccacagt	180
gactacact ttgtcgagtc ttgtgtctc aagaaggacg tgcggaaactt taaaggctgg	240
ttcctcccta tcatgtactc catcatttg ttcgtggcc tactggcaa tggctggtc	300
gtgttgacct atatctatcc caagaggcct aagaccatga ccgataccta cctgctcaac	360
ctggcggtgg cagacatcct cttccctctg accctccct tctggcccta cagcgcggcc	420
aagtccctggg ttttcgggtt ccacttttc aagctcatct ttgcacatcta caagatgagc	480
ttcttcagtg gcatgctcct acttcttgc atcagcattt accgctacgt ggcacatgtc	540
caggctgtct cagctcaccc ccaccgtgcc cgcgctcttc tcatcagcaa gctgtccctgt	600
gtgggcacatc ggatacttagc cacagtgtctc tccatcccag agctcctgtt cagtgaccc	660
cagaggagca gcagtgagca agcgatgcga tgctctctca tcacagagca tggaggcc	720
tttatcacca tccaggtggc ccagatgggt atcggcttcc tggccccct gctggccatg	780
agcttctgtt accttgcatt catccgcacc ctgctccagg cacgcaactt tgacgcac	840
aaggccatca aggtgatcat cgctgtggtc gtggcttca tagtcttcca gctgcctac	900
aatgggggtgg tccctggccca gacggtgccc aacttcaaca tcaccatgtt caccgttg	960
ctcagtaagc aactcaacat cgccctacac gtcacccata gcctggccctg cgccgcgtc	1020
tgcgtcaacc ctttcttgc tgccttcattt ggcgtcaagt tccgcacgaa tctttcaag	1080
ctcttcaggc acctgggtcg cctcaggccat gagcagatcc ggcagtggtc ttctgtcg	1140
cacatccggc gtcctccat gagtgtggag gccgagacca ccaccaccc tctccatag	1200
gcgacttttc tgcctggact agaggggactt ctcgggggtt ccctgggtt gggataggaa	1260
gcagatgcaa tgactcagga catccccccg ccaaaaagctg ctcaggaaa agcagcttc	1320
ccctcagagt gcaaggccct gtcacccat atagttcac cccaaatccca gtcacccat	1380
ccaatgccaa aaaaagacag ggctgataag ctaacaccag acagacaaca ctggaaaca	1440
gaggctatttgc tccctaaac caaaaactga aagtggaaatgt ccagaacttgc ttccac	1500
ctggagtgaa gggcccaagg agggtgatgtt caagggcggt gggatggcc tgaagatcc	1560

---

-continued

---

tctgaatgaa	ccttctggcc	tcccacagac	tcaaatgctc	agaccagctc	ttccgaaaac	1620
caggccttat	ctccaagacc	agagatagt	gggagactc	ttggcttgg	gaggaaaagc	1680
ggacatcagc	tggtcaaaca	aactctctga	accctccct	ccatcgttt	cttcactgtc	1740
ctccaagcca	gcggaaatgg	cagctgccac	gccccttaa	aagcacactc	atcccctcac	1800
tgcccggtc	gccctccag	gtctcaaca	ggggagagt	tggtgttcc	tgcaggccag	1860
gccagctgcc	tccgcgtat	caaagccaca	ctctggctc	cagagtgggg	atgacatgca	1920
ctcagctctt	ggctccactg	ggatgggagg	agaggacaag	ggaaatgtca	ggggcgggga	1980
gggtgacagt	ggccgccccaa	ggcccacgag	cttgttctt	gttctttgtc	acagggactg	2040
aaaacctctc	ctcatgttct	gcttaagaga	gcaacatttt	acccacacac		2100
agataaaagtt	ttcccttgag	gaaacaacag	ctttaaaaga	aaaagaaaaa	aaaagtcttt	2160
ggtaaatggc	aaaaaaaaaa	aaaaaaaaaa				2188

&lt;210&gt; SEQ ID NO 2

&lt;211&gt; LENGTH: 378

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Homo sapiens

&lt;400&gt; SEQUENCE: 2

Met	Asp	Leu	Gly	Lys	Pro	Met	Lys	Ser	Val	Leu	Val	Val	Ala	Leu	Leu
1						5			10			15			

Val	Ile	Phe	Gln	Val	Cys	Leu	Cys	Gln	Asp	Glu	Val	Thr	Asp	Asp	Tyr
						20		25			30				

Ile	Gly	Asp	Asn	Thr	Thr	Val	Asp	Tyr	Thr	Leu	Phe	Glu	Ser	Leu	Cys
						35		40			45				

Ser	Lys	Lys	Asp	Val	Arg	Asn	Phe	Lys	Ala	Trp	Phe	Leu	Pro	Ile	Met
						50		55			60				

Tyr	Ser	Ile	Ile	Cys	Phe	Val	Gly	Leu	Leu	Gly	Asn	Gly	Leu	Val	Val
						65		70			75			80	

Leu	Thr	Tyr	Ile	Tyr	Phe	Lys	Arg	Leu	Lys	Thr	Met	Thr	Asp	Thr	Tyr
						85		90			95				

Leu	Leu	Asn	Leu	Ala	Val	Ala	Asp	Ile	Leu	Phe	Leu	Leu	Thr	Leu	Pro
						100		105			110				

Phe	Trp	Ala	Tyr	Ser	Ala	Ala	Lys	Ser	Trp	Val	Phe	Gly	Val	His	Phe
						115		120			125				

Cys	Lys	Leu	Ile	Phe	Ala	Ile	Tyr	Lys	Met	Ser	Phe	Phe	Ser	Gly	Met
						130		135			140				

Leu	Leu	Leu	Leu	Cys	Ile	Ser	Ile	Asp	Arg	Tyr	Val	Ala	Ile	Val	Gln
						145		150			155			160	

Ala	Val	Ser	Ala	His	Arg	His	Arg	Ala	Arg	Val	Leu	Leu	Ile	Ser	Lys
						165		170			175				

Leu	Ser	Cys	Val	Gly	Ile	Trp	Ile	Leu	Ala	Thr	Val	Leu	Ser	Ile	Pro
						180		185			190				

Glu	Leu	Leu	Tyr	Ser	Asp	Leu	Gln	Arg	Ser	Ser	Ser	Glu	Gln	Ala	Met
						195		200			205				

Arg	Cys	Ser	Leu	Ile	Thr	Glu	His	Val	Glu	Ala	Phe	Ile	Thr	Ile	Gln
						210		215			220				

Val	Ala	Gln	Met	Val	Ile	Gly	Phe	Leu	Val	Pro	Leu	Leu	Ala	Met	Ser
						225		230			235			240	

Phe	Cys	Tyr	Leu	Val	Ile	Ile	Arg	Thr	Leu	Leu	Gln	Ala	Arg	Asn	Phe
						245		250			255				

---

-continued

---

```

Glu Arg Asn Lys Ala Ile Lys Val Ile Ile Ala Val Val Val Val Phe
260           265           270

Ile Val Phe Gln Leu Pro Tyr Asn Gly Val Val Leu Ala Gln Thr Val
275           280           285

Ala Asn Phe Asn Ile Thr Ser Ser Thr Cys Glu Leu Ser Lys Gln Leu
290           295           300

Asn Ile Ala Tyr Asp Val Thr Tyr Ser Leu Ala Cys Val Arg Cys Cys
305           310           315           320

Val Asn Pro Phe Leu Tyr Ala Phe Ile Gly Val Lys Phe Arg Asn Asp
325           330           335

Leu Phe Lys Leu Phe Lys Asp Leu Gly Cys Leu Ser Gln Glu Gln Leu
340           345           350

Arg Gln Trp Ser Ser Cys Arg His Ile Arg Arg Ser Ser Met Ser Val
355           360           365

Glu Ala Glu Thr Thr Thr Phe Ser Pro
370           375

```

```

<210> SEQ ID NO 3
<211> LENGTH: 11185
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

```

```

<400> SEQUENCE: 3

gctgccccga gcctttctgg ggaagaactc caggcgtcg gacgcaacag ccgagaacat      60
taggtgttgt ggacaggagc tgggaccaaag atcttcggcc agccccgcatt cctcccgcat     120
cttccagcac cgccccgcac cttccgcata cttcccccggg ccaccacgct tcctatgtga     180
cccgccctggg caacgcccggaa cccagtegccc cagcgtcgca gtgaattttc cccccaact     240
gcaataagcc gccttccaag gccaagatgt tcataaatat aaagagcata ttatggatgt     300
gttcaacctt aatagtaacc catgcgtact ataaagtcaa agtggggaaa agcccacccgg     360
tgaggggctc cctctctggaa aaagtcaacc taccttgcata tttttcaacg atgcctactt     420
tgccacccag ttacaacacc agtgaatttc tccgcataatc atggctcaatc attgaagtgg     480
acaaaaatgg aaaagatttgg aaagagacta ctgtccttgtt ggcccaaat ggaaatatca     540
agatggtca ggactacaaa gggagagtgt ctgtgccac acatcccgag gctgtggcg     600
atgcctccct cactgtggtc aagctgtcg caagtgtatgc gggctttac cgctgtgacg     660
tcatgtacgg gattgaagac acacaagaca cgggtgtact gactgtggat ggggttgtgt     720
ttcactacag ggcggcaacc agcaggtaca cactgaattt tgaggctgtc cagaaggctt     780
gtttggacgt tggggcagtc atagcaactc cagagcagct ctttgcgtcc tatgaagatg     840
gatttgagca gtgtgacgca ggctggctgg ctgatcagac tgtcagatat cccatccggg     900
ctcccccaggt aggctgttat ggagataaga tgggaaaaggc aggagtcagg acttatggat     960
tccgttctcc ccaggaaact tacgtatgtt attgttatgtt ggatcatctg gatgggtatg    1020
tgttccaccc cactgtcccc agtaattca ctttcgagga ggctgcaaaa gagtgtgaaa    1080
accaggatgc caggctggca acagtggggg aactccagggc ggcattggagg aacgggtttg    1140
accagtgcga ttacgggtgg ctgtcggatg ccagcgtcgcc ccaccctgtg actgtggcca    1200
ggggccagtg tggaggtgtt ctacttgggg tgagaaccct gtatcgatggat gagaaccaga    1260
caggcttccc tccccctgtat agcagatgg atgcctactg ctttaaacct aaagaggctt    1320

```

---

-continued

---

caaccatcgat	tttgagttatc	ctcgccggaaa	ctgcgtatcc	cagtttatcc	aaagaaccac	1380
aatgggttcc	tgtatagaact	acaccaatca	tcccttttagt	tgtatgaatta	cctgtcatcc	1440
caacagagtt	ccctccccgtg	ggaaatatttgc	tcaagtttgc	acagaaagcc	acagtccaaac	1500
ctcaggctat	cacagatagt	ttagccacca	aattacccac	acctactggc	agtaccaaga	1560
agccctggaa	tatggatgac	tactcacctt	ctgcttcagg	acctcttgg	aagctagaca	1620
tatcagaaat	taaggaagaa	gtgctccaga	gtacaactgg	cgtctctcat	tatgtctacgg	1680
attcatggaa	tggtgtcggt	gaagataaac	aaacacaaga	atcggttaca	cagattgaac	1740
aatatagaatg	gggtcccttg	gtaacatcta	tggaaatctt	aaagcacatt	ccttccaagg	1800
aattccctgt	aactgaaaca	ccattggtaa	ctgcaagaat	gatccctggaa	tccaaaaactg	1860
aaaagaaaaat	ggtaagact	gtttctgaat	tggttaaccac	aggtcactat	ggattcacct	1920
tgggagaaga	ggatgatgaa	gacagaacac	ttacagttgg	atctgatgag	agcaccttgc	1980
tctttgacca	aattcctgaa	gtcattacgg	tgtcaaagac	ttcagaagac	accatccaca	2040
ctcattttaga	agacttggag	tcaagtctcag	catccacaac	tgtttccctt	ttaatttatgc	2100
ctgataataa	tggatcatcc	atggatgact	gggaagagag	acaaactagt	ggtaggataa	2160
cggaaagagtt	tcttggcaaa	tatctgtcta	ctacacccctt	tccatcacag	catcgtacag	2220
aaatagaatt	gtttccttat	tctggtgata	aaatattagt	agaggaaatt	tccacagtt	2280
tttaccccttc	tctacaaaca	gaaatgacac	atagaagaga	aagaacagaa	acactaatac	2340
cagagatgag	aacagatact	tatacagatg	aaatacaaga	agagatcaat	aaaagtccat	2400
ttatgggaaa	aacagaagaa	gaagtcttct	ctggggatgaa	actctctaca	tctctctcag	2460
agccaaattca	tgttacagag	tcttctgtgg	aatgaccaa	gtctttgtat	ttcccaacat	2520
tgataacaaa	gttaagtgc	gagccaaacag	aaagtaagaga	tatggaggaa	gactttacag	2580
caactccagg	tactacaaa	tatgtgaaa	atattacaac	agtgttttg	gcccatggta	2640
cttaaagtgt	tgaaggcagcc	actgtatcaa	aatggtcatg	ggatgaagat	aataacaacat	2700
ccaaggcctt	agagtctaca	gaacccctcag	cctcttcaaa	attgccccct	gccttactca	2760
caactgtggg	gatgaatgaa	aggataaag	acatcccaag	tttcactgaa	gatggaggcag	2820
atgaatttac	tcttatttcca	gatagtactc	aaaagcagtt	agaggagggtt	actgtatgaa	2880
acatagcagc	ccatggaaaa	ttcacaatta	gatttcagcc	aactacatca	actggatttg	2940
cagaaaagtc	aactttgaga	gattctacaa	ctgaagaaaa	agttccaccc	atcacaagca	3000
ctgaaggcga	agtttatgca	accatggaa	gaagtgcctt	gggtgaagta	gaagatgtgg	3060
acctctctaa	gccagtatct	actgttcccc	aatttgacca	cacttcagag	gtggaggat	3120
tagcatttgt	tagttatagt	agcacccaa	agcctactac	ttatgtac	tcttccatata	3180
ccatccctct	ttctgttaatt	cccaagacag	actggggagt	gttagtacct	tctgttccat	3240
cagaagatga	agttcttaggt	gaaccctctc	aagacatact	tgtcattgtat	cagactcgcc	3300
ttgaagcgac	tatttctcca	gaaactatga	gaacaacaaa	aatcacagag	ggaacaactc	3360
aggaagaatt	cccttggaaa	gaacagactg	cagagaaacc	agttccctgt	ctcagttcta	3420
cagttggac	tcccaaggag	gcagtaacac	cactggatga	acaagaggc	gatggatcag	3480
catatacagt	ctctgttgc	gaattgttgc	cagttctgc	gagggtccca	gttttagaaa	3540
caactccagt	tggaaaaatt	gatcacatgt	tgtcttatcc	accagggtgt	gtactgagc	3600
acaaagtgaa	aacagatgaa	gtggtaacac	taacaccacg	cattggggcc	aaagtatctt	3660

---

-continued

---

taagtccagg gcctgaacaa aaatatgaaa cagaaggttag tagtacaaca ggatttacat	3720
catcttttagat tccttttagt acccacatta cccagttat ggaagaaaacc actactgaga	3780
aaacatccct agaggatatt gattnaggt caggattatt tgaaaagccc aaagecoacag	3840
aactcataga atttcaaca atcaaagtca cagttccaag tgatattacc actgcottca	3900
gttcagtaga cagacttcac acaacttcag cattcaagcc atcttcogcg atcactaaga	3960
aaccacctct catcgacagg gaacctgggt aagaaacaac cagtgcacatg gtaatcattg	4020
gagaatcaac atctcatgtt cctccacta cccttgaaga tattttagcc aaggaaacag	4080
aaaccgatata ttagatagag tatttcacgaa cttcaagtcc tcctgttaca cagccaaacaa	4140
gaccacccac tgtggaaagac aaagaggcct ttggacctca ggcgccttct acgccacagc	4200
ccccagcaag cacaaaattt caccctgaca ttaatgttta tattatttag gtcagaaaaa	4260
ataagacagg tcgaatgagt gattnaggt taattggtca tccaatagat tcagaatcta	4320
aagaagatga acctttagt gaagaaacag atccagtgc ttagtcaatg gctgaaattt	4380
tacctgaatt ccctgacata attgaaatag acctataccca cagtgaagaa aatgaagaag	4440
aagaagaaga gtgtgcaaat gctactgtat tgacaaccac cccatctgtg cagtacataa	4500
atgggaagca tctcggttacc actgtgcccagg accctgagaa gctaggcgtg	4560
gccagtttga aagtgttgc cttctcaga atttctcgaa cagctctgaa agtgataactc	4620
atccatgtt aatagccaaa acggaattgt ctactgtgt gcaacctaataat gatctacag	4680
aaacaactga gtccttgaa gttacatggaa agcctgagac ttaccctgaa acatcagaac	4740
attttcagg tggtgagct gatgtttcc ccacagtccc attccatgag gaatttggaaa	4800
gtggAACAGC caaaaaaggc gcagaatcag tcacagagag agataactgaa gttggtcattc	4860
aggcacatga acatactgaa cctgtatctc tgtttctgaa agagtcttca ggagagattt	4920
ccattgacca agaatctcag aaaatagcc ttgcaaggc tacagaagta acatttggtg	4980
aagaggtaga aaaaagtact tctgtcacat acactcccac tatagttcca agttctcat	5040
cagcatatgt ttcagaggaa gaagcgttca ccctaataagg aaatccttgg ccagatgacc	5100
tgttgtctac caaagaaagc tgggtagaag caactcttag acaagttgtt gagctctcag	5160
ggagttcttc gattccaatt acagaaggct ctggagaagc agaagaagat gaagatacaa	5220
tgttcacat ggttaactgtt gatcacaga gaaatactac tgataactc attacttag	5280
acactagcag gataatcaca gaaagctttt ttgaggttcc tgcaaccacc atttatccag	5340
tttctgaaca accttctgca aaagtgggtc ctaccaagtt tgtaagtggaa acagacactt	5400
ctgagttggat ttccagttacc actgttgagg aaaagaaaag gaaggaggag gagggacta	5460
caggtacggc ttctacattt gaggtatatt catctacaca gagatcggat caattaattt	5520
taccctttga attagaaagt ccaaatactgat ctacatcttag tgattcaggat accagaaaa	5580
gttttatgtc cttgacaaca ccaacacagt ctgaaaggga aatgcacat tctactcctg	5640
tctttacaga aacaaataca ttagaaaattt tggggcaca gaccactgag cacagcgtt	5700
tccatcaacc tggggttcag gaagggtgtca ccactctccc acgtatgttcc gctctgtct	5760
ttatggagca gggctctggaa gaagctgtgtt ccgacccttcc aaccaccact gtttcttcat	5820
tttcattaaa cgttagatgtt gcaatttcaag ccgaaaagga agtagctggc actttgtctc	5880
cgcgttggaa aactacattc tccactgagc caacaggact ggtttgagt acagtaatgg	5940

---

-continued

---

acagagtagt tgctgaaaat ataacccaa catccaggga aatagtgatt tcagagcgat	6000
taggagaacc aaattatggg gcagaaataa ggggcgtttc cacagggttt cctttggagg	6060
aagatttcag tggtgacttt agagaatact caacagtgtc tcatccata gcaaaaagaag	6120
aaacggtaat gatggaaggc tctggagatg cagcatttag ggacaccagg acttcaccat	6180
ctacagtacc tacttcagtt cacatcgtc acatatctga ctcagaagga cccagtagca	6240
ccatggtcag cacttcagcc ttccccctggg aagagttac atcctcagct gagggtcag	6300
gtgagcaact ggtcacagtc agcagctctg ttgttccagt gttcccagt gctgtgcaaa	6360
agtttctgg tacagcttcc tccattatcg acgaaggatt gggagaagtg ggtactgtca	6420
atgaaattga tagaagatcc accatttac caacagcaga agtggaaaggt acgaaaagctc	6480
cagtagagaa ggaggaagta aaggtcagtg gcacagttc aacaaacttt ccccaaacta	6540
tagagccagc caaattatgg tctaggcaag aagtcaaccc tgtaagacaa gaaattgaaa	6600
gtgaaacaac atcagaggaa caaattcaag aagaaaagtc atttgaatcc cctcaaaact	6660
ctccctgcaac agaacaaaca atctttgatt cacagacatt tactgaaact gaactcaaaa	6720
ccacagatta ttctgtacta acaacaaaga aaacttacag tgatgataaa gaaatgaagg	6780
aggaagacac ttcttttagtt aacatgtcta ctccagatcc agatgcaaat ggcttggaaat	6840
cttacacaac tctccctgaa gctactgaaa agtcacatTT tttcttagct actgcattag	6900
taactgaatc tataccagct gaacatgttag tcacagatc accaatcaaa aaggaagaaa	6960
gtacaaaaca tttccgaaa ggcattgagac caacaattca agagtcaag actgagctct	7020
tattctctgg actgggatca ggagaagaag ttttacctac tctaccaaca gagtcagtga	7080
attttactga agtggaaacaa atcaataaca cattatatcc ccacacttct caagtggaaa	7140
gtacctcaag tgacaaaattt gaagacttta acagaatgga aatgtggca aaagaagttt	7200
gaccacttgt atctcaaaca gacattttg aaggtatgg gtcagtaacc agcacaacat	7260
taatagaaat tttaagtgc actggagcag aaggaccac ggtggccacct ctcccttct	7320
ccacggacat cggacatcct caaaatcaga ctgtcaggtg ggcagaagaa atccagacta	7380
gttagaccaca aaccataact gaacaagact ctaacaagaa ttcttcaaca gcagaaatta	7440
acgaaacaac aacctcatct actgattttc tggctagagc ttatggttt gaaatggcca	7500
aagaatttgt tacatcagca cccaaaccat ctgacttgta ttatgaacct tctggagaag	7560
gatctggaga agtggatatt gttgattcat ttcacacttc tgcaactact caggcaacca	7620
gacaagaaag cagcaccaca ttgtttctg atgggtccct gaaaaacat cctgaggtgc	7680
caagcgctaa agctgttaact gctgatggat tcccaacagt ttcaatgtatg ctgccttct	7740
attcagagca gaacaaaaggc tcccctgatc caactagcac actgtcaat acagtgtcat	7800
atgagaggtc cacagacggt agttccaag accgtttcag ggaattcgag gattccacct	7860
taaaacctaa cagaaaaaaa cccactgaaa atattatcat agacctggac aaagaggaca	7920
aggatTTaat attgacaattt acagagatca ccatttttgc aattctacat gagctgacat	7980
cggttttttt tactatcata gatattgtatc atactaaacc tttgtatgaa gacatttttgc	8040
gaatgcaaac agatatacat acagaggtac catcagaacc acatgacagt aatgtatgaaa	8100
gttatgtatc cagcactcaa gttcaagaga tttatggatc agctgtcaac ctttctttaa	8160
cttgaggaaac atttgaggc tctgctgatc ttctggatc ctacactcag gcaacacatg	8220
atgaatcaat gacttatgaa gatagaagcc aactagatca catgggtttt cacttcacaa	8280

