



US 20250249024A1

(19) **United States**(12) **Patent Application Publication** (10) **Pub. No.: US 2025/0249024 A1**

Pan et al.

(43) **Pub. Date:** Aug. 7, 2025(54) **METHODS AND COMPOSITIONS RELATED TO GLUCOCORTICOID RECEPTOR ANTAGONISTS AND BREAST CANCER**(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)(73) Assignee: **The University of Chicago**, Chicago, IL (US)(21) Appl. No.: **19/081,384**(22) Filed: **Mar. 17, 2025****Related U.S. Application Data**

(63) Continuation of application No. 18/541,675, filed on Dec. 15, 2023, now Pat. No. 12,280,066, which is a continuation of application No. 17/577,882, filed on Jan. 18, 2022, now Pat. No. 11,883,417, which is a 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.**

| | |
|---------------------|-----------|
| <i>A61K 31/575</i> | (2006.01) |
| <i>A61J 1/00</i> | (2023.01) |
| <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 31/567</i> | (2006.01) |
| <i>A61K 31/58</i> | (2006.01) |
| <i>A61K 31/7068</i> | (2006.01) |
| <i>A61K 33/243</i> | (2019.01) |
| <i>A61K 39/395</i> | (2006.01) |
| <i>A61K 45/06</i> | (2006.01) |
| <i>A61N 5/00</i> | (2006.01) |

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

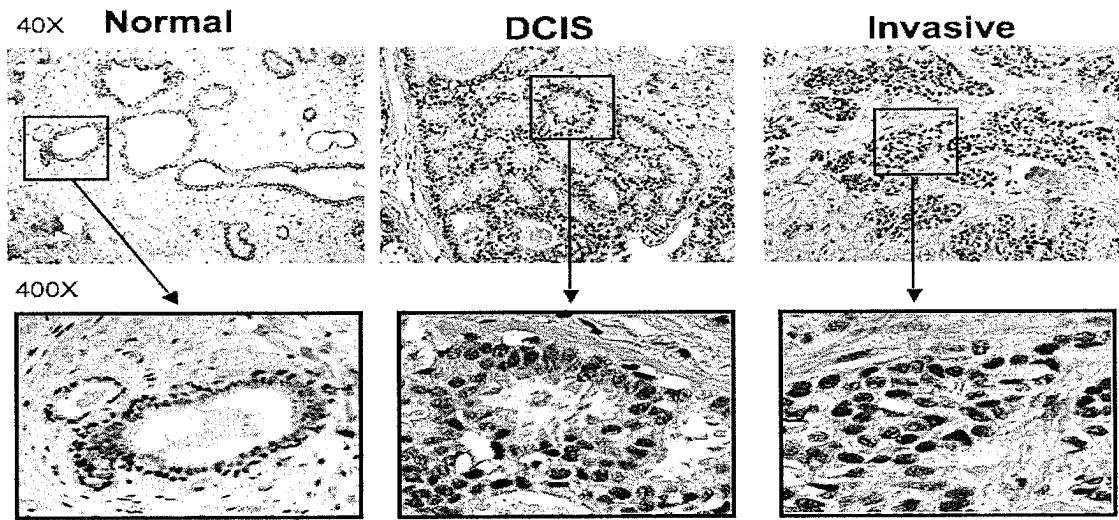
| | |
|-----------|---|
| CPC | <i>A61K 31/575</i> (2013.01); <i>A61J 1/00</i> (2013.01); <i>A61K 31/282</i> (2013.01); <i>A61K 31/337</i> (2013.01); <i>A61K 31/357</i> (2013.01); <i>A61K 31/4745</i> (2013.01); <i>A61K 31/567</i> (2013.01); <i>A61K 31/58</i> (2013.01); <i>A61K 31/7068</i> (2013.01); <i>A61K 33/243</i> (2019.01); <i>A61K 39/3955</i> (2013.01); <i>A61K 45/06</i> (2013.01); <i>A61N 5/00</i> (2013.01) |
|-----------|---|

(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**



Bolova et al. *Breast Cancer Treatment and Research*, 2008

FIG. 1

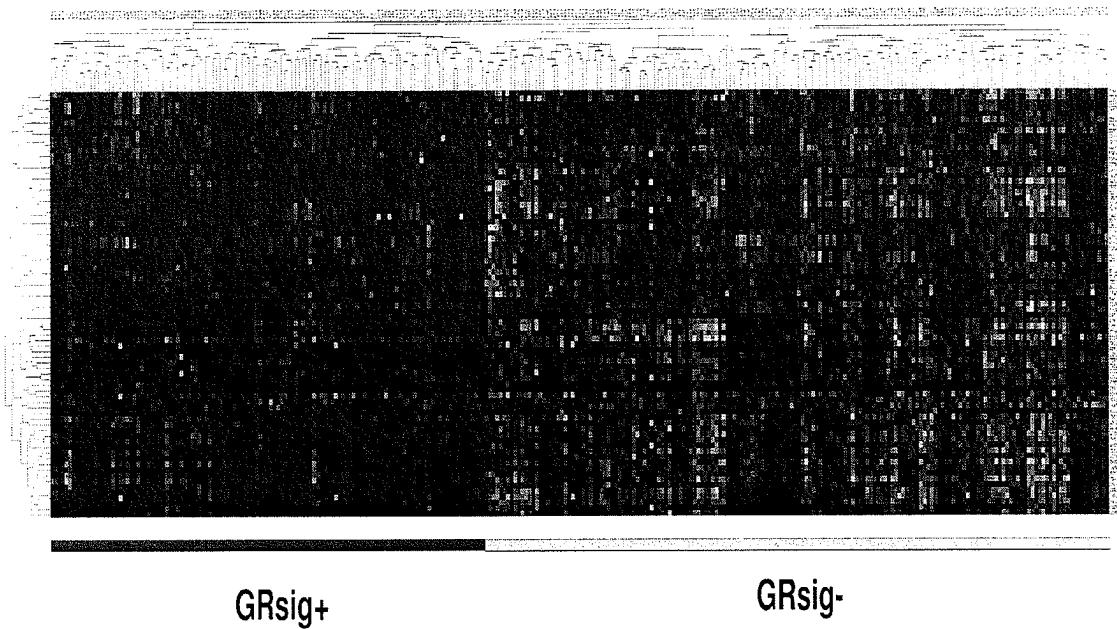


FIG. 2

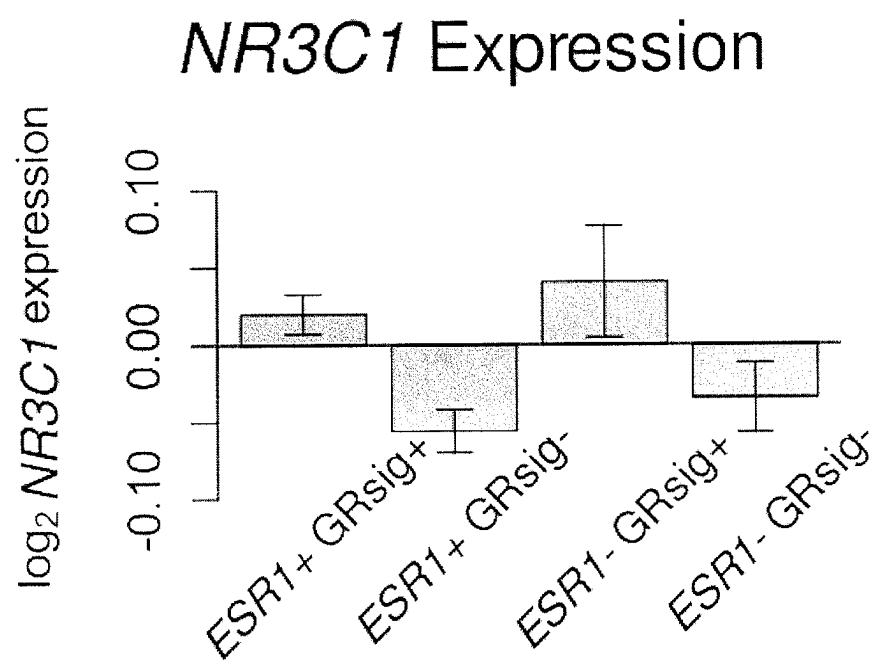


FIG. 3

NKI-295 RFS by Signature

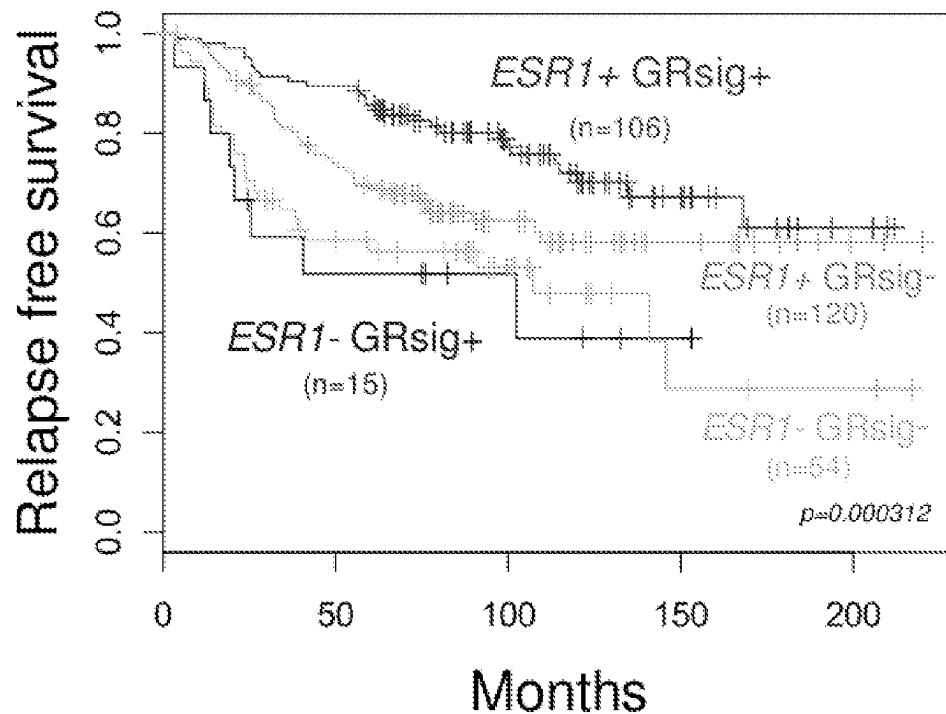


FIG. 4

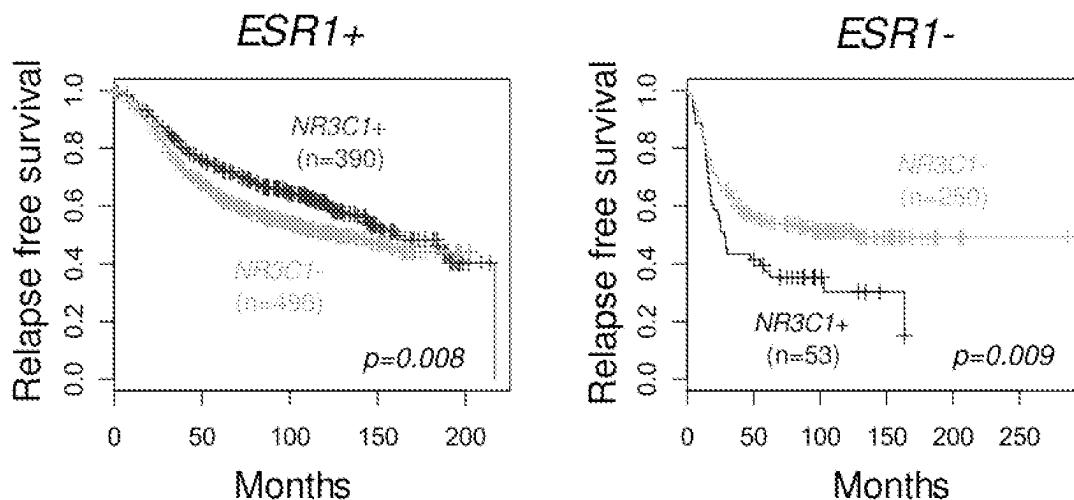


FIG. 5

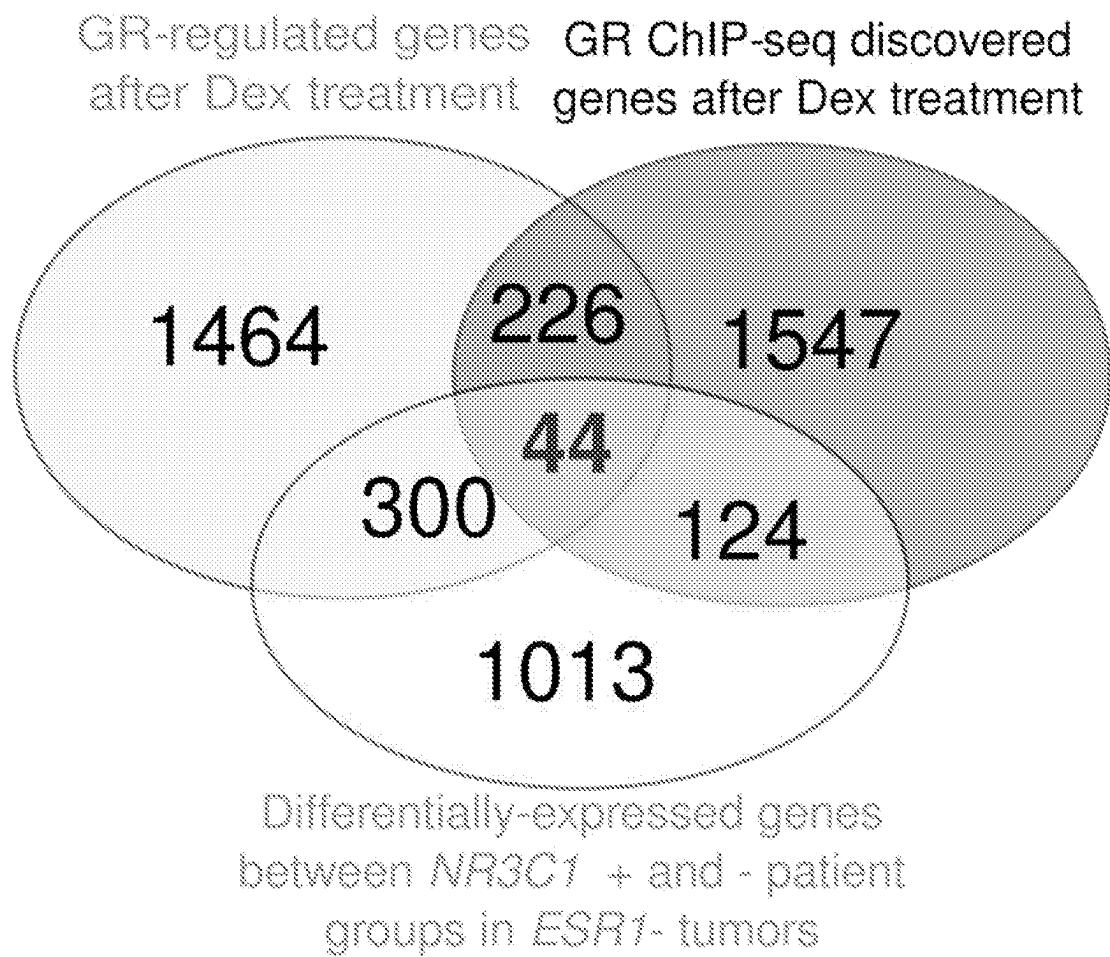
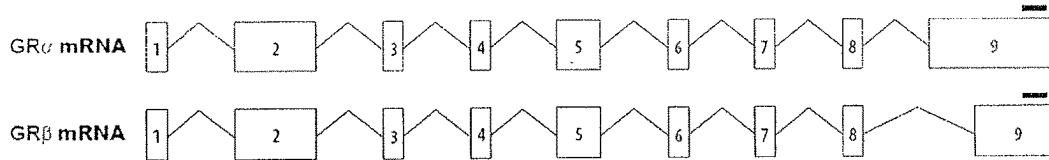


FIG. 6



Query = GR alpha
Length=6784

18665 = GR beta

ALIGNMENTS

| | | | |
|-------|-----|---|-----|
| Query | 1 | GGCGCCGCCCTCCACCCGCTCCCCGCTCGGTCCCCTCGCTCGCCCAGGCCGGCTGCCCT | 60 |
| 18665 | 1 | GGCGCCGCCCTCCACCCGCTCCCCGCTCGGTCCCCTCGCTCGCCCAGGCCGGCTGCCCT | 60 |
| Query | 61 | TTCGCGTGTCCCGCGCTCTCTCCCTCCGGCGCCCTCCTCCATTTCGAGCTCGTGTTC | 120 |
| 18665 | 61 | TTCGCGTGTCCCGCGCTCTCTCCCTCCGGCGCCCTCCTCCATTTCGAGCTCGTGTTC | 120 |
| Query | 121 | TGTGACGGGAGGCCCGAGTCACCGCCTGCCGTGGGGACGGATTCTGTGGGTGGAAGGAG | 180 |
| 18665 | 121 | TGTGACGGGAGGCCCGAGTCACCGCCTGCCGTGGGGACGGATTCTGTGGGTGGAAGGAG | 180 |
| Query | 181 | ACGCCGCAGCCGGAGCGGCCGAAGCAGCTGGGACCGGGACGGGACCGCCGGAAAC | 240 |
| 18665 | 181 | ACGCCGCAGCCGGAGCGGCCGAAGCAGCTGGGACCGGGACGGGACCGCCGGAAAC | 240 |
| Query | 241 | CTCGACCCCGCGGAGCCCGCGCAGGGCTGGCTTGCAAGCTGGCAATGGGAGA | 300 |
| 18665 | 241 | CTCGACCCCGCGGAGCCCGCGCAGGGCTGGCTTGCAAGCTGGCAATGGGAGA | 300 |
| Query | 301 | CTTTCTAAATAGGGGCTCTCCCCCACCATGGAGAAAGGGGCGCTGTTACTTCCTt | 360 |
| 18665 | 301 | CTTTCTAAATAGGGGCTCTCCCCCACCATGGAGAAAGGGGCGCTGTTACTTCCTT | 360 |
| Query | 361 | tttttAGaaaaaaaaaaaaTATATTCCTCTGCTCCTCTGCGTTACAAGCTAAGTTGT | 420 |
| 18665 | 361 | TTTTTAGAAAAAAAAATAATTTCCCTCTGCTCCTCTGCGTTACAAGCTAAGTTGT | 420 |
| Query | 421 | TTATCTCGGCTGCCGGAACTGCGGACGGTGGCGGGCGAGCGGCTCCTCTGCCAGAGT | 480 |
| 18665 | 421 | TTATCTCGGCTGCCGGAACTGCGGACGGTGGCGGGCGAGCGGCTCCTCTGCCAGAGT | 480 |
| Query | 481 | TGATATTCACTGATGGACTCAAAGAACATTAACTCCTGGTAGAGAACAAAAACCCAGC | 540 |
| 18665 | 481 | TGATATTCACTGATGGACTCAAAGAACATTAACTCCTGGTAGAGAACAAAAACCCAGC | 540 |
| Query | 541 | AGTGTGCTTGCTCAGGAGAGGGGACATGTGATGGACTCTATAAAACCTAAGAGGAGGA | 600 |
| 18665 | 541 | AGTGTGCTTGCTCAGGAGAGGGGACATGTGATGGACTCTATAAAACCTAAGAGGAGGA | 600 |
| Query | 601 | GCTACTGTGAAGGTTCTGCGTCTCACCTCACTGGCTGTCGCTCTCAATCAGACTCC | 660 |
| 18665 | 601 | GCTACTGTGAAGGTTCTGCGTCTCACCTCACTGGCTGTCGCTCTCAATCAGACTCC | 660 |
| Query | 661 | AAGCAGCGAAGACTTTGGTGATTTCCAAGGCTCAGTAAGCAATGCGCAGCAGCCA | 720 |
| 18665 | 661 | AAGCAGCGAAGACTTTGGTGATTTCCAAGGCTCAGTAAGCAATGCGCAGCAGCCA | 720 |
| Query | 721 | GATCTGTCCAAAGCAGTTCACTCTCAATGGGACTGTATATGGGAGAGACAGAAACAAA | 780 |
| 18665 | 721 | GATCTGTCCAAAGCAGTTCACTCTCAATGGGACTGTATATGGGAGAGACAGAAACAAA | 780 |
| Query | 781 | GTGATGGAAATGACCTGGGATTCCCACAGCAGGGCAAATCAGCCTTCCTGGGGAA | 840 |
| 18665 | 781 | GTGATGGAAATGACCTGGGATTCCCACAGCAGGGCAAATCAGCCTTCCTGGGGAA | 840 |
| Query | 841 | ACACACTTAAAGCTTTGGAGAAAGCAATTGCAAACCTCAATAGGTGACCAGTGTCCA | 900 |
| 18665 | 841 | ACACACTTAAAGCTTTGGAGAAAGCAATTGCAAACCTCAATAGGTGACCAGTGTCCA | 900 |

FIG. 7A

| | | | | | | | | | | | |
|----------------|------|--|--|--|----------------------------------|------------------------|--------------------------|----------------------|----------------------|--------------------|--------------|
| Query 18665 | 901 | GAGAACCCCAAGAGTTCA GAGAACCCCAAGAGTTCA | GAGCATTCCACTGCTGTCTGCTGCC GAGCATTCCACTGCTGTCTGCTGCC | CACAGAGAAGGAGTTT CACAGAGAAGGAGTTT | 960 960 | | | | | | |
| Query 18665 | 961 | CCAAAAACTCACTCTGATG CCAAAAACTCACTCTGATG | TATCTTCAGAACAGCA TATCTTCAGAACAGCA | ACATTTGAAGGGCCAGACTGG ACATTTGAAGGGCCAGACTGG | 1020 1020 | | | | | | |
| Query 18665 | 1021 | AACGGTGGCAATGT AACGGTGGCAATGT | GAATAACCACAGAC GAATAACCACAGAC | CAAAGCACC CAAAGCACC | 1080 1080 | | | | | | |
| Query 18665 | 1081 | TTGGAGTTTCTCTGG TTGGAGTTTCTCTGG | CCCCCAGGTA CCCCCAGGTA | AAAGAGACGA AAAGAGACGA | 1140 1140 | | | | | | |
| Query 18665 | 1141 | CTGTTGATAGAT CTGTTGATAGAT | GAAAACGT GAAAACGT | TTTGCTTCTGC TTTGCTTCTGC | CTGGCAGGATCATT CTGGCAGGATCATT | 1200 1200 | | | | | |
| Query 18665 | 1201 | TT'GGAAGGAAACTCG TT'GGAAGGAAACTCG | AAATGAGCA AAATGAGCA | CTGCAAGCCT CTGCAAGCCT | CTCATTTC CTCATTTC | 1260 1260 | | | | | |
| Query 18665 | 1261 | ATTAAGGATAATGG ATTAAGGATAATGG | GAGAATCTGG GAGAATCTGG | TTTGTCAAG TTTGTCAAG | CCCCCAGTA CCCCCAGTA | 1320 1320 | | | | | |
| Query 18665 | 1321 | AAAACAGAAAAAGA AAAACAGAAAAAGA | AGATTTCA AGATTTCA | TGCAACTCTG CGAACACTCTG | CACCCCTGGG CACCCCTGGG | TAACACTAA TAACACTAA | 1380 1380 | | | | |
| Query 18665 | 1381 | CTGGGCACAGTT CTGGGCACAGTT | TACTGT TACTGT | CGAGCTTCTG CGAGCTTCTG | GGCAAATATA GGCAAATATA | TTGGTAATA TTGGTAATA | 1440 1440 | | | | |
| Query 18665 | 1441 | TCTGCCATT TCTGCCATT | TCTGTTCATGG TCTGTTCATGG | GTTGAGTAC GTAGTACCT | CTCTGGAG CTCTGGAG | AGACAGAT AGACAGAT | GTACCACTAT GTACCACTAT | 1500 1500 | | | |
| Query 18665 | 1501 | AATAACAGCAT AATAACAGCAT | CCCTTTCT CCCTTTCT | CAACAGC CAACAGC | AGGATC AGGATC | AGAC AGAC | AGCTT AGCTT | 1560 1560 | | | |
| Query 18665 | 1561 | ATTCCCCTGG ATTCCCCTGG | TTCCGAAA TTCCGAAA | ATTGGA ATTGGA | ATAGGT ATAGGT | GCCAA GCCAA | AGGATCTG AGGATCTG | ACTTGACT ACTTGACT | 1620 1620 | | |
| Query 18665 | 1621 | TCTCTGGG TCTCTGGG | ACTCTGA ACTCTGA | ACATTCC ACATTCC | CTGTCGA CTGTCGA | ACAG ACAG | TTCTCA TTCTCA | AGCCCC AGCCCC | 1680 1680 | | |
| Query 18665 | 1681 | AGCATGAG AGCATGAG | ACCAGAT ACCAGAT | GTAAGCT GTAAGCT | CTCCTCC CTCCTCC | ACAGCT ACAGCT | CTCAAC CTCAAC | AGCAACA AGCAACA | 1740 1740 | | |
| Query 18665 | 1741 | CCTCCCAA CCTCCCAA | ACTCTGC ACTCTGC | CTGGCT CTGGCT | GTCTGAT GTCTGAT | GAAGCT GAAGCT | TTCAGG TTCAGG | ATGTCATT ATGTCATT | 1800 1800 | | |
| Query 18665 | 1801 | ACTTGTGG ACTTGTGG | AAAGCTG AAAGCTG | TAAAGTT TAAAGTT | CTCTCAA CTCTCAA | AGAGC AGAGC | AGCAG AGCAG | TCAGG TCAGG | ACAGCACA ACAGCACA | 1860 1860 | |
| Query 18665 | 1861 | TGTGCTGG TGTGCTGG | AAAGGAAT AAAGGAAT | GATTGC GATTGC | CATCATG CATCATG | CATA CATA | TCGATA TCGATA | AGCCAG AGCCAG | CATGC CATGC | 1920 1920 | |
| Query 18665 | 1921 | CGCTATCG CGCTATCG | AAAAATGT AAAAATGT | CTTCAGG CTTCAGG | CTGGA CTGGA | ACAG ACAG | CTGCA CTGCA | AGCTCG AGCTCG | Gaaaaaaa Gaaaaaaa | 1980 1980 | |
| Query 18665 | 1981 | ataaaaCG ATAAAAAGG | AAATTCA AAATTCA | GAGGCC GAGGCC | ACTAC ACTAC | ACAGG ACAGG | AGCTC AGCTC | TCTCA TCTCA | AAACCT AAACCT | CTGGT CTGGT | 2040 2040 |
| Query 18665 | 2041 | AAACAAA AAACAAA | ACAATAG ACAATAG | TTCTG TTCTG | CAACG CAACG | TTC TTC | CTG CTG | AGCT AGCT | TTG TTG | TCAGT TCAGT | 2100 2100 |
| Query 18665 | 2101 | GAGCTTATTG GAGCTTATTG | AAACCTGA AAACCTGA | AGGT AGGT | TATATG TATATG | CAGG CAGG | ATG ATG | AGCT AGCT | GT GT | TCAGACT TCAGACT | 2160 2160 |

