



US 20220133746A1

(19) **United States**

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

(10) **Pub. No.: US 2022/0133746 A1**

(43) **Pub. Date: May 5, 2022**

(54) **METHODS AND COMPOSITIONS RELATED TO GLUCOCORTICOID RECEPTOR ANTAGONISTS AND BREAST CANCER**

<i>A61K 33/243</i>	(2006.01)
<i>A61K 31/282</i>	(2006.01)
<i>A61K 31/337</i>	(2006.01)
<i>A61K 31/357</i>	(2006.01)
<i>A61K 31/4745</i>	(2006.01)
<i>A61K 31/7068</i>	(2006.01)
<i>A61K 39/395</i>	(2006.01)
<i>A61N 5/00</i>	(2006.01)
<i>A61J 1/00</i>	(2006.01)
<i>A61K 31/567</i>	(2006.01)
<i>A61K 31/58</i>	(2006.01)

(71) Applicant: **The University of Chicago**, Chicago, IL (US)

(72) Inventors: **Deng PAN**, Chicago, IL (US); **Masha KOCHERGINSKY**, Chicago, IL (US); **Suzanne D. CONZEN**, Park Ridge, IL (US)

(21) Appl. No.: **17/577,882**

(52) **U.S. Cl.**

(22) Filed: **Jan. 18, 2022**

CPC *A61K 31/575* (2013.01); *A61K 45/06* (2013.01); *A61K 33/243* (2019.01); *A61K 31/282* (2013.01); *A61K 31/337* (2013.01); *A61K 31/357* (2013.01); *A61K 31/58* (2013.01); *A61K 31/7068* (2013.01); *A61K 39/395* (2013.01); *A61N 5/00* (2013.01); *A61J 1/00* (2013.01); *A61K 31/567* (2013.01); *A61K 31/4745* (2013.01)

Related U.S. Application Data

(63) Continuation of application No. 16/596,342, filed on Oct. 8, 2019, now Pat. No. 11,234,990, which is a continuation of application No. 15/448,827, filed on Mar. 3, 2017, now Pat. No. 10,441,596, which is a continuation of application No. 14/296,127, filed on Jun. 4, 2014, now Pat. No. 9,623,032, which is a continuation of application No. 14/172,051, filed on Feb. 4, 2014, now Pat. No. 9,149,485, which is a continuation of application No. 13/071,363, filed on Mar. 24, 2011, now Pat. No. 8,710,035.

(60) Provisional application No. 61/317,182, filed on Mar. 24, 2010.

Publication Classification

(51) **Int. Cl.**
A61K 31/575 (2006.01)
A61K 45/06 (2006.01)

(57) **ABSTRACT**

Embodiments of the invention are directed to methods of determining the prognosis of a breast cancer patient by evaluating the activity of the glucocorticoid receptor in tumor cells. Other embodiment include methods of treating breast cancer cells, particularly, chemo-resistant cells, with a glucocorticoid receptor antagonist and an anticancer agent or compound.

Specification includes a Sequence Listing.

**Primary human breast ductal epithelium, DCIS (60%)
invasive human cancers (~30-40%) exhibit significant
glucocorticoid receptor expression**

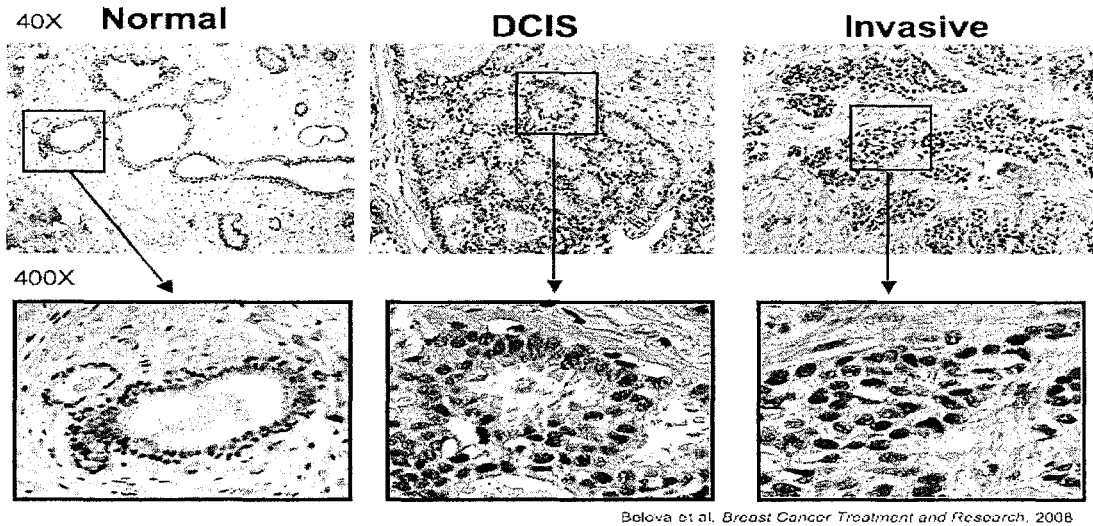


FIG. 1

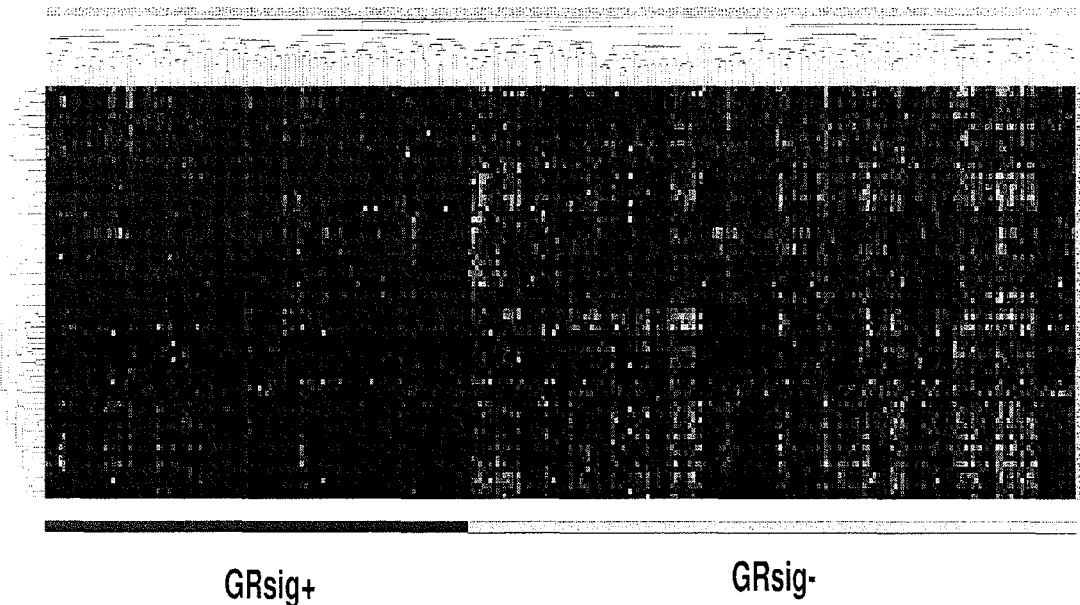


FIG. 2

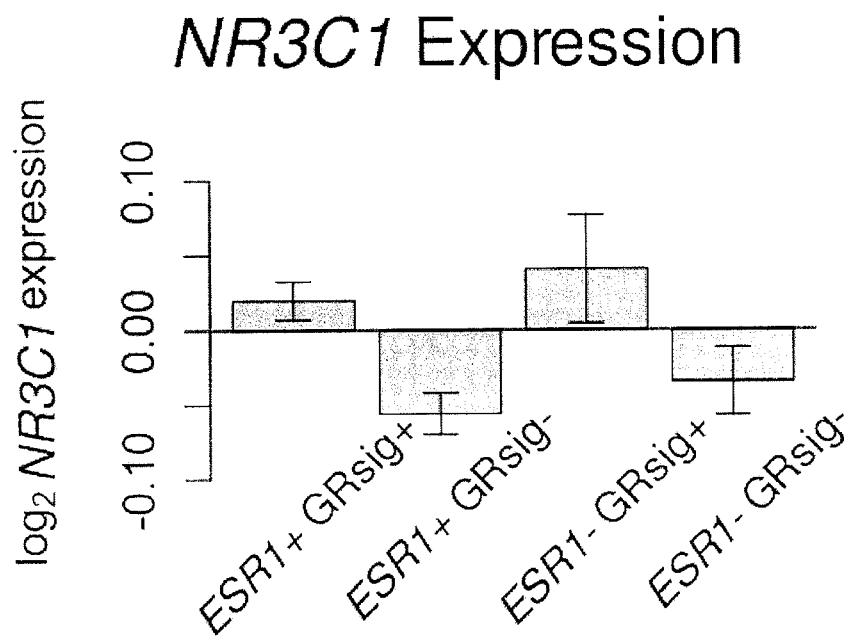


FIG. 3

NKI-295 RFS by Signiture

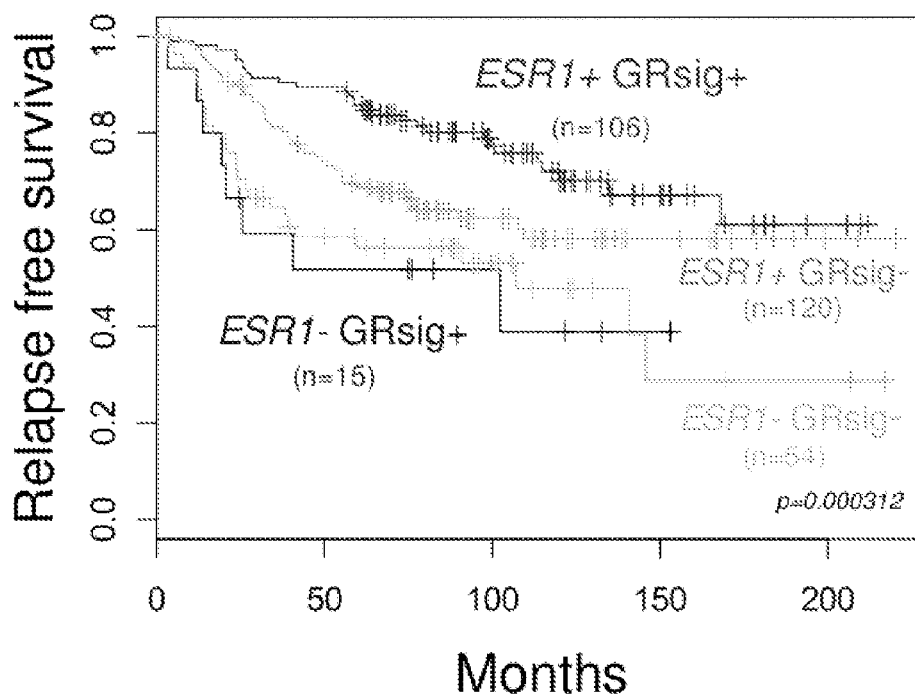


FIG. 4

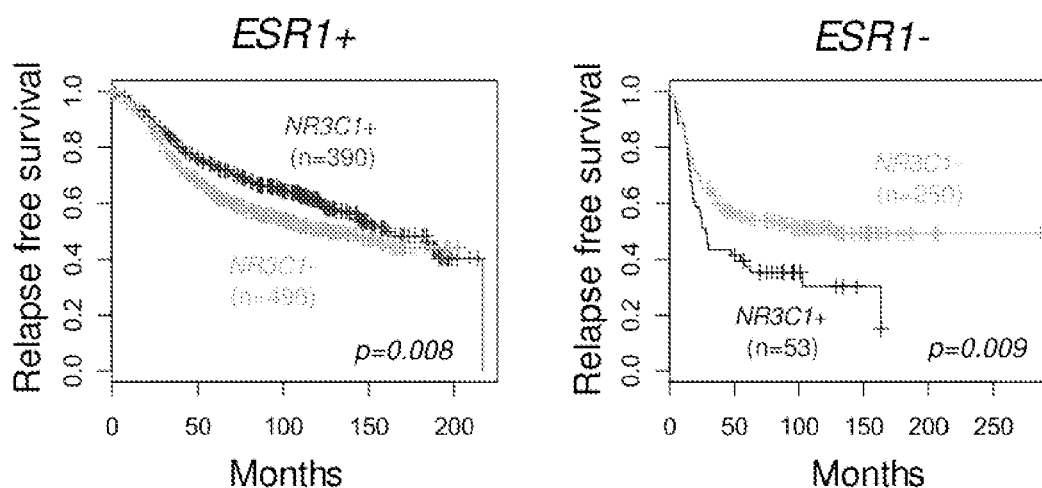


FIG. 5

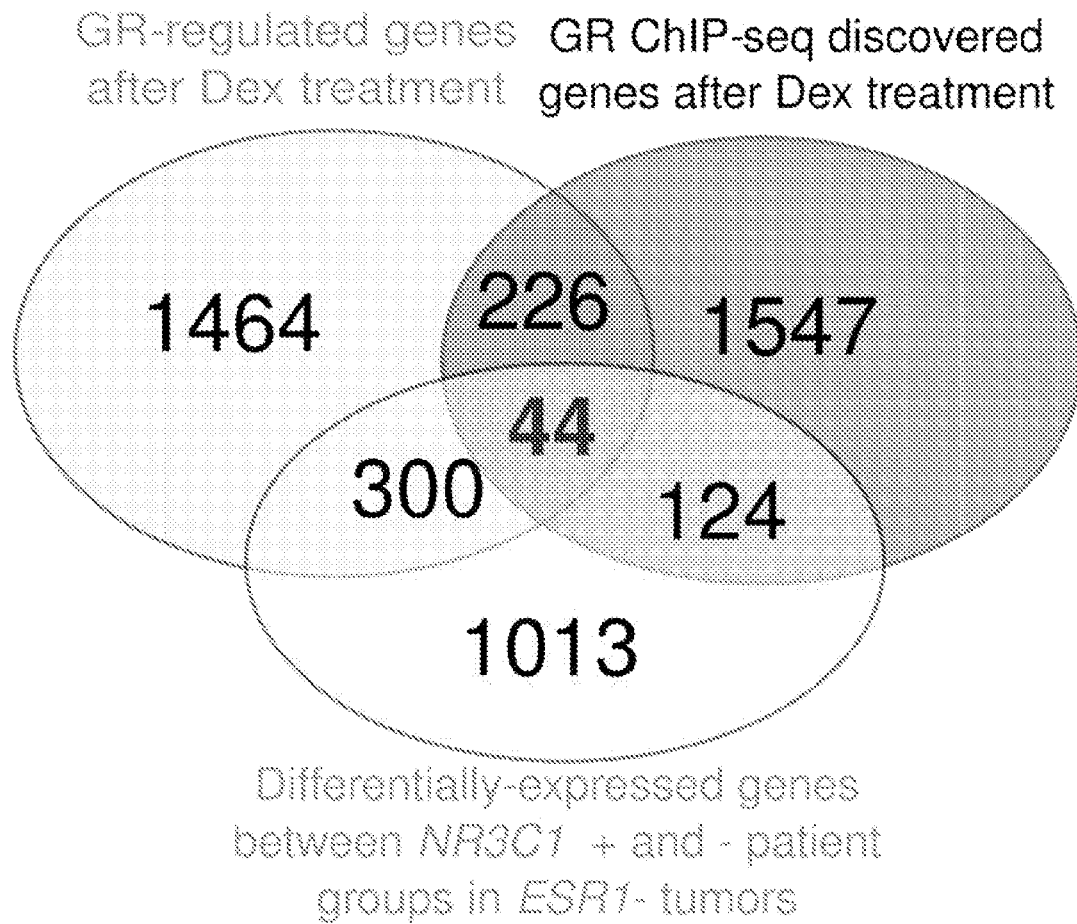
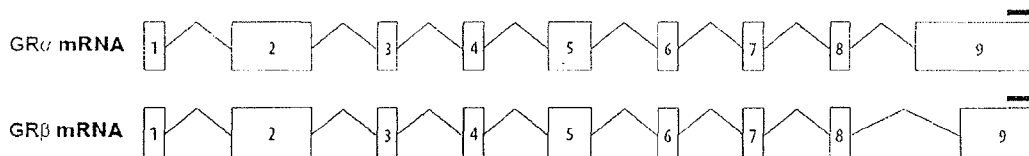


FIG. 6



Query = GR alpha
Length=6784

18665 = GR beta

ALIGNMENTS

Query	1	GGGCGCGCCTCCACCCGCTCCCCGCTCGGTCCCCTCGCTCGCCAGGCCGGGCTGCCCT	60
18665	1	GGGCGCGCCTCCACCCGCTCCCCGCTCGGTCCCCTCGCTCGCCAGGCCGGGCTGCCCT	60
Query	61	TTCGCGTGTCCGCGCTCTCTTCCCTCCGCGCGCCCTCCATTCATTTGCGAGCTCGTGTC	120
18665	61	TTCGCGTGTCCGCGCTCTCTTCCCTCCGCGCGCCCTCCATTCATTTGCGAGCTCGTGTC	120
Query	121	TGTGACGGGAGCCCGAGTCACCGCTGCCCGTCGGGGACGGATTCTGTGGGTGGAAGGAG	180
18665	121	TGTGACGGGAGCCCGAGTCACCGCTGCCCGTCGGGGACGGATTCTGTGGGTGGAAGGAG	180
Query	181	ACGCCGACGCCGGAGCGGCCGAAGCAGCTGGGACCGGGACGGGGCACGCGCCCGGAAC	240
18665	181	ACGCCGACGCCGGAGCGGCCGAAGCAGCTGGGACCGGGACGGGGCACGCGCCCGGAAC	240
Query	241	CTCGACCCCGGAGCCCGCGCGGGCGGAGGGCTGGCTTGTTCAGCTGGGCAATGGGAGA	300
18665	241	CTCGACCCCGGAGCCCGCGCGGGCGGAGGGCTGGCTTGTTCAGCTGGGCAATGGGAGA	300
Query	301	CTTCTTAAATAGGGGCTCTCCCCCACCCTGGAGAAAGGGCGGCTGTTTACTTCCTT	360
18665	301	CTTCTTAAATAGGGGCTCTCCCCCACCCTGGAGAAAGGGCGGCTGTTTACTTCCTT	360
Query	361	tttttAGaaaaaaaaaTATATTTCCCTCCTGCTCCTTCTGCGTTCACAAGCTAAGTTGT	420
18665	361	TTTTTAGAAAAAAAAAATATATTTCCCTCCTGCTCCTTCTGCGTTCACAAGCTAAGTTGT	420
Query	421	TTATCTCGGCTGCGCGGGAACTGCGGACGGTGGCGGGGAGCGGCTCCTCTGCCAGAGT	480
18665	421	TTATCTCGGCTGCGCGGGAACTGCGGACGGTGGCGGGGAGCGGCTCCTCTGCCAGAGT	480
Query	481	TGATATTCAGTGGACTCCAAAGAATCATTAACCTGGTAGAGAAGAAAACCCAGC	540
18665	481	TGATATTCAGTGGACTCCAAAGAATCATTAACCTGGTAGAGAAGAAAACCCAGC	540
Query	541	AGTGTGCTTGTCTCAGGAGAGGGGAGATGTGATGGACTTCTATAAAACCTAAGAGGAGGA	600
18665	541	AGTGTGCTTGTCTCAGGAGAGGGGAGATGTGATGGACTTCTATAAAACCTAAGAGGAGGA	600
Query	601	GCTACTGTGAAGGTTTCTGCGTCTTACCCTCACTGGCTGTGCGTCTCAATCAGACTCC	660
18665	601	GCTACTGTGAAGGTTTCTGCGTCTTACCCTCACTGGCTGTGCGTCTCAATCAGACTCC	660
Query	661	AAGCAGCGAAGACTTTTGGTTGATTTCCAAAAGGCTCAGTAAGCAATGCGCAGCAGCCA	720
18665	661	AAGCAGCGAAGACTTTTGGTTGATTTCCAAAAGGCTCAGTAAGCAATGCGCAGCAGCCA	720
Query	721	GATCTGTCAAAGCAGTTTCACTCTCAATGGGACTGTATATGGGAGAGACAGAAACAAA	780
18665	721	GATCTGTCAAAGCAGTTTCACTCTCAATGGGACTGTATATGGGAGAGACAGAAACAAA	780
Query	781	GTGATGGGAAATGACCTGGGATTCCACAGCAGGGCCAAATCAGCCTTTCCTCGGGGAA	840
18665	781	GTGATGGGAAATGACCTGGGATTCCACAGCAGGGCCAAATCAGCCTTTCCTCGGGGAA	840
Query	841	ACAGACTTAAAGCTTTTGGAAAGAAAGCATTGCAAACCTCAATAGGTGACACAGTGTTC	900
18665	841	ACAGACTTAAAGCTTTTGGAAAGAAAGCATTGCAAACCTCAATAGGTGACACAGTGTTC	900

FIG. 7A

Query	901	GAGAACCCCAAGAGTTCAGCATCCACTGCTGTGTCTGCTGCCCCACAGAGAAGGAGTTT	960
18665	901	GAGAACCCCAAGAGTTCAGCATCCACTGCTGTGTCTGCTGCCCCACAGAGAAGGAGTTT	960
Query	961	CCAAAACTCACCTCGATGTAATCTCAGAACAGCAACATTTGAAGGGCCAGACTGGCACC	1020
18665	961	CCAAAACTCACCTCGATGTAATCTCAGAACAGCAACATTTGAAGGGCCAGACTGGCACC	1020
Query	1021	AACGGTGGCAATGTGAAATTTGTATACCACAGACCAAAGCACCTTTGACATTTTGCAGGAT	1080
18665	1021	AACGGTGGCAATGTGAAATTTGTATACCACAGACCAAAGCACCTTTGACATTTTGCAGGAT	1080
Query	1081	TTGGAGTTTCTTCTGGGTCCCCAGGTAAAAGAGACGAATGAGAGTCTTTGGAGATCAGAC	1140
18665	1081	TTGGAGTTTCTTCTGGGTCCCCAGGTAAAAGAGACGAATGAGAGTCTTTGGAGATCAGAC	1140
Query	1141	CTGTGATAGATGAAAACGTTCCTTCTCTGCGGGAGAAGACGATTTCATTCCTT	1200
18665	1141	CTGTGATAGATGAAAACGTTCCTTCTCTGCGGGAGAAGACGATTTCATTCCTT	1200
Query	1201	TTGGAAGGAAAACTCGAATGAGGACTGCAAGCCTCTCATTTTACCGGACACTAAACCCAAA	1260
18665	1201	TTGGAAGGAAAACTCGAATGAGGACTGCAAGCCTCTCATTTTACCGGACACTAAACCCAAA	1260
Query	1261	ATTAAGGATAATGGAGATCTGGTTTTGTCAAGCCCCAGTAATGTAACACTGCCCAAGTG	1320
18665	1261	ATTAAGGATAATGGAGATCTGGTTTTGTCAAGCCCCAGTAATGTAACACTGCCCAAGTG	1320
Query	1321	AAAACAGAAAAGAAGATTCATCGAACTCTGCACCCCTGGGGTAATTAAGCAAGAGAAA	1380
18665	1321	AAAACAGAAAAGAAGATTCATCGAACTCTGCACCCCTGGGGTAATTAAGCAAGAGAAA	1380
Query	1381	CTGGGCACAGTTTACTGTCAGGCAAGCTTCTCTGGAGCAAATATAAATGGTAATAAAAATG	1440
18665	1381	CTGGGCACAGTTTACTGTCAGGCAAGCTTCTCTGGAGCAAATATAAATGGTAATAAAAATG	1440
Query	1441	TCTGCCATTTCTGTTCATGGTGTGAGTACCTCTGGAGGACAGATGTACCACATATGACATG	1500
18665	1441	TCTGCCATTTCTGTTCATGGTGTGAGTACCTCTGGAGGACAGATGTACCACATATGACATG	1500
Query	1501	AAATACAGCATCCCCTTCTCAACAGCAGGATCAGAAGCCTATTTTAAATGTCAATCCACCA	1560
18665	1501	AAATACAGCATCCCCTTCTCAACAGCAGGATCAGAAGCCTATTTTAAATGTCAATCCACCA	1560
Query	1561	ATTCCTGGTGGTTCCGAAAATTTGGAATAGGTGCCAAGGATCTGGAGATGACAACCTTGACT	1620
18665	1561	ATTCCTGGTGGTTCCGAAAATTTGGAATAGGTGCCAAGGATCTGGAGATGACAACCTTGACT	1620
Query	1621	TCTCTGGGACTCTGAACTTCCCTGGTCCGAACAGTTTTTTCTAATGGCTATTCAGCCCC	1680
18665	1621	TCTCTGGGACTCTGAACTTCCCTGGTCCGAACAGTTTTTTCTAATGGCTATTCAGCCCC	1680
Query	1681	AGCATGAGACCAGATGTAAGCTCTCCTCCATCCAGCTCCTCAACAGCAACAACAGGACCA	1740
18665	1681	AGCATGAGACCAGATGTAAGCTCTCCTCCATCCAGCTCCTCAACAGCAACAACAGGACCA	1740
Query	1741	CCTCCCAAACCTCTGCCTGGTGTGCTCTGATGAAGCTTCAGGATGTCATTATGGAGTCTTA	1800
18665	1741	CCTCCCAAACCTCTGCCTGGTGTGCTCTGATGAAGCTTCAGGATGTCATTATGGAGTCTTA	1800
Query	1801	ACTTGTGGAAGCTGTAAAGTTTTCTTCAAAGAGCAGTGGGAGGACAGCACAATTACCTA	1860
18665	1801	ACTTGTGGAAGCTGTAAAGTTTTCTTCAAAGAGCAGTGGGAGGACAGCACAATTACCTA	1860
Query	1861	TGTGCTGGAAGGAATGATTGCATCATCGATAAAAATTCGAAGAAAAAACTGCCAGCATGC	1920
18665	1861	TGTGCTGGAAGGAATGATTGCATCATCGATAAAAATTCGAAGAAAAAACTGCCAGCATGC	1920
Query	1921	CGCTATCGAAAATGTCTTCAGGCTGGAATGAACCTGGAAGCTCGAAAAcaaaagaaaaa	1980
18665	1921	CGCTATCGAAAATGTCTTCAGGCTGGAATGAACCTGGAAGCTCGAAAAcaaaagaaaaa	1980
Query	1981	ataaaaGGAATTCAGCAGGCCACTACAGGAGTCTCACAAGAAACCTCTGAAAATCCTGGT	2040
18665	1981	ATAAAAAGGAATTCAGCAGGCCACTACAGGAGTCTCACAAGAAACCTCTGAAAATCCTGGT	2040
Query	2041	AACAAAACAATAGTTCCTGCAACGTTACCACAACCTCACCCCTACCCCTGGTGTCACTGTTG	2100
18665	2041	AACAAAACAATAGTTCCTGCAACGTTACCACAACCTCACCCCTACCCCTGGTGTCACTGTTG	2100
Query	2101	GAGGTTATGAACTGAAGTGTATATGCAGGATATGATAGCTCTGTTCAGACTCAACT	2160
18665	2101	GAGGTTATGAACTGAAGTGTATATGCAGGATATGATAGCTCTGTTCAGACTCAACT	2160

FIG. 7B

Query	2161	TGGAGGATCATGACTACGCTCAACATGTTAGGAGGGCGGCAAGTGATTCAGCAGTGAAA	2220
18665	2161	TGGAGGATCATGACTACGCTCAACATGTTAGGAGGGCGGCAAGTGATTCAGCAGTGAAA	2220
Query	2221	TGGCAAAGGCAATACCAGGTTTCAGGAACTTACACCTGGATGACCAAATGACCCTACTG	2280
18665	2221	TGGCAAAGGCAATACCAGGTTTCAGGAACTTACACCTGGATGACCAAATGACCCTACTG	2280
Query	2281	CAGTACTCCTGGATGTTTCTTATGGCATTTGCTCTGGGGTGGAGATCATATAGACAATCA	2340
18665	2281	CAGTACTCCTGGATGTTTCTTATGGCATTTGCTCTGGGGTGGAGATCATATAGACAATCA	2340
Query	2341	AGTGCAAACCTGCTGTGTTTTGCTCCTGATCTGATTATTAATGAGCAGAGAATGACTCTA	2400
18665	2341	AGTGCAAACCTGCTGTGTTTTGCTCCTGATCTGATTATTAATGAGCAGAGAATGACTCTA	2400
Query	2401	CCCTGCATGTACGACCAATGTAACACATGCTGTATGTTTCCCTCGAGTTACACAGGCTT	2460
18665	2401	CCCTGCATGTACGACCAATGTAACACATGCTGTATGTTTCCCTCGAGTTACACAGGCTT	2460
Query	2461	CAGGTATCTTTATGAAGAGTATCTCTGTATGAAAACCTTACTGCTTCTCTCTCAGTTCCT	2520
18665	2461	CAGGTATCTTTATGAAGAGTATCTCTGTATGAAAACCTTACTGCTTCTCTCTCAGTTCCT	2520
Query	2521	AAGGACGGTCTGAAGAGCCAAGAGCTATTTTGATGAAATTAGAATGACCTACATCAAAGAG	2580
18665	2521	AAGGACGGTCTGAAGAGCCAAGAGCTATTTTGATGAAATTAGAATGACCTACATCAAAGAG	2580
Query	2581	CTAGGAAAAGCCATTTCTCAAGAGGGAAGGAAAACCCAGCCAGAACTGGCAGCGGTTTTAT	2640
18665	2581	CTAGGAAAAGCCATTTCTCAAGAGGGAAGGAAAACCCAGCCAGAACTGGCAGCGGTTTTAT	2640
Query	2641	CAACTGACAAAACCTTTGGATTTCTATGCATGAAGTGGTTGAAAATCTTCTTAACTATTC	2700
18665	2641	CAACTGACAAAACCTTTGGATTTCTATGCATGAAGTGGTTGAAAATCTTCTTAACTATTC	2673
Query	2701	TTCCAAACATTTTTGGATAAGACCATGAGTATTGAATTCCTCCGAGATGTTAGCTGAAATC	2760
Query	2761	ATCACCAATCAGATACCAAAATATTCAAATGGAAATATCAAAAAACTTCTGTTTCATCAA	2820
Query	2821	AAGTGACTGCCTTAATAAGAATGGTTGCCTTAAAGAAAGTCGAATTAATAGCTTTTATTG	2880
Query	2881	TATAAACTATCAGTTTGTCCGTAGAGgttttgttgtttatttttattgttttcatct	2940
Query	2941	gttgttttgttttAAATACGCCTACAATGTGGTTTATAGAGGGCCAAGACTTGGCAACAG	3000
Query	3001	AAGCAGTTGAGTCGTATCCTTTTCAGTGTATGGGAGAGTAGATGGTGAATTTATTAGT	3060
Query	3061	TAAATATATCCAGAAATTAGAAACCTTAATATGTGGACGTAATCTCCACAGTCAAAGAAG	3120
Query	3121	GATGGCACCTAAACCACCAGTGCCCAAAGTCTGTGTGATGAACTTTCTCTTCATACTttt	3180
Query	3181	tttCACAGTTGGCTGGATGAAATTTTCTAGACTTTCTGTGGTGTATcccccccTGTAT	3240
Query	3241	AGTTAGGATAGCATTTTIGATTTATGCATGGAAACCTGaaaaaaGTTTACAAGTGATA	3300
Query	3301	TCAGAAAAGGGAGTTTGTGCCTTTTATAGCTATTAAGTGTCTGGTTTAAACAATTTCTTT	3360
Query	3361	ATATTTAGTGAACACGCTTGCATTTTCTTACATAAATTTTTATTCAGTTATTTGT	3420
Query	3421	ACAGCTGTTTAAGATGGGCAGCTAGTTCGTAGCTTTCCCAAATAAACTCTAAACATTAAT	3480
Query	3481	CAATCATCTGTGTGAAAATGGGTTGGTGTCTTCTAACCTGATGGCACTTAGCTATCAGAAG	3540
Query	3541	ACCACAAAAATGACTCAAATCTCCAGTATTTCTTGTCAaaaaaaaaaaaaaaaaGCICA	3600
Query	3601	TATTTTGTATATATCTGCTTCAGTGGAGAATTATATAGGTTGTGCAAATTAACAGTCTTA	3660
Query	3661	ACTGGTATAGAGCACCTAGTCCAGTGACCTGCTGGGTAAACTGTGGATGATGGTTGCAAA	3720
Query	3721	AGACTAATTTAAAAATAACTACCAAGAGGCCCTGTCTGTACCTAACGCCCTATTTTTC	3780

FIG. 7C

Query	3781	AAATGGCTATATGGCAAGAAAGCTGGTAAACTATTTGTCTTTCAGGACCTTTTGAAGTAGT	3840
Query	3841	TTGTATAAAGCTCTTAAAAGTTGTGATTCCAGATAACCAGCTGTAACACAGCTGAGAGACT	3900
Query	3901	TTTAATCAGACAAAGTAATTCCTCTCACTAAACTTTACCCAAAACTAAATCTCTAATAT	3960
Query	3961	GGCAAAAATGGCTAGACACCCATTTTCACATTCCCATCTGTCCCAATFGGTTAATCTTT	4020
Query	4021	CCIGATGGTACAGGAAAGCTCAGCTACTGATTTTTGTGATTTAGAACTGTATGTCAGACA	4080
Query	4081	TCCATGTTTGTAAAACCTACACATCCCTAATGTGTGCCATAGAGTTTAAACACAAGTCCCTGT	4140
Query	4141	GAAATTCCTCACCTGTGAAAATTATTTTAAACAAAATAGAAGCTGTAGTAGCCCTTTCCTG	4200
Query	4201	TGTGCACCTTACCAACTTTCTGTAAACTCAAACTTAACATATTTACTAAGCCACAAGAA	4260
Query	4261	ATTTGATTTCTATTTCAAGGTGGCCAAAATATTTGTGTAATAGAAAACGAAAATCTAATA	4320
Query	4321	TTAAAAATATGGAACTTCTAatataatattttatatttagttatagtttcagatatatatca	4380
Query	4381	tatTGGTATTCACCTAATCTGGGAAGGGAAGGGCTACTGCAGCTTTACATGCAATTTATTA	4440
Query	4441	AAAATGATTCGTAATAAGCTTGTATAGTGTAAAATAAGAATGATTTTATAGATGAGATTGTT	4500
Query	4501	TTATCATGACATGTTATATAATTTTTGTAGGGGTCAAAGAAATGCTGATGGATAACCTAT	4560
Query	4561	ATGATTTATAGTTTGTACATGCATTCATACAGGCAGCGATGGTCTCAGAAACCAACAGT	4620
Query	4621	TTGCTCTAGGGGAAGAGGGAGATGGAGACTGGTCTCTGTGTGCAGTGAAGGTGCTGAGGC	4680
Query	4681	TCGACCCAGTGAGATTTACAGAGGAAGTTATCCCTCGCCCTCCCATTCGACCACCTTCT	4740
Query	4741	CATTCCAACAGTGAGTCTGTACAGCGCAGGTTTAGTTACTCAATCTCCCCCTGCACTAAA	4800
Query	4801	GTATGTAAGIATGTAAAACAGGACAGGAAGGTGGTGCTTACATCCCTAAAGGCACCAAT	4860
Query	4861	CTAATAGCGGGTTACTTTACATACAGCCCTCCCCAGCAGTTGAATGACAACAGAAGCT	4920
Query	4921	TCAGAAGTTTGGCAATAGTTTGCATAGAGGTACCAGCAATATGTAAATAGTGCAGAATCT	4980
Query	4981	CATAGGTTGCCAATAATACACTAATTCCTTTCTATCCTACAACAAGAGTTTATTTCCAAA	5040
Query	5041	TAAAATGAGGACATgtttttggttttctttgaaatgctttttgaaatgtaatttggtaattttc	5100
Query	5101	agtattttggagaaattatttAA'aaaaaaCAATCATTTGCTTTTTGAATGCTCTCTAA	5160
Query	5161	AAGGGAATGTAATATTTTAAGATGGTGTGTAACCCGGCTGCATAAAATTTTTGGTGCCTAA	5220
Query	5221	GAAAACCTGCTTGAATATTTCTTATCAATGACAGTGTAAAGTTTCAAAAAGAGCTTCTAAAA	5280
Query	5281	CGTAGATTATCATTCCTTTATAGAATGTTATGTGGTTAAAACAGAAAGCACATCTCACA	5340
18665	2674	AAATGTTATGTGGTTAAAACAGAAAGCACATCTCACA	2710
Query	5341	CATTAATCTGATTTTCATCCCAACAATCTTGGCGCTCAAAAAATAGAACTCAATGAGAAA	5400
18665	2711	CATTAATCTGATTTTCATCCCAACAATCTTGGCGCTCAAAAAATAGAACTCAATGAGAAA	2770
Query	5401	AAGAAGATTTATGTGCACCTTCGTTGTCAATAATAAGTCAACTGATGCTCATCGACAAC'AT	5460
18665	2771	AAGAAGATTTATGTGCACCTTCGTTGTCAATAATAAGTCAACTGATGCTCATCGACAAC'AT	2830
Query	5461	AGGAGGCTTTTTCAT'AAAATGGGAAAAGAAGCTGTGCCCTTTTAGGATACGTGGGGGAAAA	5520
18665	2831	AGGAGGCTTTTTCAT'AAAATGGGAAAAGAAGCTGTGCCCTTTTAGGATACGTGGGGGAAAA	2890
Query	5521	GAAAGTCAATCTTAAT'ATATGTTAAT'GTGGATTTAAGTGTATATGGTGGTGTCTTTGA	5580

FIG. 7D

18665	2891	GAAAGTCATCTTAAATATGTTTAAATGTGGATTAAAGTGCTATATGGTGGTGCCTGTTTGA	2950
Query	5581	AAGCAGATTTATTTCCCTATGTATGTGTATCTGGCCATCCCAACCCAAACTGTTGAAGTT	5640
18665	2951	AAGCAGATTTATTTCCCTATGTATGTGTATCTGGCCATCCCAACCCAAACTGTTGAAGTT	3010
Query	5641	TGTAGTAACCTCAGTGAGAGTTGGTTACTCACAACAAATCCTGAAAAGTATTTTGTAGTGT	5700
18665	3011	TGTAGTAACCTCAGTGAGAGTTGGTTACTCACAACAAATCCTGAAAAGTATTTTGTAGTGT	3070
Query	5701	TTGTAGGTATTTCTGTGGGATACTATACAAAGCAGAACTGAGGCACTTAGGACATAACACTT	5760
18665	3071	TTGTAGGTATTTCTGTGGGATACTATACAAAGCAGAACTGAGGCACTTAGGACATAACACTT	3130
Query	5761	TTGGGGTATATATATCCAAATGCCTAAAACCTATGGGAGGAAACCTTGGCCACCCAAAAG	5820
18665	3131	TTGGGGTATATATATCCAAATGCCTAAAACCTATGGGAGGAAACCTTGGCCACCCAAAAG	3190
Query	5821	GAAAACFAACATGATTTGTGTCTATGAAGTGCCTGGATAAATAGCATGGGATGAGCTCTGG	5880
18665	3191	GAAAACFAACATGATTTGTGTCTATGAAGTGCCTGGATAAATAGCATGGGATGAGCTCTGG	3250
Query	5881	GCATGCCATGAAGGAAAGCCACGCTCCCTTCAGAATTCAGAGGCAGGGAGCAATTCAGT	5940
18665	3251	GCATGCCATGAAGGAAAGCCACGCTCCCTTCAGAATTCAGAGGCAGGGAGCAATTCAGT	3310
Query	5941	TTACACCTAAGTCTCATAATTTTAGTTCCCTTTTAAAAACCTGAAAACCTACATCACCATG	6000
18665	3311	TTACACCTAAGTCTCATAATTTTAGTTCCCTTTTAAAAACCTGAAAACCTACATCACCATG	3370
Query	6001	GAATGAAAAATATTTGTATATACAATACATFGATCTGTCAAACCTCCAGAACCATGGTAGCC	6060
18665	3371	GAATGAAAAATATTTGTATATACAATACATFGATCTGTCAAACCTCCAGAACCATGGTAGCC	3430
Query	6061	TTCAGTGAGATTTCCATCTTGGCTGGTCACTCCCTGACTGTAGCTGTAGGTGAATgtgtt	6120
18665	3431	TTCAGTGAGATTTCCATCTTGGCTGGTCACTCCCTGACTGTAGCTGTAGGTGAATGTGTT	3490
Query	6121	tttgtgtgtgtgtgtctggTTTTAGTGTGAGAAGGGAAAATAAAGTGTAAAGGAGGACACT	6180
18665	3491	TTTTGTGTGTGTGTCTGGTTTTAGTGTGAGAAGGGAAAATAAAGTGTAAAGGAGGACACT	3550
Query	6181	TTAAACCCTTTGGGTGGAGTTTCGTAATTTCCAGACTATTTTCAAGCAACCTGGTCCAC	6240
18665	3551	TTAAACCCTTTGGGTGGAGTTTCGTAATTTCCAGACTATTTTCAAGCAACCTGGTCCAC	3610
Query	6241	CCAGGATTAGTGACCAGGTTTTCCAGGAAAGGATTTGCTTCTCTCTAGAAAATGTCTGAAA	6300
18665	3611	CCAGGATTAGTGACCAGGTTTTCCAGGAAAGGATTTGCTTCTCTCTAGAAAATGTCTGAAA	3670
Query	6301	GGATTTTTATTTCTGATGAAAGGCTGTATGAAAATACCCCTCCTCAAATAACTTGCCTAAC	6360
18665	3671	GGATTTTTATTTCTGATGAAAGGCTGTATGAAAATACCCCTCCTCAAATAACTTGCCTAAC	3730
Query	6361	TACATATAGATTCAAGTGTGTCAATATTTCTATTTTGTATATATAATGCTATATAATGGGG	6420
18665	3731	TACATATAGATTCAAGTGTGTCAATATTTCTATTTTGTATATATAATGCTATATAATGGGG	3790
Query	6421	ACAAAATCFATATTTACTGTGTATGGCATTTATTAAGAAGCTTTTTCATTAATTTTATCA	6480
18665	3791	ACAAAATCFATATTTACTGTGTATGGCATTTATTAAGAAGCTTTTTCATTAATTTTATCA	3850
Query	6481	CAGTAAATTTAAAAATGTGTAAAAATTTAAACCAGTGACTCCCTGTTTTAAAAATAAAGTTG	6540
18665	3851	CAGTAAATTTAAAAATGTGTAAAAATTTAAACCAGTGACTCCCTGTTTTAAAAATAAAGTTG	3910
Query	6541	TAGTTTTTTTATTTCATGCTGAATAATAATCTGTAGTTAAAAAAGTGTCTTTTACCTA	6600
18665	3911	TAGTTTTTTTATTTCATGCTGAATAATAATCTGTAGTTAAAAAAGTGTCTTTTACCTA	3970
Query	6601	CGCAGTGAAATGTCAGACTGTAAAACCTTGTGTGGAAATGTTTAACTTTTATTTTTCAT	6660
18665	3971	CGCAGTGAAATGTCAGACTGTAAAACCTTGTGTGGAAATGTTTAACTTTTATTTTTCAT	4030
Query	6661	TTAAATTTGCTGTCTGGTATFACCAAACCACACATTTGTACCGAATGGCAGTAAATGT	6720
18665	4031	TTAAATTTGCTGTCTGGTATFACCAAACCACACATTTGTACCGAATGGCAGTAAATGT	4090
Query	6721	TAGCCAATTTACAGCAATGCCAAATATGGAGAAACATCATAAATAAAAAATCTGCTTTTTTC	6780
18665	4091	TAGCCAATTTACAGCAATGCCAAATATGGAGAAACATCATAAATAAAAAATCTGCTTTTTTC	4150
Query	6781	ATTA 6784	

FIG. 7E

FIG. 7F

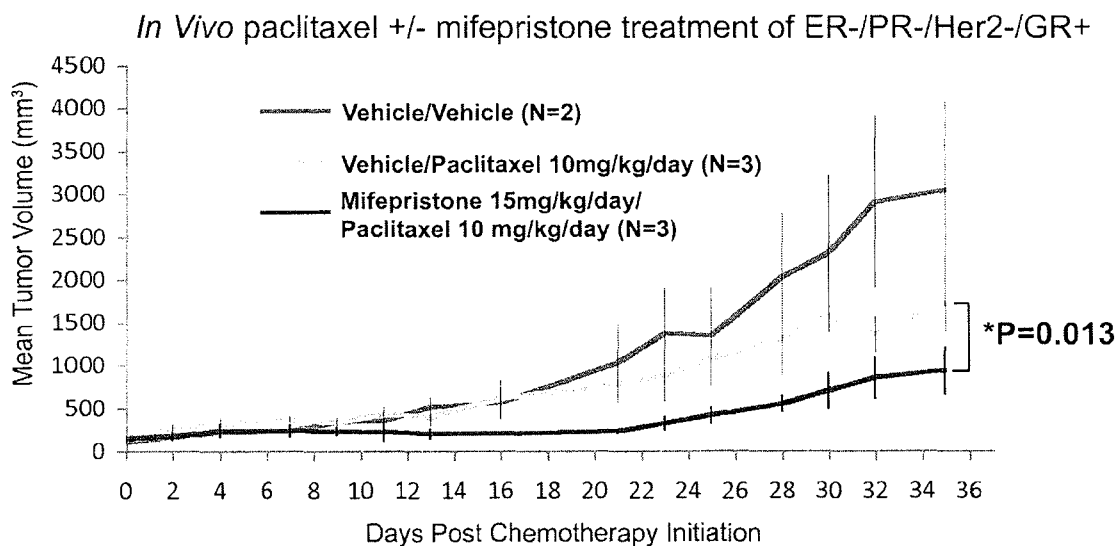


FIG. 8

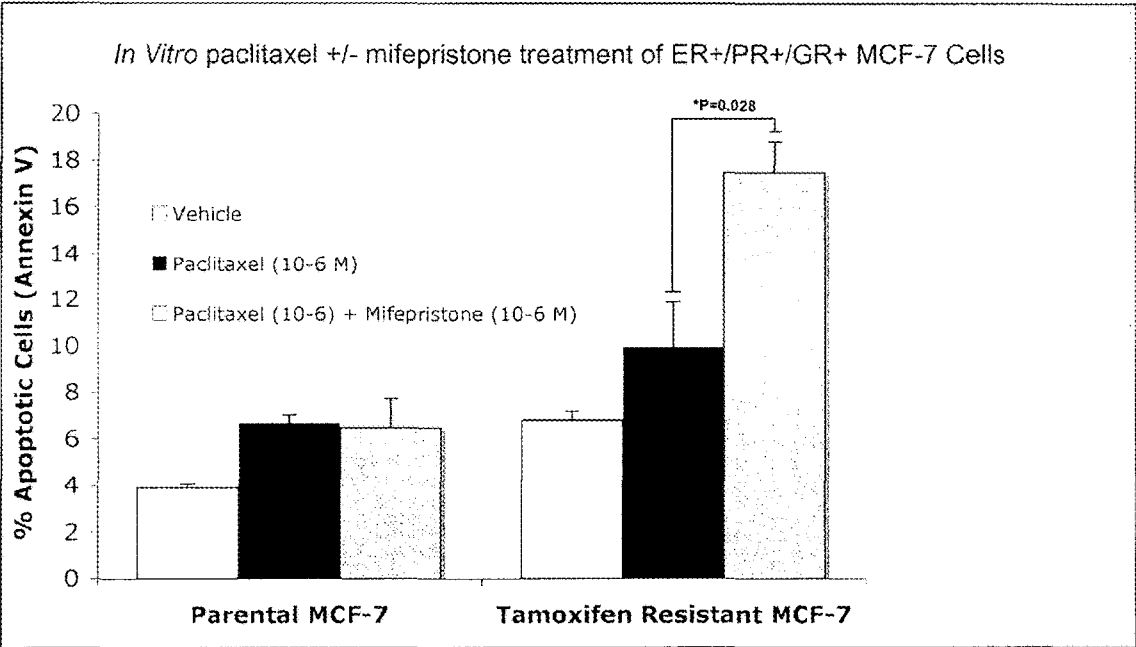


FIG. 9

**METHODS AND COMPOSITIONS RELATED
TO GLUCOCORTICOID RECEPTOR
ANTAGONISTS AND BREAST CANCER**

**CROSS-REFERENCES TO RELATED
APPLICATIONS**

[0001] This application is a Continuation of U.S. application Ser. No. 16/596,342, filed Oct. 8, 2019, which is a Continuation of U.S. application Ser. No. 15/448,827, filed Mar. 3, 2017 (now U.S. Pat. No. 10,441,596, issued Oct. 15, 2019), which is a Continuation of U.S. application Ser. No. 14/296,127, filed Jun. 4, 2014 (now U.S. Pat. No. 9,623,032, issued Apr. 18, 2017), which is a Continuation of U.S. application Ser. No. 14/172,051, filed Feb. 4, 2014 (now U.S. Pat. No. 9,149,485, issued Oct. 6, 2015), which is a Continuation of U.S. application Ser. No. 13/071,363, filed Mar. 24, 2011 (now U.S. Pat. No. 8,710,035, issued Apr. 29, 2014), which claims priority to U.S. Provisional Application No. 61/317,182, filed on Mar. 24, 2010, the disclosures of which are hereby incorporated by reference in their entireties.

[0002] This invention was made with government support under CA089208 awarded by the National Institutes of Health. The government has certain rights in the invention.

**REFERENCE TO A "SEQUENCE LISTING," A
TABLE, OR A COMPUTER PROGRAM LISTING
APPENDIX SUBMITTED ON AN ASCII TEXT
FILE**

[0003] The Sequence Listing written in file "SeqListing096487-1293223.TXT", created on Jan. 14, 2022, 225, 045 bytes, machine format IBM-PC, MS-Windows operating system, is hereby incorporated by reference in its entirety for all purposes.

BACKGROUND OF THE INVENTION

I. Field of the Invention

[0004] Embodiments of this invention are directed generally to biology and medicine. In certain aspects methods involve determining the prognosis for a breast cancer patient. In other embodiments, there are methods and compositions for treating a breast cancer patient with a glucocorticoid antagonist.

II. Background

[0005] There are over 1 million cases of breast cancer per year on a global basis, of which around 0.5 million are in the US, 40,000 are in the UK and nearly 2,000 in Ireland. It is the leading cause of cancer deaths among women (Keen and Davidson, 2003). Although the overall incidence of the disease is increasing within the western world, wider screening and improved treatments have led to a gradual decline in the fatality rate of about 1% per year since 1991. Inheritance of susceptibility genes, such as BRCA1 and BRCA2, account for only 5% of breast cancer cases and the factors responsible for the other 95% remain obscure (Grover and Martin, 2002). In the absence of a strategy to reduce causative agents of breast cancer, early detection remains the best approach to reducing the mortality rate of this disease. It is widely held that breast cancer initiates as the pre-malignant stage of atypical ductal hyperplasia (ADH), progresses into the pre-invasive stage of ductal carcinoma in

situ (DCIS), and culminates in the potentially lethal stage of invasive ductal carcinoma (IDC). This linear model of breast cancer progression has been the rationale for the use of detection methods such as mammography in the hope of diagnosing and treating breast cancer at earlier clinical stages (Ma et al., 2003).

[0006] As more molecular information is being collated, diseases such as breast cancer are being sub-divided according to genetic signatures linked to patient outcome, providing valuable information for the clinician. Emerging novel technologies in molecular medicine have already demonstrated their power in discriminating between disease subtypes that are not recognizable by traditional pathological criteria (Sorlie et al., 2001) and in identifying specific genetic events involved in cancer progression (Srinivas et al., 2002).

[0007] Endocrine therapy is a popular mode of treatment for all stages of breast cancer. A majority of breast cancers belong to the type in which growth is stimulated by the female sex hormones, estrogens and progesterone. Therefore some of the therapies are based on depriving the tumor of the hormone-induced growth stimulus. Some of the current modes of endocrine treatments include blockade of the estrogen receptor with an antiestrogen, e.g. tamoxifen; hormonal ablation by surgery (oophorectomy, adrenalectomy or hypophysectomy), radiotherapy or medically by administration of a luteinizing hormone-releasing hormone analogue (LH-RHa), e.g., goserelin; suppression of estrogen synthesis with aromatase inhibitors, e.g., anastrozole; pharmacological doses of estrogens and progestagens, e.g., megestrol acetate.

[0008] Despite recent advances, the challenge of cancer treatment, including breast cancer therapy remains. Progress is limited with respect to the development of specific treatment regimens to clinically distinct tumor types, and to personalize tumor treatment in order to maximize outcome and efficiency. Moreover, a number of patients exhibit chemotherapy resistance.

[0009] Mere classification of breast cancers into a few subgroups characterized by low to absent gene expression of the estrogen receptor (ER) alone may not reflect the cellular and molecular heterogeneity of breast cancer, and may not allow the design of treatment strategies maximizing patient response. Once a patient is diagnosed with cancer, such as breast or ovarian cancer, or an individual wants predisposition analysis, there is a strong need for methods that allow the physician to predict the expected course of disease, including the likelihood of cancer recurrence, long-term survival of the patient, and the like, and accordingly select an appropriate treatment option that is effective.

SUMMARY OF THE INVENTION

[0010] Embodiments concern methods, compositions, and apparatuses related to assessing, prognosing, and/or treating breast cancer patients. It concerns using information related to glucocorticoid receptor (GR) activity and/or expression in conjunction with information related to estrogen receptor (ER) activity or expression to identify patients with the least favorable prognosis based on current standards of care for breast cancer. Patients with relatively low levels of estrogen receptor expression and relatively high levels of glucocorticoid expression fall into a group of breast cancer patients with the least favorable prognosis (i.e., mortality rate).

[0011] Accordingly, methods concern evaluating a patient with breast cancer. Embodiments include evaluating a biological sample from a patient; evaluating breast cancer cells from a patient; evaluating a biological sample from a breast cancer patient; assessing a breast cancer patient; testing a breast cancer sample or biopsy; testing a breast tumor; prognosing a breast cancer patient; treating a breast cancer patient, particularly a patient with a particular profile related to ER and GR; determining a treatment for a breast cancer patient; altering a treatment plan for a breast cancer patient; reporting prognosis of a breast cancer patient; determining a prognosis score for a breast cancer patient; generating a prognosis score for a breast cancer patient; assessing the risk of mortality of a breast cancer patient generally or within a certain time frame, such as 150 months from end of cancer treatment; generating an ER and GR expression profile for a breast cancer patient; comparing a patient's ER and GR expression profile to a standardized profile; and/or, determining a breast cancer patient has a poor prognosis based on the patient's ER and GR status.

[0012] Embodiments also cover apparatuses, kits, and computer readable medium and systems for assessing the level or activity of ER and/or GR in a patient's breast cancer sample and determining a prognosis; and/or treating the patient accordingly. It is specifically contemplated that a breast cancer patient is a human. Accordingly, in human patients, ER refers to an estrogen receptor in a human and GR refers to a glucocorticoid receptor in a human.

[0013] Some embodiments include generating an expression profile for glucocorticoid receptor, which means obtaining the level of expression of GR directly or indirectly by measuring or assaying activity or expression. Methods include directly measuring or assaying the level of expression or activity refers to measuring or assaying a sample to determine the level of GR expression (protein or transcript) in the cell. Indirectly obtaining the level of expression includes measuring or assaying expression or activity of a gene or protein that correlates with GR expression or activity. In some embodiments, the level of GR expression can be indirectly obtained by measuring or assaying expression of a GR-responsive gene, which refers to a gene whose expression is affected in a dose-dependent manner by GR expression or activity. Expression refers to either protein expression or RNA (transcript) expression. Methods may involve either type of expression and a variety of assays are well known to those of skill in the art. For example, quantitative PCR may be performed to obtain RNA expression levels. The Affymetrix chip used in the Examples also provides information regarding RNA expression levels. Alternatively, reagents to detect protein expression levels may be employed in embodiments. Methods may involve probes, primers, and/or antibodies that are specific to GR or ER in order to assess expression levels.

[0014] In some embodiments, the activity level of GR is measured by assaying the level of GR expression. In additional embodiments, GR expression is GR transcript expression. In other embodiments, GR expression is GR protein expression. As discussed above, in some embodiments, the activity level of GR is measured by assaying the expression level of one or more GR-responsive genes. A GR-responsive gene may be one or more of the following: MCL1, SAP30, DUSP1, SGK1, SMARCA2, PTGDS, TNFRSF9, SFN, LAPTM5, GPSM2, SORT1, DPT, NR1P1, ACSL5, BIRC3, NNMT, IGF1BP6, PLXNC1, SLC46A3, C14orf139, PIAS1,

IDH2, SERPINFL, ERBB2, PECAM1, LBH, ST3GAL5, IL1R1, BIN1, WIPF1, TFPI, FN1, FAM134A, NR1P1, RAC2, SPP1, PHF15, BTN3A2, SESN1, MAP3K5, DPYSL2, SEMA4D, STOM, or MAOA.

[0015] In some embodiments, there is a step of assaying or measuring the activity level of glucocorticoid receptor (GR) in a biological sample from the patient containing breast cancer cells. As discussed above, the activity level of GR can be obtained directly or indirectly. It is specifically contemplated that levels of glucocorticoid activity or expression refers to activity or expression of GR α , GR β , or both. Unless specifically stated otherwise, the terms "glucocorticoid receptor" or "GR" refer to both forms. Embodiments discussed with respect to glucocorticoid receptor or GR may also be implemented solely with GR α or solely with GR β .

[0016] Methods may also include obtaining a level of estrogen receptor (ER) expression in breast cancer cells from the patient. The level can be obtained by obtaining the results of an assay that measured the level of ER expression. In some embodiments, the level is obtained by measuring or assaying the level of ER expression.

[0017] In some embodiments, the level of estrogen receptor expression in breast cancer cells from patient is obtained by measuring the level of estrogen receptor expression from the biological sample from the patient. In other embodiments, the level is obtained by receiving qualitative and/or quantitative data regarding the level.

[0018] In some embodiments, methods include identifying the patient as having or not having a risk factor for cancer recurrence based on the levels of ER and GR expression. Methods may involve categorizing the patient as ER+ or ER- based the level of estrogen receptor expression and a predetermined threshold value for ER expression. The term "ER+" refers to a classification of ER expression that indicates the patient expresses estrogen receptor in breast cancer cells at or above a certain level. The term "ER-" refers to a classification of ER expression that indicates the patient expresses estrogen receptor at a relatively low level in breast cancer cells, meaning at or below a certain level. In embodiments of the invention, that certain level or a predetermined threshold value is at, below, or above 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 percentile, or any range derivable therein.

[0019] Methods may involve measuring the activity level of glucocorticoid receptor in a biological sample from the patient containing breast cancer cells and measuring the expression level of estrogen receptor in the biological sample.

[0020] In certain embodiments, the predetermined threshold value for ER expression identifies a patient as ER+ if the patient's ER expression level is in the 25th percentile or greater compared to a normalized sample. This means the patient may be designated as having a level of ER expression that is at or above 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100 percentile, or any range derivable therein. It is contemplated

plated that in some cases, a patient may be designated as ER+ if the patient's ER expression level is at or above 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, or any range derivable therein. The patient may also be referred to as having a normal or high ER expression level. The higher the percentile, the higher the relative expression level.

[0021] In embodiments, methods may also involve categorizing the patient as GR+ or GR- based on a predetermined threshold value for GR activity. In some cases, a predetermined threshold value for GR activity is dependent on whether the patient is categorized as ER+ or ER-. Embodiments may involve a predetermined threshold value for GR activity that identifies a patient as GR+ if the patient is ER- and GR activity level is in the 65th percentile or greater compared to a normalized sample. It is contemplated that in some cases, a patient may be designated as GR+ if the patient's GR expression level is at or above 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, or any range derivable therein. The threshold value may or may not be dependent on ER expression levels or status. In some embodiments, the threshold value depends on whether the patient is ER- or not. The higher the percentile, the higher the relative expression level.

[0022] Methods may involve the use of a normalized sample or control that is based on one or more breast cancer samples that are not from the patient being tested.

[0023] In some embodiments, methods involve calculating a prognosis score for the patient based on the levels of ER and/or GR expression. Methods may alternatively or additionally involve reporting a prognosis score or report the levels of ER and/or GR expression. The score or report may contain or reflect raw data regarding expression levels or it may reflect a categorization of the expression levels obtained. A score could indicate the risk factor for mortality, recurrence, and/or both. The score could be a number within a numeric scale in which one end of the scale is most favorable and the other end is the least favorable with respect to a prognosis for breast cancer.

[0024] In certain embodiments, methods may involve identifying the patient as having a poor prognosis if the patient is determined to have a glucocorticoid receptor activity level at or above a certain threshold level and a level of estrogen receptor that is at or below a second threshold level. In each case, the threshold levels are specific for each of GR and ER. In certain embodiments, it is contemplated that a GR level in the 65th percentile or above based on breast cancer patients whose are in the 35th percentile or below is indicative of a poor prognosis. In some embodiments, patients with a poor prognosis include a population of breast cancer patients that numbers approximately 10% or less.

[0025] Methods also include identifying the patient as having a poor prognosis if the patient is determined to have i) an activity level of glucocorticoid receptor that is higher

than the activity level of glucocorticoid receptor in normalized control sample and ii) a expression level of estrogen receptor expression that is lower than the expression level of estrogen receptor in a normalized control sample. Consequently, methods of the invention include prognosing a breast cancer patient. In some cases, a patient is identified as having a relatively good prognosis.