---

-continued

---

ctggatccc	tgtccctagc	acagaaaacag	aatttagacgt	tttacttccc	acggcaacat	8340
ccctgccaat	tcctcgtaag	tctgccacag	ttattccaga	gattgaagga	ataaaagctg	8400
aagcaaaaagc	cctggatgac	atgtttgaat	caagcactt	gtctgatgg	caagetattg	8460
cagacaaaaag	tgaataata	ccaacattgg	gcacatttg	aaggactcg	gaggagttat	8520
aagacaaaaaa	acatgctggt	ccttctttc	agccagaatt	ctcttcagga	gctgaggagg	8580
cattagtaga	ccatactccc	tatctaagt	ttgctactac	ccacccat	tgatcagatg	8640
taacagaggt	gcctgatgt	atggaaggat	ccaatcccc	atattacact	gataacaat	8700
tagcagttc	aacatttgcg	aagttgtctt	ctcagacacc	atcatctccc	ctcactatct	8760
actcaggcag	tgaaggctct	ggacacacag	agatccccca	gcccagtgt	ctgcccaggaa	8820
tagacgtcgg	ctcatctgt	atgtccccc	aggattctt	taaggaaatt	catgttaata	8880
ttgaagcaac	tttcaaacc	tcaagtgagg	aataccctca	cataactgag	cctccctctt	8940
tatctcctga	cacaaaatta	gaacccctcg	aagatgtatgg	taaacctgag	ttattagaag	9000
aaatggaagc	ttctccca	gaacttattt	ctgtggagg	aactgagatt	ctccaagatt	9060
tccaaaacaa	aaccgatgt	caagttctg	gagaagcaat	caagatttt	cccaccatta	9120
aaacacctga	ggctggaact	gttattacaa	ctgcccgt	aattgaatta	gaaggtgcta	9180
cacagtggcc	acactctact	tctgcttctg	ccacccatgg	ggtcgaggca	ggtgtgggtc	9240
cttggctaag	tccacagact	tctgagaggc	ccacgcttcc	ttcttctcca	gaaataaacc	9300
ctgaaactca	agcagctta	atcagagggc	aggattccac	gatagcagca	tcagaacagc	9360
aagtggcagc	gagaatttt	gattccaatg	atcaggcaac	agtaaaccc	gtggatttt	9420
atactgaggt	tgcaacacca	ccatccccc	ttctggagac	ttctaatgaa	acagattcc	9480
tgattggcat	taatgaagag	tcaagtggagg	gcacggcaat	ctatccat	ggacccgtat	9540
gctgcaaaat	gaacccgtgc	cttaacggag	gcacccgtt	tcctactgaa	acttccatcg	9600
tatgcacctg	tgtgccagga	tacagcggag	accagtgt	acttgcattt	gatgaatgt	9660
actctaatacc	ctgtcgtaat	ggagccactt	gtgttgcattt	ttttacaca	ttcaggtgcc	9720
tctgcctcc	aagttatgtt	ggtgcaactt	gtgagcaaga	taccgagaca	tgtgactatg	9780
gctggcacaa	attccaaggg	cagtgcata	aatacttgc	ccatcgacgc	acatggatg	9840
cagctgaacg	ggaatggcgt	ctgcagggtt	cccatctcac	aagcatctg	tctcacagaa	9900
aacaaatgtt	tgttaatcg	gtggccat	attatcgt	gataggcctc	aatgacaaga	9960
tgtttgagca	tgaatccgt	tggactgtat	gcagcacact	gcaatacgag	aattggagac	10020
ccaaccagcc	agacagcttc	ttttctgtgt	gagaagactg	tgttgtaatc	atttggcatg	10080
agaatggcca	gtggaatgt	gttccctgca	attaccatct	cacctatacg	tgcaagaaag	10140
gaacagttgc	ttgcggccag	ccccctgtt	tagaaaatgc	caagccctt	ggaaagatga	10200
aacctcgta	tgaatcaac	tccctgatta	gataccactg	caaagatgt	ttcattcaac	10260
gtcacccctcc	aactatccgg	tgcttaggaa	atggaagatg	ggctatacc	aaaattacct	10320
gcatgaaccc	atctgcatac	caaaggactt	attctatgaa	atactttaaa	aattccat	10380
cagcaaagga	caattcaata	aatacatcca	aacatgtca	tcgttggagc	cgagggtggc	10440
aggagtcgag	gcgcgtatcc	ctaaaatggc	gaacatgtgt	tttcatcatt	tcagccaaag	10500
tcctaacttc	ctgtgcctt	cctatcacct	cgagaagtaa	ttatcagtt	gtttggattt	10560

---

-continued

---

```

ttggaccacc gttcagtcat tttgggttgc cgtgctcca aaacattta aatgaaagt 10620
ttggcattca aaaagacagc agacaaaatg aaagaaaatg agagcagaaa gtaagcattt 10680
ccagcctatc taatttctt agtttctat ttgcctccag tgcagtcacat ttcctaatgt 10740
ataccagcct actgtactat ttaaaatgct caatttcagc accgatggcc atgtaaataa 10800
gatgatttaa tggtgattt aatcctgtat ataaaataaa aagtcacaat gagtttggc 10860
atatthaatg atgattatgg agccttagag gtcttaatc attggttcg ctgctttat 10920
gtagtttagg ctggaaatgg tttcacttgc tctttgactg tcagcaagac tgaagatggc 10980
tttcctgga cagctagaaa acacaaaatc ttgttaggtca ttgcacctat ctcagccata 11040
ggtgcaagttt gcttctacat gatgctaaag gctgcgaatg ggatcctgat ggaactaagg 11100
actccaatgt cgaactcttc tttgctgcat tccttttct tcacttacaa gaaaggcctg 11160
aatggaggac ttttctgtaa ccagg 11185

```

```

<210> SEQ ID NO 4
<211> LENGTH: 3376
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

```

```
<400> SEQUENCE: 4
```

```

Leu His Lys Val Lys Val Gly Lys Ser Pro Pro Val Arg Gly Ser Leu
1 5 10 15

```

```

Ser Gly Lys Val Ser Leu Pro Cys His Phe Ser Thr Met Pro Thr Leu
20 25 30

```

```

Pro Pro Ser Tyr Asn Thr Ser Glu Phe Leu Arg Ile Lys Trp Ser Lys
35 40 45

```

```

Ile Glu Val Asp Lys Asn Gly Lys Asp Leu Lys Glu Thr Thr Val Leu
50 55 60

```

```

Val Ala Gln Asn Gly Asn Ile Lys Ile Gly Gln Asp Tyr Lys Gly Arg
65 70 75 80

```

```

Val Ser Val Pro Thr His Pro Glu Ala Val Gly Asp Ala Ser Leu Thr
85 90 95

```

```

Val Val Lys Leu Leu Ala Ser Asp Ala Gly Leu Tyr Arg Cys Asp Val
100 105 110

```

```

Met Tyr Gly Ile Glu Asp Thr Gln Asp Thr Val Ser Leu Thr Val Asp
115 120 125

```

```

Gly Val Val Phe His Tyr Arg Ala Ala Thr Ser Arg Tyr Thr Leu Asn
130 135 140

```

```

Phe Glu Ala Ala Gln Lys Ala Cys Leu Asp Val Gly Ala Val Ile Ala
145 150 155 160

```

```

Thr Pro Glu Gln Leu Phe Ala Ala Tyr Glu Asp Gly Phe Glu Gln Cys
165 170 175

```

```

Asp Ala Gly Trp Leu Ala Asp Gln Thr Val Arg Tyr Pro Ile Arg Ala
180 185 190

```

```

Pro Arg Val Gly Cys Tyr Gly Asp Lys Met Gly Lys Ala Gly Val Arg
195 200 205

```

```

Thr Tyr Gly Phe Arg Ser Pro Gln Glu Thr Tyr Asp Val Tyr Cys Tyr
210 215 220

```

```

Val Asp His Leu Asp Gly Asp Val Phe His Leu Thr Val Pro Ser Lys
225 230 235 240

```

```

Phe Thr Phe Glu Glu Ala Ala Lys Glu Cys Glu Asn Gln Asp Ala Arg
245 250 255

```

---

-continued

---

Leu Ala Thr Val Gly Glu Leu Gln Ala Ala Trp Arg Asn Gly Phe Asp  
     260                 265                 270  
  
 Gln Cys Asp Tyr Gly Trp Leu Ser Asp Ala Ser Val Arg His Pro Val  
     275                 280                 285  
  
 Thr Val Ala Arg Ala Gln Cys Gly Gly Leu Leu Gly Val Arg Thr  
     290                 295                 300  
  
 Leu Tyr Arg Phe Glu Asn Gln Thr Gly Phe Pro Pro Pro Asp Ser Arg  
     305                 310                 315                 320  
  
 Phe Asp Ala Tyr Cys Phe Lys Pro Lys Glu Ala Thr Thr Ile Asp Leu  
     325                 330                 335  
  
 Ser Ile Leu Ala Glu Thr Ala Ser Pro Ser Leu Ser Lys Glu Pro Gln  
     340                 345                 350  
  
 Met Val Ser Asp Arg Thr Thr Pro Ile Ile Pro Leu Val Asp Glu Leu  
     355                 360                 365  
  
 Pro Val Ile Pro Thr Glu Phe Pro Pro Val Gly Asn Ile Val Ser Phe  
     370                 375                 380  
  
 Glu Gln Lys Ala Thr Val Gln Pro Gln Ala Ile Thr Asp Ser Leu Ala  
     385                 390                 395                 400  
  
 Thr Lys Leu Pro Thr Pro Thr Gly Ser Thr Lys Lys Pro Trp Asp Met  
     405                 410                 415  
  
 Asp Asp Tyr Ser Pro Ser Ala Ser Gly Pro Leu Gly Lys Leu Asp Ile  
     420                 425                 430  
  
 Ser Glu Ile Lys Glu Glu Val Leu Gln Ser Thr Thr Gly Val Ser His  
     435                 440                 445  
  
 Tyr Ala Thr Asp Ser Trp Asp Gly Val Val Glu Asp Lys Gln Thr Gln  
     450                 455                 460  
  
 Glu Ser Val Thr Gln Ile Glu Gln Ile Glu Val Gly Pro Leu Val Thr  
     465                 470                 475                 480  
  
 Ser Met Glu Ile Leu Lys His Ile Pro Ser Lys Glu Phe Pro Val Thr  
     485                 490                 495  
  
 Glu Thr Pro Leu Val Thr Ala Arg Met Ile Leu Glu Ser Lys Thr Glu  
     500                 505                 510  
  
 Lys Lys Met Val Ser Thr Val Ser Glu Leu Val Thr Thr Gly His Tyr  
     515                 520                 525  
  
 Gly Phe Thr Leu Gly Glu Glu Asp Asp Glu Asp Arg Thr Leu Thr Val  
     530                 535                 540  
  
 Gly Ser Asp Glu Ser Thr Leu Ile Phe Asp Gln Ile Pro Glu Val Ile  
     545                 550                 555                 560  
  
 Thr Val Ser Lys Thr Ser Glu Asp Thr Ile His Thr His Leu Glu Asp  
     565                 570                 575  
  
 Leu Glu Ser Val Ser Ala Ser Thr Thr Val Ser Pro Leu Ile Met Pro  
     580                 585                 590  
  
 Asp Asn Asn Gly Ser Ser Met Asp Asp Trp Glu Glu Arg Gln Thr Ser  
     595                 600                 605  
  
 Gly Arg Ile Thr Glu Glu Phe Leu Gly Lys Tyr Leu Ser Thr Thr Pro  
     610                 615                 620  
  
 Phe Pro Ser Gln His Arg Thr Glu Ile Glu Leu Phe Pro Tyr Ser Gly  
     625                 630                 635                 640  
  
 Asp Lys Ile Leu Val Glu Gly Ile Ser Thr Val Ile Tyr Pro Ser Leu  
     645                 650                 655  
  
 Gln Thr Glu Met Thr His Arg Arg Glu Arg Thr Glu Thr Leu Ile Pro

---

-continued

---

660	665	670
Glu Met Arg Thr Asp Thr Tyr Thr Asp Glu Ile Gln Glu Glu Ile Thr		
675	680	685
Lys Ser Pro Phe Met Gly Lys Thr Glu Glu Glu Val Phe Ser Gly Met		
690	695	700
Lys Leu Ser Thr Ser Leu Ser Glu Pro Ile His Val Thr Glu Ser Ser		
705	710	715
720		
Val Glu Met Thr Lys Ser Phe Asp Phe Pro Thr Leu Ile Thr Lys Leu		
725	730	735
Ser Ala Glu Pro Thr Glu Val Arg Asp Met Glu Glu Asp Phe Thr Ala		
740	745	750
Thr Pro Gly Thr Thr Lys Tyr Asp Glu Asn Ile Thr Thr Val Leu Leu		
755	760	765
Ala His Gly Thr Leu Ser Val Glu Ala Ala Thr Val Ser Lys Trp Ser		
770	775	780
Trp Asp Glu Asp Asn Thr Thr Ser Lys Pro Leu Glu Ser Thr Glu Pro		
785	790	795
800		
Ser Ala Ser Ser Lys Leu Pro Pro Ala Leu Leu Thr Thr Val Gly Met		
805	810	815
Asn Gly Lys Asp Lys Asp Ile Pro Ser Phe Thr Glu Asp Gly Ala Asp		
820	825	830
Glu Phe Thr Leu Ile Pro Asp Ser Thr Gln Lys Gln Leu Glu Glu Val		
835	840	845
Thr Asp Glu Asp Ile Ala Ala His Gly Lys Phe Thr Ile Arg Phe Gln		
850	855	860
Pro Thr Thr Ser Thr Gly Ile Ala Glu Lys Ser Thr Leu Arg Asp Ser		
865	870	875
880		
Thr Thr Glu Glu Lys Val Pro Pro Ile Thr Ser Thr Glu Gly Gln Val		
885	890	895
Tyr Ala Thr Met Glu Gly Ser Ala Leu Gly Glu Val Glu Asp Val Asp		
900	905	910
Leu Ser Lys Pro Val Ser Thr Val Pro Gln Phe Ala His Thr Ser Glu		
915	920	925
Val Glu Gly Leu Ala Phe Val Ser Tyr Ser Ser Thr Gln Glu Pro Thr		
930	935	940
Thr Tyr Val Asp Ser Ser His Thr Ile Pro Leu Ser Val Ile Pro Lys		
945	950	955
960		
Thr Asp Trp Gly Val Leu Val Pro Ser Val Pro Ser Glu Asp Glu Val		
965	970	975
Leu Gly Glu Pro Ser Gln Asp Ile Leu Val Ile Asp Gln Thr Arg Leu		
980	985	990
Glu Ala Thr Ile Ser Pro Glu Thr Met Arg Thr Thr Lys Ile Thr Glu		
995	1000	1005
Gly Thr Thr Gln Glu Glu Phe Pro Trp Lys Glu Gln Thr Ala Glu		
1010	1015	1020
Lys Pro Val Pro Ala Leu Ser Ser Thr Ala Trp Thr Pro Lys Glu		
1025	1030	1035
Ala Val Thr Pro Leu Asp Glu Gln Glu Gly Asp Gly Ser Ala Tyr		
1040	1045	1050
Thr Val Ser Glu Asp Glu Leu Leu Thr Gly Ser Glu Arg Val Pro		
1055	1060	1065

-continued

Val	Leu	Glu	Thr	Thr	Pro	Val	Gly	Lys	Ile	Asp	His	Ser	Val	Ser
1070						1075					1080			
Tyr	Pro	Pro	Gly	Ala	Val	Thr	Glu	His	Lys	Val	Lys	Thr	Asp	Glu
1085						1090					1095			
Val	Val	Thr	Leu	Thr	Pro	Arg	Ile	Gly	Pro	Lys	Val	Ser	Leu	Ser
1100						1105					1110			
Pro	Gly	Pro	Glu	Gln	Lys	Tyr	Glu	Thr	Glu	Gly	Ser	Ser	Thr	Thr
1115						1120					1125			
Gly	Phe	Thr	Ser	Ser	Leu	Ser	Pro	Phe	Ser	Thr	His	Ile	Thr	Gln
1130						1135					1140			
Leu	Met	Glu	Glu	Thr	Thr	Thr	Glu	Lys	Thr	Ser	Leu	Glu	Asp	Ile
1145						1150					1155			
Asp	Leu	Gly	Ser	Gly	Leu	Phe	Glu	Lys	Pro	Lys	Ala	Thr	Glu	Leu
1160						1165					1170			
Ile	Glu	Phe	Ser	Thr	Ile	Lys	Val	Thr	Val	Pro	Ser	Asp	Ile	Thr
1175						1180					1185			
Thr	Ala	Phe	Ser	Ser	Val	Asp	Arg	Leu	His	Thr	Thr	Ser	Ala	Phe
1190						1195					1200			
Lys	Pro	Ser	Ser	Ala	Ile	Thr	Lys	Lys	Pro	Pro	Leu	Ile	Asp	Arg
1205						1210					1215			
Glu	Pro	Gly	Glu	Glu	Thr	Thr	Ser	Asp	Met	Val	Ile	Ile	Gly	Glu
1220						1225					1230			
Ser	Thr	Ser	His	Val	Pro	Pro	Thr	Thr	Leu	Glu	Asp	Ile	Val	Ala
1235						1240					1245			
Lys	Glu	Thr	Glu	Thr	Asp	Ile	Asp	Arg	Glu	Tyr	Phe	Thr	Thr	Ser
1250						1255					1260			
Ser	Pro	Pro	Ala	Thr	Gln	Pro	Thr	Arg	Pro	Pro	Thr	Val	Glu	Asp
1265						1270					1275			
Lys	Glu	Ala	Phe	Gly	Pro	Gln	Ala	Leu	Ser	Thr	Pro	Gln	Pro	Pro
1280						1285					1290			
Ala	Ser	Thr	Lys	Phe	His	Pro	Asp	Ile	Asn	Val	Tyr	Ile	Ile	Glu
1295						1300					1305			
Val	Arg	Glu	Asn	Lys	Thr	Gly	Arg	Met	Ser	Asp	Leu	Ser	Val	Ile
1310						1315					1320			
Gly	His	Pro	Ile	Asp	Ser	Glu	Ser	Lys	Glu	Asp	Glu	Pro	Cys	Ser
1325						1330					1335			
Glu	Glu	Thr	Asp	Pro	Val	His	Asp	Leu	Met	Ala	Glu	Ile	Leu	Pro
1340						1345					1350			
Glu	Phe	Pro	Asp	Ile	Ile	Glu	Ile	Asp	Leu	Tyr	His	Ser	Glu	Glu
1355						1360					1365			
Asn	Glu	Glu	Glu	Glu	Glu	Cys	Ala	Asn	Ala	Thr	Asp	Val	Thr	
1370						1375					1380			
Thr	Thr	Pro	Ser	Val	Gln	Tyr	Ile	Asn	Gly	Lys	His	Leu	Val	Thr
1385						1390					1395			
Thr	Val	Pro	Lys	Asp	Pro	Glu	Ala	Ala	Glu	Ala	Arg	Arg	Gly	Gln
1400						1405					1410			
Phe	Glu	Ser	Val	Ala	Pro	Ser	Gln	Asn	Phe	Ser	Asp	Ser	Ser	Glu
1415						1420					1425			
Ser	Asp	Thr	His	Pro	Phe	Val	Ile	Ala	Lys	Thr	Glu	Leu	Ser	Thr
1430						1435					1440			
Ala	Val	Gln	Pro	Asn	Glu	Ser	Thr	Glu	Thr	Thr	Glu	Ser	Leu	Glu
1445						1450					1455			

---

-continued

---

Val	Thr	Trp	Lys	Pro	Glu	Thr	Tyr	Pro	Glu	Thr	Ser	Glu	His	Phe
1460				1465								1470		
Ser	Gly	Gly	Glu	Pro	Asp	Val	Phe	Pro	Thr	Val	Pro	Phe	His	Glu
1475					1480							1485		
Glu	Phe	Glu	Ser	Gly	Thr	Ala	Lys	Lys	Gly	Ala	Glu	Ser	Val	Thr
1490					1495							1500		
Glu	Arg	Asp	Thr	Glu	Val	Gly	His	Gln	Ala	His	Glu	His	Thr	Glu
1505					1510							1515		
Pro	Val	Ser	Leu	Phe	Pro	Glu	Glu	Ser	Ser	Gly	Glu	Ile	Ala	Ile
1520					1525							1530		
Asp	Gln	Glu	Ser	Gln	Lys	Ile	Ala	Phe	Ala	Arg	Ala	Thr	Glu	Val
1535					1540							1545		
Thr	Phe	Gly	Glu	Glu	Val	Glu	Lys	Ser	Thr	Ser	Val	Thr	Tyr	Thr
1550					1555							1560		
Pro	Thr	Ile	Val	Pro	Ser	Ser	Ala	Ser	Ala	Tyr	Val	Ser	Glu	Glu
1565					1570							1575		
Glu	Ala	Val	Thr	Leu	Ile	Gly	Asn	Pro	Trp	Pro	Asp	Asp	Leu	Leu
1580					1585							1590		
Ser	Thr	Lys	Glu	Ser	Trp	Val	Glu	Ala	Thr	Pro	Arg	Gln	Val	Val
1595					1600							1605		
Glu	Leu	Ser	Gly	Ser	Ser	Ile	Pro	Ile	Thr	Glu	Gly	Ser	Gly	
1610					1615							1620		
Glu	Ala	Glu	Glu	Asp	Glu	Asp	Thr	Met	Phe	Thr	Met	Val	Thr	Asp
1625					1630							1635		
Leu	Ser	Gln	Arg	Asn	Thr	Thr	Asp	Thr	Leu	Ile	Thr	Leu	Asp	Thr
1640					1645							1650		
Ser	Arg	Ile	Ile	Thr	Glu	Ser	Phe	Phe	Glu	Val	Pro	Ala	Thr	Thr
1655					1660							1665		
Ile	Tyr	Pro	Val	Ser	Glu	Gln	Pro	Ser	Ala	Lys	Val	Val	Pro	Thr
1670					1675							1680		
Lys	Phe	Val	Ser	Glu	Thr	Asp	Thr	Ser	Glu	Trp	Ile	Ser	Ser	Thr
1685					1690							1695		
Thr	Val	Glu	Glu	Lys	Arg	Lys	Glu	Glu	Glu	Gly	Thr	Thr	Gly	
1700					1705							1710		
Thr	Ala	Ser	Thr	Phe	Glu	Val	Tyr	Ser	Ser	Thr	Gln	Arg	Ser	Asp
1715					1720							1725		
Gln	Leu	Ile	Leu	Pro	Phe	Glu	Leu	Glu	Ser	Pro	Asn	Val	Ala	Thr
1730					1735							1740		
Ser	Ser	Asp	Ser	Gly	Thr	Arg	Lys	Ser	Phe	Met	Ser	Leu	Thr	Thr
1745					1750							1755		
Pro	Thr	Gln	Ser	Glu	Arg	Glu	Met	Thr	Asp	Ser	Thr	Pro	Val	Phe
1760					1765							1770		
Thr	Glu	Thr	Asn	Thr	Leu	Glu	Asn	Leu	Gly	Ala	Gln	Thr	Thr	Glu
1775					1780							1785		
His	Ser	Ser	Ile	His	Gln	Pro	Gly	Val	Gln	Glu	Gly	Leu	Thr	Thr
1790					1795							1800		
Leu	Pro	Arg	Ser	Pro	Ala	Ser	Val	Phe	Met	Glu	Gln	Gly	Ser	Gly
1805					1810							1815		
Glu	Ala	Ala	Ala	Asp	Pro	Glu	Thr	Thr	Thr	Val	Ser	Ser	Phe	Ser
1820					1825							1830		
Leu	Asn	Val	Glu	Tyr	Ala	Ile	Gln	Ala	Glu	Lys	Glu	Val	Ala	Gly