FIG. 7B

| | | | |
|-------------|------|--|------|
| Query 18665 | 2161 | TGGAGGATCATGACTACGCTAACATGTTAGGAGGGCGGCAAGTGATTGCAGCAGTGAAA | 2220 |
| | 2161 | TGGAGGATCATGACTACGCTAACATGTTAGGAGGGCGGCAAGTGATTGCAGCAGTGAAA | 2220 |
| Query 18665 | 2221 | TGGGCAAAGGCAATACCAGGTTTCAGGAACCTACACCTGGATGCCAAATGACCTACTG | 2280 |
| | 2221 | TGGGCAAAGGCAATACCAGGTTTCAGGAACCTACACCTGGATGCCAAATGACCTACTG | 2280 |
| Query 18665 | 2281 | CAGTACTCCTGGATGTTCTATGGCATTGCTCTGGGTGGAGATCATATAGACAATCA | 2340 |
| | 2281 | CAGTACTCCTGGATGTTCTATGGCATTGCTCTGGGTGGAGATCATATAGACAATCA | 2340 |
| Query 18665 | 2341 | AGTGCAAACCTGCTGTGTTTGCTCCTGATCTGATTATAATGAGCAGAGAATGACTCTA | 2400 |
| | 2341 | AGTGCAAACCTGCTGTGTTTGCTCCTGATCTGATTATAATGAGCAGAGAATGACTCTA | 2400 |
| Query 18665 | 2401 | CCCTGCATGTACGACCAAATGAAACACATGCTGTATGTTCTCTGAGTTACACAGGCTT | 2460 |
| | 2401 | CCCTGCATGTACGACCAAATGAAACACATGCTGTATGTTCTCTGAGTTACACAGGCTT | 2460 |
| Query 18665 | 2461 | CAGGTATCTTATGAAGAGTATCTCTGATGAAAACCTTAAGCTGCTCTCTCAGTTCT | 2520 |
| | 2461 | CAGGTATCTTATGAAGAGTATCTCTGATGAAAACCTTAAGCTGCTCTCTCAGTTCT | 2520 |
| Query 18665 | 2521 | AAGGACGGTCTGAAGAGCCAAGAGCTATTTGATGAAATTAGAATGACCTACATCAAAGAG | 2580 |
| | 2521 | AAGGACGGTCTGAAGAGCCAAGAGCTATTTGATGAAATTAGAATGACCTACATCAAAGAG | 2580 |
| Query 18665 | 2581 | CTAGGAAAAGCCATTGTCAGAGGGAAAGGAAACTCCAGCCAGAACTGGCAGCGGTTTAT | 2640 |
| | 2581 | CTAGGAAAAGCCATTGTCAGAGGGAAAGGAAACTCCAGCCAGAACTGGCAGCGGTTTAT | 2640 |
| Query 18665 | 2641 | CAACTGACAAAACCTCTGGATTCTATGCATGAAACTGGTTGAAATCTCTTAACATATG | 2700 |
| | 2641 | CAACTGACAAAACCTCTGGATTCTATGCATGAA | 2673 |
| Query | 2701 | TTCACAAACATTGGATAAGACCATGAGTATTGAAATTCCCCGAGATGTTAGCTGAAATC | 2760 |
| Query | 2761 | ATCACCAATCAGATACCAAAATATTCAAATGGAAATATCAAAAAACTCTGTTCATCAA | 2820 |
| Query | 2821 | AAGTGACTGCCTTAATAAGAATGGTGCCTTAAAGAAAGTCGAATTAAAGCTTTATTG | 2880 |
| Query | 2881 | TATAAACTATCAGTTGCTCTGTAGAGgtttgtgtttattttattgtttcatct | 2940 |
| Query | 2941 | gttgtttgtttAAATACGCACTACATGTGGTTATAGAGGGCAAGACTGGCAACAG | 3000 |
| Query | 3001 | AAGCAGTTGAGTCGTCACTTTCACTGATGGGAGAGTAGATGGTGAATTTATTAGT | 3060 |
| Query | 3061 | TAATATATCCCAGAAATTAGAAACCTTAATATGTGGACGTAATCTCCACAGTCAAAGAAG | 3120 |
| Query | 3121 | GATGGCACCTAACCAACCAAGTGCCCCAAAGTCTGTGTGATGAACTTCTCTCATACTttt | 3180 |
| Query | 3181 | tttCACAGTTGGCTGGATGAAATTCTAGACTTTCTGTTGGTGTATccccccctGTAT | 3240 |
| Query | 3241 | AGTTAGGATAGCATTGATATGCATGGAAACCTGaaaaaaaGTTACAAGTGTATA | 3300 |
| Query | 3301 | TCAAGAAAAGGGAAAGTTGTCCTTTATAGCTATTACTGTCTGGTTTAACAATTCCATT | 3360 |
| Query | 3361 | ATATTTAGTGAACATACGCTTGCTCATTTTTCTACATAATTTCATTCAAGTTATTGT | 3420 |
| Query | 3421 | ACAGCTGTTAACGATGGCAGCTAGTCGTAGCTTCCAAATAAACTCTAACACATTAAT | 3480 |
| Query | 3481 | CAATCATCTGTGTGAAAATGGTTGGTGCCTAACCTGATGGCACTTAGCTATCAGAAG | 3540 |
| Query | 3541 | ACCACAAAAATTGACTCAAATCTCCAGTATTCTGTCAaaaaaaaGCTCA | 3600 |
| Query | 3601 | TATTTGTATATCTGCTTCAGTGGAGAATTATAGGTTGTGCAAATTAAACAGTCCTA | 3660 |
| Query | 3661 | ACTGGTATAGAGCACCTAGTCCAGTGACCTGCTGGTAAACTGTGGATGATGGTIGCAA | 3720 |
| Query | 3721 | AGACTAATTAAAAATAACTACCAAGAGGCCCTGTCTGTACCTAACGCCATTTTTIGC | 3780 |

FIG. 7C

| | | | |
|-------|------|---|------|
| Query | 3781 | AATGGCTATAATGGCAAGAAAGCTGGTAAACTATTGTCTTCAGGACCTTGAGTAGT | 3840 |
| Query | 3841 | TGTGATAACCTCTAAAAGTTGTGATTCCAGATAACCAGCTGTAACACAGCTGAGAGACT | 3900 |
| Query | 3901 | TTAACATCAGACAAAGTAATTCCCTCACTAAACTTACCCAAAAACTAAATCTCTAATAT | 3960 |
| Query | 3961 | GGCAAAAAATGGCTAGACACCCATTTCACATTCCCCTGTCACCAATTGGTTAATCTTT | 4020 |
| Query | 4021 | CCTGATGGTACAGGAAAGCTCAGCTACTGATTTGTGATTAGAACTGTATGTCAGACA | 4080 |
| Query | 4081 | TCCATGTTGTAAAACACACATCCCTAATGTGTGCCATAGAGTTAACACAAGTCCTGT | 4140 |
| Query | 4141 | GAATTTCTTCACTGTTGAAAATTATTTAAACAAAATAGAAGCTGTAGTAGGCCCTTCTG | 4200 |
| Query | 4201 | TGTGCACCTTACCAACTTCTGTAAACTCAAAACTAACATATTACTAACGCCACAAGAA | 4260 |
| Query | 4261 | ATTTGATTTCTATTCAAGGTGGCCAATTATTTGTGTAATAGAAAATGAAATCTAATA | 4320 |
| Query | 4321 | TTAAAAATATGGAACCTCTAatataattttatatttagttatagttcagatataatca | 4380 |
| Query | 4381 | tatTCGTATTCACTAATCTGGGAAGGGAAAGGGCTACTGCAGCTTACATGCAATTATTA | 4440 |
| Query | 4441 | AAATGATTGTAAAATAGCTGTATAGTGTAAAATAAGAATGATTTAGATGAGATTGTT | 4500 |
| Query | 4501 | TTATCATGACATGTTATATATTGTTGTAGGGCTAAAGAAAATGCTGATGGATAACCTAT | 4560 |
| Query | 4561 | ATGATTTATAGTTGTACATGCATTACAGGCAGCGATGGCTCAGAAACCAAACAGT | 4620 |
| Query | 4621 | TTGCTCTAGGGGAAGAGGGAGATGGAGACTGGCCTGTGCAAGGTTGCTGAGGC | 4680 |
| Query | 4681 | TCTGACCCAGTGAGATTACAGAGGAAGTTACCTCTGCCTCCATTCTGACCACCCCTCT | 4740 |
| Query | 4741 | CATTCCAACAGTGAGTCAGCAGCGAGTTAGTTACTCAATCTCCCCCTGCACAAACAA | 4800 |
| Query | 4801 | GTATGTAAAGTATGTAACACAGGAGACAGGAAGGTGGTGCTTACATCCTAAAGGCACCAT | 4860 |
| Query | 4861 | CTAATAGCGGGTTACTTCACATACAGCCCTCCCCCAGCAGTTGAATGACAACAGAACGCT | 4920 |
| Query | 4921 | TCAGAAGTTGGCAATAGTTGCATAGAGGTACCAAGCAATATGTAATAGTCAGAACATCT | 4980 |
| Query | 4981 | CATAGGTTGCCAATAACACTAATTCCCTTCTATCCTACAAACAAGAGTTATTCCAAA | 5040 |
| Query | 5041 | TAAAATGAGGACATgttttggatgttttttttttttttttttttttttttttttttttttc | 5100 |
| Query | 5101 | agtatttggagaaaattatTTAAaaaaaaaCAATCATTGCTTTGAATGCTCTCTAA | 5160 |
| Query | 5161 | AAGGGAAATGTAATATTTAAGATGGTGTGTAACCCGGCTGGATAAATTGGTGCCTAA | 5220 |
| Query | 5221 | GAAAAGCTGCTGAATATTCTTATCAATGACAGTGTAAAGTTCAAAAAGAGCTCTAAA | 5280 |
| Query | 5281 | CGTAGATTATCATTCCCTTATAGAAATGTTATGTGGTAAACCAAGAACGACATCTCAC | 5340 |
| 18665 | 2674 | AATGTTATGTGGTTAAACCAAGAACGACATCTCAC | 2710 |
| Query | 5341 | CATTAATCTGATTTCATCCCAACAACTTGGCGCTCAAAAATAGAACCTCAATGAGAAA | 5400 |
| 18665 | 2711 | CATTAATCTGATTTCATCCCAACAACTTGGCGCTCAAAAATAGAACCTCAATGAGAAA | 2770 |
| Query | 5401 | AAGAAGATTATGTGCACCTCGTTGTCAATAATAAGTCACAGTGTGCTCATCGACAACTAT | 5460 |
| 18665 | 2771 | AAGAAGATTATGTGCACCTCGTTGTCAATAATAAGTCACAGTGTGCTCATCGACAACTAT | 2830 |
| Query | 5461 | ACGAGCCTTTTCAATTAAATGGGAAAAGAAGCTGTGCCCTTTAGGATACGTGGGGAAA | 5520 |
| 18665 | 2831 | AGGAGGCTTTCAATTAAATGGGAAAAGAAGCTGTGCCCTTTAGGATACGTGGGGAAA | 2890 |
| Query | 5521 | GAAAGTCATCTTAATTATGTTAATGTGGATTAAAGTGCATATGGTGGTGTGTTGA | 5580 |

FIG. 7D

| | | | |
|-------------|------|--|------|
| 18665 | 2891 | GAAAGTCATCTAATTATGTTAATTGTGGATTAAAGTGCTATACTGGTGGTGCTGTTGA | 2950 |
| Query 18665 | 5581 | AAGCAGATTATTCCTATGTATGTGTTATCTGCCATCCAAACCAAACCTGTTGAAGTT | 5640 |
| 18665 | 2951 | AAGCAGATTATTCCTATGTATGTGTTATCTGCCATCCAAACCAAACCTGTTGAAGTT | 3010 |
| Query 18665 | 5641 | TGTAGTAACCTCAGTGAGAGTTGGTTACTCACAACAAATCCTGAAAAGTATTTTAGTGT | 5700 |
| 18665 | 3011 | TGTAGTAACCTCAGTGAGAGTTGGTTACTCACAACAAATCCTGAAAAGTATTTTAGTGT | 3070 |
| Query 18665 | 5701 | TTGTAGGTATTCTGTGGGATACTATAAACAGCAGAACTGAGGCACCTAGGACATAACACTT | 5760 |
| 18665 | 3071 | TTGTAGGTATTCTGTGGGATACTATAAACAGCAGAACTGAGGCACCTAGGACATAACACTT | 3130 |
| Query 18665 | 5761 | TTGGGGTATATATATCCTAAATGCCCTAAACTATGGGAGGAAACCTTGGCACCCCCAAAAG | 5820 |
| 18665 | 3131 | TTGGGGTATATATATCCTAAATGCCCTAAACTATGGGAGGAAACCTTGGCACCCCCAAAAG | 3190 |
| Query 18665 | 5821 | GAANAACTAACATGATTGTGCTATGAAAGTGCTGGATAATTAGCATGGATGAGCTCTGG | 5880 |
| 18665 | 3191 | GAANAACTAACATGATTGTGCTATGAAAGTGCTGGATAATTAGCATGGATGAGCTCTGG | 3250 |
| Query 18665 | 5881 | GCATGCCATGAAGGAAAGCCACGCTCCCTCAGAATTCAAGAGGCAGGGAGCAATTCCAGT | 5940 |
| 18665 | 3251 | GCATGCCATGAAGGAAAGCCACGCTCCCTCAGAATTCAAGAGGCAGGGAGCAATTCCAGT | 3310 |
| Query 18665 | 5941 | TTCACCTAACGCTCTATAATTAGTCCCTTTAAAAACCTGAAAACATCACATCACCATG | 6000 |
| 18665 | 3311 | TTCACCTAACGCTCTATAATTAGTCCCTTTAAAAACCTGAAAACATCACATCACCATG | 3370 |
| Query 18665 | 6001 | GAATGAAAAATATTGTTATACAATACATTGATCTGCTAAACCTCCAGAACATGGTAGCC | 6060 |
| 18665 | 3371 | GAATGAAAAATATTGTTATACAATACATTGATCTGCTAAACCTCCAGAACATGGTAGCC | 3430 |
| Query 18665 | 6061 | TTCAGTGAGATTCCATCTGGCTGGTCACTCCCTGACTGTAGCTGTAGGTGAAtgtgtt | 6120 |
| 18665 | 3431 | TTCAGTGAGATTCCATCTGGCTGGTCACTCCCTGACTGTAGCTGTAGGTGAAtgtgtt | 3490 |
| Query 18665 | 6121 | ttttgtgtgtgtgtgtCTGGTTTTAGTGTCAAGAGGGAAATAAAAGTGTAAAGGAGGACACT | 6180 |
| 18665 | 3491 | TTTGTGTGTGTGTCTGGTTTTAGTGTCAAGAGGGAAATAAAAGTGTAAAGGAGGACACT | 3550 |
| Query 18665 | 6181 | TTAAACCCTTGGGTGGAGTTCGTAATTCCCAGACTATTCAAGCAACCTGGTCCAC | 6240 |
| 18665 | 3551 | TTAAACCCTTGGGTGGAGTTCGTAATTCCCAGACTATTCAAGCAACCTGGTCCAC | 3610 |
| Query 18665 | 6241 | CCAGGATTAGTGACCAGGTTTCAGGAAAGGATTGCTCTCTCTAGAAAATGTCTGAAA | 6300 |
| 18665 | 3611 | CCAGGATTAGTGACCAGGTTTCAGGAAAGGATTGCTCTCTCTAGAAAATGTCTGAAA | 3670 |
| Query 18665 | 6301 | GGATTTTATTTCTGATGAAAGGCTGTATGAAAATACCCCTCTCAAATAACTGCTTAAC | 6360 |
| 18665 | 3671 | GGATTTTATTTCTGATGAAAGGCTGTATGAAAATACCCCTCTCAAATAACTGCTTAAC | 3730 |
| Query 18665 | 6361 | TACATATAGATTCAAGTGTCATATTCTATTGTTATTAATGCTATATAATGGGG | 6420 |
| 18665 | 3731 | TACATATAGATTCAAGTGTCATATTCTATTGTTATTAATGCTATATAATGGGG | 3790 |
| Query 18665 | 6421 | ACAAATCTATATTATACTGTGTATGGCATTATTAAAGAAGCTTTTCATTAATTTTATCA | 6480 |
| 18665 | 3791 | ACAAATCTATATTATACTGTGTATGGCATTATTAAAGAAGCTTTTCATTAATTTTATCA | 3850 |
| Query 18665 | 6481 | CAGTAATTAAAATGTGTAAAAAATAAAACCAGTGACTCCGTAAACCTGGTAAAGTTG | 6540 |
| 18665 | 3851 | CAGTAATTAAAATGTGTAAAAAATAAAACCAGTGACTCCGTAAACCTGGTAAAGTTG | 3910 |
| Query 18665 | 6541 | TAGTTTTTATTCATGCTGAATAATAATCTGTAGTTaaaaaaaaaaGTGTCTTTTACCTA | 6600 |
| 18665 | 3911 | TAGTTTTTATTCATGCTGAATAATAATCTGTAGTTaaaaaaaaaaGTGTCTTTTACCTA | 3970 |
| Query 18665 | 6601 | CGCAGTGAAATGTCAGACTGTAAAACCTTGTGGAAATGT"TTAACCTTATTTTACAT | 6660 |
| 18665 | 3971 | CGCAGTGAAATGTCAGACTGTAAAACCTTGTGGAAATGTAACTTTATTTTACAT | 4030 |
| Query 18665 | 6661 | TTAAATTGCTGTTCTGGTATTACCAAACCACACATTGTACCGAATTGGCAGTAAATGT | 6720 |
| 18665 | 4031 | TTAAATTGCTGTTCTGGTATTACCAAACCACACATTGTACCGAATTGGCAGTAAATGT | 4090 |
| Query 18665 | 6721 | TAGGCCATTACAGCAAATGCCAAATATGGGAGAACATGCAATTaaaaaaaCTGCTTTTC | 6780 |
| 18665 | 4091 | TAGGCCATTACAGCAAATGCCAAATATGGGAGAACATGCAATTaaaaaaaCTGCTTTTC | 4150 |
| Query | 6781 | ATTA 6784 | |

18665 4151 ATTA 4154

FIG. 7F

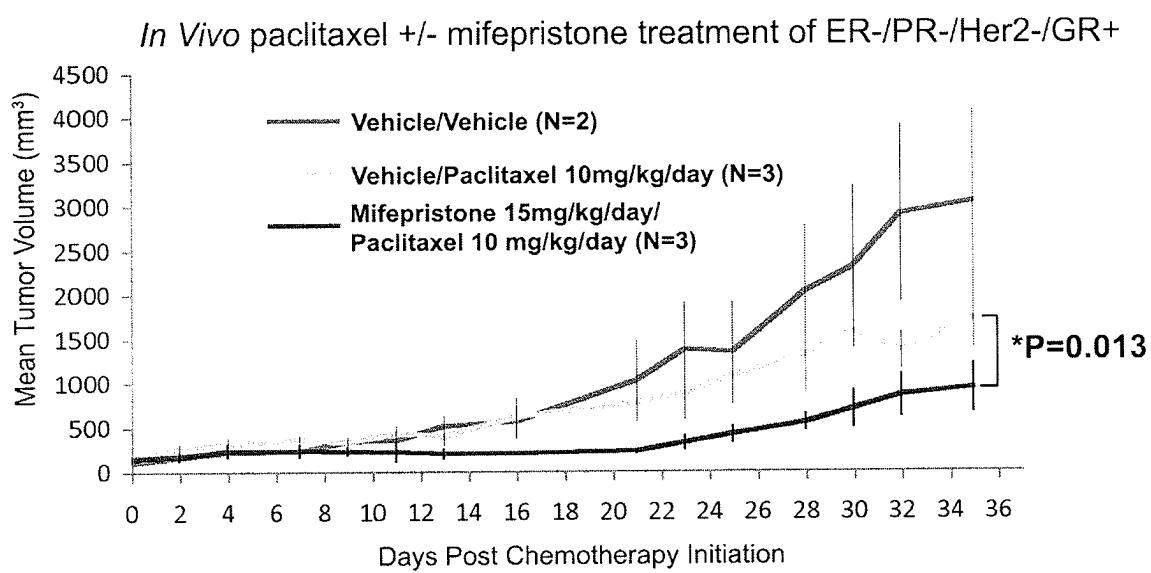


FIG. 8

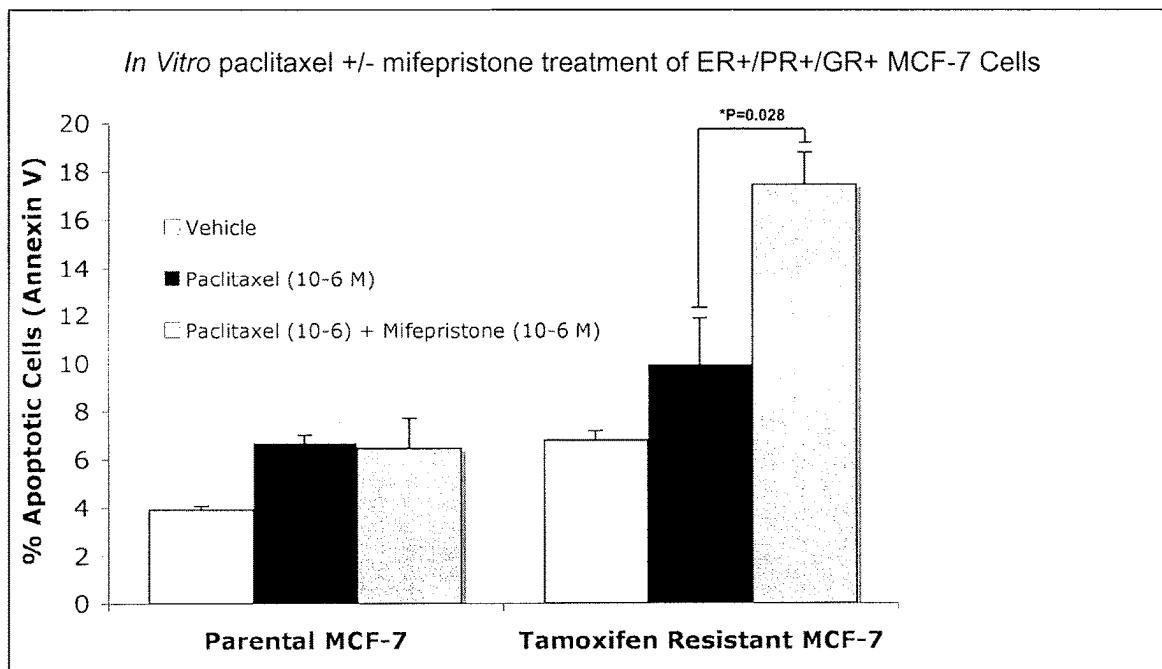


FIG. 9

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. 18/541,675, filed Dec. 15, 2023, which is a Continuation of U.S. application Ser. No. 17/577,882, filed Jan. 18, 2022, (now U.S. Pat. No. 11,883,417, issued Jan. 30, 2024), which is a Continuation of U.S. application Ser. No. 16/596,342, filed Oct. 8, 2019 (now U.S. Pat. No. 11,234,990, issued Feb. 1, 2022), 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 XML FILE

[0003] The Sequence Listing written in file "096487-1417565-000180US", created on Dec. 12, 2023, is 209,468 bytes in size and, 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 glucocor-

ticoid 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, NRP1, ACSL5, BIRC3, NNMT, IGFBP6, PLXNC1, SLC46A3, C14orf139, PIAS1, IDH2, SERPINF1, ERBB2, PECAM1, LBH, ST3GAL5, ILIR1, BIN1, WIFP1, TFPI, FN1, FAM134A, NRIP1, 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 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 practitioner 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 ESRI 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 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, beclometasone, 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 PR 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 (Cytoxan), 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, 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 a 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, including 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%) 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 30m-24h 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 NR3C vs. lower expres-

sion, respectively. For ESRI+ tumors (orange) the P< 0.0001 and for ESRI-tumors (green) p= .7 (t test). Error bars are+/- SD.

[0077] FIG. 4. RFS of GR gene expression signature. The GR signature predicts a differential prognosis for ESRI+ patients and ESRI- pts with respect to GR-signature expression. ESRI-/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 ESRI- 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 steroid 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, osteoar-

thritis, 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 mecha-

nism(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 heterogenous 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, NRP1, ACSL5, BIRC3, NNMT, IGFBP6, PLXNC1, SLC46A3, C14orf139, PIAS1, IDH2, SERPINF1, ERBB2, PECAM1, LBH, ST3GAL5, ILIR1, BIN1, WIPF1, TFPI, FN1, FAM134A, NRIP1, 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.

A. Nucleic Acids

[0114] 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.

[0115] 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.

[0116] 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.

[0117] 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.

[0118] 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.

[0119] 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 complementary 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 one 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.

[0120] 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.

[0121] 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.

[0122] 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.

[0123] 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.

[0124] 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.

[0125] 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.

[0126] 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.

[0127] 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.

[0128] 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.

[0129] 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.

[0130] 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.

[0131] 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.

[0132] 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.

[0133] 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.

[0134] 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.

[0135] 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.).

B. Proteins and Polypeptides

[0136] 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.

[0137] 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.

[0138] 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.

[0139] 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.

[0140] 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.

[0141] 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.

[0142] 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.

[0143] 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).

[0144] 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-ethyl)-benzhiazoline-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.

[0145] Another suitable method is RIA (radioimmunoassay). An example of RIA 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, I^{125} . In one embodiment, a fixed concentration of I^{125} -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 I^{125} -polypeptide that binds to the antibody is decreased. A standard curve can therefore be constructed to represent the amount of anti-

body-bound I^{125} -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.

[0146] Suitable antibodies for this invention include, but are not limited to, polyclonal antibodies, monoclonal antibodies, chimeric antibodies, humanized antibodies, single chain antibodies, Fab fragments, and fragments produced by a Fab expression library.

[0147] 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.

[0148] 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.

[0149] 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.

[0150] 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).

[0151] 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 inves-

tigating 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.

[0152] 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

[0153] 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).

[0154] 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.

[0155] 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.

[0156] 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).

[0157] 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.

[0158] 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 steroid. 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 an-

lines, 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 steroid 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 8a-benzyl isoquinolones, N-substituted derivatives, bridgehead alcohol and ethers, bridgehead amines). Additional specific examples include, but are not limited to the following specific antagonists: beclometasone, 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.

[0159] 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, nitrosourea, dacatinomycin, daunorubicin, doxorubicin, bleomycin, plicomycin, mitomycin, etoposide (VP16), tamoxifen, raloxifene, estrogen receptor binding agents, taxol, gemcitabine, navelbine, farnesyl-protein transferase inhibitors, transplatinium, 5-fluorouracil, vincristin, vinblastin and methotrexate, or any analog or derivative variant of the foregoing.

[0160] 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.

[0161] Certain chemotherapeutics are well known for use against breast cancer. These breast cancer chemotherapeutics are capecitabine, carboplatin, cyclophosphamide (Cytoxan), 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.

[0162] 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, piposulfan, nab-5404, nab-5800, nab-5801, Irinotecan, HKP, Orataxel, gemcitabine, Herceptin®, vinorelbine, Doxil®, capecitabine, Gleevec®, Alimta®, Avastin®, Velcade®, Tarceva®, Neulasta®, Lapatinib, STI-571, ZD1839, Iressa® (gefitinib), SH268, genistein, CEP2563, SU6668, SU11248, EMD121974, and Sorafenib.

[0163] In some embodiments, the chemotherapeutic agent is a composition comprising nanoparticles comprising a thiocolchicine derivative and a carrier protein (such as albumin).

[0164] 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.

[0165] 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 "B":

[0166] 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.

[0167] 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/theronin 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.

[0168] 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-1 α), CCL5, TGF- β , beta-, lipoxygenase, 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-decatetraenoic acid, mono [3R,4S,5S,6R]-5-methoxy-4-[(2R,3R)-2-methyl-3-(3-methyl-2-but enyl) 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); Tramast, 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 REVIMID.