[0026] Other embodiments include methods of treating a patient for breast cancer comprising: treating the patient for breast cancer after a biological sample from the patient containing breast cancer cells is analyzed for i) the activity level of glucocorticoid receptor and ii) the expression level of estrogen receptor. A patient may be treated with a different treatment protocol than the patient would have been treated with if the patient's biological sample had not been analyzed. In some embodiments, the patient is categorized as ER- and GR+ based on the activity level of the glucocorticoid receptor and the expression level of estrogen receptor. In some cases, the patient is treated with a more aggressive therapy than the patient would have been treated with if the patient had not been categorized as ER- and GR+. The term "more aggressive" refers to a treatment regimen that may include more drugs or drugs with more severe side effects and/or it may include an increased dosage or increased frequency of drugs. It may also include radiation or a combination of therapies. In some cases, the therapy includes one or more chemotherapeutics and/or biologics. In some embodiments, the patient is treated with a therapy comprising an anti-angiogenic agent. In additional embodiments, the therapy further comprises a chemotherapeutic agent in addition to the anti-angiogenic agent. Embodiments also include administering a glucocorticoid receptor antagonist and/or tyrosine kinase inhibitor.

[0027] Embodiments may also include where the patient is treated with more than one type of cancer therapy. This may be after the patient is determined to have a particular prognosis or after the status of the patient's GR and ER expression profile is known. In some embodiments, certain treatments are provided to an ER-/GR+ breast cancer patient who might have otherwise been treated with a less aggressive treatment for breast cancer. In some embodiments, a patient is treated with at least two of the following: radiation, chemotherapy, or a biologic. In particular embodiments, the patient may be treated with a kinase inhibitor and/or anti-angiogenic agent.

[0028] Methods may also involve obtaining a biological sample comprising breast cancer cells from the patient and categorizing the patient as i) GR+ or GR- based on the level of glucocorticoid activity assayed in the sample and compared to a predetermined threshold value for GR activity; and ii) ER+ or ER- based on the level of estrogen receptor expression assayed in the sample and compared to a predetermined threshold value for ER expression.

[0029] Any method may also include treating the patient for breast cancer, which may include directly administering or providing a cancer therapy. In some embodiments, a practitioner or doctor may prescribe a cancer therapy that the patient administers to herself.

[0030] To achieve these methods, a doctor, medical practitioner, or their staff may retrieve a biological sample from a patient for evaluation. The sample may be a biopsy, such as a breast tissue or tumor biopsy. The sample may be analyzed by the practitioner or their staff, or it may be sent to an outside or independent laboratory. The medical prac-

itioner may be cognizant of whether the test is providing information regarding the patient's level of GR and/or ER expression or activity, or the medical practitioner may be aware only that the test indicates directly or indirectly that the test reflects that the patient has a particular prognosis or can be given a particular prognosis score. Furthermore, the practitioner may know the patient's ER or GR status, such as ER+ or ER-, or GR+ or GR-. Alternatively, she may be aware only that the test or assay indicates the patient has a poor prognosis, or the worst prognosis.

[0031] Embodiments also concern kits to determine glucocorticoid receptor status in breast cancer cells comprising: (a) one or more reagents for determining expression levels of NR3C1 in a biological sample; and (b) an algorithm and software encoding the algorithm for calculating a risk factor index from the expression of NR3C1 in a sample and the estrogen receptor status of the breast cancer cells to determine a prognosis or a prognosis score. Kits may also include one or more reagents for determining expression levels of ESR1 in the biological sample to provide estrogen receptor status.

[0032] Other embodiments include a computer readable medium having software modules for performing a method comprising the acts of: (a) comparing glucocorticoid receptor data obtained from a patient's breast cancer sample with a reference; and (b) providing an assessment of glucocorticoid receptor status to a physician for use in determining an appropriate therapeutic regimen for a patient. In further embodiments, the computer readable medium further comprises a software module for assessing estrogen receptor status of the patient's breast cancer sample.

[0033] Computer systems are also included. In some embodiments, they have a processor, memory, external data storage, input/output mechanisms, a display, for assessing glucocorticoid receptor activity, comprising: (a) a database; (b) logic mechanisms in the computer generating for the database a GR-responsive gene expression reference; and (c) a comparing mechanism in the computer for comparing the GR-responsive gene expression reference to expression data from a patient sample using a comparison model to determine a GR gene expression profile of the sample.

[0034] Other embodiments include an internet accessible portal for providing biological information constructed and arranged to execute a computer-implemented method for providing: (a) a comparison of gene expression data of one or more GR-responsive genes in a patient sample with a calculated reporter index; and (b) providing an assessment of GR activity or expression to a physician for use in determining an appropriate therapeutic regime for a patient.

[0035] In addition to compiling, collecting and/or processing data related to GR status, methods, media and systems may also include the same embodiments with respect to data related to ER status. Such aspects may be instead of or in addition to the aspects related to GR status or data.

[0036] Embodiments also include methods of killing breast cancer cells comprising administering to a breast cancer patient an effective amount of a combination of anti-cancer compounds, wherein the anticancer compounds comprise a glucocorticoid receptor antagonist and a chemotherapeutic.

[0037] In other embodiments, there are methods for treating breast cancer in a patient comprising administering to the patient an effective amount of glucocorticoid receptor antagonist and a chemotherapeutic.

[0038] In further embodiments, methods are provided for treating chemotherapy-insensitive breast cancer cells comprising administering to a breast cancer patient an effective amount of a glucocorticoid receptor antagonist followed by chemotherapy.

[0039] Other methods include methods for treating breast cancer in a patient comprising: a) administering radiation or at least a first chemotherapeutic to the patient; b) subsequently administering an effective amount of a glucocorticoid receptor antagonist to the patient; and, c) administering radiation again or at least a second chemotherapeutic to the patient after the glucocorticoid receptor antagonist is administered to the patient.

[0040] In some embodiments, there are methods for treating breast cancer in a patient comprising: a) administering an effective amount of a glucocorticoid receptor antagonist to the patient, wherein the patient expresses detectable levels of GR prior to administration of the GR antagonist; b) then administering an effective amount of radiation or at least one chemotherapeutic.

[0041] It is contemplated that in methods described herein, breast cancer cells may undergo apoptosis following treatment set forth herein. Moreover, in some embodiments, the combination of a glucocorticoid receptor antagonist and an anticancer agent or compound induces more apoptosis than treatment with just the anticancer treatment alone. In other methods, it is specifically contemplated to exclude treatment with a synthetic glucocorticoid, such as dexamethasone.

[0042] Glucocorticoid receptor antagonists are known to those of skill in the art. It refers to a compound or substance that that does not provoke a biological response itself upon binding to the glucocorticoid receptor, but blocks or dampens agonist-mediated responses. Examples include, but are not limited to, bclometasone, betamethasone, budesonide, ciclesonide, flunisolide, fluticasone, mifepristone, mometasone, and triamcinolone. In additional embodiments, the glucocorticoid receptor antagonist has undetectable level or a lower level of activity as a progesterone receptor antagonist. In certain embodiments, the glucocorticoid receptor antagonist has greater than 10-fold, 50-fold, 100-fold, 200-fold, 300-fold, 400-fold, 500-fold, 1000-fold lower binding activity (or any range derivable therein) for another hormone receptor compared to its binding activity for glucocorticoid receptor. In specific embodiments the hormone receptor is estrogen receptor or progesterone receptor.

[0043] In some embodiments, a patient had been previously treated with an anti-cancer therapy, such as radiation, chemotherapy, or immunotherapy (or a combination or multiple therapies thereof). In certain embodiments, a first anti-cancer therapy prior to therapy with glucocorticoid receptor antagonist was last administered more than two weeks prior to the glucocorticoid receptor antagonist or its combination with a second anti-cancer therapy. In certain embodiments, this first anti-cancer therapy that does not include a glucocorticoid receptor antagonist was last administered to the breast cancer patient at least 7, 8, 9, 10, 11, 12, 13, 14 days, and/or 1, 2, 3, 4, or 5 weeks, and/or 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 months prior to treatment with a glucocorticoid receptor antagonist. Treatment methods may be applied to breast cancer or breast cancer cells that are chemo-resistant or breast cancer cells that are not chemo-sensitive. Moreover, treatment may be applied to breast

cancer or to breast cancer cells that were previously administered a first apoptosis inducing agent, but were resistant to apoptosis.

[0044] In some embodiments, the breast cancer cells are determined to be resistant to apoptosis. In additional embodiments, the breast cancer or the breast cancer cells are determined not to be chemo-sensitive or are determined to be chemo-resistant. This determination may be based on the results of a genetic test or based on information obtained from an assessment of a tumor or the breast cancer after treatment with a first anti-cancer therapy. In specific embodiments, the first anti-cancer therapy is a chemotherapeutic, Herceptin®, radiation, a combination of chemotherapeutics, or a combination of one or more chemotherapeutic agents and Herceptin®.

[0045] In additional embodiments, the breast cancer cells express a detectable level of glucocorticoid receptor or its transcript. In some embodiments, the patient is determined to have breast cancer cells that express a detectable level of glucocorticoid receptor or its transcript. This may be determined directly or indirectly.

[0046] It is contemplated that breast cancer cells may be treated with a glucocorticoid receptor antagonist regardless of estrogen receptor status. Therefore, breast cancer cells may be estrogen receptor-negative (ER-) or estrogen receptor-positive (ER+), accordingly to a standardized and industry accepted test for ER status. In certain embodiments, the breast cancer cells do not express any detectable levels of ER; in other embodiments, ER expression is detectable in the breast cancer cells.

[0047] It is contemplated that breast cancer cells may be treated with a glucocorticoid receptor antagonist depending on or regardless of progesterone receptor status. Therefore, breast cancer cells may be progesterone receptor-negative (PR-) or progesterone receptor-positive (PR+), accordingly to a standardized and industry accepted test for ER status. In certain embodiments, the breast cancer cells do not express any detectable levels of PR; in other embodiments, PR expression is detectable in the breast cancer cells.

[0048] Methods involve treating breast cancer, particularly a chemo-resistant breast cancer, with a combination of therapies that includes a glucocorticoid receptor antagonist and an anticancer therapy that induces apoptosis (together they may be referred to as a combination of anti-cancer agents or compounds), such as a chemotherapeutic. In some embodiments, the chemotherapeutic is capecitabine, carboplatin, cyclophosphamide (Cytosan), daunorubicin, docetaxel (Taxotere), doxorubicin (Adriamycin), epirubicin (Elice), fluorouracil (also called 5-fluorouracil or 5-FU), gemcitabine, iribulin, ixabepilone, methotrexate, mitomycin C, mitoxantrone, paclitaxel (Taxol), thiotepa, vincristine, or vinorelbine, or a combination of these agents. In other embodiments, therapy with a glucocorticoid receptor antagonist is combined Herceptin®, radiation, chemotherapeutic(s) and radiation, a combination of chemotherapeutics, or a combination of one or more chemotherapeutic agents and Herceptin®.

[0049] It is contemplated that in some embodiments of the combination therapy the glucocorticoid receptor antagonist is administered within 5, 10, 30, 45, 60 minutes, and/or 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 hours, and/or 1, 2, 3, 4, 5, 6, 7 days, or any combination thereof within administration of at least one or the combination of the anti-cancer agents or compounds. In

specific embodiments, the glucocorticoid receptor antagonist is administered within 2 hours, 12 hours or 24 hours of administration of an anticancer agent or compound (or a combination of such agents or compounds).

[0050] It is specifically contemplated that treatment may continue or be repeated. In some embodiments, once treated with the combination of a glucocorticoid receptor antagonist and at least one anticancer agent or compound, all or part of the treatment may be repeated alone or in combination with a different anticancer agent or compound.

[0051] In certain embodiments, the glucocorticoid receptor antagonist is administered prior to as the other agent or therapy included in the combination therapy. In certain embodiments, the glucocorticoid receptor antagonist is administered 5, 10, 30, 45, 60 minutes, and/or 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 hours, and/or 1, 2, 3, 4, 5, 6, 7 days, or any combination thereof prior to administration of at least one or the combination of the anti-cancer agents or compounds. It is specifically contemplated that in some embodiments, the glucocorticoid receptor antagonist is given prior to administration of the anticancer agent or compound but that the glucocorticoid receptor antagonist is also given concurrently with or after administration of the initial or a subsequent dose of the anticancer agent or compound. As discussed throughout, the anticancer agent or compound may be in a combination of such agents or compounds. In certain embodiments, the glucocorticoid receptor antagonist is administered up to three days prior to administering the anticancer agent or compound.

[0052] Additionally or alternatively, the glucocorticoid receptor antagonist is administered after administration of the other agent or therapy included in the combination therapy. In certain embodiments, the glucocorticoid receptor antagonist is administered 5, 10, 30, 45, 60 minutes, and/or 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 hours, and/or 1, 2, 3, 4, 5, 6, 7 days, or any combination thereof after administration of at least one or the combination of the anti-cancer agents or compounds. It is specifically contemplated that in some embodiments, the glucocorticoid receptor antagonist is given after to administration of the anticancer agent or compound; such administration may be repeated. As discussed throughout, the anticancer agent or compound may be in a combination of such agents or compounds. In certain embodiments, the glucocorticoid receptor antagonist is administered up to three days after administering the anticancer agent or compound.

[0053] In certain embodiments, the breast cancer is an unresectable breast cancer. In further embodiments, the breast cancer is inflammatory breast cancer.

[0054] It is specifically contemplated that in some methods, dexamethasone has not been administered to the patient within 24 hours of administration of the glucocorticoid receptor antagonist.

[0055] Compositions are contemplated to include a glucocorticoid receptor antagonist and any other anticancer compound discussed herein, such as Herceptin or one or more chemotherapeutic compounds. In some embodiments, the composition is in a pharmaceutically acceptable formulation.

[0056] Use of the one or more compositions may be employed based on methods described herein. Other embodiments are discussed throughout this application. Any

embodiment discussed with respect to one aspect of the invention applies to other aspects of the invention as well and vice versa. The embodiments in the Example section are understood to be embodiments that are applicable to all aspects of the technology described herein.

[0057] “Cancer prognosis” generally refers to a forecast or prediction of the probable course or outcome of the cancer. As used herein, cancer prognosis includes the forecast or prediction of any one or more of the following: duration of survival of a patient susceptible to or diagnosed with a cancer, duration of recurrence-free survival, duration of progression free survival of a patient susceptible to or diagnosed with a cancer, response rate in a group of patients susceptible to or diagnosed with a cancer, and/or duration of response in a patient or a group of patients susceptible to or diagnosed with a cancer.

[0058] In certain aspects, prognosis is an estimation of the likelihood of metastasis free survival of said patient over a predetermined period of time, e.g., over a period of 5 years.

[0059] In further aspects, prognosis is an estimation of the likelihood of death of disease of said patient over a predetermined period of time, e.g., over a period of 5 years.

[0060] The term “recurrence” refers to the detection of breast cancer in form of metastatic spread of tumor cells, local recurrence, contralateral recurrence or recurrence of breast cancer at any site of the body of the patient after breast cancer had been substantially undetectable or responsive to treatments.

[0061] As used herein, “prognostic for cancer” means providing a forecast or prediction of the probable course or outcome of the cancer. In some embodiments, “prognostic for cancer” comprises providing the forecast or prediction of (prognostic for) any one or more of the following: duration of survival of a patient susceptible to or diagnosed with a cancer, duration of recurrence-free survival, duration of progression free survival of a patient susceptible to or diagnosed with a cancer, response rate in a group of patients susceptible to or diagnosed with a cancer, and/or duration of response in a patient or a group of patients susceptible to or diagnosed with a cancer.

[0062] By “gene” is meant any polynucleotide sequence or portion thereof with a functional role in encoding or transcribing a protein or regulating other gene expression. The gene may consist of all the nucleic acids responsible for encoding a functional protein or only a portion of the nucleic acids responsible for encoding or expressing a protein. The polynucleotide sequence may contain a genetic abnormality within exons, introns, initiation or termination regions, promoter sequences, other regulatory sequences or unique adjacent regions to the gene.

[0063] As used herein, “treatment” or “therapy” is an approach for obtaining beneficial or desired clinical results. This includes: reduce the number of cancer cells; reduce the tumor size; inhibit (i.e., slow to some extent and/or stop) cancer cell infiltration into peripheral organs; inhibit (i.e., slow to some extent and/or stop) tumor metastasis; inhibit, to some extent, tumor growth; and/or relieve to some extent one or more of the symptoms associated with the disorder, shrinking the size of the tumor, decreasing symptoms resulting from the disease, increasing the quality of life of those suffering from the disease, decreasing the dose of other medications required to treat the disease, delaying the progression of the disease, and/or prolonging survival of patients.

[0064] The term “therapeutically effective amount” refers to an amount of the drug that may reduce the number of cancer cells; reduce the tumor size; inhibit (i.e., slow to some extent and preferably stop) cancer cell infiltration into peripheral organs; inhibit (i.e., slow to some extent and preferably stop) tumor metastasis; inhibit, to some extent, tumor growth; and/or relieve to some extent one or more of the symptoms associated with the disorder. To the extent the drug may prevent growth and/or kill existing cancer cells, it may be cytostatic and/or cytotoxic. For cancer therapy, efficacy in vivo can, for example, be measured by assessing the duration of survival, time to disease progression (TTP), the response rates (RR), duration of response, and/or quality of life.

[0065] The terms “overexpress”, “overexpression”, “over-expressed”, “up-regulate”, or “up-regulated” interchangeably refer to a biomarker that is transcribed or translated at a detectably greater level, usually in a cancer cell, in comparison to a non-cancer cell or cancer cell that is not associated with the worst or poorest prognosis. The term includes overexpression due to transcription, post transcriptional processing, translation, post-translational processing, cellular localization, and/or RNA and protein stability, as compared to a non-cancer cell or cancer cell that is not associated with the worst or poorest prognosis. Overexpression can be detected using conventional techniques for detecting mRNA (i.e., RT-PCR, PCR, hybridization) or proteins (i.e., ELISA, immunohistochemical techniques, mass spectroscopy). Overexpression can be 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more in comparison to a normal cell or cancer cell that is not associated with the worst or poorest prognosis. In certain instances, overexpression is 1-fold, 2-fold, 3-fold, 4-fold, 5, 6, 7, 8, 9, 10, or 15-fold or more higher levels of transcription or translation in comparison to a non-cancer cell or cancer cell that is not associated with the worst or poorest prognosis.

[0066] “Biological sample” includes sections of tissues such as biopsy and autopsy samples, and frozen sections taken for histologic purposes. Such samples include breast cancer tissues, cultured cells, e.g., primary cultures, explants, and transformed cells. A biological sample is typically obtained from a mammal, such as a primate, e.g., human.

[0067] A “biopsy” refers to the process of removing a tissue sample for diagnostic or prognostic evaluation, and to the tissue specimen itself. Any biopsy technique known in the art can be applied to the diagnostic and prognostic methods of the present invention. The biopsy technique applied will depend on the tissue type to be evaluated (e.g., breast), the size and type of the tumor, among other factors. Representative biopsy techniques include, but are not limited to, excisional biopsy, incisional biopsy, needle biopsy, and surgical biopsy. An “excisional biopsy” refers to the removal of an entire tumor mass with a small margin of normal tissue surrounding it. An “incisional biopsy” refers to the removal of a wedge of tissue that includes a cross-sectional diameter of the tumor. A diagnosis or prognosis made by endoscopy or fluoroscopy can require a “core-needle biopsy”, or a “fine-needle aspiration biopsy” which generally obtains a suspension of cells from within a target tissue. Biopsy techniques are discussed, for example, in Harrison’s Principles of Internal Medicine, 2005. Obtaining a biopsy includes both direct and indirect methods, includ-

ing obtaining the biopsy from the patient or obtaining the biopsy sample after it is removed from the patient.

[0068] The use of the word “a” or “an” when used in conjunction with the term “comprising” in the claims and/or the specification may mean “one,” but it is also consistent with the meaning of “one or more,” “at least one,” and “one or more than one.”

[0069] Throughout this application, the term “about” is used to indicate that a value includes the standard deviation of error for the device or method being employed to determine the value.

[0070] The use of the term “or” in the claims is used to mean “and/or” unless explicitly indicated to refer to alternatives only or the alternatives are mutually exclusive, although the disclosure supports a definition that refers to only alternatives and “and/or.” It is also contemplated that anything listed using the term “or” may also be specifically excluded.

[0071] As used in this specification and claim(s), the words “comprising” (and any form of comprising, such as “comprise” and “comprises”), “having” (and any form of having, such as “have” and “has”), “including” (and any form of including, such as “includes” and “include”) or “containing” (and any form of containing, such as “contains” and “contain”) are inclusive or open-ended and do not exclude additional, unrecited elements or method steps.

[0072] Other objects, features and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating specific embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

DESCRIPTION OF THE DRAWINGS

[0073] The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

[0074] FIG. 1. Primary human breast ductal epithelium, DCIS (60%) in invasive human cancers (~30-40%) exhibit significant glucocorticoid receptor expression.

[0075] FIG. 2. Unsupervised cluster analysis identifies GR target gene signature (Sig+) vs Sig- tumors (n=68 genes). A GR-regulated gene expression set from MCF10A-Myc (ER-/GR+) cells treated +/-Dex from 30 m-24 h was used to perform a two dimensional unsupervised clustering analysis on the NKI-295 early breast cancer gene expression data set (n=2034 starting genes). GR-regulated genes (n=68) that separate these tumors into two groups (GRsig+—Red and GRsig—Green) are shown in rows while each column represents a patient. Several EMT genes (e.g. Snail) and known anti-apoptotic genes are included.

[0076] FIG. 3. NR3C1 expression correlates with GR signature gene expression. The GRsig+ vs. GRsig- tumor designations correlate with higher NR3C1 vs. lower expression, respectively. For ESR1+ tumors (orange) the P<0.00001 and for ESR1- tumors (green) p=0.7 (t test). Error bars are +/-SD.

[0077] FIG. 4. RFS of GR gene expression signature. The GR signature predicts a differential prognosis for ESR1+ patients and ESR1- pts with respect to GR-signature expression. ESR1-/GR+ signature patients have the worst prognosis.

[0078] FIG. 5. Meta-analysis of NR3C1 expression and RFS.

[0079] FIG. 6. Common genes differentially expressed in ESR1- and NR3C1+/- tumors. ChIP-seq and gene expression in Dex-treated MCF10A-Myc cells.

[0080] FIG. 7A-7F. Schematic of glucocorticoid receptor (GR) isoforms.

[0081] FIG. 8. Administration of mifepristone increases MDA-MB-231 tumor susceptibility to paclitaxel treatment in vivo.

[0082] FIG. 9. Mifepristone pretreatment increases tamoxifen-resistant MCF-7 (T-R-MCF-7), but not parental MCF-7 cell susceptibility to paclitaxel in vitro.

DETAILED DESCRIPTION OF THE INVENTION

[0083] Glucocorticoid receptor (GR) activation initiates a potent cell survival signal in ER- breast cancer models. However, GR activity has not been previously examined in primary human breast cancers. Because anti-apoptotic signaling is believed to be an important determinant of breast cancer viability and relapse, the inventors contemplate that early stage primary human breast cancer demonstrates a correlation between high GR (NR3C1) and GR- mediated gene expression and cancer recurrence.

[0084] The Dutch NKI 295 data set was examined and the inventors determined that a gene expression signature of 68 GR-regulated genes (based on in vitro data) could cluster patients into different groups with differential outcome. In addition, it was found that GR-mediated gene expression correlated with NR3C1 expression levels. The inventors examined NR3C1 tumor expression in a much larger meta-dataset and again found that ER-/GR (NR3C1)+ patients did the worst. Moreover, key cell survival genes identified as GR gene targets from ChIP-seq experiments were differentially expressed.

I. Hormone Receptor Status of Breast Cancer

[0085] Intracellular receptors (IRs) form a class of structurally-related genetic regulators scientists have named “ligand dependent transcription factors” (R. M. Evans, Science, 240:889, 1988). Steroid receptors are a recognized subset of the IRs, including androgen receptor (AR), progesterone receptor (PR), estrogen receptor (ER), glucocorticoid receptor (GR), and mineralocorticoid receptor (MR). Regulation of a gene by such factors requires both the IR itself and a corresponding ligand, which has the ability to selectively bind to the IR in a way that affects gene transcription.

[0086] Naturally occurring as well as synthetic steroidal glucocorticoids (e.g., cortisol, cortisone, prednisolone, dexamethasone) have been widely used for over fifty years for the treatment of acute and chronic inflammatory and immune disorders. In particular, glucocorticoids have been prescribed for the treatment of rheumatoid arthritis, osteoarthritis, rheumatic fever, asthma, allergic rhinitis, systemic lupus erythematosus, chronic obstructive pulmonary disease, Crohn’s disease, inflammatory bowel disease, and

ulcerative colitis. However, the use of glucocorticoids is often associated with severe and sometimes irreversible side effects such as bone loss/osteoporosis, hyperglycemia, diabetes mellitus, hypertension, glaucoma, muscle atrophy, Cushing's syndrome, and psychosis.

[0087] Glucocorticoids exert their pharmacological effects by regulating gene transcription after the formation of a complex with the glucocorticoid receptor (GR). GR-glucocorticoid complex affects gene transcription by translocating to the nucleus after binding of the glucocorticoid where it acts as a dimer in binding to DNA glucocorticoid hormone response elements (GREs) in the promoter regions of particular genes. The GR-glucocorticoid/GRE complex then, in turn, activates (transactivation) or inhibits transcription of proximally located genes. Conversely, the GR-glucocorticoid complex may negatively regulate gene transcription by a process that does not involve binding to DNA. In this process, termed transrepression, following binding of the glucocorticoid, the complexed GR enters the nucleus where it acts as a monomer to directly interact (via protein-protein interaction) with other transcription factors, repressing their ability to induce gene transcription and thus protein expression.

[0088] Estrogen, mediated through the estrogen receptor (ER), plays a major role in regulating the growth and differentiation of normal breast epithelium (Pike et al. *Epidemiologic Reviews* (1993) 15(1):17-35; Henderson et al. *Cancer Res.* (1988) 48:246-253). It stimulates cell proliferation and regulates the expression of other genes, including the progesterone receptor (PgR). PgR then mediates the mitogenic effect of progesterone, further stimulating proliferation (Pike et al., 1993; Henderson et al., 1988). The molecular differences between estrogen receptor ("ER") negative and ER positive tumors are significant in light of clinical observations which indicate that the nature and biological behavior of ER positive and ER negative tumors are distinct even in the absence of hormonal therapy. For example, ER negative cancers tend to recur sooner and show a different rate of recurrence in distant organ sites compared to ER positive tumors. Clinical observations and molecular profiling data suggest that tumors not expressing both ER and PgR represent a different clinical entity in terms of chemotherapy responsiveness. (Colleoni et al., *Annals of Oncology* 11(8):1057 (2000)). Thus, ER negative and ER positive breast cancers are two distinct disease entities rather than phenotypic variations of the same disease.

[0089] Relatively increased expression of these genes in primary ER-negative human breast tumors is associated with high GR expression and with an earlier relapse in ER-negative breast cancer patients (described herein). Activation of the glucocorticoid receptor (GR) in epithelial cells has been shown to initiate an anti-apoptotic (i.e., cell survival) signaling pathway that prevents breast (Wu et al, 2004) and ovarian cancer (Melhem et al, 2009) cell death in vitro and in vivo (Pang et al, 2006). Blocking or antagonizing GR activation with a GR antagonist such as mifepristone reverses cell survival signaling pathways initiated by the GR (Moran et al., 2000). Other GR antagonists (e.g., dexamethasone oxetanone) also reverse GR-mediated cell survival and potentiate apoptosis in response to cell stressors such as growth factor withdrawal (Mikosz et al, 2001). The mechanism(s) whereby GR activation protects from cell death includes the transcriptional upregulation of genes encoding anti-apoptotic proteins such as SGK1, MKP1, MCL1, and

BIRC3. However, experiments with a glucocorticoid receptor antagonist, RU486, in conjunction with dexamethasone did not increase the number of apoptotic cells induced by paclitaxel, compared to paclitaxel alone (Wu et al., 2004).

II. Biomarkers and Evaluating Levels of Biomarkers

[0090] Biomarkers for prognosing human breast cancer patients have been identified. They include estrogen receptor (ER) in combination with the activity of the glucocorticoid receptor (GR) activity. It is contemplated that these biomarkers may be evaluated based on their gene products. In some embodiments, the gene product is the RNA transcript. In other embodiments, the gene product is the protein expressed by the RNA transcript. In still another embodiment is the evaluation of surrogate genes or gene targets of ER, GR, or ER and GR.

[0091] In certain aspects a meta-analysis of expression or activity can be performed. In statistics, a meta-analysis combines the results of several studies that address a set of related research hypotheses. This is normally done by identification of a common measure of effect size, which is modeled using a form of meta-regression. Generally, three types of models can be distinguished in the literature on meta-analysis: simple regression, fixed effects meta-regression and random effects meta-regression. Resulting overall averages when controlling for study characteristics can be considered meta-effect sizes, which are more powerful estimates of the true effect size than those derived in a single study under a given single set of assumptions and conditions. A meta-gene expression value, in this context, is to be understood as being the median of the normalized expression of a marker gene or activity. Normalization of the expression of a marker gene is preferably achieved by dividing the expression level of the individual marker gene to be normalized by the respective individual median expression of this marker genes, wherein said median expression is preferably calculated from multiple measurements of the respective gene in a sufficiently large cohort of test individuals. The test cohort preferably comprises at least 3, 10, 100, 200, 1000 individuals or more including all values and ranges thereof. Dataset-specific bias can be removed or minimized allowing multiple datasets to be combined for meta-analyses (See Sims et al. *BMC Medical Genomics* (1:42), 1-14, 2008, which is incorporated herein by reference in its entirety).

[0092] The calculation of a meta-gene expression value is performed by: (i) determining the gene expression value of at least two, preferably more genes (ii) "normalizing" the gene expression value of each individual gene by dividing the expression value with a coefficient which is approximately the median expression value of the respective gene in a representative breast cancer cohort (iii) calculating the median of the group of normalized gene expression values.

[0093] A gene shall be understood to be specifically expressed in a certain cell type if the expression level of said gene in said cell type is at least 2-fold, 5-fold, 10-fold, 100-fold, 1000-fold, or 10000-fold higher than in a reference cell type, or in a mixture of reference cell types. Reference cell types include non-cancerous breast tissue cells or a heterogeneous population of breast cancers.

[0094] In certain algorithms a suitable threshold level is first determined for a marker gene. The suitable threshold level can be determined from measurements of the marker gene expression in multiple individuals from a test cohort.

The median expression of the marker gene in said multiple expression measurements is taken as the suitable threshold value.

[0095] Comparison of multiple marker genes with a threshold level can be performed as follows:

[0096] 1. The individual marker genes are compared to their respective threshold levels.

[0097] 2. The number of marker genes, the expression level of which is above their respective threshold level, is determined.

[0098] 3. If a marker genes is expressed above its respective threshold level, then the expression level of the marker gene is taken to be "above the threshold level".

[0099] "A sufficiently large number", in this context, means preferably 30%, 50%, 80%, 90%, or 95% of the marker genes used.

[0100] In certain aspects, the determination of expression levels is on a gene chip, such as an Affymetrix™ gene chip.

[0101] In another aspect, the determination of expression levels is done by kinetic real time PCR.

[0102] In certain aspects, the methods can relate to a system for performing such methods, the system comprising (a) apparatus or device for storing data on the ER or nodal status of the patient; (b) apparatus or device for determining the expression level of at least one marker gene or activity; (c) apparatus or device for comparing the expression level of the first marker gene or activity with a predetermined first threshold value; (d) apparatus or device for determining the expression level of at least one second marker gene or activity; and (e) computing apparatus or device programmed to provide a unfavorable or poor prognosis if the data indicates a negative ER status and an increased or decreased expression level of said first marker gene or activity (e.g., GR expression or activity) with the predetermined first threshold value and, alternatively, the expression level of said second marker gene is above or below a predetermined second threshold level.

[0103] The person skilled in the art readily appreciates that an unfavorable or poor prognosis can be given if the expression level of the first marker gene with the predetermined first threshold value indicates a tumor that is likely to recur or not respond well to standard therapies.

[0104] The expression patterns can also be compared by using one or more ratios between the expression levels of different breast cancer biomarkers. Other suitable measures or indicators can also be employed for assessing the relationship or difference between different expression patterns.

[0105] The GR nucleic acid and protein sequences are provided in GenBank accession number AY436590. The ER nucleic acid and protein sequences are provided in GenBank accession number NG_008493. The content of all of these GenBank Accession numbers is specifically incorporated herein by reference as of the filing date of this application.

[0106] The following biomarkers are provided for implementation with embodiments discussed herein. All of them designate nucleic acid sequences for the particular gene identifier. Nucleic acid sequences related to these gene designation can be found in the Genbank sequence databases. Additional biomarkers include the MCL1, SAP30, DUSP1, SGK1, SMARCA2, PTGDS, TNFRSF9, SFN, LAPTM5, GPSM2, SORT1, DPT, NRPI, ACSL5, BIRC3, NNMT, IGFBP6, PLXNC1, SLC46A3, C14orf139, PIAS1, IDH2, SERPINF1, ERBB2, PECAM1, LBH, ST3GAL5, IL1R1, BIN1, WIPF1, TFPI, FN1, FAM134A, NRPI1,

RAC2, SPP1, PHF15, BTN3A2, SESN1, MAP3K5, DPYSL2, SEMA4D, STOM, and MAOA genes.

[0107] One or more of the biomarkers can be used to prognose a human patient with breast cancer. The expression pattern of these biomarkers in breast cancer cells may be used to evaluate a patient to determine whether they are likely to respond to standard chemotherapy, likely not to respond to standard chemotherapy, or likely to relapse after standard chemotherapy.

[0108] The expression levels of breast cancer biomarkers can be compared to reference expression levels using various methods. These reference levels can be determined using expression levels of a reference based on all breast cancer patients or all breast cancer patients determined to be ER+ and/or ER-. Alternatively, it can be based on an internal reference such as a gene that is expressed in all cells. In some embodiments, the reference is a gene expressed in breast cancer cells at a higher level than any biomarker. Any comparison can be performed using the fold change or the absolute difference between the expression levels to be compared. One or more breast cancer biomarkers can be used in the comparison. It is contemplated that 1, 2, 3, 4, 5, 6, 7, 8, and/or 9 biomarkers may be compared to each other and/or to a reference that is internal or external. A person of ordinary skill in the art would know how to do such comparisons.

[0109] Comparisons or results from comparisons may reveal or be expressed as x-fold increase or decrease in expression relative to a standard or relative to another biomarker or relative to the same biomarker but in a different class of prognosis. In some embodiments, patients with a poor prognosis have a relatively high level of expression (overexpression) or relatively low level of expression (underexpression) when compared to patients with a better or favorable prognosis, or vice versa.

[0110] Fold increases or decreases may be, be at least, or be at most 1-, 2-, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 11-, 12-, 13-, 14-, 15-, 16-, 17-, 18-, 19-, 20-, 25-, 30-, 35-, 40-, 45-, 50-, 55-, 60-, 65-, 70-, 75-, 80-, 85-, 90-, 95-, 100- or more, or any range derivable therein. Alternatively, differences in expression may be expressed as a percent decrease or increase, such as at least or at most 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 300, 400, 500, 600, 700, 800, 900, 1000% difference, or any range derivable therein.

[0111] Other ways to express relative expression levels are by normalized or relative numbers such as 0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, 7.0, 7.1, 7.2, 7.3, 7.4, 7.5, 7.6, 7.7, 7.8, 8.0, 8.1, 8.2, 8.3, 8.4, 8.5, 8.6, 8.7, 8.8, 8.9, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5, 9.6, 9.7, 9.8, 9.9, 10.0, or any range derivable therein.

[0112] Algorithms, such as the weighted voting programs, can be used to facilitate the evaluation of biomarker levels. In addition, other clinical evidence can be combined with the biomarker-based test to reduce the risk of false evaluations. Other cytogenetic evaluations may be considered in some embodiments of the invention.

[0113] Any biological sample from the patient that contains breast cancer cells may be used to evaluate the expression pattern of any biomarker discussed herein. In some embodiments, a biological sample from a breast tumor is used. Evaluation of the sample may involve, though it need not involve, panning (enriching) for cancer cells or isolating the cancer cells.

[0114] A. Nucleic Acids

[0115] Screening methods based on differentially expressed gene products are well known in the art. In accordance with one aspect of the present invention, the differential expression patterns of breast cancer biomarkers can be determined by measuring the levels of RNA transcripts of these genes, or genes whose expression is modulated by the these genes, in the patient's breast cancer cells. Suitable methods for this purpose include, but are not limited to, RT-PCR, Northern Blot, in situ hybridization, Southern Blot, slot-blotting, nuclease protection assay and oligonucleotide arrays.

[0116] In certain aspects, RNA isolated from breast cancer cells can be amplified to cDNA or cRNA before detection and/or quantitation. The isolated RNA can be either total RNA or mRNA. The RNA amplification can be specific or non-specific. Suitable amplification methods include, but are not limited to, reverse transcriptase PCR, isothermal amplification, ligase chain reaction, and Qbeta replicase. The amplified nucleic acid products can be detected and/or quantitated through hybridization to labeled probes. In some embodiments, detection may involve fluorescence resonance energy transfer (FRET) or some other kind of quantum dots.

[0117] Amplification primers or hybridization probes for a breast cancer biomarker can be prepared from the gene sequence or obtained through commercial sources, such as Affymatrix. In certain embodiments the gene sequence is identical or complementary to at least 8 contiguous nucleotides of the coding sequence.

[0118] Sequences suitable for making probes/primers for the detection of their corresponding breast cancer biomarkers include those that are identical or complementary to all or part of genes or SEQ ID NOs described herein. These sequences are all nucleic acid sequences of breast cancer biomarkers.

[0119] The use of a probe or primer of between 13 and 100 nucleotides, preferably between 17 and 100 nucleotides in length, or in some aspects of the invention up to 1-2 kilobases or more in length, allows the formation of a duplex molecule that is both stable and selective. Molecules having complementary sequences over contiguous stretches greater than 20 bases in length are generally preferred, to increase stability and/or selectivity of the hybrid molecules obtained. One will generally prefer to design nucleic acid molecules for hybridization having one or more complementary sequences of 20 to 30 nucleotides, or even longer where desired. Such fragments may be readily prepared, for example, by directly synthesizing the fragment by chemical means or by introducing selected sequences into recombinant vectors for recombinant production.

[0120] In one embodiment, each probe/primer comprises at least 15 nucleotides. For instance, each probe can comprise at least or at most 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 400 or more nucleotides (or any range derivable therein). They may have these lengths and have a sequence that is identical or complemen-

tary to a gene or SEQ ID NO described herein. Preferably, each probe/primer has relatively high sequence complexity and does not have any ambiguous residue (undetermined "n" residues). The probes/primers preferably can hybridize to the target gene, including its RNA transcripts, under stringent or highly stringent conditions. In some embodiments, because each of the biomarkers has more than one human sequence, it is contemplated that probes and primers may be designed for use with each on of these sequences. For example, inosine is a nucleotide frequently used in probes or primers to hybridize to more than one sequence. It is contemplated that probes or primers may have inosine or other design implementations that accommodate recognition of more than one human sequence for a particular biomarker.

[0121] For applications requiring high selectivity, one will typically desire to employ relatively high stringency conditions to form the hybrids. For example, relatively low salt and/or high temperature conditions, such as provided by about 0.02 M to about 0.10 M NaCl at temperatures of about 50° C. to about 70° C. Such high stringency conditions tolerate little, if any, mismatch between the probe or primers and the template or target strand and would be particularly suitable for isolating specific genes or for detecting specific mRNA transcripts. It is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide.

[0122] In another embodiment, the probes/primers for a gene are selected from regions which significantly diverge from the sequences of other genes. Such regions can be determined by checking the probe/primer sequences against a human genome sequence database, such as the Entrez database at the NCBI. One algorithm suitable for this purpose is the BLAST algorithm. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold. These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence to increase the cumulative alignment score. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. These parameters can be adjusted for different purposes, as appreciated by one of ordinary skill in the art.

[0123] In one embodiment, quantitative RT-PCR (such as TaqMan, ABI) is used for detecting and comparing the levels of RNA transcripts in breast cancer samples. Quantitative RT-PCR involves reverse transcription (RT) of RNA to cDNA followed by relative quantitative PCR (RT-PCR). The concentration of the target DNA in the linear portion of the PCR process is proportional to the starting concentration of the target before the PCR was begun. By determining the concentration of the PCR products of the target DNA in PCR reactions that have completed the same number of cycles and are in their linear ranges, it is possible to determine the relative concentrations of the specific target sequence in the original DNA mixture. If the DNA mixtures are cDNAs synthesized from RNAs isolated from different tissues or

cells, the relative abundances of the specific mRNA from which the target sequence was derived may be determined for the respective tissues or cells. This direct proportionality between the concentration of the PCR products and the relative mRNA abundances is true in the linear range portion of the PCR reaction. The final concentration of the target DNA in the plateau portion of the curve is determined by the availability of reagents in the reaction mix and is independent of the original concentration of target DNA. Therefore, the sampling and quantifying of the amplified PCR products preferably are carried out when the PCR reactions are in the linear portion of their curves. In addition, relative concentrations of the amplifiable cDNAs preferably are normalized to some independent standard, which may be based on either internally existing RNA species or externally introduced RNA species. The abundance of a particular mRNA species may also be determined relative to the average abundance of all mRNA species in the sample.

[0124] In one embodiment, the PCR amplification utilizes one or more internal PCR standards. The internal standard may be an abundant housekeeping gene in the cell or it can specifically be GAPDH, GUSB and β -2 microglobulin. These standards may be used to normalize expression levels so that the expression levels of different gene products can be compared directly. A person of ordinary skill in the art would know how to use an internal standard to normalize expression levels.

[0125] A problem inherent in clinical samples is that they are of variable quantity and/or quality. This problem can be overcome if the RT-PCR is performed as a relative quantitative RT-PCR with an internal standard in which the internal standard is an amplifiable cDNA fragment that is similar or larger than the target cDNA fragment and in which the abundance of the mRNA encoding the internal standard is roughly 5-100 fold higher than the mRNA encoding the target. This assay measures relative abundance, not absolute abundance of the respective mRNA species.

[0126] In another embodiment, the relative quantitative RT-PCR uses an external standard protocol. Under this protocol, the PCR products are sampled in the linear portion of their amplification curves. The number of PCR cycles that are optimal for sampling can be empirically determined for each target cDNA fragment. In addition, the reverse transcriptase products of each RNA population isolated from the various samples can be normalized for equal concentrations of amplifiable cDNAs.

[0127] Nucleic acid arrays can also be used to detect and compare the differential expression patterns of breast cancer biomarkers in breast cancer cells. The probes suitable for detecting the corresponding breast cancer biomarkers can be stably attached to known discrete regions on a solid substrate. As used herein, a probe is "stably attached" to a discrete region if the probe maintains its position relative to the discrete region during the hybridization and the subsequent washes. Construction of nucleic acid arrays is well known in the art. Suitable substrates for making polynucleotide arrays include, but are not limited to, membranes, films, plastics and quartz wafers.

[0128] A nucleic acid array of the present invention can comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250 or more different polynucleotide probes, which may hybridize to different and/or the same biomarkers. Multiple probes for the same gene can be used on a single nucleic acid array.

Probes for other disease genes can also be included in the nucleic acid array. The probe density on the array can be in any range. In some embodiments, the density may be 50, 100, 200, 300, 400, 500 or more probes/cm².

[0129] Specifically contemplated by the present inventors are chip-based nucleic acid technologies such as those described by Hacia et al. (1996) and Shoemaker et al. (1996). Briefly, these techniques involve quantitative methods for analyzing large numbers of genes rapidly and accurately. By tagging genes with oligonucleotides or using fixed probe arrays, one can employ chip technology to segregate target molecules as high density arrays and screen these molecules on the basis of hybridization (see also, Pease et al., 1994; and Fodor et al., 1991). It is contemplated that this technology may be used in conjunction with evaluating the expression level of one or more breast cancer biomarkers with respect to diagnostic, prognostic, and treatment methods of the invention.

[0130] The present invention may involve the use of arrays or data generated from an array. Data may be readily available. Moreover, an array may be prepared in order to generate data that may then be used in correlation studies.

[0131] An array generally refers to ordered macroarrays or microarrays of nucleic acid molecules (probes) that are fully or nearly complementary or identical to a plurality of mRNA molecules or cDNA molecules and that are positioned on a support material in a spatially separated organization. Macroarrays are typically sheets of nitrocellulose or nylon upon which probes have been spotted. Microarrays position the nucleic acid probes more densely such that up to 10,000 nucleic acid molecules can be fit into a region typically 1 to 4 square centimeters. Microarrays can be fabricated by spotting nucleic acid molecules, e.g., genes, oligonucleotides, etc., onto substrates or fabricating oligonucleotide sequences in situ on a substrate. Spotted or fabricated nucleic acid molecules can be applied in a high density matrix pattern of up to about 30 non-identical nucleic acid molecules per square centimeter or higher, e.g. up to about 100 or even 1000 per square centimeter. Microarrays typically use coated glass as the solid support, in contrast to the nitrocellulose-based material of filter arrays. By having an ordered array of complementing nucleic acid samples, the position of each sample can be tracked and linked to the original sample. A variety of different array devices in which a plurality of distinct nucleic acid probes are stably associated with the surface of a solid support are known to those of skill in the art. Useful substrates for arrays include nylon, glass and silicon. Such arrays may vary in a number of different ways, including average probe length, sequence or types of probes, nature of bond between the probe and the array surface, e.g. covalent or non-covalent, and the like. The labeling and screening methods of the present invention and the arrays are not limited in its utility with respect to any parameter except that the probes detect expression levels; consequently, methods and compositions may be used with a variety of different types of genes.

[0132] Representative methods and apparatus for preparing a microarray have been described, for example, in U.S. Pat. Nos. 5,143,854; 5,202,231; 5,242,974; 5,288,644; 5,324,633; 5,384,261; 5,405,783; 5,412,087; 5,424,186; 5,429,807; 5,432,049; 5,436,327; 5,445,934; 5,468,613; 5,470,710; 5,472,672; 5,492,806; 5,525,464; 5,503,980; 5,510,270; 5,525,464; 5,527,681; 5,529,756; 5,532,128; 5,545,531; 5,547,839; 5,554,501; 5,556,752; 5,561,071;

5,571,639; 5,580,726; 5,580,732; 5,593,839; 5,599,695; 5,599,672; 5,610,287; 5,624,711; 5,631,134; 5,639,603; 5,654,413; 5,658,734; 5,661,028; 5,665,547; 5,667,972; 5,695,940; 5,700,637; 5,744,305; 5,800,992; 5,807,522; 5,830,645; 5,837,196; 5,871,928; 5,847,219; 5,876,932; 5,919,626; 6,004,755; 6,087,102; 6,368,799; 6,383,749; 6,617,112; 6,638,717; 6,720,138, as well as WO 93/17126; WO 95/11995; WO 95/21265; WO 95/21944; WO 95/35505; WO 96/31622; WO 97/10365; WO 97/27317; WO 99/35505; WO 09923256; WO 09936760; WO0138580; WO 0168255; WO 03020898; WO 03040410; WO 03053586; WO 03087297; WO 03091426; WO03100012; WO 04020085; WO 04027093; EP 373 203; EP 785 280; EP 799 897 and UK 8 803 000; the disclosures of which are all herein incorporated by reference.

[0133] It is contemplated that the arrays can be high density arrays, such that they contain 100 or more different probes. It is contemplated that they may contain 1000, 16,000, 65,000, 250,000 or 1,000,000 or more different probes. The probes can be directed to targets in one or more different organisms. The oligonucleotide probes range from 5 to 50, 5 to 45, 10 to 40, or 15 to 40 nucleotides in length in some embodiments. In certain embodiments, the oligonucleotide probes are 20 to 25 nucleotides in length.

[0134] The location and sequence of each different probe sequence in the array are generally known. Moreover, the large number of different probes can occupy a relatively small area providing a high density array having a probe density of generally greater than about 60, 100, 600, 1000, 5,000, 10,000, 40,000, 100,000, or 400,000 different oligonucleotide probes per cm². The surface area of the array can be about or less than about 1, 1.6, 2, 3, 4, 5, 6, 7, 8, 9, or 10 cm².

[0135] Moreover, a person of ordinary skill in the art could readily analyze data generated using an array. Such protocols include information found in WO 9743450; WO 03023058; WO 03022421; WO 03029485; WO 03067217; WO 03066906; WO 03076928; WO 03093810; WO 03100448A1, all of which are specifically incorporated by reference.

[0136] In one embodiment, nuclease protection assays are used to quantify RNAs derived from the breast cancer samples. There are many different versions of nuclease protection assays known to those practiced in the art. The common characteristic that these nuclease protection assays have is that they involve hybridization of an antisense nucleic acid with the RNA to be quantified. The resulting hybrid double-stranded molecule is then digested with a nuclease that digests single-stranded nucleic acids more efficiently than double-stranded molecules. The amount of antisense nucleic acid that survives digestion is a measure of the amount of the target RNA species to be quantified. An example of a nuclease protection assay that is commercially available is the RNase protection assay manufactured by Ambion, Inc. (Austin, Tex.).

[0137] B. Proteins and Polypeptides

[0138] In other embodiments, the differential expression patterns of breast cancer biomarkers can be determined by measuring the levels of polypeptides encoded by these genes in breast cancer cells. Methods suitable for this purpose include, but are not limited to, immunoassays such as ELISA, RIA, FACS, dot blot, Western Blot, immunohistochemistry, and antibody-based radioimaging. Protocols for carrying out these immunoassays are well known in the art.

Other methods such as 2-dimensional SDS-polyacrylamide gel electrophoresis can also be used. These procedures may be used to recognize any of the polypeptides encoded by the breast cancer biomarker genes described herein.

[0139] One example of a method suitable for detecting the levels of target proteins in peripheral blood samples is ELISA. In an exemplifying ELISA, antibodies capable of binding to the target proteins encoded by one or more breast cancer biomarker genes are immobilized onto a selected surface exhibiting protein affinity, such as wells in a polystyrene or polyvinylchloride microtiter plate. Then, breast cancer cell samples to be tested are added to the wells. After binding and washing to remove non-specifically bound immunocomplexes, the bound antigen(s) can be detected. Detection can be achieved by the addition of a second antibody which is specific for the target proteins and is linked to a detectable label. Detection may also be achieved by the addition of a second antibody, followed by the addition of a third antibody that has binding affinity for the second antibody, with the third antibody being linked to a detectable label. Before being added to the microtiter plate, cells in the peripheral blood samples can be lysed using various methods known in the art. Proper extraction procedures can be used to separate the target proteins from potentially interfering substances.

[0140] In another ELISA embodiment, the breast cancer cell samples containing the target proteins are immobilized onto the well surface and then contacted with the antibodies of the invention. After binding and washing to remove non-specifically bound immunocomplexes, the bound antigen is detected. Where the initial antibodies are linked to a detectable label, the immunocomplexes can be detected directly. The immunocomplexes can also be detected using a second antibody that has binding affinity for the first antibody, with the second antibody being linked to a detectable label.

[0141] Another typical ELISA involves the use of antibody competition in the detection. In this ELISA, the target proteins are immobilized on the well surface. The labeled antibodies are added to the well, allowed to bind to the target proteins, and detected by means of their labels. The amount of the target proteins in an unknown sample is then determined by mixing the sample with the labeled antibodies before or during incubation with coated wells. The presence of the target proteins in the unknown sample acts to reduce the amount of antibody available for binding to the well and thus reduces the ultimate signal.

[0142] Different ELISA formats can have certain features in common, such as coating, incubating or binding, washing to remove non-specifically bound species, and detecting the bound immunocomplexes. For instance, in coating a plate with either antigen or antibody, the wells of the plate can be incubated with a solution of the antigen or antibody, either overnight or for a specified period of hours. The wells of the plate are then washed to remove incompletely adsorbed material. Any remaining available surfaces of the wells are then "coated" with a nonspecific protein that is antigenically neutral with regard to the test samples. Examples of these nonspecific proteins include bovine serum albumin (BSA), casein and solutions of milk powder. The coating allows for blocking of nonspecific adsorption sites on the immobilizing surface and thus reduces the background caused by nonspecific binding of antisera onto the surface.

[0143] In ELISAs, a secondary or tertiary detection means can also be used. After binding of a protein or antibody to the well, coating with a non-reactive material to reduce background, and washing to remove unbound material, the immobilizing surface is contacted with the control and/or clinical or biological sample to be tested under conditions effective to allow immunocomplex (antigen/antibody) formation. These conditions may include, for example, diluting the antigens and antibodies with solutions such as BSA, bovine gamma globulin (BGG) and phosphate buffered saline (PBS)/Tween and incubating the antibodies and antigens at room temperature for about 1 to 4 hours or at 49° C. overnight. Detection of the immunocomplex then requires a labeled secondary binding ligand or antibody, or a secondary binding ligand or antibody in conjunction with a labeled tertiary antibody or third binding ligand.

[0144] After all of the incubation steps in an ELISA, the contacted surface can be washed so as to remove non-complexed material. For instance, the surface may be washed with a solution such as PBS/Tween, or borate buffer. Following the formation of specific immunocomplexes between the test sample and the originally bound material, and subsequent washing, the occurrence of the amount of immunocomplexes can be determined.

[0145] To provide a detecting means, the second or third antibody can have an associated label to allow detection. In one embodiment, the label is an enzyme that generates color development upon incubating with an appropriate chromogenic substrate. Thus, for example, one may contact and incubate the first or second immunocomplex with a urease, glucose oxidase, alkaline phosphatase or hydrogen peroxidase-conjugated antibody for a period of time and under conditions that favor the development of further immunocomplex formation (e.g., incubation for 2 hours at room temperature in a PBS-containing solution such as PBS-Tween).

[0146] After incubation with the labeled antibody, and subsequent to washing to remove unbound material, the amount of label is quantified, e.g., by incubation with a chromogenic substrate such as urea and bromocresol purple or 2,2'-azido-di-(3-cethyl)-benziazolinc-6-sulfonic acid (ABTS) and hydrogen peroxide, in the case of peroxidase as the enzyme label. Quantitation can be achieved by measuring the degree of color generation, e.g., using a spectrophotometer.

[0147] Another suitable method is MA (radioimmunoassay). An example of MA is based on the competition between radiolabeled-polypeptides and unlabeled polypeptides for binding to a limited quantity of antibodies. Suitable radiolabels include, but are not limited to, I¹²⁵. In one embodiment, a fixed concentration of I¹²⁵-labeled polypeptide is incubated with a series of dilution of an antibody specific to the polypeptide. When the unlabeled polypeptide is added to the system, the amount of the I¹²⁵-polypeptide that binds to the antibody is decreased. A standard curve can therefore be constructed to represent the amount of antibody-bound I¹²⁵-polypeptide as a function of the concentration of the unlabeled polypeptide. From this standard curve, the concentration of the polypeptide in unknown samples can be determined. Various protocols for conducting RIA to measure the levels of polypeptides in breast cancer cell samples are well known in the art.

[0148] Suitable antibodies for this invention include, but are not limited to, polyclonal antibodies, monoclonal anti-

bodies, chimeric antibodies, humanized antibodies, single chain antibodies, Fab fragments, and fragments produced by a Fab expression library.

[0149] Antibodies can be labeled with one or more detectable moieties to allow for detection of antibody-antigen complexes. The detectable moieties can include compositions detectable by spectroscopic, enzymatic, photochemical, biochemical, bioelectronic, immunochemical, electrical, optical or chemical means. The detectable moieties include, but are not limited to, radioisotopes, chemiluminescent compounds, labeled binding proteins, heavy metal atoms, spectroscopic markers such as fluorescent markers and dyes, magnetic labels, linked enzymes, mass spectrometry tags, spin labels, electron transfer donors and acceptors, and the like.

[0150] Protein array technology is discussed in detail in Pandey and Mann (2000) and MacBeath and Schreiber (2000), each of which is herein specifically incorporated by reference. These arrays typically contain thousands of different proteins or antibodies spotted onto glass slides or immobilized in tiny wells and allow one to examine the biochemical activities and binding profiles of a large number of proteins at once. To examine protein interactions with such an array, a labeled protein is incubated with each of the target proteins immobilized on the slide, and then one determines which of the many proteins the labeled molecule binds. In certain embodiments such technology can be used to quantitate a number of proteins in a sample, such as a breast cancer biomarker proteins.

[0151] The basic construction of protein chips has some similarities to DNA chips, such as the use of a glass or plastic surface dotted with an array of molecules. These molecules can be DNA or antibodies that are designed to capture proteins. Defined quantities of proteins are immobilized on each spot, while retaining some activity of the protein. With fluorescent markers or other methods of detection revealing the spots that have captured these proteins, protein microarrays are being used as powerful tools in high-throughput proteomics and drug discovery.

[0152] The earliest and best-known protein chip is the ProteinChip by CIPHERGEN Biosystems Inc. (Fremont, Calif.). The ProteinChip is based on the surface-enhanced laser desorption and ionization (SELDI) process. Known proteins are analyzed using functional assays that are on the chip. For example, chip surfaces can contain enzymes, receptor proteins, or antibodies that enable researchers to conduct protein-protein interaction studies, ligand binding studies, or immunoassays. With state-of-the-art ion optic and laser optic technologies, the ProteinChip system detects proteins ranging from small peptides of less than 1000 Da up to proteins of 300 kDa and calculates the mass based on time-of-flight (TOF).

[0153] The ProteinChip biomarker system is the first protein biochip-based system that enables biomarker pattern recognition analysis to be done. This system allows researchers to address important clinical questions by investigating the proteome from a range of crude clinical samples (i.e., laser capture microdissected cells, biopsies, tissue, urine, and serum). The system also utilizes biomarker pattern software that automates pattern recognition-based statistical analysis methods to correlate protein expression patterns from clinical samples with disease phenotypes.

[0154] In other aspects, the levels of polypeptides in samples can be determined by detecting the biological

activities associated with the polypeptides. If a biological function/activity of a polypeptide is known, suitable *in vitro* bioassays can be designed to evaluate the biological function/activity, thereby determining the amount of the polypeptide in the sample.

III. Breast Cancer Therapy

[0155] Certain embodiments are directed to methods of treating breast cancer based on GR status of the breast cancer tissue. In some embodiments, the hormone receptor status is determined based on the expression of a hormone receptor such as the estrogen receptor (ER) in combination with the glucocorticoid receptor (GR).

[0156] In certain aspects, the hormone receptor status is high for GR and may also be low for one or more other hormone receptors such as the estrogen receptor. An individual having an elevated GR and low ER is likely to have a poor prognosis. In the event of a poor prognosis the physician may pursue a more aggressive therapy for those patients. In some embodiments, the method comprises identifying a breast cancer patient based on a hormone receptor status of patients having tumor tissue with elevated levels of GR expression.

[0157] In certain aspects, there may be provided methods for treating a subject determined to have cancer and with a predetermined expression profile of one or more biomarkers disclosed herein.

[0158] In a further aspect, biomarkers and related systems that can establish a prognosis of cancer patients in this invention can be used to identify patients who may get benefit of conventional single or combined modality therapy. In the same way, the invention can identify those patients who do not get much benefit from such conventional single or combined modality therapy and can offer them alternative treatment(s).

[0159] In certain aspects of the present invention, conventional cancer therapy may be applied to a subject wherein the subject is identified or reported as having a good prognosis based on the assessment of the biomarkers as disclosed. On the other hand, at least an alternative cancer therapy may be prescribed, as used alone or in combination with conventional cancer therapy, if a poor prognosis is determined by the disclosed methods, systems, or kits.

[0160] Embodiments concern a glucocorticoid receptor antagonist. In some embodiments, the glucocorticoid receptor antagonist is a selective glucocorticoid receptor antagonist, as set forth in Clark, 2008, which is hereby incorporated by reference. In other embodiments, the glucocorticoid receptor antagonist is a non-selective glucocorticoid receptor antagonist, such as mifepristone. In certain embodiments, the glucocorticoid receptor antagonist is steroidal. In other embodiments, the glucocorticoid receptor antagonist is nonsteroidal. A glucocorticoid receptor antagonist includes those in the following classes of chemical compounds: octahydrophenanthrenes, spirocyclic dihydropyridines, triphenylmethanes and diaryl ethers, chromenes, dibenzyl anilines, dihydroisoquinolines, pyrimidinediones, azadecalins, and aryl pyrazolo azadecalins, and which are described in more detail in Clark, 2008, which is hereby incorporated by reference. Some embodiments of steroidal antagonists from Clark, 2008 are: RU-486, RU-43044, 11-monoaryl and 11,21 bisaryl steroids (including 11 β -substituted steroids), 10 β -substituted steroids, 11 β -aryl conjugates of mifepristone, and phosphorous-containing mifepristone analogs.