---

-continued

---

1835	1840	1845
Thr Leu Ser Pro His Val Glu	Thr Thr Phe Ser	Thr Glu Pro Thr
1850	1855	1860
Gly Leu Val Leu Ser Thr Val	Met Asp Arg Val Val	Ala Glu Asn
1865	1870	1875
Ile Thr Gln Thr Ser Arg Glu	Ile Val Ile Ser Glu	Arg Leu Gly
1880	1885	1890
Glu Pro Asn Tyr Gly Ala Glu	Ile Arg Gly Phe Ser	Thr Gly Phe
1895	1900	1905
Pro Leu Glu Glu Asp Phe Ser	Gly Asp Phe Arg Glu	Tyr Ser Thr
1910	1915	1920
Val Ser His Pro Ile Ala Lys	Glu Glu Thr Val Met	Met Glu Gly
1925	1930	1935
Ser Gly Asp Ala Ala Phe Arg	Asp Thr Gln Thr Ser	Pro Ser Thr
1940	1945	1950
Val Pro Thr Ser Val His Ile	Ser His Ile Ser Asp	Ser Glu Gly
1955	1960	1965
Pro Ser Ser Thr Met Val Ser	Thr Ser Ala Phe Pro	Trp Glu Glu
1970	1975	1980
Phe Thr Ser Ser Ala Glu Gly	Ser Gly Glu Gln Leu	Val Thr Val
1985	1990	1995
Ser Ser Ser Val Val Pro Val	Leu Pro Ser Ala Val	Gln Lys Phe
2000	2005	2010
Ser Gly Thr Ala Ser Ser Ile	Ile Asp Glu Gly Leu	Gly Glu Val
2015	2020	2025
Gly Thr Val Asn Glu Ile Asp	Arg Arg Ser Thr Ile	Leu Pro Thr
2030	2035	2040
Ala Glu Val Glu Gly Thr Lys	Ala Pro Val Glu Lys	Glu Glu Val
2045	2050	2055
Lys Val Ser Gly Thr Val Ser	Thr Asn Phe Pro Gln	Thr Ile Glu
2060	2065	2070
Pro Ala Lys Leu Trp Ser Arg	Gln Glu Val Asn Pro	Val Arg Gln
2075	2080	2085
Glu Ile Glu Ser Glu Thr Thr	Ser Glu Glu Gln Ile	Gln Glu Glu
2090	2095	2100
Lys Ser Phe Glu Ser Pro Gln	Asn Ser Pro Ala Thr	Glu Gln Thr
2105	2110	2115
Ile Phe Asp Ser Gln Thr Phe	Thr Glu Thr Glu Leu	Lys Thr Thr
2120	2125	2130
Asp Tyr Ser Val Leu Thr Thr	Lys Lys Thr Tyr Ser	Asp Asp Lys
2135	2140	2145
Glu Met Lys Glu Glu Asp Thr	Ser Leu Val Asn Met	Ser Thr Pro
2150	2155	2160
Asp Pro Asp Ala Asn Gly Leu	Glu Ser Tyr Thr	Leu Pro Glu
2165	2170	2175
Ala Thr Glu Lys Ser His Phe	Phe Leu Ala Thr Ala	Leu Val Thr
2180	2185	2190
Glu Ser Ile Pro Ala Glu His	Val Val Thr Asp Ser	Pro Ile Lys
2195	2200	2205
Lys Glu Glu Ser Thr Lys His	Phe Pro Lys Gly Met	Arg Pro Thr
2210	2215	2220

---

-continued

---

Ile	Gln	Glu	Ser	Asp	Thr	Glu	Leu	Leu	Phe	Ser	Gly	Leu	Gly	Ser
2225						2230					2235			
Gly	Glu	Glu	Val	Leu	Pro	Thr	Leu	Pro	Thr	Glu	Ser	Val	Asn	Phe
2240						2245					2250			
Thr	Glu	Val	Glu	Gln	Ile	Asn	Asn	Thr	Leu	Tyr	Pro	His	Thr	Ser
2255						2260					2265			
Gln	Val	Glu	Ser	Thr	Ser	Ser	Asp	Lys	Ile	Glu	Asp	Phe	Asn	Arg
2270						2275					2280			
Met	Glu	Asn	Val	Ala	Lys	Glu	Val	Gly	Pro	Leu	Val	Ser	Gln	Thr
2285						2290					2295			
Asp	Ile	Phe	Glu	Gly	Ser	Gly	Ser	Val	Thr	Ser	Thr	Thr	Leu	Ile
2300						2305					2310			
Glu	Ile	Leu	Ser	Asp	Thr	Gly	Ala	Glu	Gly	Pro	Thr	Val	Ala	Pro
2315						2320					2325			
Leu	Pro	Phe	Ser	Thr	Asp	Ile	Gly	His	Pro	Gln	Asn	Gln	Thr	Val
2330						2335					2340			
Arg	Trp	Ala	Glu	Glu	Ile	Gln	Thr	Ser	Arg	Pro	Gln	Thr	Ile	Thr
2345						2350					2355			
Glu	Gln	Asp	Ser	Asn	Lys	Asn	Ser	Ser	Thr	Ala	Glu	Ile	Asn	Glu
2360						2365					2370			
Thr	Thr	Thr	Ser	Ser	Thr	Asp	Phe	Leu	Ala	Arg	Ala	Tyr	Gly	Phe
2375						2380					2385			
Glu	Met	Ala	Lys	Glu	Phe	Val	Thr	Ser	Ala	Pro	Lys	Pro	Ser	Asp
2390						2395					2400			
Leu	Tyr	Tyr	Glu	Pro	Ser	Gly	Glu	Gly	Ser	Gly	Glu	Val	Asp	Ile
2405						2410					2415			
Val	Asp	Ser	Phe	His	Thr	Ser	Ala	Thr	Thr	Gln	Ala	Thr	Arg	Gln
2420						2425					2430			
Glu	Ser	Ser	Thr	Thr	Phe	Val	Ser	Asp	Gly	Ser	Leu	Glu	Lys	His
2435						2440					2445			
Pro	Glu	Val	Pro	Ser	Ala	Lys	Ala	Val	Thr	Ala	Asp	Gly	Phe	Pro
2450						2455					2460			
Thr	Val	Ser	Val	Met	Leu	Pro	Leu	His	Ser	Glu	Gln	Asn	Lys	Ser
2465						2470					2475			
Ser	Pro	Asp	Pro	Thr	Ser	Thr	Leu	Ser	Asn	Thr	Val	Ser	Tyr	Glu
2480						2485					2490			
Arg	Ser	Thr	Asp	Gly	Ser	Phe	Gln	Asp	Arg	Phe	Arg	Glu	Phe	Glu
2495						2500					2505			
Asp	Ser	Thr	Leu	Lys	Pro	Asn	Arg	Lys	Lys	Pro	Thr	Glu	Asn	Ile
2510						2515					2520			
Ile	Ile	Asp	Leu	Asp	Lys	Glu	Asp	Lys	Asp	Leu	Ile	Leu	Thr	Ile
2525						2530					2535			
Thr	Glu	Ser	Thr	Ile	Leu	Glu	Ile	Leu	Pro	Glu	Leu	Thr	Ser	Asp
2540						2545					2550			
Lys	Asn	Thr	Ile	Ile	Asp	Ile	Asp	His	Thr	Lys	Pro	Val	Tyr	Glu
2555						2560					2565			
Asp	Ile	Leu	Gly	Met	Gln	Thr	Asp	Ile	Asp	Thr	Glu	Val	Pro	Ser
2570						2575					2580			
Glu	Pro	His	Asp	Ser	Asn	Asp	Glu	Ser	Asn	Asp	Asp	Ser	Thr	Gln
2585						2590					2595			
Val	Gln	Glu	Ile	Tyr	Glu	Ala	Ala	Val	Asn	Leu	Ser	Leu	Thr	Glu
2600						2605					2610			

---

-continued

---

Glu	Thr	Phe	Glu	Gly	Ser	Ala	Asp	Val	Leu	Ala	Ser	Tyr	Thr	Gln
2615														2625
Ala	Thr	His	Asp	Glu	Ser	Met	Thr	Tyr	Glu	Asp	Arg	Ser	Gln	Leu
2630														2640
Asp	His	Met	Gly	Phe	His	Phe	Thr	Thr	Gly	Ile	Pro	Ala	Pro	Ser
2645														2655
Thr	Glu	Thr	Glu	Leu	Asp	Val	Leu	Leu	Pro	Thr	Ala	Thr	Ser	Leu
2660														2670
Pro	Ile	Pro	Arg	Lys	Ser	Ala	Thr	Val	Ile	Pro	Glu	Ile	Glu	Gly
2675														2685
Ile	Lys	Ala	Glu	Ala	Lys	Ala	Leu	Asp	Asp	Met	Phe	Glu	Ser	Ser
2690														2700
Thr	Leu	Ser	Asp	Gly	Gln	Ala	Ile	Ala	Asp	Gln	Ser	Glu	Ile	Ile
2705														2715
Pro	Thr	Leu	Gly	Gln	Phe	Glu	Arg	Thr	Gln	Glu	Glu	Tyr	Glu	Asp
2720														2730
Lys	Lys	His	Ala	Gly	Pro	Ser	Phe	Gln	Pro	Glu	Phe	Ser	Ser	Gly
2735														2745
Ala	Glu	Glu	Ala	Leu	Val	Asp	His	Thr	Pro	Tyr	Leu	Ser	Ile	Ala
2750														2760
Thr	Thr	His	Leu	Met	Asp	Gln	Ser	Val	Thr	Glu	Val	Pro	Asp	Val
2765														2775
Met	Glu	Gly	Ser	Asn	Pro	Pro	Tyr	Tyr	Thr	Asp	Thr	Thr	Leu	Ala
2780														2790
Val	Ser	Thr	Phe	Ala	Lys	Leu	Ser	Ser	Gln	Thr	Pro	Ser	Ser	Pro
2795														2805
Leu	Thr	Ile	Tyr	Ser	Gly	Ser	Glu	Ala	Ser	Gly	His	Thr	Glu	Ile
2810														2820
Pro	Gln	Pro	Ser	Ala	Leu	Pro	Gly	Ile	Asp	Val	Gly	Ser	Ser	Val
2825														2835
Met	Ser	Pro	Gln	Asp	Ser	Phe	Lys	Glu	Ile	His	Val	Asn	Ile	Glu
2840														2850
Ala	Thr	Phe	Lys	Pro	Ser	Ser	Glu	Glu	Tyr	Leu	His	Ile	Thr	Glu
2855														2865
Pro	Pro	Ser	Leu	Ser	Pro	Asp	Thr	Lys	Leu	Glu	Pro	Ser	Glu	Asp
2870														2880
Asp	Gly	Lys	Pro	Glu	Leu	Leu	Glu	Glu	Met	Glu	Ala	Ser	Pro	Thr
2885														2895
Glu	Leu	Ile	Ala	Val	Glu	Gly	Thr	Glu	Ile	Leu	Gln	Asp	Phe	Gln
2900														2910
Asn	Lys	Thr	Asp	Gly	Gln	Val	Ser	Gly	Glu	Ala	Ile	Lys	Met	Phe
2915														2925
Pro	Thr	Ile	Lys	Thr	Pro	Glu	Ala	Gly	Thr	Val	Ile	Thr	Thr	Ala
2930														2940
Asp	Glu	Ile	Glu	Leu	Glu	Gly	Ala	Thr	Gln	Trp	Pro	His	Ser	Thr
2945														2955
Ser	Ala	Ser	Ala	Thr	Tyr	Gly	Val	Glu	Ala	Gly	Val	Val	Pro	Trp
2960														2970
Leu	Ser	Pro	Gln	Thr	Ser	Glu	Arg	Pro	Thr	Leu	Ser	Ser	Ser	Pro
2975														2985
Glu	Ile	Asn	Pro	Glu	Thr	Gln	Ala	Ala	Leu	Ile	Arg	Gly	Gln	Asp

---

-continued

---

2990	2995	3000
Ser Thr Ile Ala Ala Ser Glu Gln Gln Val Ala Ala Arg Ile Leu		
3005	3010	3015
Asp Ser Asn Asp Gln Ala Thr Val Asn Pro Val Glu Phe Asn Thr		
3020	3025	3030
Glu Val Ala Thr Pro Pro Phe Ser Leu Leu Glu Thr Ser Asn Glu		
3035	3040	3045
Thr Asp Phe Leu Ile Gly Ile Asn Glu Glu Ser Val Glu Gly Thr		
3050	3055	3060
Ala Ile Tyr Leu Pro Gly Pro Asp Arg Cys Lys Met Asn Pro Cys		
3065	3070	3075
Leu Asn Gly Gly Thr Cys Tyr Pro Thr Glu Thr Ser Tyr Val Cys		
3080	3085	3090
Thr Cys Val Pro Gly Tyr Ser Gly Asp Gln Cys Glu Leu Asp Phe		
3095	3100	3105
Asp Glu Cys His Ser Asn Pro Cys Arg Asn Gly Ala Thr Cys Val		
3110	3115	3120
Asp Gly Phe Asn Thr Phe Arg Cys Leu Cys Leu Pro Ser Tyr Val		
3125	3130	3135
Gly Ala Leu Cys Glu Gln Asp Thr Glu Thr Cys Asp Tyr Gly Trp		
3140	3145	3150
His Lys Phe Gln Gly Gln Cys Tyr Lys Tyr Phe Ala His Arg Arg		
3155	3160	3165
Thr Trp Asp Ala Ala Glu Arg Glu Cys Arg Leu Gln Gly Ala His		
3170	3175	3180
Leu Thr Ser Ile Leu Ser His Glu Glu Gln Met Phe Val Asn Arg		
3185	3190	3195
Val Gly His Asp Tyr Gln Trp Ile Gly Leu Asn Asp Lys Met Phe		
3200	3205	3210
Glu His Asp Phe Arg Trp Thr Asp Gly Ser Thr Leu Gln Tyr Glu		
3215	3220	3225
Asn Trp Arg Pro Asn Gln Pro Asp Ser Phe Phe Ser Ala Gly Glu		
3230	3235	3240
Asp Cys Val Val Ile Ile Trp His Glu Asn Gly Gln Trp Asn Asp		
3245	3250	3255
Val Pro Cys Asn Tyr His Leu Thr Tyr Thr Cys Lys Lys Gly Thr		
3260	3265	3270
Val Ala Cys Gly Gln Pro Pro Val Val Glu Asn Ala Lys Thr Phe		
3275	3280	3285
Gly Lys Met Lys Pro Arg Tyr Glu Ile Asn Ser Leu Ile Arg Tyr		
3290	3295	3300
His Cys Lys Asp Gly Phe Ile Gln Arg His Leu Pro Thr Ile Arg		
3305	3310	3315
Cys Leu Gly Asn Gly Arg Trp Ala Ile Pro Lys Ile Thr Cys Met		
3320	3325	3330
Asn Pro Ser Ala Tyr Gln Arg Thr Tyr Ser Met Lys Tyr Phe Lys		
3335	3340	3345
Asn Ser Ser Ser Ala Lys Asp Asn Ser Ile Asn Thr Ser Lys His		
3350	3355	3360
Asp His Arg Trp Ser Arg Arg Trp Gln Glu Ser Arg Arg		
3365	3370	3375

---

-continued

---

<210> SEQ ID NO 5  
<211> LENGTH: 7219  
<212> TYPE: DNA  
<213> ORGANISM: Homo sapiens  
  
<400> SEQUENCE: 5

ggaccccggc aagcccgccc acttggcagg agctgttagt accgccccgtcc gcgcctccaa 60  
gtttcacgg ctteccctcaggc agagactcgg gtcgtccgc catgtccgcc gcagacgagg 120  
ttgacgggct gggcgtggcc cggccgcact atggctctgt cctggataat gaaagactta 180  
ctgcagagga gatggatgaa aggagacgtc agaacgtggc ttatgagtagt ctttgtcatt 240  
tggagaaggc gaagaggtgg atggaagcat gccttagggga agatctggct cccaccacag 300  
aactggagga ggggcttagg aatggggct accttgccaa actggggaaac ttcttctc 360  
ccaaagttagt gtccctgaaa aaaatctatg atcgagaaca gaccagatac aaggcgactg 420  
gcctccactt tagacacact gataatgtga ttcagtgggtt gaatgccatg gatgagattt 480  
gattgcctaa gatTTTtac ccagaaacta cagatatcta tgatcggaaag aacatgcca 540  
gatgtatcta ctgtatccat gcacttagt tgtagctttt caagcttaggc ctggccctc 600  
agattcaaga cctatatggaa aagggttact tcacagaaga agaaaatcaac aacatgaaga 660  
ctgagttgga gaagtatggc atccagatgc ctgcctttag caagattggg ggcattttgg 720  
ctaatgaact gtcagtggat gaagccgcat tacatgctgc tgtagtttctt attaatgaag 780  
ctattgaccg tagaattcca gccgacacat ttgcagctt gaaaaatccg aatgcctatgc 840  
ttgtaaatct tgaagagccc ttggcattcca cttaccagga tatactttac caggctaagc 900  
aggacaaaat gacaaatgtt aaaaacagga cagaaaaactc agagagagaa agagatgttt 960  
atgaggagct gtcacgcaa gctgaaattc aaggcaatataa aacaaatgtc aatacatttt 1020  
ctgcatttagt aaatactgac ctggctttagt aacaaggaga tgcactggcc ttgttcagg 1080  
ctctgcagtc accagccctg gggcttcggag gactgcagca acagaatagc gactggact 1140  
tgaagcagct cctgagtgat aaacagcaga agagacagag tggtcagact gacccctgc 1200  
agaaggagga gctgcagttt ggagtggatg ctgcaaacag tgctgcccag caatatcaga 1260  
gaagattggc agcagtagca ctgattaatg ctgcaatcca gaagggtgtt gctgagaaga 1320  
ctgttttggaa actgatgaat cccgaagcccc agctgccccca ggtgtatcca ttgcggcc 1380  
atctctatca gaaggagctg gctaccctgc agcgacaaag tccctgaacat aatctcacc 1440  
acccagagct ctctgtcgca gtggagatgt tgtagctgggtt ggcctgtatc aacaggcat 1500  
tggaatcagg agatgtgaat acagtgtggaa agcaatttgc cagttcagttt actgggttta 1560  
ccaatattga ggaagaaaaac tgtagctggatgtt atctcgatga gttgtatggaa ctgaaggctc 1620  
aggcacatgc agagaataat gaattcatta catggatga tatccaagct tgctggacc 1680  
atgtgaacctt ggtgggtgca gaggaacatgc agaggattt agccatgggtt ttaattatg 1740  
aageccctggaa tgaaggtgat gcccggaaacta ctctgcaggc cctacagattt cctgcagct 1800  
aacttggggg agtccttgca gaagtggccc agcattacca agacacgtctt attagagcga 1860  
agagagagaa agcccgaggaa atccaggatg agtcagctgtt gttatgggtt gatgaaattc 1920  
aagggtggat ctggcagttcc aacaaagaca cccaaagacg acagaagttt gccttaggaa 1980  
tctttgcatt taatgaggca gtagaaatgtt gttatgggtt cttttttttt gttgtggaaact taccacagtg 2040  
gctccccctgaa tttttttttt tatggatgtca tccctgagtg tggtgaaact taccacagtg 2100

---

-continued

---

atcttgctga	agccaagaag	aaaaaaactgg	cagtaggaga	taataacagc	aagtgggtga	2160
agcactgggt	aaaagggtgga	tattattttt	accacaatct	ggagaccagg	gaaggaggat	2220
gggatgaacc	tccaaatttt	gtgcaaaattt	ctatgeagct	ttctcgggag	gagatccaga	2280
gttctatctc	tggggtgact	gcogcatata	accgagaaca	gctgtggctg	gc当地atgaag	2340
gcctgatcac	caggctgcag	gctcgctgcc	gtggatactt	agttcgacag	gaattccgat	2400
ccaggatgaa	tttcoctgaag	aaacaaatcc	ctgcccattcac	ctgcatttcg	tcacagtgg	2460
gaggatataaa	gcagaagaag	gcatatcaag	atcggttagc	ttacctgcgc	tcccaaaag	2520
atgaagttgt	aaagattcag	tccctggcaa	ggatgcacca	agtcgaaag	cgctatcgag	2580
atcgccctgca	gtacttccgg	gaccatataa	atgacattat	caaaatccag	gcttttattc	2640
gggcaaaacaa	agetcggat	gactacaaga	ctctcatcaa	tgctgaggat	cctcctatgg	2700
ttgtggtcgg	aaaatttgc	cacctgctgg	accaaagtga	ccaggatttt	caggaggagc	2760
ttgaccttat	gaagatgegg	gaagaggta	tcaccctcat	tcgttcta	cagcagctgg	2820
agaatgacct	caatctcatg	gatataaaaa	ttggactgct	agtaaaaat	aagattacgt	2880
tgcaggatgt	ggtttcccac	agtaaaaaac	ttacaaaaaa	aaataaggaa	cagttgtctg	2940
atatgatgat	gataaataaa	cagaaggag	gtctcaaggc	tttgagcaag	gagaagagag	3000
agaagttgga	agcttaccag	cacctgtttt	atttattgca	aaccaatccc	acctatctgg	3060
ccaagctcat	tttcagatg	ccccagaaca	agtccaccaa	gttcatggac	tctgtatct	3120
tcacactcta	caactacg	tccaaaccagc	gagaggagta	cctgctcctg	cggtcttta	3180
agacagact	ccaagaggaa	atcaagtca	aggttagatca	gattcaagag	atttgtacag	3240
gaaatcctac	ggtttattaa	atgggtgtaa	gtttcaaccg	tggtgcccgt	ggccagaatg	3300
ccctgagaca	gatcttggcc	ccagtcgtga	aggaaattat	ggatgacaaa	tctctcaaca	3360
tcaaaaactga	ccctgtggat	atttacaat	cttgggtta	tcagatgg	tctcagacag	3420
gagaggcaag	caaactgccc	tatgatgtga	cccctgagca	ggcgctagct	catgaagaag	3480
tgaagacacg	gctagacacg	tccatcagg	acatgcgggc	tgtgacacag	aagtttctct	3540
cagccattgt	cagctctgt	gacaaaatcc	cttatggat	gcgcttcatt	gc当地aaagtgc	3600
tgaaggactc	gttgcattg	aagttccctg	atgctggtga	ggatgagct	ctgaagatta	3660
ttggtaactt	gctttattat	cgatacatga	atccagccat	tggtgctcct	gatgcctttg	3720
acatcattga	cctgtcagca	ggaggccagc	ttaccacaga	ccaacgccc	aatctgggct	3780
ccattgcaaa	aatgcttcag	catgctgtt	ccaataagat	gtttctggga	gataatgccc	3840
acttaagcat	cattaatgaa	tatctttccc	agtcctacca	gaaattcaga	cggttttcc	3900
aaactgcttg	tgtatgtccca	gagcttcagg	ataaaatttaa	tgtggatgag	tactctgatt	3960
tagtaaccct	caccaaacca	gtaatctaca	tttccattgg	tgaatcatc	aacaccacaca	4020
ctctcctgtt	ggatcaccag	gatgccattg	ctccggagca	caatgatcca	atccacgaac	4080
tgctggacga	cctcggcgag	gtgcccacca	tcgagtcct	gatagggaa	agctctggca	4140
atttaaatga	cccaaataag	gaggcactgg	ctaagacgga	agtgtctctc	accctgacca	4200
acaagttcga	cgtgcctgga	gatgagaatg	cagaaatgga	tgctcgaacc	atcttactga	4260
atacaaaaacg	ttaatttgc	gatgtcatcc	ggttccagcc	aggagagacc	ttgactgaaa	4320
tcctagaaac	accagccacc	agtgaacagg	aagcagaaca	tcagagagcc	atgcagagac	4380