[0169] Radiation therapy that cause DNA damage and have been used extensively include what are commonly known as γ -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.

[0170] The terms "contacted" and "exposed," when applied to a cell, are used herein to describe the process by

| | | | | | | | |
|---------|---------|---------|---------|---------|---------|---------|---------|
| 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/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 | | |

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.

[0171] 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.

[0172] 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.

[0173] 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.

[0174] 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.

[0175] 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).

[0176] 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.

[0177] 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.

[0178] 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.

[0179] Particularly antibodies for use in the present invention include zalutumumab (2F8,), cetuximab (Erbitux), 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.

[0180] 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.

[0181] 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.

[0182] 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 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.

[0183] 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

[0184] 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.

[0185] 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.

[0186] 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

[0187] 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

A. Results

[0188] 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+ (ESRI-high) subset of patients suggests a slight better outcome, implying a crosstalk between the ER and the GR that is absent in ER-tumors.

[0189] 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, LAPTMS, NNMT, SERPINF1, NRIP1, WIF1, BIN1, IL1R1, ST3GAL5, SEMA4D, MAP3K5, SMARCA2, DPT, BIRC3, PTGDS, PHF15, MAOA, TFP1, 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 | |

B. Materials and Methods

[0190] Cell culture and glucocorticoid treatment: 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^{-6} M) and ethanol of the same volume as a control.

[0191] Microarray gene expression: MCF10A-Myc Cells: Time course (0.5h, 2 h, 4 h and 24h) 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.

[0192] GR ChIP-Seq experiment and analysis for MCF10A-Myc Cells: Cells were collected for the ChIP

assay following 1 hour of Dex (10^{-6} M) 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 Maq 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.

[0193] Human Primary Breast Cancer Analysis: 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 ESRI and NR3C1 positivity: Expression data of tumors with known ER IHC status were analyzed using ROC analysis. The Youden Index of the best ESRI probe's ROC curve was used as the cut-off point to separate ESRI+ and ESRI-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: Un-supervised clustering was performed by Cluster using Pearson 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.

[0194] Tumor assessment. 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 ESRI 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 ESRI 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 $n=227 \leq -1.257434$. (ESR1-) or i.e. 77.3% quantile

[0195] This cut-off was applied to the entire dataset, n=898 (ESR+), n=308 (ESR-). In addition to the method, the ACTUAL Log2 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.

[0196] The ESR1 probe ID from Affymetrix is 205225_at.

[0197] The NR3C1 probe ID from Affymetrix is 216321_s_at

[0198] 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.

[0199] 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.

[0200] 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

[0201] Xenografted ER-/PR-/HER2-(GR+) MDA-MB-231 human breast cancer cells (1×10^7 cells in 50 μ l 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=S2×L×0.52. 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.

[0202] 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

```

TTTTAGAAAAAAAAATATTTCCCTCCTGCTCCTCTGCAGTCACAAGCTAAGTTGTTATCTCGGC
TGCGGGGGAACTGCCGACGGTGGCGGGCGAGCGGCTCCTCTGCCAGAGTTGATATTCACTGATGGACTC
CAAAGAACATTAACCTCCTGGTAGAGAACAAAAACCCAGCAGTGTGCTGCTCAGGAGAGGGAGATGTG
ATGGACTTCTATAAAACCTAAGAGGAGGAGCTACTGTGAAGGTTCTCGCTCTCACCTCACTGGCTG
TCGCTCTCAATCAGACTCCAAGCAGCGAAGACTTTGGTTGATTTCCAAAAGGCTCAGTAAGCAATGC
GCAGCAGCCAGATCTGCTCAAAGCAGTTCACTCTCAATGGACTGTATATGGGAGAGACAGAAAACAAA
GTGATGGAAATGACCTGGGATTCACAGCAGGGCAAATCAGCCTTCCTGGGGAAACAGACTTAA
AGCTTTGGAAGAACGATTGCAAACCTCAATAGGTCGACCAGTGTCCAGAGAACCCAAAGAGTTCA
ATCCTACTGCTGTCTGCTGCCCAACAGAGAAGGAGTTCCAAAACACTCTGATGTATCTCAGAA
CAGCAACATTTGAAGGCCAGACTGGCACCAACGGTGGCAATGTGAAATTGTATACCACAGACCAAAGCA
CCTTGTACATTTGCAAGGATTGGAGTTCTCTGGTCCCCAGGTAAGAGACGAATGAGAGTCCTT
GAGATCAGACCTGTTGATAGATGAAAAGCTGTTGCTTCTCCTCTGGGGAGAAGACGATTCTT
TTGGAAGGAAACTCGAATGAGGACTGCAAGCCTCTATTTACCGACACTAAACCCAAAATTAAGGATA
ATGGAGATCTGGTTTGTCAGCCCCAGTAATGTAACACTGCCCAAGTGAACAGAAAAGAGAAGATT
CATCGAACTCTGCACCCCTGGGTATTAAGCAAGAGAAACTGGCACAGTTACTGTCAGGCAAGCTT

```

- continued

CCTGGAGCAAATATAATTGTAATAAAATGTCGCCATTCTGTTCATGGTGTGAGTACCTCTGGAGGAC
AGATGTACCACTATGACATGAATACAGCATCCCTTCTAACAGCAGGATCAGAAGCCTATTAAATGT
CATTCCACCAATTCCCGTTGGTCCGAAAATTGGAATAGGTGCAAGGATCTGGAGATGACAACCTGACT
TCTCTGGGGACTCTGAACCTCCCTGGTCGAAAGTTTCTAATGGTATTCAAGCCCCAGCATGAGAC
CAGATGTAAGCTCCTCCATCAGCTCTAACAGCAACACAGGACCCTCCAAACTCTGCCTGGT
GTGCTCTGATGAAGCTTCAGGATGTCATTATGGAGTCTTAACCTGTGAAAGCTGTAAAGTTCTCAA
AGAGCAGTGGAGGGACAGCACAATTACCTATGCTGGAGGAAGTATTGATCATCGATAAAATCGAA
GAAAAAAACTGCCAGCATGCCGTATCGAAAATGCTTCAGGCTGGAATGAACCTGGAAGCTCGAAAAAC
AAAGAAAAAAATAAAGGAATTCAAGCAGGCAACTACAGGAGTCTCACAAAGAAACCTCTGAAAATCTGGT
AACAAAACAATAGTCCTGCAACGTTACCAACTCACCCCTACCCGGTGTACTGTTGGAGGTTATTG
AACCTGAAGTGTATATGCAGGATATGATAGCTCTGTCAGACTCAACTGGAGGATCATGACTACGCT
CAACATGTTAGGAGGGCGCAAGTGATTGAGCAGTGAAATGGCAAGGCAATACAGGTTTCAGGAAC
TTACACCTGGATGACCAATGACCCACTGCTGACTCTGGATTTCTATGGCATTGCTCTGGGT
GGAGATCATATAGACAATCAAGTGCAACCTGCTGTGTTTGTCTGATCTGATTATAATGAGCAGAG
AATGACTCTACCCCTGCATGTACGACCAATGAAACACATGCTGTATGTTCTCTGAGTTACAGGCTT
CAGGTATCTTATGAAGAGTATCTGTATGAAAACCTTAUTGCTCTCTTCAGTTCTAACGGACGGTC
TGAAGAGCCAAGAGCTTTGATGAAATTAGAATGACCTACATCAAAGAGCTAGGAAAGCCATTGCAA
GAGGGAAAGGAAACTCCAGCCAGAACTGGCAGCGGTTTATCAACTGACAAAATCTGGATTCTATGCAT
GAAGTGGTTGAAAATCTCTTAUTTGCTTCAACTATTGCTTCAAACATTGGATAAGACCATGAGTATTGAAATTCC
CCGAGATGTTAGCTGAAATCATCACCAATCAGATAACCAAAATATTCAAATGGAAATATCAAAACCTT
GTTTCATCAAAGTGAUTGCTGCTTAATAAGAATGGTGCCTAAAGAAAGTCGAATTAATGCTTTATTG
TATAAAACTATCAGTTGCTCTGTAGAGGTTTGTGTTTATTGTTCTGATCTGTTGTTG
TTTAAATACGCACATGTTGTTTAGAGGGCCAAGACTTGGCAACAGCAGTTGAGTCGTGATCA
CTTTCACTGATGGAGAGTAGTGGTAAATTAGTTAATATCCCAGAAATTAGAAACCTTAAT
ATGTGGACGTAATCTCCACAGTCAGGAAAGGATGGCACCTAACACCAGTGGCCAAAGTCTGTGATG
AACTTTCTCTCATACTTTTCACAGTTGGCTGGATGAAATTCTAGACTTCTGTTGGTGTATCCC
CCCCCTGTATAGTTAGGATAGCATTGATGCTTAACTGAAACCTGAAAAAGTTACAAGTGTAT
ATCAGAAAAGGAAGGTTGCTGCTTATAGCTTACATAATTGTTATTCAAGTATTGATGCTTAAAGATGGC
GAACTACGCTTGTCTATTCTTACATAATTGTTATTCAAGTATTGATGCTTAAAGATGGC
AGCTAGTCGTAGCTTCCAAATAAAACTCTAAACATTAATCAATCATCTGTGAAAATGGTTGGTGC
TTCTAACCTGATGGCACTTAGTCAAGGACCAACAAAATTGACTCAAATCTCCAGTATTCTGTCAA
AAAAAAAAAAAAAGCTCATTTGTATATCTGCTTCAGTGGAGAATTATAGGTTGTGCAAAT
TAAACAGTCTTAUTGGTATAGGCACTTAGTCAGTGGCTGGTAAACTGTGGATGATGGTTGCAA
AAAGACTAATTAAAAAAACTACCAAGAGGCCCTGTCTGACCTAACGCCATTGGCAATGGCTAT
ATGGCAAGAAAGCTGGTAAACTATTGCTTCAAGGACCTTGTGAAAGTAGTTGTATAACTCTTAAAG
TTGTGATTCCAGATAACCAAGCTGTAACACAGCTGAGAGACTTTAATCAGACAAAGTAATTCTCTCACT
AAACTTTACCCAAAATCAAATCTCTAATATGGCAAAATGGCTAGACACCCATTTCACATCCCAC
GTCACCAATTGGTAACTTCTGATGGTACAGGAAAGCTCAGCTACTGATTGTGATTAGAACTG
TATGATGTCAGACATCCATGTTGAAAACACACATCCCTAATGTCAGGTTAACACAAGT

- continued

```

CCTGTGAATTCTTCACTGTTGAAAATTATTTAAACAAATAGAAGCTGTAGTAGCCCTTCTGTGTGC
ACCTTACCAACTTCTGTAAACTCAAACCTAACATATTACTAACAGCCACAAGAAATTGATTCTATT
AAGGTGGCCAATTATTTGTGTAAATAGAAAATCTAAGAAATCTAATATTAAAAATATGAACTTCTAAATATA
TTTTATATTAGTTAGTTAGTTCAAGATATATCATATTGGTATTCACTAATCTGGGAAGGGAAAGGGCTA
CTGCAGCTTACATGCAATTATTAAATGATTGTAAAATAGCTTGTATAGTGTAAAATAAGAATGATT
TTAGATGAGATTGTTTATCATGACATGTTATATTGGTAGGGTCAAAGAAATGCTGATGGATAA
CCTATATGATTATAGTTGTACATGCATTACAGGCAGCGATGGTCTCAGAAACCAAACAGTTGCT
CTAGGGGAAGAGGGAGATGGAGACTGGTCCTGTGCAGTGAGGTTGCTGAGGCTCTGACCCAGTGAGA
TTACAGAGGAAGTTATCCTCTGCCCTTCTGACCACCCCTCTCATTCCAAACAGTGAGTCAGCG
CAGGTTAGTTACTCAATCTCCCTGCACTAAAGTATGTAAGTATGTAACAGGAGACAGGAAGGTG
GTGCTTACATCCTTAAAGGCACCATCTAATAGCGGGTACTTCACATACAGCCCTCCCCAGCAGTTGA
ATGACAACAGAACGCTTCAGAAGTTGCAATAGTTGCATAGAGGTACCGCAATATGTAATAGTGCAG
AATCTCATAGGGTGCAATAATACACTAATTCTTCTATCCTACAAACAAGAGTTATTTCAAATAAAA
TGAGGACATGTTTGTGTTCTTGAATGCTTTGAATGTTATTGTTAGTATTGGAGAAA
TTATTTAATAAAAAAAACAATCATTGCTTTG

```

ESR1 mRNA

SEQ ID NO: 2

```

AGGAGCTGGC GGAGGGCGTT CGTCCTGGGA CTGCACTTGC TCCCCTGGG TCGCCCGGCT
TCACCGGACC CGCAGGCTCC CGGGGCAGGG CCGGGGCCAG AGCTCGCGTG TCGGCGGGAC
ATGCGCTGCG TCGCCTCTAA CCTCGGGCTG TGCTTTTT CCAGGTGGCC CGCCGGTTTC
TGAGCCTCT GCCTGCGGG GACACGGTCT GCACCCCTGCC CGCGGCCACG GACCATGACC
ATGACCCCTCC ACACCAAAGC ATCTGGGATG GCCCTACTGC ATCAGATCCA AGGGAACGAG
CTGGAGCCCC TGAACCGTCC GCAGCTCAAG ATCCCCCTGG AGCGGCCCT GGGCGAGGTG
TACCTGGACA GCAGCAAGCC CGCCGTGTAC AACTACCCCC AGGGGCCCGC CTACGAGTTTC
AACGCCGCGG CGGCCGCCAA CGCGCAGGTC TACGGTCAGA CGGCCCTCCC CTACGCC
GGGTCTGAGG CTGCGCGTT CGGCTCCAAC GGCCTGGGGG GTTCCCCCCC ACTCAACAGC
GTGTCTCCGA GCCCGCTGAT GCTACTGCAC CGCCCGCCGC AGCTGTGCC TTTCTGCAG
CCCCACGGCC AGCAGGTGCC CTACTACCTG GAGAACGAGC CGAGCGCTA CACGGTGC
GAGGCCGCC CGCCGGCATT CTACAGGCC AATTCAAGATA ATCGACGCCA GGGTGGCAGA
GAAAAGATTGG CCAGTACCAA TGACAAGGG AGTATGGCTA TGGAATCTGC CAAGGAGACT
CGCTACTGTG CAGTGTGCAA TGACTATGCT TCAGGCTACC ATTATGGAGT CTGGCTCTGT
GAGGGCTGCA AGGCCTCTT CAAGAGAAAT ATTCAAGGAC ATAACGACTA TATGTGTCCA
GCCACCAACC AGTGCACCAT TGATAAAAAC AGGAGGAAGA GCTGCCAGGC CTGCCGCTC
CGCAAATGCT ACGAAGTGGG AATGATGAAA GGTGGGATAC GAAAAGACCG AAGAGGAGGG
AGAATGTTGA AACACAAGCG CCAGAGAGAT GATGGGGAGG GCAGGGGTGA AGTGGGGTCT
GCTGGAGACA TGAGAGCTGC CAACCTTGG CCAAGCCCGC TCATGATCAA ACGCTCTAAG
AAGAACAGCC TGGCCTTGTC CCTGACGGCC GACCAGATGG TCAGTGCCTT GTGGATGCT
GAGCCCCCA TACTCTATTC CGAGTATGAT CCTACCAGAC CCTTCAGTGA AGCTTCGATG
ATGGGCTTAC TGACCAACCT GGCAGACAGG GAGCTGGTTC ACATGATCAA CTGGCGAAG
AGGGTGCCAG GCTTGTGGA TTTGACCCCTC CATGATCAGG TCCACCTTCT AGAATGTGCC

```

-continued

TGGCTAGAGA TCCTGATGAT TGGTCTCGTC TGGCGCTCCA TGGAGCACCC AGGGAAGCTA
CTGTTTGCTC CTAACATTGCT CTTGGACAGG AACCAAGGAA AATGTGTAGA GGGCATGGTG
GAGATCTTCG ACATGCTGCT GGCTACATCA TCTCGGTTCC GCATGATGAA TCTGCAGGGA
GAGGAGTTG TGTGCCCTCAA ATCTATTATT TTGCTTAATT CTGGAGTGTA CACATTTCTG
TCCAGCACCC TGAAGTCTCT GGAAGAGAAAG GACCATATCC ACCGAGTCCT GGACAAGATC
ACAGACACTT TGATCCACCT GATGGCCAAG GCAGGCCTGA CCCTGCAGCA GCAGCACCAAG
CGGCTGGCCC AGCTCCCTCAT CCTCTCTCC CACATCAGGC ACATGAGTAA CAAAGGCATG
GAGCATCTGT ACAGCATGAA GTGCAAGAAC GTGGTGCCCC TCTATGACCT GCTGCTGGAG
ATGCTGGACG CCCACCGCCT ACATGCGCCC ACTAGCCGTG GAGGGGCATC CGTGGAGGAG
ACGGACCAAA GCCACTTGGC CACTGCGGGC TCTACTTCAT CGCATTCCCTT GCAAAAGTAT
TACATCACGG GGGAGGCAGA GGGTTCCCT GCCACGGTCT GAGAGCTCCC TGGCTCCAC
ACGGTTCAAGA TAATCCCTGC TGCACTTTAC CCTCATCATG CACCACTTTA GCCAAATTCT
GTCTCCTGC A TACACTCCGG CATGCATCCA ACACCAATGG CTTCTAGAT GAGTGGCCAT
TCATTTGCTT GCTCAGTTCT TAGTGGCACA TCTTCTGTCT TCTGTTGGGA ACAGCCAAAG
GGATTCCAAG GCTAAATCTT TGTAACAGCT CTCTTCCCC CTTGCTATGT TACTAACCGT
GAGGATTCCC GTAGCTCTTC ACAGCTGAAC TCAGTCTATG GGTTGGGCT CAGATAACTC
TGTGCATTTA AGCTACTTGT AGAGACCCAG GCCTGGAGAG TAGACATTG GCCTCTGATA
AGCACTTTT AAATGGCTCT AAGAATAAGC CACAGCAAAG AATTAAAGT GGCTCCTTTA
ATTGGTACT TGGAGAAAGC TAGGTCAAGG GTTTATTATA GCACCCCTTT GTATTCCAT
GGCAATGCAT CCTTTATGA AAGTGGTACA CCTTAAAGCT TTTATATGAC TGTAGCAGAG
TATCTGGTGA TTGTCATTC ATTCCCCCTA TAGGAATACA AGGGGCACAC AGGGAAGGCA
GATCCCCTAG TTGGCAAGAC TATTTTAACT TGATACACTG CAGATTCAAGA TGTGCTGAAA
GCTCTGCCTC TGGCTTCCCG GTCACTGGTT CCAGTTAATT CATGCCTCCC ATGGACCTAT
GGAGAGCAGC AAGTTGATCT TAGTTAAGTC TCCCTATATG AGGGATAAGT TCCTGATTT
TGTGTTTATT TTTGTGTTAC AAAAGAAAGC CCTCCCTCCC TGAACCTGCA GTAAGGTAG
CTTCAGGACC TGTCCAGTG GGCACGTAC TTGGATCTTC CCGCGTGTG TGTGCCTTAC
ACAGGGGTGA ACTGTTCACT GTGGTGTGCA ATGATGAGGG TAAATGGTAG TTGAAAGGAG
CAGGGGCCTT GGTGTTGCAT TTAGCCCTGG GGCATGGAGC TGAACAGTAC TTGTGCAGGA
TTGTTGTTGC TACTAGAGAA CAAGAGGGAA AGTAGGGCAG AACTGGATA CAGTTCTGAG
GCACAGCCAG ACTTGCTCAG GGTGGCCCTG CCACAGGCTG CAGCTACCTA GGAACATTCC
TTGCAGACCC CGCATTGCC TTTGGGGGTG CCCTGGGATC CCTGGGGTAG TCCAGCTCTT
CTTCATTCTC CAGCGTGGCC CTGGTTGGAA GAAGCAGCTG TCACAGCTGC TGTAGACAGC
TGTGTTCTA CAATTGGCCC AGCACCCCTGG GGCACGGGG AAGGGTGGGG ACCGTTGCTG
TCACTACTCA GGCTGACTGG GGCCTGGTCA GATTACGTAT GCCCTTGGTG GTTGTAGAGAT
AAATCCAAAAT CAGGGTTGG TTTGGGAAG AAAATCCTCC CCCTTCCCTCC CCCGCCCCGT
TCCCTACCGC CTCCACTCCT GCCAGCTCAT TTCCCTCAAT TTCCCTTGAC CTATAGGCTA
AAAAAGAAAG GCTCATTCCA GCCACAGGGC AGCCTTCCCT GGGCCTTGC TTCTCTAGCA
CAATTATGGG TTACTTCCCTT TTCTTAACA AAAAGAATG TTGATTTC TCTGGGTGAC
CTTATTGTCT GTAATTGAAA CCCTATTGAG AGGTGATGTC TGTGTTAGCC AATGACCCAG

-continued

GTGAGCTGCT CGGGCTTCTC TTGGTATGTC TTGTTGGAA AAGTGGATT CATTCAATTG
TGATTGTCCA GTTAAGTGAT CACCAAAGGA CTGAGAATCT GGGAGGGCAA AAAAAAAA
AAAGTTTTA TGTGCACTTA AATTGGGG CAATTTATG TATCTGTGTT AAGGATATGT
TTAAGAACAT AATTCTTTG TTGCTGTTG TTTAAGAAC ACCTTAGTT GTTTAAGAAG
CACCTTATAT AGTATAATAT ATATTTTTT GAAATTACAT TGCTGTTA TCAGACAATT
GAATGTAGTA ATTCTGTTCT GGATTTAATT TGACTGGTT AACATGCAA AACCAAGGAA
AAATATTAG TTTTTTTT TTTTTGTA TACTTTCAA GCTACCTGT CATGTATA
GTCATTTATG CCTAAAGCCT GGTGATTATT CATTAAATG AAGATCACAT TTCATATCAA
CTTTGTATC CACAGTAGAC AAAATAGCAC TAATCCAGAT GCCTATTGTT GGATACTGAA
TGACAGACAA TCTTATGTAG CAAAGATTAT GCCTGAAAAG GAAAATTATT CAGGGCAGCT
AATTTGCTT TTACAAAAT ATCAGTAGTA ATATTTTG ACAGTAGCTA ATGGGTCA
GGGTTCTTT TAATGTTAT ACTTAGATT TCTTTAAA AAATAAAAT AAAACAAAAA
AAAATTCATA GGACTAGACG ATGTAATACC AGCTAAAGCC AAACAATTAT ACAGTGGAA
GTTTACATT ATTCACTCAA TGTGTTCTA TTCATGTTAA GATACTACTA CATTGA
GGCAGAGAA CATCAGATGA TTGAAATGTT CGCCCAGGGG TCTCCAGCAA CTTTGGAA
CTCTTGTAT TTTTACTTGA AGTGCCTA ATGGACAGCA GATATTTCT GGCTGATGTT
GGTATTGGT GTAGGAACAT GATTTAAAAA AAAACTCTTG CCTCTGCTT CCCCCACTCT
GAGGCAAGTT AAAATGTAAA AGATGTGATT TATCTGGGG GCTCAGGTAT GGTGGGAAG
TGGATTCAAGG AATCTGGGA ATGGCAAATA TATTAAGAAG AGTATTGAAA GTATTGGAG
GAAAATGGTT AATCTGGGT GTGCACCAGG GTTCAGTAGA GTCCACTTCT GCCCTGGAGA
CCACAAATCA ACTAGCTCA TTTACAGCCA TTTCTAAAAT GGCAGCTTCA GTTCTAGAGA
AGAAAGAACAA ACATCAGCAG TAAAGTCCAT GGAATAGCTA GTGGTCTGTG TTTCTTTC
CCATTGCCA GCTTGCGTA ATGATTCTAT AATGCCATCA TGCAGCAATT ATGAGAGGCT
AGGTCACTCA AAGAGAACAC CCTATCAATG TAGGTTGCAA AATCTAACCC CTAAGGAAGT
GCAGTCTTG ATTTGATTTC CCTAGTAACC TTGCAGATAT GTTTAACCAA GCCATAGCCC
ATGCCTTTG AGGGCTGAAC AAATAAGGG AACTACTGATA ATTTACTTT GATCACATTA
AGGTGTTCTC ACCTTGAAAT CTTATACACT GAAATGGCCA TTGATTTAGG CCACGGCTT
AGAGTACTCC TTCCCTGCA TGACACTGAT TACAAATACT TTCTATTCA TACTTCCAA
TTATGAGATG GACTGTGGGT ACTGGGAGTG ATCACTAACCA CCATAGTAAT GTCTAATATT
CACAGGCAGA TCTGCTTGGG GAAGCTAGTT ATGTGAAAGG CAAATAGAGT CATACTAG
CTCAAAAGGC AACCATAAATT CTCTTGGTG CAGGTCTTG GAGCGTGATC TAGATTACAC
TGCACCATTC CCAAGTTAAT CCCCTGAAAA CTTACTCTCA ACTGGAGCAA ATGAACTTTG
GTCCAAATAA TCCATTTT CAGTAGCGTT AATTATGCTC TGTTTCAAC TGCATTTCT
TTCCAATTGA ATTAAAGTGT GGCCTCGTT TTAGTCATT AAAATTGTT TCTAAGTAAT
TGCTGCCTCT ATTATGGCAC TTCAATTG CACTGTCTT TGAGATTCAA GAAAATTTC
TATTCTTTT TTTGCATCCA ATTGTGCCTG AACTTTAAA ATATGTAAT GCTGCCATGT
TCCAAACCCA TCGTCAGTGT GTGTGTTAG AGCTGTGCAC CCTAGAACAA ACATATTGTC
CCATGAGCAG GTGCCTGAGA CACAGACCCCC TTTGCATTCA CAGAGAGGTC ATTGGTTATA
GAGACTTGAA TTAATAAGTG ACATTATGCC AGTTTCTGTT CTCTCACAGG TGATAAACAA

-continued

TGCTTTTGT GCACTACATA CTCTTCAGTG TAGAGCTCTT GTTTATGGG AAAAGGCTCA

AATGCCAAAT TGTGTTTGAT GGATTAATAT GCCCTTTGC CGATGCATAC TATTACTGAT

GTGACTCGGT TTTGTCGAG CTTTGCTTG TTTAATGAAA CACACTGTA AACCTCTTT

GCACCTTGAA AAAGAATCCA CGGGGATGCT CGAGCACCTG TAAACAATT 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, ST3GAL5, IL1R1,
BIN1, WIPF1, TFP1, FN1, FAM134A, NRIP1, RAC2, SPP1, PHF15, BTN3A2,
SESN1, MAP3K5, DPYSL2, SEMA4D, STOM, and MAOA gene.