Further embodiments of nonsteroidal antagonists from Clark, 2008 are: octahydrophenanthrenes, spirocyclic dihydropyridines, triphenylmethanes and diaryl ethers, chromenes, dibenzyl anilines, dihydroquinolines, pyrimidinediones, azadecalins, aryl pyrazolo azadecalins (including 8 α -benzyl isoquinolones, N-substituted derivatives, bridgehead alcohol and ethers, bridgehead amines). Additional specific examples include, but are not limited to the following specific antagonists: beclomethasone, betamethasone, budesonide, ciclesonide, flunisolide, fluticasone, mifepristone, mometasone, and triamcinolone. Other examples include those described and/or depicted in U.S. Patent Application Publication 2010/0135956, which is hereby incorporated by reference. Even further examples include ORG-34517 (Merck), RU-43044, dexamethasone mesylate (Dex-Mes), dexamethasone oxetanone (Dex-Ox), deoxycorticosterone (DOC) (Peeters et al., 2008, which is hereby incorporated by reference in its entirety and Cho et al. 2005, which is hereby incorporated by reference in its entirety). In additional embodiments the glucocorticoid receptor antagonist may be CORT 0113083 or CORT 00112716, which are described in Belanoff et al. (2011), which is hereby incorporated by reference. It is specifically contemplated that one or more of the antagonists discussed herein or in the incorporated references may be excluded in embodiments of the invention. It is also contemplated that in some embodiments, more than one glucocorticoid receptor antagonist is employed, while in other embodiments, only one is employed as part of the therapeutic method (though it may be administered multiple times). It is contemplated that the second one may be administered concurrently with the first one or they may be administered at different times.

[0161] Conventional cancer therapies include one or more selected from the group of chemical or radiation based treatments and surgery. Chemotherapies include, for example, cisplatin (CDDP), carboplatin, procarbazine, mechlorethamine, cyclophosphamide, camptothecin, ifosfamide, melphalan, chlorambucil, busulfan, nitrosurea, dacarbazine, daunorubicin, doxorubicin, bleomycin, plicomycin, mitomycin, etoposide (VP16), tamoxifen, raloxifene, estrogen receptor binding agents, taxol, gemcitabine, navelbine, farnesyl-protein transferase inhibitors, transplatin, 5-fluorouracil, vincristin, vinblastin and methotrexate, or any analog or derivative variant of the foregoing.

[0162] Suitable therapeutic agents include, for example, *vinca* alkaloids, agents that disrupt microtubule formation (such as colchicines and its derivatives), anti-angiogenic agents, therapeutic antibodies, EGFR targeting agents, tyrosine kinase targeting agent (such as tyrosine kinase inhibitors), serine kinase targeting agents, transitional metal complexes, proteasome inhibitors, antimetabolites (such as nucleoside analogs), alkylating agents, platinum-based agents, anthracycline antibiotics, topoisomerase inhibitors, macrolides, therapeutic antibodies, retinoids (such as all-trans retinoic acids or a derivatives thereof); geldanamycin or a derivative thereof (such as 17-AAG), and other standard chemotherapeutic agents well recognized in the art.

[0163] Certain chemotherapeutics are well known for use against breast cancer. These breast cancer chemotherapeutics are capecitabine, carboplatin, cyclophosphamide (Cytosan), daunorubicin, docetaxel (Taxotere), doxorubicin (Adriamycin), epirubicin (Ellence), fluorouracil (also called 5-fluorouracil or 5-FU), gemcitabine, eribulin, ixabepilone,

methotrexate, mitomycin C, mitoxantrone, paclitaxel (Taxol), thiotepa, vincristine, vinorelbine.

[0164] In some embodiments, the chemotherapeutic agent is any of (and in some embodiments selected from the group consisting of) adriamycin, colchicine, cyclophosphamide, actinomycin, bleomycin, daunorubicin, doxorubicin, epirubicin, mitomycin, methotrexate, mitoxantrone, fluorouracil, carboplatin, carmustine (BCNU), methyl-CCNU, cisplatin, etoposide, interferons, camptothecin and derivatives thereof, phenesterine, taxanes and derivatives thereof (e.g., paclitaxel and derivatives thereof, taxotere and derivatives thereof, and the like), topotecan, vinblastine, vincristine, tamoxifen, pposulfan, nab-5404, nab-5800, nab-5801, Irinotecan, HKP, Ortataxel, gemcitabine, Herceptin®, vinorelbine, capecitabine, Gleevec®, Alimta®, Avastin®, Velcade®, Tarceva®, Neulasta®, Lapatinib, STI-571, ZD1839, Iressa® (gefitinib), SH268, genistein, CEP2563, SU6668, SU11248, EMD121974, and Sorafenib.

[0165] In some embodiments, the chemotherapeutic agent is a composition comprising nanoparticles comprising a thiocolchicine derivative and a carrier protein (such as albumin).

[0166] In further embodiments a combination of chemotherapeutic agents is administered to breast cancer cells. The chemotherapeutic agents may be administered serially (within minutes, hours, or days of each other) or in parallel; they also may be administered to the patient in a pre-mixed single composition. The composition may or may not contain a glucocorticoid receptor antagonist. Combinations of breast cancer therapeutics include, but are not limited to the following: AT (Adriamycin and Taxotere), AC±T: (Adriamycin and Cytoxan, with or without Taxol or Taxotere), CMF (Cytoxan, methotrexate, and fluorouracil), CEF (Cytoxan, Ellence, and fluorouracil), FAC (fluorouracil, Adriamycin, and Cytoxan), CAF (Cytoxan, Adriamycin, and fluorouracil) (the FAC and CAF regimens use the same medicines but use different doses and frequencies), TAC (Taxotere, Adriamycin, and Cytoxan), and GET (Gemzar, Ellence, and Taxol). In some embodiments trastuzumab (Herceptin®) is administered to a breast cancer patient with a glucocorticoid receptor antagonist, which may be with or without a chemotherapeutic or a combination of chemotherapeutics.

[0167] Various combinations with a glucocorticoid receptor antagonist and an anticancer agent or compound (or a combination of such agents and/or compounds) may be employed, for example glucocorticoid receptor antagonist is “A” and the anticancer agent or compound (or a combination of such agents and/or compounds) given as part of an anticancer therapy regime, is

A/B/A B/A/B B/B/A A/A/B A/B/B B/A/A A/B/B/B B/A/B/B
 B/B/B/A B/B/A/B A/A/B/B A/B/A/B A/B/B/A B/B/A/A
 B/A/B/A B/A/A/B A/A/A/B B/A/A/A A/B/A/A A/A/B/A

[0168] Administration of the therapeutic compounds or agents to a patient will follow general protocols for the administration of such compounds, taking into account the toxicity, if any, of the therapy. It is expected that the treatment cycles would be repeated as necessary. It also is contemplated that various standard therapies, as well as surgical intervention, may be applied in combination with the described therapy.

[0169] The term “a serine/threonine kinase inhibitor”, as used herein, relates to a compound which inhibits serine/threonine kinases. An example of a target of a serine/threonine kinase inhibitor includes, but is not limited to, dsRNA-dependent protein kinase (PKR). Examples of indirect targets of a serine/threonine kinase inhibitor include, but are not limited to, MCP-1, NF-kappaB, eIF2alpha, COX2, RANTES, IL8, CYP2A5, IGF-1, CYP2B1, CYP2B2, CYP2H1, ALAS-1, HIF-1, erythropoietin and/or CYP1A1. An example of a serine/threonine kinase inhibitor includes, but is not limited to, Sorafenib and 2-aminopurine, also known as 1H-purin-2-amine(9CI). Sorafenib is marketed as NEXAVAR.

[0170] The term “an angiogenesis inhibitor”, as used herein, relates to a compound which targets, decreases or inhibits the production of new blood vessels. Targets of an angiogenesis inhibitor include, but are not limited to, methionine aminopeptidase-2 (MetAP-2), macrophage inflammatory protein-1 (MIP-1a), CCL5, TGF-beta., lipoxigenase, cyclooxygenase, and topoisomerase. Indirect targets of an angiogenesis inhibitor include, but are not limited to, p21, p53, CDK2 and collagen synthesis. Examples of an angiogenesis inhibitor include, but are not limited to, Fumagillin, which is known as 2,4,6,8-decatetraenedioic acid, mono[3R,4S,5S,6R]-5-methoxy-4-[(2R,3R)-2-methyl-3-(3-methyl-2-butenyl)oxi-ranyl]-1-oxaspiro[2.5]oct-6-yl] ester, (2E,4E,6E,8E)-(9CI); Shikonin, which is also known as 1,4-naphthalenedione, 5,8-dihydroxy-2-[(1R)-1-hydroxy-4-methyl-3-pentenyl]-(9CI); Tranilast, which is also known as benzoic acid, 2-[[3-(3,4-dimethoxyphenyl)-1-oxo-2-propenyl]amino]-(9CI); ursolic acid; suramin; thalidomide and lenalidomide, and marketed as REVLIMID.

[0171] Radiation therapy that cause DNA damage and have been used extensively include what are commonly known as gamma-rays, X-rays, and/or the directed delivery of radioisotopes to tumor cells. Other forms of DNA damaging factors are also contemplated such as microwaves and UV-irradiation. It is most likely that all of these factors effect a broad range of damage on DNA, on the precursors of DNA, on the replication and repair of DNA, and on the assembly and maintenance of chromosomes. Dosage ranges for X-rays range from daily doses of 50 to 200 roentgens for prolonged periods of time (3 to 4 wk), to single doses of 2000 to 6000 roentgens. Dosage ranges for radioisotopes vary widely, and depend on the half-life of the isotope, the strength and type of radiation emitted, and the uptake by the neoplastic cells.

[0172] The terms “contacted” and “exposed,” when applied to a cell, are used herein to describe the process by which a therapeutic construct and a chemotherapeutic or radiotherapeutic agent are delivered to a target cell or are placed in direct juxtaposition with the target cell. To achieve cell killing or stasis, both agents are delivered to a cell in a combined amount effective to kill the cell or prevent it from dividing.

[0173] Approximately 60% of persons with cancer will undergo surgery of some type, which includes preventative, diagnostic or staging, curative and palliative surgery. Curative surgery is a cancer treatment that may be used in conjunction with other therapies, such as the treatment of the present invention, chemotherapy, radiotherapy, hormonal therapy, gene therapy, immunotherapy and/or alternative therapies.

[0174] Curative surgery includes resection in which all or part of cancerous tissue is physically removed, excised, and/or destroyed. Tumor resection refers to physical removal of at least part of a tumor. In addition to tumor resection, treatment by surgery includes laser surgery, cryosurgery, electrosurgery, and microscopically controlled surgery (Mohs' surgery). It is further contemplated that the present invention may be used in conjunction with removal of superficial cancers, precancers, or incidental amounts of normal tissue.

[0175] Laser therapy is the use of high-intensity light to destroy tumor cells. Laser therapy affects the cells only in the treated area. Laser therapy may be used to destroy cancerous tissue and relieve a blockage in the esophagus when the cancer cannot be removed by surgery. The relief of a blockage can help to reduce symptoms, especially swallowing problems.

[0176] Photodynamic therapy (PDT), a type of laser therapy, involves the use of drugs that are absorbed by cancer cells; when exposed to a special light, the drugs become active and destroy the cancer cells. PDT may be used to relieve symptoms of esophageal cancer such as difficulty swallowing.

[0177] Upon excision of part of all of cancerous cells, tissue, or tumor, a cavity may be formed in the body. Treatment may be accomplished by perfusion, direct injection or local application of the area with an additional anti-cancer therapy. Such treatment may be repeated, for example, every 1, 2, 3, 4, 5, 6, or 7 days, or every 1, 2, 3, 4, and 5 weeks or every 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 months. These treatments may be of varying dosages as well. A patient may be administered a single compound or a combination of compounds described herein in an amount that is, is at least, or is at most 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 mg/kg (or any range derivable therein). A patient may be administered a single compound or a combination of compounds described herein in an amount that is, is at least, or is at most 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 441, 450, 460, 470, 480, 490, 500 mg/kg/day (or any range derivable therein).

[0178] Alternative cancer therapy include any cancer therapy other than surgery, chemotherapy and radiation therapy in the present invention, such as immunotherapy, gene therapy, hormonal therapy or a combination thereof. Subjects identified with poor prognosis using the present methods may not have favorable response to conventional treatment(s) alone and may be prescribed or administered one or more alternative cancer therapy per se or in combination with one or more conventional treatments.

[0179] For example, the alternative cancer therapy may be a targeted therapy. The targeted therapy may be an anti-EGFR treatment. In one embodiment of the method of the invention, the anti-EGFR agent used is a tyrosine kinase inhibitor. Examples of suitable tyrosine kinase inhibitors are the quinazoline derivatives described in WO 96/33980, in particular gefitinib (Iressa). Other examples include quinazoline derivatives described in WO 96/30347, in particular erlotinib (Tarceva), dual EGFR/HER2 tyrosine kinase inhibitors, such as lapatinib, or pan-Erb inhibitors. In a preferred embodiment of the method or use of the invention, the anti-EGFR agent is an antibody capable of binding to EGFR, i.e. an anti-EGFR antibody.

[0180] In a further embodiment, the anti-EGFR antibody is an intact antibody, i.e. a full-length antibody rather than a fragment. An anti-EGFR antibody used in the method of the present invention may have any suitable affinity and/or avidity for one or more epitopes contained at least partially in EGFR. Preferably, the antibody used binds to human EGFR with an equilibrium dissociation constant (K_D) of 10^{-8} M or less, more preferably 10^{-10} M or less.

[0181] Particularly antibodies for use in the present invention include zalutumumab (2F8), cetuximab (Erbix), nimotuzumab (h-R3), panitumumab (ABX-EGF), and matuzumab (EMD72000), or a variant antibody of any of these, or an antibody which is able to compete with any of these, such as an antibody recognizing the same epitope as any of these. Competition may be determined by any suitable technique. In one embodiment, competition is determined by an ELISA assay. Often competition is marked by a significantly greater relative inhibition than 5% as determined by ELISA analysis.

[0182] Immunotherapeutics, generally, rely on the use of immune effector cells and molecules to target and destroy cancer cells. The immune effector may be, for example, an antibody specific for some marker on the surface of a tumor cell. The antibody alone may serve as an effector of therapy or it may recruit other cells to actually effect cell killing. The antibody also may be conjugated to a drug or toxin (chemotherapeutic, radionuclide, ricin A chain, cholera toxin, pertussis toxin, etc.) and serve merely as a targeting agent. Alternatively, the effector may be a lymphocyte carrying a surface molecule that interacts, either directly or indirectly, with a tumor cell target. Various effector cells include cytotoxic T cells and NK cells.

[0183] Gene therapy is the insertion of polynucleotides, including DNA or RNA, into an individual's cells and tissues to treat a disease. Antisense therapy is also a form of gene therapy in the present invention. A therapeutic polynucleotide may be administered before, after, or at the same time of a first cancer therapy. Delivery of a vector encoding a variety of proteins is encompassed within the invention. For example, cellular expression of the exogenous tumor suppressor oncogenes would exert their function to inhibit excessive cellular proliferation, such as p53, p16 and C-CAM.

[0184] Additional agents to be used to improve the therapeutic efficacy of treatment include immunomodulatory agents, agents that affect the upregulation of cell surface receptors and GAP junctions, cytostatic and differentiation agents, inhibitors of cell adhesion, or agents that increase the sensitivity of the hyperproliferative cells to apoptotic inducers. Immunomodulatory agents include tumor necrosis factor; interferon alpha, beta, and gamma; IL-2 and other

cytokines; F42K and other cytokine analogs; or MIP-1, MIP-1beta, MCP-1, RANTES, and other chemokines. It is further contemplated that the upregulation of cell surface receptors or their ligands such as Fas/Fas ligand, DR4 or DR5/TRAIL would potentiate the apoptotic inducing abilities of the present invention by establishment of an autocrine or paracrine effect on hyperproliferative cells. Increases in intercellular signaling by elevating the number of GAP junctions would increase the anti-hyperproliferative effects on the neighboring hyperproliferative cell population. In other embodiments, cytostatic or differentiation agents can be used in combination with the present invention to improve the anti-hyperproliferative efficacy of the treatments. Inhibitors of cell adhesion are contemplated to improve the efficacy of the present invention. Examples of cell adhesion inhibitors are focal adhesion kinase (FAKs) inhibitors and Lovastatin. It is further contemplated that other agents that increase the sensitivity of a hyperproliferative cell to apoptosis, such as the antibody c225, could be used in combination with the present invention to improve the treatment efficacy.

[0185] Hormonal therapy may also be used in the present invention or in combination with any other cancer therapy previously described. The use of hormones may be employed in the treatment of certain cancers such as breast, prostate, ovarian, or cervical cancer to lower the level or block the effects of certain hormones such as testosterone or estrogen. This treatment is often used in combination with at least one other cancer therapy as a treatment option or to reduce the risk of metastases.

II. Kits

[0186] Certain aspects of the present invention also encompass kits for performing the diagnostic and prognostic methods of the invention. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise any one or more of the following materials: enzymes, reaction tubes, buffers, detergent, primers, probes, antibodies. In a preferred embodiment, these kits allow a practitioner to obtain samples of neoplastic cells in blood, tears, semen, saliva, urine, tissue, serum, stool, sputum, cerebrospinal fluid and supernatant from cell lysate. In another preferred embodiment these kits include the needed apparatus for performing RNA extraction, RT-PCR, and gel electrophoresis. Instructions for performing the assays can also be included in the kits.

[0187] In a particular aspect, these kits may comprise a plurality of agents for assessing the differential expression of a plurality of biomarkers, for example, GR and/or ER, wherein the kit is housed in a container. The kits may further comprise instructions for using the kit for assessing expression, means for converting the expression data into expression values and/or means for analyzing the expression values to generate prognosis. The agents in the kit for measuring biomarker expression may comprise a plurality of PCR probes and/or primers for qRT-PCR and/or a plurality of antibody or fragments thereof for assessing expression of the biomarkers. In another embodiment, the agents in the kit for measuring biomarker expression may comprise an array of polynucleotides complementary to the mRNAs of the biomarkers of the invention. Possible means for converting the expression data into expression values and for analyzing the expression values to generate scores that predict survival or prognosis may be also included.

[0188] Kits may comprise a container with a label. Suitable containers include, for example, bottles, vials, and test tubes. The containers may be formed from a variety of materials such as glass or plastic. The container may hold a composition which includes a probe that is useful for prognostic or non-prognostic applications, such as described above. The label on the container may indicate that the composition is used for a specific prognostic or non-prognostic application, and may also indicate directions for either in vivo or in vitro use, such as those described above. The kit of the invention will typically comprise the container described above and one or more other containers comprising materials desirable from a commercial and user standpoint, including buffers, diluents, filters, needles, syringes, and package inserts with instructions for use.

EXAMPLES

[0189] The following examples are given for the purpose of illustrating various embodiments of the invention and are not meant to limit the present invention in any fashion. One skilled in the art will appreciate readily that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those objects, ends and advantages inherent herein. The present examples, along with the methods described herein are presently representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the invention. Changes therein and other uses which are encompassed within the spirit of the invention as defined by the scope of the claims will occur to those skilled in the art.

Example 1

Tumor Biomarker Status

[0190] A. Results

[0191] The glucocorticoid receptor (GR) is highly expressed in the myoepithelium of the normal human breast and in a subset of both ERalpha-positive and negative human breast cancers. In vitro and in vivo experiments suggest that activation of the GR in ER- pre-malignant breast epithelial and cancer cells triggers cell survival pathways under stress conditions (e.g. chemotherapy) that usually induce apoptosis. The inventors examined the association between NR3C1 gene expression and GR target gene expression in human ER- breast cancers and found that ER- breast cancers with high NR3C1 expression also express GR target genes associated with EMT and anti-apoptotic signaling, and that those ER- patients with high NR3C1 gene expression have a significantly worse outcome than NR3C1-low patients. Interestingly, the high NR3C1 gene expression in the ER+(ESR1-high) subset of patients suggests a slight better outcome, implying a crosstalk between the ER and the GR that is absent in ER- tumors.

[0192] Using a global approach of gene expression studies merged with data from GR ChIP-sequencing in ER- pre-malignant breast cells (MCF10A-Myc), the inventors have identified direct GR target genes are significantly associated with cell survival signaling pathways. Interestingly, a meta-analysis of the high NR3C1-expressing ER- tumors reveals that many genes identified by ChIP-sequencing/gene expression analysis are indeed differentially expressed in high

versus low NR3C1-primary breast cancers. These results suggest that GR expression may be a functional biomarker in ER- breast cancer.

TABLE 1

Clinical studies used for meta-analysis		
GEO ID	# of pts	Reference
GSE9195	77	Loi S. et al
GSE7390	189	Desmedt C, et al
GSE6532	212	Loi S, et al
GSE2603	73	Minn AJ, et al
GSE2990	183	Sotiriou C, et al
GSE2034	280	Wang YX, et al
TOTAL	1206	

TABLE 2

Differentially expressed genes with concordant expression by all three methods (33/44 genes)			
Gene expression after Dex-treatment in MCF10A-Myc	Gene expression in NR3C1 + vs. - tumors	GR-binding within distance to TSS after Dex-treatment in MCF10A-Myc	Genes
Up	Up	10 kb 10-100 kb	DUSP1, SGK1, SMARCA2, PTGDS, MCL1, DPYSL2, STOM, LAPTM5, NNMT, SERPINF1, NRIP1, WIPF1, BIN1, IL1R1, ST3GAL5, SEMA4D, MAP3K5, SMARCA2, DPT, BIRC3, PTGDS, PHF15, MAOA, TFPI, SLC46A3, PIAS1, ACSL5, SESN1, C14orf139, LBH
Down	Down	10 kb 10-100 kb	NONE SFN, SPP1, ERBB2
Overlapping genes with NKI-295 gene signature			DUSP1, DPT, NNMT, SERPINF1, IL1R1, FN1, DPYSL2

[0193] B. Materials and Methods

[0194] Cell Culture and Glucocorticoid Treatment:

[0195] MCF10A-Myc cells were cultured in a 1:1 mixture of DMEM and Hams/F12 medium supplemented with 10% fetal bovine serum, hydrocortisone (0.5 µg/ml), EGF (10 ng/ml), insulin (5 ng/ml) and 100 U/ml penicillin/streptomycin were also added. The cells were then starved for three days of all growth factors and treated with dexamethasone (10-6M) and ethanol of the same volume as a control.

[0196] Microarray Gene Expression: MCF10A-Myc Cells:

[0197] Time course (0.5 h, 2 h, 4 h and 24 h) microarray data were obtained using Affymetrix gene arrays (HG-U133A) (Wu et al., 2006). Genes that were induced or repressed >1.5 fold-change were considered to be regulated.

[0198] GR ChIP-Seq Experiment and Analysis for MCF10A-Myc Cells:

[0199] Cells were collected for the ChIP assay following 1 hour of Dex (10-6M) or EtOH treatment. The ChIP assay was done basically following Millipore's ChIP Assay Kit instructions. The DNA input (1%) was also sequenced using Illumina's Solexa Sequencer. Short-tag reads (36 bp) were mapped to the Human Genome (UCSC, hg18) by using Macq aligner. GR-binding peaks were called by using MACS software. Known SGK1 and GILZ promoter GR binding-regions (GBRs) were used as positive controls to determine the FDR threshold for retrieving significant GBRs.

[0200] Human Primary Breast Cancer Analysis:

[0201] 1) Data Collection: All the clinical data and raw CEL files (all Affymetrix HU-133A and HU-133+2) were obtained from GEO (see Table 1). Low quality arrays were removed by AffyPLM. Arrays were normalized by using RMA and then centered by mean within each study and pooled together. 2) Determination of ESR1 and NR3C1 positivity: Expression data of tumors with known ER IHC status were analyzed using ROC analysis. The Youden Index of the best ESR1 probe's ROC curve was used as the cut-off point to separate ESR1+ and ESR1- tumors. Due to the lack of tumors with both GR IHC and NR3C1 gene expression information, we were unable to use ROC analysis to determine the NR3C1 cutoff. Therefore, based on published and our unpublished GR IHC data, we used the percentiles of NR3C1 gene expression levels that correspond to the observed proportion of GR+ patients. 3) Clustering: Unsupervised clustering was performed by Cluster using Pear-

son correlation distance and complete-linkage method. Heat-maps were plotted by Treeview. 4) Statistical analysis: Relapse-free survival (RFS) Kaplan-Meier plot and log-rank test were done by using R's "survival" package. Microarray SAM analysis was performed by using R's "siggenes" package.

[0202] Tumor Assessment.

[0203] pAUC areas were calculated for all the probes on the chip by setting $p=0.2$ (meaning can separate at least 80% patients) for tumors with known ER status ($n=1000$). A probe was then selected that has biggest pAUC area, which is the ESR1 probe 205225_at. So, this probe is the best one that can separate ER IHC + versus -. Using the 205225_at probe, the Youden Index of its ROC curve was calculated, that is the max (sensitivity+specificity-1) as the cut-off value for ESR1+ and -. The range of ESR1 expression after normalization was [-5.223868-3.944120]. The Youden Index, i.e. the cut-off is -1.257434. In the $n=1000$, training set, $n=773 > -1.257434$ (ESR1+), and $n=227 < -1.257434$ (ESR1-) or i.e. 77.3% quantile

[0204] This cut-off was applied to the entire dataset, $n=898$ (ESR+), $n=308$ (ESR-). In addition to the method, the ACTUAL Log 2 value cutoff is needed for ESR1 positivity in normalized meta-dataset, as well as the range of ESR1 values encountered following batched mean normalization. If in one study, samples are obtained from different

hospitals, they were normalized separately. So, to be precisely accurate, the normalization is done within the samples from the same source.

[0205] The ESR1 probe ID from Affymetrix is 205225_at.

[0206] The NR3C1 probe ID from Affymetrix is 216321_s_at

[0207] The range for NR3C1 probe (216321_s_at) is [-3.145456 to 2.158716] for the entire data set. For ESR1+, the range is [-3.009359 2.158716] and for ESR1-, the range is [-3.145456 1.917823] Thus, the cut-off for ESR1+, is 0.172189, 55.98% quantile (or about 44% NR3C1+ percentage) and the cut-off for ESR1-, is 0.47332, 82.51% quantile (or about 17.5% NR3C1+ percentage). All the cut-off are log 2 values.

[0208] The cutoffs used are the best cut-off that can separate patients with a p<0.01. If the p-value is loosened to 0.05, the range can be widened.

[0209] For ESR1+ patients, NR3C1+ patients can be from about 35% to 60% (about 44% is the best). For ESR1- patients, NR3C1+ patients can be from about 30% to 15% (about 17.5% is the best)

Example 2

Mifepristone Pretreatment Enhances Paclitaxel Anti-Tumor Effectiveness in Models of Human Breast Cancer

[0210] Xenografted ER-/PR-/HER2- (GR+) MDA-MB-231 human breast cancer cells (1x10⁷ cells in 50 µl of PBS) were injected into the mammary fat pad of female Severe

Combined Immunodeficient Mice (SCID) mice and allowed to grow until reaching approximately 100 mm³. Mice were then injected intraperitoneally with either both vehicles, paclitaxel (10 mg/kg)+the mifepristone vehicle, or the combination of mifepristone (15 mg/kg) administered two hours prior to paclitaxel (10 mg/kg) for five successive days. The longest (L) and shortest (S) diameters of the tumors were measured bi-weekly with electronic calipers and tumor volume was calculated using the formula for an ellipsoid sphere: volume=S2xLx0.52. Mifepristone pretreatment significantly decreased tumor volume over time (P=0.013) compared to treatment with paclitaxel alone (FIG. 8).

Example 3

Mifepristone Pretreatment Increases Tamoxifen-Resistant MCF-7 (T-R-MCF-7), but not Parental MCF-7 Cell Susceptibility to Paclitaxel In Vitro

[0211] Parental MCF-7 (ER+/PR+/GR+) and T-R MCF-7 (ER+/PR+/GR+) cells were treated with the appropriate vehicle (ethanol for mifepristone and castor oil/saline for paclitaxel), paclitaxel alone (10⁻⁶ M), and paclitaxel/mifepristone (10⁻⁶ M). Apoptosis was measured using FITC conjugated-anti-Annexin V antibody labeling followed FACS analysis to determine the percentage of the total cell population undergoing apoptosis after 20 hours of treatment. Mean+/-SE is shown. Significantly more apoptosis (P=0.028) was observed in the T-R MCF-7 cells when treated with mifepristone/paclitaxel compared to paclitaxel alone (FIG. 9). No difference was seen in the parental MCF-7 cells.

Sequence Listing

NR3C1 GenBank AY436590 - 127687 bp, incorporated herein by reference
 ESR1 GenBank NG_008493 - 419779 bp, incorporated herein by reference
 NR3C1 mRNA

SEQ ID NO: 1

TTTTTAGAAAAAATATATATTTCCCTCCTGCTCCTTCTGCGTTCACAAGCTAAGTTGTTTATCTCGGC
 TCGCGCGGAACTCGCGGACGGTGGCGGCGAGCGGCTCCTCTGCCAGAGTTGATATTACTGATGGACTC
 CAAAGAATCATTAACCTCTGGTAGAGAAGAAAACCCAGCAGTGTGCTTGCTCAGGAGAGGGGAGATGTG
 ATGGACTTCTATAAAACCTAAGAGGAGGAGCTACTGTGAAGTTTCTGCGTCTTACACCTCACTGGCTG
 TCGCTTCTCAATCAGACTCCAAGCAGCGAAGACTTTTGGTTGATTTTCCAAAAGGCTCAGTAAGCAATGC
 GCAGCAGCCAGATCTGTCCAAGCAGTTTCACTCTCAATGGACTGTATATGGGAGAGACAGAAACAAAA
 GTGATGGGAAATGACCTGGGATTTCCACAGCAGGGCCAAATCAGCCTTTCTCGGGGAAAACAGACTTAA
 AGCTTTTGAAGAAAGCATTGCAAACTCAATAGGTCGACCAGTGTTCAGAGAACCCCAAGAGTTCAGC
 ATCCACTGCTGTCTGCTGCCCCACAGAGAAGGAGTTTCCAAAACCTCACTCTGATGTATCTTCAGAA
 CAGCAACATTTGAAGGGCCAGACTGGCACCACCGTGGCAATGTGAAATTGTATACCACAGACC AAGCA
 CCTTTGACATTTTGAGGATTTGGAGTTTCTTCTGGGTCCCAGGTAAGAGACGAATGAGAGTCTTTG
 GAGATCAGACCTGTTGATAGATGAAAACCTGTTTGTCTTCTCCTCTGCGGGAGAAGACGATTTCATCCTT
 TTGGAAGGAACTCGAATGAGGACTGCAAGCCTCTCATTTTACCGGACACTAAACCCAAAATTAAGGATA
 ATGGAGATCTGGTTTTGTCAAGCCCCAGTAATGTAACACTGCCCAAGTAAAAACAGAAAAAGAGATTT
 CATCGAACTCTGCACCCCTGGGTAATTAAGCAAGAGAACTGGGCAGTTTTACTGTGAGGCAAGCTTT
 CCTGGAGCAAAATAAATGGTAATAAAATGTCTGCCATTTCTGTTTCAATGTTGAGTACCTCTGGAGGAC
 AGATGTACCACATGACATGAATACAGCATCCCTTTCTCAACAGCAGGATCAGAAGCCTATTTTTAATGT

-continued

CATTCCACCAATTCCTGGTTCCGAAAATTGGAATAGGTGCCAAGGATCTGGAGATGACAACCTTGACT
TCTCTGGGGACTCTGAACCTCCCTGGTCGAACAGTTTTTTCTAATGGCTATTCAAGCCCCAGCATGAGAC
CAGATGTAAGCTCTCCTCCATCCAGCTCCTCAACAGCAACAACAGGACCACCTCCCAAACCTGCTGCTGGT
GTGCTCTGATGAAGCTTCAGGATGTCATTATGGAGTCTTAACCTGTGGAAGCTGTAAAGTTTTCTTCAA
AGAGCAGTGGAAAGACAGCACAATTACCTATGTGCTGGAAGGAATGATTGCATCATCGATAAAAATTCGAA
GAAAAAATGCCAGCATGCCGCTATCGAAAAATGCTTTCAGGCTGGAATGAACCTGGAAGCTCGAAAAAC
AAAGAAAAATAAAAGGAATTCAGCAGGCCACTACAGGAGTCTCACAGAAACCTCTGAAAATCCTGGT
AACAAAAACAATAGTTCTTGCACAGTTACCCACAACCTACCCCTACCTGGTGTACTGTTGGAGGTTATTG
AACCTGAAGTGTATATGCAGGATATGATAGCTCTGTTCCAGACTCAACTGGAGGATCATGACTACGCT
CAACATGTTAGGAGGGCGGCAAGTGATTGCAGCAGTGAATGGGCAAGGCAATACCAGGTTTCAGGAAC
TTACACCTGGATGACCAAAATGACCTACTGCAGTACTCCTGGATGTTTTCTTATGGCATTGCTCTGGGT
GGAGATCATATAGACAATCAAGTGCAAACCTGCTGTGTTTTGCTCCTGATCTGATTATTAATGAGCAGAG
AATGACTCTACCTGCATGTACGACCAATGTAACACATGCTGTATGTTTTCTCTGAGTTACACAGGCTT
CAGGTATCTTATGAAGAGTATCTCTGTATGAAAACCTTACTGCTTCTCTCTCAGTTCTTAAGGACGGTC
TGAAGAGCCAAAGAGCTATTTGATGAAATTAGAATGACCTACATCAAAGAGCTAGGAAAAGCCATTGTCAA
GAGGGAAGGAAACTCCAGCCAGAAGTGGCAGCGGTTTTATCAACTGACAAAACCTTTGGATTCTATGCAT
GAAGTGGTTGAAAATCTCCTTAACCTATTGCTTCCAAACATTTTTGGATAAGACCATGAGTATTGAATTC
CCGAGATGTTAGCTGAAATCATACCAATCAGATACCAAATATTCAAATGGAAATATCAAAAACCTTCT
GTTTTATCAAAGTACTGCTTAAATAAGAATGGTTCCTTAAAGAAAGTCAATTAATAGCTTTTATTG
TATAAATATCAGTTTGTCTGTAGAGGTTTTGTTGTTTTATTTTTATTTGTTTTATCTGTTGTTTTGT
TTAAATACGCACTACATGTGGTTTATAGAGGGCCAAAGACTTGGCAACAGAAGCAGTTGAGTCGTACATCA
CTTTTCACTGATGGGAGAGTAGATGGTGAATTTATAGTTAATATATCCAGAAATAGAAACCTTAAT
ATGTGGACGTAATCTCCACAGTCAAAGAAGGATGGCACCTAAACCACCAGTCCCAAAGTCTGTGTGATG
AACTTTCTCTCATACTTTTTTTCACAGTTGGCTGGATGAAATTTTCTAGACTTTCTGTTGGTGTATCCC
CCCCCTGTATAGTTAGGATAGCATTTTTGATTTATGCATGGAACCTGAAAAAAGTTTACAAGTGTAT
ATCAGAAAAGGAAAGTTGTGCTTTTATAGCTATTACTGTCTGGTTTTAACAAATTTCTTTATATTTAGT
GAACTACGCTTGCTCATTTTTTCTTACATAATTTTTATTCAAGTTATTGTACAGCTGTTAAGATGGGC
AGCTAGTTCGTAGCTTTCCAAATAAATCTAAACATTAATCAATCATCTGTGTGAAAATGGGTGGTGC
TTCTAACCTGATGGCACTTAGCTATCAGAAGACCACAAAAATGACTCAAACTCCAGTATTCTGTCAA
AAAAAAGAAAAAAGCTCATATTTGTATATATCTGCTTCAGTGGAGAATTATATAGGTTGTGCAAT
TAACAGTCTAACTGGTATAGAGCACCTAGTCCAGTGACCTGCTGGGTAACCTGTGGATGATGGTTGCAA
AAGACTAATTTAAAAATAACTACCAAGAGCCCTGTCTGTACCTAACGCCCTATTTTTGCAATGGCTAT
ATGGCAAGAAAGCTGCTAAACTATTGCTTTTTCAGGACCTTTTGAAGTAGTTTGTATAACTTCTTAAAG
TTGTGATCCAGATAACCAGCTGTAACACAGCTGAGAGACTTTAATCAGACAAAGTAATCTCTCACT
AACTTTTACCAAAAACTAAATCTTAATATGGCAAAAATGGCTAGACACCCATTTTACATTTCCATCT
GTCACCAATTTGGTTAATCTTTCTGATGTTACAGGAAAGCTCAGCTACTGATTTTTGTGATTTAGAAGT
TATGTATGTACAGATCCATGTTTGTAAAACCTACACATCCCTAATGTGTGCATAGAGTTAACACAAGT
CCTGTGAATTTCTCACTGTTGAAAATATTTTTAAACAAAATAGAAGCTGTAGTAGCCCTTTCTGTGTGC
ACCTTACCAACTTTCTGTAAACTCAAACTTAACATATTTACTAAGCCACAAGAAATTTGATTTCTATTC

-continued

AAGGTGGCCAAATTATTTGTGTAATAGAAAACGAAAATCTAATATATAAATATGGAACCTCTAATATA
TTTTTATATTTAGTTATAGTTTCAGATATATATCATATTGGTATTCACCTAATCTGGGAAGGGAAGGGCTA
CTGCAGCTTTACATGCAATTTATTAATAAGTATGTAATAAGCTTGTATAGTGTAAAAATAAGAATGATTT
TTAGATGAGATTGTTTTATCATGACATGTTATATATTTTTTGTAGGGGTCAAAGAAATGCTGATGGATAA
CCTATATGATTTATAGTTTGTACATGCATTCATACAGGCAGCGATGGTCTCAGAAACCAAACAGTTTGCT
CTAGGGGAAGAGGAGATGGAGACTGGTCCTGTGTGCAGTGAAGGTTGCTGAGGCTCTGACCCAGTGAGA
TTACAGAGGAAGTTATCCTCTGCCTCCCATTCTGACCACCCTTCTCATTCACACAGTGAGTCTGTCAGCG
CAGGTTTAGTTTACTCAATCTCCCTTGCCTAAAGTATGTAAAGTATGTAAACAGGAGACAGGAAGGTG
GTGCTTACATCCTTAAAGGCACCATCTAATAGCGGGTTACTTTCACATACAGCCCTCCCCAGCAGTTGA
ATGACAACAGAAGCTTCAAGTTTGGCAATAGTTTGCATAGAGGTACCAGCAATATGTAATAAGTGCAG
AATCTCATAGGTTGCCAATAATACACTAATCTCTTTCTATCTACACACAGAGTTTATTTCCAATAAAA
TGAGGACATGTTTTTGTCTTTTGAATGCTTTTTGAATGTTATTGTTATTTTCAGTATTTTGAGAGAA
TTATTTAATAAAAAAACAATCATTGCTTTTTTG

ESR1 mRNA

SEQ ID NO: 2

AGGAGCTGGC GGAGGGCGTT CGTCCTGGGA CTGCACTTGC TCCCGTCGGG TCGCCCGGCT
TCACCGGACC CGCAGGCTCC CGGGGCAGGG CCGGGGCCAG AGCTCGCGTG TCGGCGGGAC
ATGCGCTGCG TCGCCTCTAA CCTCGGGCTG TGCTCTTTTT CCAGTGGCC CGCCGGTTTC
TGAGCCTTCT GCCCTGCGGG GACACGGTCT GCACCCTGCC CGCGGCCACG GACCATGACC
ATGACCCTCC ACACCAAAGC ATCTGGGATG GCCCTACTGC ATCAGATCCA AGGGAACGAG
CTGGAGCCCC TGAACCGTCC GCAGCTCAAG ATCCCCCTGG AGCGGCCCTT GGGCGAGGTG
TACCTGGACA GCAGCAAGCC CGCCGTGTAC AACTACCCCG AGGGCGCCGC CTACGAGTTC
AACGCCGCGG CCGCCGCCAA CGCGCAGGTC TACGGTCAGA CCGGCCTCCC CTACGGCCCC
GGGTCTGAGG CTGCGGCGTT CGGCTCCAAC GGCTGGGG GTTTCCTCCC ACTCAACAGC
GTGTCTCCGA GCCCGCTGAT GCTACTGCAC CCGCCGCCGC AGCTGTGCC TTTCTGCAG
CCCCACGGCC AGCAGGTGCC CTACTACCTG GAGAACGAGC CCAGCGGCTA CACGGTGCCG
GAGGCCGGCC CGCCCGCATT CTACAGGCCA AATTCAGATA ATCGACGCCA GGGTGGCAGA
GAAAGATTGG CCAGTACCAA TGACAAGGA AGTATGGCTA TGAATCTGC CAAGGAGACT
CGCTACTGTG CAGTGTGCAA TGAATATGCT TCAGGCTACC ATTATGGAGT CTGGTCCTGT
GAGGGCTGCA AGGCCCTTCT CAAGAGAAGT ATTCAAGGAC ATAACGACTA TATGTGTCCA
GCCACCAACC AGTGCACCAT TGATAAAAAC AGGAGGAAGA GCTGCCAGGC CTGCCGGCTC
CGCAAATGCT ACGAAGTGGG AATGATGAAA GGTGGGATAC GAAAAGACCG AAGAGGAGGG
AGAATGTTGA AACACAAGCG CCAGAGAGAT GATGGGGAGG GCAGGGGTGA AGTGGGTCT
GCTGGAGACA TGAGAGCTGC CAACCTTTGG CCAAGCCCGC TCATGATCAA ACGCTCTAAG
AAGAACAGCC TGGCCTTGTG CCTGACGGCC GACCAGATGG TCAGTGCCTT GTTGGATGCT
GAGCCCCCA TACTCTATTG CGAGTATGAT CCTACCAGAC CCTTCAGTGA AGCTTCGATG
ATGGGCTTAC TGACCAACCT GGCAGACAGG GAGCTGGTTC ACATGATCAA CTGGCCGAAG
AGGGTGCCAG GCTTTGTGGA TTTGACCCTC CATGATCAGG TCCACCTTCT AGAATGTGCC
TGGCTAGAGA TCCTGATGAT TGGTCTCGTC TGGCGCTCCA TGGAGCACC AGGGAAGCTA
CTGTTTGCTC TAACTTGCT CTTGGACAGG AACAGGGAA AATGTGTAGA GGGCATGCTG

-continued

GAGATCTTCG ACATGCTGCT GGCTACATCA TCTCGGTTCC GCATGATGAA TCTGCAGGGA
GAGGAGTTTG TGTGCCTCAA ATCTATTATT TTGCTTAATF CTGGAGTGTA CACATTTCTG
TCCAGCACCC TGAAGTCTCT GGAAGAGAAG GACCATATCC ACCGAGTCCT GGACAAGATC
ACAGACACTT TGATCCACCT GATGGCCAAG GCAGGCCTGA CCCTGCAGCA GCAGCACCAG
CGGCTGGCCC AGCTCCTCCT CATCCTCTCC CACATCAGGC ACATGAGTAA CAAAGGCATG
GAGCATCTGT ACAGCATGAA GTGCAAGAAC GTGGTGCCCC TCTATGACCT GCTGCTGGAG
ATGCTGGACG CCCACCGCCT ACATGCGCCC ACTAGCCGTG GAGGGGCATC CGTGGAGGAG
ACGGACCAAA GCCACTTGGC CACTGCGGGC TCTACTTCAT CGCATTCCCTT GCAAAAAGTAT
TACATCACGG GGGAGGCAGA GGGTTTCCCT GCCACGGTCT GAGAGCTCCC TGGCTCCAC
ACGGTTCAGA TAATCCCTGC TGCATTTTAC CCTCATCATG CACCACTTTA GCCAAATTCT
GTCTCCTGCA TACACTCCGG CATGCATCCA ACACCAATGG CTTTCTAGAT GAGTGGCCAT
TCATTTGCTT GCTCAGTTCT TAGTGGCACA TCTTCTGTCT TCTGTTGGGA ACAGCCAAAG
GGATTCCAAG GCTAAATCTT TGTAACAGCT CTCTTTCCCC CTTGCTATGT TACTAAGCGT
GAGGATTCCT GTAGCTCTTC ACAGCTGAAC TCAGTCTATG GGTGGGGCT CAGATAACTC
TGTGCATTTA AGTACTTGT AGAGACCCAG GCCTGGAGAG TAGACATTTT GCCTCTGATA
AGCACTTTTT AAATGGCTCT AAGAATAAGC CACAGCAAAG AATTTAAAGT GGCTCCTTTA
ATTGGTGACT TGGAGAAAAGC TAGGTCAAGG GTTTATTATA GCACCCTCTT GTATTCCTAT
GGCAATGCAT CCTTTTATGA AAGTGGTACA CCTTAAAGCT TTTATATGAC TGTAGCAGAG
TATCTGGTGA TTGTCAATTC ATTCCCCCTA TAGGAATACA AGGGGCACAC AGGGAAGGCA
GATCCCCTAG TTGGCAAGAC TATTTTAACT TGATACACTG CAGATTCAGA TGTGCTGAAA
GCTCTGCCTC TGGCTTCCG GTCATGGGTT CCAGTTAATF CATGCCTCCC ATGGACCTAT
GGAGAGCAGC AAGTTGATCT TAGTTAAGTC TCCCTATATG AGGGATAAGT TCCTGATTTT
TGTTTTTATT TTTGTGTTAC AAAAGAAAAGC CCTCCCTCCC TGAACCTGCA GTAAGGTCAG
CTTCAGGACC TGTTCAGTG GGCACCTGTAC TTGGATCTTC CCGGCGTGTG TGTGCCTTAC
ACAGGGGTGA ACTGTTCACT GTGGTGATGC ATGATGAGGG TAAATGGTAG TTGAAAGGAG
CAGGGGCCCT GGTGTTGCAT TTAGCCCTGG GGCATGGAGC TGAACAGTAC TTGTGCAGGA
TTGTTGTGGC TACTAGAGAA CAAGAGGGAA AGTAGGGCAG AAACCTGGATA CAGTTCTGAG
GCACAGCCAG ACTTGCTCAG GGTGGCCCTG CCACAGGCTG CAGCTACCTA GGAACATTC
TTGCAGACCC CGCATTGCCC TTTGGGGGTG CCCTGGGATC CCTGGGGTAG TCCAGCTCTT
CTTCATTTCC CAGCGTGGCC CTGGTTGGAA GAAGCAGCTG TCACAGCTGC TGTAGACAGC
TGTGTTCCCTA CAATTGGCCC AGCACCCCTG GGCACGGGAG AAGGGTGGGG ACCGTTGCTG
TCACTACTCA GGCTGACTGG GGCCTGGTCA GATTACGTAT GCCCTTGGTG GTTTAGAGAT
AATCCAAAAT CAGGGTTTGG TTTGGGGAAG AAAATCCTCC CCCTTCCTCC CCGCCCCCGT
TCCCTACCGC CTCCACTCCT GCCAGCTCAT TTCCTTCAAT TTCCTTTGAC CTATAGGCTA
AAAAAGAAAAG GCTCATTTCA GCCACAGGGC AGCCTTCCCT GGGCCTTTC TTCTCTAGCA
CAATTATGGG TTAATTCCTT TTTCTTAAAC AAAAAGAATG TTTGATTTCC TCTGGGTGAC
CTTATTGTCT GTAATTGAAA CCCTATTGAG AGGTGATGTC TGTGTTAGCC AATGACCCAG
GTGAGCTGCT CGGGCTTCTC TTGGTATGTC TTGTTTGGAA AAGTGGATTT CATTCAATTC
TGATTGTCCA GTTAAGTGAT CACCAAAGGA CTGAGAATCT GGGAGGGCAA AAAAAAAAAA

-continued

AAAGTTTTTA TGTGCACTTA AATTTGGGGA CAATTTTATG TATCTGTGTT AAGGATATGT
TTAAGAACAT AATTCTTTTG TTGCTGTTTG TTTAAGAAGC ACCTTAGTTT GTTTAAGAAG
CACCTTATAT AGTATAATAT ATATTTTTTT GAAATTACAT TGCTTGTTTA TCAGACAATT
GAATGTAGTA ATTCTGTTCT GGATTTAATT TGA CTGGGTT AACATGCAAA AACCAAGGAA
AAATATTTAG TTTTTTTTTT TTTTTTTGTA TACTTTTCAA GCTACCTTGT CATGTATACA
GTCATTTATG CCTAAAGCCT GGTGATTATT CATTTAAATG AAGATCACAT TTCATATCAA
CTTTTGATC CACAGTAGAC AAAATAGCAC TAATCCAGAT GCCTATTGTT GGATACTGAA
TGACAGACAA TCTTATGTAG CAAAGATTAT GCCTGAAAAG GAAAATTATT CAGGGCAGCT
AATTTGCTT TTACCAAAAT ATCAGTAGTA ATATTTTGG ACAGTAGCTA ATGGGTCAGT
GGGTTCTTTT TAATGTTTAT ACTTAGAITT TCTTTTAAAA AAATTAATAA AAAACAAAAA
AAAATTTCTA GGACTAGACG ATGTAATACC AGCTAAAGCC AAACAATTAT ACAGTGGAAG
GTTTTACATT ATTCATCCAA TGTGTTTCTA TTCATGTTAA GATACTACTA CATTGGAAGT
GGCAGAGAA CATCAGATGA TTGAAATGTT CGCCAGGGG TCTCCAGCAA CTTTGAAAT
CTCTTTGAT TTTTACTTGA AGTGCCACTA ATGGACAGCA GATATTTTCT GGCTGATGTT
GGTATTGGGT GTAGGAACAT GATTTAAAA AAAACTCTTG CCTCTGCTTT CCCCCACTCT
GAGGCAAGTT AAAATGTAAA AGATGTGATT TATCTGGGG GCTCAGGTAT GGTGGGGAAG
TGGATTCAAG AATCTGGGGA ATGGCAAATA TATTAAGAAG AGTATTGAAA GTATTTGGAG
GAAAATGGTT AATTCTGGGT GTGCACCAGG GTTCAGTAGA GTCCACTTCT GCCCTGGAGA
CCACAATCA ACTAGTCCA TTTACAGCCA TTTCTAAAA GGCAGCTTCA GTTCTAGAGA
AGAAAAGAACA ACATCAGCAG TAAAGTCCAT GGAATAGCTA GTGGTCTGTG TTTCTTTTCG
CCATTGCCTA GCTTGCCGTA ATGATTCTAT AATGCCATCA TGCAGCAATT ATGAGAGGCT
AGGTCATCCA AAGAGAAGAC CCTATCAATG TAGGTTGCAA AATCTAACC CTAAGGAAGT
GCAGTCTTTG ATTTGATTTT CCTAGTAACC TTGCAGATAT GTTTAACCAA GCCATAGCCC
ATGCCTTTTG AGGGCTGAAC AAATAAGGGA CTTACTGATA ATTTACTTTT GATCACATTA
AGGTGTTCTC ACCTTGAAAT CTTATACACT GAAATGGCCA TTGATTTAGG CCACTGGCTT
AGAGTACTCC TTCCCCTGCA TGACACTGAT TACAAATACT TTCCTATTCA TACTTTCCAA
TTATGAGATG GACTGTGGGT ACTGGGAGTG ATCACTAACA CCATAGTAAT GTCTAATATT
CACAGGCAGA TCTGCTTGGG GAAGCTAGTT ATGTGAAAG CAAATAGAGT CATAAGTAG
CTCAAAAGGC AACATAAAT CTCTTTGGTG CAGGTCTTGG GAGCGTGATC TAGATTACAC
TGCACCATT CCAAGTTAAT CCCCTGAAA CTTACTCTCA ACTGGAGCAA ATGAACTTTG
GTCCCAAATA TCCATCTTTT CAGTAGCGTT AATTATGCTC TGTTCCTAAC TGCAATTCCT
TTCCAATTGA ATTAAGTGT GGCCTCGTTT TTAGTCATTT AAAATTGTTT TCTAAGTAAT
TGCTGCCTCT ATTATGGCAC TTCAATTTT CACTGTCTTT TGAGATTCAA GAAAAATTTT
TATTTCTTTT TTTGCATCCA ATTTGCTCTG AACTTTTAAA ATATGTAAT GCTGCCATGT
TCCAAACCCA TCGTCAGTGT GTGTGTTIAG AGCTGTGCAC CCTAGAAAACA ACATATTGTC
CCATGAGCAG GTGCCTGAGA CACAGACCCC TTTGCATTCA CAGAGAGGTC ATTGGTTATA
GAGACTTGAA TTAATAAGTG ACATTATGCC AGTTTCTGTT CTCTCACAGG TGATAAACAA
TGCTTTTTGT GCACTACATA CTCTTCAGTG TAGAGCTCTT GTTTTATGGG AAAAGGCTCA
AATGCCAAAT TGTGTTTGAT GGATTAATAT GCCCTTTTGC CGATGCATAC TATTACTGAT

-continued

GTGACTCGGT TTTGTCGCAG CTTTGCTTTG TTTAATGAAA CACACTTGTA AACCTCTTTT

GCACCTTGAA AAAGAATCCA GCGGGATGCT CGAGCACCTG TAAACAATTT TCTCAACCTA

SEQ ID NO: 3-46

MCL1, SAP30, DUSP1, SGK1, SMARCA2, PTGDS, TNFRSF9, SFN, LAPTM5,

GPSM2, SORT1, DPT, NRP1, ACSL5, BIRC3, NNMT, IGFBP6, PLXNC1, SLC46A3,

C14orf139, PIAS1, IDH2, SERPINF1, ERBB2, PECAM1, LBH, ST3GALS5, IL1RL1,

BIN1, WIPF1, TPPI, FN1, FAM134A, NRIP1, RAC2, SPP1, PHF15, BTN3A2,

SESN1, MAP3K5, DPYSL2, SEMA4D, STOM, and MAOA gene.