---

-continued

---

gtgttatccg	tgatgccaaa	acacctgaca	agataaaaaa	gtcaaatct	gtaaaggaaag	4440
acagcaacct	cactttcaa	gagaagaaag	agaagatcca	gacagggtta	aagaagctaa	4500
cagagcttgg	aaccgtggac	ccaaagaaca	aataccagga	actgatcaac	gacattgcc	4560
gggatattcg	gaatcagcg	aggtaccgac	agagggagaaa	ggccgaacta	gtgaaactgc	4620
aacagacata	cgtgctctg	aactctaagg	ccaccttta	tggggagcag	gtggattact	4680
ataaaagcta	tatcaaaaacc	tgcttggata	acttagccag	caagggcaaa	gtctccaaaa	4740
agcctaggaa	aatgaaagga	aagaaaagca	aaaagatttc	tctgaaatat	acagcagcaa	4800
gactacatga	aaaaggagtt	cttctggaaa	ttgaggacct	gcaagtgaat	cagttaaaa	4860
atgttatatt	tgaatcagt	ccaacagaag	aagttggaga	cttcgaatgc	aaagccaaat	4920
tcatggagt	tcaaattggag	acttttatgt	tacattatca	ggacctgctg	cagctacagt	4980
atgaaggagt	tgcagtcatg	aaattatgg	atagagctaa	agtaaatgtc	aacctcctga	5040
tctcccttct	caacaaaaag	ttctacggg	agtaattgt	cgtttgcgc	cagccagaa	5100
ggatgaagga	aagaagcacc	tcacagctcc	tttcttaggtc	cttctttct	cattggaaagc	5160
aaagacctag	ccaacaacag	cacctaatac	tgatacactc	ccgatgccac	attttaact	5220
cctctcgctc	tgatggaca	tttgttacc	ttttttcata	gtgaaattgt	gtttcaggct	5280
tagtctgacc	tttctgggtt	cttcattttc	ttccattact	taggaaagag	tggaaactcc	5340
actaaaattt	ctctgtgtt	ttacagtctt	agaggttgca	gtactatatt	gtaagctttt	5400
gtgtttgttt	aattagcaat	agggatggta	ggattcaat	gtgtgtcatt	tagaagtgg	5460
agctatttagc	accaatgaca	taaatacata	caagacacac	aactaaaatg	tcatgttatt	5520
aacagttatt	aggttgtcat	ttaaaaataa	agttccttta	tatttctgtc	ccatcaggaa	5580
aactgaagga	tatgggaat	cattggttat	cttccattgt	gtttttctt	atggacagga	5640
gtaatggaa	gtgacagtca	tgttcaaagg	aagcatttct	agaaaaaagg	agataatgtt	5700
tttaatttcc	attatcaaac	ttgggcattt	ctgtttgtgt	aactccccga	ctagtgatgt	5760
ggagagtccc	attgctaaaa	ttcagctact	cagataaatt	cagaatgggt	caaggcacct	5820
gcctgtttt	gttgggtgcac	agagattgac	ttgattcaga	gagacaattc	actccatccc	5880
tatggcagag	gaatgggtt	gccctaattgt	agaatgtcat	tgtttttaaa	actgttttat	5940
atcttaagag	tgcotttatta	aagtatagat	gtatgttcta	aatatgtgggt	gataggaatt	6000
ttaaagattt	atataatgca	tcaaaagct	tagaataaga	aaagcttttt	ttaaattgtct	6060
ttatctgtat	atctgaactc	ttgaaacttta	tagctaaaac	actaggattt	atctgoagt	6120
ttcagggaga	taattctgoc	tttaattgtc	taaaacaaaa	acaaaaccag	ccaaacctatg	6180
ttcacacgtga	gattaaaacc	aatttttcc	ccattttttc	tcttttttc	tcttgctgcc	6240
cacattgtgc	ctttatTTTA	tgagccccag	ttttctgggc	ttagtttaaa	aaaaaaatca	6300
agtctaaaca	ttgcatttag	aaagctttt	ttcttggata	aaaagtcata	cactttaaaa	6360
aaaaaaaaaa	cttttccag	gaaaatata	tgaaatcatg	ctgctgagcc	tctattttct	6420
ttctttgtat	tttgatttca	gtattctttt	atcataaatt	tttagcattt	aaaaattcac	6480
tgtatgtacat	taagccaata	aactgcttta	atgaataaca	aactatgtag	tgtgtcccta	6540
ttataaatgc	attggagaag	tatTTTATG	agactcttta	ctcaggtgca	tggttacagc	6600
ccacagggag	gcatggagtg	ccatggaaagg	attcgccact	acccagacct	tgtttttgt	6660
tgtatTTGG	aagacaggtt	ttttaagaa	acatTTTCT	cagattaaaa	gatgtatgt	6720

---

-continued

---

```

ttacaactag cattgcctca aaaactggga ccaaccaaag tgtgtcaacc ctgtttcctt 6780
aaaagaggct ataatccca aaggccatc ccaagacagg caataatgag cagagttaac 6840
agtccttta ataaaaatgtg tcagtaattt taaggtttat agtccctca acacaattgc 6900
taatgcagaa tagtgtaaaa tgcgcttcaa gaatgttgat gatgtatgata tagaattgtg 6960
gcttagtag cacagaggat gccccaaacaa actcatggcg ttgaaaccac acagttctca 7020
ttactgttat ttatttagtgc tagcatttctc tgcctcctct ctctcctctt ttgacottct 7080
cctcgaccag ccatcatgac atttaccatg aatttacttc ctcccaagag tttggactgc 7140
ccgtcagatt gttgctgcaat atagttgcct ttgtatctct gtatgaaata aaaggtcatt 7200
tgttcatgtt aaaaaaaaaa 7219

```

&lt;210&gt; SEQ ID NO 6

&lt;211&gt; LENGTH: 1657

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Homo sapiens

&lt;400&gt; SEQUENCE: 6

```

Met Ser Ala Ala Asp Glu Val Asp Gly Leu Gly Val Ala Arg Pro His
1 5 10 15

```

```

Tyr Gly Ser Val Leu Asp Asn Glu Arg Leu Thr Ala Glu Glu Met Asp
20 25 30

```

```

Glu Arg Arg Arg Gln Asn Val Ala Tyr Glu Tyr Leu Cys His Leu Glu
35 40 45

```

```

Glu Ala Lys Arg Trp Met Glu Ala Cys Leu Gly Glu Asp Leu Pro Pro
50 55 60

```

```

Thr Thr Glu Leu Glu Glu Gly Leu Arg Asn Gly Val Tyr Leu Ala Lys
65 70 75 80

```

```

Leu Gly Asn Phe Phe Ser Pro Lys Val Val Ser Leu Lys Lys Ile Tyr
85 90 95

```

```

Asp Arg Glu Gln Thr Arg Tyr Lys Ala Thr Gly Leu His Phe Arg His
100 105 110

```

```

Thr Asp Asn Val Ile Gln Trp Leu Asn Ala Met Asp Glu Ile Gly Leu
115 120 125

```

```

Pro Lys Ile Phe Tyr Pro Glu Thr Thr Asp Ile Tyr Asp Arg Lys Asn
130 135 140

```

```

Met Pro Arg Cys Ile Tyr Cys Ile His Ala Leu Ser Leu Tyr Leu Phe
145 150 155 160

```

```

Lys Leu Gly Leu Ala Pro Gln Ile Gln Asp Leu Tyr Gly Lys Val Asp
165 170 175

```

```

Phe Thr Glu Glu Glu Ile Asn Asn Met Lys Thr Glu Leu Glu Lys Tyr
180 185 190

```

```

Gly Ile Gln Met Pro Ala Phe Ser Lys Ile Gly Gly Ile Leu Ala Asn
195 200 205

```

```

Glu Leu Ser Val Asp Glu Ala Ala Leu His Ala Ala Val Ile Ala Ile
210 215 220

```

```

Asn Glu Ala Ile Asp Arg Arg Ile Pro Ala Asp Thr Phe Ala Ala Leu
225 230 235 240

```

```

Lys Asn Pro Asn Ala Met Leu Val Asn Leu Glu Glu Pro Leu Ala Ser
245 250 255

```

```

Thr Tyr Gln Asp Ile Leu Tyr Gln Ala Lys Gln Asp Lys Met Thr Asn
260 265 270

```

---

-continued

---

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

---

-continued

---

675	680	685
Trp Val Lys Gly Gly Tyr Tyr Tyr Tyr His Asn Leu Glu Thr Gln Glu		
690	695	700
Gly Gly Trp Asp Glu Pro Pro Asn Phe Val Gln Asn Ser Met Gln Leu		
705	710	715
Ser Arg Glu Glu Ile Gln Ser Ser Ile Ser Gly Val Thr Ala Ala Tyr		
725	730	735
Asn Arg Glu Gln Leu Trp Leu Ala Asn Glu Gly Leu Ile Thr Arg Leu		
740	745	750
Gln Ala Arg Cys Arg Gly Tyr Leu Val Arg Gln Glu Phe Arg Ser Arg		
755	760	765
Met Asn Phe Leu Lys Lys Gln Ile Pro Ala Ile Thr Cys Ile Gln Ser		
770	775	780
Gln Trp Arg Gly Tyr Lys Gln Lys Lys Ala Tyr Gln Asp Arg Leu Ala		
785	790	795
Tyr Leu Arg Ser His Lys Asp Glu Val Val Lys Ile Gln Ser Leu Ala		
805	810	815
Arg Met His Gln Ala Arg Lys Arg Tyr Arg Asp Arg Leu Gln Tyr Phe		
820	825	830
Arg Asp His Ile Asn Asp Ile Ile Lys Ile Gln Ala Phe Ile Arg Ala		
835	840	845
Asn Lys Ala Arg Asp Asp Tyr Lys Thr Leu Ile Asn Ala Glu Asp Pro		
850	855	860
Pro Met Val Val Val Arg Lys Phe Val His Leu Leu Asp Gln Ser Asp		
865	870	875
Gln Asp Phe Gln Glu Glu Leu Asp Leu Met Lys Met Arg Glu Glu Val		
885	890	895
Ile Thr Leu Ile Arg Ser Asn Gln Gln Leu Glu Asn Asp Leu Asn Leu		
900	905	910
Met Asp Ile Lys Ile Gly Leu Leu Val Lys Asn Lys Ile Thr Leu Gln		
915	920	925
Asp Val Val Ser His Ser Lys Lys Leu Thr Lys Lys Asn Lys Glu Gln		
930	935	940
Leu Ser Asp Met Met Met Ile Asn Lys Gln Lys Gly Gly Leu Lys Ala		
945	950	955
Leu Ser Lys Glu Lys Arg Glu Lys Leu Glu Ala Tyr Gln His Leu Phe		
965	970	975
Tyr Leu Leu Gln Thr Asn Pro Thr Tyr Leu Ala Lys Leu Ile Phe Gln		
980	985	990
Met Pro Gln Asn Lys Ser Thr Lys Phe Met Asp Ser Val Ile Phe Thr		
995	1000	1005
Leu Tyr Asn Tyr Ala Ser Asn Gln Arg Glu Glu Tyr Leu Leu Leu		
1010	1015	1020
Arg Leu Phe Lys Thr Ala Leu Gln Glu Glu Ile Lys Ser Lys Val		
1025	1030	1035
Asp Gln Ile Gln Glu Ile Val Thr Gly Asn Pro Thr Val Ile Lys		
1040	1045	1050
Met Val Val Ser Phe Asn Arg Gly Ala Arg Gly Gln Asn Ala Leu		
1055	1060	1065
Arg Gln Ile Leu Ala Pro Val Val Lys Glu Ile Met Asp Asp Lys		
1070	1075	1080

---

-continued

---

Ser	Leu	Asn	Ile	Lys	Thr	Asp	Pro	Val	Asp	Ile	Tyr	Lys	Ser	Trp
1085					1090							1095		
Val	Asn	Gln	Met	Glu	Ser	Gln	Thr	Gly	Glu	Ala	Ser	Lys	Leu	Pro
1100					1105							1110		
Tyr	Asp	Val	Thr	Pro	Glu	Gln	Ala	Leu	Ala	His	Glu	Glu	Val	Lys
1115					1120						1125			
Thr	Arg	Leu	Asp	Ser	Ser	Ile	Arg	Asn	Met	Arg	Ala	Val	Thr	Asp
1130					1135						1140			
Lys	Phe	Leu	Ser	Ala	Ile	Val	Ser	Ser	Val	Asp	Lys	Ile	Pro	Tyr
1145					1150						1155			
Gly	Met	Arg	Phe	Ile	Ala	Lys	Val	Leu	Lys	Asp	Ser	Leu	His	Glu
1160					1165						1170			
Lys	Phe	Pro	Asp	Ala	Gly	Glu	Asp	Glu	Leu	Leu	Lys	Ile	Ile	Gly
1175					1180						1185			
Asn	Leu	Leu	Tyr	Tyr	Arg	Tyr	Met	Asn	Pro	Ala	Ile	Val	Ala	Pro
1190					1195						1200			
Asp	Ala	Phe	Asp	Ile	Ile	Asp	Leu	Ser	Ala	Gly	Gly	Gln	Leu	Thr
1205					1210						1215			
Thr	Asp	Gln	Arg	Arg	Asn	Leu	Gly	Ser	Ile	Ala	Lys	Met	Leu	Gln
1220					1225						1230			
His	Ala	Ala	Ser	Asn	Lys	Met	Phe	Leu	Gly	Asp	Asn	Ala	His	Leu
1235					1240						1245			
Ser	Ile	Ile	Asn	Glu	Tyr	Leu	Ser	Gln	Ser	Tyr	Gln	Lys	Phe	Arg
1250					1255						1260			
Arg	Phe	Phe	Gln	Thr	Ala	Cys	Asp	Val	Pro	Glu	Leu	Gln	Asp	Lys
1265					1270						1275			
Phe	Asn	Val	Asp	Glu	Tyr	Ser	Asp	Leu	Val	Thr	Leu	Thr	Lys	Pro
1280					1285						1290			
Val	Ile	Tyr	Ile	Ser	Ile	Gly	Glu	Ile	Ile	Asn	Thr	His	Thr	Leu
1295					1300						1305			
Leu	Leu	Asp	His	Gln	Asp	Ala	Ile	Ala	Pro	Glu	His	Asn	Asp	Pro
1310					1315						1320			
Ile	His	Glu	Leu	Leu	Asp	Asp	Leu	Gly	Glu	Val	Pro	Thr	Ile	Glu
1325					1330						1335			
Ser	Leu	Ile	Gly	Glu	Ser	Ser	Gly	Asn	Leu	Asn	Asp	Pro	Asn	Lys
1340					1345						1350			
Glu	Ala	Leu	Ala	Lys	Thr	Glu	Val	Ser	Leu	Thr	Leu	Thr	Asn	Lys
1355					1360						1365			
Phe	Asp	Val	Pro	Gly	Asp	Glu	Asn	Ala	Glu	Met	Asp	Ala	Arg	Thr
1370					1375						1380			
Ile	Leu	Leu	Asn	Thr	Lys	Arg	Leu	Ile	Val	Asp	Val	Ile	Arg	Phe
1385					1390						1395			
Gln	Pro	Gly	Glu	Thr	Leu	Thr	Glu	Ile	Leu	Glu	Thr	Pro	Ala	Thr
1400					1405						1410			
Ser	Glu	Gln	Glu	Ala	Glu	His	Gln	Arg	Ala	Met	Gln	Arg	Arg	Ala
1415					1420						1425			
Ile	Arg	Asp	Ala	Lys	Thr	Pro	Asp	Lys	Met	Lys	Lys	Ser	Lys	Ser
1430					1435						1440			
Val	Lys	Glu	Asp	Ser	Asn	Leu	Thr	Leu	Gln	Glu	Lys	Lys	Glu	Lys
1445					1450						1455			
Ile	Gln	Thr	Gly	Leu	Lys	Lys	Leu	Thr	Glu	Leu	Gly	Thr	Val	Asp
1460					1465						1470			

---

-continued

---

Pro	Lys	Asn	Lys	Tyr	Gln	Glu	Leu	Ile	Asn	Asp	Ile	Ala	Arg	Asp
1475					1480						1485			
Ile	Arg	Asn	Gln	Arg	Arg	Tyr	Arg	Gln	Arg	Arg	Lys	Ala	Glu	Leu
1490					1495						1500			
Val	Lys	Leu	Gln	Gln	Thr	Tyr	Ala	Ala	Leu	Asn	Ser	Lys	Ala	Thr
1505					1510						1515			
Phe	Tyr	Gly	Glu	Gln	Val	Asp	Tyr	Tyr	Lys	Ser	Tyr	Ile	Lys	Thr
1520					1525						1530			
Cys	Leu	Asp	Asn	Leu	Ala	Ser	Lys	Gly	Lys	Val	Ser	Lys	Lys	Pro
1535					1540						1545			
Arg	Glu	Met	Lys	Gly	Lys	Lys	Ser	Lys	Lys	Ile	Ser	Leu	Lys	Tyr
1550					1555						1560			
Thr	Ala	Ala	Arg	Leu	His	Glu	Lys	Gly	Val	Leu	Leu	Glu	Ile	Glu
1565					1570						1575			
Asp	Leu	Gln	Val	Asn	Gln	Phe	Lys	Asn	Val	Ile	Phe	Glu	Ile	Ser
1580					1585						1590			
Pro	Thr	Glu	Glu	Val	Gly	Asp	Phe	Glu	Val	Lys	Ala	Lys	Phe	Met
1595					1600						1605			
Gly	Val	Gln	Met	Glu	Thr	Phe	Met	Leu	His	Tyr	Gln	Asp	Leu	Leu
1610					1615						1620			
Gln	Leu	Gln	Tyr	Glu	Gly	Val	Ala	Val	Met	Lys	Leu	Phe	Asp	Arg
1625					1630						1635			
Ala	Lys	Val	Asn	Val	Asn	Leu	Leu	Ile	Phe	Leu	Leu	Asn	Lys	Lys
1640					1645						1650			
Phe	Tyr	Gly	Lys											
1655														

&lt;210&gt; SEQ ID NO 7

&lt;211&gt; LENGTH: 847

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Homo sapiens

&lt;400&gt; SEQUENCE: 7

acagagtaaa	ctttgctgg	gttccaatg	accggccata	gtttattata	aagggtactg	60
caccctgcag	ccaccagcac	tgcctggc	cacgtgcctc	ctggcttc	tatggcgctg	120
tcctgggttc	ttacagtct	gagcctcota	cctctgttgg	aagcccagat	cccatgttgt	180
gccaacctag	taccggtgcc	catcaccaac	gccaccctgg	accggatcac	tggcaagtgg	240
ttttatatacg	catcgccctt	tgcggaaacgag	gagtcacaata	agtcgggtca	ggagatccaa	300
gcaacacctt	tttacttca	ccccaaacaag	acagaggaca	cgatcttct	cagagactac	360
cagacccgac	aggaccatgt	catctataac	accacatcacc	tgaatgtcca	gcggggaaaat	420
gggaccatct	ccagatacgt	gggaggccaa	gagcatttcg	ctcaatttgc	gatcctcagg	480
gacaccaaga	cctacatgct	tgcgtttgc	gtgaacatg	agaagaactg	ggggctgtct	540
gtctatgctg	acaagccaga	gacgaccaag	gagcaactgg	gagatctcta	cgaagctctc	600
gactgcttgc	gcattccaa	gtcagatgtc	gtgtacaccg	attggaaaaaa	ggataagtgt	660
gagccactgg	agaagcagca	cgagaaggag	aggaaacagg	aggaggggaa	atccttagcag	720
gacacagcct	tggatcagga	cagagacttg	ggggccatcc	tgccttc	acccgacatg	780
tgtacctcag	ctttttccct	cacttgcac	aataaagctt	ctgtgtttgg	aacagactaaa	840
aaaaaaaa						847

---

-continued

---

```

<210> SEQ ID NO 8
<211> LENGTH: 201
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 8

Met Ala Leu Ser Trp Val Leu Thr Val Leu Ser Leu Leu Pro Leu Leu
1           5           10          15

Glu Ala Gln Ile Pro Leu Cys Ala Asn Leu Val Pro Val Pro Ile Thr
20          25          30

Asn Ala Thr Leu Asp Arg Ile Thr Gly Lys Trp Phe Tyr Ile Ala Ser
35          40          45

Ala Phe Arg Asn Glu Glu Tyr Asn Lys Ser Val Gln Glu Ile Gln Ala
50          55          60

Thr Phe Phe Tyr Phe Thr Pro Asn Lys Thr Glu Asp Thr Ile Phe Leu
65          70          75          80

Arg Glu Tyr Gln Thr Arg Gln Asp Gln Cys Ile Tyr Asn Thr Thr Tyr
85          90          95

Leu Asn Val Gln Arg Glu Asn Gly Thr Ile Ser Arg Tyr Val Gly Gly
100         105         110

Gln Glu His Phe Ala His Leu Ile Leu Arg Asp Thr Lys Thr Tyr
115         120         125

Met Leu Ala Phe Asp Val Asn Asp Glu Lys Asn Trp Gly Leu Ser Val
130         135         140

Tyr Ala Asp Lys Pro Glu Thr Thr Lys Glu Gln Leu Gly Glu Phe Tyr
145         150         155         160

Glu Ala Leu Asp Cys Leu Arg Ile Pro Lys Ser Asp Val Val Tyr Thr
165         170         175

Asp Trp Lys Asp Lys Cys Glu Pro Leu Glu Lys Gln His Glu Lys
180         185         190

Glu Arg Lys Gln Glu Glu Gly Glu Ser
195         200

```

```

<210> SEQ ID NO 9
<211> LENGTH: 1447
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

tgtcaactgag ggttgactga ctggagagct caagtgcagc aaagagaagt gtcagagcat      60
gagcgccaaag tccagaaacca tagggattat tggagctctt ttctcaaagg gacagccacg      120
aggaggggtg gaagaaggcc ctacagtatt gagaaaaggtt ggtctgcttg agaaaacttaa      180
agaacaagag tgtgatgtga aggattatgg ggacctgccc tttgctgaca tccctaata      240
cagtcctttt caaattgtga agaatccaag gtctgtggga aaagcaagcg agcagctggc      300
tggcaagggtg gcagaagtca agaagaacgg aagaatcagc ctggtgctgg gcggagacca      360
cagtttggca attgaaagca tctctggcca tgccagggtc caccctgatc ttggagtc      420
ctgggtggat gctcacactg atatcaacac tccactgaca accacaagtg gaaacttgca      480
tggacaacct gtatcttcc tcctgaagga actaaaagga aagattcccg atgtgcagg      540
attctccctg gtgactccct gtatatctgc caaggatatt gtgtatattg gcttgagaga      600
cgtggaccct ggggaacact acatttgaa aactcttaggc attaaataact tttcaatgac      660

```

---

-continued

---

tgaagtggac agacttagaa ttggcaaggt gatggaagaa acactcagct atctactagg	720
aagaagaaaa aggccaattc atctaagttt tgcgttgac ggactggacc catcttcac	780
accagctact ggcacaccag tcgtggggg tctgacatac agagaaggc tctacatcac	840
agaagaaaatc tacaaaacag ggctactctc aggatttagat ataatgaaag tgaaccatc	900
cctggggaaag acaccagaag aagtaactcg aacagtgaac acagcagttg caataacctt	960
ggcttggttc ggacttgctc gggagggtaa tcacaagcct attgactacc ttaaccacc	1020
taagtaaatg tggaaacatc cgatataaaat ctcatagtta atggcataat tagaaagcta	1080
atcattttct taagcataga gttatccctc taaagacttg ttcttcaga aaaatgtttt	1140
tccaaattgt ataaactcta caaattccct cttgggttaa aattcaagat gtggaaattc	1200
taactttttt gaaatttaaa agtttatatt ttcttaacttg gcaaaagact tatccttaga	1260
aagagaagtg tacattgatt tccaattaaa aatttgcgg cattaaaaat aagcacactt	1320
acataagccc ccatacatag agtgggactc ttggaatcag gagacaaagc taccacatgt	1380
ggaaaggtac tatgtgtcca tgtcattcaa aaaatgtgat ttttataat aaactcttta	1440
taacaag	1447

&lt;210&gt; SEQ ID NO 10

&lt;211&gt; LENGTH: 322

&lt;212&gt; TYPE: PRT

&lt;213&gt; ORGANISM: Homo sapiens

&lt;400&gt; SEQUENCE: 10

Met Ser Ala Lys Ser Arg Thr Ile Gly Ile Ile Gly Ala Pro Phe Ser			
1	5	10	15

Lys Gly Gln Pro Arg Gly Gly Val Glu Glu Gly Pro Thr Val Leu Arg		
20	25	30

Lys Ala Gly Leu Leu Glu Lys Leu Lys Glu Gln Glu Cys Asp Val Lys		
35	40	45

Asp Tyr Gly Asp Leu Pro Phe Ala Asp Ile Pro Asn Asp Ser Pro Phe		
50	55	60

Gln Ile Val Lys Asn Pro Arg Ser Val Gly Lys Ala Ser Glu Gln Leu			
65	70	75	80

Ala Gly Lys Val Ala Glu Val Lys Lys Asn Gly Arg Ile Ser Leu Val		
85	90	95

Leu Gly Gly Asp His Ser Leu Ala Ile Gly Ser Ile Ser Gly His Ala		
100	105	110

Arg Val His Pro Asp Leu Gly Val Ile Trp Val Asp Ala His Thr Asp		
115	120	125

Ile Asn Thr Pro Leu Thr Thr Ser Gly Asn Leu His Gly Gln Pro		
130	135	140

Val Ser Phe Leu Leu Lys Glu Leu Lys Gly Lys Ile Pro Asp Val Pro			
145	150	155	160