REFERENCES

- [0203] 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.
- [0204] U.S. Pat. No. 5,143,854
- [0205] U.S. Pat. No. 5,202,231
- [0206] U.S. Pat. No. 5,242,974
- [0207] U.S. Pat. No. 5,288,644
- [0208] U.S. Pat. No. 5,324,633
- [0209] U.S. Pat. No. 5,384,261
- [0210] U.S. Pat. No. 5,405,783
- [0211] U.S. Pat. No. 5,412,087
- [0212] U.S. Pat. No. 5,424,186
- [0213] U.S. Pat. No. 5,429,807
- [0214] U.S. Pat. No. 5,432,049
- [0215] U.S. Pat. No. 5,436,327
- [0216] U.S. Pat. No. 5,445,934
- [0217] U.S. Pat. No. 5,468,613
- [0218] U.S. Pat. No. 5,470,710
- [0219] U.S. Pat. No. 5,472,672
- [0220] U.S. Pat. No. 5,492,806
- [0221] U.S. Pat. No. 5,503,980
- [0222] U.S. Pat. No. 5,510,270
- [0223] U.S. Pat. No. 5,525,464
- [0224] U.S. Pat. No. 5,525,464
- [0225] U.S. Pat. No. 5,527,681
- [0226] U.S. Pat. No. 5,529,756
- [0227] U.S. Pat. No. 5,532,128
- [0228] U.S. Pat. No. 5,545,531
- [0229] U.S. Pat. No. 5,547,839
- [0230] U.S. Pat. No. 5,554,501
- [0231] U.S. Pat. No. 5,556,752
- [0232] U.S. Pat. No. 5,561,071
- [0233] U.S. Pat. No. 5,571,639
- [0234] U.S. Pat. No. 5,580,726
- [0235] U.S. Pat. No. 5,580,732
- [0236] U.S. Pat. No. 5,593,839
- [0237] U.S. Pat. No. 5,599,672
- [0238] U.S. Pat. No. 5,599,695
- [0239] U.S. Pat. No. 5,610,287
- [0240] U.S. Pat. No. 5,624,711
- [0241] U.S. Pat. No. 5,631,134
- [0242] U.S. Pat. No. 5,639,603
- [0243] U.S. Pat. No. 5,654,413
- [0244] U.S. Pat. No. 5,658,734
- [0245] U.S. Pat. No. 5,661,028
- [0246] U.S. Pat. No. 5,665,547
- [0247] U.S. Pat. No. 5,667,972
- [0248] U.S. Pat. No. 5,695,940
- [0249] U.S. Pat. No. 5,700,637
- [0250] U.S. Pat. No. 5,744,305
- [0251] U.S. Pat. No. 5,800,992
- [0252] U.S. Pat. No. 5,807,522
- [0253] U.S. Pat. No. 5,830,645
- [0254] U.S. Pat. No. 5,837,196
- [0255] U.S. Pat. No. 5,847,219
- [0256] U.S. Pat. No. 5,871,928
- [0257] U.S. Pat. No. 5,876,932
- [0258] U.S. Pat. No. 5,919,626
- [0259] U.S. Pat. No. 6,004,755
- [0260] U.S. Pat. No. 6,087,102
- [0261] U.S. Pat. No. 6,368,799
- [0262] U.S. Pat. No. 6,383,749
- [0263] U.S. Pat. No. 6,617,112
- [0264] U.S. Pat. No. 6,638,717
- [0265] U.S. Pat. No. 6,720,138
- [0266] U.S. Patent Publn. 2010/0135956
- [0267] Belanoff et al., *Eur. J. Pharmacol.*, 655 (1-3): 117-20, 2011.
- [0268] Cho et al. *Biochemistry*, 44 (9): 3547-61, 2005.
- [0269] Clark, *Curr. Top. Med. Chem.* 8 (9): 813-838, 2008.
- [0270] Colleoni et al., *Annals of Oncology*, 11 (8): 1057, 2000.
- [0271] European Appln. EP 373 203
- [0272] European Appln. EP 785 280
- [0273] European Appln. EP 799 897
- [0274] Evans, *Science*, 240:889, 1988.
- [0275] Fodor et al., *Science*, 251:767-777, 1991.
- [0276] Grover and Martin, *Carcinogenesis*, 23 (7): 1095-102, 2002.
- [0277] Hacia et al., *Nature Genet.*, 14:441-449, 1996.
- [0278] Harrison's Principles of Internal Medicine, Kasper et al. (Eds.), 16th Ed . . . , Chapter 70, 2005.
- [0279] Henderson et al. *Cancer Res.*, 48:246-253, 1988.
- [0280] Keen and Davidson, *Cancer*, 97 (3 Suppl): 825-33, 2003.
- [0281] Ma et al., *J. Immunol.*, 171 (2): 608-615, 2003.
- [0282] MacBeath and Schreiber, *Science*, 289 (5485): 1760-3, 2000.
- [0283] Melhem et al, *Clin. Cancer Res.*, 15 (9): 3196-204, 2009.
- [0284] Mikosz et al., *J. Biol. Chem.*, 276:16649-54, 2001.
- [0285] Moran et al., *Cancer Res.*, 60:867-872, 2000.

- [0286] Pandey and Mann, *Nature*, 405 (6788): 837-46, 2000.
- [0287] Pang and Conzen, *Cancer Biol. Ther.* *Cancer Biol. Ther.*, 5 (8): 933-40, 2006.
- [0288] PCT Appln. WO 01/68255
- [0289] PCT Appln. WO 03/020898
- [0290] PCT Appln. WO 03/022421
- [0291] PCT Appln. WO 03/023058
- [0292] PCT Appln. WO 03/029485
- [0293] PCT Appln. WO 03/040410
- [0294] PCT Appln. WO 03/053586
- [0295] PCT Appln. WO 03/066906
- [0296] PCT Appln. WO 03/067217
- [0297] PCT Appln. WO 03/076928
- [0298] PCT Appln. WO 03/087297
- [0299] PCT Appln. WO 03/091426
- [0300] PCT Appln. WO 03/093810
- [0301] PCT Appln. WO 03/100448A1
- [0302] PCT Appln. WO 04/020085
- [0303] PCT Appln. WO 04/027093
- [0304] PCT Appln. WO 09/923256
- [0305] PCT Appln. WO 09/936760
- [0306] PCT Appln. WO 93/17126
- [0307] PCT Appln. WO 95/11995
- [0308] PCT Appln. WO 95/21265
- [0309] PCT Appln. WO 95/21944
- [0310] PCT Appln. WO 95/35505
- [0311] PCT Appln. WO 96/30347
- [0312] PCT Appln. WO 96/31622
- [0313] PCT Appln. WO 96/33980
- [0314] PCT Appln. WO 97/10365
- [0315] PCT Appln. WO 97/27317
- [0316] PCT Appln. WO 9743450
- [0317] PCT Appln. WO 99/35505
- [0318] PCT Appln. WO 01/38580
- [0319] PCT Appln. WO 03/100012
- [0320] Pease et al., *Proc. Natl. Acad. Sci. USA*, 91:5022-5026, 1994.
- [0321] Peeters et al., *Ann. NY Acad. Sci.*, 1148:536-41, 2008.
- [0322] Pike et al., *Epidemiologic Rev.*, 15 (1): 17-35, 1993.
- [0323] Shoemaker et al., *Nature Genetics*, 14:450-456, 1996.
- [0324] Sims et al. *BMC Medical Genomics*, 1 (42): 1-14, 2008.
- [0325] Sorlie et al., *Proc. Natl. Acad. Sci. USA*, 98:10869-10874., 2001.
- [0326] Srinivas et al., *Clin. Chem.*, 48 (8): 1160-9, 2002.
- [0327] UK Appln. 8 803 000
- [0328] Wu et al., *Cancer Res.*, 64:1757-64, 2004.
- [0329] Wu et al., *J. Clin. Invest.*, 114:560-568, 2004.
- [0330] Wu et al., *Mol Endocrinol.*, 2006

SEQUENCE LISTING

Sequence total quantity: 46
 SEQ ID NO: 1 moltype = DNA length = 4794
 FEATURE Location/Qualifiers
 source 1..4794
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 1

```

tttttagaaa aaaaaaatat attcccttc tgctccttgc gcggtcacaa gctaaggttg 60
ttatctcgcc tgcggccggga actgcggacg gtggcgggcg aegcgcttct ctggccaggt 120
tgatattcac tgatggactc caaagaatca ttaactctgt gtagagaaga aaaccggc 180
agtgtgcttg ctcaggagag gggagatgtg atggacttctt ataaaacctt aagaggagg 240
gtctactgtga aggtttctgc gtcttcaccc tcaactgtgtc tcgtttctca atcagactcc 300
aagcagcggaa gacttttggt tgatttcca aaaggctcag taagcaatgc gcacgcagcca 360
gatctgtcca aaggcgtttc actctcaatg ggactgtata tggggagagac agaaacaaaa 420
gtgatgggaa atgacacttggg atccccacag caggccaaaa tcagccccc ctccggggaa 480
acagacttaa agcttttggg agaaaggcatt goaaacccatca ataggtcgac cagtttccaa 540
gagaacccca agagttcagc attcactgtgt gtgtctgtgt ccccccacaga gaaggagtt 600
ccaaaaacttc actctgtatg atcttcagaa cagcaacatt tgaaggccca gactggcacc 660
aacgggtggca atgtgaaattt gtataccaca gaccaaagca cctttgacat tttgcaggat 720
ttggaggttt ctctggggc cccaggtaaa gagacgatgg agagtccctg gagatcgac 780
ctgttgatag atgaaaaactt ttgcctttt cctctggggc gagaagacgaa ttcatccctt 840
ttgaaaggaa actcgaatgaa ggactgcaag cctctcattt taccggacac taaaacccaa 900
attaaggata atggagatct gggtttgtca agcccccaga atgtaaactt gccccaaagt 960
aaaacagaaaa aagaacttcatcgaactt tgcacccctgt gggtaattaa gcaaggaaaaa 1020
ctgggcacag tttactgtca ggcaagcttt cctggggatca atataattgg taataaaaatg 1080
tctgcattt ctgttcatgg tttgtagtacc tctggggac agatgtacca ctatgacatc 1140
aatacagcat cccttctca acagcaggat cagaagccaa tttttaatgt cattccacca 1200
attcccggtt gttccggaaa ttggatagg tgccaaaggat ctggatgtca caactgtact 1260
tctctggggc ctctgtactt ctctgggtca acatgtttt ctaatggctt ttcacgcccc 1320
agcatcgacag catatgtaaat ctctccctca tccagcttctt caacagcaac aacagcggca 1380
cctcccaaac tctgccttgtt gtgtctgtat gaagctttagt gatgtcattt tgggttctt 1440
acttggaaatgg gctgttcaaa agagcgttgg aaggacagca caattaccta 1500
tgtgttggaa ggaatgtttt catcatcgat aaaattccgaa gaaaaacttcccccacatgc 1560
cgctatcgaa aatgtcttca ggctggaaatg aacctggaaatg ctggaaaaac aaagaaaaaa 1620
ataaaaaggaa ttcaacggc cactacaggaa gtctcaacaa aacccctgtaa atatctgtt 1680
aacaacaaatgg tagttccctca aacgttacca caactccccc ctacccttgtt gtcactgtt 1740
gagggttattt aacctgttggt gtttatgtca ggtatgtata gctctgttcc agactcaact 1800
tggaggatca tgactacgct caacatgtta ggagggccgc aagtgttgc agcgtgaaa 1860
tggccaaagg caataccagg ttccggaaac ttacacccgg atgacccaaat gaccctactg 1920
cgtactccctt ggtgtttctt tatggccattt gctctggggc ggagatcata tagacaatca 1980

```

-continued

| | | | | | | |
|--------------|-------------|-------------|-------------|-------------|-------------|------|
| agtgc当地 | tgctgtgttt | tgctctgtat | ctgatttata | atgagcagag | aatgactcta | 2040 |
| cctgtatgt | acgccaatg | taaacacatg | ctgtatgttt | cctctgagg | acacaggctt | 2100 |
| caggatctt | atgaagagta | tctctgtat | aaaacccttac | tgtttcttc | ttcagttcct | 2160 |
| aaggacggc | tgaagagcc | agagotattt | gatgaaatta | gaatgaccta | catcaaagag | 2220 |
| ctaggaaaag | ccattgtcaa | gagggaaagga | aactccagcc | agaactggca | gcgggtttat | 2280 |
| caactgacaa | aactcttgg | ttctatgtat | gaagtgggtt | aaaatctct | taactatgc | 2340 |
| ttccaaacat | ttttggataa | gaccatgagt | attnaattcc | cccgagatgtt | agctgaatc | 2400 |
| atcacaatc | agataccaaa | atattcaat | ggaatatac | aaaactctt | gtttcatcaa | 2460 |
| aagtgactgc | cttaataaaga | atgggtgcct | taaagaaagt | cgaattaata | gctttattt | 2520 |
| tataaactat | caagtgttc | tgtgaggtt | ttgttgtttt | atttttattt | gttttcatct | 2580 |
| gttgggtt | ttaataacag | cactatgtat | ggtttataga | ggggcaagac | ttggcaacag | 2640 |
| aaggcgttga | gtcgtcatca | cttttcgt | atggagagt | agatggtaa | attttatgt | 2700 |
| taatataatcc | cagaaattag | aaacctaata | atgtggacgt | aatctccaca | gtcaaaagaag | 2760 |
| gatggcacat | aaaccaccc | tgcaccaatg | ctgtgtgtat | aactttctt | tcatacttt | 2820 |
| tttcacagg | ggctggat | aattttctag | actttctgtt | ggtgtatccc | ccccctgtta | 2880 |
| tagttaggt | agcattttt | attatgtat | gaaacctgt | aaaaaaatgtt | acaagtgtat | 2940 |
| atcagaaaag | ggaagtgt | cctttatag | ctattactgt | ctggttttaa | caattctt | 3000 |
| tatattttatgt | gaactacgt | tgctcatttt | ttcttataca | atttttattt | caagttattt | 3060 |
| tacagctgtt | taagatgggc | agcttagtgc | tagctttccc | aaataaaactc | taaacattaa | 3120 |
| tcaatcatct | gtgtaaaaat | gggttgggt | ttcttaacctg | atggacttca | gctatcagaa | 3180 |
| gaccacaaaa | attgactcaa | atctccagta | ttcttgcata | aaaaaaaaaa | aaaaaaagtc | 3240 |
| atattttgtt | tatatactgt | tcagtggaga | atttatatagg | tttgtcaata | taacagtctt | 3300 |
| aactggat | gaggcacat | tccatgttgc | tgtgtggat | actgtggat | atgtgtca | 3360 |
| aagactaaat | taaaaaataa | ctaccacag | gccctgtctg | ttactaaacgc | cctattttt | 3420 |
| caatggctat | atggcaagaa | agctggtaaa | ctatttgc | ttcaggactt | tttgaagtag | 3480 |
| ttgttataac | ttctttaaaag | tttggtattt | agataaccag | ctgttaacaca | gctgagagac | 3540 |
| ttttaatccat | acaaagata | tccttctact | aaacccatca | atcttataa | 3600 | |
| tggcaaaat | ggcttagacac | ccatccat | atccctat | gtcaccat | ggtaatctt | 3660 |
| tcctgtatgt | acagaaagc | tcagacttgc | atttttgt | ttttagactg | tatgtatgc | 3720 |
| agacatccat | gtttgtaaa | ctacacatcc | ctaatgtgt | ccatagagg | taacacaat | 3780 |
| cctgtgattt | tcttcactgt | tgaaaaattt | ttttaaacaa | atagaagctg | tagtagcc | 3840 |
| ttctgtgt | acccatccaa | ctttctgt | actccaaat | taatactt | actaagccac | 3900 |
| aaaaaattt | atttctattt | aagggtggca | atttatttgc | gtaatagaaa | actgaaaatc | 3960 |
| taatattaaa | aatatggaa | ttctaatata | ttttttat | tagttatag | ttcagatata | 4020 |
| tatcatatgt | gtatttacta | atctggaa | ggaaagggtca | ctgcagottt | acatgcattt | 4080 |
| tattaaatgt | atgtttaaaat | agcttgcata | gtgttataa | agaatgtttt | tttagttaga | 4140 |
| ttgtttttatc | atgatcatgtt | atataatttt | tgtaggggtt | aaagaaatgt | tgatgtata | 4200 |
| ccttatatgt | ttatagttt | tacatgcatt | catacaggca | gcgtatgtt | cagaaaccaa | 4260 |
| acagttgtt | ctaggggaa | aggggatgg | agactgggt | tgtgtgat | gaagggttgc | 4320 |
| gaggctgt | cccaatgt | ttacacgtt | agttatcttc | tgcctccat | tctgaccacc | 4380 |
| ctcttcattt | caacacgt | tctgtcaggg | cagggtttagt | ttactcaatc | tcccttgc | 4440 |
| ctaaaatgtat | taaagtatgt | aaacaggaga | caggaagggt | gtgcttacat | ccttaaaggc | 4500 |
| accatctaata | acgggggttac | tttcacat | agccctccc | cagcagttga | atgacaacag | 4560 |
| aagcttcaga | atgttggca | tagttgc | agaggatcca | gcaatgtt | aatagtgcag | 4620 |
| aatctcatag | gttgccata | atactata | tcctttctat | cctacacaa | gagtttattt | 4680 |
| ccaaataaaa | tgaggacat | tttttgc | tttgc | tttttgc | tttgc | 4740 |
| ttttcgtat | tttgggaaa | tttataata | aaaaaaacaa | tcatttgctt | tttgc | 4794 |

| | | | | | | |
|--------------|-------------------------|---------------|-------------|-------------|-------------|------|
| SEQ ID NO: 2 | moltype = DNA | length = 6330 | | | | |
| FEATURE | Location/Qualifiers | | | | | |
| source | 1..6330 | | | | | |
| | mol_type = other DNA | | | | | |
| | organism = Homo sapiens | | | | | |
| SEQUENCE: 2 | | | | | | |
| aggagctggc | ggggggcggtt | cgtectggga | ctgcacttgc | tccctgtcggg | tcggccggct | 60 |
| tcacccggacc | cgccaggctcc | cgggggcagg | ccggggccag | agctcgcgtg | tcggcgccg | 120 |
| atgcgtcg | tgcctctaa | cctctgggt | tgctctttt | ccaggtggcc | cgccggttt | 180 |
| tgagccttct | ccccctcggtt | gacacggct | gcacccctgc | cgccggccac | gaccatgacc | 240 |
| atgaccctcc | acaccaaaagc | atctgggtat | gcctactgc | atcagatcca | agggaaacgag | 300 |
| ctggagccccc | tgaacccgtcc | gcagatca | atccccctgg | agccggccct | gggcggagg | 360 |
| taacctggaca | gcaacaaagcc | cgccgtgtac | aactacccccc | agggccgcgc | ctacggat | 420 |
| aacggcccg | ccggccccaa | cgccgcagg | tacggcaga | ccggccctcc | ctacggcccc | 480 |
| gggtctgagg | ctggggcggtt | cggttccaa | ggccctgggg | gtttccccc | actcaacagc | 540 |
| gtgtctccga | cccccgtat | gtctgc | ccggccggcc | agctgtcggc | tttcttgc | 600 |
| ccccacggcc | acggatgtcc | ctactatgt | gagaacgcgc | ccagggctca | cacgggtc | 660 |
| gaggccggcc | ccggccgtt | ctacaggca | aatttgcata | atcgacggca | gggttgc | 720 |
| gaaaagattgg | ccagttccaa | tgacaaggga | agttggct | tggatctgc | caaggagact | 780 |
| cgctactgt | caagtgtca | tgactatgt | tcaggctacc | attatggat | ctggcttgc | 840 |
| gagggtgtca | aggcccttctt | caagggaaat | atcaacggac | ataacgcata | tatgtgtcca | 900 |
| gccaccaacc | atgtcaccat | tgataaaaaac | aggaggaga | gtgtccggc | ctggccggct | 960 |
| cgccaaatgt | acggatgtgg | aatgtataaa | ggtgggtat | gaaaacccg | aagaggagg | 1020 |
| agaatgttga | aacacaagcg | ccagagat | gatggggagg | gcaggggtga | agtgggtct | 1080 |
| gctggagaca | tgagagctgc | caacccctgg | ccaagccgc | tcatgtat | acgctctaa | 1140 |
| aagaacagcc | tggcccttgc | cctgacggcc | gaccagatgg | tcagtgcctt | gttggatgt | 1200 |
| gagccccccca | tacttattt | cgatgtat | cctaccagac | ccttcgtat | agcttcgtat | 1260 |
| atggggcttac | tgaccaaccc | ggcagacagg | gagctgggtc | acatgtat | ctggggcgaa | 1320 |

-continued

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| agggtgccag | gctttgtgga | tttgaccctc | catgatcagg | tccacccct | agaatgtgcc | 1380 |
| tggcttagaga | tctctgtat | tggctcgto | tggcgctcca | tggagcaccc | agggaaagcta | 1440 |
| ctgttgcctc | ctaactgtct | cttggacagg | aaccagggaa | aatgtgtaga | gggcattggg | 1500 |
| gagatcttcg | acatgtgtct | ggctacatca | tctcggttcc | gcatgtgaa | tctgcaggga | 1560 |
| gaggagttt | tgtgcctcaa | atcttatttt | ttgtttaatt | ctggagtgt | cacatttctg | 1620 |
| tccagcaccc | tgaagtctct | ggaagagaag | gaccatatcc | accgagtct | ggacaagatc | 1680 |
| acagacaccc | tgtatccacct | gatggccaaag | gcaggcctga | ccctgcagca | gcagcaccc | 1740 |
| cggtctggccc | agctccctct | catectctcc | catatcaggc | acatgatgaa | caaaggcatg | 1800 |
| gagcatctgt | acagcatgaa | gtgcaagaac | gtgggtgcucc | tctatgacct | gctgtggag | 1860 |
| atgctggacg | ccccacccct | acatgcgccc | actagccgt | gaggggcatc | cgtggaggag | 1920 |
| acggaccaa | gccaacttgc | caactgggg | tetacttcat | cgcatccctt | gcaaaagat | 1980 |
| tacatcacgg | gggaggccaga | gggttccct | gccacgggt | gagagctccc | ttggctccac | 2040 |
| acgggttcaga | taatccctgc | tgcatttac | cctcatcatg | caccacttta | gccaattct | 2100 |
| gttctctgca | taacactccgg | catgcatca | acaccaatgg | ctttagat | gagtgccat | 2160 |
| tcattttgtt | gtcagtctct | taggtggaca | tcttctgtct | tctgtggga | acagccaaag | 2220 |
| ggatttcaag | gtaataatctt | tgttaacagct | ctctttcccc | ttgtctatgt | tactaagct | 2280 |
| gaggattccc | gtagcttttc | acagctgaa | tcagtcat | ggttgggct | cagataactc | 2340 |
| tgtgcattta | agactactgt | agagacccca | gcctggagag | tagacattt | gcctctgata | 2400 |
| agoactttt | aaatggctct | aagaataa | acagcaaaag | aattttaaatg | ggctcttta | 2460 |
| atttggtaact | tgggaaaagc | taggtcaagg | gtttagattta | gaccccttct | gtatccct | 2520 |
| ggcaatgc | ccttttatga | aagtggta | ccttaaagct | tttata | tgttagcaag | 2580 |
| tatctgttga | ttgttcaat | atccccctca | taggaataca | aggggcacac | agggaaaggca | 2640 |
| gatcccctt | ttggcaagac | tattttact | tgatcac | cagatca | tgtgtggaaa | 2700 |
| gctctgcctc | ttggcttccgg | gtcatgggtt | ccatgat | catgcctccc | atggacattat | 2760 |
| ggagagcage | aagttgatct | tagttaa | tccttatatg | agggataa | tcctgtat | 2820 |
| tgtttttttt | tttgtgttac | aaaagaaagd | cctccctccc | tgaacttgc | gtaaggtcag | 2880 |
| tttcaggacc | ttgttccagtg | ggcactgtac | tttgatctt | ccggcgtgt | tgtgcocttac | 2940 |
| acagggggtga | atgttca | gttggatgt | atgtatgggg | taatggtag | ttgaaaggag | 3000 |
| caggggccct | ggtgttgc | ttagccctgg | ggcattggac | tgaacagta | tttgccagga | 3060 |
| tttgtgtgc | tactagagaa | caaggggaa | agtagggcag | aaactggata | cagttctgag | 3120 |
| gcacagccag | acttgcgt | gggtggccct | ccacaggctg | cagtcacca | ggaacattcc | 3180 |
| tttcagacc | ccatggccc | tttgggg | ccctgggatc | cttgggtag | tccagctt | 3240 |
| cttcattttc | cagctggcc | ctgttggaa | gaagcagtc | tcacagtc | tgttagac | 3300 |
| tgtgttcc | caattggccc | agcacccctgg | ggcacgggag | aagggtggg | accgtgt | 3360 |
| tcactactca | ggctgactgg | ggcctggta | gattacgtat | gcccctgg | gtttagagat | 3420 |
| aatccaaat | cagggttgg | tttgggg | aaaatccctc | cccttctcc | cccgccccgt | 3480 |
| tccttacccg | ctccactctc | gcca | ttcattcaat | ttccttgc | ctataggt | 3540 |
| aaaaagaaag | gtcatttcca | gccacagg | agccttccct | gggccttgc | ttctctagca | 3600 |
| caattatgg | ttaacttcc | tttcaat | aaaaagaatg | tttgat | ttctggg | 3660 |
| cttatttgc | gttatttgg | ccctttag | aggtgatgt | tgtgttag | aatgacc | 3720 |
| gtgagctgt | ccgggttctc | ttgttat | ttttttggaa | aagtggattt | catttatt | 3780 |
| tgattgtca | gttaagtgt | cacaaagga | ctgagaatct | gggaggccaa | aaaaaaaaaa | 3840 |
| aaagttttta | tgtgactt | aatttggg | caattttat | tatctgtgtt | aaggatatgt | 3900 |
| ttaagaacat | aattttttt | ttgtgtt | ttaagaag | accttagtt | tttaagaag | 3960 |
| caccttat | atataat | atattttt | gaaattat | tgcttgc | tcagacatt | 4020 |
| gaatgttag | atctgttct | ggat | ttttttt | tgactgggt | aacatgc | 4080 |
| aaatatttag | ttttttttt | ttttttt | tacttttca | gtcacccct | catgtat | 4140 |
| gtcattttat | ctaaaggct | gtgttatt | cattttat | aaagatcac | ttcatat | 4200 |
| cttttgc | cacagttagac | aaaaatag | taatccat | gccttatt | ggatactg | 4260 |
| tgacagacaa | tcttgc | caaagatt | gcctgaa | aaaaattt | cagggcag | 4320 |
| aattttgtt | ttacaaaat | atca | atattttt | ggacat | atgggtc | 4380 |
| gggttcttt | taatgtttat | acttggat | tcttttttt | aaaaat | aaaaaaa | 4440 |
| aaaatcttca | ggactagac | atgtatacc | agctaaag | aaacaattat | acatgg | 4500 |
| gttttacatt | attcatccaa | tgtttct | ttcatgtt | atactacta | catttga | 4560 |
| gggcagagaa | catcgat | ttgaaat | cgcccagg | tccacg | cttggaa | 4620 |
| ctttttgtt | ttttactgt | agtgc | atggacag | catat | ggctgtat | 4680 |
| gttatttgg | gttagaa | gat | ttttttttt | cctctgtt | ccccact | 4740 |
| gaggcaagg | aaaatgt | ttttttttt | atgtgtt | ttttttttt | ttttttttt | 4800 |
| tggatttgg | aatctgggg | atggcaaa | tattaaga | agttttttt | ttttttttt | 4860 |
| aaaaatgtt | atttctgg | gtgc | tttccat | gttgc | tttccat | 4920 |
| ccacaaatca | atagcttca | tttca | atggacag | catat | tttccat | 4980 |
| agaaagaaca | acatcagc | taa | tttca | gttgc | tttccat | 5040 |
| ccattgcct | gttgcgt | atgat | tttca | tgca | tttccat | 5100 |
| aggcattca | aaagaaag | cctat | tttca | tttca | tttccat | 5160 |
| gcagtctt | atttgc | tctat | tttca | tttca | tttccat | 5220 |
| atgcctt | agggt | aaata | tttca | tttca | tttccat | 5280 |
| agggttctc | acctt | tttca | tttca | tttca | tttccat | 5340 |
| agagacttcc | ttccct | tgac | tttca | tttca | tttccat | 5400 |
| ttatgat | gtactgt | acttgg | tttca | tttca | tttccat | 5460 |
| cacaggcaga | tctgtt | gaag | tttca | tttca | tttccat | 5520 |
| ctctaaaggc | aaccat | tttca | tttca | tttca | tttccat | 5580 |
| tgcaccat | ccaa | tttca | tttca | tttca | tttccat | 5640 |
| gtcccaata | ttcat | tttca | tttca | tttca | tttccat | 5700 |
| ttccaatt | attaa | tttca | tttca | tttca | tttccat | 5760 |
| tgctgcct | attatgg | tttca | tttca | tttca | tttccat | 5820 |
| tattttttt | tttgc | tttca | tttca | tttca | tttccat | 5880 |