REFERENCES

- [0212] The following references, to the extent that they provide exemplary procedural or other details supplementary to those set forth herein, are specifically incorporated herein by reference.
- [0213] U.S. Pat. No. 5,143,854
 [0214] U.S. Pat. No. 5,202,231
 [0215] U.S. Pat. No. 5,242,974
 [0216] U.S. Pat. No. 5,288,644
 [0217] U.S. Pat. No. 5,324,633
 [0218] U.S. Pat. No. 5,384,261
 [0219] U.S. Pat. No. 5,405,783
 [0220] U.S. Pat. No. 5,412,087
 [0221] U.S. Pat. No. 5,424,186
 [0222] U.S. Pat. No. 5,429,807
 [0223] U.S. Pat. No. 5,432,049
 [0224] U.S. Pat. No. 5,436,327
 [0225] U.S. Pat. No. 5,445,934
 [0226] U.S. Pat. No. 5,468,613
 [0227] U.S. Pat. No. 5,470,710
 [0228] U.S. Pat. No. 5,472,672
 [0229] U.S. Pat. No. 5,492,806
 [0230] U.S. Pat. No. 5,503,980
 [0231] U.S. Pat. No. 5,510,270
 [0232] U.S. Pat. No. 5,525,464
 [0233] U.S. Pat. No. 5,525,464
 [0234] U.S. Pat. No. 5,527,681
 [0235] U.S. Pat. No. 5,529,756
 [0236] U.S. Pat. No. 5,532,128
 [0237] U.S. Pat. No. 5,545,531
 [0238] U.S. Pat. No. 5,547,839
 [0239] U.S. Pat. No. 5,554,501
 [0240] U.S. Pat. No. 5,556,752
 [0241] U.S. Pat. No. 5,561,071
 [0242] U.S. Pat. No. 5,571,639
 [0243] U.S. Pat. No. 5,580,726
 [0244] U.S. Pat. No. 5,580,732
 [0245] U.S. Pat. No. 5,593,839
 [0246] U.S. Pat. No. 5,599,672
 [0247] U.S. Pat. No. 5,599,695
 [0248] U.S. Pat. No. 5,610,287
 [0249] U.S. Pat. No. 5,624,711
 [0250] U.S. Pat. No. 5,631,134
 [0251] U.S. Pat. No. 5,639,603
 [0252] U.S. Pat. No. 5,654,413
 [0253] U.S. Pat. No. 5,658,734
 [0254] U.S. Pat. No. 5,661,028
 [0255] U.S. Pat. No. 5,665,547
 [0256] U.S. Pat. No. 5,667,972
- [0257] U.S. Pat. No. 5,695,940
 [0258] U.S. Pat. No. 5,700,637
 [0259] U.S. Pat. No. 5,744,305
 [0260] U.S. Pat. No. 5,800,992
 [0261] U.S. Pat. No. 5,807,522
 [0262] U.S. Pat. No. 5,830,645
 [0263] U.S. Pat. No. 5,837,196
 [0264] U.S. Pat. No. 5,847,219
 [0265] U.S. Pat. No. 5,871,928
 [0266] U.S. Pat. No. 5,876,932
 [0267] U.S. Pat. No. 5,919,626
 [0268] U.S. Pat. No. 6,004,755
 [0269] U.S. Pat. No. 6,087,102
 [0270] U.S. Pat. No. 6,368,799
 [0271] U.S. Pat. No. 6,383,749
 [0272] U.S. Pat. No. 6,617,112
 [0273] U.S. Pat. No. 6,638,717
 [0274] U.S. Pat. No. 6,720,138
 [0275] U.S. Patent Publ. 2010/0135956
 [0276] Belanoff et al., *Eur. J. Pharmacol.*, 655(1-3):117-20, 2011.
 [0277] Cho et al. *Biochemistry*, 44(9):3547-61, 2005.
 [0278] Clark, *Curr. Top. Med. Chem.* 8(9):813-838, 2008.
 [0279] Colleoni et al., *Annals of Oncology*, 11(8):1057, 2000.
 [0280] European Appln. EP 373 203
 [0281] European Appln. EP 785 280
 [0282] European Appln. EP 799 897
 [0283] Evans, *Science*, 240:889, 1988.
 [0284] Fodor et al., *Science*, 251:767-777, 1991.
 [0285] Grover and Martin, *Carcinogenesis*, 23(7):1095-102, 2002.
 [0286] Hacia et al., *Nature Genet.*, 14:441-449, 1996.
 [0287] Harrison's Principles of Internal Medicine, Kasper et al. (Eds.), 16th Ed., Chapter 70, 2005.
 [0288] Henderson et al. *Cancer Res.*, 48:246-253, 1988.
 [0289] Keen and Davidson, *Cancer*, 97(3 Suppl):825-33, 2003.
 [0290] Ma et al., *J. Immunol.*, 171(2):608-615, 2003.
 [0291] MacBeath and Schreiber, *Science*, 289(5485):1760-3, 2000.
 [0292] Melhem et al, *Clin. Cancer Res.*, 15(9):3196-204, 2009.
 [0293] Mikosz et al., *J. Biol. Chem.*, 276:16649-54, 2001.
 [0294] Moran et al., *Cancer Res.*, 60:867-872, 2000.
 [0295] Pandey and Mann, *Nature*, 405(6788):837-46, 2000.
 [0296] Pang and Conzen, *Cancer Biol. Ther. Cancer Biol. Ther.*, 5(8):933-40, 2006.
 [0297] PCT Appln. WO 01/68255

- [0298] PCT Appln. WO 03/020898
 [0299] PCT Appln. WO 03/022421
 [0300] PCT Appln. WO 03/023058
 [0301] PCT Appln. WO 03/029485
 [0302] PCT Appln. WO 03/040410
 [0303] PCT Appln. WO 03/053586
 [0304] PCT Appln. WO 03/066906
 [0305] PCT Appln. WO 03/067217
 [0306] PCT Appln. WO 03/076928
 [0307] PCT Appln. WO 03/087297
 [0308] PCT Appln. WO 03/091426
 [0309] PCT Appln. WO 03/093810
 [0310] PCT Appln. WO 03/100448A1
 [0311] PCT Appln. WO 04/020085
 [0312] PCT Appln. WO 04/027093
 [0313] PCT Appln. WO 09/923256
 [0314] PCT Appln. WO 09/936760
 [0315] PCT Appln. WO 93/17126
 [0316] PCT Appln. WO 95/11995
 [0317] PCT Appln. WO 95/21265
 [0318] PCT Appln. WO 95/21944
 [0319] PCT Appln. WO 95/35505
 [0320] PCT Appln. WO 96/30347
 [0321] PCT Appln. WO 96/31622
 [0322] PCT Appln. WO 96/33980
 [0323] PCT Appln. WO 97/10365
 [0324] PCT Appln. WO 97/27317
 [0325] PCT Appln. WO 97/43450
 [0326] PCT Appln. WO 99/35505
 [0327] PCT Appln. WO 01/38580
 [0328] PCT Appln. WO 03/100012
 [0329] Pease et al., *Proc. Natl. Acad. Sci. USA*, 91:5022-5026, 1994.
 [0330] Peeters et al., *Ann. NY Acad. Sci.*, 1148:536-41, 2008.
 [0331] Pike et al., *Epidemiologic Revi.*, 15(1):17-35, 1993.
 [0332] Shoemaker et al., *Nature Genetics*, 14:450-456, 1996.
 [0333] Sims et al. *BMC Medical Genomics*, 1(42):1-14, 2008.
 [0334] Sorlie et al., *Proc. Natl. Acad. Sci. USA*, 98:10869-10874., 2001.
 [0335] Srinivas et al., *Clin. Chem.*, 48(8):1160-9, 2002.
 [0336] UK Appln. 8 803 000
 [0337] Wu et al., *Cancer Res.*, 64:1757-64, 2004.
 [0338] Wu et al., *J. Clin. Invest.*, 114:560-568, 2004.
 [0339] Wu et al., *Mol Endocrinol.*, 2006

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 46

<210> SEQ ID NO 1

<211> LENGTH: 4794

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

```

tttttagaaa aaaaaaatat atttccctcc tgctccttct gcggtcacia gctaagttgt    60
ttatctcggg tgcggcgagg actgctggagc gtggcgggcg agcggctcct ctgccagagt    120
tgatattcac tgatggactc caaagaatca ttaactctcg gtagagaaga aaaccccagc    180
agtgtgcttg ctcaggagag gggagatgtg atggacttct ataaaacct aagaggagga    240
gctactgtga aggtttctgc gtcttcaccc tcactggctg tcgcttctca atcagactcc    300
aagcagcgaa gacttttggg tgattttcca aaaggctcag taagcaatgc gcagcagcca    360
gatctgtcca aagcagtttc actctcaatg ggactgtata tgggagagac agaaacaaaa    420
gtgatgggaa atgacctggg attcccacag cagggccaaa tcagcctttc ctcgggggaa    480
acagacttaa agcttttggg agaaagcatt gcaaacctca ataggtcgac cagtgttcca    540
gagaacccca agagtccagc atccactgct gtgtctgctg cccccacaga gaaggagttt    600
ccaaaaactc actctgatgt atcttcagaa cagcaacatt tgaaggcca gactggcacc    660
aacggtggca atgtgaaatt gtataccaca gaccaaagca cctttgacat tttgcaggat    720
ttggagtttt cttctgggct cccaggtaaa gagacgaatg agagtccttg gagatcagac    780
ctgttgatag atgaaaactg tttgctttct cctctggcgg gagaagacga ttcattcctt    840
ttggaaggaa actcgaatga ggactgcaag cctctcattt taccggacac taaacccaaa    900
attaaggata atggagatct ggttttgtca agccccagta atgtaacct gccccaagtg    960
aaaacagaaa aagaagattt catcgaactc tgcacccctg gggtaattaa gcaagagaaa   1020
ctgggcacag tttactgtca ggcaagcttt cctggagcaa atataattgg taataaaatg   1080

```

-continued

tctgccattt ctgttcatgg tgtgagtacc tctggaggac agatgtacca ctatgacatg	1140
aatacagcat cectttctca acagcaggat cagaagccta tttttaatgt cattccacca	1200
attcccgttg gttccgaaaa ttggaatagg tgccaaggat ctggagatga caacttgact	1260
tctctgggga ctctgaactt ccctggtcga acagtttttt ctaatggcta tccaagcccc	1320
agcatgagac cagatgtaag ctctcctcca tccagctcct caacagcaac aacaggacca	1380
cctcccaaac tctgectggg gtgctctgat gaagcttcag gatgtcatta tggagtotta	1440
acttgaggaa gctgtaaaat tttcttcaaa agagcagtgg aaggacagca caattaccta	1500
tgtgctggaa ggaatgattg catcatcgat aaaattcgaa gaaaaactg cccagcatgc	1560
cgctatcgaa aatgtcttca ggctggaatg aaacctggaag ctcgaaaaac aaagaaaaaa	1620
ataaaaggaa ttcagcaggc cactacagga gtctcacaag aaacctctga aaatcctggt	1680
aacaaaaaa tagttcctgc aacgttacca caactcacco ctaccctggt gtcactggtg	1740
gaggttattg aaacctgaagt gttatatgca ggatatgata gctctgttcc agactcaact	1800
tggaggatca tgactacgct caacatgta ggaggggggc aagtgattgc agcagtgaaa	1860
tgggcaaagg caataccagg tttcaggaac ttacacctgg atgaccaaat gaccctactg	1920
cagtactcct ggatgtttct tatggcattt gctctggggg ggagatcata tagacaatca	1980
agtgcaaacc tgctgtgttt tgctcctgat ctgattatta atgagcagag aatgactcta	2040
ccctgcatgt acgaccaatg taaacacatg ctgtatgttt cctctgagtt acacaggett	2100
caggtatcct atgaagagta tctctgtatg aaaaccttac tgcttctctc ttcagttcct	2160
aaggacgggc tgaagagcca agagctattt gatgaaatta gaatgaccta catcaaagag	2220
ctaggaaaag ccattgtcaa gagggaagga aactccagcc agaactggca gcggttttat	2280
caactgacaa aactcttgga ttctatgcat gaagtgggtg aaaatctcct taactattgc	2340
ttccaaacat ttttgataa gaccatgagt attgaattcc ccgagatggt agctgaaatc	2400
atcaccaatc agataccaaa atattcaaat ggaaatatca aaaaacttct gtttcatcaa	2460
aagtgactgc cttaataaga atggttgccct taaagaaagt cgaattaata gcttttattg	2520
tataaactat cagtttgtcc tgtagaggtt ttgttgtttt attttttatt gttttcatct	2580
gttgttttgt tttaaatacg cactacatgt ggtttataga gggccaagac ttggcaacag	2640
aagcagttga gtcgtcatca cttttcagtg atgggagagt agatggtgaa atttattagt	2700
taatatatcc cagaaattag aaaccttaat atgtggacgt aatctccaca gtcaaagaag	2760
gatggcaact aaaccaccag tgccccaaagt ctgtgtgatg aactttctct tcatactttt	2820
tttcacagtt ggctggatga aatttctag actttctggt ggtgtatecc cccccctgta	2880
tagttaggat agcatttttg atttatgcat ggaaacctga aaaaaagttt acaagtgtat	2940
atcagaaaag ggaagtgtg ctttttatag ctattactgt ctggttttaa caatttcctt	3000
tatatattag gaactacgct tgctcatttt ttcttacata attttttatt caagttattg	3060
tacagctggt taagatgggc agctagttcg tagctttccc aaataaactc taaacattaa	3120
tcaatcatct gtgtgaaaat gggttggtgc ttctaacctg atggcactta gctatcagaa	3180
gaccacaaaa attgactcaa atctccagta ttcttgtaaa aaaaaaaaaa aaaaaagctc	3240
atattttgta tatatctgct tcagtggaga attatatagg ttgtgcaaat taacagtcct	3300
aactggtata gagcacctag tccagtgacc tgctgggtaa actgtggatg atggttcaa	3360

-continued

```

aagactaatt taaaaaataa ctaccaagag gccctgtctg tacctaacgc cctatTTTTg 3420
caatggctat atggcaagaa agctggtaaa ctatttGtct ttcaggacct tttgaagtag 3480
tttgTataac ttctTaaaag ttgtgattcc agataaccag ctgtaacaca gctgagagac 3540
ttttaatcag acaaaGtaat tccctcTact aaactTtacc caaaaactaa atctcTaaTa 3600
tggcaaaaat ggctagacac ccattTtTcac attcccatct gtcaccaatt ggtTaatctt 3660
tccTgatggT acaggaaagc tcagctactg attTtTgtga tttagaactg tatgtatgTc 3720
agacatccat gtttGtaaaa ctacacatcc cTaatgtgtg ccatagagtt taacacaagt 3780
cctgtgaatt tcttCactgt Tgaaaattat ttTaaacaaa atagaagctg tagtagccct 3840
ttctgtgtgc accttaccaa ctttctgtaa actcaaaact taacatattt actaagccac 3900
aagaaatttg atttctattc aaggTggcca aattatttgt gTaatagaaa actgaaaatc 3960
taatattaaa aatattggaac ttctaatata tttttatatt tagttatagt ttcagatata 4020
tatcatattg gtattcacta atctgggaag ggaagggcta ctgcagcttt acatgcaatt 4080
tattaaaatg attgTaaaat agctTgtata gtgTaaaata agaatgattt ttagatgaga 4140
ttgttttate atgacatgtt atatattttt tgtaggggtc aaGaaatgc Tgatggataa 4200
cctatatgat ttatagtttg tacatgcatt catacaggca gcgatggTct cagaaaccaa 4260
acagtttGct ctaggggaag agggagatgg agactggTcc tgtgtgcagt gaaggttGct 4320
gaggctctga cccagtgaga ttacagagga agttatcctc Tgectcccat tctgaccacc 4380
cttctcattc caacagtGag tctgtcagcg caggTttagt ttactcaatc tccccttGca 4440
ctaaagtatg taaagtatgt aaacaggaga caggaaggtg gtgctTacat cctTaaagGc 4500
accatcTaat agcgggttac ttTcacatac agccctcccc cagcagttga atgacaacag 4560
aagcttcaga agtttggcaa tagtttGcat agaggtacca gcaatatgta aatagtgcag 4620
aatctcatag gttGcCaata atacactaat tccTttctat cctacaacaa gagtttattt 4680
ccaaaataaaa Tgaggacatg tttttgtttt ctttgaatgc tttttgaatg ttatttGtTa 4740
ttttcagTat tttggagaaa ttattTaaTa aaaaaaCaAa tcatttGctt tttg 4794

```

<210> SEQ ID NO 2

<211> LENGTH: 6330

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

```

aggagctggc ggagggcgctt cgtcctggga ctgcacttgc tcccgtcggg tcgcccggct 60
tcaccggacc cgcaggctcc cggggcaggg ccggggccag agctcgcgtg tggcgggac 120
atgcgctgcg tcgcctctaa cctcgggctg tgctcttttt ccaggTggcc cgcgggttTc 180
Tgagccttct gccctgcggg gacacggTct gcaccctgcc cgcggccacg gaccatgacc 240
atgaccctcc acaccaagc atctgggatg gccctactgc atcagatcca agggaacgag 300
ctggagcccc tgaaccgtcc gcagctcaag atccccctgg agcggcccct gggcgaggtg 360
Tacctggaca gcagcaagcc cgcctgttac aactaccccg agggcgcgcg ctacgagttc 420
aacgccgcgg ccgcccGcaa cgcgcaggtc tacggTcaga ccggcctccc ctacggcccc 480
gggtctgagg ctgcggcgctt cggctccaac ggctggggg gtttcccccc actcaacagc 540
gtgtctccga gcccgctgat gctactgcac ccgcgcgcgc agctgtcgcC tttcctgcag 600

```

-continued

ccccacggcc agcaggtgcc ctactacctg gagaacgagc ccagcggcta cacgggtgcgc 660
gaggccggcc cgcgggcatt ctacaggcca aattcagata atcgacgcca gggtagcaga 720
gaaagattgg ccagtaccaa tgacaaggga agtatggeta tggaaatctgc caaggagact 780
cgctactgtg cagtgtgcaa tgactatgct tcaggctacc attatggagt ctggctectgt 840
gagggctgca aggccttctt caagagaagt attcaaggac ataacgacta tatgtgtcca 900
gccaccaacc agtgcaccat tgataaaaac aggaggaaga gctgccaggc ctgccggctc 960
cgcaaatgct acgaagtggg aatgatgaaa ggtgggatac gaaaagaccg aagaggaggg 1020
agaatgttga aacacaagcg ccagagagat gatggggagg gcagggtga agtggggtct 1080
gctggagaca tgagagctgc caacctttgg ccaagccgc tcatgatcaa acgctctaag 1140
aagaacagcc tggccttgtc cctgacggcc gaccagatgg tcagtgcctt gttgatgct 1200
gagccccca tactctatct cgagtatgat cctaccagac ccttcagtga agcttcgatg 1260
atgggcttac tgaccaacct ggcagacagg gagctggctc acatgatcaa ctgggcgaag 1320
agggtgccag gctttgtgga tttgaccctc catgatcagg tccacctctc agaatgtgcc 1380
tggetagaga tctgatgat tggctctgct tggcctcca tggagcacc agggaagcta 1440
ctgtttgctc ctaacttgct cttggacagg aaccaggaa aatgtgtaga gggcatggtg 1500
gagatcttcg acatgtgct ggctacatca tctcggttcc gcatgatgaa tctgcaggga 1560
gaggagttag tgtgcctcaa atctattatt ttgcttaatt ctggagtgtg cacatttctg 1620
tccagcacc tgaagtctct ggaagagaag gaccatatcc accgagtctt ggacaagatc 1680
acagacactt tgatccacct gatggccaag gcaggcctga cctcgagca gcagcaccag 1740
cggttgccc agctcctcct catcctctcc cacatcaggc acatgagtaa caaaggcatg 1800
gagcatctgt acagcatgaa gtgcaagaac gtggtgcccc tctatgacct gctgctggag 1860
atgctggacg cccaccgctc acatgcgccc actagcctg gaggggcatc cgtggaggag 1920
acggacaaa gccacttggc cactgcgggc tctacttcat cgcattcctt gcaaaagtat 1980
tacatcacgg gggaggcaga gggtttccct gccacggtct gagagetccc tggctccca 2040
acggttcaga taatcctctc tgcattttac cctcatcatg caccacttta gcaaaattct 2100
gtctcctgca tacactccgg catgcatcca acaccaatgg ctttctagat gagtggccat 2160
tcatttgctt gctcagttct tagtggcaca tcttctgtct tctgttggga acagccaaag 2220
ggattccaag gctaaatctt tgtaacagct ctctttcccc cttgctatgt tactaagcgt 2280
gaggattccc gtagctcttc acagctgaac tcagtctatg ggtggggct cagataactc 2340
tgtgcattta agctacttgt agagaccag gctggagag tagacatttt gectctgata 2400
agcacttttt aaatggctct aagaataagc cacagcaaag aatttaaagt ggctccttta 2460
attggtgact tggagaaagc taggtcaagg gtttattata gcacctctt gtattcctat 2520
ggcaatgcat ccttttatga aagtgtgaca ccttaagct tttatatgac tgtagcagag 2580
tatctggtga ttgtcaatc attcccccta taggaataca aggggcacac agggaaggca 2640
gatccccatg ttggcaagac tattttaact tgatacactg cagattcaga tgtgctgaaa 2700
gctctgcctc tggccttccg gtcattgggt ccagttaatt catgctccc atggacctat 2760
ggagagcagc aagtgtatct tagttaagtc tccctatag agggataagt tectgatttt 2820
tgtttttatt tttgtgttac aaaagaaagc cctcctccc tgaacttga gtaaggctag 2880

-continued

cttcaggacc tgttccagtg ggcactgtac ttggatcttc cggcgtgtg tgtgccttac 2940
acaggggtga actgttcaact gtggtgatgc atgatgaggg taaatggtag ttgaaaggag 3000
caggggcect ggtgttgcac ttagecctgg ggcattggagc tgaacagtac ttgtgcagga 3060
ttgttgtggc tactagagaa caagagggaa agtagggcag aaactggata cagttctgag 3120
gcacagccag acttgtctcag ggtggccctg ccacaggctg cagctaccta ggaacattcc 3180
ttgcagaacc cgcattgccc tttgggggtg ccctgggato cctggggtag tccagctctt 3240
cttcatttcc cagcgtggcc ctggttgaa gaagcagctg tcacagctgc tgtagacagc 3300
tgtgttccca caattggccc agcacccctg ggcacgggag aaggggtggg accggttctg 3360
tcaactactca ggctgactgg ggcctggta gattacgtat gcccttgggt gtttagagat 3420
aatccaaaat cagggtttgg tttggggaag aaaatcctcc ccttctctcc cccgccccgt 3480
tccctaccgc ctccactcct gccagctcat ttcctcaat ttcctttgac ctataggcta 3540
aaaaaagaag gctcattcca gccacagggc agccttcctt gggcctttgc tctctagca 3600
caattatggg ttacttctct tttcttaaca aaaaagaatg tttgatttcc tctgggtgac 3660
cttattgtct gtaattgaaa ccttattgag aggtgatgct tgtgttagcc aatgaccag 3720
gtgagctgct cgggcttctc ttggtatgct ttgtttgaa aagtggattt cattcatttc 3780
tgattgtcca gtttaagtga caccaaagga ctgagaatct gggagggcaa aaaaaaaaaa 3840
aaagttttta tgtgcactta aatttgggga caattttatg tatctgtgtt aaggatatgt 3900
ttaagaacat aattcttttg ttgctgtttg ttaagaagc accttagttt gtttaagaag 3960
caccttatat agtataatat atattttttt gaaattacat tgcttgttta tcagacaatt 4020
gaatgtagta attctgttct ggatttaatt tgactgggtt aacatgcaa aaccaaggaa 4080
aaatatttag tttttttttt tttttttgta tacttttcaa gctacctgtt catgtataca 4140
gtcatttatg cctaaagcct ggtgattatt catttaaag aagatcacat ttcatatcaa 4200
cttttgtatc cacagtagac aaaatagcac taatccagat gcctattggt ggatactgaa 4260
tgacagacaa tcttatgtag caaagattat gctgaaaag gaaaattatt cagggcagct 4320
aattttgctt ttaccaaaat atcagtagta atatttttgg acagtagcta atgggtcagt 4380
gggttctttt taatgtttat acttagattt tcttttaaaa aaattaaaat aaaacaaaaa 4440
aaaatttcta ggactagacg atgtaatacc agctaagcc aaacaattat acagtggaag 4500
gttttacatt attcatccaa tgtgtttcta ttcattgtaa gataactacta catttgaaat 4560
gggcagagaa catcagatga ttgaaatggt cggccagggg tctccagcaa ctttgaaat 4620
ctctttgtat ttttacttga agtgccacta atggacagca gatattttct ggctgatggt 4680
ggattgggt gtaggaacat gatttaaaaa aaaactcttg cctctgcttt cccccactct 4740
gaggcaagtt aaaatgtaaa agatgtgatt tatctggggg gctcaggtat ggtggggaag 4800
tggattcagg aatctgggga atggcaaata tattaagaag agtattgaaa gtatttggag 4860
gaaaaatggt aattctgggt gtgcaccagg gttcagtaga gtccacttct gccctggaga 4920
ccacaaatca actagctcca tttacagcca tttctaaaat ggcagcttca gttctagaga 4980
agaaagaaca acatcagcag taaagtccat ggaatagcta gtggtctgtg tttctttctg 5040
ccattgccta gottgcccga atgattctat aatgcatca tgcagcaatt atgagaggct 5100
aggtcatcca aagagaagac cctatcaatg taggttgcaa aatctaacc ctaaggaagt 5160

-continued

```

gcagtctttg atttgatttc cctagtaacc ttgcagatat gtttaaccaa gccatagccc 5220
atgccttttg agggctgaac aaataaggga cttactgata atttactttt gatcacatta 5280
agggtttctc accttgaat cttatacaact gaaatggcca ttgatttagg ccaactggett 5340
agagtactcc tccccctgca tgacactgat tacaaatact ttcctattca tactttccaa 5400
ttatgagatg gactgtgggt actgggagtg atcactaaca ccatagtaat gtctaataat 5460
cacaggcaga tctgcttggg gaagctagtt atgtgaaagg caaatagagt catacagtag 5520
ctcaaaaggc aaccataaatt ctctttggtg caggctttgg gagcgtgatc tagattacac 5580
tgcaccattc ccaagtaat cccctgaaa cttactctca actggagcaa atgaactttg 5640
gtcccaata tccatctttt cagtagcgtt aattatgctc tgtttccaac tgcatttctc 5700
ttccaattga attaaagtgt ggctcgttt ttagtcattt aaaattgttt tctaagtaat 5760
tgctgcctct attatggcac ttcaattttg cactgtcttt tgagattcaa gaaaaatttc 5820
tattcttttt tttgcatcca attgtgcctg aacttttaaa atatgtaaat gctgccatgt 5880
tccaaaccca tcgtcagtggt gtgtgtttag agctgtgcac cctagaaaca acatattgtc 5940
ccatgagcag gtgcctgaga cacagacccc tttgcatcca cagagaggtc attggttata 6000
gagacttgaa ttaataagtg acattatgcc agtttctggt ctctcacagg tgataaacia 6060
tgctttttgt gcaactacata ctcttcagtg tagagctctt gttttatggg aaaaggctca 6120
aatgccaaat tgtgtttgat ggattaatat gcccttttgc cgatgcatac tattactgat 6180
gtgactcggg tttgtgcgag ctttgccttg tttaatgaaa cacacttgta aacctctttt 6240
gcactttgaa aaagaatcca cggggatgct cgagcacctg taaacaattt tctcaaccta 6300
tttgatgttc aaataaagaa ttaactaaa 6330

```

<210> SEQ ID NO 3

<211> LENGTH: 4107

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 3

```

gcgcaacct ceggaagctg ccgccccctt ccccttttat gggaatactt tttttaaaaa 60
aaaagagttc gctggcgcca ccccgtagga ctggcgccc taaaaccgtg ataaaggagc 120
tgctcgccac ttctacttc cgcttccttc cagtaaggag tgggggtctt cccagtttt 180
ctcagccagg cggcggcggc gactggcaat gtttgccctc aaaagaaacg cggtaatcgg 240
actcaacctc tactgtgggg gggccggctt gggggcggc agcggcggc ccaccgccc 300
gggagggcga cttttggcta cggagaagga ggcctcgcc cggcgagaga tagggggagg 360
ggaggccggc gcggtgattg gcggaagcgc cggcgcaage ccccgcca cctcacgcc 420
agactcccgg agggtegcgc ggcgcggcc cattggcgcc gaggtcccgc acgtcacccg 480
gacccccgcg aggctgcttt tcttcgcgc caccgcccgc gcggcgccgc ttgaggagat 540
ggaagccccg gccctgacg ccatcatgtc gccgaagag gagctggacg ggtacgagcc 600
ggagcctctc ggaagcggc cggctgctct gccgctgctg gagttggtcg ggaatctgg 660
taataacacc agtacggacg ggtcactacc ctgacgccg ccgcccagcag aggaggagga 720
ggacgagttg taccggcagt cgctggagat tatctctcgg taccttcggg agcaggccac 780
cggcgccaag gacacaaagc caatgggcag gtctggggcc accagcagga aggcgctgga 840

```

-continued

gaccttacga cggggttgggg atggcgtgca gcgcaaccac gagacggcct tccaaggcat	900
gcttcggaaa ctggacatca aaaacgaaga cgatgtgaaa tcgttgtctc gagtgatgat	960
ccatgttttc agcgacggcg taacaaactg gggcaggatt gtgactctca tttcttttgg	1020
tgcctttgtg gctaaacact tgaagacat aaaccaagaa agctgcatcg aaccatttagc	1080
agaaagtatc acagacgttc tcgtaaggac aaaacgggac tggctagtta aacaaagagg	1140
ctgggatggg tttgtggagt tcttccatgt agaggaccta gaaggtaggca tcaggaatgt	1200
gctgctggct tttgcagggt ttgctggagt aggagctggt ttggcatatc taataagata	1260
gccttactgt aagtgaata gttgactttt aaccaaccac caccaccacc aaaaccagtt	1320
tatgcagttg gactccaagc tgtaacttcc tagagttgca ccctagcaac ctagccagaa	1380
aagcaagtgg caagaggatt atggctaaca agaataaata catgggaaga gtgctcccca	1440
ttgattgaag agtcaactgc tgaagaagc aaagttcagt ttcagcaaca aacaaacttt	1500
gtttgggaag ctatggagga ggacttttag atttagtgaa gatggtaggg tggaaagact	1560
taatttcctt gttgagaaca gaaagtggc cagtagccag gcaagtcata gaattgatta	1620
cccgccgaat tcattaattt actgtagtgt taagagaagc actaagaatg ccagtgacct	1680
gtgtaaaagt tacaagtaat agaactatga ctgtaagcct cagtactgta caagggagac	1740
ttttcctctc tctaatttag tttcccagta tacttcttag aaagtccaag tgttcaggac	1800
ttttatacct gttatacttt ggcttggttt ccatgattct tactttatta gcctagttta	1860
tcaccaataa tacttgacgg aaggtcagt aattagttat gaatatgat atcctcaatt	1920
cttaagacag cttgtaaatg tatttgtaaa aattgtatat atttttacag aaagtctatt	1980
tctttgaaac gaaggaagta tcgaatttac attagttttt ttcataccct tttgaacttt	2040
gcaacttcg taattaggaa cctgtttctt acagcttttc tatgctaaac tttgttctgt	2100
tcagttctag agtgtataca gaacgaattg atgtgtaact gtatgcagac tggttgtagt	2160
ggaacaaac tgataactc gcaggtttaa attttcttat ctgattttgg taagtattcc	2220
ttagataggt ttttcttga aaacctggga ttgagaggt gatgaatgga aattcttca	2280
cttcattata tgcaagtttt caataattag gtctaagtg agttttaagg ttactgatga	2340
cttacaata atgggctctg attgggcaat actcatttga gttccttcca tttgacctaa	2400
tttaactggt gaaatttaaa gtgaattcat gggctcatct ttaaagcttt tactaaaaga	2460
ttttcagctg aatggaactc attagctgtg tgcataaaa aagatccat caggtggatg	2520
gagagacatt tgatcccttg tttgcttaat aaattataaa atgatggctt ggaaaagcag	2580
gctagtctaa ccatgggtgt attattagge ttgcttgta cacacacagg tctaagccta	2640
gtatgtcaat aaagcaata ctactgttt tgtttctatt aatgattccc aaacctgtt	2700
gcaagttttt gcattggcat ctttgattt cagtcttgat gtttgttcta tcagacttaa	2760
ccttttattt cctgtccttc cttgaaattg ctgattgttc tgctccctct acagatattt	2820
atatcaatc ctacagcttt cccctgccat cctgaaactc tttctagccc ttttagattt	2880
tggcactgtg aaaccctgc tggaaacctg agtgaccctc cctccccacc aagagtccac	2940
agaccttca tctttcacga acttgatcct gttagcaggt ggtaatacca tgggtgctgt	3000
gacactaaca gtcattgaga ggtgggagga agtccctttt ccttggactg gtatcttttc	3060
aactattgtt ttatcctgtc tttgggggca atgtgtcaaa agtccccca ggaattttca	3120

-continued

```

gaggaaagaa cattttatga ggctttctct aaagtttctt ttgtatagga gtatgctcac 3180
ttaaatttac agaaagaggt gagctgtggt aaacctcaga gtttaaaagc tactgataaa 3240
ctgaagaaag tgtctatatt ggaactaggg tcatttgaaa gcttcagtct cggaacatga 3300
cctttagtct gtggactcca tttaaaaata ggtatgaata agatgactaa gaatgtaatg 3360
gggaagaact gccctgctcg cccatctcag agccataagg tcatctttgc tagagctatt 3420
ttacctatg tatttatcgt tcttgatcat aagccgctta tttatatcat gtatctctaa 3480
ggacctaaaa gcactttatg tagtttttaa ttaactctaa gatctgggta cggtaactaa 3540
aaaagcctgt ctgccaaatc cagtggaaac aagtgcatag atgtgaattg gtttttaggg 3600
gccccacttc ccaattcatt aggtatgact gtggaaatac agacaaggat cttagttgat 3660
atthtgggct tggggcagtg agggcttagg acacccaag tggtttggga aaggaggagg 3720
ggagtgtggt gtttataggg ggaggaggag gcaggtgggc taagtgtgta ctggctacgt 3780
agttcgggca aatcctccaa aagggaaagg gaggtttgca ttagaaggat ggcgctccca 3840
gtgactactt tttgacttct gtttgtctta cgcttctctc agggaaaaac atgcagtcct 3900
ctagtgttcc atgtacatcc tgtggggggg gaacaccttg gttctgggta aacagctgta 3960
cttttgatag ctgtgccagg aagggttagg accaactaca aattaatgtt gttgtgcaaa 4020
tgtagtgtgt ttccctaaat ttctgttttt cctgagaaaa aaaaataaat cttttattca 4080
aatacaggga aaaaaaaaaa aaaaaaa 4107

```

<210> SEQ ID NO 4

<211> LENGTH: 1126

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4

```

tccccatgtg acagtgagcg gggteccgc tccaggagac gctcagctct gcgtcccggc 60
cctcagcact gtccactggt tcggtgccag cagagaccag caggccccgg acagttgggtg 120
tttgccgtg ccgctgtcta acttggtgtg cagagtgaat tgccgctgcc ggagcggaga 180
gaggcggagc ggcaggaga gaggggatt ctgtcagcgc cggcctcggg agctcggaga 240
catgaacggc ttcacgctcg acgagatgag ccgcgccggg gatgcggccg ccgcagtggc 300
cgcagtggtc gctgccgctg ccgcccgcgc ctccgggggg aacgggaccg gcgcgggac 360
cggggctgag gtgccgggct cggggggcgt ctccagcggc gggccccggg gggcggcgg 420
gccgggcccc gggcaactgt gctgctcgc ggaggatggt gagcggctgc gccggcggc 480
aggcaacgcc agcttcagca agaggatcca gaagagcacc tcccagaaga aggtgaagat 540
cgagctggat aagagcga aagcatttta catatgtgat taccataaaa acttaattca 600
gagtgttcca aacagaagaa agagaaaagg gagtgatgat gatggagggt attcacctgt 660
tcaagatatt gatccccag aggttgattt ataccaatta caagtaaata cacttaggag 720
atacaaaaga cacttcaagc taccaaccag accaggactt aataaagcac aacttgttga 780
gatagttggt tgccacttta ggtctattcc agtgaatgaa aaagacacct taacatattt 840
catctactca gtgaagaatg acaagaacaa atcagatctc aaggttgata gtgggttcca 900
ctaggagacg tggaattgag actaataact tggatgttaa cactgtttac tgttttttca 960
catgtagaaa tgtctcttgt gtatttttct tacagaggat tttctctgat tttattttct 1020

```

-continued

```

ttgtttctga ctctaataat tagttgaaa ctcatataaa atgagctttc ctaaattaaa 1080
tctattttta ataaagggtta ttactattaa aaaaaaaaaa aaaaaa 1126

<210> SEQ ID NO 5
<211> LENGTH: 2040
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 5
tcgctgcgaa ggacatttgg gctgtgtgtg cgacgcgggt cggaggggca gtcgggggaa 60
ccgcgaagaa gccgaggagc ccggagcccc gcgtgacgct cctctctcag tccaaaagcg 120
gcttttgggt cggcgcagag agaccgggg gtctagcttt tcctcgaaaa gcgccgcct 180
gcccttgccc ccgagaacag acaaagagca ccgcagggcc gatcacgctg ggggcgctga 240
ggccggccat ggtcatggaa gtgggcaccc tggacgctgg aggcctgcgg gcgctgctgg 300
gggagcgagc ggcgcaatgc ctgctgctgg actgccgctc cttcttcgct tccaacgcgg 360
gccacatcgc cggctctgtc aacgtgcgct tcagcaccat cgtgcggcgc cgggccaagg 420
gcgccatggg cctggagcac atcgtgccc acgccgagct ccgcggcgcg ctgctggcgg 480
gcgcctaaca cgcctggtg ttgctggac agcgcagcgc cgcctcggac ggcgccaagc 540
gcgacggcac cctggccctg gcggccggcg cgtctgcgg cgaggcgcgc gccgcgcaag 600
tcttcttctc caaaggagga tacgaagcgt ttctggtctc ctgcccggag ctgtgcagea 660
aacagtcgac ccccatgggg ctacgccttc ccctgagtac tagcgtccct gacagcgcgg 720
aatctgggtg cagttcctgc agtaccacc tctacgatca gggtgccccg gtggaaatcc 780
tgccctttct gtacctgggc agtgcgtatc acgcttccc caaggacatg ctggatgcct 840
tgggcatcac tgccctgatc aacgtctcag ccaattgtcc caaccatttt gagggtaact 900
accagtacaa gagcatccct ttggaggaca accacaagge agacatcagc tcctggttca 960
acgaggccat tgacttcata gactccatca agaatgctgg aggaaggggtg tttgtccact 1020
gccaggcagg catttcccgg tcagccacca tetgccttgc ttacctatg aggactaatc 1080
gagtcaagct ggacgagggc tttgagttg tgaagcagag gcgaagcatc atctctccca 1140
acttcagctt catgggccag ctgctgcagt ttgagtcca ggtgctggct ccgcactgtt 1200
cggcagaggc tgggagcccc gccatggctg tgctcgaccg aggcacctcc accaccaccg 1260
tgttcaactt ccccgctctc atccctgtcc actccacgaa cagtgcgctg agctaccttc 1320
agagccccc atcgacctct cccagctgct gaaaggccac gggaggtgag gctcttcaca 1380
tcccattggg actccatgct ccttgagagg agaaatgcaa taactctggg aggggctcga 1440
gagggctggg ccttatttat ttaacttcac ccgagttcct ctgggtttct aagcagttat 1500
ggtgatgact tagcgtcaag acatttctgt aactcagcac attcgggacc aatatatagt 1560
gggtacatca agtccatctg acaaaatggg gcagaagaga aaggactcag tgtgtgatcc 1620
ggtttctttt tgctcgcccc tgttttttgt agaatctctt catgcttgac atacctacca 1680
gtattattcc cgacgacaca tatacatatg agaataatcc ttatttttt ttgtgtaggt 1740
gtctgccttc acaaatgtca ttgtctactc ctagaagaac caaatacctc aatttttggt 1800
tttgagtact gtactatcct gtaaataat ctttaagcagg tttgttttca gcaactgatgg 1860
aaaataccag tgttgggttt ttttttagtt gccaacagtt gtatgtttgc tgattattta 1920

```

-continued

```

tgacctgaaa taatatatctt cttcttctaa gaagacattt tgttacataa ggatgacttt 1980
tttatacaat ggaataaatt atggcatttc tattgaaatt tcaaaaaaaaa aaaaaaaaaa 2040

```

```

<210> SEQ ID NO 6
<211> LENGTH: 3208
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

```

```

<400> SEQUENCE: 6

```

```

agatattcat gaaccgttgc ttcttcacgc ctgccttctt cgctccctct gcctttctgg 60
cgctgttctc cctccctccc tctgcttctt gctctttctt actccttctc tcagctgctt 120
aactacagct cccactggaa cttgcacaat caaaaaaac tctcctctct caagccgctt 180
ccaggagcgc atcacctgga gaagagcgc tcgctccccg cgccggccgc ggaagagcag 240
ccaggtagct gggggcgggg aggcgtacc ttctcccgtt cggtaagagc cacagcatct 300
ccccggagat tggccgtatc ccaccgtccg gcccccaggg tctcgcagcg gtgatgcata 360
tgtttcggag caatgatgga aggagaaaaag ccgctgtcgg tggcaactga aagtggggag 420
aggttgctgc agtagctggt gctgcagaat gcgagagtgga agaactgagc cccgctagat 480
totccatccc gctcagctctt cattaactgt ctgcaggagg taaaccgggg aaacagatat 540
gcactaacca ggggggtgcc aacctggatc tataactgtg aattcccac ggtggaaaat 600
ggtaaaaaa gacatgaatg gattcccagt caagaaatgc tcagccttcc aattttttaa 660
gaagcgggta cgaaggtgga tcaagagccc aatggtcagt gtggacaagc atcagagtec 720
cagcctgaag tacaccggct cctccatggt gcacatccct ccaggggagc cagacttoga 780
gtcttccttg tgtcaaacat gcctgggtga acatgcttct caaagagggg ttctccctca 840
ggagaacgag tcagtgtcat gggaaactca atctgggtgt gaagtgaag agccatgtaa 900
tcatgccaac atcctgacca agcccagtc aagaaccttc tggactaatg atgatccagc 960
tttcatgaag cagaggagga tgggtctgaa cgactttatt cagaagattg ccaataactc 1020
ctatgcatgc aaacaccctg aagttcagtc catcttgaag atctcccac ctcaggagcc 1080
tgagcttatg aatgccaac cttctcctcc accaagtcct tctcagcaa tcaaccttgg 1140
cccgtctgtc aatcctcatg ctaaacctc tgactttcac ttcttgaaag tgatcggaaa 1200
gggcagtttt ggaaggttcc ttctagcaag acacaaggca gaagaagtgt tctatgcagt 1260
caaagtttta cagaagaaag caatcctgaa aaagaagag gagaagcata ttatgtcggg 1320
gcggaatggt ctgttgaaga atgtgaagca cctttctctg gtgggccttc acttctcttt 1380
ccagactgct gacaaattgt actttgtcct agactacatt aatgggtggag agttgttcta 1440
ccatctccag agggaaacgt gcttctctgga accacgggct cgtttctatg ctgctgaaat 1500
agccagtgcc ttgggtacc tgcattcact gaacatgctt tatagagact taaaaccaga 1560
gaatattttg ctgattcac agggacacat tgtccttact gacttcggac tctgcaagga 1620
gaacattgaa cacaacagca caacatccac cttctgtggc acgcccggagt atctcgcacc 1680
tgagggtgct cataagcagc cttatgacag gactgtggac tgggtggtgcc tgggagctgt 1740
cttgatgag atgctgtatg gcctgccgct tttttatagc cgaaacacag ctgaaatgta 1800
cgacaacatt ctgaacaagc ctctccagct gaaaccaaatt attacaaatt ccgcaagaca 1860
cctcctggag ggcctcctgc agaaggacag gacaaagcgg ctcggggcca aggatgactt 1920

```

-continued

```

catggagatt aagagtcacg tcttcttctc ctttaattaac tgggatgatc tcattaataa 1980
gaagattact ccccctttta acccaaatgt gagtgggccc aacgacctac ggcactttga 2040
ccccgagttt accgaagagc ctgtcccaaa ctccattggc aagtcccctg acagcgtcct 2100
cgtcacagcc agcgtcaagg aagctgccga ggctttccta ggcttttctt atgcgcctcc 2160
cacggactct ttcctctgaa ccctgttagg gcttggtttt aaaggatltt atgtgtgttt 2220
ccgaatgttt tagttagcct tttgggtggag ccgccagctg acaggacatc ttacaagaga 2280
atttgacatc ctctggaagc ttagcaatct tattgcacac tgttcgctgg aagctttttg 2340
aagagccatc tctcctcagt gagctcatga ggttttcatt tttattcttc cttccaacgt 2400
ggtgctatct ctgaaacgag cgtagagtg ccgccttaga cggaggcagg agtttctgta 2460
gaaagcggac gctgttctaa aaaaggtctc ctgcagatct gtctgggctg tgatgacgaa 2520
tattatgaaa tgtgcctttt ctgaagagat tgtgttagct ccaaagcttt tcctatcgca 2580
gtgtttcagt tctttatltt cccttggtga tatgctgtgt gaaccgtcgt gtgagtgtgg 2640
tatgcctgat cacagatgga ttttggtata agcatcaatg tgacacttgc aggacactac 2700
aacgtgggac attgtttgtt tcttccatat ttggaagata aatttatgtg tagacttttt 2760
tgtaagatag ggttaataac taaaatltat tgaatggto ttgcaatgac tctattcag 2820
atgcttaaaag aaagcattgc tgcatacaat atttctatlt ttgaaaaggg tttttatgga 2880
ccaatgcccc agttgtcagt cagagccgtt ggtgttttct attgtttaaa atgtcacctg 2940
taaaatgggc attatlttatg tttttttttt tgcattcctg ataattgtat gtattgtata 3000
aagaacgtct gtacattggg ttataacact agtatatltta aacttacagg cttatlttga 3060
atgtaaaacca ccattlttaat gtactgtaat taacatggtt ataatacgtta caatccttcc 3120
ctcatcccat cacacaactt tttttgtgtg tgataaaactg atlttggttt gcaataaaac 3180
cttgaaaaaat atttaccatat aaaaaaaaa 3208

```

<210> SEQ ID NO 7

<211> LENGTH: 5758

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 7

```

tttctgtact ctgggtgact cagagagggg agagattcag ccagcacact cctcgcgagc 60
aagcattact ctactgactg gcagagacag gagaggtaga tgtccacgcc cacagaccct 120
ggtgcgatgc cccaccagg gccttcgccg gggcctgggc cttcccctgg gccaatctt 180
gggcctagtc caggaccagg accatcccca ggttccgtcc acagcatgat ggggccaagt 240
cctggacctc caagtgtctc ccatectatg ccgacgatgg ggtccacaga cttcccacag 300
gaaggcatgc atcaaatgca taagccatc gatggtatac atgacaaggg gattgtagaa 360
gacatccatt gtggatccat gaagggcact ggtatgagac cacctcacc aggcatgggc 420
cctccccaga gtccaatgga tcaacacagc caaggttata tgtcaccaca cccatctcca 480
ttaggagccc cagagcacgt ctccagccct atgtctggag gaggccaac tccacctcag 540
atgccaccaa gccagccggg ggcctcctc ccaggtgato cgcaggccat gagccagccc 600
aacagaggtc cctcaccttt cagtctctgc cagctgcatc agcttcgagc teagatltta 660
gcttataaaa tctgtggccc aggccagccc ctccccgaaa cgtgcagct tgcagtccag 720

```


-continued

gggaaaagga cgttgccctgg cttgcagcaa caacagcagc agcaacagca gcagcagcag	780
cagcagcagc agcagcagca gcagcaacag cagccgcagc agcagccgcc gcaaccacag	840
acgcagcaac aacagcagcc ggccttctgt aactacaaca gaccatctgg cccggggccg	900
gagctgagcg gcccgagcac cccgcagaag ctgcccgtgc ccgcccggcg cggcccggccc	960
tcgcccggcg cccccgcagc cgcgcagccg cccgcccggcg cagtgccggg gccctcagtg	1020
ccgcagccgg cccccgggca gccctcgccc gtctctcagc tgcagcagaa gcagagccgc	1080
atcagcccca tcagaaaa ccaaggcctg gaccccgctg aaattctgca agagcgggaa	1140
tacagacttc agccccgat agctcatagg atacaagaac tggaaaatct gcctggctct	1200
ttgccaccag atttaagaac caaagcaacc gtggaactaa aagcacttcg gttactcaat	1260
ttccagcgtc agctgagaca ggaggtggtg gcctgcctgc gcagggacac gaccctggag	1320
acggctctca actcctaaagc atacaacagg agcaagccgc agactctgag agaagctcgc	1380
atgaccgaga agctggagaa gcagcagaag attgagcagg agaggaaacg ccgtcagaaa	1440
caccaggaat acctgaacag tattttgcaa catgcaaaag attttaagga atatcatcgg	1500
tctgtggccg gaaagatcca gaagctctcc aaagcagtgg caacttgca tgccaacact	1560
gaaagagagc agaagaagga gacagagccg attgaaaagg agagaatcgc gcgactgatg	1620
gctgaagatg aggagggtta tagaaaactg attgatcaaa agaaagacag gcgtttagct	1680
taccttttgc agcagaccga tgagtatgta gccaatctga ccaatctggt ttgggagcac	1740
aagcaagccc aggcagccaa agagaagaag aagaggagga ggaggaaagaa gaaggctgag	1800
gagaatgcag aggggtgggga gtctgcctcg ggaccggatg gagagcccat agatgagagc	1860
agccagatga gtgacctccc tgtcaaagtg actcacacag aaaccggcaa ggttctgttc	1920
ggaccagaag caccctaaagc aagtcagctg gacgcctggc tggaaatgaa tcctggttat	1980
gaagttgccc ctgatctgca cagtgaagag agtgattctg attatgagga agaggatgag	2040
gaagaagagt ccagtaggca ggaaccgaa gagaaaatac tcctggatcc aaatagcgaa	2100
gaagtttctg agaaggatgc taagcagatc attgagacag ctaagcaaga cgtggatgat	2160
gaatacagca tgcagtacag tgcccaggggc tcccagtcct actacaccgt ggctcatgcc	2220
atctcggaga ggggtggagaa acagtctgccc ctcttaatta atgggaccct aaagcattac	2280
cagctccagg gcctggaatg gatggtttcc ctgtataata acaacttgaa cggaatctta	2340
gccgatgaaa tggggcttgg aaagaccata cagaccattg cactcatcac ttatctgatg	2400
gagcacaata gactcaatgg cccctatctc atcattgttc ccccttcgac tctatctaac	2460
tggacatagc aatttgacaa atgggctcct tctgtgggta agatttctta caagggtact	2520
cctgccatgc gtcgctccct tgtccccag ctacggagtg gcaaatcaa tgcctcttg	2580
actacttatg agtatattat aaaagacaag cacattcttg caaagattcg gtggaatac	2640
atgatagtgg acgaaggcca ccgaatgaag aatcaccact gcaagctgac tcaggtcttg	2700
aacactcact atgtggcccc cagaaggatc ctcttgactg ggaccccgct gcagaataag	2760
ctccctgaac tctgggcccct cctcaacttc ctctcccaa caatttttaa gagctgcagc	2820
acatttgaac aatggttcaa tgctccattt gccatgactg gtgaaagggt ggacttaaat	2880
gaagaagaaa ctatattgat catcaggcgt ctacataagg tgtaagacc atttttacta	2940
aggagactga agaagaaggt tgaatcccag cttcccgaag aagtgaata tgtgatcaag	3000

-continued

tgtagcatgt cagctctgca gaagattctg tatcgccata tgcaagccaa ggggatcctt 3060
ctcacagatg gttctgagaa agataagaag gggaaaggag gtgctaagac acttatgaac 3120
actattatgc agttgagaaa aatctgcaac cacccatata tgtttcagca cattgaggaa 3180
tcctttgtctg aacacctagg ctattcaaat ggggtcatca atggggctga actgtatcgg 3240
gcctcagga agtttgagct gcttgatcgt attctgcca aattgagagc gactaatcac 3300
cgagtgtgc tttctgcca gatgacatct ctcagacca tcatggagga ttattttgct 3360
tttcggaact tcctttacct acgccttgat ggcaccacca agtctgaaga tcgtgtgct 3420
ttgctgaaga aattcaatga acctggatcc cagtatttca ttttcttctg gacacaaga 3480
gctggtggcc tgggcttaaa tcttcaggca gctgatacag tggatcattt tgacagcgac 3540
tggaatcctc atcaggatct gcaggcccaa gaccgagctc accgcatcgg gcagcagaac 3600
gaggtccggg tactgaggct ctgtaccgtg aacagcgtgg aggaaaagat cctcgcggcc 3660
gcaaaataca agctgaacgt ggatcagaaa gtgatccagg cgggcatggt tgacaaaaag 3720
tcttcaagcc acgagcggag ggcattcctg caggccatct tggagcatga ggaggaaaat 3780
gaggaagaag atgaagtacc ggacgatgag actctgaacc aaatgattgc tcgacgagaa 3840
gaagaatttg acctttttat cgggatggac atggaccggo ggagggaaga tgcccggaac 3900
ccgaaacgga agccccgttt aatggaggag gatgagctgc cctcctggat cattaaggat 3960
gacgctgaag tagaaaggct cacctgtgaa gaagaggagg agaaaatatt tgggaggggg 4020
tcccgccagc gccgtgacgt ggactacagt gacgccctca cggagaagca gtggctaagg 4080
gccatcgaag acggcaattt ggaggaaatg gaagaggaag tacggcttaa gaagcgaaaa 4140
agacgaagaa atgtggataa agatcctgca aaagaagatg tggaaaaagc taagaagaga 4200
agaggccgcc ctcccgtgaa gaaaactgca ccaaatcccc ccaaaactgac aaagcagatg 4260
aacgctatca tcgatactgt gataaactac aaagataggt gtaacgtgga gaaggtgccc 4320
agtaattctc agttgaaat agaaggaaac agttcagggc gacagctcag tgaagtcttc 4380
attcagttac cttcaaggaa agaattacca gaatactatg aattaattag gaagccagtg 4440
gatttcaaaa aataaagga aaggattcgt aatcataagt accggagcct aggcgacctg 4500
gagaaggatg tcattgctct ctgtcacaac gctcagacgt tcaacctgga gggatcccag 4560
atctatgaag actccatcgt cttacagtca gtgtttaaga gtgcccggca gaaaattgcc 4620
aaagaggaag agagtgagga tgaaagcaat gaagaggagg aagaggaaga tgaagaagag 4680
tcagagtcgg aggcaaaatc agtcaagggt aaaattaagc tcaataaaaa agatgacaaa 4740
ggccgggaca aagggaaaagg caagaaaagg ccaaatcgag gaaaagccaa acctgtagtg 4800
agcgattttg acagcgatga ggagcaggat gaacgtgaac agtcagaag aagtgggacg 4860
gatgatgagt gatcagtatg gaccttttct cttggtagaa ctgaattcct tcctcccctg 4920
tctcatttct acccagtgag ttcatttctc atataggcac tgggttggtt ctatatcctc 4980
atcgtctata aactagcttt aggatagtgc cagacaaaca tatgatatac tgggttaaaa 5040
aacacacaca tacacaaata tttgtaacat attgtgacca aatgggcttc aaagattcag 5100
attgaaacaa acaaaaagct tttgatggaa aatatgtggg tggatagtat atttctatgg 5160
gtgggtctaa tttggtaacg gtttgattgt gcctggtttt atcacctggt cagatgagaa 5220
gatttttctc tttttagca ctgataacca ggagaagcca ttaaaagcca ctggttattt 5280

-continued

```

tatttttcat caggcaattt tcgaggtttt tatttggtcg gtattgtttt tttacactgt 5340
ggtacatata agcaacttta ataggtgata aatgtacagt agttagattt cacctgcata 5400
tacatttttc cattttatgc tctatgatct gaacaaaage tttttgaatt gtataagatt 5460
tatgtctact gtaaacattg ctttaatttt ttgctcttga tttaaaaaaa agttttgttg 5520
aaagcgctat tgaatattgc aatctatata gtgtattgga tggcttcttt tgtcacctcg 5580
atctcctatg ttaccaatgt gtatcgtctc cttctccta aagtgtaact aatctttgct 5640
ttctttgcac aatgtctttg gttgcaagtc ataagcctga ggcaataaaa attccagtaa 5700
tttcgaagaa tgtggtgttg gtgctttctt aataaagaaa taatttagct tgacaaaa 5758

```

```

<210> SEQ ID NO 8
<211> LENGTH: 837
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

```

```

<400> SEQUENCE: 8

```

```

gctcctcctg cacacctccc tcgctctccc acaccactgg caccaggccc cggacacccc 60
ctctgctgca ggagaatggc tactcatcac acgctgtgga tgggactggc cctgctgggg 120
gtgctggggc acctgcaggc agcaaccggag gcccaggctc cgtgcagcc caacttccag 180
caggacaagt tctggggcgt ctggttcagc gggggcctcg cctccaaact gagctggctc 240
cgggagaaga agggcgcggt gtccatgtgc aagtctgtgg tggccctgc cacgatggt 300
ggcctcaacc tgacctccac cttcctcagg aaaaaccagt gtgagaccg aacctgctg 360
ctgcagcccg cggggtccct cggtcctac agctaccgga gtcccactg gggcagcacc 420
tactccgtgt cagtgggtgga gaccgactac gaccagtaag cgctgctgta cagccagggc 480
agcaagggcc ctggcgagga cttccgcatg gccaccctct acagccgaac ccagaccccc 540
agggtgagat taaaggagaa attcaccgcc ttctgcaagg cccagggtt cacagaggat 600
accattgtct tctgcccga aaccgataag tgcattgacg aacaatagga ctcccagggg 660
ctgaagctgg gatcccggcc agccaggtga cccccacgct ctggatgtct ctgctctgtt 720
cctccccga gccctgccc cggctcccc ccaaagcaac cctgcccact caggettcat 780
cctgcacaat aaactccgga agcaagtcag taaaaaaaa aaaaaaaaa aaaaaaa 837

```

```

<210> SEQ ID NO 9
<211> LENGTH: 6001
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

```

```

<400> SEQUENCE: 9

```

```

caaggaggga tcccacagat gtcacagggc tgtcacagag ctgtggtggg aatttcccat 60
gagaccccgc cctggctgga gtcaccgcac tcctgtgttt gacctgaagt cctctcgagc 120
tgcagaagcc tgaagaccaa ggagtggaaa gttctccggc agcctgaga tctcaagagt 180
gacatttggt agaccagcta atttgattaa aattctcttg gaatcagett tgetagtatc 240
atacctgtgc cagatttcat catgggaaac agctgttaca acatagtagc cactctgttg 300
ctggtcctca actttgagag gacaagatca ttgcaggatc cttgtagtaa ctgccagct 360
ggtacattct gtgataataa caggaatcag atttgagtc cctgtcctcc aatagtttc 420
tccagcgcag gtggacaaag gacctgtgac atatgcaggc agtgtaaagg tgttttcagg 480

```

-continued

accaggaagg agtgttccct caccagcaat gcagagtgtg actgcactcc agggtttcac	540
tgcttggggg caggatgcag catgtgtgaa caggattgta aacaaggta agaactgaca	600
aaaaaaggtt gtaaagactg ttgctttggg acatttaacg atcagaaaacg tggcatctgt	660
cgaccctgga caaactgttc tttggatgga aagtctgtgc ttgtgaatgg gacgaaggag	720
agggacgtgg tctgtggacc atctccagcc gacctctctc cgggagcacc ctctgtgacc	780
cgcctgccc ctgagagaga gccaggacac tctccgcaga tcatctcctt ctttcttgcg	840
ctgacgtcga ctgctgtgct ttctctgctg ttcttctca cgctccgttt ctctgttgtt	900
aaacggggca gaaagaaact cctgtatata ttcaaacac catttatgag accagtacaa	960
actactcaag aggaagatgg ctgtagctgc cgatttccag aagaagaaga aggaggatgt	1020
gaactgtgaa atggaagtca atagggtgtg tgggacttct ttgaaaagaa gcaaggaaat	1080
atgagtcatc cgctatcaca gctttcaaaa gcaagaacac catctacat aataccagg	1140
attcccccaa cacacgttct tttctaaatg ccaatgagtt ggcctttaa aatgcaccac	1200
ttttttttt tttttgacag ggtctcactc tgtcaccag getggagtgc agtggcacca	1260
ccatggctct ctgagcctt gacctctggg agctcaagt atcctctgc ctcaagtctc	1320
tgagtagctg gaactacaag gaagggccac cacacctgac taactttttt gtttttgtt	1380
tggtaaagat ggcatttcac catgtgtgac aggtctgtct caaactccta ggttcacttt	1440
ggctcccaa agtctgtgga ttacagacat gaactgccag gcccgccaa aataatgac	1500
cacttttaac agaacagaca gatgaggaca gagctggtga taaaaaaaa aaaaaaaaaag	1560
cattttctag ataccactta acaggtttga gctagttttt ttgaaatcca aagaaaatta	1620
tagtttaaat tcaattacat agtccagtgg tccaactata attataatca aatcaatgc	1680
aggtttgttt tttggtgcta atatgacata tgacaataag ccacgagtg cagtaagtac	1740
ccgactaaag tttccgtggg ttctgtcatg taacacgaca tgctccaccg tcagggggga	1800
gtatgacag agtgcctgag tttagggtca aggacaaaa acctcaggcc tggaggaaat	1860
tttgaaaga gttcaagtgt ctgtatatcc tatggtcttc tccatctca cacctctgc	1920
ctttgtctg ctccctttta agccaggta cattctaaaa attcttaact ttaacataa	1980
tattttatc caaagccaat aaatgaactg catatgatag gtatgaagta cagtgagaaa	2040
attaacacct gtgagctcat tgtcctacca cagcaactaga gtgggggccc caaaactccc	2100
atggccaaac ctgggtgacc atttgcctt gttgtctgt tggtttgett gagacagtct	2160
tgctctgtt cccaggtctg aatggagtgg ctattcacag gcacaatcat agcacacttt	2220
agccttaaac tectgggctc aagtgatcca cccgctcag tctcccaagt agctgggatt	2280
acaggtgcaa acctggcatg cctgccattg tttggcttat gatctaagga tagctttta	2340
aattttattc attttattt tttttgagac agtgtctcac tctgtctccc aggtggaggt	2400
acagtgtgac aatcttggat caccgcctcc cagtttcaag tgatctccct gctcagcct	2460
cctaagtagc tgggactaca ggtatgtgcc accacgctg gctaattttt atatttttag	2520
tagagacggg gtttcacat gttgtccagg ctggtctcaa actcctgacc tcaggtgatc	2580
tgcccacctc tgctcccaa agtctgtgga ttacaggcat gagccaccat gcttgccat	2640
ttcttacct tttgtatgac atgctattg caagcttgcg tgcctctgct ccatgttatt	2700
ttactctggg atttagggtg agggagcagc ttctatttgg aacattggcc atcgcattgg	2760

-continued

aaatgggtat	ctgtcacttc	tgetcctatt	tagttgggtc	tactataacc	tttagagcaa	2820
atcctgcagc	caagccaggc	atcaataggg	cagaaaagta	tattctgtaa	ataggggtga	2880
ggagaagata	tttctgaaca	atagtctact	gcagtaacca	attgcttttc	aaagtggctg	2940
ttctaagtga	ctcecgtcag	tcatataagt	gtcatgtaag	tatcccattg	atccacatcc	3000
ttgctaccct	ctggctactat	caggtgccct	taattttgcc	aagccagtgg	gtatagaatg	3060
agatctcact	gtggctcttag	tttgcatttg	cttggttact	gatgagcacc	ttgtcaaata	3120
tttatatacc	atttgtgttt	atTTTTTTaa	ataaaatgct	tgctcatgct	tttttgccca	3180
tttgcaaaaa	aacttggggc	cggtgcagct	ggctcatgcc	tgtagtccca	gctctttggg	3240
aggccaaggt	gggcagatcg	cttgagccca	ggagtccgag	accagccttg	gcaacatggc	3300
gaaacctgt	ctttacaaaa	aatacaaaaa	ttagccgggt	gtggtgggtg	gcacctgaag	3360
tcccagctac	tcagtagggt	cgctttgagc	ctgggaggca	gaggttgtag	tgagctggga	3420
ccgcatcact	acacttcagc	ctgggcaaca	gagaaaaacc	ttttctcaga	aacaaacaaa	3480
cccaaagtgt	gttgtttgtc	ctgattccta	aaaggtcttt	atgtattcta	gataataatc	3540
tttggtcagt	tatatgtgtt	aaaaaatatc	ttctttgtgg	ccaggcacgg	tagctcacac	3600
ctgtaatccc	agcactttgc	gggctgaggg	tgggtggatc	atctgaggtc	aagagttcaa	3660
gatcagcctg	gccaacacag	tgaaacccca	tctctactaa	acatgtacaa	aacttagctg	3720
ggtaggtgg	cggtgcctcg	taaccccagc	tgctccagag	gctgtggcag	aagaatcgct	3780
tgaacccagg	aggcagaggt	tgcagcgcgc	caagattgtg	ccattgcact	ccagactggg	3840
tgacaagagt	gaaattctgc	ctatctatct	atctatctat	ctatctatct	atatatatat	3900
atatatatcc	tttgaattt	atTTTTTccct	ttttaaatt	ttttataaaa	ttctttttta	3960
ttttatttt	tagcagaggt	gaggtttctg	aggtttcatt	atggtgcccc	ggtgggtctt	4020
gaactcctga	gctcaagtga	tcctcccacc	tcagccttcc	aaagtgtctg	aattgcagac	4080
atgagccacc	gcccctctcc	tgTTTTTctc	taattaatgg	tgtctttctt	tgtctttctg	4140
gtaataagca	aaaagtctct	catttgattt	ggttaaattt	ataactgttt	tctcatatgg	4200
ttaacatttt	ttcttgcctg	gctaagaaa	tccttttctg	cccaatacta	taaagaggtt	4260
tgcccacatt	ttattccaaa	agttttaagt	tttgtcttcc	atcttgaagt	ctaatgtatc	4320
aggaactggc	ttttgtgcct	gttgggaggt	agtgatccaa	ttccatgtct	tgcatgtagg	4380
taaccactgg	tcctctgcgc	atgtattcaa	taegtctctc	ttctctctcg	ggtctgcaat	4440
ctcacctacc	atccatcaag	tttccatagg	gccatgggtc	tgcttctggg	ctccctgttc	4500
tgttccattg	teaatttgtc	tatcctgtgc	cagtatccca	ctgtgtttat	tacaatagct	4560
ttgtaacagc	tctcgatata	cggtaggaca	tctccctcca	ccttcttttt	ctacttcaga	4620
agtgtcttag	ctaggtcagg	cacggtggct	caegcctgta	atcccagcac	tttgggaggc	4680
cgacgcggat	ggatccctcg	aggtcaggag	ttttgagaca	gcctggccaa	catggtgaaa	4740
ccccatctct	actaaaaaat	acaaaaatta	gtcaggcatg	gtggcatgtg	cctgtaatcc	4800
cagctatttg	ggaggctgag	cccggagaat	tgcttgaacc	cggggggcgg	aggttgcagt	4860
gagccgagat	cgtaccattg	cactccagcc	tgggtgacag	agcgaactc	tgtctcagga	4920
aaaaaaagaa	aagagatgtc	ttggttattc	ttggttcttt	attattcaat	ataaatTTTA	4980
gaagctgaat	ttgaaaagat	ttggattgga	atTtcattaa	atctacaggt	caatttaggg	5040

-continued

```

agagttgata attttacaga attgagtcac ctggtgttcc aataagaata agagaacaat 5100
tattggctgt acaattcttg ccaaatagta ggcaaagcaa agcttaggaa gtatactggt 5160
gccatttcag gaacaaagct aggtgcgaat atttttgtct ttctgaatca tgatgctgta 5220
agttctaaag tgatttctcc tcttggtctt ggacacatgg tgtttaatta cctactgctg 5280
actatccaca aacagaaaga gactggtcat gccccacagg gttgggggat ccaagataat 5340
ggagcgaggc tctcatgtgt cctaggttac acaccgaaaa tccacagttt attctgtgaa 5400
gaaaggaggc tatgtttatg atacagactg tgatattttt atcatagcct attctgggat 5460
catgtgcaaa agctataaat gaaaacaca ggaacttggc atgtgagtca ttgctcccc 5520
taaatgacaa ttaataagga aggaacattg agacagaata aatgatccc cttctgggtt 5580
taatttagaa agttccataa ttaggtttta tagaaataaa tgtaaattc tatgattaaa 5640
aataaattag cacatttagg gatacacaaa ttataaatca ttttctaaat gctaaaaaca 5700
agctcaggtt ttttccagaa gaaagtttta attttttttc tttagtgga gatatactc 5760
tgacggaaag ttttgatgtg aggggcccgat gactataaag tgggcatctt cccccacag 5820
aagatgttc catctgtggg tgagaggtgc ccaccgcgc tagggcaggt tacatgtgcc 5880
ctgtgtgtgg taggacttgg agagtgatct ttatcaacgt ttttatttaa aagactatct 5940
aataaacac aaaactatga tgttcacagg aaaaaagaa taagaaaaaa agaaaaaaa 6000
a 6001