Gly Phe Ser Trp Val Thr Pro Cys Ile Ser Ala Lys Asp Ile Val Tyr		
165	170	175

Ile Gly Leu Arg Asp Val Asp Pro Gly Glu His Tyr Ile Leu Lys Thr		
180	185	190

Leu Gly Ile Lys Tyr Phe Ser Met Thr Glu Val Asp Arg Leu Gly Ile		
195	200	205

Gly Lys Val Met Glu Glu Thr Leu Ser Tyr Leu Leu Gly Arg Lys Lys	
---	--

---

-continued

---

210	215	220
Arg Pro Ile His Leu Ser Phe Asp Val Asp Gly Leu Asp Pro Ser Phe		
225	230	235
Thr Pro Ala Thr Gly Thr Pro Val Val Gly Gly Leu Thr Tyr Arg Glu		
245	250	255
Gly Leu Tyr Ile Thr Glu Glu Ile Tyr Lys Thr Gly Leu Leu Ser Gly		
260	265	270
Leu Asp Ile Met Glu Val Asn Pro Ser Leu Gly Lys Thr Pro Glu Glu		
275	280	285
Val Thr Arg Thr Val Asn Thr Ala Val Ala Ile Thr Leu Ala Cys Phe		
290	295	300
Gly Leu Ala Arg Glu Gly Asn His Lys Pro Ile Asp Tyr Leu Asn Pro		
305	310	315
Pro Lys		

<210> SEQ ID NO 11  
<211> LENGTH: 619  
<212> TYPE: DNA  
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 11

agttaaatct tttctgctta ctgaaaagga agagtctgat gattagttac tgatcccttt	60
tgcatttgta aagcttttgg aatatttata catgttacca tttctgtttt ttccaccct	120
gttttcttcc atatttactg aagctcagaa gcagtttgg gtctgcaact catccgatgc	180
aagtatttca tacacctact gtgataaaat gcaataccca atttcaatta atgttaaccc	240
ctgtatagaa ttgaaaagat ccaaaggatt attgcacatt ttctacattt caaggagaga	300
tttaaagcaa ttatatttca atctctatat aactgtcaac accatgaatc ttccaaagcg	360
caaagaagtt atttgcgag gatctgtatca cgattactt ttttgcagag ctctgaaggg	420
agagactgtg aataacaacaa tatcattctc cttcaaggaa ataaaaatttt ctaaggaaaa	480
atacaaatgt gttgttgaag ctatttctgg gagcccagaa gaaatgctt tttgttttgg	540
gtttgtcatc ctacaccaac ctaattcaaa tttagataaa ttgagtattt aaaaaaaaaa	600
aaaaaaaaaaaa aaaaaaaaaa	619

<210> SEQ ID NO 12  
<211> LENGTH: 2387  
<212> TYPE: DNA  
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 12

agacacctct gcccctcacca tgagcctctg gcagccccctg gtcctgggtgc tcctggtgct	60
gggctgctgc tttgctgecc ccagacagcg ccagtccacc cttgtgtct tcctggaga	120
cctgagaacc aatctcaccc acaggcagct ggcagaggaa tacctgtacc gctatggta	180
cactcgggtg gcagagatgc gtggagagtc gaaatctctg gggcctgcgc tgctgtttct	240
ccagaagcaa ctgtccctgc ccgagaccgg tgagctggat agcgccacgc tgaaggccat	300
gcgaacccca cggtgccggg tcccagacct gggcagatcc caaaccttg agggcgacct	360
caagtggcac caccacaaca tcacctattt gatccaaaac tactcggaa acttgcgcgc	420
ggcggtgatt gacgacgcct ttgccccgcgc cttcgcactg tggagcgcgg tgacgcgcgt	480
cacccctact cgcgtgtaca gcccggacgc agacatgc tc atccagttt gttgtcgccga	540

---

-continued

---

gcacggagac	gggttatccct	tcgacggaa	ggacgggctc	ctggcacacg	ccttccctcc	600
tggccccggc	attcaggag	acgcccatt	cgacgatgac	gagttgttgt	ccctgggcaa	660
gggcgtcg	gttccaactc	ggttgaaa	cgcagatgac	gcccgtcg	acttccctt	720
catcttcgag	ggcggctct	actctgcct	caccacggac	ggtcgtccg	acggcttgcc	780
ctggtgca	accacggca	actacgacac	cgacgacccg	tttggttct	gccccagcga	840
gagactctac	accoaggacg	gcaatgtca	tggaaaccc	tgccagtttc	cattcatctt	900
ccaaggccaa	tcctactcg	cctgcaccac	ggacggtcgc	tccgacggct	accgtggtg	960
cgccaccacc	gccaactacg	accgggacaa	gtcttcggc	ttctgeccga	cccgagctga	1020
ctcgacggtg	atggggggca	actcggeggg	ggagctgtgc	gtctccct	tcactttcct	1080
gggttaaggag	tactcgacct	gtaccagcga	ggggcgggga	gatgggggcc	tctggtgccg	1140
taccacctcg	aactttgaca	gcaacaagaa	gtggggctc	tgcccgacc	aaggatacag	1200
tttgttccctc	gtggccggc	atgagttcgg	ccacgcgtcg	ggcttagatc	attcctcagt	1260
gccggagggcg	ctcatgtacc	ctatgtacc	cttcactgag	ggggccccct	tgcataagga	1320
cgacgtgaat	ggcatccggc	acctctatgg	tcctcgccct	gaacctgagc	cacggctcc	1380
aaccaccacc	acacccgc	ccacggctcc	cccgacggc	tgccccaccc	gaccccccac	1440
tgtccacccc	tcagagcgc	ccacagctgg	ccccacaggt	ccccctcag	ctggcccccac	1500
aggccccccc	actgctggcc	cttctaeggc	cactactgtg	cctttgagtc	cggtggacga	1560
tgcctgcaac	gtgaacatct	tcgacgccc	cgccggagatt	gggaaccaggc	tgtatttgtt	1620
caaggatggg	aagtactggc	gattctctga	gggcagggggg	agccggccgc	aggcccctt	1680
ccttatacgcc	gacaagtggc	cegcgcgtcc	ccgcaagctg	gactcggtct	ttgaggagcg	1740
gctctccaag	aagctttct	tcttctctgg	gcccggatgt	tgggtgtaca	caggcgcgtc	1800
ggtgctggc	ccgaggcg	tggacaagct	gggcctggga	gccgacgtgg	cccaggatgac	1860
cggggccctc	cggagtggc	ggggaaagat	gtgtgtgttc	agcggggccgc	gcctctggag	1920
gttcgacgtg	aaggcgcaga	tgggtgtacc	ccggagcgcc	agcgaggatgg	accggatgtt	1980
ccccggggtg	ccttggaca	cgcacgacgt	cttccagtac	cgagagaaag	cctatttctg	2040
ccaggaccgc	ttctactggc	cgctgagttc	ccggaggatgg	ttgaaccagg	tggaccaagt	2100
gggtacgtg	acatatgaca	tcctgcgt	ccctgaggac	tagggctccc	gtcctgtctt	2160
ggcagtgc	tgtaaatccc	cactggacc	aaccctgggg	aggaggccag	tttgcggat	2220
acaaacttgt	attctgttct	ggaggaaagg	gaggagtgga	ggtgggttgg	gcccctcttt	2280
ctcacctttg	tttttgttg	gagtgtttct	aataaacttg	gattctctaa	cctttaaaaa	2340
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	2387

<210> SEQ ID NO 13  
<211> LENGTH: 707  
<212> TYPE: PRT  
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 13

Met	Ser	Leu	Trp	Gln	Pro	Leu	Val	Leu	Val	Leu	Leu	Val	Lys	Cys	
1					5			10			15				
Cys	Phe	Ala	Ala	Pro	Arg	Gln	Arg	Gln	Ser	Thr	Leu	Val	Leu	Phe	Pro
								20			25			30	

---

-continued

---

Gly	Asp	Leu	Arg	Thr	Asn	Leu	Thr	Asp	Arg	Gln	Leu	Ala	Glu	Glu	Tyr
35						40					45				
Leu	Tyr	Arg	Tyr	Gly	Tyr	Thr	Arg	Val	Ala	Glu	Met	Arg	Gly	Glu	Ser
50						55				60					
Lys	Ser	Leu	Gly	Pro	Ala	Leu	Leu	Leu	Gln	Gln	Leu	Ser	Leu		
65						70			75		80				
Pro	Glu	Thr	Gly	Glu	Leu	Asp	Ser	Ala	Thr	Leu	Lys	Ala	Met	Arg	Thr
85						90			95						
Pro	Arg	Cys	Gly	Val	Pro	Asp	Leu	Gly	Arg	Phe	Gln	Thr	Phe	Glu	Gly
100						105				110					
Asp	Leu	Lys	Trp	His	His	Asn	Ile	Thr	Tyr	Trp	Ile	Gln	Asn	Tyr	
115						120					125				
Ser	Glu	Asp	Leu	Pro	Arg	Ala	Val	Ile	Asp	Asp	Ala	Phe	Ala	Arg	Ala
130						135				140					
Phe	Ala	Leu	Trp	Ser	Ala	Val	Thr	Pro	Leu	Thr	Phe	Thr	Arg	Val	Tyr
145						150			155		160				
Ser	Arg	Asp	Ala	Asp	Ile	Val	Ile	Gln	Phe	Gly	Val	Ala	Glu	His	Gly
165						170				175					
Asp	Gly	Tyr	Pro	Phe	Asp	Gly	Lys	Asp	Gly	Leu	Leu	Ala	His	Ala	Phe
180						185				190					
Pro	Pro	Gly	Pro	Gly	Ile	Gln	Gly	Asp	Ala	His	Phe	Asp	Asp	Asp	Glu
195						200			205						
Leu	Trp	Ser	Leu	Gly	Lys	Gly	Val	Val	Val	Pro	Thr	Arg	Phe	Gly	Asn
210						215				220					
Ala	Asp	Gly	Ala	Ala	Cys	His	Phe	Pro	Phe	Ile	Phe	Glu	Gly	Arg	Ser
225						230			235		240				
Tyr	Ser	Ala	Cys	Thr	Thr	Asp	Gly	Arg	Ser	Asp	Gly	Leu	Pro	Trp	Cys
245						250			255						
Ser	Thr	Thr	Ala	Asn	Tyr	Asp	Thr	Asp	Asp	Arg	Phe	Gly	Phe	Cys	Pro
260						265			270						
Ser	Glu	Arg	Leu	Tyr	Thr	Gln	Asp	Gly	Asn	Ala	Asp	Gly	Lys	Pro	Cys
275						280			285						
Gln	Phe	Pro	Phe	Ile	Phe	Gln	Gly	Gln	Ser	Tyr	Ser	Ala	Cys	Thr	Thr
290						295			300						
Asp	Gly	Arg	Ser	Asp	Gly	Tyr	Arg	Trp	Cys	Ala	Thr	Thr	Ala	Asn	Tyr
305						310			315		320				
Asp	Arg	Asp	Lys	Leu	Phe	Gly	Phe	Cys	Pro	Thr	Arg	Ala	Asp	Ser	Thr
325						330			335						
Val	Met	Gly	Gly	Asn	Ser	Ala	Gly	Glu	Leu	Cys	Val	Phe	Pro	Phe	Thr
340						345				350					
Phe	Leu	Gly	Lys	Glu	Tyr	Ser	Thr	Cys	Thr	Ser	Glu	Gly	Arg	Gly	Asp
355						360			365						
Gly	Arg	Leu	Trp	Cys	Ala	Thr	Thr	Ser	Asn	Phe	Asp	Ser	Asp	Lys	Lys
370						375			380						
Trp	Gly	Phe	Cys	Pro	Asp	Gln	Gly	Tyr	Ser	Leu	Phe	Leu	Val	Ala	Ala
385						390			395		400				
His	Glu	Phe	Gly	His	Ala	Leu	Gly	Leu	Asp	His	Ser	Ser	Val	Pro	Glu
405						410			415						
Ala	Leu	Met	Tyr	Pro	Met	Tyr	Arg	Phe	Thr	Glu	Gly	Pro	Pro	Leu	His
420						425			430						
Lys	Asp	Asp	Val	Asn	Gly	Ile	Arg	His	Leu	Tyr	Gly	Pro	Arg	Pro	Glu
435						440			445						

---

-continued

---

```

Pro Glu Pro Arg Pro Pro Thr Thr Thr Pro Gln Pro Thr Ala Pro
450          455          460

Pro Thr Val Cys Pro Thr Gly Pro Pro Thr Val His Pro Ser Glu Arg
465          470          475          480

Pro Thr Ala Gly Pro Thr Gly Pro Pro Ser Ala Gly Pro Thr Gly Pro
485          490          495

Pro Thr Ala Gly Pro Ser Thr Ala Thr Thr Val Pro Leu Ser Pro Val
500          505          510

Asp Asp Ala Cys Asn Val Asn Ile Phe Asp Ala Ile Ala Glu Ile Gly
515          520          525

Asn Gln Leu Tyr Leu Phe Lys Asp Gly Lys Tyr Trp Arg Phe Ser Glu
530          535          540

Gly Arg Gly Ser Arg Pro Gln Gly Pro Phe Leu Ile Ala Asp Lys Trp
545          550          555          560

Pro Ala Leu Pro Arg Lys Leu Asp Ser Val Phe Glu Glu Arg Leu Ser
565          570          575

Lys Lys Leu Phe Phe Ser Gly Arg Gln Val Trp Val Tyr Thr Gly
580          585          590

Ala Ser Val Leu Gly Pro Arg Arg Leu Asp Lys Leu Gly Leu Gly Ala
595          600          605

Asp Val Ala Gln Val Thr Gly Ala Leu Arg Ser Gly Arg Gly Lys Met
610          615          620

Leu Leu Phe Ser Gly Arg Arg Leu Trp Arg Phe Asp Val Lys Ala Gln
625          630          635          640

Met Val Asp Pro Arg Ser Ala Ser Glu Val Asp Arg Met Phe Pro Gly
645          650          655

Val Pro Leu Asp Thr His Asp Val Phe Gln Tyr Arg Glu Lys Ala Tyr
660          665          670

Phe Cys Gln Asp Arg Phe Tyr Trp Arg Val Ser Ser Arg Ser Glu Leu
675          680          685

Asn Gln Val Asp Gln Val Gly Tyr Val Thr Tyr Asp Ile Leu Gln Cys
690          695          700

Pro Glu Asp
705

```

```

<210> SEQ ID NO 14
<211> LENGTH: 160
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 14

Met Leu Pro Phe Leu Phe Ser Thr Leu Phe Ser Ser Ile Phe Thr
1          5          10          15

Glu Ala Gln Lys Gln Tyr Trp Val Cys Asn Ser Ser Asp Ala Ser Ile
20         25          30

Ser Tyr Thr Tyr Cys Asp Lys Met Gln Tyr Pro Ile Ser Ile Asn Val
35         40          45

Asn Pro Cys Ile Glu Leu Lys Arg Ser Lys Gly Leu Leu His Ile Phe
50         55          60

Tyr Ile Pro Arg Arg Asp Leu Lys Gln Leu Tyr Phe Asn Leu Tyr Ile
65         70          75          80

Thr Val Asn Thr Met Asn Leu Pro Lys Arg Lys Glu Val Ile Cys Arg
85         90          95

```

---

-continued

---

Gly Ser Asp Asp Asp Tyr Ser Phe Cys Arg Ala Leu Lys Gly Glu Thr  
100 105 110

Val Asn Thr Thr Ile Ser Phe Ser Phe Lys Gly Ile Lys Phe Ser Lys  
115 120 125

Gly Lys Tyr Lys Cys Val Val Glu Ala Ile Ser Gly Ser Pro Glu Glu  
130 135 140

Met Leu Phe Cys Leu Glu Phe Val Ile Leu His Gln Pro Asn Ser Asn  
145 150 155 160

<210> SEQ ID NO 15

<211> LENGTH: 1165

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 15

cgcataaaaa cccaggccgg caggatcgct gcacccgggg cggcctcctc ggtgcggcac	60
ccccggctca gaggacttt tgctgtcccc caagatgccc atgctgtgg cgcttcgtgc	120
ccttcggcgg gcgcggccat cggccagtgc agagtcacac tggtgctacg aggttcaagc	180
cggatcccttc aactaccctt gcttggtgcc agtcaagtgg ggtggaaact gccagaagga	240
ccggccaggccc cccatcaaca tcgtcaccac caaggcaaag gtggacaaaa aactgggacg	300
ctttttcttc tctggctacg ataagaagca aacgtggact gtccaaaata acgggcactc	360
agtgtatgtt ttgctggaga acaaggccag catttctgga ggaggactgc ctgccccata	420
ccaggccaaa cagttgcacc tgcactggc cgacttgcata tataagggtt cggagcacag	480
cctcgatggg gagcactttt ccatggagat gcacatagta catgagaaag agaaggggac	540
atcgaggaat gtgaaagagg cccaggaccc tgaagacaa attgcgggtc tggccttct	600
ggtgtggggct ggaacccagg tgaacgaggg cttccagcca ctggtgagg cactgtctaa	660
tatccccaaa cctgagatga gcactacgat ggcagagac agcctgtgg acctgtccc	720
caaggaggag aaactgaggc actacttccg ctacctggc tcactcacca caccgacctg	780
cgatgagaag gtcgtcttga ctgtgttccc ggagccatt cagttcaca gagaacagat	840
cctggcatcc ttcagaagc tgtactacga caaggaacag acagtggca tgaaggacaa	900
tgtcaggccc ctgcagcagc tggggcagcg cacgggtata aagtccgggg cccgggtcg	960
ggcgctgccc tggccctgc ctggcctgct gggcccccatt ctggcctgcc tgctggccgg	1020
cttcctgcga tgatggctca cttctgcacg cagcctctct gttgcctcag ctctccaaat	1080
tccaggcttc cggtccttag cttccagg tgggacttta ggcatgatta aaatatggac	1140
atatttttgg agaaaaaaaaaaaaaaa	1165

<210> SEQ ID NO 16

<211> LENGTH: 312

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 16

Met Arg Met Leu Leu Ala Leu Leu Ala Leu Ser Ala Ala Arg Pro Ser  
1 5 10 15

Ala Ser Ala Glu Ser His Trp Cys Tyr Glu Val Gln Ala Glu Ser Ser  
20 25 30

Asn Tyr Pro Cys Leu Val Pro Val Lys Trp Gly Gly Asn Cys Gln Lys  
35 40 45

---

-continued

---

Asp Arg Gln Ser Pro Ile Asn Ile Val Thr Thr Lys Ala Lys Val Asp  
 50 55 60  
 Lys Lys Leu Gly Arg Phe Phe Ser Gly Tyr Asp Lys Lys Gln Thr  
 65 70 75 80  
 Trp Thr Val Gln Asn Asn Gly His Ser Val Met Met Leu Leu Glu Asn  
 85 90 95  
 Lys Ala Ser Ile Ser Gly Gly Leu Pro Ala Pro Tyr Gln Ala Lys  
 100 105 110  
 Gln Leu His Leu His Trp Ser Asp Leu Pro Tyr Lys Gly Ser Glu His  
 115 120 125  
 Ser Leu Asp Gly Glu His Phe Ala Met Glu Met His Ile Val His Glu  
 130 135 140  
 Lys Glu Lys Gly Thr Ser Arg Asn Val Lys Glu Ala Gln Asp Pro Glu  
 145 150 155 160  
 Asp Glu Ile Ala Val Leu Ala Phe Leu Val Glu Ala Gly Thr Gln Val  
 165 170 175  
 Asn Glu Gly Phe Gln Pro Leu Val Glu Ala Leu Ser Asn Ile Pro Lys  
 180 185 190  
 Pro Glu Met Ser Thr Thr Met Ala Glu Ser Ser Leu Leu Asp Leu Leu  
 195 200 205  
 Pro Lys Glu Glu Lys Leu Arg His Tyr Phe Arg Tyr Leu Gly Ser Leu  
 210 215 220  
 Thr Thr Pro Thr Cys Asp Glu Lys Val Val Trp Thr Val Phe Arg Glu  
 225 230 235 240  
 Pro Ile Gln Leu His Arg Glu Gln Ile Leu Ala Phe Ser Gln Lys Leu  
 245 250 255  
 Tyr Tyr Asp Lys Glu Gln Thr Val Ser Met Lys Asp Asn Val Arg Pro  
 260 265 270  
 Leu Gln Gln Leu Gly Gln Arg Thr Val Ile Lys Ser Gly Ala Pro Gly  
 275 280 285  
 Arg Pro Leu Pro Trp Ala Leu Pro Ala Leu Leu Gly Pro Met Leu Ala  
 290 295 300  
 Cys Leu Leu Ala Gly Phe Leu Arg  
 305 310

```

<210> SEQ_ID NO 17
<211> LENGTH: 466
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 17

accactgctg gcttttgct gtagctccac attcctgtgc attgaggggt taacattagg 60
ctggaaagat gacaaaactt gaagagcatc tggagggaat tgtcaatatac ttccaccaat 120
actcagttcg gaaggggcat tttgacaccc tctctaaggg tgagctgaag cagctgctta 180
caaaggagct tgcaaacacc atcaagaata tcaaagataa agctgtcatt gatgaaatat 240
tccaaggcct ggatgctaat caagatgaac aggtcgactt tcaagaattc atatccctgg 300
tagccattgc gctgaaggct gcccattacc acacccacaa agagtaggta gctctctgaa 360
ggcttttac ccagcaatgt cctcaatgag ggtctttct ttccctcacc aaaaccagc 420
cttgcccgta gggagtaaga gttaataaac acactcacga aaagt 466
  
```

---

-continued

---

```

<210> SEQ ID NO 18
<211> LENGTH: 92
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 18

Met Thr Lys Leu Glu Glu His Leu Glu Gly Ile Val Asn Ile Phe His
1 5 10 15

Gln Tyr Ser Val Arg Lys Gly His Phe Asp Thr Leu Ser Lys Gly Glu
20 25 30

Leu Lys Gln Leu Leu Thr Lys Glu Leu Ala Asn Thr Ile Lys Asn Ile
35 40 45

Lys Asp Lys Ala Val Ile Asp Glu Ile Phe Gln Gly Leu Asp Ala Asn
50 55 60

Gln Asp Glu Gln Val Asp Phe Gln Glu Phe Ile Ser Leu Val Ala Ile
65 70 75 80

Ala Leu Lys Ala Ala His Tyr His Thr His Lys Glu
85 90

```

```

<210> SEQ ID NO 19
<211> LENGTH: 10
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
    oligonucleotide

<400> SEQUENCE: 19

tttttttttt 10

```

---

What is claimed is:

**1.** A method for the diagnosis of acute ischemic stroke in a subject, comprising detecting in a biological sample obtained from the subject one or more biomarkers selected from the group consisting of:

- (a) chemokine receptor 7 (CCR7);
- (b) chondroitin sulfate proteoglycan 2 (CSPG2);
- (c) IQ motif-containing GTPase activation protein 1 (IQ-GAP1); and
- (d) orosomucoid 1 (ORM1).

**2.** A method for the diagnosis of acute ischemic stroke in a subject, comprising detecting in a biological sample obtained from the subject two or more biomarkers selected from the group consisting of:

- (a) chemokine receptor 7 (CCR7);
  - (b) chondroitin sulfate proteoglycan 2 (CSPG2);
  - (c) IQ motif-containing GTPase activation protein 1 (IQ-GAP1);
  - (d) orosomucoid 1 (ORM1)
  - (e) arginase 1 (ARG1);
  - (f) lymphocyte antigen 96 (LY96);
  - (g) matrix metalloproteinase 9 (MMP9);
  - (h) carbonic anhydrase 4 (CA4); and
  - (i) s100 calcium binding protein A12 (s100A12),
- wherein at least one of the biomarkers is (a), (b), (c) or (d).

**3.** The method of claim **1** or **2**, wherein the sample is whole peripheral blood obtained from the subject.

**4.** The method of claim **1** or **2**, wherein the method is executed on the subject no more than 3 hours after onset of presenting acute ischemic stroke symptoms.

**5.** The method of claim **1** or **2**, wherein the method is executed on the subject no more than 4.5 hours after onset of presenting acute ischemic stroke symptoms.

**6.** The method of claim **1** or **2**, wherein the detecting step further comprises contacting the biological sample with a detection means capable of detecting the biomarker.

**7.** The method of claim **6**, wherein the biomarker is a nucleic acid molecule.

**8.** The method of claim **7**, wherein the nucleic acid molecule is mRNA.

**9.** The method of claim **6**, wherein the biomarker is a polypeptide.

**10.** The method of claim **6**, wherein the detection means is an antibody.

**11.** The method of claim **6**, wherein the detection means is an oligonucleotide probe.

**12.** The method of claim **6**, wherein the detection means is a filament-based diagnostic system capable of detecting a polypeptide biomarker.