-continued

| | | | | | | |
|------------|-------------|-------------|------------|------------|-------------|------|
| tccaaaccca | tcgtcaagtgt | gtgtgtttag | agctgtgcac | cctagaaca | acatattgtc | 5940 |
| ccatgagoag | gtgcctgaga | cacagacccc | tttgcattca | cagagaggc | attggttata | 6000 |
| gagacttgaa | ttaataagt | acattatgcc | agtttctgtt | ctctcacagg | tgataaaacaa | 6060 |
| tgtttttgt | goactacata | ctcttcagtg | tagagcttt | gttttatggg | aaaaggctca | 6120 |
| aatgccaat | tgtgtttgtat | ggattaatat | gcccctttgc | cgatgcatac | tattactgtat | 6180 |
| gtgactcggt | tttgtcgca | ctttgccttg | ttaatgaaa | cacactgtt | aacctcttt | 6240 |
| gcactttgaa | aaagaatcca | cgggatgtct | cgagcacctg | taaacaattt | tctcaaccta | 6300 |
| tttgatgttc | aaataaagaa | tttaactaaaa | | | | 6330 |

| | | | | | | |
|--------------|-------------------------|---------------|-------------|-------------|-------------|------|
| SEQ ID NO: 3 | moltype = DNA | length = 4107 | | | | |
| FEATURE | Location/Qualifiers | | | | | |
| source | 1..4107 | | | | | |
| | mol_type = other DNA | | | | | |
| | organism = Homo sapiens | | | | | |
| SEQUENCE: 3 | | | | | | |
| ggccaacccct | ccggaagctg | ccggcccttt | ccccctttat | gggaataactt | tttttaaaaa | 60 |
| aaaagagttc | gctggcgcca | ccccgttagga | ctggccgccc | taaaacogtg | ataaaggagc | 120 |
| tgtctccac | ttctcaactc | cgcttccttc | cgatcaaggag | tcggggctt | ccccagttt | 180 |
| ctcagccagg | cgccggcgcc | gactggcaat | gtttggccct | aaaagaaacg | cggtatcgg | 240 |
| actcaaccc | tactgtgggg | ggccggctt | ggggggccgc | acggccggcg | ccacccccc | 300 |
| gggagggcga | ctttggcta | cggagaagga | ggccctggcc | cgccgagaga | tagggggagg | 360 |
| ggagggccgc | gggggttattt | cgccgcaagc | ccccctgtca | cccttcacgcc | 420 | |
| agactccccc | agggtcgccg | ggccggccgc | cattggcgcc | gagggtcccc | acgttcacgc | 480 |
| gaccggcccg | aggctgttt | tcttcggcc | caccggccgc | ggggccgcgc | ttgaggagat | 540 |
| ggaagccccc | gcccgtacg | ccatcatgtc | gcccgaagag | gagctggacg | gttacggac | 600 |
| ggagcccttc | gggaagcgcc | cggtgttct | gcccgtgtt | gagtttgcg | ggaaatctgg | 660 |
| taataaaccc | agtaacggac | ggtcaactac | ctggacgcgc | ccggccacgc | aggaggagg | 720 |
| ggacgagttt | taccggcagt | cgctggagat | tatcttcgg | tacccctggg | agcaggccac | 780 |
| cgccgcggaa | gacacaaagc | caatggggag | gtctggggcc | accageagga | aggcgcttga | 840 |
| gaccttaacg | cggggtgggg | atggcgtca | gcccgaacc | gagacggct | tccaaggcat | 900 |
| gcttcggaaa | ctggggatcc | aaaacgttca | cgatgttca | tcgttgc | gagtgtat | 960 |
| ccatgtttt | agcgcacggc | taacaaactg | ggccggaggat | gtgacttca | tttcttttgg | 1020 |
| tgcccttgg | gtctaaact | tgaagaccat | aaaccaagaa | agctgtatcg | aaccattagc | 1080 |
| agaaagtatc | acagacgtt | tcgttaaggac | aaaacgggac | tggctatgtt | aacaaagagg | 1140 |
| ctggggatgg | tttgcgttgc | tcttcatgt | agaggacca | gaagggttgc | tccggaaatgt | 1200 |
| gctgtctgg | tgttcggat | tttgcggat | tttgcgttgc | tttgcataatc | taataagata | 1260 |
| gccttactgt | aagtgtata | gttgcatttt | aaccaaccac | caccaccacc | aaaaccagg | 1320 |
| tatgcagtt | gacttcaagc | tgtaacttcc | tagagttca | cccttagcaac | ctagccagaa | 1380 |
| aaagcaagtgg | caagggattt | atggcataa | agaataaaa | catggaaaga | gtgtccccc | 1440 |
| ttgattgttgc | atgtactgtc | tggaaagac | aaagtttgc | ttcgtcaaca | aacaaacattt | 1500 |
| gtttggggag | ctatggagga | ggacttttag | atttagtga | gatggtaggg | tggaaagact | 1560 |
| taatttcctt | gttggaaaca | ggaaaatggc | cgttagccg | gcaagtcat | gaattgatta | 1620 |
| cccgccgaaat | tcattaaatc | actgtttgtt | taagggaaac | actaaatgt | ccagtgaccc | 1680 |
| gtgtaaaatgt | tacaaatgt | aaactgtat | ctgttttttt | cgtactgtt | caagggaaac | 1740 |
| tttccttc | ttaattttgc | tttcccaat | tactttttt | aaactgttca | tttttttttt | 1800 |
| ttttataactt | gttataactt | ggcttggttt | ccatgtttt | tacttttata | gccttagttt | 1860 |
| tcaccaataa | tacttgcacgg | aagggttgc | tttgcgttgc | tttgcataatc | taatgtat | 1920 |
| cttaagggatc | tttgcgttgc | tattttttat | ttttttttat | ttttttttat | ttttttttat | 1980 |
| tctttgttgc | aaagggatgtt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2040 |
| gcaacttccg | taattttggaa | cctgtttttt | acagttttt | tttgcgttgc | tttgcgttgc | 2100 |
| tcagtttgc | agtttgcata | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2160 |
| ggaacaaatc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2220 |
| tttagatgtt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2280 |
| cttcattata | tgcaagtttt | caataattag | gtctaaatgg | agttttttag | ttactgtat | 2340 |
| cttacaaata | atgggtctgt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2400 |
| tttaactgtt | gaaatttttt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2460 |
| tttcgttgc | aatggaaatc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2520 |
| gagagacatt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2580 |
| gttagtctaa | ccatgtttgtt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2640 |
| gtatgttcaat | aaagggatgtt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2700 |
| gcaatgtttt | gttggatgtt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2760 |
| ccttttattt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2820 |
| atatcaatc | tcacagtttt | ccccgtccat | ccccgtccat | tttgcgttgc | tttgcgttgc | 2880 |
| ttggcaactgt | aaaccccttc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2940 |
| agaccccttc | tcttcacgca | acttgcattt | tttgcgttgc | tttgcgttgc | tttgcgttgc | 3000 |
| gacactaaac | gtcatttttt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 3060 |
| aactattgtt | ttatcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 3120 |
| gaggaaagaa | cattttatgt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 3180 |
| ttaaattttac | aaaaaaatgt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 3240 |
| ctgaagaaat | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 3300 |
| cctttttttt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 3360 |
| ggaagaact | cccttgcctg | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 3420 |
| tttacctatg | tattttatgt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 3480 |
| ggacctttat | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 3540 |
| aaaaggctgt | ctgccaat | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 3600 |
| ggcccaacttc | ccaatttcatt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 3660 |

-continued

| | | | | | | |
|-------------|------------|-------------|------------|------------|------------|------|
| attttggcgt | tggggcagtg | agggcttagg | acaccccaag | tggtttggga | aaggaggagg | 3720 |
| ggagtgggg | gtttataggg | ggaggaggag | gcaggtggtc | taagtgtga | ctggctacgt | 3780 |
| agttcggggca | aatccctcaa | aaggaaaagg | gaggatttg | ttagaaggat | ggcgctccca | 3840 |
| gtgactactt | tttacttct | gtttgtcta | cgttctctc | agggaaaaac | atgcagtcct | 3900 |
| ctagtgttcc | atgtacattc | tgtgggggt | gaacaccttg | gttctggta | aacagctgta | 3960 |
| cttttgcata | ctgtgcagg | aagggtttagg | accaactaca | aattaatgtt | ggttgtcaaa | 4020 |
| tgtatgtgt | ttccctaact | ttctgtttt | cctgagaaaa | aaaaataaat | cttttattca | 4080 |
| aatacaggaa | aaaaaaaaaa | aaaaaaaaaa | | | | 4107 |

| | | | | | | |
|--------------|-------------------------|---------------|------------|-------------|-------------|------|
| SEQ ID NO: 4 | moltype = DNA | length = 1126 | | | | |
| FEATURE | Location/Qualifiers | | | | | |
| source | 1..1126 | | | | | |
| | mol_type = other DNA | | | | | |
| | organism = Homo sapiens | | | | | |
| SEQUENCE: 4 | | | | | | |
| tccccatgtg | acagtgagcg | gggtccccgc | tccaggagac | gctcgagct | qcgtccccgc | 60 |
| cctcagca | gtccactgtt | tcgggtccag | cagagaccag | caggccccc | acagtgggt | 120 |
| tttggccgtg | ccgcgtctca | acttgggtgt | cagagtgaat | tgccgtgtcc | ggagccggaga | 180 |
| gaggccggac | ggccaggaga | gaggggattt | ctgtcagcgc | cggccctcg | agctccggaga | 240 |
| catgaacggc | ttcaagcctg | acgagatgag | ccgcggccgg | gatggccgc | ccgcagttgc | 300 |
| cgcgtggc | gtcgccgg | ccgcggccgc | ctcgccgggg | aacgggaccc | gcgcgggac | 360 |
| cggggctgag | gtgcggccgc | cggggccgtt | ctcagcggct | gggccccccc | gggcggccgg | 420 |
| ggccggccccc | gggcactgt | gtctgcgtcg | ggaggatgg | gagcgtgtgc | gcccggccgc | 480 |
| aggcaacgc | agcttcagca | agaggatcca | gaagagcatc | tcccagaaga | aggtgaagat | 540 |
| cgcgtgtat | aagagcgc | ggcatctta | catatgtat | tatcataaaa | acttaattca | 600 |
| gagtgttgc | aacagaagaa | agagaaaagg | gagtgtat | gatggagggt | attcacctgt | 660 |
| tcaagatatt | gatacccccag | agggttattt | ataccaatta | caagtaataa | cacttaggg | 720 |
| atacaaaaga | cacttcaagc | taccaaccag | accaggactt | aataaaagcac | aactttgtta | 780 |
| gatagttgtt | tgccacttta | ggtctattcc | agtgaatgaa | aaagacactt | taacatattt | 840 |
| catctactca | gtgaagaatg | acaagaacaa | atcagatctc | agggttata | gtgggttca | 900 |
| ctaggagac | gttgcatttt | actaataatc | tggatgttta | cactgtttac | tgtttttca | 960 |
| catgttagaaa | tgttcttgc | gtatattttt | tacagggat | tttctctgt | tttatttttca | 1020 |
| ttgtttctga | ctctaataat | tagtggaaa | ctcatataaa | atgagcttc | ctaaattaaa | 1080 |
| tctatatttaa | ataaagggtt | ttactattaa | aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | 1126 |

| | | | | | | |
|--------------|-------------------------|---------------|------------|------------|-------------|------|
| SEQ ID NO: 5 | moltype = DNA | length = 2040 | | | | |
| FEATURE | Location/Qualifiers | | | | | |
| source | 1..2040 | | | | | |
| | mol_type = other DNA | | | | | |
| | organism = Homo sapiens | | | | | |
| SEQUENCE: 5 | | | | | | |
| tcgctgcgaa | ggacatttgg | gctgtgtgt | cgacgcgggt | cgaggggca | gtcggggaa | 60 |
| ccgcgaagaa | ggcgaggagg | ccggaggccc | gcgtgcgcgt | cctctctcag | tccaaaagcg | 120 |
| gcttttgggt | ccgcgtcgag | agacccgggg | gtctagcttt | tcctcgaaaa | ggccgcgcct | 180 |
| ggcccttggcc | ccgaaacacag | acaaagagca | ccgcggccgc | gatcgcgtcg | ggggcgctga | 240 |
| ggccggccat | ggtcatggaa | gtggccacc | tggacgtcg | aggcgtcg | gctgcgtcg | 300 |
| ggggaggcgc | ggcgcatacg | ctgtgtcg | actgcgcgtc | cctcttcgc | ttcaacgcgc | 360 |
| gccaacatgc | ccggctctgc | aaegtggcg | tcagcacat | cggtggccgc | cgggccaaag | 420 |
| ggcccatggg | cttggagcac | atcgccca | acgcccgcgt | ccgcggccgc | ctgtggccg | 480 |
| gcgccttacca | ccgcgtgtgt | ttgctggacg | agggcgcgc | ccgcctggac | ggcgccaaag | 540 |
| gegacgcgcac | cttggccctg | ggggccggcg | cgtctcg | cgagggcg | ccgcgcgcag | 600 |
| tcttcattcc | caaaaggaga | tacgaagcgt | tttcggccgt | ctggccggag | ctgtcagca | 660 |
| aacagtcgc | ccccatgggg | ctcagcctc | ccctgagatc | tagcgttcc | gacagcgcgg | 720 |
| aatctgggt | cgttctgc | agtacccac | tctacgtatc | gggtggcccg | gtgaaatcc | 780 |
| tgccttttct | gtacccctggc | agtgcgtatc | acgttccccc | caaggacatc | ctggatgcct | 840 |
| tgggcatac | tgccttgc | aaacgtctcg | ccatttgc | caacatttt | gagggtcaat | 900 |
| accagtttac | gagcatccct | gtggagaca | accacaaggc | agacatcgc | tccgtgtca | 960 |
| acggggccat | tgacttcata | gactccatca | agaatgtcg | aggaagggtg | tttgcact | 1020 |
| ggcaggcagg | cattttccgg | tcgcgcacca | tctgccttc | ttaccttata | aggactatc | 1080 |
| gagtcaagct | ggacggggcc | tttgatgtt | tgaagcagag | gcaagacatc | atctctccca | 1140 |
| acttcagttt | catggggccag | tcgtgtcg | tttgcgttca | ccatatttt | gagggtcaat | 1200 |
| ccgcagggc | ttggggccccc | gcatggctg | tgcgcaccc | aggcacctcc | accaccacgg | 1260 |
| tgttcaactt | ccccgtctcc | atccctgtcc | actccacgaa | cagtgcgtg | agtcaccc | 1320 |
| agagccatct | tacgaccc | cccaagctgt | gaaaggccac | ggggagggtg | gtcttcaca | 1380 |
| tcccatttttgc | actccatgt | ccttgcggat | tttgcgttca | ttactctgg | aggggctcga | 1440 |
| gagggtgtgt | ccttattttat | ttaacttca | cccgatcc | ctgggtttt | aaggcgttat | 1500 |
| ggtgtatgt | tagcgtcaag | acatttgc | aactcagcac | attcgggacc | aatatatagt | 1560 |
| gggtatcatca | agtccatctg | acaaaatggg | gcagaagaga | aggactcag | tgtgtatcc | 1620 |
| ggtttttttt | tgctgcggcc | ttttttttgt | agaatcttc | catgcgtac | ataccatcca | 1680 |
| gttattttcc | cgacgcacaca | tatacatatc | agaatatacc | tttattttat | ttgtgtatgt | 1740 |
| gtctgccttc | acaaatgtca | ttgttactc | ctagaagac | caaatactc | atttttttca | 1800 |
| tttgatgtact | gtactatcc | gtaaaatat | cttaagcagg | tttgcgttca | gcactgtatgg | 1860 |
| aaaataccag | tgttgggttt | tttttttagtt | gccaacatgt | gtatgttgc | tgattattta | 1920 |
| tgacctgaaa | taataatattt | cttcttctaa | gaagacattt | tgttacataa | ggatgacttt | 1980 |
| tttatacaat | ggaataaaatt | atggcatttc | tattgaaatt | tcaaaaaaaa | aaaaaaaaaa | 2040 |

-continued

SEQ ID NO: 6 moltype = DNA length = 3208
 FEATURE Location/Qualifiers
 source 1..3208
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 6

```

agatattcat gaaccgttgc ttcttcaga ctcgccttct cgctccctct gcctttctgg 60
cgctgttcc ctcctcctcc tctggcttct gctctttctt actccttctc tcagctgttt 120
aactacagct cccactggaa ctgcacaat caaaaacaac tcteectctt caagccgctt 180
ccaggagcgc atcacctggaa gaagagcgc acgcgtccccgg ccggggccgc ggaagagcag 240
ccaggtagct ggggggggg aggcttaccc ttctcccgctt cggtaagagc cacagcatct 300
ccccggagat tggccgttcc ccacccgttcc gcccccaaggg tcttgacggc gtgtatgtata 360
tgtttccggag caatgttggaa aggagaaaaag cgcgtgttggc tggcaacttga aagtggggag 420
aggttgttgc agtagctgtt gctgttggaa ggcgcgttggc agaaatggc cccgttagat 480
tctccatccc gtcgttctt cattaactgt ctgcagggtt taaaccgggg aaacagatata 540
gactaaccac ggccgggttcc aacctggatc tataacttggt aattccccac ggtggaaaat 600
ggtaaacaaa gacatgttgc gatcccgtt caagaaatgc tcagcccttc aattttttaa 660
gaagggggttca aaggttggaa tcaagagccc aatggttagt gtggacaagc atcagatgtcc 720
cagcccttggaa tacacccgtt ctcctatgtt gcacatccctt ccaggggggc cagacttgcg 780
gttctccctt tgcataacat gcctgggttca acatgttttcc caaaaggggg ttctccctca 840
ggagaacggc tcatgttcat gggaaacttca atctgggtgtt gaagtggagag agccatgtta 900
tcatgccaac atccgttccca agccggatcc aagaacccccc tggactaatg atgatccacg 960
tttcatgttgc agaggaggaa tgggttggaa cgactttatgtt cagaagatgtt ccaataactt 1020
ctatgttgc aacacccctt aagtttgcgtt catcttgcgtt atctcccaac ctcaggagcc 1080
tgagctttagt aatgttccacc cttcttcttcc accaagtctt tctcgttgc tcaaccttgg 1140
cccgctgtcc aatcccttcatg cttaaccatcc tgcactttccat ttcttgcgtt tgatcgaaa 1200
ggcgagttt gggaaagggtt ttctgttgcg acacaaggccg gaagaagttt tctatgttgcgtt 1260
caaagtttta cagaaggaaag caatcttgcgaa aaagaaaggag gagaaggatata ttatgttgcg 1320
ggcgaaatgtt ctgttgcgtt atgttgcgtt ccctttcttgc gtggggccctt accttctttt 1380
ccagactgtt gacaaatattgtt acctttgttcc agactatcatgtt aatggggag agttgttcta 1440
ccatcttcccg agggaaacgtt gcttcttgcg accacccgggtt cgtttctatgtt ctgttgcgtt 1500
agccgttgcg ttgggttccatc tgcattttactt gaaatgttgcgtt tataatgttgcgtt 1560
gaatattttt ttagattttccatc agggacacat tgccttactt gacttggac tcttgcgtt 1620
gaacatttgcg cacaacacggc caacatccatc ctcttgcgttcc acggccgggtt atctccgttcc 1680
tgagggttgcgtt ctaaagcgcg cttatgttgcgtt gacttggac tgggttggcc tggggactgtt 1740
cttgcgttgcgtt atgttgcgtt gcttgcgttcc tttttatgttgcgtt cggaaacacatgtt 1800
cgacaacatccatc tgcataatgttgcgtt gaaatccaaat attacaatttccgcaacatgtt 1860
ccctccttgggg ggccttccgttcc agaaggacacgg gacaaaggccg ctcggggccca aggatgttcc 1920
catggagatgtt aaggttgcgtt tcttcttccatc cttatgttgcgttcc tgcattatataa 1980
gaagatttttccatc aacccatgttgcgtt gatgttgcgttcc aacggacttccatc ggcacttttgcgtt 2040
cccccggatgtt accggaaacggc ctgttcccaatc ctccatgttgcgttcc aagtcccttgcgtt 2100
cgtcacacggc aagcttgcgttcc ggccttccatc ggcctttccatc atgcgttccatc 2160
caaggacttccatc ttcgttgcgttcc cccttgcgttcc gcttgggtttt aaaggattttt atgttgcgtt 2220
ccgaatgttccatc tagttttccatc tttttttccatc ccggccatgttcc acggacatccatc ttacaagaga 2280
atttgcacatccatc ttgttgcgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 2340
aagagccatccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 2400
gttgcgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 2460
gaaaggccgac tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 2520
tattatgaaatccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 2580
gttgcgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 2640
tatgcgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 2700
aacgttggacatccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 2760
tgcgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 2820
atgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 2880
ccaaatgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 2940
taaaatgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 3000
aagaacgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 3060
atgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 3120
cttgcgttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 3180
tttgaaaaatccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 3208

```

SEQ ID NO: 7 moltype = DNA length = 5758
 FEATURE Location/Qualifiers
 source 1..5758
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 7

```

tttctgtact ctgggttactt cagagaggaa agagatttgcgttcc acggccatctt ccttcgttcc 60
aaacattactt ctacttgcgttcc acggccatctt cagagaggaa agagatttgcgttcc acggccatctt 120
ggtgcgttcc cccacccagg gcttgcgttcc gggcccttggcc ctteccctgg gccaattttt 180
gggccttgcgttcc acggccatctt ctttgcgttcc acggccatctt cagatgttcc gggcccaatgtt 240
cctggacccctt ccacttgcgttcc acggccatctt ctttgcgttcc acggccatctt ctttgcgttcc 300
gaaggccatgttcc acggccatctt ctttgcgttcc acggccatctt ctttgcgttcc acggccatctt 360
gacatccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 420
cctcccccggaa tttttttccatc tttttttccatc tttttttccatc tttttttccatc tttttttccatc 480
tttggggccatgttcc acggccatctt ctttgcgttcc acggccatctt ctttgcgttcc acggccatctt 540

```

-continued

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| atgcacccaa | gccagccggg | ggccctcata | ccaggtgatc | cgcaggccat | gagccagccc | 600 |
| aacagaggtc | cctcacctt | cagtctgtc | cagctgcatc | agcttcgagc | tcaagattta | 660 |
| gcttataaaa | tgttgtggcg | aggccagccc | ctccccgaaa | cgtcgagct | tgcaagtcag | 720 |
| ggggaaaaagg | cgttgcctgg | cttgagcaa | caacagcagc | agcaacagca | gcagcagcag | 780 |
| cagcagcagc | agcagcagca | gcagcaacag | cagccgcagc | agcageccgc | gcaaccacag | 840 |
| acgcagcaac | aacagcagcc | ggcccttgtt | aactacaaca | gaccatctgg | cccgggggcg | 900 |
| gagctgagcg | gccccgagac | cccgagaag | ctgcccgtgc | ccgegeccgg | ccgcggggcc | 960 |
| tcggccgcgc | cccccgccgc | cgcgagccg | ccccggggcg | cagtgeccgg | gcccctcaigt | 1020 |
| ccgcageccg | ccccggggca | gcccctgccc | gtcttccagc | tgcageagaa | gcagagccgc | 1080 |
| atcagccccca | tccagaaacc | gcaaggccgt | gaccctgtgg | aaattctgca | agagcgggaa | 1140 |
| tacagactt | aggccgcat | agetcataagg | atacaagaac | tggaaaatct | gcctggctt | 1200 |
| ttggccaccc | attaagaac | caaaagcaacc | gttgaactaa | aagacttgc | gttactcaat | 1260 |
| tttcagegtc | agctgagaca | ggagggtggt | gcctgcattgc | gcagggacac | gaccctggag | 1320 |
| acggcttc | atcccaaaagc | atacaacacgg | agcaagccgc | agactctgag | agaagctcgc | 1380 |
| atgaccgaga | agctggagaa | gcagcagaag | attgagcagg | agaggaaacg | ccgtcagaaa | 1440 |
| caccaggaaat | acctaacaag | tattttgca | catgcääaa | atttttaagga | atatacatcg | 1500 |
| tctgtggccg | gaaagatcca | gaagctctcc | aaagcagtgg | caacttggca | tgccaaacact | 1560 |
| gaaaagagagc | agaagaaagg | gacagagcgg | attgaaaagg | agagaatgcg | gcgactgtat | 1620 |
| gctgaaatgt | aggagggtta | tagaaaact | attgatcaaa | aaaagacag | gcgtttagct | 1680 |
| taccttttc | agcagaccga | tqagtatgt | gccaatctga | ccaatctgtt | ttgggagcac | 1740 |
| aagcaagccc | aggcagccaa | agagaagaag | aagaggagga | ggaggaagaa | gaaggctgag | 1800 |
| gagaatgcg | agggtgggg | gtctgcccgt | ggaccggatg | gagagcccat | agatgagagc | 1860 |
| agccagatga | gtgacccccc | tgtcaaaatg | actcacacag | aaaccggcaa | ggttctgttc | 1920 |
| ggaccaggaa | cacccaaagc | aagtcaatgc | gacgcctggc | tggaaatgaa | tcctgttat | 1980 |
| gaagttgccc | ctagatctga | cagtgaagag | agtgattctg | attatgagga | agaggatgag | 2040 |
| gaagaagagt | ccagttaggc | ggaaccgcgg | gagaaaatac | tcctggatcc | aaatagcgaa | 2100 |
| gaagtttgc | agaaggatgc | taagoagat | attgagacag | ctaagcaaga | cgtggatgat | 2160 |
| gaatacagca | tgcagatcac | tgccagggg | tcccagtc | actacacgt | ggctcatgccc | 2220 |
| atctcgaga | gggtggagaa | acagtctgc | ctctctaatta | atgggacct | aaagcataac | 2280 |
| cagctccagg | ggcttggaaat | gatggttcc | ctgtataata | acaacttggaa | cggaatctta | 2340 |
| gcgcgatgaa | ttggggcttg | aaaggaccata | caagaccattt | cactcatcac | ttatctgtat | 2400 |
| gagcacaatgg | gactcaatgg | ccccatctc | atcttgc | ccctttcgac | tctatctaa | 2460 |
| tggacatatg | aatttgacaa | atgggctct | tctgtggta | agatttctt | caagggtact | 2520 |
| cctgccatgc | gtcgctcc | tgtccccca | ctacggagt | gcaatttca | tgtctctt | 2580 |
| actactatgt | atgatattat | aaaagacaca | cacatttctt | caaaggatcg | gtggaaatac | 2640 |
| atgatagttg | acgaaggcga | ccgaatgg | aatcaccat | gcaagctgac | tcaggctt | 2700 |
| aacactact | atgtggcccc | cagaaggatc | ctcttgc | ggaccggcgt | gcagaataa | 2760 |
| ctccctgaac | tctggggcc | cctcaactt | ctcttccaa | caattttaa | gagctgcagc | 2820 |
| acatttgc | aatgggtca | tgctccat | gcccattact | gtgaaagggt | ggacttaat | 2880 |
| gaagaagaaa | ctatattgt | catcggcgt | ctacataagg | tgtttagacc | atttttacta | 2940 |
| aggagactga | agaaaagatg | tgaatccag | cttccggaa | aagtggata | tgtgtatca | 3000 |
| tgtgacatgt | cgactctgc | gaagattctg | tatgc | tttgcata | tgcaagccaa | 3060 |
| ctcacagatg | gttctggagaa | agataagaag | ggggaaaggag | gtgttaaag | actttatgaa | 3120 |
| actattatgc | atgttgtggaa | aatctgc | acttcgaa | tttttgc | tttttgc | 3180 |
| tccttgcgt | acacccatgg | ctattcaat | ggggatcatca | atggggctga | actgtatcg | 3240 |
| gcctcaggga | attttgagct | gcttgcgt | attctgc | atttgc | atttgc | 3300 |
| cagatgtgc | ttttctgc | gatgcacat | ctcatgc | ttatgc | ttatgc | 3360 |
| tttcggaaat | tcctttactt | acgcctgt | ggcaccacca | agtcgttgc | tcgtgtgc | 3420 |
| ttgctgaaga | aattcaatga | acctggatc | cagtatttca | tttttgc | tttttgc | 3480 |
| gtctggggcc | ttggggctt | tettcaggca | gctgatac | tttttgc | tttttgc | 3540 |
| tggatcttc | atcaggat | gcggccca | gaccggatc | accgcatttgc | tttttgc | 3600 |
| gagggtccgg | tactggggct | ctgtac | aaacgcgtt | ggggatcc | tttttgc | 3660 |
| gcaaaataca | agcttaacat | ggatc | tttttgc | tttttgc | tttttgc | 3720 |
| tcttcgaacc | acgagccgg | ggcattctg | caggccat | tttttgc | tttttgc | 3780 |
| gagggaaag | atgaagtacc | ggcgcgt | actctgaa | tttttgc | tttttgc | 3840 |
| gaagaat | tttttgc | tttttgc | tttttgc | tttttgc | tttttgc | 3900 |
| ccgaaacccg | acccccgtt | aatgggg | tttttgc | tttttgc | tttttgc | 3960 |
| gacgctgaa | tagaaaaggct | cacctgt | tttttgc | tttttgc | tttttgc | 4020 |
| tcggccacgc | gcgcgtacgt | ggatcact | tttttgc | tttttgc | tttttgc | 4080 |
| cccatcgac | acggcaat | ggggat | tttttgc | tttttgc | tttttgc | 4140 |
| agacgaaag | atgtgtat | tttttgc | tttttgc | tttttgc | tttttgc | 4200 |
| agaggccgc | ctcccgct | tttttgc | tttttgc | tttttgc | tttttgc | 4260 |
| aacgctatca | tcgatactgt | tttttgc | tttttgc | tttttgc | tttttgc | 4320 |
| acttgcata | tttttgc | tttttgc | tttttgc | tttttgc | tttttgc | 4380 |
| tttcatttgc | atgttgtt | tttttgc | tttttgc | tttttgc | tttttgc | 4440 |
| gatttcaaaa | aaataaaagg | tttttgc | tttttgc | tttttgc | tttttgc | 4500 |
| gagaaggatg | tcatgttct | tttttgc | tttttgc | tttttgc | tttttgc | 4560 |
| atctatgt | tttttgc | tttttgc | tttttgc | tttttgc | tttttgc | 4620 |
| aaagaggg | agagtggag | tttttgc | tttttgc | tttttgc | tttttgc | 4680 |
| tcagatgtcc | tttcaggaa | tttttgc | tttttgc | tttttgc | tttttgc | 4740 |
| ggccgggaca | aaggaaagg | tttttgc | tttttgc | tttttgc | tttttgc | 4800 |
| agcgatttt | acagcgat | tttttgc | tttttgc | tttttgc | tttttgc | 4860 |
| gatgtatgt | tttttgc | tttttgc | tttttgc | tttttgc | tttttgc | 4920 |
| tctcatttct | acccagt | tttttgc | tttttgc | tttttgc | tttttgc | 4980 |
| atcgctata | aactagttt | tttttgc | tttttgc | tttttgc | tttttgc | 5040 |
| aacacacaca | tacacaaata | tttttgc | tttttgc | tttttgc | tttttgc | 5100 |