```

<210> SEQ ID NO 10

<211> LENGTH: 1336

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 10

```

gagagacaca gagtccggca ttggtcccag gcagcagtta gccccccgcc cgctgtgtg 60
tccccagagc catggagaga gccagtctga tccagaagcc caagctggca gagcaggccg 120
aacgctatga ggacatggca gccctcatga aaggcgcctg ggagaaggcc gagagctct 180
cctgcgaaga gcgaaacctg ctctcagtag cctataagaa cgtggtgggc gcccagaggg 240
ctgcctggag ggtgctgtcc agtattgagc agaaaagcaa cgaggaggcc tccgaggaga 300
aggggcccga ggtgctgtgag taccgggaga aggtggagac tgagctccag ggcgtgtgct 360
acaccgtgct gggcctgctg gacagccacc tcatcaagga ggccggggac gccgagagcc 420
gggtcttcta cctgaagatg aagggtgact actaccgcta cctggccgag gtggccaccg 480
gtgacgacaa gaagcgcac c attgactcag cccggtoagc ctaccaggag gccatggaca 540
tcagcaagaa ggagatgccg cccaccaacc ccatccgctt gggcctggcc ctgaactttt 600
ccgtcttcca ctacgagatc gccaacagcc ccgaggagcc catctctctg gccaaagacca 660
ctttcgacga ggcctatggt gatctgcaca cctcagcga ggactcctac aaagacagca 720
ccctcatcat gcagctgctg cgagacaacc tgacactgtg gacggccgac aacgccgggg 780
aagagggggg cgaggctccc caggagcccc agagctgagt gttgcccgcc accgccccgc 840
cctgccccct ccagtccccc accctgccga gaggactagt atggggtggg aggccccacc 900
cttctccccct aggcctgttt cttgctccaa agggctccgt ggagagggac tggcagagct 960
gaggccacct ggggctgggg atcccactct tcttcagct gttgagcga cctaaccact 1020

```

-continued

```

ggtcgatgcc ccaccctgc tctccgcacc cgcttctccc cgaccccagg accaggctac 1080
tttccccctc ctcttgctcc cctcctgccc ctgctgctcc tgatcgtagg aattgaggag 1140
tgtcccgcct tgtggctgag aactggacag tggcaggggc tggagatggg tgtgtgtgtg 1200
tgtgtgtgtg tgtgtgtgtg tgtgcccgcg cgccagtgca agaccgagat tgagggaaaag 1260
catgtctgct ggggtgtgacc atgtttctcc tcaataaagt tcccctgtga cactcaaaaa 1320
aaaaaaaaa aaaaaa 1336

```

<210> SEQ ID NO 11

<211> LENGTH: 2240

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 11

```

ggagggcagc cagcagcttc cccttctctg ccctgctcca ggcaccaggc tctttcccct 60
tcagtgtctc agaggagggg acggcagcac catggacccc cgcttgctcca ctgtccgcca 120
gacctgtgct tgtttcaatg tccgcactgc aaccacggcc ctggccatct accatgtgat 180
catgagcgtc ttgttttcca tcgagcactc agtagagggt gcccatggca aggcgtcctg 240
caagctctcc cagatgggct acctcaggat cgctgacctg atctccagct tccctgctcat 300
caccatgctc ttcateatca gcctgagcct actgatcggc gtagtcaaga accgggagaa 360
gtacctgctg cccttccctg ccctgcaaat catggactat ctctgtgccc tgctcacctc 420
gctgggctcc tacattgagc tgcccgccta cctcaagttg gcctcccgga gccgtgctag 480
ctctccaag tccccctgta tgacgctgca gctgctggac ttctgctgta gcatcctgac 540
cctctgcagc tctacatgg aagtgcaccac ctatctcaac ttcaagtcca tgaaccacat 600
gaattacctc cccagccagg aggatattgc tcataaccag ttcatacaaga tgatgatcat 660
cttttccatc gccctcatca ctgtccttat cttcaaggto tacatgttca agtgcgtgtg 720
gcggtgctac agattgatca agtgcatgaa ctgggtggag gagaagagaa actccaagat 780
gctccagaag gtggctcctg cgtcctacga ggaagccctg tctttgcat cgaagacccc 840
agaggggggc ccagcaccac ccccaatact agaggtgtga ccctcgccag gccccagccc 900
cagtgtctgg aggggtggag ctgcctcata atctgctttt ttgctttggt gggccctgtg 960
gcctgggtgg gccctcccgc ccctccctgg caggacaatc tgcttgtgct tccctcgctg 1020
gcctgctcct cctgcagggc ctgtgagctg ctcaaaactg ggtcaacgct ttaggctgag 1080
tcactcctcg ggtctctcca taattcagcc caacaatgct tggtttattt caatcagctc 1140
tgacacttgt ttagacgatt ggccattcta aagttgggtg gtttgcgaag caactatcga 1200
cttgatcagt tcagccaagc aactgacaaa tcaaaaacc ccttgtcagt tcagtaaaat 1260
aatttggtca aacaacagtc tattgcattg atttataaat agttgtcagt tcacatagca 1320
atttaatcaa gtaatcatta attagtacc ccctatataa aaatatatgt aatcaatttc 1380
ttcaaatagc ttgcttcat gataatcaat tagccaacca tgagtcattt agaatagtga 1440
taaatagaat acacagaata gtgatgaaat tcaatttaa aatcacggtt agcctccaaa 1500
ccatttaatt caaatgaacc catcaactgg atgccaactc tggcgaatgt aggacctctg 1560
agtggctgta taattgtaa ttcaaatgaa attcatttaa acagttgaca aactgtcatt 1620
caacaattag ctccaggaaa taacagttat ttcataataa aacagtcctc tcaaacacac 1680

```

-continued

```

aattgttctg ctgaagagtt gtcatacaaca atccaatgct cacctattca gttgctctgt 1740
ggtcagtgtg gctgcataac agtggattcc atgaaaggag tcatttttagt gatgagctgc 1800
cagtccattc ccaggccagg ctgtegetgg ccatccattc agtcgattca gtcataggcg 1860
aatctgttct gcccgagggt tgtgggtcaag caaaaattca gccctgaaat caggcacatc 1920
tgttcgttgg actaaaccca caggttagtt cagtcaaagc aggcaacccc cttgtgggca 1980
ctgaccctgc cactggggtc atggcggttg tggcagctgg ggaggtttgg ccccaacagc 2040
cctcctgtgc ctgcttccct gtgtgtcggg gtctccagg gagctgacc agaggtggag 2100
gccacggagg cagggtctct ggggactgtc ggggggtaca gagggagaag gctctgcaag 2160
agtcctctgg caataccccc ttgtgtaatt gctttgtgtg cgacaggag gaagtttcaa 2220
taaagcagca acaagcttct 2240

```

<210> SEQ ID NO 12

<211> LENGTH: 3039

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 12

```

aggcgcagag gagggcgggtg ttgagaccgg cggagcggcg ggaccctag gtggcggagg 60
gacgctccgg gaaagcgagg ggcgctacga gctctggccc acgtgacctg cegggggcgg 120
gagcaggggg cgcgcggggc tcttgcgggtg cccctgcctt ggggaggggc cgtgaccacc 180
cgtctgtcgc ccgaggcggc gcgccctgca ccttcaccgc gtaccggga cccgccgc 240
cgcgggagaa atgttgctga agtctgctg aaagggccag agatgcaagg atttgggata 300
cattttgaac cttaagctg tctgacattg acctccttct attattaata aagaagaatc 360
aggagcttag gatgtattaa caccaactca ttaataact aaccggacaa tgttctacaa 420
acaattctac attgtaaagg actggattgg cacaaaataa aataatttta ttttattcag 480
cttataatat gactcgatgg aggaaaattt gataagcatg agagaagacc attcctttca 540
tgttcgttac agaatggaag cttcttgcct agagctggcc ttggaagggg aacgtctatg 600
taaatcagga gactgccgcg ctggcgtgtc attccttgaa gctgcagttc aagttggaac 660
tgaagaccta aaaacactta cgcctattta cagccagttg ggcaatgctt atttctattt 720
gcatgattat gccaaagcat tagaataatca ccatcatgat ttaacccttg caaggactat 780
tggagaccag ctggggggaag cgaaagctag tggtaatctg ggaaacacct taaaagttct 840
tgggaatttt gacgaagcca tagtttgttg tcagcgacac ctagatattt ccagagagct 900
taatgacaag gtgggagaag caagagcact ttacaatctt gggaatgtgt atcatgccaa 960
agggaaaagt tttggttgc ctggtcccca ggatgtagga gaatttccag aagaagtgag 1020
agatgctctg caggcagccg tggattttta tgaggaaaac ctatcattag tgactgcttt 1080
gggtgaccga ggggcacaag gacgtgcctt tggaaatctt ggaaacacac attacctcct 1140
tggcaacttc agggatgag ttatagetca tgagcagcgt ctcttattg caaaagaatt 1200
tggagataaa gcagctgaaa gaagagcata tagcaacctt ggaaatgcat atatatttct 1260
tggtgaattt gaaactgcct cggaatacta caagaagaca ctactgttgg cccgacagct 1320
taaagaccga gctgtagaag cacagtcttg ttacagtctt ggaaatacat atactttact 1380
tcaagactat gaaaaggcca ttgattatca tctgaagcac ttagcaattg ctcaagagct 1440

```


-continued

```

gaatgataga attggtgaag gaagagcatg ttggagctta ggaaatgcat acacagcact 1500
aggaatcat gatcaagcaa tgcattttgc tgaaaagcac ttggaaatct caagagaggt 1560
tggggataaa agtgggtgaac taacagcacg acttaatctc tcagaccttc aaatggttct 1620
tggctcagc tacagcacia ataactccat aatgtctgaa aatactgaaa ttgatagcag 1680
tttgaatggt gtacgccccca agttgggacg ccggcatagt atggaaaata tggaaacttat 1740
gaagttaaca ccagaaaagg tacagaactg gaacagtga attcttgcta agcaaaaacc 1800
tcttattgcc aaacctctcg caaagctact ctttgcacac agactgaagg ggaaaaata 1860
caaaacgaat tctccacta aagtctcca agatgccagt aattctattg accaccgaat 1920
tccaaattct cagaggaaaa tcagtgcaga tactattgga gatgaagggt tctttgactt 1980
attaagccga tttcaagca ataggatgga tgatcagaga tgttgcttac aagaaaagaa 2040
ctgccataca gcttcaacaa caactctctc cactccccct aaaatgatgc taaaaacatc 2100
atctgttctc gtggatccc ccaacacgga tgagtttta gatctctctg ccagctcaca 2160
gagtcgctg ctggatgacc agagggctag tttcagtaat ttgccagggc ttegtctaac 2220
acaaaacagc cagtcggatc ttagccacct gatgactaat gacaacaag aggctgatga 2280
agatttcttt gacatccttg taaaatgtca aggatccaga ttagatgatc aaagatgtgc 2340
tccaccacct gctaccacia agggctccgac agtaccagat gaagactttt tcagccttat 2400
tttacggtcc cagggaaaga gaatggatga acagagaggt cttttacaaa gagatcaaaa 2460
cagagacact gactttgggc taaaggactt tttgcaaat aatgctttgt tggagttaa 2520
aaattcaggg aaaaaatcgg cagaccatta gttactatgg atttattttt tttcctttca 2580
aacacggtaa ggaacaatc tattactttt ttccttaaaa ggagaattta tagcactgta 2640
atcacgctta aaatattttt agaatgatgt aaatagttaa ccttcagtag tctattaagg 2700
cattaatact tctctggaca tgcgcgcttg aggggtgagg ggtcctgtaa ggtgcttcat 2760
cgtctgtgat tactgcttg gatgtgttct ttggcagctt gtgagattac tttacctagt 2820
gtttataaag taggaagtta agtgaatcat agattagaat ttaatactct tatggaaata 2880
attttttaac atcttaattg acaatggcgt ttttttatac ataaccatgg atgtagtggg 2940
aaacaatggt gtttggtaaa aataatgtac ttgatcaatg taaaaaagta tataaaatag 3000
tcttactaaa aatctaggtt ttttttctc ccaaaaaaa 3039

```

<210> SEQ ID NO 13

<211> LENGTH: 7018

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 13

```

ggcgggcgcg ccgggcgcca ggtgtcggcg tcggcggcat tcggcggcga tggagcggcc 60
ctggggagct ggggacggcc tctcgcgctg gccccatggc ctggcctcc tctcctcct 120
gcagctgctg ccgcctcga cctcagcca ggaccggctg gacgcgccgc cgcgcgccgc 180
tgcgccgctg ccgcctggt ctggcccat cggggtgagc tgggggctgc gggcggccgc 240
agccgggggc gcttttccc gcggcgccgc ttggcgtcgc agcgcgccgc gcgaggacga 300
ggagtgcggc cgggtccggc acttcgtcgc caagctggcc aacaacacgc accagcatgt 360
gtttgatgat ctcagaggct cagtatcctt gtcctgggtt ggagatagca ctggggctcat 420

```

-continued

tctagtcttg actaccttcc atgtaccact ggtaattatg acttttgac agtccaagct	480
ataticgaagt gaggattatg ggaagaactt taaggatatt acagatctca tcaataaac	540
ctttatctgg actgaatttg gcatggctat tggctctgag aactctggaa aggtggtgtt	600
aacagcagag gtgtctggag gaagctgtgg aggaagaatc ttcagatcat cagattttgc	660
gaagaatttt gtgcaaacag atctcccttt tcatcctctc actcagatga tgtatagccc	720
tcagaattct gattatcttt tagctctcag cactgaaaat ggctgtggg tgtccaagaa	780
ttttggggga aaatgggaag aaatccacaa agcagtatgt ttggccaaat ggggatcaga	840
caacaccatc tcttttacia cctatgcaaa tggctcctgc aaagctgacc ttggggctct	900
ggaattatgg agaacttcag acttgggaaa aagcttcaaa actattggtg tgaatatcta	960
ctcatttggg cttgggggac gtttctcttt tgcctctgtg atggctgata aggataccac	1020
aagaaggatc cacgtttcaa cagatcaagg ggacacatgg agcatggccc agctcccctc	1080
cgtgggacag gaacagttct attctattct ggcagcaaat gatgacatgg tattcatgca	1140
tgtagatgaa cctggagaca ctgggtttgg cacaatcttt acctcagatg atcgaggcat	1200
tgtctattcc aagtcttttg accgacatct ctacactacc acaggcggag agacggactt	1260
taccacogtg acctccctcc gcggcgtcta cataacaago gtgctctccg aagataatc	1320
tatccagacc atgateactt ttgaccaagg aggaagggtg acgcacctga ggaagcctga	1380
aaacagtcaa tgtgatgcta cagcaaaaa caagaatgag tgcagccttc atattcatgc	1440
ttcctacagc atctcccaga aactgaatgt tccaatggcc ccactctcag agccgaatgc	1500
cgtaggcatt gtcattgctc atggtagcgt gggggatgcc atctcagtga tggttccaga	1560
tgtgtacatc tcagatgatg ggggttactc ctggacaaa atgctggaag gacccacta	1620
ttacaccatc ctggattctg gaggcatcat tgtggccatt gagcacagca gcegtcctat	1680
caatgtgatt aagtctcca cagacgaagg tcaatgctgg caaacctaca cgttcaccag	1740
ggacccatc tatttactg gcctagcttc agaacctgga gctaggctca tgaatatcag	1800
catttggggc ttcacagaat ctttctgac cagccagtgg gtctctaca ccattgattt	1860
taaagatc cttgaaagga actgtgaaga gaaggactat accatatggc tggcacactc	1920
cacagacctt gaagattatg aagatggctg cattttgggc tacaagaac agtttctgag	1980
gctacgcaag tcatcctgtg ttcagaatgg tcgagactat gttgtgacca agcagcctc	2040
catctgctc tgttccctgg aggaactttt ctgtgatctt ggctactacc gtccagaaaa	2100
tgactccaag tgtgtggaac agccagaact gaaggccac gacctggagt tttgtctgta	2160
cggaaagaaa gaacaccta caacaaatgg gtaccggaaa attccagggg acaaatgcca	2220
gggtggggta aatccagttc gagaagtaaa agacttgaaa aagaatgca caagcaactt	2280
tttgagtcg gaaaaacaga attccaagtc aaattctggt ccaattatcc tggccatcgt	2340
gggatgtatg ctggctcacg tcgtagcagg agtgcctatt gtgaagaaat atgtctgtgg	2400
gggaaggttc ctggctcacc gatactctgt gctgcagcag catgcagagg ccaatggtgt	2460
ggatggtgtg gatgctttgg acacagcctc ccacactaat aaaagtgtt atcatgatga	2520
ctcagatgag gacctcttg aatagctctt cagaggagct ggaccagca tggatggtgg	2580
aaccacagta cctcttacac tccctgtggc tccaactca ggaataaat tcccattgc	2640
gagggaccca gctctgttcc tgcctgtccc atcaagcca aaaggaccta cactaaagaa	2700

-continued

atgcagggtg ggggtgggga accctgagca cttttttaca attggctctg agaaaaaggg 2760
agacatttta aattctttaa cttcttattt ctctgctctg ctctttgcaa agtatgggct 2820
tttttgtttt tgttttttaa gggaaacgaa atggaattcg aaggacctt ttcactaacc 2880
ccaactctgt gtgttttgca tgggcctctg cccagggcat ctgccaaactc cagtatcagc 2940
tctcacagtg tacttgggtac catccctggg ctctgctggc gagacgaaac agctgtagag 3000
atgaaaacag gctgcagagg ctggcacagc ctggccggct tttctccatc tggggacagt 3060
cctactccaa gaacactgca caccagctcc tcacacagat cccacttact cttttttttt 3120
ttttcagaga ccacagacca cagtgatattt tcttttcctt tgtttaatta ggcaatacce 3180
ttgttaattg ccttttggca actaacttaa ccatgtgctt cccacacagt acatcaggaa 3240
aacttacagg gcaatatttt taacttgggg caggaagaag ggagcagcag agaattgact 3300
agatatagca cctattaaaa gagaactctt gcttcttctg agatttttca agctgtgctt 3360
tgtgtgtgtg ccagtagact tacgcaagga cagggtacaa acttagctgg aagtctgcc 3420
aggctgaatg atctcttccc tagagttgat tgtcgggtac acagtgtgaa cccccgaaga 3480
cggaaacctc cagtcttcca tgttcccttc ttaactgtcg tgtggctcgt tgetaaatca 3540
tgacaatggc tgccatctcg ctgctcttta ggttgctggt gtacatggaa ccaggactag 3600
agatttttcc agatttatag acttaaaaaa ttagaatttt attaccagge tttccttctc 3660
accctttttt tctgactttg ccaagtaatt tgttgacacg aaaattttgg aggaaccaat 3720
tgaaaacaca cttccagtct agatgatgct ttgtgtgata cattaagtcc ttattttgga 3780
ttaaagaag ttttccattt gatacttctc taaattaaat aaattataga atgtagtgg 3840
gtggattttg ggggtggccat atagtaatgg aaagctgcaa taattagttt taatacagct 3900
tgaatatttg ctatatagaa atatagtatg gaaagttttt ggtcttaatg tagctactgt 3960
gogggtcaca gtttctccca atgattatga ctgggacatt ctttggtaga taccatttgc 4020
tactagttta ttttgtggct agaaagttag ttttgtgtgt tttttttttt ttttatttga 4080
agtgccaaat taactttagt cagaatgtga gcagatggct aagttctctc ctcccagaa 4140
tggattaaca gctgcgtgga aagtggggga gagagtggat ggagactttt agagatgtta 4200
aaactgcagt agaataaat gagtcaggga gcttcagtta gaaaataaag ttgaggcagt 4260
ttttgtgaag ataatatggt tagggctgga gtgcaactagt ctttttgctt attcattttg 4320
catggtttta aaatataaaa taattccgaa gatcacaccag ctcaaaaatg aaaacgtcag 4380
cctctgcccc accctccctc ctgcccaag tgaatttggc actcagaaaa gaactgttta 4440
taccactcac ctttctccca gcatgtactc actgtgggca gatgcaccaa tacatggtaa 4500
tctcttact cattttaaga cgtaggaaac tcaatattct tctctaacca tatacgatag 4560
ggctctctgc ttttaatgat atctgggatt tctgtggaac ttggcaaatt ttcagagcac 4620
cttactcac ataatgtcat ttgaacctca caatgttctt gggatggagt cagtgttca 4680
gggtccccgt gtgtgtgata agcagtctg gctggctgtc ttcagaactc ttggaaatct 4740
ttacacatgc gagtgtaac cactttgagc aaggctgcct tctttagat gacttctgt 4800
tctttatgac agggatcagt ggcatttgtt tcctagcagt atttagcacc tttttgocac 4860
cttggtaaac agaaaattgt attttctgtt ctttcatggc tgaaaacaaa agtaatggga 4920
attttaata cgtttgcaga aactgccccct cccctcattg agggctactg ctcaagagtg 4980

-continued

```
caggagtgga ctctccactg atgggtctcc ctccccatcc tggtttccac cccgggctgg 5040
ctagctctgt tggtttgaag actgacagcc agcctggctc attctcatta ttggctagtt 5100
agetttcttt atcaacctgc tcaactcaca atgtgtgccc tcagccagag agtaagaaaag 5160
cccaaatctg ttacagcttc taaaaaaaaa gatttctaata ttgtcctact catggttagga 5220
gcattatctt tgaaggtaaa acatagtgta tcattgtgta aactcccagg cttgatgtag 5280
cagaagagat cttttctgga ggcttcagca atggaattta gcattataag agagattgga 5340
caaaccagtc caaagtggtc cgagttctta aatccaggta gggaactcac tcttctttct 5400
tctctggacc taattgggca ttgggcttta gtgagaccac agaccaggcc cgtctctcct 5460
gtaggctttt aattcaatgg caactctatt tcaaagaata aaagccttg gagagttgag 5520
gcagttctgg gggcgggctc aggagagtcc atagatcagc cgtaactgga acgtagaatc 5580
tacgtctgcc tctgaatgga cttcccacct cctctctctt gctctgatgc ttgcctctgg 5640
gcctctccat gcccaagggtg gtctttcctc cttgacaggc tggtaatgtg ctggccacct 5700
ccagctcctg catcgagtct gtaaaccaga gctggttctc atggccttcg tcacgatacc 5760
aggatacggg ggggagccca gggccatcca taccacccc agggtaacgg ggetggcctg 5820
gcattagtca ttatttagtt tccaggccaa ccatccagat agagattccc tctttccttt 5880
gagcagtgct ctcaagagct ccgtgcctgt ccacaatgac ctagagtgca tectgctcat 5940
tgteagtgta gccctctgcc cctatattca tccaggatac ttggaagtgc taaaatagga 6000
agggattcgg ctttcaactt tgctaccatc ttcctgaag caggaaaatg aacatggact 6060
taaatgttct ttgaaaaaac caaagtttta agatttgctg tgtgatgaag tgacagggag 6120
ggccggagtc agcaggtgcc agactttctg ttctgtctgc catgggttg tccagctcag 6180
gtagctctag gagcaccatc ctgcccctagc agagcccagg ccttgccctc atgaagcatc 6240
attgaaatag caggagcatg ttgatttctt ggttaggttg cattataata acaagagtca 6300
gaacattaat tggaaacaac ttgcagtatg cttttcttca caccagtaca ttcttaagtg 6360
tacttgttta taaggaataa cataaactaa tctgtacctt tatatatatg tgtgtgtaca 6420
tatatacata tataaactgt atagtgtaca tggtaatgat ttattgctat gcccagatc 6480
cttaatgtag ttctcatcct ccgcatgccc tcagccacaa gcgggtgact gactgttccc 6540
tgatgatttg gcccaacctc tgtgtttgga cctctagga ggagggtttt ggtcactactc 6600
tccttacctc cgtgcacaga aatgctcagg gtcccctatg gcctgttgtt cagccctctc 6660
tcttgctccc tttctgagca tgtggctcct ccccaggctg tgggacagct gccttcccac 6720
gaaagtgtaa agcagtatta agatcattac tgcattgtgc ctaaaaaacc aagttttcta 6780
ttcccttagg acagaaaatt gcatgtgagg tgggataatc gagtttcagt gaccacgctc 6840
agttacacat taaagccaga ccccatgata aaattccaca aaatggaaat aaaactcaaa 6900
tttctttagc attgtgtaaa taaactgaa tgtgtttaac tttgtactgg taattttctg 6960
tatatttgga atatttgggt taaaaataaa acagactgga ctttgttacc tgacctac 7018
```

<210> SEQ ID NO 14

<211> LENGTH: 1749

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 14

-continued

```

gtgacattgt ttgcaaaat cccaggcagc atggacctca gtcttctctg ggtacttctg    60
cccctagtca ccatggcctg gggccagtat ggcgattatg gatacccata ccagcagtat    120
catgactaca gcgatgatgg gtgggtgaat ttgaaccggc aaggcttcag ctaccagtgt    180
ccccaggggc aggtgatagt ggccgtgagg agcatcttca gcaagaagga aggttctgac    240
agacaatgga actacgcctg catgcccacg ccacagagcc tcggggaacc cacggagtgc    300
tgggtggagg agatcaacag ggctggcatg gaatggtacc agacgtgctc caacaatggg    360
ctggtggcag gattccagag ccgctacttc gagtcagtgc tggatcgga gtggcagttt    420
tactgttctc gctacagcaa gaggtgccca tattcctgct ggctaacaac agaatatcca    480
ggctactatg gtgaggaaat ggacatgatt tcctacaatt atgattacta tatccgagga    540
gcaacaacca ctttctctgc agtggaaaag gatcgccagt ggaagtcat aatgtgccgg    600
atgactgaat acgactgtga atttgcaaat gtttagattt gccacatacc aaatctgggt    660
gaaaggaaag gggccgggga caggagggtg tccacatag ttaacatcag ttggatctcc    720
tatagaagtt tctgctgctc tctttccttc tcctgagct ggtaactgca atgccactt    780
cctgggcctt tctgactagt atcacacttc taataaaatc cacaattaaa ccatgtttct    840
cacttttctc atgtttcata gcaactgctt tatatgactg atgatggctt ccttgccacac    900
cacatataca gtgcgatgc ttacagccgg gcttctggag caccagctgc agcctggcta    960
ctgcttttta ctgcagaatg aactgcaagt tcagcatagt ggaggggaga ggcagaactg   1020
gaggagaggt gcagtgaagg ttctctacag ctaagcctgt ttgaatgata cgtaggttcc   1080
ccaccaaaag caggetttct gccctgaggg acatcttccc actcccctgc tccacatgag   1140
ccatgcatgc ttagcaatcc aagtgcagag ctctttgctc caggagtgag gagactggga   1200
ggtgaaatgg ggaatggaa gggtttgag gcagagctga aaacagggtt ggaaggattt   1260
cctgaattag aagacaaacg ttagcatacc cagtaaggaa aatgagtga ggggccaggg   1320
gaaccctgga ggatcactct caaatgagat taaaaacaag gaagcagaga atggtcagag   1380
aatgggatc agattgggaa cttgtgggga tgagagtga caggttgaac tgggaagtgg   1440
aaaaaggagt ttgagtcact ggcacctaga agcctgccc aattcctag gaaggctggc   1500
agacaccctg gaaccctggg gagctactgg caaactctcc tggattgggc ctgattttt   1560
tgggtggaaa ggctgcctc gggatcaact ttctctctgt gtgtggctca ggagttctc   1620
tgcagagatg gcgctatctt tctctctct gtgatgtct gctcccaccc atttgtactc   1680
ttcattacaa aagaaataaa aatattaacg ttcactatgc tgaataataa aaaaaaaaaa   1740
aaaaaaaaa                                     1749

```

<210> SEQ ID NO 15

<211> LENGTH: 2478

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 15

```

gcagttgggt aaactcctct gtctcccgt catcttttca ttgctcgctc ccctccttc    60
cgcagacacc cggacctccc ctggcgccca gctccgoggc tccaacgggt ccagaacaa    120
gccggatttt tttttttct tcttggaat tggctttggt gtgtgttgc ctacctcct    180
cctcccctc caccacacag cccccccc gcctttttt tttttttt tttttttgag    240

```

-continued

acatggcccg ggcagtggt cctggaagag gaacaagtgt gggaaaaggg agaggaagcc	300
ggagctaaat gacaggatgc aggcgacttg agacacaaaa agagaagcgt tectctcgga	360
tccaggcatt gctcgcctgc tttcttttct ccaagacggg ctgaggattg tacagcteta	420
ggcggagtgt gggctctctg gatcgcttag attctctctt ttgctgcatt tcccccaag	480
tctcgtttct cccgcgtctg cctgcggacc cggagaaggg agaatggaga gggggctgcc	540
gctcctctgc gccgtgctcg ccctcgtcct cgccccggcc ggcgcttttc gcaacgataa	600
atgtggcgat actataaaaa ttgaaagccc cgggtacctt acatctctctg gttatcctca	660
ttcttatcac ccaagtgaat aatgcgaatg gctgattcag gctccggacc cataccagag	720
aattatgac aacttcaacc ctcacttca tttggaggac agagactgca agtatgacta	780
cgtggaagtc ttcgatggag aaaatgaaaa tggacatttt aggggaaagt tctgtgaaa	840
gatagcccct cctcctggtg tgtcttcagg gccatttctt tttatcaaat ttgtctctga	900
ctacgaaaca catgggtcag gattttccat acgttatgaa attttcaaga gaggtcctga	960
atgttcccag aactacacaa cacctagtgg agtgataaag tccccggat tcctgaaaa	1020
atatcccaac agccttgaat gcacttatat tgtctttgcg ccaaagatgt cagagattat	1080
cctggaattt gaaagctttg acctggagcc tgactcaaat cctccagggg ggatgttctg	1140
tcgctacgac cggctagaaa tctgggatgg attccctgat gttggccctc acattgggag	1200
ttactgtgga cagaaaacac caggtcgaat ccgacctca tcgggcattc tctccatggt	1260
tttttacacc gacagcgcga tagcaaaaaga aggtttctca gcaactaca gtgtcttgca	1320
gagcagtgtc tcagaagatt tcaaatgtat ggaagctctg ggcatggaat caggagaaat	1380
tcattctgac cagatecacag cttcttccca gtatagcacc aactggtctg cagagcgtc	1440
ccgcctgaac taccctgaga atgggtggac tcccgagag gattcctacc gagagtggat	1500
acaggtagac ttgggccttc tgcctttgt cacggctgtc gggacacagg gcgccatttc	1560
aaaagaaacc aagaagaaat attatgtcaa gacttacaag atcgactta gctccaacgg	1620
ggaagactgg atcaccataa aagaaggaaa caaacctgtt ctctttcagg gaaacacca	1680
ccccacagat gttgtggtg cagtattccc caaacactg ataactcgat ttgtccgaat	1740
caagcctgca acttgggaaa ctggcatatc tatgagattt gaagtatacg gttgcaagat	1800
aacagattat ccttgctctg gaatgttggg tatggtgtct ggacttattt ctgactocca	1860
gatcacatca tccaaccaag gggacagaaa ctggatgcct gaaaaacatcc gctggtaac	1920
cagtcgctct ggctgggac tccaccgc acctcattcc tacatcaatg agtggctcca	1980
aatagacctg ggggaggaga agatcgtgag gggcatcctc attcagggtg ggaagcaccg	2040
agagaacaag gtgttcatga ggaagttaa gatcgggtac agcaacaacg gctcggactg	2100
gaagatgatc atggatgaca gcaaacgcaa ggcgaagtct tttgagggca acaacaacta	2160
tgatacacct gagctgcgga cttttccage tctctccacg cgattcatca ggatctacc	2220
cgagagagcc actcatggcg gactggggct cagaatggag ctgctgggct gtgaagtgga	2280
agcccctaca gctggaccga ccactcccaa cgggaacttg gtggatgaat gtgatgacga	2340
ccaggccaac tgcccacgtg gaacaggtga tgacttccag ctcacaggtg gcaccactgt	2400
gctggccaca gaaaagccca cggtcataga cagcaccata caatcaggta tcaataaaaa	2460
tacgaaatgt gacagatt	2478

-continued

<210> SEQ ID NO 16

<211> LENGTH: 3372

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 16

taaaaccagg aagtgaagtc cccgagcacg ttagaaagcc tgacatggcc tgactcggga 60
cagctcagag cagggcagaa ctggggacac tctgggocgg ccttctgect gcatggacgc 120
tctgaagcca cctgtctctt ggaggaacca cgagcgaggg aagaaggaca gggactcgtg 180
tggcaggaag aactcagagc cgggaagccc ccattcacta gaagcactga gagatgcggc 240
cccctcgcag ggtctgaatt tctctgctgt gttcacaag atgcttttta tctttaactt 300
tttgttttcc ccaactccga ccccggcgtt gatctgeatc ctgacatttg gagctgceat 360
cttcttgagg ctgatcacca gacctcaacc cgtcttacct cttcttgacc tgaacaatca 420
gtctgtggga attgagggag gagcacggaa gggggtttcc cagaagaaca atgacctaac 480
aagttgtgct ttctcagatg ccaagactat gtatgaggtt ttccaagag gactcgtctg 540
gtctgacaat gggccctgct tgggatatag aaaaccaaac cagccctaca gatggctatc 600
ttacaacacag gtgtctgata gagcagagta cctgggttcc tgtctcttgc ataaagggta 660
taaatcatca ccagaccagt ttgtcggcat ctttgcctcag aataggccag agtggatcat 720
ctccgaattg gcttgttaca cgtactctat ggtagctgta cctctgtatg acacctggg 780
accagaagcc atcgtacata ttgtcaacaa ggctgatata gccatggtga tctgtgacac 840
accccaaaag gcattgggtg tgatagggaa tgtagagaaa ggcttcaccc cgagcctgaa 900
ggatgatcat cttatggacc cctttgatga tgacctgaag caaagagggg agaagagtgg 960
aattgagatc ttatccctat atgatgctga gaacctagga aaagagcact tcagaaaacc 1020
tgtgcctcct agcccagaag acctgagcgt catctgcttc accagtggga ccacaggtga 1080
ccccaaagga gccatgataa cccatcaaaa tattgtttca aatgctgctg cctttctcaa 1140
atgtgtggag catgcttatg agcccactcc tgatgatgtg gccatatact acctccctct 1200
ggctcatatg tttgagagga ttgtacagggc tgttgtgtac agctgtggag ccagagttgg 1260
attcttccaa ggggatattc ggttgcctgc tgacgacatg aagactttga agcccacatt 1320
gttcccgcg gtgcctcgac tcttaacag gatctacgat aaggtacaaa atgaggccaa 1380
gacacccttg aagaagtctt tgttgaagct ggctgtttcc agtaaattca aagagcttca 1440
aaagggtatc atcaggcatg atagtctctg ggacaagctc atctttgcaa agatccagga 1500
cagcctgggc ggaagggttc gtgtaattgt cactggagct gccccatgt ccaactcagt 1560
catgacattc ttccgggcag caatgggatg tcaggtgat gaagcttatg gtcaaacaga 1620
atgcacaggt ggctgtacat ttacattacc tggggactgg acatcaggtc acgttggggt 1680
gccccctggct tgcaattacg tgaagctgga agatgtggct gacatgaact actttacagt 1740
gaataatgaa ggagaggtct gcatcaaggg tacaacgtg ttcaaaggat acctgaagga 1800
ccctgagaag acacaggaag ccttgacagc tgatggctgg cttcacacag gagacattgg 1860
tcgctggctc ccgaatggaa ctctgaagat catcgaccgt aaaaagaaca ttttcaagct 1920
ggcccaagga gaatacattg caccagagaa gatagaaaat atctacaaca ggagtcacac 1980
agtgttacia atttttgtac acggggagag cttacggtca tccttagtag gagtgggtgt 2040

-continued

```

tctgacaca gatgtacttc cctcatttgc agccaagctt ggggtgaagg gctccttga 2100
ggaactgtgc caaaaccaag ttgtaagga agccatttta gaagacttgc agaaaattgg 2160
gaaagaaagt ggccttaaaa cttttgaaca ggtcaaagcc atttttcttc atccagagcc 2220
atthtccatt gaaaatgggc tcttgacacc aacattgaaa gcaaagcgag gagagcttcc 2280
caaatacttt cggaccctaaa ttgacagcct gtatgagcac atccaggatt aggataaggt 2340
acttaagtac ctgccggccc actgtgcact gcttgtgaga aaatggatta aaaactatcc 2400
ttacatttgt tttgccttcc ctccattttt tttttaacct gttaaactct aaagccatag 2460
cttttgtttt atattgagac atataatgtg taaacttagt tcccaataa atcaatcctg 2520
tctttcccat ctctgatgtt gctaataatta aggcttcagg gctactttta tcaacatgcc 2580
tgtcttcaag atcccagttt atgttctgtg tcttctctca tgatttccaa ccttaatact 2640
attagtaacc acaagttaa gggctcaaagg gaccctctgt gccttctctt ttgttttgtg 2700
ataaacataa cttgccaca gtctctatgc ttatttacct cttctactgt tcaactaag 2760
agatttttaa attctgaaaa actgcttaca attcatgttt tctagccact ccacaacca 2820
ctaaaatttt agtttttagcc tctactcat gtcaatcata tctatgagac aaatgtctcc 2880
gatgctcttc tgcgtaaat aaattgtgta ctgaaggaa aagtttgatc ataccaaca 2940
tttctaaac tctctagtta gatctctgac ttgggagat taaaaattgg gtctatgaca 3000
tattgtccaa aaggaatgct gttcttaaag cattatttac agtaggaact ggggagtaa 3060
tctgttccct acagtttgc tctgagctgg aagctgtggg ggaaggagtt gacaggtggg 3120
cccagtgaa ctttccagta aatgaagcaa gcaactgaata aaaacctcct gaactgggaa 3180
caaagatcta caggcaagca agatgcccac acaacaggct tattttctgt gaaggacca 3240
actgatctcc cccaccttg gattagagtt cctgctctac cttaccaca gataacacat 3300
gttgtttcta cttgtaaatg taaagtcttt aaaataaact attacagata cttaaaaaaa 3360
aaaaaaaaaa aa 3372

```

<210> SEQ ID NO 17

<211> LENGTH: 5243

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 17

```

agcgtgagac tcgcccctc cggcacgaa aaggccaggc gacaggtgct gcttgaagaag 60
actgggcttg tcttctctgg tgcctgcctc gtcggcctct gggcagcagg tttacaaagg 120
aggaaaaaga cttctctctag attttttttt cagtttcttc tataaatcaa aacatctcaa 180
aatggagacc taaaatcctt aaagggactt agtctaactc cgggaggtag ttttctgcat 240
gggtaacaaa attaagtatt aactggtggt ttactatcca aagaatgcta attttataaa 300
catgatcgag ttatataagg tataaccataa tgagtttgat tttgaatttg atttgtggaa 360
ataaaggaaa agtgattcta gctggggcat attgttaaag cattttttcc agagttggcc 420
aggcagctct ctaactggc atctctccat tatgtagaat agaaatagta cctgtgtttg 480
ggaaagatth taaaatgagt gacagttatt tggaacaaag agctaataat caatccactg 540
caaattaaag aaacatgcag atgaaagttt tgacacatta aaatacttct acagtgacaa 600
agaaaaatca agaacaaagc tttttgatat gtgcaacaaa ttttagaggaa gtaaaaagat 660

```


-continued

aaatgtgatg attggccaag aaattatcca gttatttaca aggccactga tattttaaac	720
gtccaaaagt ttgtttaaat gggctgttac cgctgagaat gatgaggatg agaatgatgg	780
ttgaaggtta catttttagga aatgaagaaa cttagaaaat taatataaag acagtgatga	840
atacaaaaga gatttttata acaatgtgta aaatttttgg ccagggaag gaatattgaa	900
gtagatata attacttacc tttagaggaa ataattgttg gtaatgagat gtgatgtttc	960
tctgcccacc tggaaacaaa gcattgaagt ctgcagtga aaagcccaac gtctgtgaga	1020
tccaggaaac catgcttgca aaccactggt aaaaaaaaa aaaaaaaaa aaaaaagcca	1080
cagtgacttg cttattggto attgctagta ttatcgactc agaacctctt tactaatggc	1140
tagtaaatca taattgagaa attctgaatt ttgacaaggt ctctgctgtt gaaatggtaa	1200
atattattatt tttttgtca tgataaatc tggttcaagg tatgctatcc atgaaataat	1260
ttctgaccaa aactaaatg atgcaatttg attatccatc ttagcctaca gatggcatct	1320
ggtaactttt gactgtttta aaaaaataat ccactatcag agtagatttg atgttgctt	1380
cagaaacatt tagaaaaaca aaagttcaaa aatgttttca ggaggtgata agttgaataa	1440
ctctacaatg ttagtctctt gagggggaca aaaaatttaa aatctttgaa aggtcttatt	1500
ttacagccat atctaaatta tcttaagaaa atttttaaca aagggatga aatataatc	1560
atgattctgt ttttccaaa gtaacctgaa tatagcaatg aagttcagtt ttgttattgg	1620
tagtttgggc agagtctctt ttgacagcac ctgtttgtca ccataattac agaggacatt	1680
tccatgttct agccaagtat actattagaa taaaaaaact taacattgag ttgcttcaac	1740
agcatgaaac tgagtccaaa agaccaaatg aacaaacaca ttaatctctg attatttatt	1800
ttaaatagaa tatttaattg tgtaagatct aatagatca ttatactta gcaatcatat	1860
tctgatgat ctatgggaaa taactattat ttaattaata ttgaaaccag gttttaagat	1920
gtgtagcca gtctgttac tagtaaatct ctttatttgg agagaaattt tagattggtt	1980
tgttctcctt attagaagga ttgtagaaag aaaaaaatga ctaattggag aaaaattggg	2040
gatataatcat atttcaactga attcaaaatg tcttcagttg taaatcttac cattatttta	2100
cgtacctcta agaaataaaa gtgcttctaa ttaaaatag atgtcattaa ttatgaaata	2160
cttcttgata acagaagttt taaaatagcc atcttagaat cagtgaaata tggtaatgta	2220
ttattttcct cctttgagtt aggtcttggc ctttttttcc ctggccacta aatttcacaa	2280
tttccaaaa gcaaaataaa catattctga atatttttgc tgtgaaacac ttgacagcag	2340
agctttccac catgaaaaga agcttcatga gtcacacatt acatcttgg gttgattgaa	2400
tgccactgaa acattctagt agcctggaga agttgacct cctgtggaga tgccctccat	2460
taaatggcat cctgatggct taatacacat cactctctg tgaagggtt taattttcaa	2520
cacagcttac tctgtagcat catgtttaca ttgtatgtat aaagattata caaagggtgca	2580
attgtgtatt tcttctttaa aatgtatcag tataggattt agaactctca tgttgaact	2640
ctaaatgcat agaaataaaa ataataaaaa atttttcatt ttggcttttc agcctagtat	2700
taaaactgat aaaagcaaag ccatgcacaa aactacctc ctagagaaag gctagtccct	2760
tttcttcccc attcatttca ttatgaacat agtagaaaac agcatattct tatcaaat	2820
gatgaaaagc gccaacacgt ttgaaactgaa atacgacttg tcatgtgaaac tgtaccgaat	2880
gtctacgtat tccacttttc ctgctggggt tctgtctca gaaaggagtc ttgctcgtgc	2940

-continued

tggtttctat tacactgggtg tgaatgacaa ggtcaaatgc ttctgttggtg gcctgatgct 3000
ggataactgg aaaagaggag acagtcctac tgaaaagcat aaaaagttgt atcctagctg 3060
cagattcgtt cagagtctaa attccgttaa caacttgaa gctacctctc agcctacttt 3120
tccttcttca gtaacaaatt ccacacactc attacttccg ggtacagaaa acagtggaata 3180
tttccgtggc tcttattcaa actctccate aaatcctgta aactccagag caaatcaaga 3240
tttttctgcc ttgatgagaa gttcctacca ctgtgcaatg aataacgaaa atgccagatt 3300
acttactttt cagacatggc cattgacttt tctgtcgcca acagatctgg caaaagcagg 3360
cttttactac ataggacctg gagacagagt ggcttgcttt gcctgtgggtg gaaaattgag 3420
caattgggaa ccgaaggata atgctatgtc agaacacctg agacatttcc ccaaatgccc 3480
atttatagaa aatcagcttc aagacacttc aagatacaca gtttctaate tgagcatgca 3540
gacacatgca gcccgcttta aaacattctt taactggccc tctagtgttc tagttaatcc 3600
tgagcagctt gcaagtgcgg gtttttatta tgtgggtaac agtgatgatg tcaaatgctt 3660
ttgtgtgat ggtggactca ggtgtggga atctggagat gatccatggg tcaacatgc 3720
caagtggttt ccaaggtgtg agtacttgat aagaattaaa ggacaggagt tcatccgtca 3780
agttcaagcc agttaccctc atctacttga acagctgcta tccacatcag acagcccagg 3840
agatgaaaat gcagagatca caattatcca ttttgaacct ggagaagacc attcagaaga 3900
tgcaatcatg atgaatactc ctgtgattaa tgctgccgtg gaaatgggct ttagtagaag 3960
cctggtaaaa cagacagttc agagaaaaat cctagcaact ggagagaatt atagactagt 4020
caatgatctt gtgttagact tactcaatgc agaagatgaa ataagggag aggagagaga 4080
aagagcaact gaggaaaaag aatcaaatga tttattatta atccggaaga atagaatggc 4140
acttttcaa catttgactt gtgtaattcc aatcctggat agtctactaa ctgccggaat 4200
tattaatgaa caagaacatg atgtatttaa acagaagaca cagacgtctt tacaagcaag 4260
agaactgatt gatacgattt tagtaaaagg aaatattgca gccactgtat tcagaaactc 4320
tctgcaagaa gctgaagctg tgttatatga gcatttattt gtgcaacagg acataaaata 4380
tattcccaca gaagatgttt cagatctacc agtggagaa caattgcgga gactacaaga 4440
agaaagaaca tgtaaagtgt gtatggacaa agaagtgtcc atagtgttta ttccttgggtg 4500
tcatctagta gtatgcaaag attgtgctcc ttctttaaga aagtgtccta tttgtaggag 4560
tacaatcaag ggtacagttc gtacatttct tcatgaaga agaaccaaaa catcgtctaa 4620
actttagaat taatttatta aatgtattat aactttaact tttatcctaa tttggtttcc 4680
ttaaaatfff tttttatfta caactcaaaa aacattgttt tgtgtaacat atttatatat 4740
gtatctaacc catatgaaca tatatftttt agaaactaag agaatgatag gcttttgttc 4800
ttatgaacga aaaagaggta gcactacaaa cacaatattc aatcaaaatt tcagcattat 4860
tgaaaattgta agtgaagtaa aacttaagat atttgagtta accttaaga attttaata 4920
ttttggcatt gtactaatac cgggaacatg aagccagggtg tgggtggtatg tgccgttagt 4980
cccaggctga ggcaagagaa ttacttgagc ccaggagttt gaatccatcc tgggcagcat 5040
actgagaccc tgccctttaa aacaacaga acaaaaaaa aacaccaggg acacatttct 5100
ctgtctfttt tgatcagttc cctatacatc gaaggtgtgc atatagtty aatgacattt 5160
tagggacatg gtgtttttat aaagaattct gtgagaaaa atfttaataaa gcaacaaaa 5220

-continued

 ttactcttaa aaaaaaaaaaaa aaa 5243

<210> SEQ ID NO 18

<211> LENGTH: 1579

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 18

gaggaggatgc ttgccagaca ctgggtcatg gcagtggtcg gtgaagctgc agttgcctag 60
 ggcagggatg gagagagagt ctgggcatga ggagagggtc tcgggatggt tggtggact 120
 agattttaca gaaagcctta tccaggcttt taaaattact ctttcagac ttcactgag 180
 actccttctt cagccaacat tccttagccc tgaatacatt tcctatcctc atctttccct 240
 tctttttttt cctttctttt acatgtttaa atttaaacca ttcttcgtga ccccttttct 300
 tgggagattc atggcaagaa cgagaagaat gatggtgctt gttaggggat gtcctgtctc 360
 tctgaacttt ggggtcctat gcattaaata attttcctga cgagctcaag tctcctctc 420
 ggtctacaat cctggcggc tggccttcat ccttgggca agcattgcat acagctcatg 480
 gccctcctc taccataccc tccacccccc ttegcetaag ctcccttctc cgggaatttc 540
 atcatttctc agaacagcca gaacatttgt ggtctatttc tctgttagtg ttaaccaac 600
 catctgttct aaaagaaggg ctgaactgat ggaaggaatg ctgtagcct gagactcagg 660
 aagacaactt ctgcagggtc actcctctgc ttctggagga aagagaagga gggcagtgtc 720
 ccagtggtac agaagtgaga cataatggaa tcaggcttca cctccaagga cacctatcta 780
 agccatttta accctcggga ttacctagaa aaatattaca agtttggttc taggcactct 840
 gcagaaagcc agattcttaa gcacctctg aaaaatcttt tcaagatatt ctgcctagac 900
 ggtgtgaagg gagacctgtc gattgacatc ggtctgtgcc ccactatcta tcagctcctc 960
 tctgtctgtg aatcctttaa ggagatcgtc gtcactgact actcagacca gaacctgcag 1020
 gagctggaga agtggctgaa gaaagagcca gaggcctttg actggteccc agtggtgacc 1080
 tatgtgtgtg atcttgaagg gaacagagtc aagggtccag agaaggagga gaagttgaga 1140
 caggcggta agcagggtgt gaagtgtgat gtgactcaga gccagccact gggggccgtc 1200
 ccttaccctc cggctgactg cgtgctcagc aactgtgtc tggatgccgc ctgccagac 1260
 ctccccact actgcagggc gctcaggaac ctggcagcc tactgaagcc agggggcttc 1320
 ctggtgatca tggatgcgct caagagcagc tactacatga ttggtgagca gaagttctcc 1380
 agcctcccc tggcccgga ggcagtagag gctgctgtga aagaggctgg ctacacaatc 1440
 gaatggtttg aggtgatctc gcaaagtat tcttcacca tggccaacaa cgaaggactt 1500
 ttctccttg tggcgaggaa gctgagcaga cccctgtgat gctgtgacc tcaattaaag 1560
 caattccttt gacctgtca 1579

<210> SEQ ID NO 19

<211> LENGTH: 980

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 19

gggggggcgg gcagcagctg cgctgcgact gctctggaag gagaggacgg ggcacaaacc 60
 ctgaccatga cccccacag cgtgctgcca ccgctgctgc tgctgctagc tctgctgctc 120

-continued

```

gctgccagcc caggaggcgc cttggcgcgg tgcccaggct gcgggcaagg ggtgcaggcg 180
ggttgccag ggggctgctt ggaggaggag gatggggggt cgccagccga gggctgcgcg 240
gaagctgagg gctgtctcag gaggggggg caggagtgcg gggctacac ccctaactgc 300
gccccaggac tgcaagtcca tccgcccag gacgacgagg egcctttgcg ggcgctgctg 360
ctcgcccgag gccgctgctt tccggcccgc gcgctgctg ttgcagagga gaatcctaag 420
gagagtaaac cccaagcagg cactgcccgc ccacaggatg tgaaccgcag agaccaacag 480
aggaatccag gcacctctac cagccctcc cagcccaatt ctgcggtgtt ccaagacact 540
gagatgggcc catgccgtag acatctggac tcagtgtgc agcaactcca gactgaggtc 600
taccgagggg ctcaaacact ctacgtgccc aattgtgacc atcgaggctt ctaccggaag 660
cggeagtgcc gctcctccca ggggcagcgc cgaggtcctt gctggtgtgt ggatcggatg 720
ggcaagtccc tgccagggtc tccagatggc aatggaagct cctcctgccc cactgggagt 780
agcggctaaa gctgggggat agaggggctg cagggccact ggaaggaaca tggagctgtc 840
atcactcaac aaaaaaccga ggcctcaat ccacctcag gcccgcgcc atgggccctt 900
caccgctggt tggaaagagt gttggtgttg gctggggtgt caataaagct gtgcttgggg 960
tcgctgaaaa aaaaaaaaaa 980

```

<210> SEQ ID NO 20

<211> LENGTH: 7346

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 20

```

gcgaggagga aacggtgccg gagcgcgcag ggcttgcctc cgccaccgcc gctgcacagg 60
ctgccggagc gagcctgccg gcgcgcgcc tcccgcctct ccttctggg cgagctgagg 120
ggatggggcg gccgcgggag cccgagcgcg cgcaggaacc gccgcgcgc cgcgccgctt 180
ctcgttgcg gcgcgcctga gccgcctgc cgcgcgcgc ccctgcgcgc gggcgccccc 240
cccagcccca tggaggtctc ccgaggaag gcgcgcgcg gccccgcgc cccgcagcgc 300
ccaactgccc tgctgccta tctgtggca ctggcggtc ccggccggg gcgagcagag 360
cccgtgtggc ggtcggagca agccatcggg gccatcgcgc cgagccagga ggacggcgtg 420
tttgtggcga gcggcagctg cctggaccag ctggactaca gctggagca cagcctctcg 480
cgctgtacc gggaccaagc gggcaactgc acagagccgc tctcgtggc gccccgcgc 540
cggccccggc ccgggagcag cttcagcaag ctgctgtgc cctaccgcg gggggcgccc 600
ggcctcgggg ggtgctgct caccgctgg accttegacc ggggcgcctg cgaggtgccc 660
cccctgggca acctgagcgc caactcctg cgcaacggca ccgaggtggt gtcgtgccac 720
ccgagggct cgacggccgc cgtggtgtac cgcgcgggcc ggaacaacc ctggtacctg 780
gcggtggccc ccacctacgt gctgcctgag ccggagacgg cgagccgctg caaccccgcg 840
gcatccgacc acgacacggc catcgcgctc aaggacacgg aggggcgcag cctggccacg 900
caggagctgg ggcgcctcaa gctgtgcgag ggcgcgggca gcctgcaact cgtggacgcc 960
tttctctgga acggcagcat ctacttccc tactaccctt acaactacac gagcggcgtt 1020
gccaccggct ggcceagcat ggccgcgcat gcgcagagca ccgaggtgct gttccagggc 1080
caggcatccc tcgactgcgg ccacggccc cccgacggcc gccgctgct cctctcctcc 1140

```

-continued

agcctagtgg aggccctgga cgtctgggcg ggagtgttca gcgcggccgc tggagagggc 1200
caggagcggc gctccccac caccacggcg ctctgcctct tcagaatgag tgagatccag 1260
gcgcgcgcca agagggtcag ctgggacttc aagacggcgg agagccactg caaagaaggg 1320
gatcaacctg aaagagtcca accaatcgca tcattctacct tgatccattc cgacctgaca 1380
tccgtttatg gcacctgggt aatgaacagg actgttttat tcttggggac tggagatggc 1440
cagttactta aggttattct tgggtgagaat ttgacttcaa attgtccaga ggttatctat 1500
gaaattaaag aagagacacc tgtttctac aaactcgttc ctgatcctgt gaagaatata 1560
tacatttatc taacagctgg gaaagaggtg aggagaatto gtgttgcaaa ctgcaataaa 1620
cataaatcct gttcggagtg tttaacagcc acagaccctc actgcggttg gtgccattcg 1680
ctacaaaggt gcacttttca aggagattgt gtacattcag agaacttaga aaactggctg 1740
gatatttcgt ctggagcaaa aaagtgcctt aaaattcaga taattcgaag cagtaaagaa 1800
aagactacag tgactatggg ggggaagctt tctccaagac actcaaatg catggtgaag 1860
aatgtggact ctagcagggg cgtctgccag aataaaagtc agcccaaccg gacctgcacc 1920
tgtagcatcc caaccagagc aacctacaaa gatgtttcag ttgtcaacgt gatgttctcc 1980
ttcggttctt ggaatttata agacagattc aactttacca actgctcacc attaaaagaa 2040
tgcccagcat gcgtagaaa tggctgcgcg tgggtgtaaaa gtgcaagaag gtgtatccac 2100
cccttcacag cttgcgaccc ttctgattat gagagaaacc aggaacagtg tccagtggct 2160
gtcgagaaga catcaggagg aggaagaccc aaggagaaca aggggaacag aaccaaccag 2220
gctttacagg tctctacat taagtccatt gagccacaga aagtatcgac attagggaaa 2280
agcaacgtga tagtaacggg agcaaaactt acccgggcat cgaacatcac aatgatcctg 2340
aaaggaaaca gtacctgtga taaggatgtg atacaggtta gccatgtgct aaatgacacc 2400
cacatgaaat tctctcttcc atcaagccgg aaagaaatga aggatgtgtg tatccagttt 2460
gatggtgagg actgctcttc tgtgggatcc ttatctaca ttgctctgce acattgttcc 2520
cttatatttc ctgctaccac ctggatcagt ggtggtcaaa atataaccat gatgggcaga 2580
aattttgatg taattgacaa cttaatcatt tcacatgaat taaaaggaaa cataaatgtc 2640
tctgaatatt gtgtggcgac ttactgcggg tttttagccc ccagtttaa gagttcaaaa 2700
gtgcgcacga atgtcactgt gaagctgaga gtacaagaca cctacttga ttgtggaacc 2760
ctgcagtatc gggaggaccc cagattcacg gggtatcggg tggaaatcga ggtggacaca 2820
gaactggaag tgaataatca aaaagaaaat gacaacttca acatttcaa aaaagacatt 2880
gaaattactc tcttccatgg ggaaaatggg caattaaatt gcagttttga aaatattact 2940
agaaatcaag atcttaccac catcctttgc aaaattaaag gcatacaagac tgcaagcacc 3000
attgccaact cttctaagaa agttcgggtc aagctgggaa acctggagct ctacgtcgag 3060
caggagtccg ttccttcac atggtatttt ctgattgtgc tccctgtctt gctagtgatt 3120
gtcatttttg cggccgtggg ggtgaccagg cacaaatcga aggagctgag tcgcaaacag 3180
agtcaacaac tagaattgct ggaagcggag ctccggaaag agatacgtga cggctttgct 3240
gagctgcaga tgataaatt ggatgtgggt gatagttttg gaactgttcc ctctctgac 3300
tacaacatt ttgctctgag aactttcttc cctgagtcag gtggcttcc ccacatcttc 3360
actgaagata tgcataacag agacgccaac gacaagaatg aaagtctcac agctttggat 3420

-continued

gccctaactct gtaataaaaag ctttcttggt actgtcatcc acacccttga aaagcagaag 3480
aacttttctg tgaaggacag gtgtctgttt gectccttcc taaccattgc actgcaaacc 3540
aagetggctc acctgaccag catcctagag gtgctgacca gggacttgat ggaacagtgt 3600
agtaacatgc agccgaaaact catgctgaga cgcacggagt ccgctcgtcg aaaactcctc 3660
acaaactgga tgtccgtctg ctttctgga tttctccggg agactgtcgg agagcccttc 3720
tatttgctgg tgacgactct gaaccagaaa attaacaagg gtcccgtgga tgtaactact 3780
tgcaaagccc tgtacacact taatgaagac tggctgtgt ggcaggttcc ggaattcagt 3840
actgtggcat taaacgtcgt ctttgaaaa atcccggaaa acgagagtgc agatgtctgt 3900
cggaatattt cagtcaatgt tctcgtactg gacaccattg gccaaagcaa agaaaagatt 3960
ttccaagcat tcttaagcaa aaatggtctc ccttatggac ttcagcttaa tgaattggt 4020
cttgagcttc aaatgggac acgacagaaa gaacttctgg acatcgacag ttctcctg 4080
attcttgaag atggaatcac caagctaaac accattggcc actatgagat atcaaatgga 4140
tccactataa aagtctttaa gaagatagca aattttactt cagatgtgga gtactcggat 4200
gaccactgcc atttgatttt accagattcg gaagcattcc aagatgtgca aggaaagaga 4260
catcgaggga agcacaagtt caaagtaaaa gaaatgtatc tgacaaagct gctgtcgacc 4320
aagggtgcaa ttcattctgt gcttgaaaaa ctttttagaa gcatttgag tttacccaac 4380
agcagagctc catttctat aaaatacttt tttgactttt tggacgcca ggctgaaaaac 4440
aaaaaaaaatca cagatcctga cgtcgtacat atttgaaaa caaacagcct tcctcttcgc 4500
ttctgggtaa acatcctgaa gaaccctcag tttgtcttg acattaagaa gacaccacat 4560
atagacggct gtttgcagt gattgcccag gcattcatgg atgcattttc tctcacagag 4620
cagcaactag ggaaggaagc accaactaat aagcttctct atgccaagga tatcccaacc 4680
tacaagaag aagtaaaatc ttattacaaa gcaatcaggg atttgcctcc attgtcatcc 4740
tcagaaatgg aagaattttt aactcaggaa tctaagaaac atgaaatga atttaatgaa 4800
gaagtggcct tgacagaaat ttacaaatac atcgtaaaat atttgatga gattctaaat 4860
aaactagaaa gagaacgagg gctggaagaa gctcagaaac aactcttga tgtaaaagtc 4920
ttatttgatg aaaagaagaa atgcaagtgg atgtaagcac tctggggcct ggcttaactc 4980
ggcaaagtcc tcagacgac ttgggagcaa aatggctgct tgagetactc tgtgtcgtta 5040
atttgtgtt tgcacatagg ttccacttg ggcactgtct ttttaagaga ccaaggcaca 5100
tgcacagctt ttagaaagca taccaccctc tgtgcctgtg tgtataccgt gggaaacctt 5160
ctgtaaatag agttgaagt gttgttgcaa acagcctcct tgtttacaga gaatacaagg 5220
ccagtaagcg aatgtcagta ttgtaactac agtctccact taagcacaat gatataagt 5280
gttttgttg aaaaactacag ctatgtagca cttgtgtac actgcacctc tgcattgtaa 5340
agggatactg ccagtgctca aaacaaaatg tgaatgagt catttgaaa caaggtgggg 5400
gtgttagggc aacctcgagg atttgacgca ttgaaacttt ccccagtagt tcttgaaaa 5460
gotgaccgca gaatttggtg gtgtacactt agcatttgtg agtgtgtgtg tgtgtttaa 5520
ccaaaaacta acagtgttgc aacattgttg aaagggtcgt tgttttcag tggatcaaa 5580
ctgcactcca tcaaacac cctcatttca ccaaggagct ctaaagtaag gagagtgggc 5640
tttattttaa tgaacagcat ttaaccaga tactttgtcc taatgtatgt tcctttctt 5700