**13.** The method of claim **6**, wherein the detection means is a filament-based diagnostic system capable of detecting a nucleic acid molecule biomarker.

**14.** The method according to claim **1** or **2**, further comprising obtaining brain imaging data of the subject and evaluating the data to detect an acute ischemic stroke.

**15.** The method according to claim **14**, wherein the brain imaging data is obtained by MRI.

**16.** The method according to claim **14**, wherein the brain imaging data is obtained by computerized tomography (CT) scan.

**17.** The method according to claim **1** or **2**, further comprising treating the subject with a stroke therapy if the subject is diagnosed with an acute ischemic stroke.

**18.** The method according to claim **17**, wherein the stroke therapy is the administration of a therapeutically effective amount of recombinant plasminogen activator (rtPA).

**19.** The method according to claim **6**, wherein the one or more biomarkers detected by the detection means has at least a 1.5 fold increase or decrease in expression level as compared to the levels of the one or more biomarkers in a non-stroke subject.

**20.** The method according to claim **6**, wherein the one or more biomarkers detected by the detection means has at least a 2.0 fold increase or decrease in expression level as compared to the levels of the one or more biomarkers in a non-stroke subject.

**21.** A method for differentiating an acute ischemic stroke from a transient ischemic attack (TIA), a hemorrhagic stroke and a stroke mimic in a subject presenting symptoms characteristic of a stroke, comprising:

- (a) obtaining a biological sample from the patient;
- (b) contacting the biological sample with a detection means capable of detecting the presence of at least one biomarker selected from the group consisting of: chemokine receptor 7 (CCR7); chondroitin sulfate proteoglycan 2 (CSPG2); IQ motif-containing GTPase activation protein 1 (IQGAP1); and orosomucoid 1 (ORM1),

wherein the presence of at least one of the biomarkers in the biological sample is indicative of an acute ischemic stroke but not indicative of a transient ischemic attack (TIAs), hemorrhagic stroke or stroke mimic.

**22.** The method of claim **21**, wherein the detection means is further capable of detecting the presence of at least one additional biomarker selected from the group consisting of: arginase 1 (ARG1); lymphocyte antigen 96 (LY96); matrix metalloproteinase 9 (MMP9); carbonic anhydrase 4 (CA4); and s100 calcium binding protein A12 (s100A12).

**23.** The method of claim **21**, wherein the sample is whole peripheral blood obtained from the subject.

**24.** The method of claims **21**, wherein the method is executed on the subject no more than 3 hours after onset of the presenting of the stroke symptoms.

**25.** The method of claims **21**, wherein the method is executed on the subject no more than 4.5 hours after onset of the presenting of the stroke symptoms.

**26.** The method of claim **21**, wherein the biomarker is a nucleic acid molecule.

**27.** The method of claim **26**, wherein the nucleic acid molecule is mRNA.

**28.** The method of claim **21**, wherein the biomarker is a polypeptide.

**29.** The method of claim **21**, wherein the detection means is an antibody.

**30.** The method of claim **21**, wherein the detection means is an oligonucleotide probe.

**31.** The method of claim **21**, wherein the detection means is a filament-based diagnostic system capable of detecting a polypeptide biomarker.

**32.** The method of claim **21**, wherein the detection means is a filament-based diagnostic system capable of detecting a nucleic acid molecule biomarker.

**33.** The method of claim **21**, further comprising obtaining brain imaging data of the subject and evaluating the data to detect an acute ischemic stroke.

**34.** The method according to claim **33**, wherein the brain imaging data is obtained by MRI.

**35.** The method according to claim **33**, wherein the brain imaging data is obtained by computerized tomography (CT) scan.

**36.** The method of claim **21**, further comprising treating the subject with a stroke therapy if the subject is diagnosed with an acute ischemic stroke.

**37.** The method of claim **36**, wherein the stroke therapy is the administration of a therapeutically effective amount of recombinant plasminogen activator (rtPA).

**38.** The method of claim **21**, wherein the one or more biomarkers detected by the detection means has at least a 1.5 fold increase or decrease in expression level as compared to the levels of the one or more biomarkers in a non-stroke subject.

**39.** The method of claim **21**, wherein the one or more biomarkers detected by the detection means has at least a 2.0 fold increase or decrease in expression level as compared to the levels of the one or more biomarkers in a non-stroke subject.

**40.** A kit comprising a means for detecting one or more biomarkers diagnostic of acute ischemic stroke, said biomarkers being selected from the group consisting of:

- (a) chemokine receptor 7 (CCR7);
- (b) chondroitin sulfate proteoglycan 2 (CSPG2);
- (c) IQ motif-containing GTPase activation protein 1 (IQGAP1); and
- (d) orosomucoid 1 (ORM1).

**41.** The kit of claim **40**, wherein the biomarker is a nucleic acid molecule.

**42.** The kit of claim **41**, wherein the nucleic acid molecule is mRNA.

**43.** The kit of claim **40**, wherein the biomarker is a polypeptide.

**44.** The kit of claim **40**, wherein the means for detecting the biomarkers is an oligonucleotide capable of binding to the nucleic acid molecule biomarker.

**45.** The kit of claim **43**, wherein the means for detecting the biomarkers is an antibody capable of binding to the polypeptide biomarker.

**46.** The kit of claim **40**, wherein the detection means is a filament-based diagnostic system capable of detecting a polypeptide biomarker.

**47.** The kit of claim **40**, wherein the detection means is a filament-based diagnostic system capable of detecting a nucleic acid molecule biomarker.

**48.** The kit of claim **40**, wherein the detection means includes a surface on which is attached at known locations one or more oligonucleotides capable of hybridizing to the biomarkers.

**49.** The kit of claim **40**, wherein the detection means includes a surface on which is attached at known locations one or more antibodies capable of binding to the biomarkers.

**50.** The kit according to any one of claim **48** or **49**, wherein the surface is a microarray, microtiter plate, or membrane.

**51.** The kit of claim **40** further comprising instructions for use.

**52.** A filament-based diagnostic system comprising a panel of detectable polypeptides or functional polypeptide frag-

ments thereof each corresponding to an acute ischemic stroke biomarker selected from the group consisting of:

- (a) chemokine receptor 7 (CCR7);
- (b) chondroitin sulfate proteoglycan 2 (CSPG2);
- (c) IQ motif-containing GTPase activation protein 1 (IQ-GAP1); and
- (d) orosomucoid 1 (ORM1).

**53.** A filament-based diagnostic system comprising a panel of detectable oligonucleotides each corresponding to an acute ischemic stroke biomarker selected from the group consisting of:

- (a) chemokine receptor 7 (CCR7);
- (b) chondroitin sulfate proteoglycan 2 (CSPG2);
- (c) IQ motif-containing GTPase activation protein 1 (IQ-GAP1); and
- (d) orosomucoid 1 (ORM1).

**54.** A filament-based diagnostic system comprising a panel of detectable antibodies each capable of specifically binding an acute ischemic stroke biomarker selected from the group consisting of:

- (a) chemokine receptor 7 (CCR7);
- (b) chondroitin sulfate proteoglycan 2 (CSPG2);
- (c) IQ motif-containing GTPase activation protein 1 (IQ-GAP1); and
- (d) orosomucoid 1 (ORM1).

\* \* \* \*

## Metabolite Profiling Identifies a Branched Chain Amino Acid Signature in Acute Cardioembolic Stroke

W. Taylor Kimberly, MD, PhD; Yu Wang, MD; Ly Pham, BS; Karen L. Furie, MD;  
Robert E. Gerszten, MD

**Background and Purpose**—There is limited information about changes in metabolism during acute ischemic stroke. The identification of changes in circulating plasma metabolites during cerebral infarction may provide insight into disease pathogenesis and identify novel biomarkers.

**Methods**—We performed filament occlusion of the middle cerebral artery of Wistar rats and collected plasma and cerebrospinal fluid 2 hours after the onset of ischemia. Plasma samples from control and patients with acute stroke were also analyzed. All samples were examined using liquid chromatography followed by tandem mass spectrometry. Positively charged metabolites, including amino acids, nucleotides, and neurotransmitters, were quantified using electrospray ionization followed by scheduled multiple reaction monitoring.

**Results**—The concentrations of several metabolites were altered in the setting of cerebral ischemia. We detected a reduction in the branched chain amino acids (valine, leucine, isoleucine) in rat plasma, rat cerebrospinal fluid, and human plasma compared with respective controls (16%, 23%, and 17%, respectively;  $P<0.01$  for each). In patients, lower branched chain amino acids levels also correlated with poor neurological outcome (modified Rankin Scale, 0–2 versus 3–6;  $P=0.002$ ).

**Conclusions**—Branched chain amino acids are reduced in ischemic stroke, and the degree of reduction correlates with worse neurological outcome. Whether branched chain amino acids are in a causal pathway or are an epiphomenon of ischemic stroke remains to be determined. (*Stroke*. 2013;44:1389–1395.)

**Key Words:** cerebrospinal fluid ■ liquid chromatography ■ mass spectrometry ■ metabolomics  
■ transient ischemic attack stroke

The underlying pathogenesis of acute ischemic stroke remains poorly understood, with a paucity of biological insight translating into useful therapy in patients. Metabolomics is an emerging analytic technology for understanding disease pathogenesis that can be applied to both animal models and patient blood samples. It therefore represents an attractive translational tool to link the biology of model systems to the pathophysiology in patients. Using either nuclear magnetic resonance spectroscopy or mass spectrometry (MS),<sup>1</sup> metabolomics can measure numerous small metabolites simultaneously.<sup>2</sup> MS-based profiling methods include gas chromatography–MS and liquid chromatography coupled to MS, the most common of which is tandem MS (LC-MS/MS).<sup>3</sup> Approaches that use LC-MS/MS are increasingly used because of their sensitivity, flexibility, and quantitative capability for small molecule detection.<sup>2</sup>

Metabolomic profiling has found application in other forms of metabolic stress,<sup>4</sup> including intense exercise,<sup>5</sup> myocardial ischemia,<sup>6</sup> myocardial infarction,<sup>7</sup> and diabetes mellitus,<sup>8–10</sup> but little is known about metabolite changes in the setting of stroke.

A common strategy used, in prior metabolomics studies, was to compare the metabolome within subjects, before and after the exposure. However, baseline blood sampling is not feasible in patients with acute stroke. We therefore sought to establish a metabolomic profile in an animal model of ischemic stroke in which baseline sampling is possible, and then integrate the findings with profiling in individuals with acute ischemic stroke. Using a rodent filament occlusion model, we first identified potential candidates whose plasma and cerebrospinal fluid (CSF) were altered. We then evaluated those candidates in an analogous patient cohort in which plasma samples were collected in the acute setting. We hypothesized that we could detect a specific pattern of circulating metabolites that would reflect the chain of metabolic events that occur during cerebral ischemia. Our goal was to apply this new systematic tool as a first step to better understand the biology and pathogenesis of acute ischemic stroke. In doing so, we also explored whether these candidates might serve as potential biomarkers for diagnosis or prognosis.<sup>11–13</sup>

Received December 5, 2012; final revision received February 22, 2013; accepted February 25, 2013.

From the Center for Human Genetic Research and Division of Neurocritical Care and Emergency Neurology (W.T.K., L.P.), Martinos Center for Biomedical Imaging (W.T.K., Y.W.), and Cardiovascular Research Center (R.E.G.), Massachusetts General Hospital, Harvard Medical School, Boston, MA; and Department of Neurology, Warren Alpert Medical School of Brown University, Rhode Island, RI (K.L.F.).

Costantino Iadecola, MD, was guest editor for this article.

The online-only Data Supplement is available with this article at <http://stroke.ahajournals.org/lookup/suppl/doi:10.1161/STROKEAHA.111.000397/-DC1>.

Correspondence to W. Taylor Kimberly, MD, PhD, Lunder 644, 55 Fruit St, Boston, MA 02114. E-mail wtkimberly@partners.org; or Robert E. Gerszten, MD, Simches 3, 185 Cambridge St, Boston, MA 02114. E-mail gerszten.robert@mgh.harvard.edu

© 2013 American Heart Association, Inc.

Stroke is available at <http://stroke.ahajournals.org>

DOI: 10.1161/STROKEAHA.111.000397

## Methods

### Animals

Adult male Wistar rats weighing 275 to 350 g were obtained from Charles River Laboratories (Wilmington, MA). Animals were housed with free access to food and water. The evening before surgery, animals were given nothing per oral to avoid the effect of dietary intake on circulating metabolites. Transient filament occlusion was performed using a 4-0 siliconized suture (Dccol Corp, Sharon, MA) according to standard methods (Methods in the online-only Data Supplement).<sup>14,15</sup> Approximately 250 µL of plasma was withdrawn at baseline and at 2 hours after ischemia onset.

CSF ( $\approx$ 50 µL) was collected from the cisterna magna at 2 hours after ischemia, using a 27-gauge winged needle set attached to a 1 cm<sup>3</sup> syringe.<sup>16</sup> Animals were allowed to recover, and at 24 hours after ischemia, brains were harvested for 2,3,5-triphenyltetrazolium chloride staining to assess the size of stroke.<sup>17</sup> All experiments were approved under an institutionally approved protocol in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

### Patients

We analyzed EDTA-containing plasma samples collected at a single center, as part of a prospective 2 center biomarker study of acute ischemic stroke (Specialized Programs of Translational Research in Acute Stroke [SPOTRIAS] Network). The SPOTRIAS biomarker study enrolled consecutive patients aged  $\geq$ 18 years between January 2007 and April 2010, who presented to the Massachusetts General Hospital Emergency Department within 9 hours of symptom onset, with symptoms consistent with ischemic stroke (see Methods in the online-only Data Supplement for additional details of the cohort and the patient data and imaging collection). Ischemic stroke was defined as acute-onset focal neurological deficit with neuroimaging evidence of infarction, or symptom duration  $>$ 24 hours in the setting of negative diffusion-weighted MRI. Transient ischemic attack was defined as resolution of neurological symptoms within 24 hours that were consistent with a vascular ischemic event (n=18). The designation of “not a stroke” was reserved for subjects with a negative diffusion-weighted MRI who also had an alternative diagnosis for neurological symptoms at discharge (n=14). All subjects or their healthcare proxy provided informed consent, and this study was approved by the local institutional review board.

We applied a case-control design to mirror the animal modeling experiments. We defined 3 groups from the SPOTRIAS biomarker cohort: control, mild, and severe stroke groups. Controls included all subjects with a final diagnosis of transient ischemic attack or absence of stroke (n=32). A similar sized group of mild ischemic stroke was selected from cardioembolic stroke subjects, and 22 sequential subjects with a National Institutes of Health Stroke Scale (NIHSS)  $\geq$ 4 were used. We also selected sequential subjects with severe cardioembolic stroke, defined as those with an NIHSS  $\geq$ 15 (n=30). All subjects or their healthcare proxy provided informed consent, and this study was approved by the local institutional review board.

### High Performance Liquid Chromatography and Tandem MS

EDTA blood samples were collected and immediately centrifuged to separate cellular material. Aliquots of plasma supernatant were frozen on dry ice and stored at  $-80^{\circ}\text{C}$  until analysis. Plasma samples (10 µL) were deproteinized with 90 µL acetonitrile/methanol (3:1; v/v) containing internal standards (valine-d8 [Sigma-Aldrich] and phenylalanine-d8 [Cambridge Isotope Laboratories]). After centrifugation, the extracts were subjected to normal phase hydrophilic interaction chromatography. The chromatography system consisted high throughput screen prep and load autosampler (Leap Technologies, Carrboro, NC) connected to a high performance liquid chromatography pump (1200 Series, Agilent, Santa Clara, CA). MS data were acquired using a 4000 QTRAP triple quadrupole mass spectrometer (Applied Biosystems/Sciex, Framingham, MA) equipped with an electrospray ionization source. Positively charged amino acids,

nucleotides, and neurotransmitters were selected for targeted MS/MS analysis using selected multiple reaction monitoring conditions determined previously using reference standards.<sup>6,7</sup>

A total of 68 endogenous metabolites were monitored and detected for each sample. The metabolites were selected on the basis of a broad representation of diverse metabolic pathways as possible, balanced against compatibility with the chromatography and MS ionization method. Deuterated internal standards (valine-d8 and phenylalanine-d8, Cambridge Isotope Laboratories (Andover, MA)) were included in each sample to monitor for quality control. Any sample with internal standard values  $\geq$ 2 SD were excluded from peak integration and further analysis. In addition, pooled plasma samples were interspersed within each analytic run at standardized intervals, enabling the monitoring and correction for temporal drift in MS performance. Each of these samples were prepared, extracted, and processed as separate 10 µL aliquots from a larger pool of normal human plasma. Replicate injections of pooled plasma demonstrated that 50% of the analytes had a coefficient of variation  $\leq$ 5% (including the branched chain amino acids [BCAA]), 69% of the analytes had a coefficient of variation  $\leq$ 10%, and 91% had coefficient of variation  $\leq$ 20%, which is consistent with prior studies.<sup>9</sup>

### Statistical Analysis

#### Univariate Analysis

Differences in clinical and laboratory continuous variables were compared using Student *t* test or Mann–Whitney test, as appropriate. Categorical variables were compared using Fisher exact test. For the metabolite analysis in the animal samples, we used an uncorrected *P* value threshold of 0.05, using Mann–Whitney or Student *t* test, depending on data normality. In this exploratory phase, no correction for multiple comparisons was made.

In the human cohort analysis, we used a similar approach to our prior studies<sup>8</sup> and applied the Benjamini–Hochberg procedure<sup>18</sup> to limit the false discovery rate to  $q<0.1$ , which corresponded to a threshold of  $P<0.015$ . This would be expected to yield  $\approx$ 1 false-positive discovery in 68 metabolites analyzed, assuming independent hypotheses. Moreover, this threshold also approximates the Bonferroni correction of the combined probability between the discovery cohort ( $P<0.05$ ) and the human validation cohort ( $P<0.015$ ) (ie,  $0.05 \times 0.015 = 7.5 \times 10^{-4}$ , whereas Bonferroni correction =  $0.05/68 = 7.4 \times 10^{-4}$ ). Although many metabolites were associated with predefined groups (eg, amino acids, tryptophan derivatives, nucleotide metabolites, etc.), this is a conservative estimate because the number of independent tests was substantially lower than the nominal ones. Statistical analyses were performed using the STATA statistical software (release 12) or JMP 10 Pro (SAS Institute, Cary, NC).

#### Multivariate Analysis

To uncover the multivariate structure within the human data set, we performed principal component analysis and partial least-squares discriminant analysis using MetaboAnalyst 2.0 (Edmonton, BC, Canada).<sup>19</sup> Because each method provides slightly different insight into high-dimensional data, we performed to highlight the metabolites in common (see Methods in the online-only Data Supplement for further details).

### Results

Using LC-MS/MS, we first examined serial blood samples after filament occlusion in a rat model of ischemic stroke. We collected plasma at baseline and 2 hours after ischemia, as well as CSF at the 2-hour time point. In pilot experiments, the placement of a laser Doppler flowmetry probe led to poor recovery of CSF (data not shown). Exploiting the variability in stroke volume that would occur in the absence of Doppler flowmetry, we designed our experiment as a comparison among sham, small stroke, and large stroke animal cohorts. Of 23 animals, 2 died acutely and the volume of infarct could

not be determined. Another animal assigned to the middle cerebral artery occlusion group had no infarction at 24 hours and was therefore excluded. The remaining 20 animals were included in the analysis: 7 sham-operated animals, 6 animals with small infarction (stroke volume 9±5%), and 7 animals with large infarction (stroke volume 29±5%; Table I in the online-only Data Supplement).

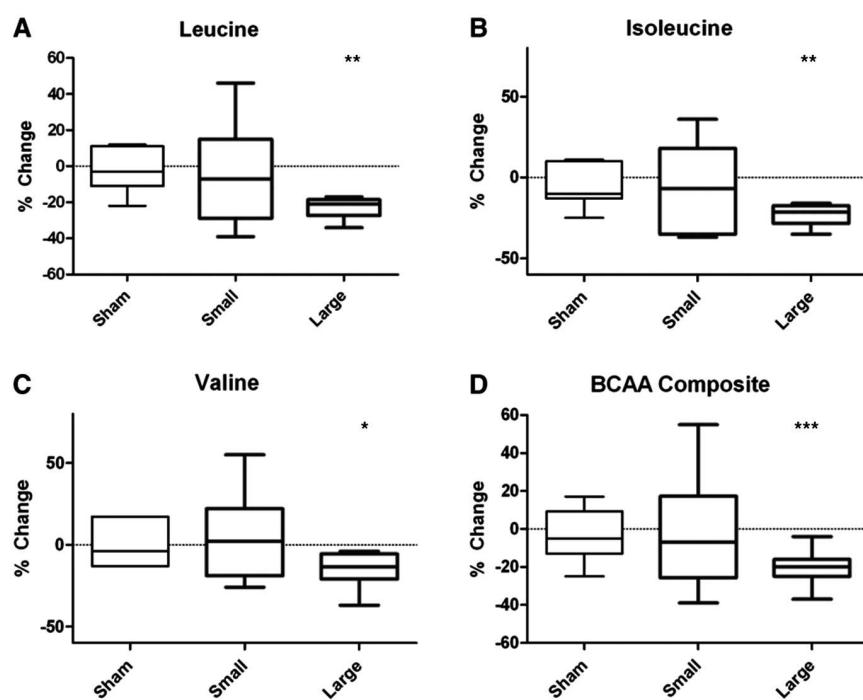
We measured a total of 68 metabolites in baseline and 2-hour follow-up plasma samples, and results were analyzed as a percent change from baseline, which adjusts for within animal variation. To eliminate any nonspecific effects of the operative technique, we compared percent metabolite changes in stroke animals to those in sham-operated animals. From baseline to 2 hours after stroke, there was a significant decrease in the concentration of BCAA leucine, isoleucine, and valine in the large stroke group ( $P=0.003$ , 0.01, and 0.04, respectively). BCAA are coordinately regulated, and the levels change in conjunction with each other.<sup>20</sup> Accordingly, a composite measurement of the BCAA showed a 16±6% decrease in large stroke ( $P=1\times 10^{-5}$ ; Figure 1D) and a nonsignificant trend in small stroke. Several other metabolites were altered in a dose-dependent manner in small and large stroke. These included stepwise increases in xanthosine (+57%;  $P<0.001$ ), carnosine (+71%;  $P<0.005$ ), and glutamate (+40%;  $P=0.01$ ), and decreases in niacinamide (-31%;  $P=0.02$ ) and phenylalanine (-18%;  $P<0.01$ ) relative to sham-operated animals.

We also measured the same metabolites in the CSF obtained through cisterna magna puncture at 2 hours after onset of ischemia. Because the concentration of most metabolites in CSF is lower than in plasma, we excluded any CSF samples with visible blood contamination (Methods and Figure I in the online-only Data Supplement). Figure 2A through 2D shows that the individual BCAAs had a consistent trend toward a decrease (leucine -21%,  $P=0.06$ ; isoleucine -23%,  $P=0.14$ ; valine -22%,  $P=0.11$ ). Moreover, a composite of BCAA

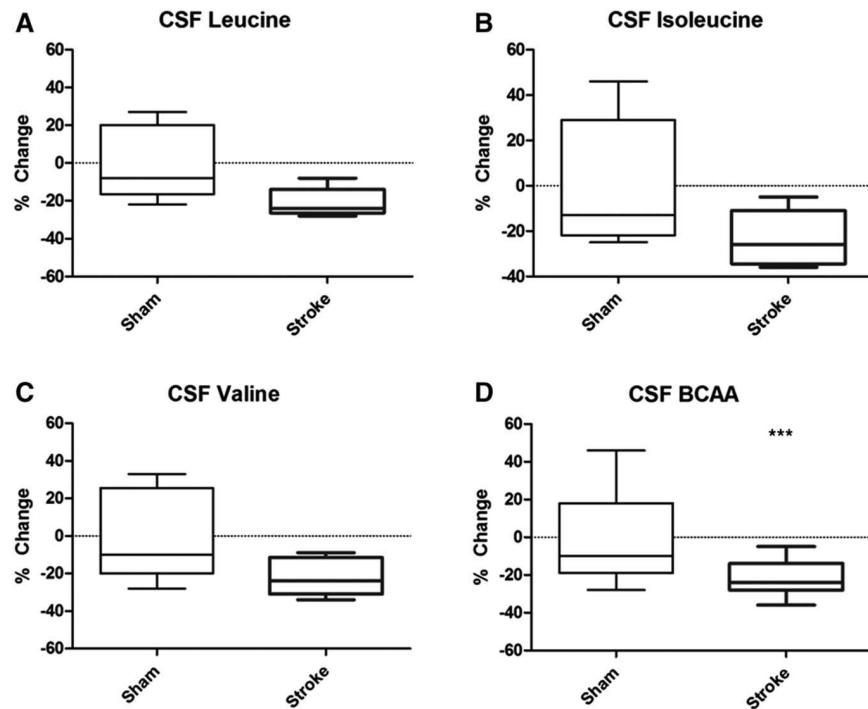
demonstrated a decrease of 23±9% compared with sham CSF (n=5 for each group;  $P<0.005$ ). Other significantly altered CSF metabolites included an accumulation of xanthosine (102%,  $P=0.01$ ) and lysine (18%,  $P=0.02$ ).