-continued

| | | | | | | |
|------------|--------------|-------------|--------------|-------------|-------------|------|
| attgaaacaa | acaaaaagct | tttgcgtggaa | aatatgtggg | tggatgttat | atttctatgg | 5160 |
| gtgggtctaa | tttgcgtacg | gtttgtatgt | gcctgggttt | atcacctgtt | cagatgagaa | 5220 |
| gatTTTGTc | tttgcgtac | ctgataacca | ggagaagccaa | ttaaaagccaa | ctgggttattt | 5280 |
| tatTTTCTat | caggcaattt | tcgggtttt | tatttgcgtc | gtattgtttt | tttacactgt | 5340 |
| ggtacatata | agcaacttta | atagggtata | aatgtacagt | agtttagattt | cacctgcata | 5400 |
| taCATTTC | cattttatgc | tctatgtatct | gaacaaaaaagc | tttttaattt | gtataagatt | 5460 |
| taTGTCTact | gtaaaacattt | cttaattttt | ttgtcttgc | ttaaaaaaaa | agttttgtt | 5520 |
| aaAGCGCTat | tgaatattgc | aatctatata | gtgtattggaa | ttggcttctt | tgtcaccctg | 5580 |
| atctctatg | ttacaatgt | gtatcgctc | cttctcccta | aagtgtactt | aatctttgt | 5640 |
| ttctttgcac | aatgttttttgc | gttgcgtac | ataaggctgaa | ggcacaataaa | attccagtaa | 5700 |
| tttcgaagaa | tttgcgtgttgc | gtgttttctt | aataaagaaa | taattttagt | tgacaaaaa | 5758 |

SEQ ID NO: 8 moltype = DNA length = 837
 FEATURE Location/Qualifiers
 source 1..837
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 8

| | | | | | | |
|-------------|-------------|-------------|--------------|-------------|------------|-----|
| gtccctctcg | cacacccc | tcgcgtctcc | acaccactgg | caccaggccc | cggacacccg | 60 |
| ctctgtcga | ggagaatggc | tactcatca | acgcgtgtgg | ttggactggc | cctgtctggg | 120 |
| gtgtgtggcg | acctgcaggc | agcaccggag | gcccaaggct | ccgtgcagcc | caacttcag | 180 |
| caggacaatg | ttctggggcg | cttgcgtcago | gcgggcctcg | cctccaaactc | gagctggctc | 240 |
| cgggagaaga | aggcggcggt | gttccatgtgc | aagtctgtgg | ttggccctgc | cacggatgtt | 300 |
| ggccctcaacc | tgacatccac | tttcttcagg | aaaacccaggat | ttgtgacccgg | aaccatgtct | 360 |
| ctgcageccg | ccccggccct | cggtcttac | agttaccggaa | gtccccactg | gggcagcacc | 420 |
| tactccgtgt | cagtgtgtgg | gaccgactac | gaccgtacg | cgctgtgt | cagccaggcc | 480 |
| agcaaggccc | cttgcggagga | cttgcgtatg | gcacccctct | acagccgaac | ccagaccccc | 540 |
| aggggctgagt | taaaggagaa | tttgcgtccgc | tttgcgtccgc | cccaggctt | cacagaggat | 600 |
| accattgtct | ttctgtccca | aaccgtataag | tgcatgtacg | aacaatagga | ctcccccagg | 660 |
| ctgaagctgg | gtatcccggc | agccagggt | ccccacgc | ctggatgtt | ctgtctgtt | 720 |
| ccttcggccca | gcccgtccgc | cggtccccc | ccaaagcaac | cctgcggact | cagggttcat | 780 |
| cctgcacaat | aaactccgg | agcaagtca | taaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | 837 |

SEQ ID NO: 9 moltype = DNA length = 6001
 FEATURE Location/Qualifiers
 source 1..6001
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 9

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| caaggaggga | tcccacagat | gtcacaggc | tgtcacagag | ctgtgggtgg | aatttccat | 60 |
| gagacccccc | ccctggctga | gtcacgcac | tccgtgtttt | gacctgaatg | cctctcgagc | 120 |
| tgcagaagcc | tgaagaccaa | ggagtggaaa | gttctccgc | agccctgaga | tctcaagagt | 180 |
| gacatttgt | agacccgtca | atttgtttaa | aatttcttgc | aatcgtacgtt | tgctgtatc | 240 |
| atacctgtgc | cagatttcat | catggaaac | agttgttaca | aatcgtatgc | cactctgtt | 300 |
| ctgtcttca | acttttgagag | gacaatgtca | ttgcaggatc | tttgcgtttaa | ctgcccagct | 360 |
| ggtacattct | gtgataataa | caggaatcat | atttgcgtc | cctgtctcc | aaatagttt | 420 |
| tccagcgcg | gttggacaaag | gaccgtgtac | atatgcggac | agtgtaaagg | tttttgcagg | 480 |
| accaggaaagg | agtgttcc | caccagcaat | gcagagtg | actgcactcc | agggtttcac | 540 |
| tgcctggggg | caggatgcag | catgtgtaa | caggattgt | aacaagggtca | agaactgaca | 600 |
| aaaaaaagggt | gtaaagactg | ttgttttggg | acattttaacg | atcagaaacg | ttggcatctgt | 660 |
| cgaccctgtg | caaactgttc | tttggatgtt | aaagtctgtc | ttgtgtatgg | gacgaaaggag | 720 |
| agggacgtgg | tctgtggacc | atctccagcc | gacctcttc | cgggacatc | ctctgtgacc | 780 |
| cgcctcgtcc | ctgcgtcgaga | gcccggac | tctccgca | tcatctc | ttttttgcg | 840 |
| ctgacgtcg | ctgcgtgtc | tttctctca | cgctccgtt | ctctgttgc | ttttttgcg | 900 |
| aaacggggca | gaaagaaact | cctgtatata | tccaaacac | catttatgag | accagtacaa | 960 |
| actactcaag | aggaaagatgg | ctgtatgtc | cgatccatgc | aagaagaaga | aggaggatgt | 1020 |
| gaactgtgaa | atgggctgt | ttggactttc | ttggactttc | ttgaaaaaaa | gcaaggaaat | 1080 |
| atgagtcatc | cgctatcaca | gtttcaaaa | gcaagaacac | catectacat | aataccagg | 1140 |
| atccccccaa | cacaacctt | tttcttaatg | ccaatgtat | ggccctttaaa | aatgcaccc | 1200 |
| ttttttttttt | tttttgacag | ggtctcact | ttgtcaccc | gttggaggtc | agtggcacca | 1260 |
| ccatggctct | ctgcgtcc | gacccctgg | agtcgtatgt | atccctctgc | ctcgtctcc | 1320 |
| tgagtagctg | gaactacaag | gaaggggccac | cacacgtac | taacttttt | ttttttgtt | 1380 |
| tggtaaagat | ggcatttcac | catgtgtac | agggtgtt | caaactcc | ggttcac | 1440 |
| ggcctcccaa | agtgtgg | tttgcgtat | taacacgaca | tgctccaccc | tcagggggga | 1500 |
| cacttttaac | agaacacaga | gttggacaca | gatgtgtt | aaaaaaa | aaaaaaaaaa | 1560 |
| cattttctat | ataccatca | acagggtt | gtatgtttt | ttgaaatcca | aagaaaatta | 1620 |
| tagtttaat | tcaattacat | agtccgttgc | tccaaactata | attataatca | aatcaatgc | 1680 |
| aggtttgtt | tttgcgtat | atatgcata | tgacaataag | ccacgggtt | cagtaagtac | 1740 |
| ccgactaaag | tttccgtgg | tttgcgtat | taacacgaca | tgctccaccc | tcagggggga | 1800 |
| gtatgacgt | agtgcgttgc | tttgcgtat | aggacaaaaaa | acctcaggcc | ttggaggt | 1860 |
| ttggaaaga | gttcaatgt | ctgtatata | tttgcgtat | tccatctca | cacccctgc | 1920 |
| cttgcgtcc | ctcccttta | agccagggtt | cattctaaa | attttaact | tttacataaa | 1980 |
| tatTTTATAC | caaagccat | aatgaactg | catatgtat | gtatgtatgt | cagtggaaaa | 2040 |
| attaacaccc | gtgacgtat | tttgcgtat | cgacactaga | gtggggccg | ccaaactccc | 2100 |
| atggccaaac | ctgggtcacc | atttgcctt | ttttgttgc | ttttttgtt | gagacatgt | 2160 |
| tgctctgttgc | ccagggttgc | aatgggttgc | ctattcacag | gcacaatcat | agccacactt | 2220 |

-continued

| | | | | | | |
|--------------|----------------|-------------|--------------|--------------|--------------|------|
| agccttaaac | tcctgggctc | aagtgtatcca | ccgcgcctca | tctccaaagt | agctgggatt | 2280 |
| acaggcgaa | acctggcatg | cctgcatttgc | tttggcttat | gatctaaggaa | tagctttta | 2340 |
| aatttttattc | atttttttt | tttttgagac | agtgtctcac | tctgtctcc | aggctggagt | 2400 |
| acagtggtagc | aatcttggat | caccgcctcc | cagttcaag | tgatctccct | gcctcagcc | 2460 |
| cctaagttagc | tgggactaca | ggtatgtgcc | accacgcctg | gctaattttt | atatttttag | 2520 |
| tagagacggg | gtttcaccat | gttgcagg | ctgggtctcaa | actcctgacc | tcagggtgatc | 2580 |
| tgcccaccc | tgccccccaa | agtgtggga | ttacaggcat | gagccacca | gcctggccat | 2640 |
| tttttacact | tttgtatgac | atgectattt | caagcttgc | tgccctgtc | ccatgttatt | 2700 |
| ttactctggg | atttaggtgg | agggagcago | ttctatttttgg | aacattggcc | atcgcatggc | 2760 |
| aaatgggtat | ttgtactatc | tgcttctat | tagttggatc | tactataacc | tttagagcaa | 2820 |
| atccctgacg | aaagccaggg | atcaatagggg | tattctgtaa | ataggggtga | 2880 | |
| ggagaagata | tttctgaaca | atagtctact | gcagttacca | attgttttc | aaagtggctg | 2940 |
| ttctaatgt | ctcccgctag | tcataatagt | gtcatgtaa | tatcccattt | atccacatcc | 3000 |
| ttgcttacct | ctggactatc | cagggtccct | taattttgc | aagccagtttgg | gtatagaatg | 3060 |
| agatctca | tggttactt | tttgcatttgc | tttgcatttgc | tgatgcaccc | ttgtcaaaaata | 3120 |
| tttataatacc | attttgtttt | ataaaatgt | tgctcatgt | tttttgc | tttttgc | 3180 |
| tttgcaaaaaa | aacttggggc | cgggtgcagt | ggctcatgc | tgttagtccca | gctttttggg | 3240 |
| aggcccaaggt | gggcagatcg | cttgcggcc | ggagttcag | accagcc | gcaacatggc | 3300 |
| gaaacccttcc | tttttttttttttt | aaataaaaaaa | ttagccgggt | tttaggtgggt | gcacctgaag | 3360 |
| tcggcggat | tcggatgtt | cggttgc | ctgggaggca | gagggtgcag | tgagctggga | 3420 |
| ccgcatca | acacttc | ctgggcaaca | gagaaaaacc | tttttcoaga | aacaaacaaa | 3480 |
| ccaaatgt | gttgttgc | ctgttccat | aaaggctttt | atgtatttca | gataataatc | 3540 |
| tttggtca | gttatgttt | aaaaatata | ttctttgtgg | ccaggcacgg | tagtgcacac | 3600 |
| ctgtatccc | aggacttgc | ggggctgagg | ttgggtggatc | atctggatc | aagagttca | 3660 |
| gatcagectg | gccaacacag | tgaaaaacc | tcttctactaa | acatgtacaa | aacttagctg | 3720 |
| ggtatgtgg | cggttgcctg | taaacc | tgctccagag | gtgtggccag | aagaatcgct | 3780 |
| tgaacccagg | aggcagagg | tgca | caagatgttgc | ccatgttgc | ccagacttggg | 3840 |
| tgacaagagt | gaaatttctgc | ctatccat | atcttatcat | ctatatat | atatatataat | 3900 |
| atatatatcc | tttgcatttt | atttttccct | ttttttttttt | ttttttttttt | ttttttttttt | 3960 |
| tttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | 4020 |
| tttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | 4080 |
| tttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | 4140 |
| tttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | 4200 |
| tttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | 4260 |
| tttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | 4320 |
| tttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | 4380 |
| taaccactgg | ttccctgc | ttttttttttt | ttttttttttt | ttttttttttt | ttttttttttt | 4440 |
| ctcaccttacc | atccatca | tttccatagg | gccatgggc | tgttttgc | ctccctgtt | 4500 |
| tgttccatttgc | tatccatgc | tatccatgc | cgatcatc | tgttttgc | tacaatag | 4560 |
| ttgttaac | atccatgc | ttgttgc | tctccctcc | ccttctttt | ctacttc | 4620 |
| agtgctttag | ctaggcttgc | cacgggttgc | cacgcctgt | atccacacac | tttgggg | 4680 |
| cgacgcggat | ggatcacctg | aggttgcagg | tttgcggaca | gcctggccaa | catgtgaaa | 4740 |
| ccccatcttct | actaaaaat | acaaaaata | gtcaggcat | gttgcattgt | cctgtatcc | 4800 |
| cgatcttttg | ggaggcttgc | ggccggaa | tgttttgc | ccgggggggg | agggtgc | 4860 |
| gagccgagat | cgatccatttgc | cactccgc | ttgggtgc | agcgcacac | tgttgc | 4920 |
| aaaaaaagaa | aagatgtc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 4980 |
| gaagctgtat | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5040 |
| agatgttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5100 |
| tatttttttt | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5160 |
| gccatccat | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5220 |
| agtttccat | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5280 |
| actatccaca | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5340 |
| ggacgcggat | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5400 |
| gaaaggaggc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5460 |
| catgtgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5520 |
| taaatgcac | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5580 |
| taatttttttt | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5640 |
| aataaaat | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5700 |
| agotcagg | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5760 |
| tgacggaa | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5820 |
| aagatgttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5880 |
| ctgtgtgttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 5940 |
| aataaaacac | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | tttgcatttgc | 6000 |

SEQ ID NO: 10 moltype = DNA length = 1336
 FEATURE Location/Qualifiers
 source 1..1336
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 10
 gagagacaca gagtccggca ttggtccag gcagcgttta gcccgcggcc cgcctgttg 60
 tccccagacatggaga gccagtctga tccagaaggca caagctggca gaggcaggcc 120
 aacgttatgttgc ggacatggca gccttcatgt aaggcgcggct ggagaaggcc gaggagctct 180
 cctgcgaaga gcgaaacatgc ctctcgttgc cttataaaggaa cttgtgtggc ggcaggagg 240
 ctgcctggat ggtgtgttgc agtattgttgc agaaaaggca cgaggaggcc tcggaggaga 300

-continued

| | | | | | | |
|------------|------------|-------------|--------------|------------|-------------|------|
| ggggggccga | ggtgcgtag | tacccggaga | aggtgtggagac | tgagctccag | ggcgtgtgcg | 360 |
| acaccgtcg | gggcctcg | gacagccaco | tcatcaagga | ggccggggac | gccgagagcc | 420 |
| gggttttcta | cctgaagatg | aagggtgact | actaccgcta | cctggccgag | gtggccacccg | 480 |
| gtgacgacaa | gaagcgcatc | attgactag | cccggtcagc | ctaccaggag | gccatggaca | 540 |
| tcagcaagaa | ggagatgccg | cccaccaacc | ccatccgccc | gggcctggcc | ctgactttt | 600 |
| cegtctcca | ctacgagatc | gccaacagcc | ccgaggagggc | catctctcg | gccaagacca | 660 |
| cttgcacga | ggccatggct | gatctgcaca | cctcagcgcg | ggactctac | aaagacagca | 720 |
| ccttcatcat | cgagctgctg | cgagacaacc | tgacactgtg | gacggccgac | aacgccccgg | 780 |
| aagagggggg | cgaggctccc | caggagcccc | agagctgagt | gttgcceggc | accgccccgc | 840 |
| cctgccccct | ccaggtcccc | accctggcc | gaggactag | atgggttggg | aggccccacc | 900 |
| cttctccccc | aggcgctt | cttgcacca | agggtcccg | ggagaggggc | tggcagagct | 960 |
| gaggccac | gggggtgggg | atccccactt | tcttgcacgt | gttgcgcga | ccttaaccact | 1020 |
| ggtcatgcc | ccacccctgc | tcttcgcacc | cggttctcc | cgacccccc | accaggctac | 1080 |
| tttctccctc | ctcttcgcct | ccttcgtccc | ctgtgcctc | tgatctgagg | aattgaggag | 1140 |
| tgtcccccet | tgtggcgtag | aactggcag | tggcaggggc | tggagatggg | tgtgtgtgt | 1200 |
| tgtgtgtgt | tgtgtgtgt | tgtgcgcgc | cgccagtgca | agaccgagat | tgaggaaag | 1260 |
| catgtctgt | gggtgtgacc | atgtttccctc | tcaataaaagt | tccctgtga | cactcaaaaa | 1320 |
| aaaaaaaaaa | aaaaaaa | | | | | 1336 |

SEQ ID NO: 11 moltype = DNA length = 2240
 FEATURE Location/Qualifiers
 source 1..2240
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 11
 ggaggggcgc cagcagcttc cccttctctg ccctgcttca ggcaccaggc tctttccct 60
 tcagtgttc agaggagggg acggcagcac catggaccac cggttgcaca ctgtccgcca 120
 gaccctgtgc ttgttcaatg tccgcacatcg aaccaccgcg ctggccatct accatgtgat 180
 catgagtcgtc ttgtgttca tcgagcactc agtagagggtg gcccattggca aggggtctcg 240
 caagctctcc cagatgggc acctcaggat cgctgacccg atctccagct tcctgctcat 300
 caccatgtcgc acgttacatca gcctgacccg actgatccggc gtatgtcaaga accgggagaa 360
 gtacacctgtc cccttctgt ccctgcacaaat catggactat ctccctgtgcc tgctcaccc 420
 gctgggtctcc tacattgagc tgccccccta cctcaagttt gcctccggc gccgtgtctag 480
 ctcttcctcaag ttccccctga tgacgctgcg gctgctggc ttctgcctga gcatcctgtac 540
 cctctgcagc ttcttacatgg aagtggccac statctcaac ttcaagtgcg tgaaccacat 600
 gaattttccctc cccagccagg agggatgtcc tcataaccag ttcatcaaga tgatgtatc 660
 cttttccatc gcttcatca ctgtccttat cttcacggcc tacatgttca agtgcgtgt 720
 gccgtgtctac agattgtatca agtgcacggc ctcgggtggag gagaagagaa acttcaagat 780
 gctcccaaaag gtgggtctcg cgttgcacgtc ggaaggccctc tctttgcacat cgaagacccc 840
 agaggggggc cccggccaccc cccatactc agagggtgtg ccctgcgcag gcccggcccc 900
 cagtgcgtgg agggttggag ctgcctcata atctgccttt ttgccttggt ggcctctgt 960
 gcctgggtgg gcctcccccgc ccctccctgg caggacaatc tgcgtgtgc tccctcgctg 1020
 gcctgcctc ctgcggggcc ctgtgcacgtc ctcaacactg ggttgcacgtc ttgggtgt 1080
 tcacttcctcg ggtcttcaca taatttgcgca caacaatgtc tggttattt caatcagctc 1140
 tgacacttgtt ttagacgatt ggccatctca aatgtgggtgaa ttgtgtcaag caactatcg 1200
 cttgtatcgt tcagccaaacg aactgacaaa tcaaaaaccctt acttgcgttgcgt tcaatcaa 1260
 aattttgttca aacaacatgc tatttgcattt atttataaat agtgcgttgcgt tccatagca 1320
 atttaatcaa gtaatcatta atttgcattt ccctatataat aatataatgt aatcaatcc 1380
 ttcaaatagc ttgccttacat gataatcaat tagccaaaccat tggtgcattt agaatatgt 1440
 taaaatagaat acacagaataa gtgtatggaaat tcaattttaaa aaatcagctt agcctccaaa 1500
 ccatttaat caaatgtaaac catcaactgg atgccaacttc tggcgaatgtt aggactctcg 1560
 agtggctgttca attttgttcaatggaa ttcaatggaa aacttgcattt aatcattttaaa 1620
 caacaatccat ccggccggaa taacgttattt ttcatcatataa aacatccctc tcaaaacac 1680
 aattttgtctcg ctgcggggat ttgcacggat atccatgttgcgtt caccattttca gttgtctgt 1740
 ggttcgtgtg gtgttcataac agtggatcc atgaaaggag tcatttttgtt gatgagctgc 1800
 cagtccttc ccaggccggc ctgtgcgttgc cccatccatc agtgcgttgc gtcataggcg 1860
 aatctgttcc tggccgggtt gttgttgcggg cccatccatc agtgcgttgc gtcataggcg 1920
 tggcgtgtgg actaaaccctt caggatgtttt caggatgttgcgtt cccatccatc aatcattttaaa 1980
 ctgcggccatc cactggggc atggccgttgc tggcagctgg ggagggtttgg ccccaacacgc 2040
 cctccctgtgc ctgtttttttt gttgttgcggg gtccctccagg gaggtgcacc agagggtggag 2100
 gccacggggc cagggttctt gggggacttc ggggggttaca gagggtggaaat gcttcgttgc 2160
 agtccctggg caataaaaaaa ttgtgttattt gttgttgcgtt cggccggggat gaaatggat 2220
 taaaacggca acaagcttctt

SEQ ID NO: 12 moltype = DNA length = 3039
 FEATURE Location/Qualifiers
 source 1..3039
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 12
 agggcgcagag gggggcggtt ttggacccggg cggagccggc ggacccctag gtggcgagg 60
 gacgcctccgg gaaagcgagg ggcgcatacg aacttgcggcc acgtgcaccc cgccgggggg 120
 gacgcgggggg cgcgcggggcc tcctgcgggtt cccctgcctt gggggggggc cgtgaccacc 180
 cgtctgtcgcc cgcggccggc cgcgcgttca ctttcacccgc gtacccggga cccggccggcc 240
 cgcggggaaat atgttgcgtt aacttgcgttgc gggggccatc agatgcacgg atttggata 300
 ctttttgcacat ctttgcacattt accttccttc atttataataa aagaagaatc 360