-continued

```

catctgtttt ttcatactaa atgtatttga tagtggacat gttggatatt atacaaaaaa 5760
atcattaatt catttctgtt ccaaaacctt tgatcagaac gatctgtgga agagtaactc 5820
catttctata tgagtgtgtg tctccttget ttagatttct ggtgaacctc gtggttatga 5880
atacttgtgt gtgatttaaa aaaaaaaga tacattttac atttcatcga attgctgttc 5940
acactggagt attatatata aatatatata tttgaggccc aaggcctgaa aaatattagt 6000
atacaacttg gtatcttagt cttactatgt actttttgaa agtattcctc gcaggagaaa 6060
gaatttaaaa taccctttt attcatgctt ttctttttaa agaattctct atccagttat 6120
actgtagtct ttttagtgct gattttttaa ttctgaatt tttgctgctc atgaccagtt 6180
ttaataccac tgtgttttcc ttctattaaa ccagaagaag taaacagcat aattggcaac 6240
tcttgagctt ttctgtggtc aggcaccttt tacccttggg gctccaaatc ccccatctag 6300
gaaagaaaat ttttcaagt caaataacat tgatcacata ttccttgaaa tcatttacca 6360
acactgtatg gagcattagg atttaaatat gaatttgtct taaaggcaat tcctttttgc 6420
ttctgtatta tctgaaaag catgagagag gtgacacctc aacaaactga tcagagaaaa 6480
taagcagtta ctaccctgat aggcaccttc ccaatcctgt tgcttttgac cattgtctgt 6540
ccaacggaca cacctcaaac aaacaaaact accaaataga tgacagatca gaataaaggt 6600
gagaggtctg gtccccattg aaggtctgta cagtcttcaa agaggtgaag gagttcataa 6660
gagaacaaca gtaggaaagt tgagagccaa gggtaggaga gttgccaaa agacttcccc 6720
tactacttta gggactgaa aactcaaagg atcagctaca gctttatcta agtatttact 6780
aaatgctaca tgaggtgtgc cctgtccagc tttctggcac atgagctctg tgtggagagt 6840
tacctcctct tccagggact gtgctgttgg gaacttggg caagtcaett acctcttgt 6900
gctcaattt ctgtataata tttctaagct acctcaactga ggtggtatga agattcacta 6960
atgtatgtag cgtgtttgtc aatcctccag tgaaaagcac tatctagatc acattttgga 7020
tcacattagc caaatgcagt aaatggccaa attagatgtg tgctgaagac aatcagtcac 7080
tgggtctata ttaaacagca accagagcaa caaatggcaa acaatttcta tttcaagtt 7140
tctttgcata ttttttggg gcaaaacctt ttataaactt ttttttctaa cactagtgtc 7200
tacagcagca ttcaaaaaaa ttctgttacc tttctgtat taggatttaa agtctatttc 7260
ttattgtata cctgattgaa gctgttcttg gagatgaatg ttttaaatgt ctatatocaa 7320
aaaataaaca ttttgatgta actgtg 7346

```

<210> SEQ ID NO 21

<211> LENGTH: 2828

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 21

```

agaacagtga cagcgccgcg gcagccgacc ccgcctcctc ggcggacagc gatgctcagc 60
tggtgcggc cgagtcctcg cctagcctcg gcagggccgc tgaccgaccg acggaggcgc 120
cgattggccg attgtccact cgcgagaagg agcagctgct ccgcgccccg ccgcgcccgc 180
ctgagggcca ggtccgcagg gccgcgggga agccgagggc tgccggagaa ccctgcaggt 240
gtcaactcggg acgcggaagt gcgcttgccg aggtttgctt tacaatacgc ttgagactcc 300
ccgacaagcg taatttggtc gagttcgacg ggaaagtact ctccccacc cagcgcgggc 360

```

-continued

cgcgtagtcc gaggttactg tccccggcgc gtccctctgtt gccccagtcc agaggctgcc	420
cttgaacccg ggcgcgcacg agcgcagggc atccgagggc acagccctg gcacggcccg	480
acctgtacc agcctggcag gaagactgta atcgtgggaa tacagctacc taccaggea	540
atatgaagat tttatttgta gaacctgcca ttttccttag tgcatttget atgactttga	600
cgggtccact gacaacgcaa tatgtttatc ggagaatatg ggaagaaact ggcaactaca	660
ctttttcatc tgatagcaat atttctgagt gtgaaaaaaaa caaaagcagc ccaatttttg	720
cattccagga ggaagttag aaaaaagtgt cacgttttaa tctgcagatg gacataagt	780
gattaattcc tggctctagt tctacattca tacttttgtc tattagtgat cactacggac	840
gaaaaattccc tatgattttg tcttcctgtg gtgctcttgc aaccagcgtt tggctctgtt	900
tgetttgcta ttttgccctt ccattccagc ttttgattgc atctaccttc attggtgat	960
tttgtggcaa ttataccaca ttttggggag cttgctttgc ctatatagtt gatcagtgt	1020
aagaacacaa acaaaaaaca attcgaatag ctatcattga ctttctactt ggacttgta	1080
ctggactaac aggaactgtca tctggctatt ttattagaga gctaggtttt gagtggctgt	1140
ttctaattat tgetgtgtct cttgctgtta atttgatcta tattttattt tttctcggag	1200
atccagttaa agagtgttca tctcagaatg ttactatgto atgtagtgaa ggcttcaaaa	1260
acctatttta ccgaacttac atgcttttta agaatgcttc tggtaagaga cgatttttgc	1320
tctgtttgtt actttttaca gtaatcactt atttttttgt ggtaattggc attgcccaca	1380
tttttatcct ttatgaattg gattcaccac tctgctggaa tgaagttttt ataggttatg	1440
gatcagcttt gggtagtgcc tcttttttga ctagtctcct aggaatatgg cttttttctt	1500
attgtatgga agatattcat atggccttca ttgggatttt taccacgatg acaggaatgg	1560
ctatgacgcg gtttgccagt acaaacctga tgatgttttt agccagggtg ccgttccttt	1620
tcactattgt gccattctct gttctacggc ccattgttgc aaaagtgggt cgttcgactg	1680
aacaaggtag cctgtttgtg tgtattgctt tcttagaaac acttggagga gtcactgcag	1740
tttctacttt taatggaatt tactcagcca ctgttgettg gtaccctggc ttcactttcc	1800
tgtgtctgc tggctctgta ctacttccag ccattcagct atgtgttgc aagtgtacca	1860
gctggaatga gggagctat gaacttctta tacaagaaga atccagttaa gatgcttcag	1920
acagagcctg ttaagctgct attgatagtc ggagcttata tactgtgact tctgaagact	1980
atacatgaat tccacaatca gtgctttggt gatacaaaat ccttaaaagg gaggcacttt	2040
aaagaatag tattttttcac ttttctaat atgtttcato ggtgacaggc atgataatat	2100
ttctatatgt aatgggtaat tgggaaaaaa tagatgataa ataaaattgc tctaaagaag	2160
ttaaaaaact gaatgaacag ctaatactgg tataaagtaa ctaatgtttg gagccaacat	2220
ttgttccttg tgcagcaaaa aggatattca cattccatga tccctggctg agaattctgc	2280
ctctagtctt tcttaccag ctgttgctca tcttgttca attataaata ctgctaaggg	2340
catttttaaa atacgatctt gtactcetta aatttgaatc cgtcagcacg gtcactcata	2400
ggaaatgat caaacaagca agccagctcat gatttgactc cttcccatct catttcttac	2460
tgccttacgc tcatcctgag gtccacctg gtctctaaaa acaccatgtg ttctcatgcc	2520
tccatgtctt ttcacacact gttccatttg ctcttctccc cacattacat tgaactttc	2580
aagcctcagt cgaaacattg cttctctctg atagcagcct tcttgacatc cctcctcact	2640

-continued

```

ccccagtcac tacaggcctt ccatagctct ttgtgtgca ctcgatccca gcattttcca 2700
tcgacttgta attgtttctg ctacctgaca atcctcgcct tgagtactgg gacaaccttt 2760
gattactcat tatatcctca ataaatattt gttgaactaa aaaaaaaaaa aaaaaaaaaa 2820
aaaaaaaaa 2828

```

```

<210> SEQ ID NO 22
<211> LENGTH: 2840
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

```

```

<400> SEQUENCE: 22

```

```

gtttttgtgc aggaacagcc cctcccgtct ttgtcctggc ggtgagcacc cagggctaag 60
cttttgaaca ctttctttgt gtttgattc agcccaggca atgcatattt gctttcattt 120
cttcttgagc ttgaggagct cctgggtgca aatcctggaa aatgaggatc tctgagcctt 180
tccagggcag ctctttgttt tgtagcagac aatgaggct ttgaaaagga aagtgggtgg 240
gggcacccca caggtggccc tcatcaccca attgccagt cctgcaggct gcttcagcag 300
aggcccagag tcaaaagagga cttaaaacca gctgtcgttt ctcccttagc ttctgtgat 360
gagagaaaac acttctgttt ttcaaagtaa gaacaaggag gaatttgttt ctaaaagAAC 420
attaaaacac aggctcgtgg tctaaaagca aatggttcag caggatgttc agggccttaa 480
agcacagtca gcaggactca gcactcctca gcacctgctc tccggttgct atggtaacat 540
catcccaaac ccaaccacct tgtccagccg agagacagca atcataagga gggacctcgg 600
tttccccga ggatcctggg ctctctttct gaaacgcttg ctctgagct cagcaaccag 660
gaacaccagg ccagcccctc cccagcact ctgtggagat gagggacaaa gtctacagt 720
ccctcttct gttctgatga gaaagggagg gaagaaaaca taccccagc gctgcaata 780
tggtcatgac actttcaaaa agcctgtgct atggagtcat gatcagaac cagagtgtgg 840
agaggtcag cagcctgcct cagagcagcc agctaggcgg ggagtggtaa atttgggact 900
tgtaaccagg catgactggc tccagccca gtgctccact ctatggaatg ttccctgggc 960
ctcagttgct ttctttctc ttgcaggccg cgggctgctg ccactctggc agctggtgag 1020
ttagctggag ggaacattc caaagcaggg gcagcatgct gctttcctcc tgtgccact 1080
cctgcgggga agtccgttga ctcccaccgc tgaagggagc tggcaacacc aggatgaggt 1140
cccaggggac gggagcagggt acccactgct tgtctacct cccactggaa aagcacggac 1200
aggccagccc ttgcgggggc aggcagagga cagagttggc tttgcgcggt ctctgcctgc 1260
tgagcagttc caatctctc catgggagaa acaaggaggc agtcgcttgt gcatgttcca 1320
gaagttttac tggggaggag gaagcggaca gaggaagctg tgtgtgcatg tgaaggggtg 1380
ggcaggggtg gagggatgca cgcgtatgtg agcatagcat gtgtgagtac tacacacatc 1440
tccatgcaga agcacaactg ggcagccctg gcttccagct ctgggcttca gcacaacaga 1500
caccagcctg tggctctctc gaagccaggg agaccacatc gggctcagga cgttttacc 1560
aaagtccaga gtttttatgc ctctccctgg cattctccat aaagaaggga aggtcagatg 1620
acccttaga tctgtgtcat ctgggaattt ccttgggctg gtttagacac gatgcctct 1680
ttttctcagg atagcagata acctgctttg aaagagggct taattctgtg ggtcctaaat 1740
tttctcctt ctctctctct ttctgtgtgt gtgtgttggg aaaatggcaa gtttccaata 1800

```

-continued

```

ccagctttgg aggaacgatt acgttttccc tccaatttca agtccgaaag accagagccc 1860
tcattccaaa gcecccccacc cagatggatt tttctgttcc atttgcate cgtcccatgg 1920
gagggcccca tgtctcctca gaaccatcc tggaggecgc aggtcgggta gagtgagttt 1980
ggcctgtcca tgacctccac ccctgagatt gtgaacaagg atgtctgggg cgatgctgag 2040
aatgtttttg aagctgctcc cagatgacgc tgatgatcac accagattga gtgctgcgat 2100
cgccttgagt ccaacctctg cataaacgag gttctcataa acaagttcac tctaccctaa 2160
gctaagtcta tgtgagcaaa cccacttcat cctttgtacc tggagacctg gttacactaa 2220
cctgatactg acctgttcat gtactgtgaa tgggtgtgtt catgcagtgt ggaccaagca 2280
atggcatggg gtgtgtgtgt gtgtgtgtgt gtgtctgtgt gtgtgtgttt gtgtatgctg 2340
tcacacttgt gtgtgtatat gtgcatgtag atgctgcata aatgattttt gatgtcaaag 2400
acaaacacat tccattgttt taaatattct attatgtaaa caatacgcag agggaccata 2460
tctactcttg tcatattatt tgtgatggta aaacatgcat ttgcaataaa ttaagctttc 2520
tgggaaggca agcagatttg gagccaaacg actgtctcgg aacatgtgtg tgttatctcg 2580
gttcataatca agtccaaagc taatggagcc tcccccgcca tccagggagg aacaccagga 2640
ccccggagtt tcttcttagt gctatatttt aaagtgcac tgacgttttc ccccccttc 2700
ttttgtgcaa gttggaagta gcagtgttct aaaagatggt ttgacgtttt tctgtttgtt 2760
ttatgttttt aaaaatgtat ctgctttgtg tttgaaata aaaatctcta tttgtgteta 2820
tgaaaaaaaa aaaaaaaaaa 2840

```

<210> SEQ ID NO 23

<211> LENGTH: 2309

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 23

```

gccccggcgg gccccggcgg gcccaggcgg gctagagggg cgggtctage ggcggccccc 60
ggcgaagttc actgcgcttg cgtgacaga cgcaagatgg cggacagtgc ggaactaaag 120
caaatggtta tgagccttag agtttctgaa ctccaagtac tgttgggcta cgcggggaga 180
aacaagcacg gacgcaaca cgaacttctc aaaaagccc tgcatttgct aaaggctggc 240
tgtagtccctg ctgtgcaaat gaaaattaag gaactctata ggccgcggtt cccacagaaa 300
atcatgacgc ctgcgactt gtccatccc aacgtacatt caagtcctat gccagcaact 360
ttgtctccat ctaccattcc acaactcact tacgatggtc accctgcac atcgccatta 420
ctccctgttt ctcttctggg acctaaacat gaactggaac tcccacatct tacatcagct 480
cttcaccag tccatccgga tataaaactt caaaaattac cattttatga tttactggat 540
gaactgataa aaccaccag tctagcatca gacaacagtc agccttttcg agaaacctgt 600
tttgcatctg ccttgacacc acaacaagtg cagcaaatca gtagtccat ggatatttct 660
gggacaaat gtgacttcc acgtacagtc cagttaaggt tttgtttatc agaaacctgt 720
tgtccacaag aagatcactt cccacccaat ctttgtgtga aagtgaatac aaaaccttgc 780
agccttccag gttaccttcc acctacaaa aatggcgtgg aaccaagcg acccagccga 840
ccaattaata tcacctcact tgtccgactg tccacaacag taccaaaac gattgtttgt 900
tcttgactg cagaaattgg aagaaactat tccatggcag tatatcttgt aaaacagttg 960

```

-continued

```

tctcaacag ttcttcttca gaggttacga gcaaagggaa taaggaatcc ggatcattct 1020
agagctttaa ttaaagagaa gttgactgcg gatccagaca gtgaaatage tacaaccagc 1080
ctaagggttt ctctactatg tccacttggg aaaatgoggc tgacaattcc gtgtcgggce 1140
ottacatggt ctcatctaca atgttttgac gcaactcttt acattcagat gaatgagaaa 1200
aaaccaacct gggtttgctc tgtctgtgat aagaaggctc catatgaaca ccttattatt 1260
gatggcttgt ttatggaaat cctaaagtac tgtacagact gtgatgaaat acaatttaag 1320
gaggatggca cttgggcacc gatgagatca aaaaaggaag tacaggaagt ttctgcctct 1380
tacaatggag tcgatggatg cttgagctcc acattggagc atcaggtagc gtctcaccac 1440
cagtcctcaa ataaaaacaa gaaagttaga gtgattgacc taaccataga cagtccatct 1500
gatgaagagg aagaagagcc atctgccaaag aggacctgtc ctccctatc tcccacatec 1560
ccactaaata ataaaggcat tttaagtctt ccacatcaag catctccagt atcccgacc 1620
ccaagccttc ctgctgtaga cacaagctac attaatacct ccctcatcca agactatagg 1680
catcctttcc acatgacacc catgccttac gacttacaag gattagattt ctttccttcc 1740
ttatcaggag acaatcagca ttacaacacc tccttgcttg ccgctgcagc agcagcagtt 1800
tcagatgac aagacctcct acactcgtct cggtttttcc cgtatacctc ctacagatg 1860
tttcttgatc agttaagtgc aggaggcagt acttctctgc caaccaccaa tggaagcagt 1920
agtggcagta acagcagcct ggtttcttcc aacagcctaa gggaaagcca tagccacacc 1980
gtcacaaaac ggagcagcac ggacacggca tccatcttgg gcatcatacc agacattatt 2040
tcattggact gattcccagg ccctgctgct cccatcccca cccagatcg aatgaacttg 2100
gcagaaagaa gagaactttg tgctctgttt taccttactc tgtttagaaa agtatacaag 2160
cgtgtttttt ttctttttt tagggaaaaa attaaagaa atgtacagag aacaaaacta 2220
tattttcagt tttacttttg tatataatc taagactgcc tgtgtgataa aacacttgtt 2280
taaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2309

```

```

<210> SEQ ID NO 24
<211> LENGTH: 1740
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 24

```

```

ccagcgttag ccgcggcca ggcagccggg aggagcggcg cgcgctcggc cctctcccgc 60
cctgctcgtt cgctctccag cttgggatgg ccggtacct cggggtcgtg cgctcgtctt 120
gcagagcctc aggctcggcg ccggcctggg cgcgcggcgc cctgacagcc cccacctcgc 180
aagagcagcc gcggcggccac tatgccgaca aaaggatcaa ggtggcgaag cccgtggtgg 240
agatggatgg tgatgagatg acccgatata tctggcagtt catcaaggag aagctcatcc 300
tgccccacgt ggacatccag ctaaagtatt ttgacctcgg gctcccaaac cgtgaccaga 360
ctgatgacca ggtcaccatt gactctgca c tggccacca gaagtacagt gtggctgtca 420
agtgtgccac catcaccctt gatgaggccc gtgtggaaga gttcaagctg aagaagatgt 480
ggaaaagtcc caatggaact atccggaaca tcctgggggg gactgtcttc cgggagocca 540
tcctctgcaa aaacatccca cgcctagtcc ctggctggac caagccatc accattggca 600
ggcagcccca tggcgaccag tacaaggcca cagactttgt ggcagaccgg gccggcactt 660

```

-continued

```

tcaaaatggt cttcacccca aaagatggca gtggtgtcaa ggagtgggaa gtgtacaact 720
tccccgcagg cggcgtgggc atgggcatgt acaacaccga cgagtccatc tcaggttttg 780
cgcacagctg cttccagtat gccatccaga agaaatggcc gctgtacatg agcaccaaga 840
acaccatact gaaagcctac gatggcgctt tcaaggacat cttccaggag atctttgaca 900
agcactataa gaccgacttc gacaagaata agatctggta tggcaccggg ctcattgatg 960
acatggtggc tcaggctctc aagtcttcgg gtggctttgt gtgggcctgc aagaactatg 1020
acggagatgt gcagtcagac atcctggccc agggctttgg ctcccttggc ctgatgacgt 1080
ccgtcctggt ctgccctgat ggggaagcga ttgaggotga ggccgctcat gggaccgtca 1140
cccgccacta tggggagcac cagaagggcc ggcccaccag caccaacccc atcgccagca 1200
tctttgctg gacacgtggc ctggagcacc gggggaagct ggatgggaac caagacctca 1260
tcaggtttgc ccagatgctg gagaagggtg gcgtggagac ggtggagagt ggagccatga 1320
ccaaggacct ggcgggctgc attcacggcc tcagcaatgt gaagctgaac gagcacttcc 1380
tgaacaccac ggacttctcc gacaccatca agagcaacct ggacagagcc ctgggcaggc 1440
agtaggggga ggcgccacc atggtctcag tggaggggcc agggctgagc cggcgggtcc 1500
tcctgagcgc ggcagagggt gagcctcaca gccctctct ggaggccttt ctaggggatg 1560
tttttttata agccagatgt ttttaaaagc atatgtgtgt tcccctcat ggtgacgtga 1620
ggcaggagca gtgcgtttta cctcagccag tcagtatgtt ttgcatactg taatttata 1680
tgcccttggg acacatggtg ccatatttag ctactaaaaa gctcttcaca aaaaaaaaaa 1740

```

<210> SEQ ID NO 25

<211> LENGTH: 1552

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 25

```

ggtcgcttta agaaggagt agctgtaatc tgaagcctgc tggacgctgg attagaaggc 60
agcaaaaaaa gctctgtgct ggctggagcc cctcagtggt gcaggetttag agggactagg 120
ctgggtgtgg agctgcagcg tatccacagg ccccaggatg caggccctgg tgctactcct 180
ctgcattgga gccctcctcg ggcacagcag ctgccagaac cctgccagcc ccccggagga 240
gggctcccca gaacccgaca gcacaggggc gctggtggag gaggagatc ctttcttcaa 300
agtccccgtg aacaagctgg cagcggctgt ctccaacttc ggctatgacc tgtaccgggt 360
gcgatccagc acgagcccca gcaccaactg gctcctgtct cctctcagtg tggccacggc 420
cctctcggcc ctctcgctgg gagcggagca gcgaacagaa tccatcatte accgggctct 480
ctactatgac ttgatcagca gccagacat ccatggtacc tataaggagc tccttgacac 540
ggctactgcc ccccagaaga acctcaagag tgctccccgg atcgtctttg agaagaagct 600
gcgcataaaa tccagctttg tggcacctct ggaaaagtca tatgggacca ggcccagagt 660
cctgacgggc aacctcctct tggacctgca agagatcaac aactgggtgc aggcgcagat 720
gaaaggggag ctgccaggt ccacaaagga aattcccgat gagatcagca ttctccttct 780
cgggtgtggc cacttcaagg ggcagtgggt aacaaagttt gactccagaa agacttcct 840
cgaggatttc tacttggatg aagagaggac cgtgagggtc cccatgatgt cggaccctaa 900
ggctgtttta cgctatggct tggattcaga tctcagctgc aagattgccc agctgcctct 960

```

-continued

```

gaccggaagc atgagtatca tcttcttctt gccctgaaa gtgaccaga atttgacctt 1020
gatagaggag agcctcacct ccgagttcat tcatgacata gaccgagaac tgaagaccgt 1080
gcaggcggtc ctcactgtcc ccaagctgaa gctgagttat gaaggcgaag tcaccaagtc 1140
cctgcaggag atgaagctgc aatccttggt tgattcacca gactttagca agatcacagg 1200
caaacccatc aagctgactc aggtggaaca ccgggctggc tttgagtgga acgaggatgg 1260
ggcgggaacc acccccagcc cagggtgca gcctgccacc ctcacctcc cgctggacta 1320
tcaccttaac cagcctttca tcttcgtact gaggacaca gacacagggg cccttctctt 1380
cattggcaag attctggacc ccaggggcc ctaatatccc agttaaataat tccaataccc 1440
tagaagaaaa ccgaggggac agcagattcc acaggacacg aaggctgccc ctgtaaggtt 1500
tcaatgcata caataaaaga gctttatccc taaaaaaaa aaaaaaaaaa aa 1552

```

<210> SEQ ID NO 26

<211> LENGTH: 4816

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 26

```

gttcccggat ttttgtgggc gcctgccccg ccctctgtcc ccctgctgtg tccatatatc 60
gaggcgatag ggttaaggga aggcggacgc ctgatgggtt aatgagcaaa ctgaagtgtt 120
ttccatgac ttttttgagt cgcaattgaa gtaccacctc ccgaggggtga ttgettcccc 180
atgcggggta gaacctttgc tgtcctgttc accactctac ctccagcaca gaatttggtt 240
tatgcctact caatgtgaag atgatgagga tgaaaacctt tgtgatgac cacttccact 300
taatgaatgg tggcaaagca aagctatatt caagaccaca tgcaaageta ctccctgagc 360
aaagagtcac agataaaacg ggggcaccag tagaatggcc aggacaaaacg cagtgcagca 420
cagagactca gaccttgcca gccatgcctg cgcaggcagt gatgagagtg acatgtactg 480
ttgtggacat gcacaaaagt gagtgtgca cggcacagac atgaagctgc ggctccctgc 540
cagtcccag acccacctgg acatgtctcc ccacctctac cagggtgccc aggtggtgca 600
gggaaacctg gaactcacct acctgcccc caatgccagc ctgtcctcc tgcaggatat 660
ccaggagggt cagggctacg tgcctatcgc tcacaaccaa gtgaggcagg tcccactgca 720
gaggctgcgg attgtgagc gcacccagct ctttgaggac aactatgccc tggccgtgct 780
agacaatgga gaccgctga acaataacc cctgtcaca ggggctccc caggaggcct 840
gcgggagctg cagcttcgaa gcctcacaga gatcttgaag ggagggtct tgatccagcg 900
gaacccccag ctctgtacc aggcacgat tttgtggaag gacatctcc acaagaacaa 960
ccagctggct ctcactgta tagacaccaa ccgctctcgg gcctgccacc cctgttctcc 1020
gatgtgtaag ggctcccgt gctggggaga gatttctgag gattgtcaga gcctgacgcg 1080
cactgtctgt gccggtgggt gtgcccctg caaggggcca ctgccactg actgctgcca 1140
tgagcagtggt gctgccggct gcacgggccc caagcactct gactgcctgg cctgcctcca 1200
cttcaaccac agtggcatct gtgagctgca ctgccagcc ctggtcacct acaacacaga 1260
cacgtttgag tccatgccc atcccaggg ccggtataca ttcggcgcca gctgtgtgac 1320
tgctgtccc tacaactacc tttctacgga cgtgggatcc tgcacctcg tetgcccct 1380
gcacaaccaa gaggtgacag cagaggatgg aacacagcgg tgtgagaagt gcagcaagcc 1440

```

-continued

ctgtgcccga gtgtgctatg gtctgggcat ggagcacttg cgagaggtga gggcagttac 1500
cagtgccaat atccaggagt ttgtggctg caagaagatc tttgggagcc tggcatttct 1560
gcccggagagc tttgatgggg acccagcctc caaactgcc ccgctccagc cagagcagct 1620
ccaagtgttt gagactctgg aagagatcac aggttaccta tacatctcag catggccgga 1680
cagcctgctt gacctcagcg tcttccagaa cctgcaagta atccggggac gaattctgca 1740
caatggcgcc tactcgctga cctgcaagg gctgggcato agctggctgg ggetgctgc 1800
actgaggaa ctgggcagtg gactggcctt catccaccat aacaccacc tetgcttctg 1860
gcacacggtg cctgggacc agctctttcg gaaccgcac caagctctgc tccactgctc 1920
caaccggcca gaggacgagt gtgtggcgca gggcctggcc tgccaccagc tgtgccccg 1980
agggcactgc tggggtccag gggccacca gtgtgtcaac tgcagccagt tcttccggg 2040
ccaggagtgc gtggaggaat cccgagtagt gcaggggctc cccaggagat atgtgaatgc 2100
caggcactgt ttccctgctc accctgagtg tcagccccag aatggctcag tgacctgtt 2160
tggaccggag gctgaccagt gtgtggcctg tgcccactat aaggaccctc ccttctgct 2220
ggcccctgc cccagcggtg tgaaacctga cctctctac atgcccactt ggaagtttcc 2280
agatgaggag ggcgatgctc agccttgccc catcaactgc accactcct gtgtggacct 2340
ggatgacaag ggtgccccg ccgagcagag agccagcctc ctgacgtcca tcatctctgc 2400
ggtggttggc attctgctgg tctgtgctt gggggtggtc tttgggatcc tcatcaagcg 2460
acggcagcag aagatccgga agtacacgat gcggagactg ctgcagaaa cggagctggt 2520
ggagccgctg acacctagcg gacgatgctc caaccaggcg cagatgcgga tctgaaaga 2580
gacggagctg aggaaggtga aggtgcttgg atctggcctc tttggcacag tctacaagg 2640
catctggatc cctgatgggg agaattgtaa aattccagtg gccatcaaag tgttgaggga 2700
aaacacatcc cccaaagcca acaagaat cttagacgaa gcatactga tggctggtgt 2760
gggctcccca tatgtctccc gccctctggg catctgctg acatccacgg tgcagctggt 2820
gacacagctt atgcccctg gctgctctt agaccatgtc cgggaaaacc gcggacgct 2880
gggctcccag gacctgctga actggtgtat gcagattgcc aaggggatga gctacctgga 2940
ggatgtgctg ctgctacaca gggacttggc cgtctggaac gtgctggtca agagtccca 3000
ccatgtcaaa attacagact tcgggctggc tcggctgctg gacattgacg agacagagta 3060
ccatgcagat gggggcaagg tgcccataca gtggatggcg ctggagtcca ttctccgccc 3120
gcggttcacc caccagagtg atgtgtggag ttatggtgtg actgtgtggg agctgatgac 3180
ttttggggcc aaaccttacg atgggatccc agcccgggag atccctgacc tctggaaaa 3240
gggggagcgg ctgccccagc ccccactctg caccattgat gtctacatga tcatggtcaa 3300
atgttgatg attgactctg aatgctggc aagattccgg gagttggtgt ctgaattctc 3360
ccgcatggcc agggaccccc agcctttgtt ggtcatccag aatgaggact tgggcccagc 3420
cagtcccttg gacagacact tctaccgctc actgctggag gacgatgaca tgggggacct 3480
ggtgatgct gaggagtatc tggtaaccca gcagggttc ttctgtccag accctgcccc 3540
ggcgctggg ggcaggttcc accacaggca ccgacgtca tctaccagga gtggcggtgg 3600
ggacctgaca ctaggctgg agcctctga agaggaggcc cccaggtctc cactggcacc 3660
ctccgaagg gctggtctcc atgtattga tggtagctg ggaatggggg cagccaaggg 3720

-continued

```

gctgcaaagc ctccccacac atgaccccag ccctctacag cggtagcagtg aggaccccac 3780
agtaccctcg cectctgaga ctgatggcta cgttgcccc ctgacctgca gccccagcc 3840
tgaatatgtg aaccagccag atgttcggcc ccagcccctc tcgccccgag agggccctct 3900
gcctgctgcc cgacctgctg gtgccactct ggaaaggccc aagactctct ccccagggaa 3960
gaatggggtc gtcaaagacg tttttgcctt tgggggtgcc gtggagaacc ccgagtactt 4020
gacaccccag ggaggagctg cccctcagcc ccacctcct cctgecttca gcccagcctt 4080
cgacaacctc tattactggg accaggagcc accagagcgg ggggctccac ccagcacctt 4140
caaagggaca cctacggcag agaaccocaga gtacctgggt ctggacgtgc cagtgtgaa 4200
cagaaggcca agtccgcaga agcctgatg tgcctcagg gagcaggaa ggccctgactt 4260
ctgctggcat caagaggtg gagggccctc cgaccactc caggggaacc tgccatgcca 4320
ggaacctgtc ctaaggaacc ttccttcctg cttgagttcc cagatggctg gaaggggtcc 4380
agcctcgttg gaagaggaac agcactgggg agtctttgtg gattctgagg cctgoccaa 4440
tgagacteta gggctcagtg gatgccacag cccagcttgg cctttcctt ccagatcctg 4500
ggtactgaaa gccttaggga agctggcctg agaggggaag cggccctaag ggagtgtcta 4560
agaacaaaag cgaccatctc agagactgtc cctgaaacct agtactgcc cccatgagga 4620
aggaacagca atgggtgcag tatccagctt ttgtacagag tgcctttctg tttagttttt 4680
actttttttg tttgttttt ttaaagatga aataaagacc cagggggaga atgggtgttg 4740
tatggggagg caagtgtggg gggctcctct ccacaccac tttgtccatt tgcaaatata 4800
ttttgaaaa cagcta 4816

```

<210> SEQ ID NO 27

<211> LENGTH: 6831

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 27

```

ccaggcccca ttgttcccgg tttccagcca tggctgcat tacctgacca gcgccacagc 60
cggctctctc gcaggcgcgg ggagaagtga ccagagcaat ttctgctttt cacagggcgg 120
gtttctcaac ggtgacttgt gggcagtgcc ttctgctgag cgagtcattg cccgaaggca 180
gaactaactg tgectgcagt cttcactctc aggatgcagc cgaggtgggc ccaaggggcc 240
acgatgtggc ttggagtccg gctgacctt ctgctctgtt caagccttga gggtaagaa 300
aactctttca caatcaacag tgttgacatg aagagcctgc cggactggac ggtgcaaat 360
gggaagaacc tgacctgca gtgcttcgcg gatgtcagca ccacctctca cgtcaagcct 420
cagcaccaga tgctgttcta taaggatgac gtgctgtttt acaacatctc ctccatgaag 480
agcacagaga gttattttat tectgaagtc cggatctatg actcaggac atataaatgt 540
actgtgattg tgaacaacaa agagaaaacc actgcagagt accaggtgtt ggtggaagga 600
gtgcccagtc ccaggtgac actggacaag aaagaggcca tccaaggtgg gatcgtgagg 660
gtcaactgtt ctgtcccaga ggaaaaggcc ccaatacact tcacaattga aaaactgaa 720
ctaaatgaaa aaatggtaaa gctgaaaaga gagaagaatt ctcgagacca gaattttgtg 780
atactggaat tcccgttga ggaacaggac cgcgttttat ccttccgatg tcaagctagg 840
atcatttctg ggatccatat gcagacctca gaatctacca agagtgaact ggtcacctg 900

```

-continued

acggaatcct tctctacacc caagttccac atcagcccca ccggaatgat catggaagga 960
gctcagctcc acattaagtg caccattcaa gtgactcacc tggcccagga gtttcagaa 1020
atcataattc agaaggacaa ggcgattgtg gcccaaca gacatggcaa caaggctgtg 1080
tactcagtca tggccatggt ggagcacagt ggcaactaca cgtgcaaagt ggagtccagc 1140
cgcatatcca aggtcagcag catcgtggtc aacataacag aactatlttc caagcccga 1200
ctggaatctt ccttcacaca tctggaccaa ggtgaaagac tgaacctgtc ctgctccatc 1260
ccaggagcac ctccagccaa cttcaccatc cagaaggaag atacgattgt gtcacagact 1320
caagatttca ccaagatagc ctcaaagtcg gacagtggga cgtatatctg cactgcaggt 1380
attgacaaag tggtaagaa aagcaacaca gtccagatag tcgtatgtga aatgctctcc 1440
cagcccagga tttcttatga tgcccagttt gaggtcataa aaggacagac catcgaagtc 1500
cgttgcaaat cgatcagtgg aactttgctt atttcttacc aacttttaa aacaagtaa 1560
gttttgaga atagtacaa gaactcaaat gatcctgctg tattcaaaga caaccocact 1620
gaagacgtcg aataccagtg tgttcagat aattgccatt cccatgcaa aatgttaagt 1680
gaggttctga gggtaaggt gatagcccc gtggatgagg tccagatttc tatcctgtca 1740
agtaagtggt tggagtctgg agaggacatt gtgctgcaat gtgctgtgaa tgaaggatct 1800
ggtccatca cctataagtt ttacagagaa aaagagggca aacccttcta tcaaatgacc 1860
tcaaatgcca cccaggcatt ttggaccaag cagaaggcta gcaaggaaca ggaggagag 1920
tattactgca cagccttcaa cagagccaac cagcctcca gtgtcccag aagcaaaata 1980
ctgacagtca gagtcattct tgccccatgg aagaaaggac ttattgcagt ggttatcatc 2040
ggagtgatca ttgctctctt gatcattgct gccaaatggt attttctgag gaaagccaag 2100
gccaaagcaga tgccagtga aatgtccagg ccagcagtac cacttctgaa ctccaacaac 2160
gagaaaatgt cagatcccaa tatggaagct aacagtcatt acggtcacia tgacgatgtc 2220
agaaacctg caatgaaacc aataaatgat aataaagagc ctctgaaact agacgtgcag 2280
tacacggaag ttcaagtgtc ctgagctgag tctcacaag atctaggaaa gaaggacaca 2340
gagacagtgt acagtgaagt ccggaagct gtccctgatg ccgtggaag cagatactct 2400
agaacggaag gctcccttga tggaacttag acagcaaggc cagatgcaca tccctggaag 2460
gacatccatg tccgagaag aacagataat cctgtatctt caagacctct gtgcacttat 2520
ttatgaacct gccctgctcc cacagaacac agcaattcct caggctaagc tgccggttct 2580
taaaccatc ctgctaagtt aatgttgggt agaaagagat acagaggggc tgttgaattt 2640
cccacatacc ctccctccac caagttgga catccttga aattggaaga gcacaagagg 2700
agatccagg caaggccatt gggatattct gaaactgaa tattttgtt tgtgcagaga 2760
taaagacett tccatgcac cctcatacac agaaaccaat tttctttttt ataactcaatc 2820
atttctagcg catggcctgg tttagaggctg gttttttctc ttttctttg gtcctcaaa 2880
ggctttagt tttgctagt ccttgttctt tggaaataca cagtctgac cagacagcct 2940
ccccctgtcc cctctatgac ctgcccctcc acaaattgga aaaccagact acttgggagc 3000
accgctgtg aaataccaac ctgaagacac cgttcattca ggcaacgcac aaaacagaaa 3060
atgaagtggt aacaagcaca gatgttcttc aactgttttt gtctacatc tttctctttt 3120
cctctaccat gctgaaggct gaaagacagg aagatggtgc catcagcaa tattattctt 3180

-continued

aattgaaaac tgaaatgtg tatgtttctt actaattttt aaaaatgtat tccttgccag 3240
ggcaggcaag gtggctcacg cctgtaatcc cagcacttca ggaggctgag gtgggaggat 3300
cacctgaggt caggagtttg agaccagcct gatgaaacce tgtctctact aaaaatacaa 3360
gaattagccg ggcgtggtgg cgcctgcctg tagtatcagc tactcaagag gctgaggtga 3420
gattatcgc tgaaccagg aaacggaggt tgtagtgagc ggagatcgcg ccaactgcact 3480
ccagcctgag tgacagagtg agaatccatc tcaaaaaaaa caaaaaacaa aattgcttgc 3540
taaagaagtg gtctcctgag gtcttaagac attcctgaca gtgtcttgag tgggtgggag 3600
agaggctgct gtcattgcgc tgtggaattt cacagatgag aaccacgcct agccaaaatc 3660
acttttctcg tttgcctcag tgacacagct gcagggaccc tcgtggatgt tgtattaaat 3720
aaatttgacc tttgctcttt gcagatctgt gaaatgttgt cttctgaggg gccacatgca 3780
tctatagtgc tgaggactcc ttgggcctct gaagtccacag agagaaccga gcaggctcat 3840
gtttttgttt tgtgtttttg agacggagat tcgctcttgt tgccccggct ggactgcagc 3900
ggcgcaacct ctgctcactg caacctccgc ctctgggtt caagcagttc tcctgtctca 3960
gcctcccag tagctgggat tacaggcaca tgtcaccacg cctggctaata tttgtattt 4020
ttagtagaga tggggtttca ccacgttggc caggctgato tcgaatgcct gacctttggt 4080
gatctgcccg ccttgcctcc atgtgtgctc cacaggcctt tgggttggga ttgcaggcgt 4140
gagccaccat gccagccta gactcttttg acaatatgat gaaagctggt ggttcctttc 4200
cccaacacac acacaccgag ttgtatcacg aaaatgtcat acaatttcca ggttttctga 4260
gtggtgggct cagattgagg tcaaaggate agacgacctc taacgacctc catgtctctg 4320
ttgatgatct ggggacagcc agatcccctg tgtccaggga gttccttagt cccttgccac 4380
caccagagaa gggcaattgc cacgggagct gcaaagaccc tattcctact cctgggtgcct 4440
tacttatgca gcacgactga attttttgtt ttgttttgtt ttgttgagac aggggcttgc 4500
tctgttgcce aggctggagt gcagtgacc aacaatggct caccgcagcc tgaacccct 4560
gggctcaage gatcctccca tctcagcttc ctgggtagct gggaccagag gcgtgagccg 4620
ccatagctgg ctaattttta atttttttt tgcagagatg aggtttcacc atggtgccc 4680
ggctggtctc gaacttctgg gctcaagtga tcctccctcc ttggcctcgc aaagtgctgg 4740
gattgcaggc atgagccacc gcccccggcc tgtggagcac acatgagttt aaaattactt 4800
tccttctgc ctatatttcc gaggaggaaa cttcatgcgc agggatcttt cttagtggat 4860
ttaatggcta aaaggtctgt ctgaatccag gacgctggct ttagccttcc tcggcagctg 4920
ccgtaacccc ggtgtctaaa cctgaagcat cccaggagca cccactccag gagttttctc 4980
ggcgcgga ctcattagt agagccct cttgtgttct catgtggtaa tcggtcactg 5040
aaggacttaa aatggtcctt agccaacaca cagtaaaact tttcctctt ctgaccccaa 5100
gaggctagcc acccatttca tgagcatata ctggtegccc catcagcgtt ctctgattgg 5160
ctaactgaac ccactccccg acctagactc aagacaggcg aagtgcgct taggtcaaca 5220
ttcactcact aaagcaacga ctgtcgggag attttgtctc ccgctggttt tggaaatggtg 5280
tctggagaca tttttggttg tcacagctgg gtgggtgtgc tcccggcatc tgggtggtag 5340
aaaccaagca tgctcctaaa catcctacag gcacagaacc gtctcccacg accaagcatg 5400
atcaagtccc aatgccaat aatggccagg ttgagaaact ctgcacagaa gcatccagtt 5460

-continued

```

atttgtctgt ttgctcaaca agcttgtgct catcatgctc tgtgttctctg acgctgtgct 5520
gggtggttggc ggtgggaaga ttacaagagt cacatggcag ctgtcctcct ggaaggtaca 5580
accagtaga gatgcagact aacagagagc caattacaaa gcagtgtgac aagcgtcatg 5640
gtggaaaatt aaaagctcaa acaagggcac atgggagggg cttccaacac agactttggg 5700
ggatccagga aggtctaa gaaggtggg tctcaccaaa gccttgacca taggcagagg 5760
gtaccagtgg aaaaggtggg gtgaagaaca ttgaggacaa aaggaagaag tgcaggaagg 5820
ccctgaggca agggagtggg ggggtgcctg gagggatggc agcagggcag tctgtcagac 5880
ccaagtggcc tccagcccta gaagccaatt agtcctcctc aaaaagctgt cactgtcccc 5940
taagaattgc tgcaggctc ccactggcct gactcagtct ttgagagtct taaggaggag 6000
gtctctgaaa ggtacacacc aagaactctc cccagcacag ctgtttttaa gactctccac 6060
cagcgtcatt ggcgtgttgg gaagaaacc tctgccacag aggccagctt cagcctttgc 6120
ctaacaccgc aagggcaaat ggaaggtaa acgggaagga gatgtctccc cagcaggcta 6180
tttgaggaca gtcttcctctg cagaagatct caacctgggg tccacagagt ggaatgtta 6240
gagtagggag ctaggcaaac atgagcagga cagggtgagg cccccacagg aatgtcaggc 6300
taccatcagg tgatggtcag gtggtgttta aactgtctct gtaaataat aattggttgc 6360
agccagctcc aagcaaggac agtctctcaa tagatacaa acacctgat ctgggtgatea 6420
gccgcttccc gataagatct caggagctgg gcaagcagcc tggagcatgc gcaccaagag 6480
gcaaaatggc ggaattaac cagtatatga cctaccttc tctgggaacg cagcactggt 6540
aaggggaaaa atgcctcaag tgagcatgcg cgcaacttca gtaatcacac tgtgcatgcg 6600
acccttcca agtgcctggca ggtcaccaca tacgcgga gectgctgca agggaagaat 6660
caggggagat gagacgtaa tcccagaact atgccaaata cataaaacc caagttaagg 6720
gtcaggcagg gcacttagat ctctcaagtt gcctgcctga cccaagtga gtgtacttcc 6780
ttttgttctc gctctaaaac tttttaataa actctcactc ctgctctaaa a 6831

```

<210> SEQ ID NO 28

<211> LENGTH: 2956

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 28

```

ggggctgagt gctcagtgga gagcggggag ttgtgtccac cttgccgacg tcgctagccg 60
tggggctgtc ctgggaaggc ggaagggcag cgcctgggtt ccgcaactcg cgcctgccc 120
tgcccgtctg cgcctgtgtc atcctcactc gggacgcagg gacctttttt aaatcacagg 180
ggcgtgtgtc agcctgcctt aggaacttcat gtctatatat ttcccattc actgcccga 240
ctatctgaga tcggccaaga tgactgaggt gatgatgaac acccagccca tggaggagat 300
cggcctcagc ccccgcaagg atggccttcc ctaccagatc ttcccagacc cgtcagattt 360
tgaccgctgc tgcaaaactga aggaccgtct gcctccata gtggtggaac ccacagaagg 420
ggaggtggag agcggggagc tccggtggcc ccctgaggag ttctgtgtcc aggaggatga 480
gcaagataac tgcgaagaga cagcgaaga aaataaagag cagtagagtc cctgtggact 540
cccatgggtc ataccagcca gcatctgttc ctgaactgtg tttttcccat catgacggaa 600
gaagagagtg agccgcaatt gttctgaaaa tgtcaaacga ggcttctgtt ttgcacctgc 660

```

-continued

agatcaccca gttggtttcc tttcttttcc ttgccttttt ttttttttga aatttgccga	720
gcagtgaggc cctctgacaa tttgcaaggc cctctgagaa aggaagctgc ttagagccag	780
ggggtagtg ggtgagggga gcgagtgctg tttttgagat cattatctga actcaggcag	840
cctagtagag gcagtggtgg gattccaatg ggtcttggtg ggtgggaggt ggggcatgtg	900
caaagcaagc aaggaacatt tgggtaaga aaacaaacat gaggcaaaag aaaaaataca	960
tgttttaag aaaacattga gcagagaact gcagccagga tgcgctcagc agacattcac	1020
tctggctgct gggacatcag aaaacaaagt cttcatctct ctctccagtt tccccaccc	1080
caccctttgc tttcatttca ggtgtgttgg tctatatgac agggaggaga gtaaaggaga	1140
gcaggagcaa ttggctgcct gcaaagccag ctggaggatga agtgcaggaa aggaaaggtc	1200
acccattct actccatggc ctctctgctc ccagctgtgg taggctcaca tagccagtgt	1260
gatcggttt taagaggcag tgcttttccag cttttctccc tgatataatcc attttgcttc	1320
ccagcacttt ttaggagtag tgagagcact tcttgcctt gttggaagcc ccagggtgga	1380
cactcagcac gaaggtctct ccttaactg ctgcccctcc aagacttgct cccgagatgg	1440
agtgggctg gtcttccagg ctggcccttc cttctctca cgcacactt cctgccccca	1500
gccccagcag ccatgggtac atgggtcccc agctcaccta tggattcccg ccagtctgcc	1560
cagctgcagt actcaagccc catgggggat cttggtctgt tttcttctgt ggagcctagt	1620
ggagagcaga cgtggctttt tatgtgtctt gttggggagg tgacttgcac ggtggggaca	1680
aggetgtcgt ggcaaccttg ggatcgagtt tgagactaaa ggatgtcatg agatccctgg	1740
cttctcccca tgtgttccc ggacaagggc agaagggagg catggcaagg gacctctgct	1800
gtccttactc aacagtggtc ctcatccctc cccacctccc actgettctt gcaagggcac	1860
cagttgtatg agaaagtgg cctttggact taggatttct tattgtagct aagagccatc	1920
tgaagcagca ggttgcagga caaatgtctc agtcccgcga gaggcagacc gtgtggccaa	1980
gaggtggact cagagccttc cttgagctaa actcggccaa ccaaggcacg cagcatgtcc	2040
cctcaggtct ccagtcagtc caggttgacc ctcaagttctg gacgtgtgta tatagctgta	2100
tttaatacct caaggtcatt ttggctctgg ggatgccggg gcaggaggac gagggtgccc	2160
tgtggacaca gcagtcagcg gaattccgtt ctgggaagcc aatggctgcc ggcacccctt	2220
gcttctccc tctgtgtct cctgtgtgta cacacatcaa tggcaataac ttcttccaac	2280
tctctgcaga agtgggagag gccggcagcc tgcaccgaga ggggcttcc tctctcttgc	2340
tcccccttc gttctgtttt ggctgcagag agtggttcat ccatactctc attccctcgc	2400
ctcccctgt ggaagggggc cttgcccctt caattcctgt gttttggtgt ctcccttat	2460
ctgetacct gaatcacctg tctgtgtctt gctgtgtgat gggaaatgc ttgtaactg	2520
cgtaacaaat ctactttgtg tatgtgtctg tttatggggg tggtttatta tttttgctgg	2580
tccctagacc actttgtatg accgtttgca gtctgagcag gccaggggct gacagctaat	2640
gtcaggaccc tcagcgggtg agcctgctgg ggggacccag ctgctcttgg acaagtggct	2700
gagtcctat ctggcctcct cttttttttt ttttcaagta atttgtgtgt atttctaact	2760
gattgtattg aaaaaattcc tagtatttca gtaaaaaatgc ctgttgtgag atgaacctcc	2820
tgtaacttct atctgttctt ttttgaggct caggagaaaa ctageatttt ttttttcca	2880
aactactttt tgtcactgtg acagttgtaa ataaagtttg aaaatgcttt ccactctgaa	2940

-continued

aaaaaaaaaaaa 2956

<210> SEQ ID NO 29
 <211> LENGTH: 2262
 <212> TYPE: DNA
 <213> ORGANISM: Homo sapiens

<400> SEQUENCE: 29

ctggggagta atagcatggg caaccattat cctgtctcgc cgccaccag gacatggctt 60
 ctgttccaat gccaaagtgg tacacctatg tgaactgag aagtgattgc tggaggcctt 120
 ccttgcaatg gtacaccgca gctcaaagca agatgagaag gccagcttg ttattaaaag 180
 acatcctcaa atgtacattg cttgtgttg gagtgggat cctttatate ctcaagttaa 240
 attatactac tgaagaatgt gacatgaaaa aaatgcatta tgtggaccct gaccatgtaa 300
 agagagctca gaaatatgct cagcaagtct tgcagaagga atgtcgtccc aagtttgcca 360
 agacatcaat ggcgctgcta tttgagcaca ggtatagcgt ggacttactc ccttttgctc 420
 agaaggcccc caaagacagt gaagctgagt ccaagtacga tcctcctttt gggttccgga 480
 agttctccag taaagtccag accctcttgg aactcttggc agagcacgac ctccctgaac 540
 acttgaaagc caagacctgt cggcgctgtg tggttattgg aagcggagga atactgcacg 600
 gattagaact gggccacacc ctgaaccagt tcgatgtgtg gataaggcta aacagtgcac 660
 cagttgaggg atattcagaa catgttgtaa ataaaactac tataaggatg acttatccag 720
 agggcgacc actgtctgac cttgaaatatt attccaatga cttatttgtt gctgttttat 780
 ttaagagtgt tgatttcaac tggcttcaag caatggtaaa aaaggaaacc ctgccattct 840
 gggtagcact cttcttttgg aagcaggtgg cagaaaaaat cccactgcag caaaaacatt 900
 tcaggatttt gaatccagtt atcatcaaa agactgcctt tgacatcctt cagtactcag 960
 agcctcagtc aaggttctgg ggcgagata agaactccc cacaatcggg gtcattgccc 1020
 ttgtcttagc cacacactctg tgcgatgaag tcagtttggc gggttttgga tatgacctca 1080
 atcaaccagc aacacctttg cactacttct acagtcaatg catggctgct atgaacttct 1140
 agacctgca taatgtgaca acggaaacca agttcctctt aaagctggtc aaagaggag 1200
 tggtgaaaga tctcagtgga ggcattgac gtgaattttg aacacagaaa acctcagttg 1260
 aaaatgcaac tctaactctg agagetgttt ttgacagcct tcttgatgta tttctccatc 1320
 ctgcagatc tttgaagtgc agctcatggt ttttaacttt aatttaaaaa cacaaaaaaa 1380
 attttagctc ttcccacttt ttttttcta tttattgag gtcagtgttt gtttttgac 1440
 accatthtgt aaatgaaact taagaattga attggaaaga cttctcaaag agaattgtat 1500
 gtaacgatgt tgtattgatt ttaagaaag taatttaatt tgtaaaactt ctgctcgttt 1560
 aactgcaca ttgaatacag gtaactaatt ggaaggagag gggaggtcac tcttttgatg 1620
 gtggccctga acctcattct ggttccctgc tgcgctgctt ggtgtgaccc acggaggatc 1680
 cactcccagg atgactgct cctgactct gctgctgata ctgggtctgc gatgcagcgg 1740
 cgtgaggcct gggctggttg gagaaggtca caacccttct ctgittgtct gccttctgct 1800
 gaaagactcg agaaccaacc agggaaagctg tcctggaggt ccctggtcgg agaggacat 1860
 agaactctgt acctctgaca actgtgaagc caccctgggc tacagaaacc acagtcttcc 1920
 cagcaattat tacaattctt gaattccttg gggatttttt actgcccttt caaagcactt 1980

-continued

aagtgttaga tctaactgtt tccagtgctt gtctgagggt acttaaaaaa tcagaacaaa 2040
acttctatta tccagagtca tgggagagta caccctttcc aggaataatg ttttgggaaa 2100
cactgaaatg aaatcttccc agtattataa attgtgtatt taaaaaaaaa aaacttttct 2160
gaatgcctac ctggcggtgt ataccaggca gtgtgccagt taaaaaagat gaaaaagaat 2220
aaaaactttt gaggaaaaaa aaaaaaaaaa aaaaaaaaaa aa 2262

<210> SEQ ID NO 30
<211> LENGTH: 4909
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 30

tagacgcacc ctctgaagat ggtgactccc tctgagaag ctggaccctt tggtaaaaga 60
caaggccttc tccaagaaga atatgaaagt gttactcaga cttatttgtt tcatagctct 120
actgatttct tctctggagg ctgataaatg caaggaactg gaagaaaaaa taattttagt 180
gtcatctgca aatgaaattg atgttcgtcc ctgtcctctt aaccctaatg aacacaaagg 240
cactataact tggataaaag atgacagcaa gacacctgta tctacagaac aagcctccag 300
gattcatcaa cacaaagaga aactttgggt tgttcctgct aaggtggagg attcaggaca 360
ttactattgc gtggtaagaa attcatctta ctgcctcaga attaaaataa gtgcaaaatt 420
tgtggagaat gagcctaact tatgttataa tgcacaagcc atatttaagc agaaactacc 480
cgttgcagga gacggaggac ttgtgtgccc ttatatggag ttttttaaaa atgaaaataa 540
tgagttacct aaattacagt ggtataagga ttgcaaacct ctactctctg acaatataca 600
ctttagtgga gtcaaagata ggctcatcgt gatgaatgtg gctgaaaagc atagagggaa 660
ctatacttgt catgcctcct acacatactt gggcaagcaa tatcctatta cccgggtaat 720
agaatttatt actctagagg aaaaacaaac cacaaaggcct gtgattgtga gccagctaa 780
tgagacaatg gaagttagat tgggatccca gatacaattg atctgtaatg tcaccggcca 840
gttgagtgac attgcttact ggaagtggaa tgggtcagta attgatgaag atgaccctgt 900
gctaggggaa gactattaca gtgtggaaaa tcttgcaaac aaaagaagga gtaccctcat 960
cacagtgtct aatataatcg aaattgaaag tagattttat aaacatccat ttacctgttt 1020
tgccaagaat acacatggta tagatgcagc atatatccag ttaatatatc cagtcactaa 1080
ttccagaag cacatgattg gtatatgtgt cacgttgaca gtcataattg tgtgttctgt 1140
tttcctctat aaaatcttca agattgacat tgtgctttgg tacagggatt cctgctatga 1200
ttttctccca ataaaagctt cagatggaaa gacctatgac gcatatatac tgtatccaaa 1260
gactgttggg gaagggtcta cctctgactg tgatattttt gtgtttaaag tcttgctga 1320
ggctttggaa aaacagtgtg gatataagct gttcatttat ggaagggatg actacgttgg 1380
ggaagacatt gttgaggtca ttaatgaaaa cgtaaagaaa agcagaagac tgattatcat 1440
tttagtcaga gaaacatcag gcttcagctg gctgggtggg tcatctgaag agcaaatagc 1500
catgtataat gctcttgttc aggtatgaa taaagtgtgc ctgcttgagc tggagaaaat 1560
ccaagactat gagaaaatgc cagaatcgat taaatcatt aagcagaaac atggggctat 1620
ccgctggta ggggacttta cacagggacc acagtctgca aagacaaggt tctggaagaa 1680
gtcaggtac cacatgccag tccagcgagc gtcaccttca tctaacacc agttactgtc 1740

-continued

accagccact aaggagaaac tgcaaagaga ggctcacgtg cctctcgggt agcatggaga 1800
agttgccaaag agttcttttag gtgcctcctg tcttatggcg ttgcaggcca ggttatgcct 1860
catgctgact tgcagagttc atggaatgta actatatcat cctttatccc tgaggtcacc 1920
tggaatcaga ttattaaggg aataagccat gacgtcaata gcagcccag gcacttcaga 1980
gtagagggtc tgggaagatc ttttaaaaag gcagtaggcc cggtgtggtg gctcacgcct 2040
ataatcccag cactttggga ggctgaagtg ggtggatcac cagaggtcag gagttcgaga 2100
ccagcccagc caacatggca aaaccccate tctactaaaa atacaaaaat gagctaggca 2160
tggtggcaca cgctgtaat cccagctaca cctgaggctg aggcaggaga attgctttaa 2220
ccggggagac ggaggttgca gtgagccag tttgggccac tgcactctag cctggcaaca 2280
gagcaagact ccgtctcaaa aaaagggcaa taaatgcct ctctgaatgt ttgaactgcc 2340
aagaaaaggc atggagacag cgaactagaa gaaagggcaa gaaggaaata gccaccgtct 2400
acagatggct tagttaagtc atccacagcc caaggcgggg gctatgcctt gctcggggac 2460
cctgtagagt cactgaccct ggagcggctc tctgagagg tgctgcaggc aaagtgagac 2520
tgacacctca ctgaggaagg gagacatatt cttggagaac tttccatctg cttgtatttt 2580
ccatacacat cccagccag aagttagtgt ccgaagaccg aattttttt tacagagctt 2640
gaaaactcac tteaatgaac aaagggatc tccaggatc caaagtttg aagtcatctt 2700
agctttccac aggagggaga gaacttaaaa aagcaacagt agcagggaat tgatccactt 2760
cttaatgctt tctccctgg catgaccatc ctgtccttg ttattatcct gcattttacg 2820
tctttggagg aacagctccc tagtggttc ctccgtctgc aatgtccctt gcacagccca 2880
cacatgaacc atccttccca tgatgccgct cttctgtcat cccgctcctg ctgaaacacc 2940
tcccaggggc tccacctgtt caggagctga agcccatgct tcccaccag catgtcactc 3000
ccagaccacc tccctgccct gtctccagc tcccctcgc tgtcctgctg tgtgaattcc 3060
caggttgccc tggtgccat gtgcctgcc cccagcactc ctctgtctct getcctgcct 3120
cgacccttcc tctcctttg cctagagggc cttctcgcct tttctctagc tgatcagaat 3180
tttaccaaaa ttcagaacat cctccaattc cacagtctct gggagacttt ccctaagagg 3240
cgacttctc tccagccttc tctctctggt caggcccact gcagagatgg tggtagcac 3300
atctgggagg ctggtctccc tccagctgga attgctgctc tctgaggag aggctgtggt 3360
ggctgtctct gtcctcact gccttccagg agcaatttgc acatgtaaca tagatttatg 3420
taatgcttta tgttaaaaa cattcccaa ttatcttatt taatttttgc aattattcta 3480
atttatata tagagaaagt gacctatttt ttaaaaaaat cacactctaa gttctattga 3540
acctaggact tgagctcca tttctggctt ctagtctggt gttctgagta cttgatttca 3600
ggtcaataac ggtccccct cactccacac tggcacgttt gtgagaagaa atgacathtt 3660
gctaggaagt gaccgagtct aggaatgctt ttattcaaga caccaaatc caaacttcta 3720
aatgttgaa tttcaaaaa ttgtgttag attttatgaa aaactcttct actttcatct 3780
attctttccc tagaggcaaa catttcttaa aatgtttcat tttcattaaa aatgaaagcc 3840
aaatttatat gccaccgatt gcaggacaca agcacagttt taagagttgt atgaacatgg 3900
agaggacttt tggtttttat atttctcgtt tttaatatgg gtgaacacca acttttattt 3960
ggaataataa ttttctcctt aaacaaaaac acattgagtt taagtctctg actcctgcct 4020