On the basis of the animal studies, concordant metabolite changes between plasma and CSF included valine, leucine, isoleucine, and xanthosine. We next evaluated whether these candidate metabolites were altered in the plasma of patients with acute stroke to determine whether these metabolite changes represented a common alteration. We obtained plasma samples from a cohort of patients in whom blood was collected acutely, shortly after presentation to the emergency department. We selected a subset of subjects to coincide with the animal modeling design, which included a control group (patients with a diagnosis of transient ischemic attack or absence of stroke), a group with mild stroke (patients with an NIHSS 4–5), and a severe stroke group (NIHSS 15–19). To limit potential heterogeneity, we focused on subjects with a cardioembolic cause of stroke. The clinical characteristics of the cohort are listed in Table 1. As would be expected, the stroke group had an older age and higher rates of atrial fibrillation compared with the control group. In addition, the large stroke group had a higher acute stroke volume, higher acute NIHSS, and worse 3-month neurological outcome as compared with the small or control groups.

We analyzed plasma samples obtained within 6±2 hours from the last seen well time, using our metabolomics method. Heat map correlation analysis confirmed a close association of the BCAAs (Figure 3, top right), consistent with the animal modeling data and with the known coordinated metabolism of these amino acids.<sup>20</sup> Analysis of individual metabolites showed that leucine, isoleucine, and valine were all decreased in stroke compared with control, and to a greater extent in large compared with small stroke ( $P<0.01$  for each; Figure 4). Similarly the composite BCAA score demonstrated a 9±17%



**Figure 1. A–D.** Rats subjected to filament occlusion of the middle cerebral artery had plasma collected at baseline (just before filament occlusion) and 2 hours after stroke. The concentration of branched chain amino acids (BCAA) were diminished from baseline to 2 hours after stroke. \* $P<0.05$ , \*\* $P<0.01$ , and \*\*\* $P<0.001$ .



**Figure 2. A–D,** Change in branched chain amino acids (BCAA) in cerebrospinal fluid (CSF) from baseline to 2 hours after stroke ( $n=5$  for each group). Each individual BCAA showed a trend toward a decrease,  $P=0.06$ ,  $0.14$ , and  $0.11$ , respectively, whereas the composite BCAA, xanthosine and lysine were significant,  $*P<0.01$ ,  $**P<0.01$ .

decrease in small stroke ( $P=0.03$ ) and a  $17\pm23\%$  decrease in large stroke ( $P=1.1\times10^{-5}$ ). Table 2 provides a complete list of all metabolites that were altered in the setting of ischemia when compared with control patients. In addition to novel metabolites, we found that glucose showed a significant increase in stroke compared with control, which is concordant with the well-described phenomenon of acute stress hyperglycemia.<sup>21–23</sup>

To further simplify the high-dimensional metabolomics data, we next performed principal component analysis. This approach consolidates data into fewer metabolite clusters, which maximally explain the variance in the data.<sup>19</sup> Intriguingly, the first principal component (principal component, which

explained 20% variance in the data; Figure II in the online-only Data Supplement for score and loading plots) contained the BCAA metabolites. In addition to leucine, valine, and isoleucine, the first principal component also included tyrosine, lysine, and methionine. Comparing the individual subjects' scores, the first principal component also distinguished cases from controls ( $P=0.020$  comparing control versus all stroke and  $P=0.011$  for control versus large stroke).

Next, we performed partial least-squares discriminant analysis, which is a method of supervised classification that is designed to highlight metabolite differences between cases and controls. This technique is commonly used in metabolomics studies for biomarker discovery because it emphasizes the distinction between the 2 classes.<sup>19</sup> The metabolites that contributed the greatest discrimination between stroke and controls were similar to our univariate analysis presented in Table 2. These included the BCAs, carnitine, threonine, histidine, and glucose (Figure III in the online-only Data Supplement). Validation of the model was confirmed using cross-validation and permutation testing ( $P<0.01$ ; Figure III in the online-only Data Supplement).<sup>19,24</sup>

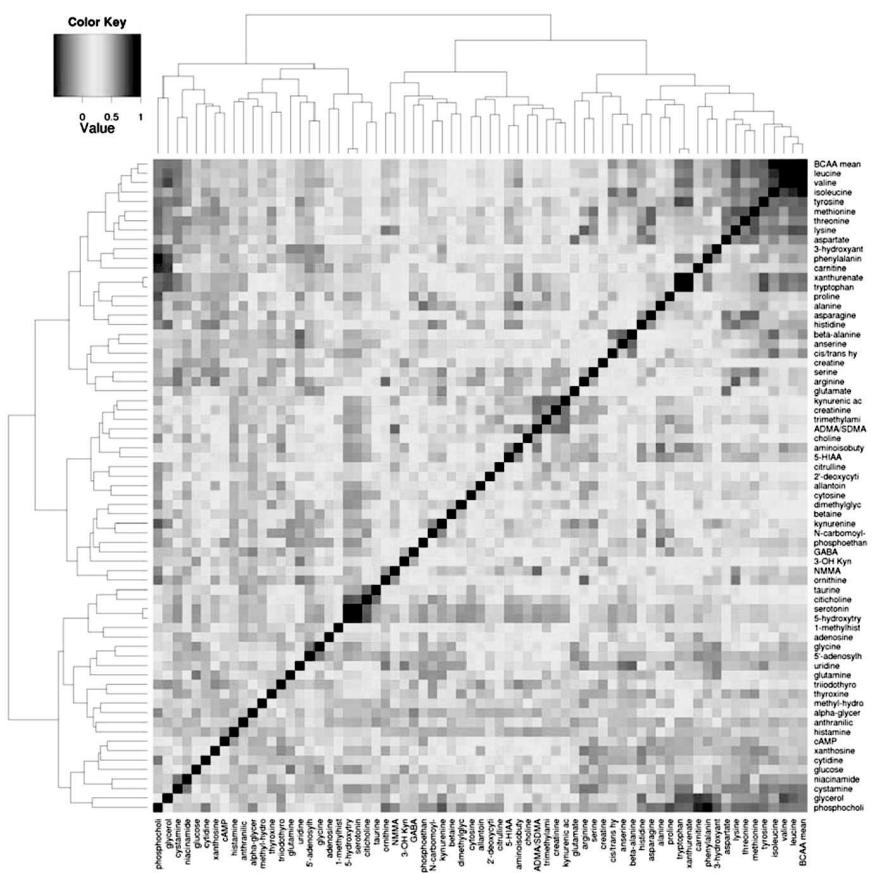
Having confirmed that BCAA were altered acutely in stroke, we next explored its association with imaging and clinical measures. Because the magnitude of BCAA change seemed to correlate with size of stroke in the animal model, we evaluated the correlation between BCAA and diffusion-weighted volume in the patient cohort. There was a nonsignificant trend in association between admission infarct volume and BCAA ( $r=-0.18$ ;  $P=0.11$ ). On the contrary, a lower concentration of BCAA was associated with increased age ( $r=-0.26$ ;  $P=0.02$ ), female sex ( $P<0.001$ ), and worse outcome at 3 months (modified Rankin Scale, 3–6;  $P=0.002$ ). Because age and sex are also recognized predictors of worse neurological outcome,<sup>25,26</sup> we explored whether BCAA predicted outcome independently of age and sex. Although the

**Table 1. Clinical Characteristics of the Stroke Cohort**

	TIA (N=32)	Stroke (N=52)	P Value
Female, N, %	16 (50%)	22 (42%)	0.51
Age $\pm$ SD, y	66 $\pm$ 16	75 $\pm$ 10	<0.01
Admission temp, °F $\pm$ SD	98.3 $\pm$ 0.7	98.2 $\pm$ 1.1	0.72
CAD, N, %	9 (28%)	17 (33%)	0.81
HTN, N, %	25 (75%)	45 (87%)	0.37
DM2, N, %	11 (34%)	12 (23%)	0.32
HL, N, %	15 (47%)	23 (44%)	0.83
Afib, N, %	5 (16%)	33 (63%)	<0.001

	Small	Large	
NIHSS, median [IQR]	3 [1, 8]	4 [4, 5]	17 [15, 19] <0.001
DWI volume, median [IQR]	0 [0, 3]	3 [1, 15]	25 [11, 59] <0.001
3 Months mRS, 0–2, %	23 (79%)	11 (65%)	5 (23%) <0.001

Afib indicates atrial fibrillation; CAD, coronary artery disease; DM2, diabetes mellitus 2; DWI, diffusion-weighted imaging; HL, hyperlipidemia; HTN, hypertension; IQR, interquartile range; mRS, modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale; TIA, transient ischemic attack.

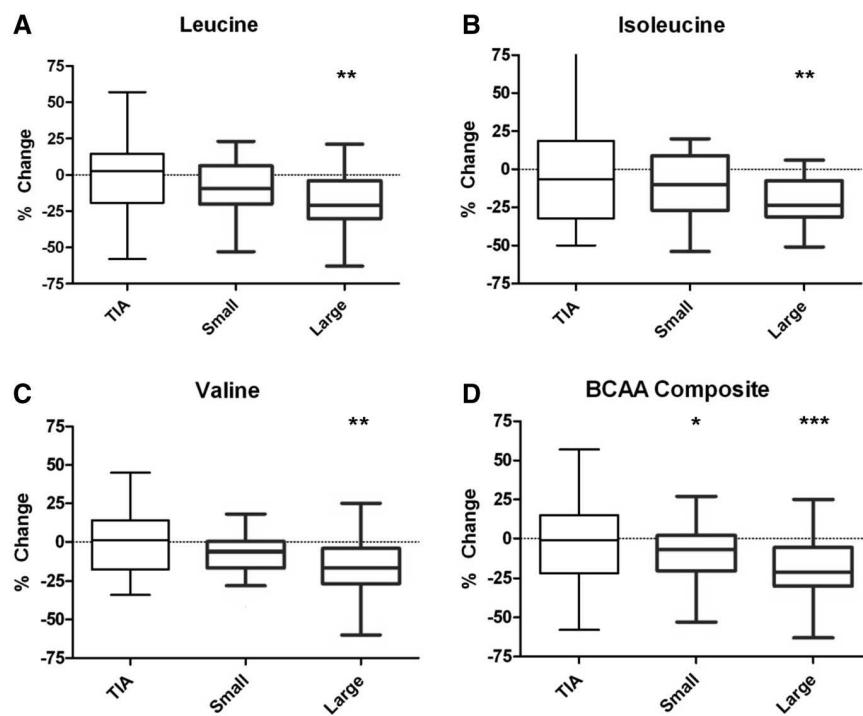


**Figure 3.** Heat map representation of metabolites highlights the tight correlation of the branched chain amino acid (BCAA), which are located in the top right. The heat map is generated from 52 patients with acute stroke who had blood samples drawn at  $6 \pm 2$  hours from the last seen well time. Analytes that are positively correlated are represented in red, whereas compounds inversely correlated are represented in blue. ADMA indicates asymmetric dimethylarginine; GABA, gamma aminobutyric acid; HIAA, 5-hydroxyindoleacetic acid; and SDMA, symmetric dimethylarginine.

cohort was limited in size and stratified on the basis of stroke severity, we performed exploratory multivariable logistic regression and found that BCAA remained an independent predictor of outcome ( $P=0.04$ ) after adjusting for age and sex.

## Discussion

Using metabolomics, we have identified specific circulating metabolites that are altered in the setting of cerebral infarction. On the basis of our systematic analysis in a well-controlled



**Figure 4. A–D.** The concentration of plasma branched chain amino acids (BCAA) in patients with stroke is reduced when compared with control subjects at the time of acute presentation. \* $P<0.05$ , \*\* $P<0.01$ , and \*\*\* $P<0.001$ .

**Table 2. All Metabolites Significantly Changed in Human Stroke Subjects Compared With Control**

Metabolite	Fold Change In Stroke	PValue	BH Procedure, Q<0.1
Carnitine	0.89	0.001	0.0015
Threonine	0.80	0.002	0.0029
Histidine	0.83	0.003	0.0044
Glucose	1.42	0.0057	0.0059
Valine	0.88	0.007	0.0074
BCAA mean	0.86	0.008	0.0088
Methionine	0.82	0.009	0.0103
Leucine	0.86	0.009	0.0117
Glycine	0.82	0.0131	0.0132
Proline	0.89	0.017	0.0147
Lysine	0.86	0.025	0.0162
Cysteamine	0.56	0.027	0.0177
Isoleucine	0.85	0.028	0.0191
Uridine	0.81	0.033	0.0258
5'-Adenosylhomocysteine	0.84	0.036	0.0221
Creatinine	0.89	0.039	0.0235
N-carbamoyl-β-alanine	1.33	0.041	0.0250
cis/trans Hydroxyproline	0.73	0.041	0.0265
Asparagine	0.89	0.043	0.0279

The false discovery threshold based on the Benjamini-Hochberg (BH) procedure<sup>18</sup> is indicated by the bold line. For completeness, additional metabolites that exceed this threshold, but with an uncorrected  $P<0.05$ , are listed below the bold line.

BCAA indicates branched chain amino acids.

animal model and linking those findings to patient samples in the acute setting, we have identified a small and interrelated subset of metabolites. Our data demonstrate a reduction in the concentration of BCAAs that associates with stroke severity and worse neurological outcome. Although our data do not point to an underlying biological mechanism, they focus future experiments on investigating candidate pathways that relate to BCAA. The notion that BCAA play an important role in the metabolic response to disease is supported by evidence of its alteration in other illnesses. For example, BCAA is reduced in critical illnesses, such as sepsis, trauma, and burn injury.<sup>27-29</sup> BCAA is also associated with the risk of incident diabetes mellitus<sup>9</sup> and can induce insulin resistance,<sup>8</sup> further suggesting a role in metabolic homeostasis. Perhaps, most interestingly, BCAA are altered also in heart disease,<sup>30</sup> suggesting that these amino acids play a critical role in bioenergetic homeostasis. Whether BCAA represent a novel link between cardiovascular and cardioembolic cerebrovascular diseases requires further investigation.

In addition to their potential role in systemic disease states, BCAA also serve a unique role in the brain.<sup>31,32</sup> For example, BCAA are integral to the glutamate/glutamine cycle between astrocytes and neurons, which is critical for the efficient uptake of glutamate during excitatory neuronal signaling.<sup>31</sup> Intriguingly, inhibition of the first step of BCAA catabolism with gabapentin reduces brain glutamate concentration.<sup>31</sup> Gabapentin has been reported to reduce stroke volume in a

rodent model,<sup>33</sup> and 1 possibility is that it may do so by limiting glutamate concentration and subsequent excitotoxicity. Although our rodent data showed an accumulation of glutamate, we did not detect a similar change in the patients. Whether this reflects inadequate power or greater complexity in the human cohort requires further study. Alternatively, the reduction in BCAA level may reflect a metabolic pathway leading to consumption or sequestration in a tissue compartment other than blood or CSF. BCAA are also known to have roles in protein metabolism and in catabolic energy metabolism.<sup>20</sup> These putative mechanisms are not mutually exclusive, and, indeed, systemic BCAA levels have been shown to influence brain neurotransmitter levels.<sup>32</sup> Nevertheless, our data raise the possibility that manipulation of BCAA may influence outcome. Future studies that focus on whether BCAA are causally related to cerebral ischemia, such as through supplementation and pharmacological inhibition, will help determine whether BCAA holds promise as a therapeutic target.

Our analysis in rodents and patients identified additional candidate metabolites, which were not shared in common between the 2 (Table 2 and Results). The similarities and differences between rodent model systems and patients are an area of substantial importance for translational therapy. Metabolomics is a technique that allows direct comparisons between the model systems and patients in a way that was not previously available. Although our findings with BCAA highlight that there are similar biological pathways in rodents and patients, the differences may offer some caution. Nevertheless, our data point to one approach to explore these similarities and differences systemically, both of which are important for novel target discovery. There is little prior metabolomics analysis of stroke, with the exception of an nuclear magnetic resonance-based study in a cohort of lacunar stroke subjects,<sup>34</sup> which analyzed blood samples collected within 72 hours of stroke onset. Of the overlapping metabolites in common with our method, valine was diminished in lacunar stroke, although leucine and isoleucine were not.<sup>34</sup> The apparent differences may reflect the increased sensitivity of LC-MS/MS compared with nuclear magnetic resonance, differing metabolomes based on stroke subtype, differences in control selection, and potentially in the timing of the blood draw.

Our study has several strengths. We used a carefully controlled model system to establish a metabolite profile and then compared it to a well-phenotyped patient cohort. We used a metabolomics technique that is well validated and possesses excellent quantitative capability and reproducibility. The patient samples were obtained in the hyperacute phase and compared with a control group of stroke mimics. However, there are several limitations to our analysis. We used a targeted metabolomics approach, which identifies a limited set of metabolites rather than a comprehensive list of known and unknown peaks. It is therefore possible that additional metabolite changes occur that we cannot detect with our current method. LC-MS/MS-based metabolomics also has limited throughput capability. Nevertheless, we have selected key sentinel metabolites that are central to several important biochemical pathways, including amino acids, nucleotides, and selected neurotransmitters. Although our data point toward a key role for BCAA in stroke, our correlation and

multivariate regression must be interpreted with caution in a small patient cohort. Most importantly, validation in a larger cohort that includes all stroke subtypes with a broad range of stroke severity will be necessary to confirm our findings and determine whether BCAA holds promise as a clinically useful biomarker or a therapeutic target.

### Sources of Funding

This study was supported by the Clinical Investigator Training Program: Beth Israel Deaconess Medical Center, Harvard Medical School, in collaboration with Pfizer Inc and Merck and Co (Dr Kimberly), by National Institute of Health (NIH) K23NS076597 (Dr Kimberly), NIH 5P50NS051343-07 (Dr Furie), and NIH R01HL096738 and R01HL098280 (Dr Gerszten).

### Disclosures

Dr Kimberly discloses a research grant (National Institute of Health [NIH]; significant) and Dr Gerszten discloses a research grant (NIH; significant).

### References

- Bain JR, Stevens RD, Wenner BR, Ilkayeva O, Muoio DM, Newgard CB. Metabolomics applied to diabetes research: moving from information to knowledge. *Diabetes*. 2009;58:2429–2443.
- Lewis GD, Asnani A, Gerszten RE. Application of metabolomics to cardiovascular biomarker and pathway discovery. *J Am Coll Cardiol*. 2008;52:117–123.
- Nicholson JK, Lindon JC. Systems biology: metabolomics. *Nature*. 2008;455:1054–1056.
- Holmes E, Wilson ID, Nicholson JK. Metabolic phenotyping in health and disease. *Cell*. 2008;134:714–717.
- Lewis GD, Farrell L, Wood MJ, Martinovic M, Arany Z, Rowe GC, et al. Metabolic signatures of exercise in human plasma. *Sci Transl Med*. 2010;2:33ra37.
- Sabatine MS, Liu E, Morrow DA, Heller E, McCarroll R, Wiegand R, et al. Metabolomic identification of novel biomarkers of myocardial ischemia. *Circulation*. 2005;112:3868–3875.
- Lewis GD, Wei R, Liu E, Yang E, Shi X, Martinovic M, et al. Metabolite profiling of blood from individuals undergoing planned myocardial infarction reveals early markers of myocardial injury. *J Clin Invest*. 2008;118:3503–3512.
- Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab*. 2009;9:311–326.
- Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E, et al. Metabolite profiles and the risk of developing diabetes. *Nat Med*. 2011;17:448–453.
- Cheng S, Rhee EP, Larson MG, Lewis GD, McCabe EL, Shen D, et al. Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation*. 2012;125:2222–2231.
- Kimberly WT. Biomarkers in neurocritical care. *Neurotherapeutics*. 2012;9:17–23.
- Foerch C, Montaner J, Furie KL, Ning MM, Lo EH. Invited article: searching for oracles? Blood biomarkers in acute stroke. *Neurology*. 2009;73:393–399.
- Kernagis DN, Laskowitz DT. Evolving role of biomarkers in acute cerebrovascular disease. *Ann Neurol*. 2012;71:289–303.
- Tatlisumak T, Fisher M. *Handbook of Experimental Neurology: Methods and Techniques in Animal Research*. Cambridge, NY: Cambridge University Press; 2006.
- Sicard KM, Fisher M. Animal models of focal brain ischemia. *Exp Transl Stroke Med*. 2009;1:7.
- Sharma AK, Schultze AE, Cooper DM, Reams RY, Jordan WH, Snyder PW. Development of a percutaneous cerebrospinal fluid collection technique in F-344 rats and evaluation of cell counts and total protein concentrations. *Toxicol Pathol*. 2006;34:393–395.
- Benedek A, Móricz K, Jurányi Z, Gigler G, Lévay G, Hásing LG Jr, et al. Use of TTC staining for the evaluation of tissue injury in the early phases of reperfusion after focal cerebral ischemia in rats. *Brain Res*. 2006;1116:159–165.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B*. 1995;57:289–300.
- Xia J, Psychogios N, Young N, Wishart DS. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res*. 2009;37:W652–W660.
- Brosnan JT, Brosnan ME. Branched-chain amino acids: enzyme and substrate regulation. *J Nutr*. 2006;136:207S–211S.
- Bruno A, Biller J, Adams HP Jr, Clarke WR, Woolson RF, Williams LS, et al. Acute blood glucose level and outcome from ischemic stroke. Trial of ORG 10172 in Acute Stroke Treatment (TOAST) Investigators. *Neurology*. 1999;52:280–284.
- Capes SE, Hunt D, Malmberg K, Pathak P, Gerstein HC. Stress hyperglycemia and prognosis of stroke in nondiabetic and diabetic patients: a systematic overview. *Stroke*. 2001;32:2426–2432.
- Scott JF, Robinson GM, French JM, O'Connell JE, Alberti KG, Gray CS. Prevalence of admission hyperglycaemia across clinical subtypes of acute stroke. *Lancet*. 1999;353:376–377.
- Bijlsma S, Bobeldijk I, Verheij ER, Ramaker R, Kochhar S, Macdonald IA, et al. Large-scale human metabolomics studies: a strategy for data (pre-) processing and validation. *Anal Chem*. 2006;78:567–574.
- Reeves MJ, Bushnell CD, Howard G, Gargano JW, Duncan PW, Lynch G, et al. Sex differences in stroke: epidemiology, clinical presentation, medical care, and outcomes. *Lancet Neurol*. 2008;7:915–926.
- Reeves MJ, Lisabeth LD. The confounding issue of sex and stroke. *Neurology*. 2010;74:947–948.
- Nawabi MD, Block KP, Chakrabarti MC, Buse MG. Administration of endotoxin, tumor necrosis factor, or interleukin 1 to rats activates skeletal muscle branched-chain alpha-keto acid dehydrogenase. *J Clin Invest*. 1990;85:256–263.
- Holecek M. Leucine metabolism in fasted and tumor necrosis factor-treated rats. *Clin Nutr*. 1996;15:91–93.
- Holecek M. Relation between glutamine, branched-chain amino acids, and protein metabolism. *Nutrition*. 2002;18:130–133.
- Huang Y, Zhou M, Sun H, Wang Y. Branched-chain amino acid metabolism in heart disease: an epiphomenon or a real culprit? *Cardiovasc Res*. 2011;90:220–223.
- Hutson SM, Lieth E, LaNoue KF. Function of leucine in excitatory neurotransmitter metabolism in the central nervous system. *J Nutr*. 2001;131:846S–850S.
- Fernstrom JD. Branched-chain amino acids and brain function. *J Nutr*. 2005;135:1539S–1546S.
- Hoffmann U, Lee JH, Qin T, Eikermann-Haerter K, Ayata C. Gabapentin reduces infarct volume but does not suppress peri-infarct depolarizations. *J Cereb Blood Flow Metab*. 2011;31:1578–1582.
- Jung JY, Lee HS, Kang DG, Kim NS, Cha MH, Bang OS, et al. 1H-NMR-based metabolomics study of cerebral infarction. *Stroke*. 2011;42:1282–1288.