-continued

| | | | | | | |
|-------------|--------------|-------------|-------------|--------------|-------------|------|
| aggagcttag | gatgttataa | caccaactca | ttaatatact | aaccggacaa | tgttctacaa | 420 |
| acaattctac | attgttaaaagg | actggatttg | cacaaaataa | aataatttttta | tttttattcag | 480 |
| cttataatat | gactcgatgg | aggaaaattt | gataagcatg | agagaagacc | attcttttca | 540 |
| tgttcgttac | agaatggaaag | cttcttgctt | agagctggcc | ttggaaagggg | aacgtctatg | 600 |
| taaatcagga | gactgcgcg | ctggcgtgtc | attctttgaa | gctgcgttc | aagtggaaac | 660 |
| tgaagacctt | aaaacactta | gchgcttta | cagccagttt | ggcaatgttt | atttctattt | 720 |
| gcatgattat | gccaauagcat | tagaataatca | ccatcatgtat | ttaacccttg | caaggactat | 780 |
| tggagaccag | ctggggggaaag | cgaaaactgt | ttggtaatctg | ggaaacacact | taaaagtct | 840 |
| tgggatttt | gacgaaGCCA | tagtttgtt | tcagcgacac | ctagatattt | ccagagagct | 900 |
| taatgacaAG | gtggggaaag | caagagact | ttacaatctt | ggaaatgtgt | atcatgcCAA | 960 |
| aggggaaaatg | tttgggttgc | ctggggccca | ggatgttagga | gaatttccag | aagaatgtgag | 1020 |
| agatgtcttg | caggcagccg | tggatttttta | tgaggaaaac | ctatcatat | tgactgtttt | 1080 |
| gggtgaccga | cgcccacaag | gacgtgcctt | tggaaatctt | ggaaacacac | attacctcct | 1140 |
| tggcaacttc | agggtgtcag | ttatagctca | tgacgcgtgt | ctcccttattt | caaaaagaaat | 1200 |
| tggagataaa | cgacgtgaaa | gaagagacta | tagcaacctt | ggaaatgtcat | atataattttt | 1260 |
| tgtgttaattt | gaaactgcct | cggaactacta | caagaagacaa | ctactgttgg | cccgacacgt | 1320 |
| taaagacgca | gctgttataa | cacagtctt | ttacagtctt | ggaaataatcat | atactttact | 1380 |
| tcaagactat | aaaaaggccca | ttgattatca | tctgaagcac | ttagcaat | ctcaagagct | 1440 |
| gaatgataga | attgggttgaag | gaagagactg | ttggagactt | ggaaatgtcat | acacagcact | 1500 |
| aggaaatcat | gtcaagcaa | tgcattttgc | tggaaatctt | caagagaggt | ttggaaat | 1560 |
| tggggataaa | agtgggttgaac | taacagcagc | acttaatctt | tcagacctt | aaatggttct | 1620 |
| tggctgttgc | tacagcacca | ataactccat | aatgttctga | aatactgtt | ttgatagcag | 1680 |
| tttgaatgtt | tgacggccca | agttggggac | ccggcatat | atggaaaata | ttggacttat | 1740 |
| gaagtttaca | ccagaaaaagg | ttagacactg | gaacagtgtt | ttttttgtct | agccaaaaacc | 1800 |
| tcttatttgc | aaaccttctg | caaagctact | cttgcacaa | agactgtt | ggaaaaataa | 1860 |
| caaaacaaat | ttcttccacta | aagtgttcc | agatgttcc | aatttctat | accaccgaat | 1920 |
| tccaaatttct | cagaggaaaa | tcagtgccaa | taacttgg | gatggagggt | tctttgtactt | 1980 |
| attaaggccg | tttccaaagca | ataggatgg | tgatcgagaa | tgttgcatt | aagaaaaagaa | 2040 |
| ctgccatata | gtttcaacaa | caacttctt | cactccccct | aaaatgttc | taaaaacatc | 2100 |
| atctgttctt | gtggatattcc | ccaaacacgg | ttagtttttgc | gttgcattt | ccagctcaca | 2160 |
| gagtgcgtgt | ctggatgtcc | agagggtgtt | tttcagttat | tttgcagggt | ttcgcttcaac | 2220 |
| acaacacgc | cagtccgtac | tttgcacccat | gatgactaat | tttgcagggt | tttgcagggt | 2280 |
| agattttctt | gacatcctt | taaaaatgtca | aggatccaga | ttagatgtat | aaagatgtc | 2340 |
| tccaccacat | gttaccacaa | agggtccgac | agtaccat | gaagactttt | tcagccttat | 2400 |
| tttacgggtcc | caggggaaaag | gaatgttgc | acagagact | cttttacaaa | gagatcaaaa | 2460 |
| cagagacact | gactttgggc | taaaggactt | tttgcaccaat | aatgttttgc | tttgcaccaat | 2520 |
| aaattcagg | aaaaaaatcg | cagaccat | tttgcaccaat | tttgcaccaat | tttgcaccaat | 2580 |
| aacacggtaa | ggaaacaatc | tattactttt | tttgcaccaat | tttgcaccaat | tttgcaccaat | 2640 |
| atacagctt | aaatattttt | agaatgtat | aaatgttgc | tttgcaccaat | tttgcaccaat | 2700 |
| cattaatact | tctctggaca | tgccgtgttgc | aggggtgggg | tttgcaccaat | tttgcaccaat | 2760 |
| cgtctgtgt | tactgttgc | gatgttgc | tttgcaccaat | tttgcaccaat | tttgcaccaat | 2820 |
| gtttataaag | taggaatgtt | agtgttgc | tttgcaccaat | tttgcaccaat | tttgcaccaat | 2880 |
| attttttac | atcttaattt | acaatgggt | tttgcaccaat | ataaccat | tttgcaccaat | 2940 |
| aaacaatgtt | tttttgtttaaa | aataatgtt | tttgcaccaat | taaaaaagta | tataaaatag | 3000 |
| tcttactaaa | aatcttagtt | ttttttctt | ccaaaaaaa | | | 3039 |

SEQ ID NO: 13 moltype = DNA length = 7018
 FEATURE Location/Qualifiers
 source 1..7018
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 13

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| ggcgccgcgc | ccggggggca | ggtgtcgccg | tcggcgccat | tcggcgccga | tggagcggcc | 60 |
| ctggggggact | ggggacggcc | tctcgctgt | gccccatggc | ctcgcccttc | tcctcttctt | 120 |
| gcagctgtgt | ccggcgctga | ccctcagcc | ggacggggcc | gacgcgcgc | ccggccgcgc | 180 |
| tgegcgcgt | ccgcgtgttt | ctggggccat | cggggtgtgc | ttgggggtgc | ggggggcgc | 240 |
| agccccggggc | gggtttttcc | ggggccggcg | ttggcggtgc | agcgcgcgc | gcgaggacga | 300 |
| ggagtgcggc | gggggtccgg | acttcgtgc | caagctggcc | aacaacacgc | accagcatgt | 360 |
| gtttgtatgt | ctcagggctt | cagtatccat | gtcctgggtt | ggagatagca | ctggggtcat | 420 |
| tctagtcttgc | actaccttcc | atgttaccat | ggtaattat | acttttgc | agtccaaat | 480 |
| atatcgaatgt | gaggatattt | ggaaaactt | taaggat | acagatctca | tcaataaacac | 540 |
| ctttatttgc | actgttgc | gtcatggat | ttgttgcgtt | aactctggaa | agggtgggtt | 600 |
| aacagcagat | gtgtctggag | gaagtcgttt | aggaaatgtt | tttgcaccaat | cagattttgc | 660 |
| gaagaattttt | gtgcaacac | atctcccttc | acttcgtat | tttgcaccaat | tttgcaccaat | 720 |
| tcagaatctt | gattatctt | tagtcgtat | cactgtttat | ggcctgtgg | tgttcaagaa | 780 |
| ttttggggaa | aatatggaaag | aaatcccaaa | agcgtatgt | tttgcaccaat | ggggatcaga | 840 |
| caacaccatc | ttctttacaa | cctatgc | tttgcaccaat | aaagcttgc | tttgcaccaat | 900 |
| ggaattatgg | agaacttcag | acttggaaa | aagcttcaa | acttgggtt | tttgcaccaat | 960 |
| ctcattttgtt | tttgcaccaat | tttgcaccaat | tttgcaccaat | tttgcaccaat | tttgcaccaat | 1020 |
| aagaaggatc | cacgtttca | cagatcaagg | ggacacatgg | agcatggcc | agctccctc | 1080 |
| cgtggggacac | gaacgtttt | attctattt | ggcagcaat | gttgcaccaat | tttgcaccaat | 1140 |
| tgttagatgtt | cctggggaca | ctgggtttgg | cacaatctt | accttcgtat | tttgcaccaat | 1200 |
| tgtctattcc | aaatcttttt | accgcacat | tttgcaccaat | acaggcggag | agacggactt | 1260 |
| taccaacgtt | acctcccttc | ggggcgctca | cataacaac | tttgcaccaat | tttgcaccaat | 1320 |
| tatccagacc | atgtatctt | tttgcaccaat | tttgcaccaat | tttgcaccaat | tttgcaccaat | 1380 |
| aaacacgtt | tttgcaccaat | tttgcaccaat | tttgcaccaat | tttgcaccaat | tttgcaccaat | 1440 |

-continued

| | | | | | | |
|---------------|---------------|---------------|---------------|---------------|---------------|------|
| tccctacagc | atctcccaga | aactgaatgt | tccaatggcc | ccactctcag | agccgaatgc | 1500 |
| cgttaggatt | gtcattgtctc | atggtagctg | gggggatgccc | atctcagtga | tggttccaga | 1560 |
| tgtgtacatc | tcatatgtat | ggggttactc | ctggacaaag | atgttggaa | gaccggacta | 1620 |
| ttacaccatc | ctggattctg | ggggatcat | tgtggccatt | gagcacagca | gccgtccat | 1680 |
| caatgtgatt | aaggatcca | cagacgaagg | tcaatgtcg | caaaccatac | cgttcaccag | 1740 |
| ggacccatc | tatccatgt | gcctagctc | agaaacctgg | gttaggtc | tgaatatcag | 1800 |
| catttgggc | tccacagaat | cttccatgt | cagccagtgt | gttccataca | ccattgtatt | 1860 |
| taaaagatatc | cttggaaagga | actgtgaaga | aaaggactat | accatatggc | tgccacactc | 1920 |
| cacagaccc | gaagattatg | aagatggctg | cattttggc | tacaaagaac | agtttctgcg | 1980 |
| gtcacgtca | tcatccgtgt | gtcaatgtgg | tcgagactat | gttggacca | agcagccctc | 2040 |
| catctgcctc | tgttccatgg | aggactatc | ctgtgattt | ggctactacc | gtccagaaaa | 2100 |
| tgactcca | tggtgtgaac | ggccagaact | gaaggggcc | gacctggat | tttgtctgt | 2160 |
| cggaaagagaa | gaacaccta | caacaatgg | gtaccggaa | attccagggg | acaatggca | 2220 |
| gggtggggta | aatccatgtc | gagaagtaa | agacttggaa | aagaatgtca | caagcaact | 2280 |
| ttttagtccg | gaaaacacaga | attccatgtc | aaatttctgtt | ccaaatattcc | tggccatctgt | 2340 |
| gggattgtat | ctggcacag | tcgtgacgg | agtgctatt | gtgaaagaa | atgtctgtgg | 2400 |
| gggaagggtc | ctgggtcatc | gatactctgt | gtcgacgg | catgcacagg | ccaaatgggt | 2460 |
| ggatgggtgt | gtatgtttgg | acacagcc | ccacactaa | aaaatgttt | atcatgtga | 2520 |
| ctcagatgt | gacccatgtt | aatagcttt | cagggatgt | ggaccggca | tggtgggtgg | 2580 |
| aaccacagta | ctcttacac | tccctgtgg | tccaaactca | ggaaataat | ttcccatatgc | 2640 |
| gaggggacca | gctctgttcc | tgctgcttcc | atcaaagcca | aaaggacca | cactaaagaa | 2700 |
| atgcagggtt | gggggtgggg | accctggac | cttttttaca | attggctctg | agaaaaagg | 2760 |
| agacatttta | aattctttaa | cttctttttt | ctcgctctgt | cttcttgc | atgtgggt | 2820 |
| ttttgtttt | tggtttttaa | ggggaaacaa | atggatattc | aaggggactt | ttcataacc | 2880 |
| ccacttctgt | gtgttctgca | ttggcctgt | cccagggtat | ctgccaactc | cgtatcage | 2940 |
| tcccacatgt | tacttggtac | cateccctgg | ctctgtgtgc | gagacgaaac | agctgttagag | 3000 |
| atgaaacacg | gtgtcgagg | ctggacacgc | ctggccgggt | tttcttcoatc | tggggacagt | 3060 |
| cctactccaa | gaacactgtca | caccacgtca | tccacacat | ccccctact | ctttttttt | 3120 |
| ttttcagaga | ccacagacca | cagtatgtt | tcttttccct | tggttaatta | ggcaatacc | 3180 |
| ttgtttaatgt | ccctttggca | actaaactaa | ccatgtgtc | cccacacat | acatcaggaa | 3240 |
| aacttacagg | gcaatatttt | taacttgggg | caggaaaga | ggagcagcag | agaattgtact | 3300 |
| agatatacgt | cttattttaa | gagaacttt | gttcttctgt | agattttca | agctgtgttt | 3360 |
| tgtgtgtgt | ccatgtact | tacgcaagga | cagggtacaa | acttagtgg | aaatgtgtcc | 3420 |
| aggctgtat | atcttctccc | tagagttgt | tgccgggtac | acagtgtaa | ccccccgaa | 3480 |
| cggaaacatca | catgttccca | tggtccctt | ttaactgtcg | tggtggctgt | tgctaaatca | 3540 |
| tgacaatggc | tgcctatctg | ctgttcttta | gggtgtgtt | gtacatggaa | ccaggactag | 3600 |
| agattttttgc | agatttatag | actttttttt | ttttttttttt | tttcccttctc | tttcccttctc | 3660 |
| accctttttt | tctgactttt | ccaaatgtt | tgttgcac | aaaatttgg | aggaaaccaat | 3720 |
| tgtaaaacaca | cttccatgtt | agatgtatgt | ttgtgtgtat | cattaaatgtc | ttatttttgg | 3780 |
| ttaaaaagaa | ttttccatgtt | gatacttctc | taaattaaat | aaattataga | atgttagttg | 3840 |
| gttgatgttt | gggtggccat | atagtaatgg | aaagctgtca | taattttttt | taatcagat | 3900 |
| tgaatatttttgc | atatatagaa | atataatgt | gaaagttttt | ggtttaatgt | tagctactgt | 3960 |
| ccggggcaca | gttttccccc | atgtatgt | ctggggat | ctttttttttt | taccatttgc | 4020 |
| tacttagtta | tttttgggt | agaaatgtcg | ttttttttttt | ttttttttttt | ttttttttttt | 4080 |
| agtgcctatgt | taactttatgt | cagaatgtca | gcagatgtgt | aaatcttctc | ctcccccagaa | 4140 |
| tggattaaaca | gtgtcggtca | gtgtgggg | gagagtggat | ggagactttt | agagatgtt | 4200 |
| aaactgcgt | agaatgtat | gagtccagg | gttccatgtt | gaaataaaag | tttggcgt | 4260 |
| ttttgtgttttgc | atataatgtt | ttttgtgttttgc | ttttttttttt | ttttttttttt | ttttttttttt | 4320 |
| catgggttttgc | aaatttttttgc | taattccgaa | gatacaccag | ctcacaat | aaaacgtcg | 4380 |
| cctctgecccc | accctccctc | ctgccccaa | tgtttttgtt | actcagaaaa | gaactgttta | 4440 |
| taccacttac | cttttccccc | gtatgtact | actgtgggg | gtatgttta | tacatgttta | 4500 |
| tcctcttact | catttttaaga | cgtaggtttt | tcataatttct | tctcttaacca | tatacgtat | 4560 |
| ggctcttcgc | ttttatgtat | atctggatt | tctgtggac | ttggcaattt | tttccatgtt | 4620 |
| cttcacttcac | ataatgttcat | tttgcaccc | caatgttctt | gggttggat | cgttgcgtca | 4680 |
| gggtccccgt | gtgtgtgtat | agcagtgtgt | gttccatgtt | ttccatgtt | tttccatgtt | 4740 |
| tttacacatgc | gagtgtatgc | cactttgtt | aagggtgtt | ttttttttttt | ttttttttttt | 4800 |
| tctttatgtat | ggggatgtat | ggccatgtt | tcttagcgt | atttagcacc | ttttttccat | 4860 |
| cttgggttgc | agaaaaattgt | attttctgt | ttttttttttt | ttttttttttt | ttttttttttt | 4920 |
| atttttatgtat | cgtttgcaga | aacttccccc | cccttcatgtt | agggttactgt | ctcaagatgt | 4980 |
| caggagtgttgc | cttccatgtt | atgggttcc | ctccccatcc | ttgtttccac | ccggggctgg | 5040 |
| ctatgttgc | tgtttttatgt | gttccatgtt | aggttgcct | ttttttttttt | ttttttttttt | 5100 |
| agtttttttgc | atcaacatgtc | tcacttcacaa | atgtgtgtcc | tcagccagag | agtaagaaag | 5160 |
| cccaaataatgt | tttacatgtt | taaaaaata | gatttcttact | ttttttttttt | ttttttttttt | 5220 |
| gcatttatgtt | tttttttttttgc | acatgtgtt | tttttttttttgc | tttttttttttgc | tttttttttttgc | 5280 |
| cagaagatgt | catttttttttgc | ggcttccatgtt | tttttttttttgc | tttttttttttgc | tttttttttttgc | 5340 |
| caaaccatgt | caaactgttgc | cgatgttca | aatccatgtt | ggggacttcc | tcttttttttgc | 5400 |
| tctctggacc | taatttttttgc | tttttttttttgc | tttttttttttgc | tttttttttttgc | tttttttttttgc | 5460 |
| gttgggttttgc | atatttttttgc | caacttctt | tcaatgtt | aaagccatgtt | tttttttttttgc | 5520 |
| gcaggatgttgc | ggggccgggtc | agggttgcgt | atagatgtt | cgtaactgtt | acgttgcgtt | 5580 |
| tacgttgc | tctgtatgtt | tttttttttttgc | tttttttttttgc | tttttttttttgc | tttttttttttgc | 5640 |
| gcctctccat | gccaatgttgc | gttccatgtt | tttttttttttgc | tttttttttttgc | tttttttttttgc | 5700 |
| ccagctctgt | cattgtatgtt | gttccatgtt | tttttttttttgc | tttttttttttgc | tttttttttttgc | 5760 |
| aggatacgttgc | ggggggatgttgc | ggggccatgtt | tacccatgtt | agggttacgg | ggcttgcgtt | 5820 |
| gcatttagtgc | ttattttttgc | tttttttttttgc | tttttttttttgc | tttttttttttgc | tttttttttttgc | 5880 |
| gacgtgtgttgc | ctcaatgttgc | ccatgtt | tttttttttttgc | tttttttttttgc | tttttttttttgc | 5940 |
| tgtcgtgtat | gccttttttttgc | cctatattca | tccaggat | tttttttttttgc | tttttttttttgc | 6000 |

-continued

```

agggattcgg ctttcaactt tgctaccatc ttccctgaag caggaaaatg aacatggact 6060
taaatgttct ttgaaaaaa caaagttta agatttgctg ttgtatggag tgacaggag 6120
ggccggagtc agcaggatgc agacttctg ttctgtctgc catgggttg tccagctcg 6180
gtagctctag gagcaccatc ctgccttagc agagcccagg cttgcctc atgaagcatc 6240
attgaaatag caggagcatg ttgatttctt ggttaggtt cattataata acaagagtca 6300
gaacattaat tcgaaaacaac ttgcagatgt catttctca caccagtaca ttcttaagt 6360
tacttgttta taaggaataa cataaaactaa tctgtaccc tatatatatg ttgtgttaca 6420
tatatacata tataaaactgt atagtgtaca tggtaatgtat ttattgtotat gccccagatc 6480
cttaatgttag ttctcatctt ccgcattgcgc tcagccacaa gcgggtact gactgttccc 6540
tgcgttgggg tggccacccctt tgggttggc cctcttaggg ggagggttt ggtcataactc 6600
tccttatactt cgtgcacaga aatgcctagg gtccccatgt gcctgttgg cagcccttc 6660
tcttgttccc ttctcgagca tgggttgcctt ccccaaggctg tgggacagct gccttcccac 6720
gaaaagtgtaa agcagtattt agatcattac tgcattgtgc ctaaaaaccc aagtttcta 6780
ttccctttagg acagaaaaat gcatgtgagg tgggataatc gagtttcaat gaccacatc 6840
agttacacat taaagccaga ccccatgata aaatccaca aatggaaat aaaaactcaaa 6900
tttcttttgc attgtgtaaa taaaatcttca tgggtttaac tttgtactgg taatttctg 6960
tatatttggg atatttgggt taaaataaa acagactgga ctttggtaacc tgacctac 7018

```

```

SEQ ID NO: 14 moltype = DNA length = 1749
FEATURE Location/Qualifiers
source 1..1749
mol_type = other DNA
organism = Homo sapiens

```

```

SEQUENCE: 14
gtgacattgt ttgccaaaat cccaggcago atggacctca gtcttctcg ggtacttctg 60
ccctctgtca ccatggctg gggccaggat ggcgattatc gataccatca ccagcaggat 120
catgactaca gogatgtgg ttgggtgaaat ttgaacccggc aaggcttcag ctaccatgt 180
ccccaggggc aggtgtatgtt ggcgcgtggg agcatctca gcaagaaggaa aggttctgac 240
agacaatggaa actacgcctg catgcccacg ccacagaggc tcggggaaacc cacggaggc 300
tgggtggggg agatcaacag ggctggcatg gaatggtacc agacgtgc caacaatggg 360
ctgggtggcag gatccagag ccgcctactt ggtcgttgc tggatcgggg gtggcagg 420
tacttgttgc gtcacagaa gagggtggca tatttctgtt ggtcaacaac agaatatcca 480
ggtcactatgt tgaggaaat ggacatgatt tccatcaattt atgattacta tatccggg 540
gcaacaacca ctttctctgc agtggaaagg gatgcggat ggaaggatcat aatgtgcgg 600
atgactgtatc acggatgtt atttggcaat gtttagattt gccacatacc aaatctgggt 660
gaaaggaaag gggccggggg caggagggtt tccacatagttt aatcatcg tggatctcc 720
tatagaatgt tctgtgtctc tctttcttc tccctgagctt ggttaactgca atgccaactt 780
cctggccctt tctgtactgtt atcacacttc taataaaatc cacaattaaa ccatgtttct 840
cacttttcac atgttgcatac gcaactgtt tatatgtactt atgtggctt ccttgcacac 900
cacatatacaca tgcgcgtgc ttacagccgg gcttctggg caccagatgc agccctggct 960
ctgttttttca ctgcacatgt aactgcaatgt tcacatgtt gggaggggg ggcacaaactg 1020
gaggagagggtt gcaatgttgc gtttcttgc tttttttttt ggtttttttt ggtttttttt 1080
ccaccaaaatc cagggttttc gcccgtgggg acatctccctgc tccacatgg 1140
ccatgtatgc ttgcacatcc aagtgcacatc ctctttgtctc caggaggatgg gagactggg 1200
gggtgaaatggg gggaaatggaa ggggttgggg gcaaggctgtt aaacagggtt ggaaggattt 1260
cctgaatttag aagacaacgc ttacatacc cagtaaggaa aatgtgttca gggggccagg 1320
gaacccgttgc gatctactt ccaaatgtt gaaaacaacaa gaagcggatggg atggtcagat 1380
aattggggatc agattggggaa ctgtggggg tgagatgttgc cagggttggaaatggg 1440
aaaaggaggat tttttttttt ggcacatgtt ggcacatgtt ggcacatgtt ggcacatgtt 1500
agacaccctg gaaaccttggg gagttactgg caaaactcttc tggattgggg ctgatctttt 1560
tgggtggaaa ggtctggctt gggatcaatc ttctttctgtt gttttttttt ggtttttttt 1620
tgcagatgtt ggcgttatctt tttttttttt gttttttttt gttttttttt gttttttttt 1680
ttcattacaa aaaaataaaat tttttttttt gttttttttt gttttttttt gttttttttt 1740
aaaaaaaaaaaaaa 1749

```

```

SEQ ID NO: 15 moltype = DNA length = 2478
FEATURE Location/Qualifiers
source 1..2478
mol_type = other DNA
organism = Homo sapiens

```

```

SEQUENCE: 15
gcagttgggtt aaactctctt gtctccgtt catcttttca ttgtctcgatc cccttcttcc 60
cgcagacacc cggacccccc ctggggcgca gtcggccgc tccaaacgggtt ccagaaacaa 120
ggccggatgtt ttttttttttctt tccatgtttttt gttttttttt gttttttttt gttttttttt 180
cctcccccttcc caccatccatc cccccccttcc gttttttttt tttttttttt tttttttttt 240
acatggccgc ggcgttgc tctggatggg gaaacaatgtt gggaaaaggaa agaggaaatgc 300
ggagcttataat gacaggatgc agggacttg agacacaaaatc agagaaggatgttcc 360
tccaggcattt ggcgttgc ttttttttttcc tccatgttttttcc gtttttttttcc 420
ggccggatgtt gggatgttgc gatctgttttttcc ttttttttttcc ttttttttttcc 480
ttctctgttcc cccgcgttgc ttttttttttcc ttttttttttcc ttttttttttcc 540
gtctctgttcc ggcgttgc ttttttttttcc ttttttttttcc ttttttttttcc 600
atgtggatgtt actataaaaaatc ttgttttttttcc ttttttttttcc ttttttttttcc 660
ttttttatcatc ccaatgttcaatc aatgttgcatgtt gtttttttttcc ttttttttttcc 720
aattatgtatc aacttcaacc ctttttttttcc ttttttttttcc ttttttttttcc ttttttttttcc 780
cgtggaaatgc ttgttttttttcc ttttttttttcc ttttttttttcc ttttttttttcc 840
gtatgttttttcc ttttttttttcc ttttttttttcc ttttttttttcc ttttttttttcc 900

```

-continued

| | | | | | | |
|-------------|-------------|--------------|--------------|------------|-------------|------|
| ctacgaaaca | catggcag | gatttccat | acgttatgaa | atttcaaga | gaggcctga | 960 |
| atgttcccac | aactcacaaa | cacccatgg | agtgtataaag | tccccggat | tccctgaaaa | 1020 |
| atatcccaac | agccctgaat | gcacttatat | tgtctttcg | ccaaagatgt | cagagattat | 1080 |
| ccttggattt | gaaagcttgc | acctggggcc | tgactcaaat | cctccagggg | ggatgttctg | 1140 |
| tcgtctacgc | cggttagaaa | tctggatgg | atcccctgat | gttggccctc | acattggcg | 1200 |
| ttactgttgc | cagaaaaaac | caggcgtgaat | ccgatctca | tcgggcatc | tctccatgg | 1260 |
| tttttacacc | gacagcgcga | tagcaaaaaga | aggtttctca | gcaactaca | gtgtcttgc | 1320 |
| gagcagtgtc | tcaagaaggat | tcaaatgtat | ggaagctctg | ggcatggat | caggagaat | 1380 |
| tcattctgc | cagatcacag | cttcttcca | gtatagcacc | aactggctg | cagacgctc | 1440 |
| ccgcctgaaac | taccctggaa | atgggtggaa | tccggagagat | gattctacc | gagagtggat | 1500 |
| acaggtagac | ttggggcttc | tgccgttgg | cacggctgtc | gggacacagg | gcgcatttc | 1560 |
| aaaagaaaaac | aagaagaaat | attatgtcaa | gacttacaag | atcgacgtta | gctccaaacgg | 1620 |
| ggaagactgg | atcaccataa | aagaaggaaa | caaaccctgtt | ctctttcagg | gaaacaccaa | 1680 |
| cccccacatg | gttgtgttgc | cagtattccc | caaaccactg | ataactcgat | ttgtccgaat | 1740 |
| caagcctgc | acttggggaa | ctggcatatc | tatgagattt | gaagataacg | ttgtcaagat | 1800 |
| aacaggattat | ccttgcgtc | qaatgttgg | ttatgttgc | ggacttattt | ctgactccca | 1860 |
| gatcacatca | tccaaaccaag | ggggacagaaa | ctggatgcct | gaaaacatcc | gcctggtaac | 1920 |
| caagtgcgtc | ggctggccac | ttccacccgc | acccatccatc | catatcaatg | agtggctcca | 1980 |
| aatagacctg | ggggaggaga | agatgttgag | gggcatcatc | attcagggtg | ggaagcacccg | 2040 |
| agagaacaag | gtgttgcata | ggaagttca | gatcgggtac | agcaacaacg | gtctggactg | 2100 |
| gaagatgatc | atggatgaca | gcaaaacgca | ggcgaagtct | tttgaggggc | acaacaacta | 2160 |
| tgatacacct | gaggtgcgg | ctttccago | tctctccac | cgattcatca | ggatctaccc | 2220 |
| cgagagagcc | actcatggcg | gacttgggttgc | caaatggag | ctgctggct | gtgaagtgga | 2280 |
| agccccatca | gtggccacga | ccacttccaa | cgggaaacttgc | gtggatgaat | gtgatgacga | 2340 |
| ccaggccaa | tgccacagtg | gaacaggtga | tgacttccat | ctcacagggt | gcaccactgt | 2400 |
| gttggccaca | aaaaagccca | cggtcataga | cagcaccata | caatcaggta | tcaaataaaa | 2460 |
| tacgaaatgt | gacagattt | | | | | 2478 |