-continued

```
ttccacctgc tttctcctgg gcccgcttgg cctgcttgaa ggaacagtgc tgttctggag 4080
ctgctgttcc aacagacagg gcctagcttt catttgacac acagactaca gccagaagcc 4140
catggagcag ggatgtcacg tcttgaaaag cctattagat gttttacaaa ttttaattttg 4200
cagattatth tagtctgtca tccagaaaaat gtgtcagcat gcatagtgtc aagaaagcaa 4260
gccaatttgg aaacttaggt tagtgacaaa attggccaga gagggggggt gatgatgacc 4320
aagaattaca agtagaatgg cagctggaat ttaaggaggg acaagaatca atggataagc 4380
gtgggtggag gaagatccaa acagaaaagt gcaaagtat tccccatctt ccaagggttg 4440
aattctggag gaagaagaca cactctagat tccccgtgaa cttccttga cttattgtcc 4500
ccactaaaac aaaacaaaaa acttttaatg ccttccacat taattagatt ttcttgagc 4560
ttttttatgg cattttttta aagatgcctt aagtgttgaa gaagagtttg caaatgcaac 4620
aaaatattta attaccggtt gttaaaactg gtttagcaca atttatattt tccctctctt 4680
gcctttctta tttgcaataa aaggatttga gccatttttt aaatgacatt tttgataaat 4740
tatgtttgta ctagtgtatg aaggagtttt ttttaacctg tttatataat tttgcagcag 4800
aagccaaatt tttgtatat taaagcaca aattcatgta cagcatgcat cacggatcaa 4860
tagactgtac ttattttcca ataaaatttt caaactttgt actgttaaa 4909
```

<210> SEQ ID NO 31

<211> LENGTH: 2210

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 31

```
cgcgccctc cctcctcgcg gacctggcgg tgccggcgcc cggagtggcc ctttaaaagg 60
cagettattg tccggagggg gcggggcggg ggccgcgacc gcggcctgag gcccgcccc 120
tccccctccc ctcctctctg ccccgctcgc ctctctggct agctctctgg ctctctcgc 180
cgctccgcgc acgctccgcg tccgtcagtt ggctccgctg tccgggtgcgc ggcgtggagc 240
ggcagccggt ctggacgcgc ggcgggggct gggggctggg agcgcggcgc gcaagatctc 300
cccgcgcgag agcggccctt gccaccgggc gaggcctcgc ccgcgatggc agagatgggc 360
agtaaagggg tgacggcggg aaagatcgcg agcaacgtgc agaagaagct caccgcgcg 420
caggagaagg ttctccagaa gctggggaag gcagatgaga ccaaggatga gcagtttgag 480
cagtgcgtcc agaatttcaa caagcagctg acggagggca cccggctgca gaaggatctc 540
cggacctacc tggcctccgt caaagccatg cagcaggctt ccaagaagct gaatgagtgt 600
ctgcaggagg tgtatgagcc cgattggccc ggcagggatg aggcacacaa gatcgcagag 660
aacaacgacc tgctgtggat ggattaccac cagaagctgg tggaccaggc gctgctgacc 720
atggacacgt acctgggcca gttccccgac atcaagtcac geattgcca gcgggggcgc 780
aagctggtgg actacgacag tgcccggcac cactacgagt cccttcaaac tgccaaaaag 840
aaggatgaag ccaaaattgc caaggccgag gaggagctca tcaaagccca gaaggtgttt 900
gaggagatga atgtggatct gcaggaggag ctgccgtccc tgtggaacag ccgcgtaggt 960
ttctacgtca acacgttcca gagcatcgcg ggccctggagg aaaacttcca caaggagatg 1020
agcaagctca accagaaact caatgatgtg ctggctggcc tggagaagca acacgggagc 1080
aacaccttca cggatcaaggc ccagcccaga aagaaaagta aactgttttc gcggctgcgc 1140
```

-continued

```

agaaagaaga acagtgacaa cgcgcctgca aaaggaaca agagcccttc gcctccagat 1200
ggctccccctg cegccacccc cgagatcaga gtcaaccacg agccagagcc ggccggcggg 1260
gccacgcccg gggccacctc cccaagtcc ccatctcagc cagcagagcc ctccggaggtg 1320
gcggttgagg cccaacctgc ggctggagcc caggagccag gggagacggc ggcaagtga 1380
gcagcctcca gctctcttc tctgtctgtg gtggagacct tcccagcaac tgtgaatggc 1440
accgtggagg gcggcagtgg ggccggggcg ttggacctgc ccccaggttt catgttcaag 1500
gtacaggccc agcagacta caccggccact gacacagacg agctgcagct caaggctggt 1560
gatgtggtgc tggatgccc cttccagaac cctgaagagc aggatgaagg ctggctcatg 1620
ggcgtgaagg agagcgactg gaaccagcac aaggagctgg agaagtgccg tggcgtcttc 1680
cccgagaact tcaactgagag ggtcccatga cggcggggcc caggcagcct ccgggcgtgt 1740
gaagaacacc tcctcccga aatgtgtgg ttcttttttt tgttttggtt tegtttttca 1800
tcttttgaag agcaaaggga aatcaagagg agacccccag gcagaggggc gttctccaa 1860
agattaggtc gttttccaaa gagccgcgct ccggcaagtc cggcgggaatt caccagtgtt 1920
cctgaagctg ctgtgtcttc tagttgagtt tctggcgccc ctgctgtgct ccgcatgtgt 1980
gcctggccgc agggcggggc tgggggctgc cgagccacca tgcttgctg aagcttcggc 2040
cgcgccaccc gggcaagggt cctcttttcc tggcagctgc tgtgggtggg gccagacac 2100
cagcctagcc tggctctgct ccgcagacgg tctgtgtgct gtttggaaaat aaatcttagt 2160
gttcaaaaca aaatgaaca aaaaaaaaaat gataaaaact ctcaaaaaa 2210

```

<210> SEQ ID NO 32

<211> LENGTH: 4664

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 32

```

gcagcctccc ggcgctgagc gcttttctcg cccgcccggc tcagccctgc ggaccccggg 60
agaagtttcc cagaaaaaat gcccagcgcg gcgcggggct gcggagtcgt ccggagccgc 120
tgcgcgattt atcagcaaga ctgttgaacg cataactgcc caagatgctt gtccctcccc 180
ctccagcacc cccgcgcgcc ccgacgtttg cactggccaa tacagagaag cctaccttga 240
ataagacaga gcaggctggg agaaatgctc tcctttctga tatcagcaaa ggaagaaac 300
taaagaagac ggtcaccaat gacagaagtg caccaatact ggacaaaact aaaggagctg 360
gtgctggagg cgggtggtgt ggctttggtg gagcggcggc atttggcggg ggaggtggtg 420
gcggaggcgg tggaaagttt ggagggggcg gacctccagg tctgggagga ttgttccagg 480
ctggaatgcc gaagctgaga tccacggcca acagggataa tgattctgga ggaagccgac 540
caccattgtt gccaccggga ggaagatcca catctgcaaa acccttttca ccccagggtg 600
gcccagggag gtttctctg ccttctccag gccacagaag tgggtcccca gagcctcaga 660
ggaaccgaat gccgccccca aggcccagc tgggctcaaa gcctgatagc attcctctc 720
cagtlacctag tactccaaga cccattcaat caagtccgca caaccggggg tccccaccag 780
tgcccggagg ccccaggcag cccagccccg ggcccactcc tcccccttcc cctggaaaac 840
gcggcactgc tttgggagga ggctcaatac gtcagtcccc cttgagctcc tectcgccct 900
tctcaaacg gcctccccct cgcctaccc ccagcagggc cttggatgac aaaccctct 960

```


-continued

caccacctcc tccagtgggc aacaggccct ccatccacag ggaagcggtt cccctcctc	1020
ctctcagaa caacaagcct ccagtgcctt ccaactccgc gccttcggcc tctcaccagg	1080
ccccacctcc gccgccacct cccagcaggc ccgggcccgc tctctgctt ccaagtcca	1140
gcggaatga cgaacccca agactccac agcggaatct gtcctcagt tegtccacgc	1200
ccccgttacc ttgccagga cgttcaggtc ctcttctcc cccgccagt gagagacccc	1260
cacctcagt gagggaccgc ccaggccgat caggccccct cccaccact cctccagtaa	1320
gcagaaacgg cagcacatct cgggccctgc ctgctacccc tcagttgcca tccaggagt	1380
gagtagacag tcccaggagt ggacccaggc ctccccttc tctgatagg cccagtgtg	1440
gggcacctcc cccacctcca ccatcaacat ctattagaaa tggcttccaa gactctccat	1500
gtgaagatga gtgggaaagc agattctact tccatccgat ttccgatttg ccacctccag	1560
agccatatgt acaaacgacc aaaagttatc ccagcaaac ggcaagaaac gaaagccgga	1620
gtggatccaa ccgaagagaa aggggtgctc caccactccc tccatcccc aggtgatctt	1680
tgctgtctct tctctacca agctcaagag ctgctctgt tgctatctaa gaactgcata	1740
ccctcctccc tgettcttc cttgtgctc atgtatgggc aggaggaaag gtgggagggg	1800
gagtgggaat atgcgtgtgt ggggtgggat cggtaagaaa tgcacctagc tttcatatt	1860
gtgtttatc tccaggctat tgcttgctc agctgcagcc tgctgtgct ggtgtgagg	1920
gtcgataggc ttttgcgta ataggcagag atgactgca tcccagcttt ccaccaacca	1980
aattcaaaac ttcactgctt atttgttaca gactgtaatt attaaagtcc ctgagagctg	2040
ttttctccc ttctttttc gcactgttg cctcctctct gtttctatga accacagacc	2100
acctaaagca gctgctgagt aagggtcac tggaaacttg cagtcacagg atgtccaatc	2160
tttggcagtc cgagcttggc tctaggacag agctgtccaa tagaaatata atgtgagccc	2220
catatacaat ttttacattt ctaatatatt ttaaacaagt gaagtaata tgcatacaaa	2280
atatttcaac ctgtaatcaa cataaaattt taatgagata ttttatatta tttttggta	2340
ctgaatcttc aaaaaccaga gtgtatttta cacttacgc acatctccat tcagactagt	2400
cacatttta agtgcctagt agccacatgt ggctggtggc tactggatta gacagcacga	2460
gtctggaaga tggaagctag tgcagaaacc tctgtttta aaaacaaaa aggcaagatg	2520
ggcttgagcg attcaagagg caactaaaa taaaattagg acccagcacc ttgtttgaca	2580
cacagttga ccttcgattt tctccctta acttcctct tcccttaata tctgtataca	2640
agtgtgctt caaagtacca aggtcagaaa ttgattoagt acggtttact aaagtcatgt	2700
ggaataaagc cattggaaac aaatggaaag cctgtcggga cttctgggct cagaaccagc	2760
tggtcagc actccacttg tcagctggac ttctgcctg tgaatggaa gcagccttg	2820
ttcctttctg gctgagcaag ctccagaggc tgggagagac taggaaggct tggtaggagg	2880
ggaaaaaagt caggaaaaa tatcaaatca gaaacatgga agaagaaggg aaccgatttg	2940
agttggtggg caaaactcta aaaatctaaa tctgatgctt atgtaagggt tgagcgaatt	3000
aggagattg ctagtggaaa ttggagggaa tttgtttgc atcatttgc taggatctat	3060
gcaaatatag ctccactaaa ggaccatagg gaagagccag ccttccttt tcttatatga	3120
ttttgttac aaaattttc tgggactttt aaatctagct atagagttgg gaaaaaatat	3180
ttccacttag atattttaca tggttttgtt taaaattacc attacttgtt ttttaaaac	3240

-continued

```

acatgaccac atatgtatat gtatatctac ctaaaccattg tatcatgggt tcagtatggt 3300
attcatgtat tactgggaga tgctaccaag aaaccaaccc aaagaaaatt ctgaaaaata 3360
catttctatt tatagaataa atgtttcatt tatataaaag caaaagaact tagagtteta 3420
ataaatggga tgtetaataa attatgaagt tactgatttg aatatattat atttttataa 3480
cttcttggc aaagtctga tttagtacat tagagaaact gtgtttcttc tctcctctac 3540
cattcatctc tcttccatac agtcatttgg gctttttact caaagagaat caagaaataa 3600
taaggataa caagcttggc aaagtgttg ctttttaaaa aaaaattttt ttaactctta 3660
gcagtttgg aatttagcag catcatttat ttgggattct tttatctgat tccaacagt 3720
aaaaacatcc ctatgataaa gcctaataac ccatttcaca aaagatggaa tttgcccttc 3780
ctagaaaata tgacggagaa aagtctgact cagagaaagt gagtctgaat tttataagg 3840
gtagtaagaa ttggacaatt cctttgcata tctgaacttg gcaggtaacc ttctaaatct 3900
gaaacagggt gatagctcaa agttgccatt catccagaat agattgtttt agaattgagt 3960
gtttaagtga ctgtttcatt aatacaccta caccctttct ttgaaagttt gcaacctaat 4020
tgcactaaa actatgaata agttctgttg taaaatctta aactatggaa aattacaaaa 4080
atgaattttt cttccctgaa atcagagctt acatgtgtgt tttttataa cattttcaga 4140
taaatgtatt caacatgtaa tacagtattt taacattcac ctcttatttt atattgaaat 4200
gtattacagt attaaaactc agtgttcagt atttatttca ctatgcattt tatttagtaa 4260
aagccaggag aaatgtttta tccaatgggt ccttactttg tgatttaaaa gaaatcaact 4320
tttttttatg tctaagtagt agattatttg catatttga aaaactgtta ggtctttata 4380
ttttaaagt taataccagt tttgttatt tagtagcaga aatgggatga ttgttaaagt 4440
tccccaaaaa tgttggcatg aaattaattt tccctcctt atagtcaagg accgtagagg 4500
aagaaaaact ttttttcat accatgcact atgtaaacag acacattttg ctatctgtgt 4560
catcaggata gtgtaagtgg taggtagag actaccctag acatctgcat ctttgtaagt 4620
tagccagaca ataaagaaaa gcagaatgaa aaaaaaaaaa aaaa 4664

```

<210> SEQ ID NO 33

<211> LENGTH: 1166

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 33

```

attcccaact gccagtgatc tctgaagccg actctgagge tccctctttg ctctaacaga 60
cagcagcgac tttaggctgg ataatagta aattcttacc tcgctcttcc actgctagta 120
agatcagatt gcgtttcttt cagttactct tcaatcgcca gtttcttgat ctgcttctaa 180
aagaagaagt agagaagata aatcctgtct tcaatacctg gaaggaaaaa caaataaacc 240
tcaactccgt tttgaaaaaa acattccaag aactttcctc agagatttta cttagatgat 300
ttacacaatg aagaaagtac atgcactttg ggcttctgta tgctgctgct ttaactcttg 360
cctgccccct cttaatgctg attctgagga agatgaagaa cacacaatta tcacagatac 420
ggagttgcca ccaactgaaac ttatgcattc attttgtgca ttcaaggcgg atgatggccc 480
atgtaaagca atcatgaaaa gatttttctt caatatttcc actcgacagt gegaagaatt 540
tatatatggg ggatgtgaag gaaatcagaa tcgattttaa agtctggaag agtgcaaaaa 600

```

-continued

```

aatgtgtaca agagataatg caaacaggat tataaagaca acattgcaac aagaaaagcc 660
agatttctgc tttttggaag aagatcctgg aatatgtcga ggttatatta ccaggatatt 720
ttataacaat cagacaaaac agtgtgaacg tttcaagat ggtggatgcc tgggcaatat 780
gaacaatfff gagacactgg aagaatgcaa gaacatttgt gaagatggtc cgaatggttt 840
ccaggtggtt aattatggaa cccagctcaa tgctgtgaat aactccctga ctccgcaatc 900
aaccaaggtt cccagccttt ttgttacaaa agaaggaaca aatgatggtt ggaagaatgc 960
ggctcatatt taccaagtct ttctgaacgc cttctgcatt catgcatcca tgttctttct 1020
aggattggat agcatttcat gcctatgta atatttgtgc ttttggcatt tccttaatat 1080
ttatatgtat acgtgatgcc tttgatagca tactgctaata aaagttttaa tattttacatg 1140
catagtaaaa aaaaaaaaaa aaaaaa 1166

```

<210> SEQ ID NO 34

<211> LENGTH: 8449

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 34

```

gccccgcgcg gctgtgctgc acagggggag gagagggaac cccagggcgcg agcgggaaga 60
ggggacctgc agccacaact tctctggtec tctgcatccc ttctgtccct ccaccgctcc 120
ccttccccac cctctggccc ccaccttctt ggaggcgaca acccccggga ggcattagaa 180
gggatttttc ccgaggttg cgaagggaa caaacttggg ggcaacttgc ctcccggtgc 240
gggcgctctc cccccaccgt ctcaacatgc ttaggggtcc ggggcccggg ctgctgctgc 300
tggccgtcca gtgcctgggg acagcgggtc cctccacggg agcctcgaag agcaagaggc 360
aggetcagca aatggttcag ccccagtcce cggtggtgtg cagtcaaaagc aagcccggtt 420
gttatgacaa tggaaaacac tatcagataa atcaacagtg ggagcggacc tacctaggca 480
atgcgttggg ttgtacttgt tatggaggaa gccgaggttt taactgcgag agtaaacctg 540
aagetgaaga gacttgcttt gacaagtaca ctgggaacac ttaccgagtg ggtgacactt 600
atgagcgtcc taaagactcc atgatctggg actgtacctg catcggggct gggcgagggga 660
gaataagctg taccatcgca aaccgctgcc atgaaggggg tcagtcctac aagattggtg 720
acacctggag gagaccacat gagactggtg gttacatggt agagtgtgtg tgtcttggtg 780
atggaaaagg agaatggacc tgcaagccca tagctgagaa gtgttttgat catgctgctg 840
ggacttccta tgtggtcggg gaaacgtggg agaagcccta ccaaggctgg atgatggtag 900
attgtacttg cctgggagaa ggcagcggac gcatcacttg cacttctaga aatagatgca 960
acgatcagga cacaaggaca tcctatagaa ttggagacac ctggagcaag aaggataatc 1020
gaggaaaact gctccagtgc atctgcacag gcaacggccg aggagagtgg aagtgtgaga 1080
ggcacacctc tgtgcagacc acatcgagcg gatctggccc cttcaccgat gttcgtgcag 1140
ctgtttacca accgcagcct cacccccagc ctctcccta tggccactgt gtcacagaca 1200
gtggtgtggt ctactctgtg gggatgcagt ggctgaagac acaaggaat aagcaaatgc 1260
tttgacgtg cctgggcaac ggagtcagct gccaaagagac agctgtaacc cagacttacg 1320
gtggcaactc aaatggagag ccatgtgtct taccattcac ctacaatggc aggacgttct 1380
actctgcac cacagaaggg cgacaggacg gacatcttgg gtgcagcaca acttcgaatt 1440

```

-continued

atgagcagga ccagaaatac tctttctgca cagaccacac tgttttggtt cagactcgag 1500
gaggaaattc caatggtgcc ttgtgccact tccccttctc atacaacaac cacaattaca 1560
ctgattgcac ttctgagggc agaagagaca acatgaagtg gtgtgggacc acacagaact 1620
atgatgccga ccagaagttt ggggtctgcc ccattgctgc ccacgaggaa atctgcacaa 1680
ccaatgaagg ggtcatgtac cgcattggag atcagtggga taagcagcat gacatgggtc 1740
acatgatgag gtgcacgtgt gttgggaatg gtcgtgggga atggacatgc attgcctact 1800
cgcagcttcg agatcagtgc attgttgatg acatcactta caatgtgaac gacacattcc 1860
acaagcgtca tgaagagggg cacatgctga actgtacatg cttcggtcag ggtcggggca 1920
ggtggaagtg tgatcccgtc gaccaatgcc aggattcaga gactgggacg ttttatcaaa 1980
ttggagattc atgggagaag tatgtgcatg gtgtcagata ccagtgtac tgctatggcc 2040
gtggcattgg ggagtggcat tgccaacctt tacagaccta tccaagctca agtggtcctg 2100
tcgaagtatt tatcactgag actccgagtc agcccaactc ccaccccatc cagtggaatg 2160
caccacagcc atctcacatt tccaagtaca ttctcaggtg gagacctaaa aattctgtag 2220
gccgttgga ggaagctacc ataccagcc acttaaaactc ctacaccatc aaaggcctga 2280
agcctggtgt ggtatacagag ggcacagctca tcagcatcca gcagtacggc caccaagaag 2340
tgactcgtt tgacttcacc accaccagca ccagcacacc tgtgaccage aacaccgtga 2400
caggagagac gactcccttt tctcctcttg tggccacttc tgaatctgtg accgaaatca 2460
cagccagtag ctttgtggtc tccctgggtc cagcttccga caccgtgtcg ggattccggg 2520
tggaaatga gctgagttag gaggagatg agccacagta cctggatctt ccaagcacag 2580
ccacttctgt gaacatccct gacctgttc ctggccgaaa atacattgta aatgtctatc 2640
agatatctga ggatggggag cagagtttga tccgtctac ttcacaaaaca acagcgctg 2700
atgcccctcc tgaccagct gtggaccaag ttgatgacac ctcaattgtt gttcgtgga 2760
gcagacccca ggtcccctc acagggtaaca gaatagtcta ttcgcatca gtagaaggta 2820
gcagcacaga actcaacctt cctgaaactg caaactccgt caccctcagt gacttgcaac 2880
ctggtgttca gtataacatc actatctatg ctgtggaaga aaatcaagaa agtacacctg 2940
ttgtcattca acaagaaacc actggcacc cagctcaga tacagtgcc tctcccaggg 3000
acctgcagtt tgtggaagtg acagacgtga aggtcaccat catgtggaca ccgcctgaga 3060
gtgcagtac cggctaccgt gtggatgtga tcccctcaa cctgctggc gagcacgggc 3120
agaggctgcc catcagcagg aacaccttg cagaagtca cgggctgtcc cctggggctca 3180
cctattactt caaagtcttt gcagtgacc atgggagga gagcaagcct ctgactgctc 3240
aacagacaac caaactggat gctcccacta acctccagt tgtaatgaa actgattcta 3300
ctgtcctggt gagatggact ccacctcggg cccagataac aggataccga ctgacctgg 3360
gccttaccgg aagaggacag cccaggcagt acaatgtggg tccctctgtc tccaagtacc 3420
cactgaggaa tctgcagcct gcactctagt acaccgtatc cctcgtggc ataaagggca 3480
accaagagag ccccaaagcc actggagtct ttaccacact gcagcctggg agctctattc 3540
caccttaca caccgagtg actgagacca ccattgtgat cacatggacg cctgctocaa 3600
gaattggtt taagctgggt gtacgacaa gccagggagg agaggacca cgagaagtga 3660
cttcagactc aggaagcctc gttgtgtccg gcttgactcc aggagtagaa tacgtctaca 3720

-continued

ccatccaagt cctgagagat ggacaggaaa gagatgogcc aattgtaaac aaagtgggta	3780
caccattgtc tccaccaaca aacttgcatc tggaggcaaa ccctgacact ggagtgtctca	3840
cagtctctcg ggagaggagc accaccccag acattactgg ttatagaatt accacaaccc	3900
ctacaaaacgg ccagcaggga aattctttgg aagaagtggc ccatgctgat cagagetcc	3960
gcacttttga taacctgagt ccggcctgg agtacaatgt cagtgtttac actgtcaagg	4020
atgacaagga aagtgtccct atctctgata ccatcatccc agctgttctc cctcccactg	4080
acctgcgatt caccaacatt ggtccagaca ccatgcgtgt cacctgggct ccaccccac	4140
ccattgattt aaccaacttc ctggcgcgtt actccactgt gaaaaatgag gaagatggtg	4200
cagagtgtc aatttctcct tcagacaatg cagtgttctt aacaaatctc ctgcctggta	4260
cagaatatgt agtgagtgtc tccagtgtct acgaacaaca tgagagcaca cctcttagag	4320
gaagacagaa aacaggctctt gattccccc a ctggcattga ctttctctgat attactgcca	4380
actcttttac tgtgcaactgg attgctcctc gagccaccat cactggctac aggatccgcc	4440
atcatcccga gcacttcagt gggagacctc gagaagatcg ggtgcccac tctcggaatt	4500
ccatcacctt caccaacctc actccaggca cagagtatgt ggtcagcatc gttgctctta	4560
atggcagaga ggaagtccc ttattgattg gccacaatc aacagttctt gatggtccga	4620
gggacctgga agttgttggc gcgaccccc cccagcctact gatcagctgg gatgctcctg	4680
ctgtcacagt gagatattac aggatcactt acggagagac aggaggaaat agccctgtcc	4740
aggagtccac tgtgcctggg agcaagtcta cagctaccat cagcggcctt aaacctggag	4800
ttgattatac catcactgtg tatgctgtca ctggcctgg agacagcccc gcaagcagca	4860
agccaatttc cattaattac cgaacagaaa ttgacaaaac atcccagatg caagtgaccg	4920
atgttcagga caacagcatt agtgtcaagt ggctgccttc aagttccctt gttactgggt	4980
acagagtaac caccactccc aaaaatggac caggaccaac aaaaactaaa actgcaggtc	5040
cagatcaaac agaagtact attgaaggct tgcagcccac agtggagtat gttggttagtg	5100
tctatgctca gaatccaagc ggagagagtc agcctctggt tcagactgca gtaaccaaca	5160
ttgatcgccc taaaggactg gcattcactg atgtggatgt cgattccatc aaaattgctt	5220
gggaaagccc acaggggcaa gtttccaggt acaggggtgac ctactcgagc cctgaggatg	5280
gaatccatga gctattccct gcacctgatg gtgaagaaga cactgcagag ctgcaaggcc	5340
tcagaccggg ttctgagtac acagtcatg tggttgcctt gcacgatgat atggagagcc	5400
agcccctgat tggaaccag tccacagcta ttctgcacc aactgacctg aagttcactc	5460
aggtcacacc cacaagcctg agcgcaccgt ggacaccacc caatgttcag ctcactggat	5520
atcgagtgcg ggtgaccccc aaggagaaga ccggaccaat gaaagaaac aacctgtctc	5580
ctgacagetc atccgtggtt gtatcaggac ttatggtggc caccaaatat gaagtgagtg	5640
tctatgctct taaggacact ttgacaagca gaccagctca gggagtgtc accactctgg	5700
agaatgtcag cccaccaaga agggctcgtg tgacagatgc tactgagacc accatcacca	5760
ttagtggag aaccaagact gagacgatca ctggcttcca agttgatgcc gttccagcca	5820
atggccagac tccaatccag agaaccatca agccagatgt cagaagctac accatcacag	5880
gtttacaacc aggcactgac tacaagatct acctgtacac cttgaatgac aatgctcgga	5940
gctcccctgt ggtcatcgac gcctccactg ccattgatgc accatccaac ctgcgtttcc	6000

-continued

tggccaccac acccaattcc ttgctggat catggcagcc gccacgtgcc aggattaccg 6060
gctacatcat caagtatgag aagcctgggt ctctcccag agaagtggte cctcggcccc 6120
gccttggtgt cacagaggct actattactg gcctggaacc gggaaaccgaa tatacaattt 6180
atgtcattgc cctgaagaat aatcagaaga gcgagcccct gattggaagg aaaaagacag 6240
acgagcttcc ccaactggta acccttcac accccaatct tcatggacca gagatcttgg 6300
atgttccttc cacagttaa aagaccctt tcgtcaccca cctgggtat gacactggaa 6360
atggtattca gcttctggc acttctggte agcaaccag tgttgggcaa caaatgatct 6420
ttgaggaaca tggttttagg cggaccacac cgcccacaac ggccaccccc ataaggcata 6480
ggccaagacc ataccggccg aatgtaggac aagaagctct ctctcagaca accatctcat 6540
gggccccatt ccaggacact tctgagtaca tcatttcctg tcatcctggt ggcactgatg 6600
aagaaccctt acagttcagg gttcctggaa cttctaccag tgccactctg acaggcctca 6660
ccagagggtgc cacctacaac atcatagtgg aggcactgaa agaccagcag aggcataagg 6720
ttcgggaaga ggttgttacc gtgggcaact ctgtcaacga aggccttgaac caacctacgg 6780
atgactctg ctttgacccc tacacagttt cccattatgc cgttggagat gagtgggaac 6840
gaatgtctga atcaggcttt aaactgtgt gccagtgtt aggccttggaa agtggctcatt 6900
tcagatgtga ttcactaga tgggtccatg acaatgggtg gaactacaag attggagaga 6960
agtgggaccg tcagggagaa aatggccaga tgatgagctg cacatgtctt gggaaacggaa 7020
aaggagaatt caagtgtgac cctcatgagg caacgtgtta tgatgatggg aagacatacc 7080
acgtaggaga acagttggcag aaggaatate tcggtgccat ttgctcctgc acatgctttg 7140
gaggccagcg gggctggcgc tgtgacaact gccgcagacc tgggggtgaa cccagtcocg 7200
aaggcactac tggccagtc tacaaccagt attctcagag ataccatcag agaacaaca 7260
ctaagttaa ttgcccatt gagtgttca tgcccttaga tgtacaggct gacagagaag 7320
attcccgaga gtaaatcctc tttccaatcc agaggaacaa gcattgtctct ctgccaaagat 7380
ccatcctaac tggagtgatg ttagcagacc cagcttagag ttctctcttc tttcttaage 7440
cctttgctct ggaggaagtt ctccagcttc agctcaactc acagcttctc caagcatcac 7500
cctgggagtt tctgagggt tttctcataa atgagggctg cacattgect gttctgcttc 7560
gaagtattca ataccgctca gtattttaa tgaagtgatt ctaagatttg gtttgggatc 7620
aataggaaag catatgcagc caaccaagat gcaaatgttt tgaatgata tgacaaaaat 7680
tttaagtagg aaagtcaccc aaacacttct gctttcactt aagtgtctgg cccgcaatac 7740
tgtaggaaca agcatgatct tggttactgtg atattttaa tatccacagt actcactttt 7800
tccaatgat cctagtaatt gcctagaaat atctttctct tacctgttat ttatcaattt 7860
ttcccagtat ttttatacgg aaaaaattgt attgaaaaca cttagtatgc agttgataag 7920
aggaatttgg tataattatg gtgggtgatt attttttata ctgtatgtgc caaagcttta 7980
ctactgtgga aagacaactg ttttaataaa agatttcat tccacaactt gaagttcatc 8040
tatttgatat aagacacctt cgggggaaat aattcctgtg aatattcttt ttcaattcag 8100
caaacatttg aaaatctatg atgtgcaagt ctaattgttg atttcagtac aagattttct 8160
aaatcagttg ctacaaaaac tgattgggtt ttgtcacttc atctctcacc taatggagat 8220
agctttacac tttctgcttt aatagattta agtggacccc aatatttatt aaaattgcta 8280

-continued

```

gtttaccggt cagaagtata atagaataa tctttagtgt ctcttttcta accattgtaa 8340
ttcttccctt ctctcccca cctttccttc attgaataaa cctctgttca aagagattgc 8400
ctgcaaggga aataaaaatg actaagatat taaaaaaaa aaaaaaaaa 8449

```

<210> SEQ ID NO 35

<211> LENGTH: 4625

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 35

```

agcgtccgc agtcacgtga cgctcgtccg caacctotgc tgtcctccgc ggcgccccct 60
tccgcctgac ggcgccccgc cggcgccgcg gcagccctgg ctctcgcgg gctcggggcg 120
cggtgcggc ggggctatgg cgagcggcgg tggcggggg aacactggcg cgggtggggg 180
gccggggaag ggcctgagcc tgggctggg tctgggtctg agcctaggca tgagtgggg 240
caccagtggg gcagaggagg aggcggccac ggcggaggcg gtgggacgcc tggccacgac 300
gctgtggctg cggctccgcg gctgggaggc ggtgctggcg gcggcgcagc ggttgctggg 360
gtgggagaag ccgctgcaca gcctggctac ggcggccgcg ctcaacggcc tcttctgggt 420
gctgtcttcc tcgtccctcc ggcctctctt cctactcagc gtctcacttt tggcctattt 480
tctgctggat ctctggcagc ctgcctttct cctgacggtt tcagcactcat ccccagagga 540
gccacactct gacagtggag gtgcggggtc aggcgccccg ccgcacctgc tgagtgtgcc 600
cgagttgtgc agatacctgg ctgagagctg gctcaccttc cagattcacc tgcaggagct 660
gctgcagtac aagaggcaga atccagctca gttctgctgt cgagttctgt ctggctgtgc 720
tgtgttggct gtgttgggac actatgttcc agggattatg atttctaca ttgtcttgtt 780
gagtatcctg ctgtggcccc tgggtggtta tcatgagctg atccagagga tgtacactcg 840
cctggagccc ctgctcatgc agctggacta cagcatgaag gcagaagcca atgccctgca 900
tcacaaacac gacaagagga agcgtcaggg gaagaatgca cccccaggag gtgatgagcc 960
actggcagag acagagagtg aaagcgaggc agagctggct ggcttctccc cagtggtgga 1020
tgtgaagaaa acagcattgg ccttggccat tacagactca gagctgtcag atgaggaggc 1080
ttctatcttg gagagtggtg gcttctccgt atccccggcc acaactccgc agctgactga 1140
tgtctccgag gatttggacc agcagagcct gccaaagtga ccagaggaga ccctaagccg 1200
ggacctaggg gagggagagg agggagagct ggcctctccc gaagacctac taggcccgtc 1260
tcaagctctg tcaaggcaag ccctggactc ggaggaagag gaagaggatg tggcagctaa 1320
ggaaaacctg ttgcggctct catccccctt ccactttgtg aacacgcact tcaatggggc 1380
agggctcccc ccagatggag tgaatgctc cctggagga ccagtggaga cactgagccc 1440
cgagacagtg agtgggtggc tcaactgctc gcccggcacc ctgtcacctc caetttgctc 1500
tgttgggaag gacccagccc cctccccctt cattctccca cctgttcccc aggactcacc 1560
ccagccccct cctgccccct aggaagaaga ggcactcacc actgaggact ttgagttgct 1620
ggatcagggg gagctggagc agctgaatgc agagctgggc ttggagccag agacaccgcc 1680
aaaaaaaaat gatgctccac ccctggggcc cgacatccat tctctggtac agtcagacca 1740
agaagctcag gccgtggcag agccatgagc cagccgttga ggaaggagct gcaggcacag 1800
tagggcttcc tggttaggag tgttctgtgt tctctcttgg cctaccactc tggggtgggg 1860

```

-continued

cagtggtgg ggaagctggc tgtcggatgg tagctattcc accctctgcc tgccctgctg	1920
cctgctgtcc tgggcctggg gcagtaacctg tgccctaggat tggttttaaa tttgtaata	1980
attttccatt tgggttagtg gatgtgaaca gggctagggg agtccttccc acagcctgcg	2040
cttgccctccc tgccctcatc ctattctcat tccaactatgc cccaagccct ggtggtctgg	2100
ccctttcttt ttccctctat cctcagggac ctgtgctgct ctgccctcat gtcccacttg	2160
gttgttagt tgaggcactt tataattttt ctcttgctct gtgttccttt ctgctttatt	2220
tcctgctgt gtccctgctc tagcagctca accccatcct ttgccagctc ctccatccc	2280
gtgggcactg gccaaagcttt agggaggctc ctggtctggg aagtaaagag taaacctggg	2340
gcagtggtgc aggccagtag ttacactctt aggtcactgt agtctgtgta acctcactg	2400
catccttgcc ccattcagcc cggcctttca tgatgcagga gagcagggat ccgcagtag	2460
atggcgccag cactggagtt ggtgagcatg tgctctctct tgagattagg agcttcccta	2520
ctgctctctc ggggtgatcca agtgtagtgg gacccccctac tagggctcagg aagtggacac	2580
taacatctgt gcaggtggtg acttgaaaaa taaagtgttg attggctaga actgctgcct	2640
ccctgactgt gagctgcctc ccacacctg cactgcaactg tgttctctcc tcaccttaa	2700
cctgcttcc tccagctctgt tctggctgtt tattaccttg ttgcataaca gggccgaagc	2760
aaggattacc ttgacaacct tagcttctcc ttagecatct tccttgacag tgtgatctgt	2820
ttagtgagat ttagcatgtg tgaataaagt atatgcagga ggaaattgct ttgtctccc	2880
aatcggtaga aattcggggc cataaaaatt gtgttttacc atgtggccta caacctaac	2940
actgctttct taagaagtct tcacctatct acatgctaac aactcactca gcttgattt	3000
atcttactg ggaagccaa acaagcaata gaggacctt acctgtgta gaaatgagtt	3060
ggagccaagg aacctgaag aaatagtatc ttaacagtta ctgagccat tgtatgtgt	3120
tggctctgct ctgagtgatt tatatgtatt aagatttttc ctcacaggtc agatatatac	3180
tgttactaac ttcattttt agacaggtta agcttctctg aggccacagg tcccagtaa	3240
ttgtggagcc agaaccaaa cccaagaagt tttggcttca gcaaatgcat cagacagccc	3300
ctgtccatta atagggcaca ggtaggaaga tgcacaagga tgtgggaact atagagaacc	3360
aatctgatgc cttggcttaa caaagagtgg acatggcaag ccttctctt tggggaagaa	3420
aagcccagaa ctgagcagat ggcctcctt atgagttcat gtcctccgcc ttcagctgga	3480
ggtaccatat gggagtgcta cctgtcttcc tgctggaggt acccatgggt aatgctgcct	3540
ggctgtctgc tggaggtacc atatggtaat gctgcctgto tttctgaggt tgacttttat	3600
gccatgtctt tectaagtgt gtaagaattt ttctgtttgc ttcacattg actgagaatc	3660
attctagggt ttgattgagc cctgtctctg tgccactaaa ggaactcgaa ctttctatca	3720
cttagagatt tcagagggga atggaaaaac agttctaato aataagcaag caattcaaga	3780
aaaatagaat taatcaggca atgactgcaa catgtctat cttaaatcta tttctttatt	3840
aagcttgagc attgacaata gaaccagaag cttgtagctg gatcaaaata ttctccatag	3900
gctggaggt tcgatggggt ctattctttt gttgtgttg ttttggttt ttgtttttg	3960
tgggtttttt ttttttttt tttgagacgg agtctgttgc tgttgcccag gctggagtg	4020
aatggtgcag tcttggttca ctgcaacctc tgccctccag gttcaaaaca ttctctgcc	4080
tcagccgtcc aagtagctgg gattacaggt gcatgccacg atgctggct attttttga	4140

-continued

```

tttttagtag aggtgggggtt tcaccatggt gccaggtg gtctcgaact cctgacctca 4200
ggtgattcac ccacctcggc ctcccaaagt gctgggatta caggtgtgag ccacggcgcc 4260
cagcctcatg agggctctatt ctttacattc accatggtct gatggttget acatgtttgt 4320
ctatgatttt tttttctat tatcaggtgt cttggccggt tcatgcecca cgatgaaagg 4380
gccagagggtt ttcatatgag taaaagaaaa aagcagaaat gtgaaacctca caattaggct 4440
aaacaaaaat caactggaaa agtacaggct gaggggagaa gagttggcta catgtttatg 4500
ttaggggagg agggagtaca ttttagctat gtattcaaac agtaaatagt ttaatgttgc 4560
tgcttataaa cttaatttta ggctgcatta ataaaagtgt agtctccaaa acaaaaaaaaa 4620
aaaaa 4625

```

<210> SEQ ID NO 36

<211> LENGTH: 7556

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 36

```

gcaggcgect tcgcgaccg agcctgacgg agccggaggc tgggagccgc ggcggcctgg 60
ggaagtgttt ggattgtgag ctatttcaga actgttctca ggactcatta ttttaacatt 120
tgggagaaac acagccagaa gatgcacact tgactgaagg aggacagga atctgaagac 180
tccggatgac atcagageta cttttcaaca gccttctcaa ttttctttct cagaaagcag 240
aggetcagag cttggagaca gacgaacct gatatttgca tttaatgggg aacaaaagat 300
gaagaaggaa aaggaatata ttcactaagg attctatctg cttactgcta cagacctatg 360
tgtaaggaa ttcttctcct cctccttgcg tagaagtga tcagcactgt ggtcagactg 420
catttatctt gtcattgcca gaagaaatct tggacagaat gtaacagtac gtctctctct 480
gattgcatg gaaggtgata aactgatact cctttattaa agttacatcg cactcaccac 540
agaaaacct tctttaaagt gaatagaaac caagcccttg tgaacactte tattgaacat 600
gactcatgga gaagagctg gctctgatgt gcaccaggat tctattgttt taacttacct 660
agaaggatta ctaatgcac aggcagcagg gggatcaggt actgccgttg acaaaaagtc 720
tgctgggcat aatgaagagg atcagaactt taacatttct ggcagtgcac tcccacctg 780
tcaaagtaat ggtccagttc tcaatacaca tacatatcag gggctctggca tgctgcacct 840
caaaaaagcc agactgttgc agtctcttga ggactggaat gcagcaaaagc ggaagaggct 900
gtctgattct atcatgaatt taaacgtaaa gaaggaagct ttgctagctg gcatggttga 960
cagtgtgect aaaggcaaac aggatagcac attactggcc tctttgctte agtcattcag 1020
ctctaggctg cagactgttg ctctgtcaca acaaatcagg cagagcctca aggagcaagg 1080
atatgccctc agtcatgatt ctttaaaagt ggagaaggat ttaaggtgct atggtgttgc 1140
atcaagtcc ttaaaaactt tgttgaagaa aagtaagtt aaagatcaaa agcctgatac 1200
gaatcttct gatgtgacta aaaacctcat cagagatagg tttgcagagt ctctcatca 1260
tgttggacaa agtgaacaa aggtcatgag tgaaccgttg tcatgtgctg caagattaca 1320
ggctgttga agcatggtgg aaaaaaggc tagtcctgcc acctcaccta aacctagtgt 1380
tgcttgtagc cagttagcat tacttctgtc aagcgaagcc catttgacgc agtattctcg 1440
agaacacgct taaaaaacgc aaaatgcaaa tcaagcagca agtgaaagac ttgctgctat 1500

```

-continued

ggccagattg caagaaaatg gccagaagga tgttggcagt taccagctcc caaaaggaat 1560
gtcaagccat cttaatgggc aggcaagaac atcatcaagc aaactgatgg ctgcaaaaag 1620
tagtgetaca gtgtttcaaa atccaatggg tatcattcct tcttccccta aaaatgcagg 1680
ttataagaac tcactggaaa gaaacaatat aaaacaagct gctaacaata gtttgetttt 1740
acatcttctt aaaagccaga ctatacctaa gccaatgaat ggacacagtc acagtgagag 1800
aggaagcatt tttgagaaa gtagtacacc tacaactatt gatgaatatt cagataacaa 1860
tcctagtttt acagatgaca gcagtgggta tgaaagtctt tattccaact gtgttcccat 1920
agacttgctc tgcaaacacc gaactgaaa atcagaatct gaccaacctg tttccctgga 1980
taacttcaact caatccttg taaacacttg ggatccaaaa gtcccagatg tagatatcaa 2040
agaagatcaa gatacctcaa agaattctaa gctaaactca caccagaaag taacacttct 2100
tcaattgcta cttggccata agaatgaaga aaatgtagaa aaaaacacca gccctcaggg 2160
agtacacaat gatgtgagca agttcaatac acaaaattat gcaaggactt ctgtgataga 2220
aagccccagt acaaatcgga ctactccagt gagcactcca cctttactta catcaagcaa 2280
agcaggtctc cccatcaact ctctcaca caactctggtc atcaaatgga attccccacc 2340
atatgtctgc agtactcagt ctgaaaagct acaaaatact gcactaacc actcaatgga 2400
ccttcaaaaa agcaaaagacc caccaggaga gaaaccagcc caaaatgaag gtgcacagaa 2460
ctctgcaacg tttagtcca gtaagctgtt acaaaatcta gcacaatgtg gaatgcagtc 2520
atccatgca gtggaagagc agagaccag caaacagctg ttaactggaa acacagataa 2580
accgataggt atgattgata gattaaatag ccttttgctc tcaataaaaa caaatgcagt 2640
tgaagaaaat aaagcattta gtagtcaacc aacaggtcct gaaccagggc tttctgggtc 2700
tgaatagaa aatctgcttg aaagacgtac tgtctccag ttgctcctgg ggaacccaa 2760
caaagggag agtgaaaaa aagagaaaac tcccttaaga gatgaaagta ctgagaaaca 2820
ctcagagaga gctttaagtg acaaaatact gatggtgaaa ataaaatctg agccttggtg 2880
tgacttaca attcctaaca caaatgtgca cttgagccat gatgctaaga gtgccccatt 2940
cttgggtatg gctcctgctg tgcagagaag cgcacctgcc ttaccagtgt ccgaagactt 3000
taaatcggag cctgtttcac ctcaggattt ttctttctcc aagaatggtc tgctaagtcg 3060
attgctaaga caaaatcaag atagttacct ggcagatgat tcagacagga gtcacagaaa 3120
taatgaaatg gcacttctag aatcaaaaga tctttgcatg gtccctaaga aaaggaagct 3180
ttatactgag ccattagaaa atccatttaa aaagatgaaa aacaacattg ttgatgctgc 3240
aaacaatcac agtgccccag aagtactgta tgggtccttg cttaccagg aagagctgaa 3300
atthagcaga aatgatcttg aatttaata tctgtctggt catggctcag ccagcgaaag 3360
tgaacacagg agttgggcca gagagagcaa aagctttaat gttctgaaac agctgcttct 3420
ctcagaaaac tgtgtgagag atttgtcccc gcacagaagt aactctgtgg ctgacagtaa 3480
aaagaaaaga cacaaaaata atgtgacca cagcaaacct gaatttagca tttcttctt 3540
aaatggactg atgtacagtt ccactcagcc cagcagttgc atggataaca ggacattttc 3600
ataccagggt gtagtaaaaa ctctgtgag tcctacttct cctgagcact tgggctgtgc 3660
agggctctaga ccagaatctg ggcttttgaa tgggtgttcc atgccagtg agaaaggacc 3720
cattaagtgg gttatcactg atgcggagaa gaatgagat gaaaaagact ctccaagatt 3780

-continued

gacccaaaacc aacccaatac tatattacat gcttcaaaaa ggaggcaatt ctgttaccag 3840
tcgagaaaaca caagacaagg acatttggag ggaggcttca tctgctgaaa gtgtctcaca 3900
ggtcacagcc aaagaagagt tacttcctac tgcagaaacg aaagcttctt tctttaattt 3960
aagaagccct tacaatagcc atatgggaaa taatgcttct egcccacaca ggc caaatgg 4020
agaagtttat ggacttctgg gaagcgtgct aacgataaag aaagaatcag aataaaatgt 4080
acctgccatc cagttttgga tctttttaa actaatgagt atgaactga gatctgtata 4140
aataagagca tgatttgaaa aaaagcatgg tataattgaa acttttttca ttttgaaaag 4200
tattggttac tgggtgatgtt gaaatagca tactaatttt tgcttaacat tagatgtcat 4260
gaggaaaacta ctgaactagc aattggttgt ttaacacttc tgmtatgcatc agataacaac 4320
tgtgagtagc ctatgaatga aattctttta taaatattag gcataaatta aaatgtaaaa 4380
ctccattcat agtggattaa tgcattttgc tgcctttatt aggggtacttt attttgcttt 4440
tcagaagtca gctacataa cacattttta aagtctaaac tgttaaacaa ctctttaaag 4500
gataattatc caataaaaaa aaacctagtg ctgattcaca gcttattatc caattcaaaa 4560
ataaattaga aaaatatatg cttacatttt tcaacttttgc taaaaagaaa aaaaaaaggt 4620
gtttattttt aactcttggga agaggttttg tggttcccaa tgtgtctgtc ccaccctgat 4680
ccttttcaat atatattttt ttaaaccttg tgcacttag taaaaattga ttacaattga 4740
gggaagtttg atagatcctt taaaaaaaaa gcagatttcc attttttgta ttttaactac 4800
tttactaat taatactcct ccttttacag aattagaaaa gttaacattt atctttaggt 4860
ggtttctga aaagttgaat atttaagaaa ttgtttttaa cagaagcaaa atggcttttc 4920
tttgacagt ttcaccatc tcttgtaaaa gtttaattctc accattcctg tggtaacctgc 4980
gagtgttatg accaggatc cttaaacctg aactcagacc acttgcatta gaaccatctg 5040
gagcacttgt tttaaaatgc agattcatag gcagcatctc agactctacag aacaagaatc 5100
tctgctaagt ggacctggaa tcttccatct gcatcttaac atgctcteta ggtgtttctt 5160
gtgtttgaga accatgactt atgactttcc tcagaacatg agactgtaaa acaaaaacaa 5220
aaaactatgt gatgcctcta ttttcccaaa tacagtcaca catcagctca aaatttgcaa 5280
tattgtagtt catatattac cgttatgtct ttggaaatcg ggttcagaac actttttatg 5340
acaaaaattg ggtggagggg ataactttca tatctggctc aacatctcag gaaaatctgt 5400
gattatttgt gtgttctaag gagtaacatc tacttagtta gccttaggga tggaaaaaca 5460
gggccactta ccaaaactcag gtgattccag gatggtttgg aaacttctcc tgaatgcac 5520
cttaaccttt attaaaacca ttgtcctaag aacaatgcca acaaaactta caacatttag 5580
tttaaaccca agaagggcac taaactcaga ttgactaaat aaaaagtaca aagggcacat 5640
atcgtgaca gaattgtaca caatcactcc attggatcct ttactttaaa gtagtgatga 5700
aaagtacatg ttgatactgt cttagaagaa ataatatata tagtgaagcc acatgggggt 5760
tcagttgcga aacaggtctg tttttatgtt cagtttgtaac aatccacaat tcattcacca 5820
gatattttgt tcttaattgt gaaccagggt agcaaatgac ctatcaaaaa ttattctata 5880
atcactacta gtttagatat tgatttaaaa ttgttctact tgaagtgggt tctaagattt 5940
ttatattaaa aatagggtgt atttcctaat atgatctaaa accctaaatg gttatttttc 6000
ctcagaatga tttgtaata gctactggaa atattatata gtaataggag tgggtattat 6060

-continued

```

gcaacatcat ggagaagtga aggcataaggc ttattctgac ataaaattcc actggccagt 6120
tgaatatatt ctattccatg tccatactat gacaatctta ttgtcaacac tatataaata 6180
agcttttaaa caagtcattt ttcttgatcg ttgtggaagg tttggagcct tagaggtatg 6240
tcagaaaaaa tatggttgga ttctcccttg ggtaggggga aatgacctt ttacaagaga 6300
gtgaaattta ggtcagggaa aagaccaagg gccagcattg ctacttttgt gtgtgtgtgt 6360
gtgggttttg tttgttttt ttggttggtt ggttgtttt gttgttgtta acaaaggaat 6420
gagaatatgt aatacttaaa taaacatgac cacgaagaat gctgttctga tttactagag 6480
aatgttccca atttgaattt aggggtgatt taaagaacag tgagaaaggg catacatcca 6540
cagattcact ttgtttatgc atatgtatg acaaggatgc acatatacac attttcaagg 6600
actattttag atatctagac aatttcttct aataaagtca tttgtgaaag ggtactacag 6660
cttattgaca tcagtaaggt agcattcatt acctgtttat tctctgctgc atcttacaga 6720
agagtaaact ggtgagagta tatattttat atatataat atataatata atataatag 6780
tatatatata tatattgact tgttacatga agatgttaaa atcggttttt aaaggtgatg 6840
taaatagtga tttccttaat gaaaaataca tattttgtat tgttctaag caacagaaaa 6900
gccttttaat ctctttgggt cctgtatatt ccatgtataa gtgtaaatat aatcagacag 6960
gtttaaaagt tgtgcagtga tgtatacagt tgcaagtctg gacaaatgta tagaataaac 7020
cttttattta agttgtgatt acctgctgca tgaaaagtgc atgggggacc ctgtgcatct 7080
gtgcatttgg caaaatgtct taacaaatca gatcagatgt tcatcctaac atgacagtat 7140
tccatttctg gacatgacgt ctgtggttta agctttgtga aagaatgtgc tttgattcga 7200
agggctctaa agaatttttt taactgtcaa ccacttttaa acataaagaa ttcacacaac 7260
tactttcatg aattttttta tcccattgca aacattatc caagagtatc ccagtattag 7320
caactactga atataggcac attaccatc atagtaagaa ttctggtgtt tacacaacca 7380
aatttgatgc gatctgctca gtaataat tggccattt tattagaaat ttaatttctt 7440
catgtgatgt catgaaactg tacatactgc agtgtgaatt tttttgttt gttttttaat 7500
cttttagtgt ttacttctc cagtgaattt gaataaatga gaaaaaatgc attgtc 7556

```

<210> SEQ ID NO 37

<211> LENGTH: 1516

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 37

```

tgccccacca cegctgctcc tcagcaggcg cctcaccagc ctccacacce cttgcccgcg 60
cagaaacgcg cctggccctg agctgtcacc accgacactc tccaggctcc ggacacgatg 120
caggccatca agtgtgtggt ggtgggagat ggggcccgtg gcaagacctg ccttctcatc 180
agctacacca ccaacgcctt tcccggagag tacatcccca ccgtgtttga caactattca 240
gccaatgtga tgggtgacag caagccagtg aacctggggc tgtgggacac tgcctggcag 300
gaggactacg accgtctccg gccctctcc tatccacaga cggacgtctt cctcatctgc 360
ttctccctcg tcagcccagc ctcttatgag aacgtccgcg ccaagtggtt cccagaagtg 420
cggcaccact gcccagcac acccatcacc ctggtgggca ccaagctgga cctgcccggc 480
gacaaggaca ccatcgagaa actgaaggag aagaagtgg ctcccatac ctaccgcgag 540

```

-continued

```

ggcctggcac tggccaagga gattgactcg gtgaaatacc tggagtgctc agctctcacc 600
cagagaggcc tgaaaaccgt gttcgacgag gccatccggg ccgtgctgtg cctcagccc 660
acgcggcagc agaagcgcgc ctgcagcctc ctctaggggt tgcaccccag cgctcccacc 720
tagatgggtc tgatctccca ggatccccac ccaaagcctg atggcacccc ggctggccat 780
gctgtcccct cctgtggcg tttcttagca gatggctgca gagcttcggt gatggtcttt 840
tctgtactgg aggccctctg aggccaggaa cgtgcaaatt tgcaggtgct gcatcccaag 900
cccctcatgc tctgccttc ctgagggcca gaggggagcc ccaggacca ttaagccacc 960
cccgtgttcc tgcgctcagt gccaaactgcc gcatgtgaa gcatctacc gttcactcca 1020
gtcccacccc acgctgact cccctctgga aactgcaggc cagatggttg ctgccacaac 1080
ttgtgtacct tcagggatgg ggtcttact cctcctgag gccagctgct ctaatatcga 1140
tggctctgct tgcagagag ttcctctacc cagcaaaaat gagtgtctca gaagtgtgct 1200
cctctggcct cagttctcct cttttggaac aacataaaac aaatttaatt ttctacgct 1260
ctggggatat ctgctcagcc aatggaaaaat ctgggttcaa ccagcccctg ccatttctta 1320
agactttctg ctgcactcac aggatctga gctgcactta cctgtgagag tcttcaaact 1380
tttaaacctt gccagtcagg acttttgcta ttgcaaatag aaaacccaac tcaacctgct 1440
taagcagaaa ataattttat tgattcaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1500
aaaaaaaaaa aaaaaa 1516

```

<210> SEQ ID NO 38

<211> LENGTH: 1641

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 38

```

ctccctgtgt tggtgaggga tgtctgcagc agcatttaa ttctgggagg gcttggttgt 60
cagcagcagc aggaggaggc agagcacagc atctgctggga ccagactcgt ctccaggccag 120
tgcaagcctt ctccagcaaa cgcgcaccaa ggaaaactca ctaccatgag aattgcagtg 180
atttgccttt gcctctagg catcacctgt gccataccag ttaaaccaggc tgattctgga 240
agttctgagg aaaagcagct ttacaacaaa taccagatg ctgtggccac atggctaaac 300
cctgacccat ctccagaagca gaattctcta gcccacaga atgctgtgtc ctctgaagaa 360
accaatgact ttaaacaaga gacccttcca agtaagtcca acgaaagcca tgaccacatg 420
gatgatatgg atgatgaaga tgatgatgac catgtggaca gccaggactc cattgactcg 480
aacgactctg atgatgtaga tgacactgat gattctcacc agtctgatga gtctcaccat 540
tctgatgaat ctgatgaact ggtcactgat ttcccacgg acctgccagc aaccgaagtt 600
ttcactccag ttgtcccac agtagacaca tatgatggcc gaggtgatag tgtggtttat 660
ggactgaggt caaaatctaa gaagtctcgc agacctgaca tccagtaccc tgatgctaca 720
gacgaggaca tcacctcaca catggaaagc gaggagtga atggtgcata caaggccatc 780
cccgttgccc aggacctgaa cgcgccttct gattgggaca gccgtgggaa ggacagttat 840
gaaaacgagtc agctggatga ccagagtgct gaaacccaca gccacaagca gtccagatta 900
tataagcgga aagccaatga tgagagcaat gagcattccg atgtgattga tagtcaggaa 960
ctttccaaag tcagccgtga attccacagc catgaatttc acagccatga agatatgctg 1020

```

-continued

```

gttgtagacc ccaaaagtaa ggaagaagat aaacacctga aatttcgtat ttctcatgaa 1080
ttagatagtg catcttctga ggtcaattaa aaggagaaaa aatacaattt ctcaactttgc 1140
athtagtcaa aagaaaaaat gctttatagc aaaatgaaag agaacatgaa atgcttcttt 1200
ctcagtttat tggttgaatg tgtatctatt tgagctgga aataactaat gtgtttgata 1260
attagtttag tttgtggctt catggaaact ccctgtaaac taaaagcttc agggttatgt 1320
ctatgttcat tctatagaag aaatgcaaac tatcaactgta ttttaattatt tgttattctc 1380
tcatgaatag aaatztatgt agaagcaaac aaaatacttt taccactta aaaagagaat 1440
ataacathtt atgtcactat aatcttttgt tttttaagtt agtgtatatt ttgttgtgat 1500
tatctttttg tgggtggaat aaatctttta tcttgaatgt aataagaatt tgggtgtgtc 1560
aattgcttat ttgttttccc acggttgtcc agcaattaat aaaacataac cttttttact 1620
gcctaaaaaa aaaaaaaaaa a 1641

```

```

<210> SEQ ID NO 39
<211> LENGTH: 6463
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

```

```

<400> SEQUENCE: 39

```

```

ctctcttctc cgctcgctcc ctctctctcc tgetggctgc ctgttctagg aagccagcgc 60
ggagaggggg gggatgcaca gcacagggga gagagattgc gcatgttggc cagtcgtgtt 120
ttaaagagta cagtgcgggg aggtgagag gggcgcatgc aacaacaact tttggaagga 180
tgaagagaa gagcgaaaa tactccatca gcagtgacaa ctctgacacc actgacagtc 240
atgcgacatc tacatccgca tcaagatgct ccaaactgcc cagcagcacc aagtcgggct 300
ggccccgaca gaacgaaaag aagcctccg aggttttccg gacagacttg atcacagcca 360
tgaagatccc ggactcatac cagctcagcc cggatgacta ctacatcctg gcagaccat 420
ggcgacagga atgggagaaa ggtgtgcagg tgcctgccgg ggcagaggcc atcccagagc 480
ccgtggtgag gatcctccca ccaactggaag gcccccctgc ccaggcatcc cagagcagca 540
ccatgcttgg tggggctcc cagcctgatt ggccaggggg cagccgctat gacttggagc 600
agattgatgc ctactggctg gagctcatca actcggagct taaggagatg gagaggccgg 660
agctggaaga gctgacatta gagcgtgtgc tggaggagct ggagaccctg tgcaccaga 720
atatggccag ggcattgag acgcaggagg ggctgggcat cgagtacgac gaggatggtg 780
totgcagct gtgtcgtct cctgagggcg aggatggcaa cgagatggtc ttctgtgaca 840
agtgcaacgt ctgtgtgcat caggcatgct acgggaccc caagtgccc acgggcagct 900
ggctgtgcc gacgtgtgcc ctgggtgtcc agccaaagtg cctgctctgc cccaagcgag 960
gaggagcctt gaagcccact agaagtggga ccaagtgggt geatgtcagc tgtgccctat 1020
ggattcctga ggtcagcatc ggctgcccag agaagatgga gcccatcacc aagatctcgc 1080
atatcccage cagccgctgg gctctgtcct gcagcctctg caaggaatgc acaggcacct 1140
gcacccagtg ttccatgcct tctctcgtca cagcgttcca tgtcacatgc gcctttgacc 1200
acggcctgga aatgcggact atattagcag acaacgatga ggtcaagtcc aagtcattct 1260
gccaggagca cagtgcaggg ggccccagta atgagccac atctgagccc acggaacca 1320
gccaggctgg cgaggacctg gaaaaggtga ccctgcgcaa gcagcggctg cagcagctag 1380