**SHORT COMMUNICATION**

# N-terminal pro-brain natriuretic peptide level determined at different times identifies transient ischaemic attack patients with atrial fibrillation

F. Purroy<sup>a</sup>, I. Suárez-Luis<sup>a</sup>, G. Mauri-Capdevila<sup>a</sup>, S. Cambray<sup>a</sup>, J. Farré<sup>a</sup>, J. Sanahuja<sup>a</sup>, G. Piñol-Ripoll<sup>a</sup>, A. Quílez<sup>a</sup>, C. González-Mingot<sup>a</sup>, R. Begué<sup>b</sup>, M. I. Gil<sup>b</sup>, E. Fernández<sup>a</sup> and I. Benabdellahak<sup>a</sup>

<sup>a</sup>Stroke Unit, Hospital Universitari Arnau de Vilanova, Grup Neurociències Clíniques IRBLleida, Lleida; and <sup>b</sup>Institut de diagnòstic per la Imatge, Hospital Universitari Arnau de Vilanova, Grup Neurociències Clíniques IRBLleida, Lleida, Spain

**Keywords:**

atrial fibrillation,  
biomarker,  
cardioembolic,  
NT-proBNP,  
transient  
ischemic attack

Received 13 February 2013

Accepted 24 May 2013

**Background and purpose:** The etiological classification of patients with transient ischaemic attack (TIA) is a difficult endeavor and the use of serum biomarkers could improve the diagnostic accuracy. The aim of this study was to correlate atrial fibrillation, the main cardioembolic etiology (CE), with different serum biomarkers measured in consecutive TIA patients.

**Methods:** The concentrations of interleukin-6 (IL-6), tumor necrosis factor-alpha, neuron-specific enolase, high-sensitivity C-reactive protein, IL-1- $\alpha$  and the N-terminal pro-B type natriuretic peptide (NT-proBNP) were quantified in the serum of 140 patients with TIA and 44 non-stroke subjects. Measurements were performed at different times throughout evolution: within 24 h of symptoms onset and at days 7 and 90.

**Results:** With the exception of IL-6, all biomarkers were higher in TIA patients than in controls. NT-proBNP was significantly related to the presence or new diagnosis of AF at all time points analyzed. Furthermore, the baseline NT-proBNP level was significantly higher than values at the 7-day and 90-day follow-up. For this reason, different cut-off values were obtained at different times: 313 pg/ml at baseline [odds ratio (OR) = 18.99,  $P < 0.001$ ], 181 pg/ml at 7 days (OR = 11.4,  $P = 0.001$ ) and 174 pg/ml (OR = 8.46,  $P < 0.001$ ) at 90 days.

**Conclusion:** High levels of NT-proBNP determined during the first 3 months after a TIA were associated with AF. Consequently, this biomarker may be useful to reclassify undetermined TIA patients as having disease of CE.

## Introduction

As in stroke patients, etiological subtype diagnosis in transient ischaemic attack (TIA) patients leads to specific therapeutic actions to reduce recurrences and it has important prognosis implications that can guide management decisions [1].

Therefore, it is of great interest to find biological markers offering precise information about TIA etiology, particularly in those patients for whom a large artery atherosclerosis is excluded. Elevated serum levels of brain natriuretic peptide (BNP) have been asso-

ciated with atrial fibrillation (AF) [2] and stroke of cardioembolic etiology (CE) [3]. N-terminal pro-BNP (NT-proBNP), which coexists in circulation with BNP in equal proportion, is a better target than BNP for diagnostic blood testing because of its longer half-life and more stable concentration in the blood [4].

The aim of our study was to correlate CE with the presence or new diagnosis of AF with different serum biomarkers measured at different times in TIA patients.

## Methods

Patients with a TIA stroke attended by a neurologist in the emergency department of the Hospital Universitari Arnau de Vilanova were prospectively recruited from January 2008 to January 2010. TIA was defined

Correspondence: Francisco Purroy, Stroke Unit, Department of Neurology, Professor Universitat de Lleida, IRBLleida, Hospital Universitari Arnau de Vilanova de Lleida, Avda Rovira Roure, 80. Lleida 25198, Spain (tel.: +34 973248100; fax: +34 973248754; e-mail: fpurroygarcia@gmail.com).

clinically as a reversible episode of neurological deficit of ischaemic origin that resolved completely within 24 h [5]. Patients with a known inflammatory or malignant disease were excluded from the final analysis. Controls were stroke-free patients from primary-care centers with the common characteristic of having at least one cardiovascular risk score.

A total of 140 TIA patients and 44 control subjects with no history of IS or TIA were included in the study. Cardiovascular risk factors were evaluated for both groups.

Duration and typology of clinical symptoms were recorded and the ABCD2 risk score was determined prospectively [6]. All patients underwent electrocardiography, cervical and transcranial duplex ultrasonography and CT. Cranial MRI with diffusion-weighted image sequences were performed in 135 patients [mean = 3.8 (interquartile range, IQR 2.5–5.4) days] [7]. The following cardiological examinations were performed to complete the study of patients with suspected cardioembolism: transthoracic echocardiography in 54 cases, transesophageal echocardiography in seven cases, Holter ECG in 34 cases, continuous ECG monitoring in 41 cases, and right-to-left shunt detection in 57 cases.

Transient ischaemic attacks were classified etiologically according to the ORG 10172 Trial definitions [8] as large-artery occlusive disease (LAA), small-vessel disease (SV), CE, or undetermined causes (UND). Patients with uncommon causes were excluded. In the CE subtype, the presence of a high- or medium-risk cardiac source of embolism is necessary. The Trial of Org 10172 in Acute Stroke Treatment (TOAST) classification was performed at discharge, when the diagnostic work-up was completed, but those patients belonging to the UND group were followed up in the outpatient clinic until the etiology of stroke was reached or UND was confirmed.

Blood samples were taken upon arrival in the emergency department arrival within 24 h of symptom onset (baseline level), at 7 days and at 90 days to test a panel of biomarkers that included high-sensitivity C-reactive protein (hs-CRP), interleukin-1-alpha (IL-1 $\alpha$ ), IL-6, tumor necrosis factor-alpha (TNF- $\alpha$ ), neuron-specific enolase (NSE) and BNP. Blood was drawn into ethylenediaminetetra-acetic acid (EDTA) tubes centrifuged at 3000 rpm for 15 min, and plasma was frozen at  $-80^{\circ}\text{C}$  until analysis. However, it was not possible to measure the plasma biomarker in three patients at 7 days, and in 14 patients at 90 days.

Data are reported using standard descriptive statistics. To calculate the sensitivity and specificity for biomarker cut-off values to predict AF, a receiver operator characteristic (ROC) curve was configured.

Variables associated in the univariate analysis were entered into a sequential logistic-regression model to identify variables independently associated with AF (ORs with their 95% CI are given).  $P < 0.05$  was considered to be statistically significant.

## Results

A total of 184 subjects were included in our study: 140 with TIAs and 44 controls. Previously excluded subjects were as follows: three for cancer, two for infectious diseases, one for rheumatoid arthritis and 10 for neurological transient episodes attributable to causes other than brain ischaemia (hypoglycemia, two; subdural hemorrhage, one; epilepsy, three; migraine, three; and hypoglycemia, one). Compared with controls, TIA patients comprised a significantly higher proportion of men, those with previous stroke and hypertension (Table S1). Moreover, they showed significantly higher median plasma levels of NT-proBNP, IL-1 $\alpha$ , NSE, TNF- $\alpha$  and hs-CRP. The distribution of TIA subtypes was as follows: LAA,  $n = 25$ ; CE,  $n = 43$ ; SV,  $n = 25$ ; and UND,  $n = 47$  (Table 1). The proportion of diabetes mellitus ( $P = 0.011$ ) and previous AF ( $P < 0.001$ ) varied significantly among the different subtypes. Patients with SV were younger than patients with other etiologies ( $P = 0.044$ ). Plasma pro-BNP levels determined at baseline and at 7 and 90 days' follow-up were significantly higher in the CE patients than in the patients with other etiologies ( $P < 0.001$ ).

Cardioembolic etiologies were as follows: detection of patent foramen ovale, seven cases (16.3%); only abnormalities related to Heart failure, one case (2.3%); valvulopathy, one case (2.3%); akinetic or hypokinetic left ventricular segment, three cases (6.9%); AF, 31 cases (72.1%). Sixteen of the 31 patients developed AF during or after admission. New AF was detected in 12 patients by performing an ECG during admission and in four cases by ECG monitoring or by Holter ECG. In all but two UND patients, a Holter ECG or ECG monitoring was carried out.

Among subjects classified as CE-only, patients with AF showed significantly higher median plasma levels of NT-proBNP and IL-1 $\alpha$  determined at different times (Table S2). Nevertheless, the area under the ROC curve analysis only demonstrated good discriminating ability of NT-proBNP to detect AF (Table S3). Patients with persistent AF ( $n = 15$ ) had higher levels of this biomarker than patients with paroxysmal AF ( $n = 16$ ) (data not shown).

The NT-proBNP level at baseline was significantly higher than at the 7- and 90-day follow-up [median (IQR): 218.7 (77.1–506.0) vs. 146.2 (51.9–406.0) vs.

**Table 1** Characteristics and biomarker levels by TOAST subtype

Variable	LAA (n = 25)	CE (n = 43)	SV (n = 25)	UND (n = 47)	Total (n = 140)	P
Risk factors						
Mean (SD) age (years)	72.8 (9.1)	72.5 (13.6)	66.4 (10.6)	67.0 (12.3)	69.6 (12.2)	0.044
Male	18 (72.0)	26 (60.5)	15 (60.0)	26 (55.3)	85 (52.8)	0.590
Previous stroke	6 (24.0)	8 (18.6)	2 (8.0)	8 (17.0)	24 (17.1)	0.500
Hypertension	12 (48.0)	28 (65.1)	16 (64.0)	33 (70.2)	89 (63.6)	0.313
Coronary artery disease	1 (4.0)	9 (20.9)	1 (4.0)	4 (8.5)	15 (10.7)	0.063
Diabetes mellitus	10 (40.0)	5 (11.6)	9 (36.0)	19 (40.4)	43 (30.7)	0.013
Smoking	4 (16.0)	5 (11.6)	6 (24.0)	7 (14.9)	22 (15.7)	0.601
Hypercholesterolemia	9 (33.3)	14 (32.6)	7 (28.0)	18 (38.3)	48 (34.3)	0.834
Previous atrial fibrillation	0 (0)	15 (34.9)	0 (0)	0 (0)	15 (10.7)	<0.001
Heart failure	2 (8.0)	4 (9.3)	0 (0)	0 (0)	6 (4.3)	0.082
Mean (SD) ABCD2	4.9 (1.4)	4.8 (1.3)	4.9 (1.4)	4.7 (1.4)	4.8 (1.3)	0.947
Clinical features						
Duration (h)	0.75 (0.33–5.00)	1.00 (0.50–4.00)	2.5 (0.38–6.75)	0.75 (0.25–2.50)	1.00 (0.33–4.00)	0.402
Cluster TIA	6 (24.0)	8 (19.0)	6 (24.0)	11 (23.4)	31 (22.3)	0.946
Weakness	14 (56.0)	25 (58.1)	18 (72.0)	20 (42.6)	77 (55.0)	0.110
Isolated sensory symptoms	0 (0)	2 (5.0)	4 (16.0)	4 (8.5)	10 (7.1)	0.144
Speech impairment	13 (52.0)	25 (58.1)	14 (56.0)	29 (63.0)	81 (58.3)	0.828
Positive DWI	13 (54.2)	24 (57.1)	12 (52.2)	18 (39.1)	67 (49.6)	0.357
NT-proBNP levels (pg/ml)						
Basal	198.2 (89.6–376.3)	541.9 (313.9–1662.0)	173.2 (49.3–305.4)	135.1 (58.7–289.3)	218.7 (771–506.0)	<0.001
7 days	151.6 (52.5–361.8)	388.2 (186.8–918.8)	97.6 (36.7–441.5)	67.8 (31.6–138.3)	148.9 (52.5–432.6)	<0.001
90 days	104.5 (76.3–194.5)	323.6 (138.8–1233.0)	76.8 (67.4–310.0)	88.4 (40.9–159.8)	110.7 (67.7–309.4)	<0.001
IL-1- $\alpha$ levels (pg/ml)						
Basal	11.1 (7.9–17.3)	9.9 (5.9–16.4)	9.0 (6.3–14.9)	9.1 (6.9–14.1)	9.8 (6.8–16.1)	0.547
7 days	12.0 (6.6–15.6)	9.8 (5.7–15.6)	8.3 (5.6–13.1)	8.0 (5.9–12.3)	8.6 (5.9–14.4)	0.303
90 days	9.7 (5.5–14.1)	8.7 (4.8–13.1)	8.1 (5.3–12.3)	7.8 (5.9–12.4)	8.1 (5.4–13.0)	0.835
IL-6 levels (pg/ml)						
Basal	4.1 (2.5–10.2)	4.2 (1.5–9.5)	3.7 (1.5–6.3)	2.5 (1.5–6.4)	3.5 (1.5–8.1)	0.299
7 days	6.5 (3.1–8.9)	3.6 (1.6–9.8)	3.2 (1.5–4.8)	2.6 (1.5–5.9)	3.3 (1.5–7.6)	0.054
90 days	3.7 (1.6–6.7)	3.3 (1.5–5.3)	1.9 (1.5–6.6)	1.5 (1.5–5.9)	2.7 (1.5–5.9)	0.304
NSE levels (pg/ml)						
Basal	6.4 (4.7–10.2)	6.9 (4.9–10.4)	7.9 (4.7–11.7)	6.7 (4.9–7.9)	6.8 (4.8–10.1)	0.670
7 days	6.2 (4.0–8.0)	7.5 (4.6–10.1)	4.4 (3.9–8.3)	6.7 (4.1–8.7)	6.6 (4.1–8.6)	0.427
90 days	7.1 (4.2–10.7)	8.3 (5.1–9.8)	8.6 (4.4–10.4)	5.8 (4.0–10.1)	7.5 (4.8–10.4)	0.557
TNF-alpha levels (pg/ml)						
Basal	11.0 (6.3–16.4)	9.0 (5.9–14.8)	8.7 (6.7–13.4)	8.1 (5.9–12.7)	8.7 (6.4–14.5)	0.713
7 days	10.9 (5.9–14.6)	8.4 (6.1–13.6)	7.4 (6.0–12.6)	7.0 (5.5–11.5)	8.0 (5.9–13.4)	0.444
90 days	7.8 (5.1–12.8)	7.3 (5.0–12.1)	7.8 (5.2–12.1)	6.8 (4.8–11.7)	7.3 (5.1–12.2)	0.835
hs-CRP levels (mg/l)						
Basal	3.5 (1.5–9.2)	3.9 (1.6–13.2)	2.7 (1.2–6.8)	4.3 (1.4–8.9)	3.2 (1.4–9.6)	0.327
7 days	3.9 (1.8–7.9)	5.3 (1.8–13.8)	1.8 (0.7–4.7)	2.3 (1.3–9.2)	3.5 (1.4–10.4)	0.067
90 days	4.2 (1.7–5.6)	3.3 (1.6–7.0)	1.5 (1.0–4.9)	4.2 (1.6–8.8)	3.6 (1.5–6.4)	0.227

LAA, large-artery occlusive disease; CE, cardioembolic etiology; SV, small-vessel disease; UND, undetermined cause; TIA, transient ischaemic attack; DWI, diffusion-weighted image; NT-proBNP, N-terminal pro-B type natriuretic peptide; IL-1, IL-6, interleukin-1, -6; NSE, neuron-specific enolase; TNF-alpha, tumor necrosis factor-alpha; hs-CRP, high-sensitivity C-reactive protein; TOAST, Trial of Org 10172 in Acute Stroke Treatment ABCD2.

110.7 (67.7–309.4) pg/ml, respectively;  $P < 0.001$ . For this reason, different cut-off values were obtained at different times: 313 pg/ml with a sensitivity of 89.7% and a specificity of 74.5% at baseline level; 181 pg/ml with a sensitivity of 88.0% and a specificity of 74.5% at 7 days; and 174 pg/ml with a sensitivity of 83.3% and a specificity of 71.9% at 90 days. Patients with AF were significantly older, with a higher proportion of previous coronary heart and speech impairment,

than those with other cardioembolism or no cardioembolism, whereas, patients without cardioembolism had a greater incidence of diabetes mellitus than patients with AF.

Table 2 shows the logistic regression analysis. In all models, NT-proBNP levels determined at different times remained as independent predictors of previous or new diagnosis of AF. No time determination was better than the others.

**Table 2** Sequential binary regression model to assess relation to the presence or new diagnosis of atrial fibrillation (AF)

Variable	Model 1		Model 2		Model 3	
	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
Age	1.05 (0.98–1.13)	0.171	1.06 (0.99–1.14)	0.110	1.08 (1.01–1.16)	0.026
Coronary disease	8.13 (1.29–51.12)	0.026	3.48 (0.67–17.98)	0.137	6.24 (1.13–34.63)	0.036
Speech impairment	3.86 (1.11–13.40)	0.033	3.17 (0.85–11.77)	0.085	3.51 (1.04–11.81)	0.042
Diabetes mellitus	0.20 (0.05–0.78)	0.021	0.14 (0.3–0.64)	0.011	0.20 (0.5–0.77)	0.019
Basal pro-BNP > 313 pg/ml	18.99 (4.39–82.18)	< 0.001	—	—	—	—
7-day pro-BNP >181 pg/ml	—	—	11.4 (2.72–44.74)	0.001	—	—
90-day pro-BNP >174 pg/ml	—	—	—	—	8.46 (2.59–27.66)	<0.001

OR, odds ratio; pro-BNP, pro-B type natriuretic peptide.

## Discussion

The results of our study are in accordance with previous reports of stroke patients. NT-proBNP was associated with CE and particularly with AF [9–12]. But, to the best of our knowledge, no previous reports have determined this biomarker at different times in TIA patients. Our results are consistent with those obtained in stroke patients. As in a recent study performed with stroke patients [13] a serial change of plasma NT-proBNP level was observed. Therefore, the best cut off point to predict AF should be adapted to time determination. However, the plasma NT-proBNP level was significantly higher in patients with AF in all phases after a TIA than in patients without AF.

Brain natriuretic peptide is mainly secreted by the ventricle and it has a stable concentration over time. The correlation between AF and high levels of BNP is well known. Recently, BNP has been defined as a predictor of delayed AF in stroke patients [11]. According to our results, NT-proBNP measurement could be a non-invasive diagnostic tool to help decide on the initiation of early anticoagulant therapy among patients with suspicious CE. Interestingly, patients who were initially classified as UND and who were then reclassified as CE after diagnosis of new AF had significantly higher levels of NT-proBNP at all times (data not shown). The correct cut-off point depends on the time at which the NT-proBNP was determined.

This study had several limitations, in particular the fact that cardioembolic explorations were not exhaustive and were made especially in patients with suspected cardioembolism stroke. Not all patients with paroxysmal AF were diagnosed by continuous ECG or 24-h Holter ECG. Moreover, the establishment of a correct cut-off point is difficult when only a small number of patients are analyzed. Previously, different cut-off points of NT-proBNP or BNP were proposed

to predict CE or new AF diagnosis that varied from 85 pg/ml [11] to 491 pg/ml [14]. This variability could be explained because there were multiple assays for its determination [15]. A multicenter replication of these results with a larger sample size than ours is of interest. Finally, controls were not matched for gender with TIA subjects. However, the NT-proBNP level, as in previous reports [15], was only correlated with age, not with gender or other vascular risk factors.

We conclude that determination of NT-proBNP level may be a feasible strategy to improve cardioembolic diagnosis, especially AF, among TIA patients, not only in the acute phase but also during the first 90 days of follow-up. More studies are needed to establish the optimal cut-off point at different evolutionary times.

## Acknowledgements

This project has been partially supported by grants from the Spanish Ministry of Health (Instituto de Salud Carlos III – PI081398 and PI1102033). Some of the diagnostic kits were kindly donated by Roche. We express our sincere gratitude to the technicians of the core laboratory, Alexandra Beà and Sonia Carrasco. Samples were obtained with the support of IRBLleida biobank and RETICS BIOBANCOS (RD09/0076/00059).

## Disclosure of conflicts of interest

The authors declare no financial or other conflict of interests.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Table S1.** Characteristics and biomarker levels for TIA patients and controls.

**Table S2.** Characteristics and biomarker levels of the atrial fibrillation group, the other cardioembolism group and the no cardioembolism group.

**Table S3.** Predictive accuracy of biomarkers for the presence or new diagnosis of atrial fibrillation.

## References

- Purroy F, Montaner J, Molina CA, Delgado P, Ribo M, Alvarez-Sabin J. Patterns and predictors of early risk of recurrence after transient ischemic attack with respect to etiologic subtypes. *Stroke* 2007; **38**: 3225–3229.
- Schnabel RB, Larson MG, Yamamoto JF, et al. Relations of biomarkers of distinct pathophysiological pathways and atrial fibrillation incidence in the community. *Circulation* 2010; **121**: 200–207.
- Rodriguez-Yanez M, Sobrino T, Blanco M, et al. High serum levels of pro-brain natriuretic peptide (pro BNP) identify cardioembolic origin in undetermined stroke. *Dis Markers* 2009; **26**: 189–195.
- Bai M, Yang J, Li Y. Serum N-terminal-pro-brain natriuretic peptide level and its clinical implications in patients with atrial fibrillation. *Clin Cardiol* 2009; **32**: E1–E5.
- Special report from the National Institute of Neurological Disorders and Stroke. Classification of cerebrovascular diseases III. *Stroke* 1990; **21**: 637–676.
- Johnston SC, Rothwell PM, Nguyen-Huynh MN, et al. Validation and refinement of scores to predict very early stroke risk after transient ischaemic attack. *Lancet* 2007; **369**: 283–292.
- Purroy F, Begue R, Gil MI, et al. Patterns of diffusion-weighted magnetic resonance imaging associated with etiology improve the accuracy of prognosis after transient ischaemic attack. *Eur J Neurol* 2011; **18**: 121–128.
- Adams HP Jr, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke* 1993; **24**: 35–41.
- Di Angelantonio E, De Castro S, Toni D, et al. Determinants of plasma levels of brain natriuretic peptide after acute ischemic stroke or TIA. *J Neurol Sci* 2007; **260**: 139–142.
- Yukiiri K, Hosomi N, Naya T, et al. Plasma brain natriuretic peptide as a surrogate marker for cardioembolic stroke. *BMC Neurol* 2008; **8**: 45.
- Okada Y, Shibasaki K, Kimura K, Iguchi Y, Miki T. Brain natriuretic peptide as a predictor of delayed atrial fibrillation after ischaemic stroke and transient ischaemic attack. *Eur J Neurol* 2010; **17**: 326–331.
- Tamura H, Watanabe T, Satoshi N, et al. Elevated plasma brain natriuretic peptide levels predict left atrial appendage dysfunction in patients with acute ischemic stroke. *J Cardiol* 2012; **60**: 126–132.
- Jensen JK, Mickley H, Bak S, Korsholm L, Kristensen SR. Serial measurements of N-terminal pro-brain natriuretic peptide after acute ischemic stroke. *Cerebrovasc Dis* 2006; **22**: 439–444.
- Fonseca AC, Matias JS, Pinho e Melo T, Falcao F, Canhao P, Ferro JM. N-terminal probrain natriuretic peptide as a biomarker of cardioembolic stroke. *Int J Stroke* 2011; **6**: 398–403.
- Wu AH, Smith A, Wieczorek S, et al. Biological variation for N-terminal pro- and B-type natriuretic peptides and implications for therapeutic monitoring of patients with congestive heart failure. *Am J Cardiol* 2003; **92**: 628–631.