```

-continued

aggaggactt ctacgagctg gtggagccgg ctgaggtggc tgagcggctg gacctggctg	1440
aggcactggg cgacttcacg taccagtact ggaagctgaa gaggaaagcc aatgccaacc	1500
agccgctgct gacccccaaag accgacgagg tggacaacct ggcccagcag gagcaggacg	1560
tcctctaccg ccgectgaag ctcttcaccg atctgcggca ggacctagag aggggttagaa	1620
atctgtgcta catggtgaca aggcgcgaga gaacgaaaca cgccatctgc aaactccagg	1680
agcagatatt ccacctgcag atgaaactta ttgaacagga tctgtgtcga ggectgtcca	1740
cctcattccc catcgatggc accttctca acagctggct ggccacagtc gtgcagatca	1800
cagcagagaa catggccatg agcgagtggc cactgaacaa tgggcaccgc gaggaccctg	1860
ctccagggct gctgtcagag gaaactgctg aggacgagga gacctgctc agcttcatgc	1920
gggacccctc gctgcgacct ggtgacctg ctaggaagge ccgaggccgc acccgctgc	1980
ctgccaagaa gaaaccacca ccaccaccac cgcaggacgg gctgggtca cggacgactc	2040
cagacaaagc ccccaagaag acctggggcc aggatgcagg cagtggcaag gggggtcaag	2100
ggccacctac caggaagcca ccacgtcggc catcttctca cttgectcc agccctgcag	2160
ccggggactg tcccactca gccaccctg aaagccccc gccactggcc cctgagacc	2220
cggacgaggc agcctcagta gctgctgact cagatgtcca agtgctggc cctgcagcaa	2280
gccctaagcc tttgggcccg ctccggccac cccgcgagag caaggtaacc cggagattgc	2340
cgggtgccag gctgatgct gggatgggac cacctcagc tgtggctgag aggcccaagg	2400
tcagcctgca ttttgacct gagactgat gctacttctc tgatggggag atgagcgaact	2460
cagatgtaga ggcaggaggc ggtgggggtg agcggggctc ccgggaggca ggggcagagg	2520
aggtggtecg catgggcgta ctggectct aactcaccgc cttcctgct ccaggccctg	2580
ccctggtecc ccacaaggc ctcagccag tcacaactgc catttccagt ctctgctgag	2640
tgtcccagac cctcaggct gccactccgt cgtggttcta ttttaatat agagagagtt	2700
ttgaattcta cactgttgc tttcctctgt gctggcctag gacattagga ttccttccac	2760
ggctccggcc gctaggacc tgcaggtcc cgcgcaccat ccctgcccctg cccacgtggt	2820
attgctgggc tctggctag atgcaagcaa ggtggacaag agctcaggac tccagccac	2880
tgccactggg tgacacagac tgcgtttgg gcattatttc atggcagatg gggcagtcca	2940
gggcctaccc cgecttgcgc ccagatccca ctggggtcca tttgggggt cctgctacac	3000
tccaccgac cccaaggaag tataataaac gataccagc cagagtctac tcaactgtcac	3060
aagcacaacg agtttatatg agaaagcact gaggggggtg agagggccg ctagttccag	3120
gggaactgaa agctgttct gatcagcccg tatcatctga ggctgctg cccacctgc	3180
cacctcccc tcccttctg ctctgcccct gccagtgccc agcccagcgg ctctgggaag	3240
gggttcccag aatccctct gagctgtgcc atttactcag gggactccca aacagccagc	3300
tgccagtgca ggtggagggc tgtaggggag ggccagtgcc cagacagggt catggggctc	3360
agaccagccc actgtagaga atcactctga ggtcccaact tccttcttc cttcggggcc	3420
agtctcggcc gaagtctggt cacgctcaga cagagctgac cagaccagac cgtttgcctt	3480
ttcaagtttc ctagtctgc tacaagatga gcttcttccg tggtttctt ttggaaactc	3540
ctccttccaa caagcagtgg gatcccggg cccagggcgg gccggtgttg gccgctgggg	3600
ctgttgtaag tcttctgga tgttcccctg ttcctgagcc ttaaccctc gcacagccat	3660

-continued

```

ccccccccc gtcctgccat cccccccgc cgtcctgcct tccccacccc acccttaggt 3720
cccaggtagt tgctctgaag agtttcagta gagtgccccc aggggtgatag ctcagggaac 3780
aacaaaaaag gaattccgtg aaaacatttt tttttctttg atgaattact cctgggtcac 3840
ttccaccact ggtaaagcca gaacttctcc aaaaagaacc ttgcaaaaag tccagtgaat 3900
cagtcgaate attctgtgga tgccaaagaa tattttgacc ataatacagc acagcctgga 3960
cctgacaact tgtcatttgg actttttttt aaatggagtt ctttagcaac aaagtataga 4020
aacatgttca ttgcacacac ccaaggagaa gagctcaagc gcttggaga ggatgctttg 4080
ctgctgctga agtgtaacctg ggtgttagat ttcagatcct gggctgagcc cactgtgagc 4140
tttctaaac tgtgagactc acagagggga aagatactga cggtgaaacc agcatggaaa 4200
acgtctttac catgtggttc cctcctcccc aaatacataa agcaataaag caggatgggg 4260
aacagcttga ccttcatcca cccctaactc caaaactatc aaggtacgac agtggcattg 4320
tcatcgacac tcaatttcat gtgaatttta gcaaaacagg aaacaaagat aatgactcag 4380
ttcagaggat cggacaaaatg tgtctagtc gggtggaact ggaggagtg gggtgggctt 4440
caaggattct gggcgttggg atggcatgag ctacctgta gagtttagtc tgectgcccg 4500
ccttggtagt agtgaccagt cagtgtcagc atcagtgtcc caaccccagt ctctgtttac 4560
tgccctttaa cagaacttct tccctcccca tgcttgggtt cacctcgggc tgcaaccctg 4620
tctgtgccag attgcccggt ctgacctgc aggaagcaaa gaggtgagct taaagaacaa 4680
ccaaactctg ccaggggtcc cagaaagccc agggccagc agtctcagca cttggcccct 4740
tgccccttca caccatcctg gggcaggggc tgggcctccc tgggtggcagg ggtgggtgga 4800
gaattagga gagggtgcaa cgagtctggc cccttgctc gggctggctg gtgttcttcc 4860
aagagcctct gctcacattg ttggcctctg gattctggcc cttcttcaat ggetgttget 4920
ttggactgga ctgttgctga gcctgtgtcc tgcagaaccc agatgtctgt taggctggct 4980
ggctgctgcg aggggagggg ggtggccttt catttgggtt gccctttcac tcccaggcca 5040
agccctggag caatcttctt caggcagctg tctccacctc caggatgtcc agcaggetgc 5100
aaggagaagg atgccagcca cccatcctcc cccagttccc agcctttccc ctgttggtea 5160
cagccgcttc tgtcttttcc cggctcactg tcccagtgat agagggtctt gctgtccctg 5220
agactgaggc aggttctctt tccaggtcag aggtggaggt agatctttct ctaaccaca 5280
tctgctcca cacacagctc ctccgcaggg aaggagaagc tgctctgtaa ctcattctgg 5340
ctatctccc ccttctcact gacctgaccg cccaccacct cctccccct catcacatga 5400
caaaggataa tgtgcaagaa aagtattttt atgtatcata aatgtatttt gaaacaaatg 5460
agaagaagaa aggtagaagg gtttatttta ttaaagagc ctgacttagt gacagtgtgt 5520
gagcatttgc aatgtaaggg cctcagcttc cttggagaag ccaccccagg tttccagaca 5580
tagatgttga attgttttgg ggggtgtgtc caggccacgt ctcgtgtgtc cgtatgcagg 5640
catgctgtg tatactgtgt atgggcacac tgggactagc tgggacaatt cctagagatt 5700
caactgccc attctaacca acattggcag cggctgaact tggcatttcc ttgctaactg 5760
ccagatgtgg ccaacctttg tccatagca aacctgaa aaatgatctg gatttctata 5820
gcaaggccct tggggagggc actctcccat gcccttggcc tcgctggcca cattggccaa 5880
tgagccaggg ctggagtctg agaccttgg ttgttcttta aggcacctcc tgccacttc 5940

```


-continued

```

tccctcagag gcacaaacac tttgtgttcc acgtcagttt gaggggacgg tggggggatg 6000
atatgaatgt cacaggagga gacaccttct gtctttgttt caaagaaagt gatgtgccat 6060
ttgttaatat acaagagaaa tattgaaaat atattgaaaa gagcaatfff aaattatfff 6120
tggettatgt tgcaatatff atffttctgt attagaaaag atffctffgt agagaaaaaa 6180
tgtatffttc attaacgcaa agacctatff ctctfftttg tacattgtcc atgtgcgcaa 6240
cccttaacga gcaatagaat gtatggtcac ctgggtgtgg ccagtgcctg ctgtgcctcg 6300
catgattctg tgttgcctg gctgcatagt tcccagcccc atctgtctct gctcaactcat 6360
gggggcttcc agccccggc cccaccaggg cttgtgtcat agggagcctt ttgcaactct 6420
cgtgtgttgg caaacgcagt taataaaagca gtgtffttctg tgc 6463

```

<210> SEQ ID NO 40

<211> LENGTH: 2828

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 40

```

catagatgaa aatggcaagt tccctggctt tccttctgct caactttcat gtctccctcc 60
tcttgggtcca gctgctcact ccttgcctag ctcaagtttct tgtgcttggc ccctctgggc 120
ccatcctggc catggtgggt gaagacgctg atctgcctg tcacctgttc ccgaccatga 180
gtgcagagac catggagctg aagtgggtaa gttccagcct aaggcaggtg gtgaacgtgt 240
atgcagatgg aaaggaagtg gaagacaggc agagtgcacc gtatcgaggg agaacttcca 300
ttctgcggga tggcatcact gcagggaagg ctgctctccg aatacacaac gtcacagcct 360
ctgacagtgg aaagtacttg tgttatttcc aagatggtga cttctatgaa aaagccctgg 420
tggagctgaa ggttgcagca ctgggttcta atcttcaact cgaagtgaag ggttatgagg 480
atggagggat ccatctggag tgcaggtcca ccggctggta cccccaacc caaatacagt 540
ggagcaacgc caagggagag aacatcccag ctgtggaagc acctgtggtt gcagatggag 600
tgggcctata tgaagtagca gcatctgtga tcatgagagg cggctccggg gaggggtgat 660
cctgcatcat cagaaattcc ctctcggcc tggaaaagac agccagcatt tccatcgag 720
acccttctt caggagcgc cagccctgga tgcagccct ggcagggacc ctgcctatct 780
tgetgctgct tctcgcggga gccagttact tcttgtggag acaacagaag gaaataactg 840
ctctgtccag tgagatagaa agtgagcaag agatgaaaga aatgggatat gctgcaacag 900
agcgggaaat aagcctaaga gagagcctcc aggaggaact caagaggaaa aaaatccagt 960
acttgactcg tggagaggag tcttctgctg ataccaataa gtcagcctga tgetctaactg 1020
gaaaaatggc cctcttcaag cctggaaaaa tggctgacct catggacacc tctcaaaact 1080
ctctgcagca gatgtaatct tgatccaga catggcaaat gccatcctcc ttgtttctga 1140
ggaccagagg agtgtacagc gtgctgagga gccccatgac ctaccagaca accctgagag 1200
atftgaatgg cgttactgtg tgettgctg tgaagcttc atgtcagaga gacactactg 1260
ggaggtggaa gtgggggaca gaaaagagt gcatattggg gtatgtagta agaactgga 1320
gaggaaaaaa gtttgggtca aaatgacacc ggagaacgga tactggacta tgggcctgac 1380
tgatgggaat aagtatcggg ctctcaactg gcccagaacc aacctgaaac tctctgagcc 1440
tctaggaaa gtgggggtca tcttgacta tgagactgga catatctctg tctacaatgc 1500

```

-continued

```

cacggatgga tctcatatct acacatttct gcacgcctct tcctctgagc ctctgtatcc 1560
tgtattcaga attttgacct tggagcccac tgcctgacc gtttgcccaa taccaaaagt 1620
agagagtcc cccgatcccg acctagtgcc tgatcattcc ctggagatac cactgacccc 1680
aggetttagct aatgaaagtg gggagcctca ggctgaagta acatctctgc ttctcctgct 1740
ccagcctgga gctaagggtc tcaccctcca caacagccag tcagaacccat aaagctacag 1800
gcacacactg aagcacttta ctgatattca ttcaattatt ccataggaca gttggttgag 1860
tttggtgcca ccttattggc ccccttatac agataaggaa actgggggtgt agaaaagtgt 1920
attgacttta caaagcagac aggaatagtg aacaacagag ctgggatctg aacaacaatg 1980
actaacatta atggagaatt taaaacgttc tgagtgtgt gttatgagct ttgggtgggtg 2040
tcactccttt aatcctcaca acacctgtc aggtagtctc atttggcaag tatggaagca 2100
gaggcagggc aacattaagt agcttacata actcacacgg taatttgtgc agttgggaga 2160
tgttcagctt cagtccctgg ccaattgccc gttcttttcc agcctgattt ttctctgatg 2220
ggaagagccc acatgtagcc ctgaggttcc ctcccagga cagctccagg atcgagatca 2280
ctgtgagtgg ttgtggagt aagacccta tggactcctt cccagctgat tatcagagcc 2340
ttagaccag cactccttgg attggctctg cagagtgtct tgggtgagag aataacggtg 2400
cagttccac agggcatgtg actttgaaag agactagagg ccacactcag ttaataatgg 2460
ggcacagatg tgtcccacc caacaaatgt gataagtgat cgtgcagcca gagccagcct 2520
tccttcagtc aaggtttcca ggcagagcaa ataccctaga gattctctgt aatattggtg 2580
atttggatga aggaagctag aagaattaca gggatgttt taatcccact atggactcag 2640
tctcctggaa aaggatctgt ccactcctgg tcattggtgg atgttaaacc catatcctt 2700
tcaactgtg cctgctaggg aaaaactgct ctcattatca tcactattat tgetcaccac 2760
tgtatccct ctactgggca agtgcttgc aagttctagt tgttcaataa atttgttaat 2820
aatgctga 2828

```

```

<210> SEQ ID NO 41
<211> LENGTH: 2698
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

```

```

<400> SEQUENCE: 41

```

```

gccgtccgtg ctgactgagg cgctgcagcc aggagccgcg gccggctgcc cagcgtctgc 60
cgcctccgcg cgtccgcagc cgtcccgcg ccgacatgcg cttggccgcc gccgcgaacg 120
aggegtacac ggcctcttgg gcggtctcgg ggctgctggg ctgcaagcag tgcggggggg 180
gccgcgacca ggacgaggaa cttggcatta gaattcctcg accactagga cagggaccaa 240
gcagattcat ccagaaaag gagatcctcc aagtggggag tgaagacgca cagatgcatg 300
ctttatttgc agattctttt cctgctttgg gccgtttgga taacattacg ttagtgatgg 360
ttttccacc acaatattta gaaagtttct taaaaactca gcactatcta ctgcaaatgg 420
atgggcccgtt acccctacat tatcgtcact acattggaat aatggctgcg gcaagacatc 480
agtgtccta cttagtgaac ctgcatgtaa atgattcct tcatgttggg ggggacccca 540
agtggctcaa tggtttagag aatgctctc aaaaactaca gaatttagga gaacttaaca 600
aagtgttagc ccatagacct tggcttatta ccaagaaca cattgagggg cttttaaag 660

```

-continued

ctgaagagca cagctggtcc cttgcggaat tggtagatgc agtagtttta ctcacacact 720
atcattctct tgcctcattc acattcggct gtggaatcag tccagaaatt cattgtgatg 780
gtggccacac attcagacct cttctgtta gcaactactg catctgtgac attacaaatg 840
gcaatcacag tgtggatgag atgccggtea actcagcaga aaatgtttct gtaagtgatt 900
ctttctttga ggttgaagcc ctcatggaaa agatgaggca gttacaggaa tgtcgagatg 960
aagaagaggc aagtcaggaa gagatggctt cacgttttga aatagaaaaa agagagagta 1020
tgtttgcctt ctcttcagat gatgaagaag ttacaccagc aagagctgta tctcgtcatt 1080
ttgaggatac tagttatggo tataaagatt tctctagaca tgggatgcat gttccaacat 1140
ttcgtgtcca ggactattgc tgggaagatc atggttatc tttggtaaat cgcctttatc 1200
cagatgtggg acagttgatt gatgaaaaat ttcacattgc ttacaatctt acttataata 1260
caatggcaat gcacaaagat gttgatacct caatgcttag acgggcaatt tggaactata 1320
ttcactgcat gtttgaata agatatgatg attatgacta tggtgaaatt aaccagctat 1380
tggatcgtag ctttaaagtt tatacaaaa ctggtgtttg cactcctgaa aaggttacca 1440
aaagaatgta tgaatagctt tggaggcagt tcaagcactc tgagaaggtt catgttaatc 1500
tgcttcttat agaagctagg atgcaagcag aactccttta tgctctgaga gccattacc 1560
gctatatgac ctgatgcctt tcttcatta aagatgatc tggaatgate agcagatata 1620
gtctacaagg ggaaggttac taagcccag gaccaatggt agacaaaata attcagaaat 1680
ccattgtgcc atgattcctt tagtttctgc tatttttctg tggaaaacca ctgctggcac 1740
aagcagtgc tgtttggcag cttcaagttt agagctgta agacaggctg ccattcacag 1800
tattttgctt tttgacagta caagatgctg tgtaactggt ttaatacagc aaatagtaac 1860
tctccaaatc ctgttgctt tatgttaaat aagataacaa gaattggagc atgcaaagaa 1920
tgggacttgg ataatgactt aagcttata tghtaagaat tttagaagat cttggtgctg 1980
ctattcctgc tggaggaatg aatagatggc tgtttcagtt aagctattag taataaaagt 2040
gaacattgct actatctgag cctacatata taacttgtgt gatttcaat taaactgca 2100
ttatgtgtta attttcttc atctaaaaaa gcatagaatt cctactcaca cagctcagca 2160
acaaccatct tgatgtaac agttaatttc tttcattagt tttttaaatt cagggttctg 2220
gatattaaat taaaatggca ttcttaaaga ttttcttcaa aaagcaatcc taaatgaaag 2280
tgtgtaaatt ataagaagct ggcatctt ttgatgctg tttccagga tctgacact 2340
ggagggcagc tgtcttctgc attacttctg tttccagcac caaagttgtg ggacatggtg 2400
ctgtagactg ctgcgcagtc ctgggtgcat tcagctctct tgcctctgce tgcctctgg 2460
tccccacttt aaaggctgtg cagctcctta aataataaag ctggaaaata ttttagtctg 2520
ggttatcaaa tttgatttac aaaaacgcta actttgtttg aaatgcaaac aggtttgaaa 2580
atatgtatta agtactttgt attctggaag cgtgaattgc ttttgaagtc tgtcagtatt 2640
actggtatct ttaataaag aagaatcttt ctccaatctt aaaaaaaaaa aaaaaaaaaa 2698

<210> SEQ ID NO 42

<211> LENGTH: 5215

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 42

-continued

cgagcgcggc gcccttgagc tgcaccgcgg cgcaggtttg cgagccgact tgtcagccgg 60
ccaagaaaag gaagctccgt cccttccgc tcaccggct tccccacccc ttgtactcta 120
aactctgcag agggcgagcg gcgcggccac ggaggcgccg agggagagcg agccgcgcgc 180
gggcagcggc gtgccctcgg gggagagggc gccggagagg agggcgccgc gggcgggcga 240
gggcgcggcg cgcgatggca gctgcttagc ccggcgggcg cggagcagcc ccgagctgtg 300
gctggccagg cggtgccgct gggcggggga cgcgcgcgc gttgctgccc ggcccggaga 360
gatgagcacg gaggcgagcg agggcatcac tttctctgtg ccacccttcg ccccctcggg 420
cttctgcacc atccccgagg gcggcatctg caggagggga ggagcggcgg cgttgggcca 480
gggcgagggg caccagctgc caccgcgcgc gccgggcagc ttctggaacg tggagagcgc 540
cgctgccctt ggcacgtggt gtccggcgcc cacctcctcg agcagtcca cccgagccg 600
gggcagctct gttggcgggg gcagccgacg gaccacgggt gcatatgtga tcaacgaagc 660
gagccaaggg caactggtgg tggccgagag cgaggccctg cagagcttgc gggagcgctg 720
cgagacagtg ggcgccaccc tggaaacctt gcattttggg aaactcgact ttggagaaac 780
caccgtgctg gaccgctttt acaatgcaga tattgctggtg gtggagatga gcgatgcctt 840
ccggcagccg tccttgtttt accaccttgg ggtgagagaa agtttcagca tggccaacaa 900
cateatctc tactgtgata ctaactcgga ctctctgcag tcactgaagg aaataatttg 960
ccagaagaat actatgtgca ctgggaacta cacctttgtt ccttacatga taactccaca 1020
taacaaagtc tactgctgtg acagcagctt catgaagggg ttgacagagc tcatgcaacc 1080
gaacttcgag ctgcttcttg gaccatctg cttacctctt gtggatcgtt ttattcaact 1140
tttgaaggtg gcacaagcaa gttctagcca gtacttcgg gaatctatac tcaatgacat 1200
caggaaagct cgtaatttat acactggtaa agaattggca gctgagttgg caagaattcg 1260
gcagcgagta gataatatcg aagtcttgac agcagatatt gtcataaatc tgttactttc 1320
ctacagagat atccaggact atgattctat tgtgaagctg gtagagactt tagaaaaact 1380
gccaaccttt gatttggcct cccatcacca tgtgaagttt cattatgcat ttgcaactgaa 1440
taggagaaat ctccctggtg acagagcaaa agctcttgat attatgattc ccatggtgca 1500
aagcgaagga caagttgctt cagatatgta ttgcctagtt ggtcgaatct acaaagatat 1560
gtttttggac tctaatttca cggacactga aagcagagac catggagctt cttggttcaa 1620
aaaggcattt gaatctgagc caaacactaca gtcaggaatt aattatgagg tcctcctcct 1680
ggcagctgga caccagtttg aatctctcct tgagctccgg aaagtgggg tgaagctaag 1740
tagtcttctt ggtaaaaagg gaaacttgga aaaaactcag agctactggg aagttggatt 1800
ttttctgggg gccagcgtcc tagccaatga ccacatgaga gtcattcaag catctgaaaa 1860
gctttttaa ctgaagacac cagcatggta cctcaagtct attgtagaga caattttaat 1920
atataagcat tttgtgaaa tgaccacaga acagcctgtg gccaaagcaag aacttgggga 1980
cttttgatg gatttctcgg tcgagggccac aaagacagat gttactgtgg ttaggtttcc 2040
agtattaata ttagaaccaa ccaaaatcta tcaaccttct tatttgccta tcaacaatga 2100
agttgaggaa aagacaatct ctatttggca cgtgcttcc gatgacaaga aaggtataca 2160
tgagtggaa tttagtgcct cttctgtcag gggagtgagt atttctaaat ttgaagaaag 2220
atgctgcttt ctttatgtgc ttcacaatc tgatgattc caaatctatt tctgtacaga 2280

-continued

acttcattgt	aaaaagtttt	ttgagatggt	gaacaccatt	accgaagaga	aggggagaag	2340
cacagaggaa	ggagactgtg	aaagtgactt	gctggagtat	gactatgaat	atgatgaaaa	2400
tggtagacaga	gtcgttttag	gaaaaggcac	ttatgggata	gtctacgcag	gtcgggactt	2460
gagcaaccaa	gtcagaattg	ctattaagga	aatcccagag	agagacagca	gatactctca	2520
gcccctgcat	gaagaaatag	cattgcataa	acacctgaag	cacaaaaata	ttgtccagta	2580
tctgggctct	ttcagtgaga	atggtttcat	taaaatcttc	atggagcagg	tccttgagg	2640
aagtctttct	gctctccttc	gttccaaatg	gggtccatta	aaagacaatg	agcaaacaaat	2700
tggcttttat	acaaagcaaa	tactggaagg	attaaaatat	ctccatgaca	atcagatagt	2760
tcaccgggac	ataaagggtg	acaatgtgtt	gattaatacc	tacagtgtg	ttctcaagat	2820
ctctgacttc	ggaacatcaa	agaggcttgc	tggcataaac	cctgtactg	aaacttttac	2880
tggtagccctc	cagtatatgg	caccagaaat	aatagataaa	ggaccaagag	gctacggaaa	2940
agcagcagac	atctggtctc	tgggctgtac	aatcattgaa	atggccacag	gaaaaccccc	3000
attttatgaa	ctgggagaac	cacaagcagc	tatgttcaag	gtgggaatgt	ttaaagtcca	3060
ccttgagatc	ccagagtcca	tgtctgcaga	ggccaaggca	ttcactactga	aatgttttga	3120
accagatcct	gacaagagag	cctgtgctaa	cgacttgctt	gttgatgagt	ttttaaagt	3180
ttcaagcaaa	aagaaaaaga	cacaacctaa	gctttcagct	ctttcagctg	gatcaaatga	3240
atatctcagg	agtatatcct	tgccggtagc	tgtgctggtg	gaggacacca	gcagcagcag	3300
tgagtacggc	tcagtttcac	ccgacacgga	gttgaaagt	gacccttct	ctttcaaac	3360
aagagccaag	tctgctggag	aaagagatgt	caaggaatt	cggacactct	tttgggcat	3420
tcagatgag	aattttgaag	atcacagtgc	tcctccttc	cctgaagaaa	aagattctgg	3480
attcttcatg	ctgaggaagg	acagtgcag	gcgagctacc	cttcacagga	tcctgacgga	3540
agaccaagac	aaaattgtga	gaaacctaat	ggaatcttta	gctcaggggg	ctgaagaacc	3600
gaaactaaaa	tgggaacaca	tcacaacct	cattgcaagc	ctcagagaat	ttgtgagatc	3660
cactgaccga	aaaatcatag	ccaccacact	gtcaaaagctg	aaactggagc	tggacttcga	3720
cagccatggc	attagccaag	tcaggtggt	actctttggt	tttcaagatg	ctgtcaataa	3780
agttcttcgg	aatcataaca	tcaagccgca	ctggatggtt	gccttagaca	gtatcattcg	3840
gaaggcggta	cagacagcca	ttaccatcct	ggttccagaa	ctaaggccac	atttcagcct	3900
tgcactctgag	agtgatctg	ctgatcaaga	agacttggat	gtagaagatg	accatgagga	3960
acagccttca	aatcaaactg	tcggaagacc	tcaggctgtc	attgaagatg	ctgtggctac	4020
ctcagcgtg	agcaegctca	gttctactgt	gtctcatgat	tcccagagtg	ctcaccggtc	4080
actgaatgta	cagcttggaa	ggatgaaaat	agaaaccaat	agattactgg	aagaattggt	4140
tggaaaagag	aaagaattac	aagcactcct	tcctcagact	attgaagaaa	aagaccaaga	4200
aattaaacac	ctgaagctta	agtcaccaacc	catagaaatt	cctgaattgc	ctgtatttca	4260
tctaaattct	tctggcacia	atactgaaga	ttctgaaact	accgactggc	tgagagtga	4320
tggagctgat	gaagacacta	taagccggtt	ttggctgaa	gattatacac	tattggatgt	4380
tctctactat	gttacacgtg	atgacttaaa	atgcttgaga	ctaaggggag	ggatgctgtg	4440
cacactgtgg	aaggctatca	ttgactttcg	aaacaaacag	acttgactgt	tgtcaatct	4500
aatcttcgat	ggaatttcta	aaaattaata	cagagctgat	cttctgggg	gtgggaaaat	4560

-continued

```

cgaagggaga ggagaaagge gctgcacttt aaatccagta tttgtttact catgttaaaa 4620
aaaaaaaaa cagacaaaac aactgaaat ttctaaacta catctatttc tataatTTTT 4680
aaggactctt cataaggact cttaaaataa tcttgaacat tagaaccta atgttcagga 4740
agattttaat ctaagcattt ttatggaaat atttttaatg cagcagctat tgcacttcag 4800
ccaaatgttt atttcacaca aaacggatgt aacatttcat gtgatcgtgc accactggaa 4860
caaaacaaa atgtgaccat aactgtttag gcttctgtgt gtttgaata tgctctaata 4920
atctgagtag aaatgcgtaa tttcaattac tgtataaagt ttatgttttt ttaagtgtgc 4980
agaatctgag agcaatgggt tttacttctc tgtgttaatt gtaatattga ctctattttg 5040
taacttaagt ttctgacctg tcgtacattt gtttgagtcg tttatgtact actgaactgt 5100
accagttgca catgctttaa ctgtagtaat gttagcttgt tctaagcta tccattgtgt 5160
catatttact ctaaaaatta aagagactct caacaaaaaa aaaaaaaaaa aaaaa 5215

```

<210> SEQ ID NO 43

<211> LENGTH: 4655

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 43

```

tctgtgcacc ttgcggtggg cggcgaacgg cagcccgccg agcagctagg gggcttgtgc 60
acacagcgag ggagacttag ggactggcag acggacggac ggacggcgag gacctacc 120
gagcccccga gccatggccg agagaaagca atccgggaag gcggcagagg acgaagaggt 180
ccctgctttt ttaaaaaacc tgggctccgg cagccccaag ccccggcaga aattctgtgg 240
catgttctgc ccggtggaag ggtcctcgga gaacaagacc atcgacttcg actcgctgtc 300
ggtagggccg ggtcggggg aggtgggtgc tcagcagcgg gacgtcgccc acttggggccc 360
ggaccccgag ccgccgtact cgcggcaggg ccggcgccgc ggccgagagc catctgttga 420
atcgggcccg aaggtggaga tccggagggc ctcgggcaaa gaagccctgc agaacatcaa 480
cgaccagagc gatcgtcttc tgatcaaagg aggtaaaatt gttaatgatg accagctcgt 540
ctatgcagac atatacatgg aagatgggtt gatcaagcaa ataggagaaa atctgattgt 600
gccaggagga gtgaagacca tcgaggccca ctcccgatg gtgatccccg gaggaattga 660
cgtccacact cgtttccaga tgccctgatc gggaatgacg tctgctgatg atttcttcca 720
aggaaccaag gcggccctgg ctgggggaa cactatgac attgaccagc ttgttctctga 780
gcctgggaca agcctgctcg ctgcctttga ccagtggagg gaatgggccc acagcaagtc 840
ctgctgtgac tactctctgc atgtggacat cagcagtggt cataagggca tccaggagga 900
gatggaagcg cttgtgaagg atcacgggtt aaattccttc ctgctgtaca tggcttcaa 960
agatcgcttc cagctaacgg attgccagat ttatgaagta ctgagtgtga tccgggatat 1020
tggcgccata gcccaagtcc acgcagaaaa tggcgacatc attgcagagg agcagcagag 1080
gatcctggat ctgggcatca cgggccccga gggacatgtg ctgagccgac ctgaggaggt 1140
cgaggccgaa gccgtgaatc gtgccatcac catcgccaac cagaccaact gcccgctgta 1200
tataccaag gtgatgagca aaagctctgc tgaggtcac gccagggcac ggaagaaggg 1260
aactgtggtg tatggcgagc ccatcactgc cagcttggga acggacggct cccattactg 1320
gagcaagaac tgggccaagg ctgtgcctt tgtcacctcc ccacccttga gccctgatcc 1380

```

-continued

aaccactcca gactttctca actccttgct gtctgtgga gacctccagg tcacgggcag 1440
tgcccattgc acgtttaaca ctgcccagaa ggctgtagga aaggacaact tcaccctgat 1500
tccggagggc accaatggca ctgaggagcg gatgtccgtc atctgggaca aggetgtggt 1560
cactgggaag atggatgaga accgatttgt ggctgtgacc agcaccaatg cagccaaagt 1620
cttcaacctt taccocgga aaggccgat tgctgtgga tcogatgccg acctggtcat 1680
ctgggacccc gacagcgta aaaccatctc tgccaagaca cacaacagct ctctcgagta 1740
caacatcttt gaaggcatgg agtgcgcgg ctcccactg gtggtcatca gccaggggaa 1800
gattgtcctg gaggacggca cctgcctgt caccgaagge tctggacgct acattccccg 1860
gaagcccttc cctgattttg tttacaagcg tatcaaggca aggagcagge tggctgagct 1920
gagaggggtt cctcgtggcc tgtatgacgg acctgtgtgt gaagtgtctg tgacgccccaa 1980
gacagtcact ccagcctcct cggccaagac gtctcctgcc aagcagcagg ccccactgt 2040
ccggaacctg caccagtctg gattcagttt gtctggtgct cagattgatg acaacattcc 2100
ccgccgacc acccagcgta tcgtggcgcc ccccggtggc cgtgccaaca tcaccagcct 2160
gggetagagc tctgggctg tgccgtccac tggggactgg ggatgggaca cctgaggaca 2220
ttctgagact tctttcttc ttcctttttt ttttttgtt tttttttta agagcctgtg 2280
atagtactg tggagcagcc agttcatggg gtccccctg gggccccaca ccccgctctt 2340
caccaagagt tactgatttt gctcatccac ttccctacac atctatgggt atcacacca 2400
agactacca ccaagctcat acagggaacc acaccaaca cttagacatg cgaacaagca 2460
gccccagcg agggctcctc tcgcttcaa cctcctagt tctgttagca tcttctttt 2520
catgggggga ggaagataa agtgaattgc ccagagctgc cttttcttt tctttttaa 2580
aattttaaga agttttctt gtggggctgg ggaggggccc gggtcaggga gagtctttt 2640
ttttttttt tttaaact aaattgaa atttaattcc atattaatac aaggggtttg 2700
aactggacat cctaagtgt caattacgtc atcaccagc tgattccggg tggttggcaa 2760
actcatcgtg tctgtcctga gaggtccac aatgcccacc cgcctcgcca tctgtagtc 2820
ttcagggtca gctgttgata aaggggcag cttgcgttat tggcctagat tttgtgcag 2880
attaatcct ttgaggatc tcttctctt taccatttt ctgctgtctc tcaactctc 2940
ttctctctc tagctttta atctatgaat atttctgtg ctgtctctc ctctctctg 3000
gtttcctca gccctgtct cggagacggt gttttctcc cttgcccct tatcttttca 3060
cctcccaggt ctaccattc atggtggtcg ttgggtccc ctaaaggatt tgagcgtttg 3120
ccattgcaag catagtgtg tgtcatcctg gtccatgtag gactgggtct aaccaectgc 3180
catcatgagg atgtgtgcta gagtgtgga cctggccaa gtgcaggaat gggccatgcc 3240
gtctcaccca cagtatcaca cgtggaaccg cagacagggc ccagaagctt tagaggtatg 3300
aggetgcaga accggagaga ttttctctg tgcagtgtc tctggctaaa gtcacgggtca 3360
aacctaaaca ccgagcctca ttaacccaag tgaaccaacc aaagtacca gttcagaagt 3420
gctaagctaa taggagtctg acccgagggc ctgctgttc ctggttaagt atcttttgag 3480
attctagaac acatgggagc tttttattt cggggaaaaa ccgtatttt ttctgttcca 3540
attatttcta aagacacact acatagaaa aggccctata aactcaaaaa gtcattggga 3600
aactaaagt ctattctact ttgcaagagg agaaatgtgt tttatgaacg atagatcaca 3660

-continued

```

tcagaactcc tgtggggagg aaaccttata aattaaacac atggccccct tagagaccac 3720
aggtgatgtc tgtctccatc cttccctctc cttttctgtc acctttccce ctagctggct 3780
cctttggacc taccctgtc cttgctgact tgtgttgcac tgtattccaa acgtgtttac 3840
aggttctctt aagcaatggt gtattttgcag gcttttctga ataccaaaac tgctttttgt 3900
aaagcgtaaa aacatcacaa agtaggtcat tccatcacca cccttgtctc tctacacatt 3960
ttgcctttgg ggatctgggt ggggttttgg gttttttggt gttgttgttt atttgttatt 4020
ttaaaggtaa attgcacttt taaaaaaaaa attggttgac ttaatatatt tgcttttttt 4080
ctcacctgca cttagaggaa atttgaacaa gttggaaaaa aacaattttt gtttcaattc 4140
taagaaacac ttgcagctct agtattcact tgagtcttcc tgtttttcct gtaccgggtc 4200
atggttaattt ttggttgttt tgggtgtttt cttaaaaaac aagttaaac ctgacgattt 4260
ctgcaggctg tgtaagcatg tttacctggt ggcttgcctt gtgtgtctgt taaatgaatg 4320
tcatatgtaa atgctaaaat aaatcgacag tgtctcagaa ctgaataact gcagtgactt 4380
gatgctctaa aacagtgtag gatttaagaa tagatggttt ttaatcctgg aaattgtgat 4440
tgtgaccocat gagtggagga actttcagtt ctaaagctga taaagtgtgt agccagaaga 4500
gtactttttt ttttgaacc actgtcttga tggcaaaata attatggtaa aaaacaagtc 4560
tcgtgtttat tattccttaa gaactctgtg ttatattacc atggaacgcc taataaagca 4620
aaatgtgggt gtttcaggaa aaaaaaaaaa aaaaa 4655

```

<210> SEQ ID NO 44

<211> LENGTH: 4417

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 44

```

gctgtaacac tcaccgtgaa ggtctgcagc ttcactcccg agccagcgag accacgaacc 60
caccagaagg aagaaactct gaacacatct gaacatcaga agggacagac tccagacgcg 120
ccaccactct gctaacacca gatagtggaa agaaaccatg tgetgaaatg tttgacgaca 180
ctgatggttt gactctgcta actggaatgg cttattgtgc aagaaagtac acctggtcgg 240
gtcctggggc tcactctctag caccagcaaa gatttctgaa gacgtcttcc tagaaatgac 300
tggaagttt caagaggcat aagatacagc atttctctct aggcctgaa gaagtatcaa 360
gtgggctttg acattgctgt ggtgagagcg acccctcctc acctggagaa ctgggaaatg 420
tggattctca gggaccgctc tgttcacgag ctccagctg tgetgctggc cctggctcctg 480
gggcgctgag cegcatctgc aatagcacac ttgcccggcc acctgctgce gtgagccttt 540
gctgctgaag cccctggggt cgcctctacc tgatgaggat gtgcaccccc attagggggc 600
tgetcatggc ccttgcagtg atgtttggga cagcgatggc atttgcaacc ataccctgga 660
tcacctggga gcacagagag gtgcacctgg tgcagtttca tgagccagac atctacaact 720
actcagcctt gctgctgagc gaggacaagg acaccttcta cataggtgcc cgggagggcg 780
tcttcgctgt gaacgcactc aacatctccg agaagcagca tgaggtgtat tggaaaggtct 840
cagaagacaa aaaagcaaaa tgtgcagaaa aggggaaatc aaaacagaca gagtgcctca 900
actacatccg ggtgctgagc ccactcagcg ccacttcctt ttactgtgtg gggaccaacg 960
cattccagcc ggcctgtgac cacctgaaat taacatcctt taagtctctg gggaaaaatg 1020

```


-continued

aagatggcaa aggaagatgt ccctttgacc cagcacacag ctacacatcc gtcatggttg	1080
atggagaact ttattcgggg acgtcgtata attttttggg aagtgaacce atcatctccc	1140
gaaattcttc ccacagtect ctgaggacag aatatgcaat cccttggctg aacgagceta	1200
gttctcgtgt tgctgacgtg atcccgaaaa gccagacag ccccagcggc gaggatgaca	1260
gggtctactt cttcttcaag gagggtgctg tggagtatga gtttctgttc aggggtgctga	1320
tcccacggat agcaagagtg tgcaaggggg accagggcgg cctgaggacc ttgcagaaga	1380
aatggacctc cttcctgaaa gcccgactca tetgctccc gccagacagc ggcttggctt	1440
tcaatgtgct ggggagtgtc ttctgtctca ggtcccggg cctgaagtg cctgtgttct	1500
atgcactctt caccacacag ctgaacaacg tgggctgtc ggcagtgtgc gcctacaacc	1560
tgccacagc cgaggaggtc ttctcccacg ggaagtacat gcagagcacc acagtggagc	1620
agtccacac caagtgggtg cgctataatg gcccggtacc caagccgagg cctggagcgt	1680
gcacgcacag cgaggcagcg gccccaact acaccagctc cttgaatttg ccagacaaga	1740
cgctgcagtt cgttaagac cacccttga tggatgactc ggtaaccca atagacaaca	1800
ggcccagggt aatcaagaaa gatgtgaact acaccagat cgtggtggac cggaccagg	1860
ccctggatgg gactgtctat gatgtcatgt ttgtcagcac agaccgggga gctctgcaca	1920
aagccatcag cctcagcagc cctgttcaca tcctcagga gaccagctc ttccaggact	1980
ttgagccagt ccagaccctg ctgctgtctt caaagaagg caacaggttt gtctatgctg	2040
gctetaactc gggcgtggtc caggcccgc tggccttctg tgggaagcac ggcacctgcg	2100
aggactgtgt gctggcggc gaccctact gcgcctggag cccgccaca gcgacctgcg	2160
tggctctgca ccagaccgag agccccagca ggggtttgat tcaggagatg agcggcgatg	2220
cttctgtgtg cccggcctcg tctcctaage ccctccctcc tcttggctcc tcttccctgt	2280
cctgtctggg ccagtgtggg gacaggaggc tttcctctcc ctggacccc tggccagcct	2340
cgggtgcggg gccagacag agctcagagg tctccttctc gccgccctc ctgagtgacc	2400
aggcacagca cgtgcacgcc ctggggaact tetacctct ctgccaggcc acaggctctg	2460
cagacattcg ctttctctgg gagaagaatg ggcgagctct ggagacctgt gtcctgtgc	2520
agaccatgc actgccgat ggcagggccc atgcactcag ctggctgcag gacgccatca	2580
gggaaagcgc tgagtatcgc tgctctgtcc tctcctcagc agggaaacaag acttcgaagg	2640
tgcaggttgc tgtgatgaga cctgaagtga cccaccagga gaggtggacc agagagctct	2700
ctgcctggag ggctgtggct ggggagcac accggatgat gcagagctgg aggaaggcgt	2760
gggaaagctg tagcaaggac accctgtagc caccaggaag gagtccctga caccgacctc	2820
aacccaaca agaccctgct gccactgacc acagccacc ccggagaagg cctggctccc	2880
cacaactgtg aactgtcttg cccaagcctg ctctgaacac agccattggg ccaccacctg	2940
atgggcagag ggggacagct ggagaagcct ggaacccaag tgggcctgtg acaggaacta	3000
agacttaaaa aattaggtgc ttacctggga cagtaagttc tgtctggcac aagcaggtaa	3060
ccaggatggc taacaggctt tgatagctgc tcgtgaacta aacagcagg gtgtgtgcag	3120
gttctcctc tacggtcagg cagcagctc tgaaggctga tctacaccg tcccagtgac	3180
tccccttgac agagtgcctc caccctctaa tagccaacag ggtagcatg gccagcacag	3240
atcgtgctt ttattgatgc aatcaagcc tgctgcttct cctcctgca gacttagcca	3300

-continued

```

aggaactcca agatgcatga ctgggacaag aaaaggtgag actccacatg gaaatgcctt 3360
gccctaaacc ttgaatgact gtgagatgcg atctgggagt gcatctgtca agtctttgtg 3420
ttttcttcac taacctcaga atactgggct ctattttate aagcgetgca gtttatgctt 3480
ctgtcccgtc aatgctcagc ttctgcaaca ggacacccaa cttgatgagc aaagccaaat 3540
aggccaatta tgcaaatctc ctggtgccat attaaatttc ttgacgatgg aatgagtctc 3600
atgagtgttt tgttctacct gctttcaagt ctctaattat taaagctgta tctctgaaga 3660
ctgtgtcact gtgtgtgtga acttgtccta aagctactca gcctttaatc ttacacacac 3720
gtctcttctt gtctgttgaa tgacagtttt catgtctatc ataaaaccaa agcctctgtt 3780
aaaagtcaag ccgcacccct ctggtgatcc tagcaaatc tgagtgtctt cccagcagtg 3840
tgacaatgac ctgttttgca tcccctcttt ctggagctgg acaaattctc taccagcctt 3900
tgtgtgggat cagcatacat cgcctgctaa ttcttcagg atccatcaca acaggtgtcc 3960
tgaagatgct ggagacaccc tggttgtctc cacacgttcc ccctccgac cccaagtcca 4020
gaggcccagc tgectgtgag gtgtgtgctt gcccatccag ccaaggatgc cagtcttctt 4080
cacggaacca tcacatactc ataacctgaa gttttctctt aaaatatcca tcagctcact 4140
gtggttcttg ctttgggtgt ggcttcaacc actacaaact gatgagtcaa atgctatggg 4200
ctttaggctt atattcttgg tctgttttct tctctcttct cctgaagtct ggatttcaag 4260
cactttcaca cttaacaaaa taattacata cttgaagttt tcgtaatgtg gagtgttcta 4320
ctgggaaatg gagttatgag gatgaatttc tgagtcttct tttgctctgc tggaaaaaat 4380
aaaaatagag ttgtacattg aaaaaaaaa aaaaaaa 4417

```

<210> SEQ ID NO 45

<211> LENGTH: 3108

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 45

```

gcctctggct cctcagggca ttcccggcgg ctccgggttt ggcaacgagc acgggggagt 60
gcgactgcgt ctccggcagc atggccgaga agcggcacac acgggactcc gaagcccagc 120
ggctccccga ctccctcaag gacagcccca gtaagggcct tggacctgac ggatggattt 180
tgggtggcgtt ctcattctta ttcaaccgta taactttccc aatctcaata tggatgtgca 240
taaagattat aaaagagtat gaaagagcca tcactcttag attgggtcgc attttacaag 300
gaggagccaa aggacctggc ttgtttttta ttctgcoatg cactgacagc ttcatcaaag 360
tggacatgag aactatttca tttgatattc ctctcagga gatcctcaca aaggattcag 420
tgacaattag cgtggatggt gtggtctatt accgcgttca gaatgcaacc ctggctgtgg 480
caaatatcac caacgctgac tcagcaaccc gtcttttggc acaaactact ctgaggaatg 540
ttctgggac caagaatctt ttctcagatc tctctgacag agaagaaatt gcacacaaca 600
tgcagtctac tctggatgat gccactgatg cctggggaat aaaggtggag cgtgtggaag 660
ttaaggatgt gaaactacct gtgcagctcc agagagctat ggctgcagaa gcagaagcgt 720
cccgcagggc ccgcgccaag gttattgac ccgaaggaga aatgaatgca tccagggctc 780
tgaagaagc ctccatggct atcactgaat ctctgcagc ccttcagctc cgatacctgc 840
agacactgac caccattgct gctgagaaaa actcaacaat tgtcttccct ctgcccatag 900

```

-continued

atatgctgca aggaatcata ggggcaaac acagccatct aggctagtgt agagatgagc	960
gctagccttc caagcatgaa gtcggggacc aaattagcct ttaactcata aagagagggt	1020
agggcttttc tttttccata tgtcaattgt ggtgttccca gaatgtatag cagttataaa	1080
aataggtgaa agaattgtta gcttgtaaat actgagagat tgggtattta tataaggtaa	1140
tctgttagtc ttaaaatagt taaaagtttg tatTTTTaga ttattatgta gtaggttaga	1200
tccctcttgt tttgacttcc actgactcat tctgaacccc ctaagcacc aggccagagg	1260
caagaacctg ggetgtaact gccacctgac accgetgact ggctaaatgc tttgcagaaa	1320
gtgatgacct tacaccacaa ccagcttctc caggtcatat gtgccttacc tccagagagt	1380
ctTTTTTTT tttttctga gatggagttt cactcttgtt gcccaggctg gagtgcaata	1440
gcatgatctc ggctcactgc aacctccgcc tctggggtc aagagattct cctgcctcag	1500
cctccccagt agctgggatt acaggctcat gccaccatgc ccagctaatt tttgtattat	1560
tattattgtt ttttagtaga gacgggggtt caccatgttg gccaggetag tcacgaactc	1620
ctaacctcag gtgateccacc cacctctgcc tcccaaagt ctgggattac aggcattgagc	1680
taccacacct ggtttggaga gtcttaatta aggaaatctc cctaatgttc atttattttc	1740
taaatccaga ccgtgtttca gaataatcct tacttgagag tagccatttt cttgcctgta	1800
cttgtcagaa cttagggaaa tagccaagac taatgaaaa gattactcta acccttaaaa	1860
gacttttaaa ttcactacta gagtggctcat tttaaaaata catccatgtt ttaacttatt	1920
tgagccttct ttatgagtaa atgattcctc cttgttctgt ctttcaaacc agctaaatat	1980
ttgtcacaaa agtgettttt tctcactgtt gcctattttc atatatcagg ttttaaatag	2040
ttttaatttt ttaataaaat tttctctacg ttctatatgc aattgttata tatctatttg	2100
aatagctgaa ggactaaaat acttttttaa gagataaact caggaaacca ttatatttta	2160
ctatctgcat gctgttaact gtggtacact gtgaaatag ttgattacaa acccattcat	2220
tacatagtat aaggaattca cagtatattg actatatagt gtctaatgat cttgggcaga	2280
tactgtcaaa cttacaatat ctatatagat gtaggtcttt ttaaatttac ctagtcatc	2340
ttctatcatg tatattgatg ctgaaagagg aactggtcag ctcctctgga caacaaatc	2400
ttagtctata atattaggag acatcttctg ttttgcaaat gtctgtgaat ctgagcaacc	2460
tggcattctg cttactggcc agaaagctgg cgggtgacat ttgtaacatt tcctctttga	2520
gactctgagt tcacctagag aagtetaage ataacagctt tctttccag cactgacctt	2580
tatagctctc tttagctcaa ccactctgtc catccagcca atggatgtcc cttccctgt	2640
accccaatth caagcttatt ttaggaagcc ttgaaactacc atgtatctg gctcctagct	2700
gagtttatta gaggtatgga gcagtgaac ttaaactcaa gttgcactta cattttgaa	2760
tttaaatga tggttttatc tgttgtgtga agtggttcac ccttgaggac caggagcctc	2820
catatctga ctgaaaacct tttctgagac ttagagtaac agtgcttttg gttccttgag	2880
ttctcctgtc tccagatacc aaatgacctt gacttttctg ccttgtgaat tegtagtcca	2940
atcagctgaa attaaatcac ttggggaggga cgcatagaag gagctctagg aacacagtgc	3000
cagtgcagaa gtttctccag gtggcctccc tttccaacaa tgtacataat aaagtgtatg	3060
cactttcact aataaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	3108

-continued

```

<210> SEQ ID NO 46
<211> LENGTH: 4090
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 46
gggcgctccc ggagtatcag caaaaggggtt cgccccgccc acagtgcceg gctccccccg      60
ggatcaaaa gaaggatcgg ctccgcccc gggtccccg ggggagttga tagaagggtc      120
cttcccaccc ttgcegtcc ccaactcctgt gcctaegacc caggagcgtg tcagccaaag      180
catggagaat caagagaagg cgagtatcgc gggccacatg ttcgacgtag tctgatcgg      240
agggtggcatt tcaggactat ctgtgcccaa actcttgact gaatatggcg ttagtgtttt      300
ggttttagaa gctcgggaca gggttggagg aagaacatat actataagga atgagcatgt      360
tgattacgta gatgttgggt gagcttatgt gggaccaacc caaacagaa tcttacgctt      420
gtctaaggag ctgggcatag agacttacia agtgaatgto agtgagcgtc tegtccaata      480
tgtcaagggg aaaacatata catttcgggg cgcccttcca ccagtatgga atcccattgc      540
atatttgat tacaataatc tgtggaggac aatagataac atggggaagg agattccaac      600
tgatgcaccc tgggaggctc aacatgctga caaatgggac aaaatgacca tgaagagct      660
cattgacaaa atctgtctga caaagactgc taggcgggtt gcttatcttt ttgtgaatat      720
caatgtgacc tctgagcctc acgaagtgtc tgcctgtgg ttcttgtggt atgtgaagca      780
gtgccccggc accactcgga tattctctgt caccaatggt ggcaggaac ggaagtttgt      840
agggtgatct ggtcaagtga cgaacggat aatggacctc ctcggagacc aagtgaagct      900
gaaccatcct gtcactcacg ttgaccagtc aagtgacaac atcatcatag agacgctgaa      960
ccatgaacat tatgagtgca aatacgtaat taatgcatc cctccgaact tgactgocaa     1020
gattcacttc agaccagagc ttccagcaga gagaaaccag ttaattcagc ggcttccaat     1080
gggagctgtc attaatgca tgatgtatta caaggaggcc ttctggaaga agaaggatta     1140
ctgtggctgc atgatcattg aagatgaaga tgctccaatt tcaataacct tggatgacac     1200
caagccagat gggtcactgc ctgccatcat gggcttcatt cttgcccgga aagctgatcg     1260
acttgctaag ctacataaag aaataaggaa gaagaaaatc tgtgagctct atgccaaagt     1320
gctgggatcc caagaagctt tacatccagt gcattatgaa gagaagaact ggtgtgagga     1380
gcagtactct gggggctgct acacggccta ctccctcct gggatcatga ctcaatatgg     1440
aagggtgatt cgtcaacccg tgggcaggat tttctttgcg ggcacagaga ctgccacaaa     1500
gtggagcggc tacatggaag gggcagttga ggctggagaa cgagcagcta gggaggtctt     1560
aaatggtctc ggaaggtga ccgagaaaaga tatctgggta caagaacctg aatcaagga     1620
cgttccagcg gtagaaatca cccacacctt ctgggaaagg aacctgcctt ctgtttctgg     1680
cctgtgaag atcattggat tttccacatc agtaactgcc ctgggggttg tgetgtacaa     1740
atacaagctc ctgccacggt cttgaagttc tgttcttatg ctctctgctc actggttttc     1800
aataccacca agaggaaaat attgacaagt ttaaaggctg tgcattggg ccatgtttaa     1860
gtgtactgga ttaactacc ttggcttaa ttccaatcat tgttaaagta aaaacaattc     1920
aaagaatcac ctaattaatt tcagtaagat caagctocat cttatttgtc agtgtagatc     1980
aactcatggt aattgataga ataaagcctt gtgatcaact tctgaaatc acaaagttaa     2040
acgtgatgtg ctcatcagaa acaatttctg tgtcctgttt ttattccctt caatgcaaaa     2100

```

-continued

tacatgatga tttcagaaac aaagcatttg actttctgtc tgtggagtg gagtaggtga	2160
aggcccagcc tgtaactgtc cttttcttc ccttaggcaa tggggaactg tcattacaga	2220
gcctagagge tcacagcctc ctggaggaag cagcctccac tttggatcag gaaatagtaa	2280
aggaaagcag tgttgggggt agcggcatgc agaccctcag accagaatgg ggacatcttg	2340
tggctctgcty cctcaggaat ctctgacca cttgtagtcc ctccgacttc tctagacatc	2400
tagtctcagt gctagcttat ttgtatthtt cctctttcac ttcttatgga ggagagtgtt	2460
taactgagtt agaatggtga aactgacttg ctgtgactta tgtgcagctt tccagttgag	2520
cagaggaaaa tagtggcagg actgtcccc aggaggactc cctgcttagc tctgtgggag	2580
accaactacg actggcatct tctcttcccc ctggaaggca gctagacacc aatggatcct	2640
tgtcagttgt aacattctat ttcaacttca ggaaagcagc agttttcttt taatttttcc	2700
tatgaccata aaattagaca tacctctcaa cttacatag tcttcaacat ggttaacctct	2760
gcataaatat tagcaaaagc tgccaatttc tcttaagtac tgaatacat atgataaatt	2820
tgactgttat ttgttgagac tatcaaacag aaaagaaatt agggctctaa tttccttaa	2880
gcaagctcac ttgctttagt tgtaagttt tataaaagac atgaaattga gtcattttat	2940
atatgaaaac taagttctct atcttaggag taatgtogge ccacaagggt gccacacctct	3000
tgttttcccc ttttaaaaa tcagatthttt aaaagccctt tccaaagttt tcaactgtaa	3060
aatacttctt tttacaatgt atcaacatat tttatthta gggaattaa caattgccag	3120
ggaaaccagc caaccacaag ttattatctc attaacctta tcataaattc aaacctaaat	3180
tgtgtgaccc tgggtgtgagg acataaatct tccaaagttt tgccatctct aagagctgca	3240
tttttctact gctctttacc ttgcatttha gctaatthtag gagttttgag aatgtattgg	3300
atagctcca gtacataaagg agttgccga tattatatca gactgctttg agaaatctca	3360
tccctagtct attgcagttg tttctattag cttactgatt aactcagtc tgacacacct	3420
tttgggaaat gctgatttha acttcttaac tggcaacagt tggaaacagta atcagtttgc	3480
taacatattt aaagtcttga atgtgaaga actcatgtga tttacccttt tcaactthtt	3540
ggaaaacgat ttaatttatt ctaattagat taacctatt aatctatgga ttgggtatca	3600
aaatgaatgc cagtcagat gtgcctagac acgaaattgg agctgaggac tctcacgata	3660
tgcaagttca tccaactgga agataccata agctttttct ctgaaccaga gaaatgaaag	3720
tcagtttaag aggctgatag atcttggccc tgttaaggca tccacttca agttctgaag	3780
gctgagtcag cccactcca cagttaggcc aagaattaga ttttaaaact tcatctgtct	3840
gtcccagtt actgttaaat aaggcctcat cctccactga agagtatgga ttgaaggatt	3900
gtgaactatg tttagtgtga ttgtgaactt ggtgcctaat gttccatgct tgaagtttgc	3960
cccagtgcta cagttggag tatacctatg tgtgtgcttt gccactgaag taagattttg	4020
cctgtatggt actgttttgt ttgttaataa agtgcactgc caccctcaat gcaaaaaaaa	4080
aaaaaaaaa	4090

What is claimed is:

1. A method for treating breast cancer in a patient having breast cancer cells, the method comprising:

a) administering an effective amount of a glucocorticoid receptor antagonist to said patient, wherein the patient

expresses a detectable level of glucocorticoid receptor (GR) prior to administration of the GR antagonist;
b) then administering an effective amount of radiation, immunotherapy, or at least one chemotherapeutic.

2. The method of claim 1, wherein said breast cancer cells undergo apoptosis.

3. The method of claim 1, wherein said breast cancer patient was previously treated with chemotherapy.

4. The method of claim 1, wherein the breast cancer patient is determined to have breast cancer cells that express a detectable level of glucocorticoid receptor.

5. The method of claim 1, wherein the breast cancer patient is determined to have breast cancer cells that do not express a detectable level of estrogen receptor but do express a detectable level of glucocorticoid receptor.

6. The method of claim 1, wherein the glucocorticoid receptor antagonist is GR- selective.

7. The method of claim 1, wherein the breast cancer patient was previously administered said effective amount of radiation, immunotherapy, or at least one chemotherapeutic more than two weeks prior to the administration of said glucocorticoid receptor antagonist.

8. The method of claim 7, wherein breast cancer cells in the breast cancer patient were resistant to apoptosis.

9. The method of claim 7, wherein breast cancer cells in the breast cancer patient express a detectable level of glucocorticoid receptor.

10. The method of claim 7, wherein the breast cancer cells in the breast cancer patient do not express a detectable level of estrogen receptor.

11. The method of claim 1, wherein the effective amount of radiation, immunotherapy, or at least one chemotherapeutic is administered within 1 week of the administration of the glucocorticoid receptor antagonist.

12. The method of claim 5, wherein the effective amount of radiation, immunotherapy, or at least one chemothera-

peutic is administered within 1 week of administration of the glucocorticoid receptor antagonist.

13. The method of claim 11, wherein the effective amount of radiation, immunotherapy, or at least one chemotherapeutic is administered within 24 hours of the administration of the glucocorticoid receptor antagonist.

14. The method of claim 12, wherein the effective amount of radiation, immunotherapy, or at least one chemotherapeutic is administered within 24 hours of the administration of the glucocorticoid receptor antagonist.

15. The method of claim 11, wherein the glucocorticoid receptor antagonist is administered up to three days prior to administering the effective amount of radiation, immunotherapy, or at least one chemotherapeutic.

16. The method of claim 12, wherein the glucocorticoid receptor antagonist is administered up to three days prior to administering the effective amount of radiation, immunotherapy, or at least one chemotherapeutic.

17. The method of claim 1, comprising further administration of the glucocorticoid receptor antagonist after administration of the effective amount of radiation, immunotherapy, or at least one chemotherapeutic.

18. The method of claim 5, comprising further administration of the glucocorticoid receptor antagonist after administration of the effective amount of radiation, immunotherapy, or at least one chemotherapeutic.

19. The method of claim 1, wherein the breast cancer cells are PR- and HER2-.

20. The method of claim 5, wherein the breast cancer cells are PR- and HER2-.

* * * * *