SEQ ID NO: 16 moltype = DNA length = 3372
 FEATURE Location/Qualifiers
 source 1..3372
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 16

| | | | | | | |
|--------------|--------------|-------------|-------------|-------------|--------------|------|
| taaaaccagg | aagtgaagtc | cccgagca | ttagaaagcc | tgacatggcc | tgactcggt | 60 |
| cagctcagag | cagggcagaa | ctggggacac | tctggggccg | ccttctgc | gcatggacgc | 120 |
| tcttgaocca | ccctgtctct | ggaggaaacc | cgacgcggg | aagaaggaca | gggactcg | 180 |
| tggcaggaaag | aactcagagc | cgggaaagcc | ccattcaacta | gaagcaactg | gagatgcggc | 240 |
| cccccctcgc | gggtctgaaat | tcctctgtct | gttcacaaaat | atgtttttta | tctttaactt | 300 |
| tttggtttcc | ccacttccga | ccccggcg | gatctgc | atc | tgacatgg | 360 |
| cttcttgcgg | ctgatcacca | gacccaacc | cgttttactt | tttcttgacc | tgaacaatca | 420 |
| gtctgtggaa | attgaggggag | gagcacggaa | gggggttcc | cagaagaaca | atgaccta | 480 |
| aatgttgc | ttctcagat | ccaaactat | gtatgttgc | tttccaaagg | gactcgct | 540 |
| gtctgacaaat | ggggccgt | tgatgttgc | aaaacccaaac | cageccat | atgcgttac | 600 |
| ttaacaacac | gtgttgcata | gaggcagata | cctgggttcc | tgtctctgc | ataaaaggta | 660 |
| taaatcatca | ccagaccagt | tttgcggat | ctttgtctca | aataggccag | agtggatcat | 720 |
| cttccgtat | gtttgttaca | cgtatctat | gtttagtgc | cctctgtat | acaccttgg | 780 |
| accagaagcc | actgtacata | tttgcacaaa | ggctgttgc | tctgtgac | ac | 840 |
| accccaaaag | gtatgggtc | tgatgggaa | tgttagagaaa | ggcttcaccc | cgagcgttca | 900 |
| gggtgtatc | ctatggacc | cctttgtat | tgacctgt | caaaggggg | agaagagtgg | 960 |
| aatttgatc | ttatccat | atgtgtca | gaacctggc | aaagacact | tcgaaaaacc | 1020 |
| tgtgcctct | acggcccaag | acctgtgc | catctgttgc | accagtgg | ccacagggt | 1080 |
| cccccaaaagg | gocatgtata | cccatcaaaa | tatttttca | aatgtgtctg | ccttctca | 1140 |
| atgtgtggag | catgttat | agccacttcc | tgatgttgc | gccatattct | acccctct | 1200 |
| ggctcatat | tttgagggaa | tttgacccgg | tgttgttgc | agctgtgg | ccagagtgt | 1260 |
| attttccaa | ggggatattc | gggtgttgc | tgacgacat | aaacttta | agcccacatt | 1320 |
| gtttcccgcc | gtgcctcgac | tccttacag | gatcttgc | aaaggtaaaa | atgaggccaa | 1380 |
| gacacccttgc | aagaagtct | tgttgcac | ggctgttcc | agtaattca | aagagctca | 1440 |
| aaagggtatc | atcggccat | atgtttctg | ggacaacttc | atcttgc | aaatccagg | 1500 |
| cagcctggcc | ggaagggttc | gtgttatttc | cactggat | gccccatgt | ccacttgc | 1560 |
| catgacatcc | ttccggccg | caatggat | tcagggtat | aaagacttgc | gtcaaacaga | 1620 |
| atgcacaggat | ggctgtat | ttacattacc | ttggggactgg | acatcagg | tc acgttgggt | 1680 |
| ccccctggct | tgcaattacg | tgaagcttgc | agatgttgc | gacatgaa | actttacat | 1740 |
| gaataatgaa | ggagagggtc | gcatcaagg | tacaaacgtt | tttcaaggat | accttgcagg | 1800 |
| ccctgagaag | acacccatgg | ccctggac | tgttgttgc | ttcacac | gagacatttg | 1860 |
| tcgtctggctc | ccgaatggaa | tcctgtat | catcgac | aaaaaaaca | ttttcaact | 1920 |
| ggcccaaggaa | gaatacatttgc | caccagagaa | gatagaaaat | atctacaaca | ggagtcaacc | 1980 |
| agtgttacaa | atttttgtac | acggggagag | tttacggta | tcctttagt | gagtgttgc | 2040 |
| tccgtacata | gtatgttgc | agccaaatgt | gggggttgc | gctcttttgc | tgacttgc | 2100 |
| ggaactgtgc | caaaaacca | tttgcaccc | tgatgttgc | aaagacttgc | agaaaatttg | 2160 |
| aaaaagaaatgt | ggccctttaaaa | tttttgcacca | tttgcaccc | attttttgc | atccagac | 2220 |
| attttccatt | aaaaatgggc | tcttgcaccc | aaatgttgc | gaaagcgg | gagacttgc | 2280 |
| caaataactt | cggtttccaa | tttgcaccc | gtatgttgc | atccagg | aggataaggt | 2340 |
| acttaagtac | ctggccggccc | actgtgcact | gtttgttgc | aaatggat | aaaacttgc | 2400 |
| ttacattttgt | tttgccttcc | tccttatttt | tttttgcaccc | tttgcaccc | aaagccat | 2460 |
| ctttttgttt | atattgagac | ataataatgt | taaaatgttgc | ttccaaataaa | atcaatct | 2520 |

-continued

-continued

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| tttttctgcc | ttgatgagaa | gttcctacca | ctgtgcaatg | aataacgaaa | atgccagatt | 3300 |
| acttactttt | cagacatggc | cattgacttt | tctgtcgcca | acagatctgg | caaaacgcagg | 3360 |
| cttttactac | ataggactg | gagacagagt | ggcttgcttt | gcctgtgggt | gaaaatttgag | 3420 |
| caattggaa | ccgaaggata | atgttatgtc | agaacacccg | agacattttc | ccaaatgccc | 3480 |
| atttatagaa | aatcagctt | aagacactt | aagatacaca | gttctaatc | tgagcatgca | 3540 |
| gacacatgca | gccccctta | aaacatttt | taactggccc | tctagtgttc | tagttaatcc | 3600 |
| tgagcagctt | gcaagtgcgg | gtttttatta | tgtgggttaac | agtgtatgt | tcaaatgtct | 3660 |
| ttgctgtat | ggtggactca | atctggat | gatccatggg | ttcaacatgc | 3720 | |
| caagtggtt | ccaagggtgt | agtacttgc | aagaattaaa | ggacaggagt | tcatccgtca | 3780 |
| agttcaccc | agtttaccct | atctacttg | acagctgtca | tccacatcg | acagcccgagg | 3840 |
| agatgaaaat | gcaagatcat | caattttca | ttttgaacct | ggagaagacc | attcagaaga | 3900 |
| tgaatcatg | atgaatactc | ctgtgattaa | tgctgcccgt | gaatgggct | tttagttagaa | 3960 |
| cctggtaaaa | cagacatgtc | agagaaaaat | cctagcaact | ggagagaatt | atagactagt | 4020 |
| caatgatttt | gtgttagact | tactatgtc | agaagatgaa | ataaggggaa | aggagagaga | 4080 |
| aagagcaact | gaggaaaaaa | aatcaaatga | tttattatta | atccggaaaga | atagaatggc | 4140 |
| acttttttca | catttgcatt | gtgttaattc | aatccgttgc | agtctactaa | ctgcccgaat | 4200 |
| tattatgaa | caagaacatg | atgttattaa | acagaagaca | cagacgttt | tacaagcaag | 4260 |
| agaactgtt | gatacattt | tagttaaaagg | aaatatttgc | gcacactgtat | tcagaaactc | 4320 |
| tctgcaagaa | gctgaagctg | tggttatatgc | gcattttat | tgcaacagg | acataaaaata | 4380 |
| tattccacca | gaagatgtt | catatcttcc | agtggaaaga | caattgcgg | gactacaaga | 4440 |
| agaaagaaca | tgtaaaatgt | gtatggacaa | agaagtgtcc | atagtgttt | ttccttgcgg | 4500 |
| tcatctaga | gtatgcaaaag | atttgcgtcc | ttctttaaga | aagtgtctt | tttgcggag | 4560 |
| tacaatcaag | ggtacagttc | gtacattttt | ttcatgttgc | agaacccaaa | catcgcttac | 4620 |
| acttttagat | taattttttt | aatgttattt | aactttaact | tttacccat | tttgcgttcc | 4680 |
| ttaaaatttt | tattttttt | caactcaaaa | aacattgttt | tgtgttacat | atttataat | 4740 |
| gtatctaac | cataatgaca | tatattttt | agaaacttaag | agaatgtat | gtctttgttc | 4800 |
| ttatgaacga | aaaagggatg | gcacttacaa | cacaatattt | aatcaaaaatt | tcagcattat | 4860 |
| tgaaatttgta | agtgtttttt | aactttaat | attttttttt | acctttttaa | attttttata | 4920 |
| ttttggcatt | gtactaatac | cgggacatgt | aagccagggt | tgggtgtatg | tgccctgtat | 4980 |
| cccaggctga | ggcaagagaa | ttacttgcgc | ccaggagggtt | gaatccatcc | tgggcaggat | 5040 |
| actgagaccc | tgccatccaa | aacaaacacaa | aacaccaggg | acacatttt | 5100 | |
| ctgtctttt | tgatcgtgt | cttataatc | gaagggtgtc | atatatgtt | aatgacattt | 5160 |
| taaggacatg | gtgttttat | aaagaatttct | gtgagaaaaaa | atthaataaaa | gcaacaaaaaa | 5220 |
| ttactcttaa | aaaaaaaaaa | aaa | | | | 5243 |

| | | | | | | |
|---------------|-------------------------|---------------|--------------|-------------|-------------|------|
| SEQ ID NO: 18 | moltype = DNA | length = 1579 | | | | |
| FEATURE | Location/Qualifiers | | | | | |
| source | 1..1579 | | | | | |
| | mol_type = other DNA | | | | | |
| | organism = Homo sapiens | | | | | |
| SEQUENCE: 18 | | | | | | |
| gaggaggatgc | ttggccagaca | ctgggtcatg | gcagtggatgc | gtgaagatgc | agttgcctag | 60 |
| ggcaggatgt | gagagagatg | ctggccatgt | ggagaggggtc | tcgggatgtt | ttggctggat | 120 |
| agattttaca | gaaaggctta | tcacggatgt | ttaaattact | ttttccatgt | ttcatctgtat | 180 |
| actccctttt | cagccaaat | tcctttagccc | tgaatatcc | tcctatcc | atctttccct | 240 |
| tctttttttt | cttttttttt | acatgtttaa | atttaaacc | tttttcgttgc | ccccctttct | 300 |
| ttgggatgtt | atggcaagaa | cgagaaatgt | gtatggatgt | tttagggatgt | gtccctgttc | 360 |
| tctgaactt | gggggtttat | gttattttat | tcgttgc | cgagctcaat | tgctccctt | 420 |
| ggtctacaat | ccctggccgc | ttggcttcat | cccttggcc | agcattgtat | acagctatgt | 480 |
| gccctccctc | taccatacc | tcaccccccg | ttcgcttca | tcctccatc | cggttttttc | 540 |
| atcattttct | agaacatggc | gaacattttt | ggtcttattt | tctgttagt | tttaaccac | 600 |
| catctgttct | aaaaggatgg | ctgaatgt | ggaaggatgt | tgttagt | gagactcagg | 660 |
| aagacaactt | ctgcagggtc | actcccttgc | ttctggggaa | aagagaagg | gggcgtgt | 720 |
| ccagtggatc | agaagtggaa | cataatggaa | tcaggcttca | ccttcaagg | cacctatcta | 780 |
| agccatatttta | accctcggg | tttccatggaa | aaatatttca | agtttgcgtc | taggcactt | 840 |
| gcagaaaggcc | agattttttt | gcacccatgt | aaaaatcttt | tcaatgtt | ctgcctatgt | 900 |
| ggtgtgtatgt | gagacgtgt | gttgcacatc | ggctctggcc | ccactatctt | tcagctctc | 960 |
| tctgtttgtt | aatcctttaa | ggagatgttgc | gtcactgtact | actcagacca | gaaccgttgc | 1020 |
| gagctggatgt | atgtgtgttt | gaaagatggc | ggggcttttt | actgggttccc | agtgtgttgc | 1080 |
| atgtgtgtt | atcttgcgg | gaacatgttgc | ggggatgttgc | ggggatgttgc | ggggatgttgc | 1140 |
| caggcggctca | acggatgtgt | gttgcgtgt | ggggatgttgc | ggggatgttgc | ggggatgttgc | 1200 |
| cccttaccc | ccggctgtact | gtgtgttgc | acactgttgc | ttggatgttgc | ctggccatgt | 1260 |
| ctcccccacct | actgcagggtc | gttgcgtgt | tttttttttt | tttttttttt | tttttttttt | 1320 |
| ctgtgtatca | tttttttttt | caaggatgttgc | tttttttttt | tttttttttt | tttttttttt | 1380 |
| agectccccc | tttttttttt | ggggccgggg | ggcgtgtgttgc | gttgcgtgt | tttttttttt | 1440 |
| gaatgttttgc | atgtgtatctc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 1500 |
| tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 1560 |
| caattttttt | gttgcgtgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 1579 |

| | | | | | | |
|---------------|-------------------------|--------------|-------------|------------|-------------|----|
| SEQ ID NO: 19 | moltype = DNA | length = 980 | | | | |
| FEATURE | Location/Qualifiers | | | | | |
| source | 1..980 | | | | | |
| | mol_type = other DNA | | | | | |
| | organism = Homo sapiens | | | | | |
| SEQUENCE: 19 | | | | | | |
| gcggcgccgg | gcagcagctg | cgctcgccact | gtcttggaaag | gagaggacgg | ggcacaaaacc | 60 |

-continued

| | | | | | | |
|-------------|--------------|------------|-------------|-------------|------------|-----|
| ctgaccatga | ccccccacag | gtctgtcca | ccgtctgtgc | tgtgtctgc | tctgtgtc | 120 |
| gtcgccaggc | caggaggcgc | cttggcgcgg | tgcccaggct | gcccccaagg | ggtcaggcg | 180 |
| ggttgtccag | ggggctgcgt | ggaggaggag | gatgggggg | cgccagccg | gggctgcgcg | 240 |
| gaagctgagg | gtgttctag | gagggggggg | caggagtgc | gggtctacac | ccctaactgc | 300 |
| gccccaggac | tgcagtgcca | tccgccccaa | gacgacgagg | cgcccttgcg | ggcgctgtc | 360 |
| ctcgcccgag | gocgtgcct | tccggcccg | gcccgtgtc | ttgcagagga | gaatcttaag | 420 |
| gagagtaaac | ccaaacgcagg | cactgcccc | ccacaggatg | tgaaccgcag | agaccaacag | 480 |
| aggaatccag | gcacctctac | caacccctcc | caacccaa | ctgcgggtgt | ccaagacact | 540 |
| gagatggggc | catgcccgt | acatctggac | tcagtgtgc | accaactcca | gactgaggtc | 600 |
| taccgggggg | ctcaaaact | ctacgtgccc | aatttgtgacc | atcgaggcct | ctacccggaa | 660 |
| cgccagtgc | gtctcccca | ggggcggcc | cgagggtcc | gtgtgtgt | ggatcgatg | 720 |
| ggcaagtccc | tgcccaagggtc | tccagatgg | aatggaaact | cctctgc | cactgggagt | 780 |
| agcggctaaa | gctgggggat | agaggggtc | cagggccact | ggaaggaaaca | tggagctgtc | 840 |
| atcaactcaac | aaaaaaacg | ggccctcaat | ccacccatcg | gcccccccc | atggggccct | 900 |
| caccgctgtt | tggaaaagagt | gttgggtt | gtctgggtgt | caataaaagct | gtgttgggg | 960 |
| tcgctgaaaa | aaaaaaaaaa | | | | | 980 |

| | | | | | | |
|---------------|-------------------------|---------------|-------------|-------------|-------------|------|
| SEQ ID NO: 20 | moltype = DNA | length = 7346 | | | | |
| FEATURE | Location/Qualifiers | | | | | |
| source | 1..7346 | | | | | |
| | mol_type = other DNA | | | | | |
| | organism = Homo sapiens | | | | | |
| SEQUENCE: 20 | | | | | | |
| gcgaggagga | aacgggtccg | gagcgcgcag | ggcttgcgtc | cgccaccggc | gctgcacagg | 60 |
| ctgcccggagc | gagcgtccg | cgcgccgccc | tccccgtct | ccttcttggg | cgagctgcgg | 120 |
| ggatggggcg | gcccggggag | cccggcgcgc | cgccaggaa | gccccccgcg | ccgccccgcgt | 180 |
| ctccgtgtcc | ggcgccgtcg | ggccgcgtcg | ccggcgcgc | ccctgcgggg | ggggcgcccc | 240 |
| cccaggcccc | tggagggtc | ccggaggaa | gccccccgcg | ccccccgcg | ccccccgcg | 300 |
| ccactgcccc | tgctcgcc | tctgtggc | ctggcggttc | ccggccgggg | cgccggacag | 360 |
| cccggtgtgc | ggtcgagac | agccatcgga | gcacatcgcg | cgagccagga | ggacggcggt | 420 |
| tttggcggca | ggggcggact | cctggaccat | ctggactaca | gcttggagca | cagccctc | 480 |
| cgcctgtacc | ggggacaa | ggggacatgc | acagagccgg | tctcggtgc | gccccccgcg | 540 |
| cggccccccg | ccggggacag | cttcagca | ctgtgtgc | cctacccg | ggggggccgc | 600 |
| ggcctcgggg | ggctgtgt | cacccggctt | acccctcgacc | ggggcgctt | cgagggtcgg | 660 |
| cccccgtggca | acctgcggc | caactccctg | cgaaacggca | ccgggggtgt | gtctgtgcac | 720 |
| ccgcaggggc | cgacggccgg | cgtgtgtac | cgccggggcc | ggaacaacgg | ctggactct | 780 |
| gcgggtggcgc | ccacccatcg | gctgtgtac | ccggagacgg | cgagccgtg | caaccccg | 840 |
| gcatccgacc | acgacacggc | catcgcc | aaggacacgg | ggggggcg | cctggccacg | 900 |
| caggagctgg | gggcgttca | gtctgtgt | ggggggcc | ccctggcactt | cggtggacgc | 960 |
| tttctctgg | acggcgttat | ctacttcc | tactaccct | acaactacac | gagcggcgct | 1020 |
| gccaccggct | ggcccgacat | ggcgcgtatc | ggcagagac | ccgggggtgt | gttccagggc | 1080 |
| caggcatccc | tcgactgtcg | ccacggcc | ccgacggcc | gccccctgt | cctctcc | 1140 |
| agctctgtt | aggcccttgg | cgcttgggg | gggggtgttca | cgccggccgc | ttggagggc | 1200 |
| caggagccgtt | gtccccccac | cacccatcg | ctctgcctt | tcataatgt | tgagatccag | 1260 |
| ggccgcgcoca | agagggctag | ctggggactt | aagacggcc | agagccactt | caaagaagg | 1320 |
| gatcaacatcg | aaagagtcca | accaatcgca | tcatctactt | tatccat | cgacatcgaca | 1380 |
| tccgtttagt | gtacgtgtt | aatggatagg | actgtttat | tcttggggac | ttggatggc | 1440 |
| cagttaactt | agggtttat | tgggtat | ttgacttca | attgtccaga | gttataatct | 1500 |
| gaaattaaag | aaagacaccc | tgttttca | aaactcg | ctgatcc | gaagaataatc | 1560 |
| tacattttatc | taacagctgg | gaaagagggt | aggagaattt | gttggcaaa | ctgcaataaa | 1620 |
| cataaaatctt | gttcgggtgt | tttacacgg | acagacccctt | actgcgggt | gtgcatttc | 1680 |
| ctacaatagg | gcactttca | aggatgtt | gtacatttc | agaatatttgc | aaactggct | 1740 |
| gatatttcgt | ctggggcaaa | aaatgtcc | aaatccaga | taatccat | cgtaaagaa | 1800 |
| aagactacag | tgactatgtt | ggggagctt | tctccaa | actcaaagt | catgttgc | 1860 |
| atgtgtggact | ctacgggg | gtctgtcc | aataaaatgc | agccaaaccc | gacctgc | 1920 |
| tgttagcatcc | caaccggc | aacccat | gatgttttgc | ttgtcaatgt | gtgttctcc | 1980 |
| ttccgttgc | ggatatttac | agacagat | aactttacca | actgttc | atattaaagaa | 2040 |
| tgcggcgtat | gtgtggaaac | tgggtgc | tgggtt | ttttgttca | gtgcaagaag | 2100 |
| cccttcacag | tttgcggaccc | tttgcgttatt | gagagaa | aggaaacat | tccagtgt | 2160 |
| gtcgagaaaga | catcgggg | aggagaccc | aaggagaa | aggaaacac | aaccaaccc | 2220 |
| gtttagtttgc | tttttcat | taatgtcc | tttgcgtt | ttgtcaatgt | gtgttctcc | 2280 |
| agcaacgtgt | tagtaacggg | agcaaaactt | accggggat | cgacatcac | aatgtatct | 2340 |
| aaaggaacca | gtacctgtga | taaggatgt | atacaggta | gcccattgt | aatgacacc | 2400 |
| cacatcaat | tctctttcc | atcaacccgg | aaagaaatgt | aggatgtgt | tatccat | 2460 |
| gatgggtggaa | actgttcc | tgtggatcc | ttatcc | ttgtgtcc | acatgttcc | 2520 |
| cttataatttc | ctgttcc | ctggatc | gttgg | ttatccat | gtatggcaga | 2580 |
| aattttgtat | taatttgc | tttgc | tttgc | ttatccat | tcacatgtat | 2640 |
| tctgaatatt | gtgtggcgac | ttactgtcc | tttttgc | ccagttt | aaatgttcaaa | 2700 |
| gtgcgcacga | atgttactgt | gaagctgtat | gtacaagaca | ccatcttgc | ttgtgtga | 2760 |
| ctgcgtatcc | ggggggaccc | cagatcc | gggtatc | ttggatcc | gggtggacaca | 2820 |
| gaactggaaag | tggaaaatca | aaaagaaaat | gacaactca | acatttccaa | aaaagacat | 2880 |
| gaaaattactt | tcttccatgg | ggaaaatggg | caattaaat | gtatgttgc | aaatattact | 2940 |
| agaaatcaag | atcttaccac | catcc | aaaat | gcatcaagac | tgcaagcacc | 3000 |
| attgccaact | cttctaa | aggctgggtc | aagctggaa | acctggact | ctacgtcg | 3060 |
| caggagtccat | ttcccttcc | atggat | tttgc | ttccctgtt | gtctgtt | 3120 |
| gtcatat | tttgcgg | gggtgacc | cacaaatcg | aggacgt | tcgcaaa | 3180 |

-continued

| | | | | | | |
|--------------|---------------|--------------|--------------|-------------|-------------|------|
| agtcaacaac | tagaattgct | ggaaaagcgag | ctccggaaag | agatacgtga | cggcttgct | 3240 |
| gagctcgaga | tggataaatt | ggatgtggtt | gatagttttg | gaactgttcc | cttccttgac | 3300 |
| tacaaacatt | ttgcgtcgag | aactttctt | cctgagtcag | gtggcttcac | ccacatett | 3360 |
| actgaagata | tgcataaacag | agacgccaad | gacaagaatg | aaagtctcac | agctttggat | 3420 |
| gccctaatact | gtaataaaaag | ctttcttgc | actgtcatcc | acacccttga | aaagcagaag | 3480 |
| aactttctg | tgaaggacag | gtgttgtt | gcctcccttc | taaccatcgc | actgcaaaacc | 3540 |
| aagctggct | acctgaccag | catcttagag | tggtgcacca | gggacttgc | ggaacagtgt | 3600 |
| agtaacatgc | agccgaaact | catgtgaga | cgacggagt | cctgtcgaa | aaaactctc | 3660 |
| acaaactgga | tgtccgtctg | ccttcttgc | tttctccggg | agactgtcg | agagcccttc | 3720 |
| tatttgcgg | tgacgactct | gaaccagaaa | attaacaagg | gtccctgtga | tgtaatcaact | 3780 |
| tgc当地 | tgacactct | taatgaagaa | tggttgttgc | ggcagggttc | ggaattcagt | 3840 |
| actgtggcat | taaacgtcgt | ctttggaaaa | atccggaaa | acgagagtgc | agatgtctgt | 3900 |
| cggaatattt | cagtcaatgt | tctcgactgt | gacaccattt | gcacagccaa | agaaaagatt | 3960 |
| ttccaagcat | tcttaacgaa | aaatggctc | ccttatggac | tccagcttaa | tgaaatttgg | 4020 |
| cttgcgcgtt | aaatggcac | acgacagaaa | gaacttgc | acatgcacag | ttccctccgt | 4080 |
| attcttgaag | atggatcac | caagtcataa | accattggcc | actatgat | atcaaatttgg | 4140 |
| tccactataa | aagtctttaa | gaagatagca | aattttactt | catatgttgc | gtactcgat | 4200 |
| gaccactgcc | atttgatttt | accagatttc | gaagcatttc | aatgttgc | agggaaagaga | 4260 |
| catcgaggga | agcacaagtt | caaagtaaa | gaaatgttac | tgacaaagct | gtgtgcgacc | 4320 |
| aagggtggaa | ttcattctgt | gctttggaaa | ctttttagaa | gcattttgg | tttacccaac | 4380 |
| agoagagctc | catttgcata | aaaatacttt | tttgactttt | tggacgocca | ggctgaaaac | 4440 |
| aaaaaaatca | catatcttgc | cgtcgatcat | atttggaaaa | caaacaccc | tccttccgc | 4500 |
| ttctgggtaa | acatccctgc | gaaccccttgc | tttgcatttt | acattaagaa | gacaccat | 4560 |
| atagacggct | gtttgtcgt | gatttgcggc | gcatttgc | atgcatttgc | tctcacagag | 4620 |
| cagcaactag | ggaaggaaagc | accaactaat | aagcttctc | atgccaagg | tatcccaacc | 4680 |
| tacaaaagaa | aagtaaaatc | tttattacaaa | gcaatcagg | atttgcctcc | attgtcatcc | 4740 |
| tcagaaatgg | agaatattttt | aactcaggaa | tctaaagaaa | atgaaaatgt | atttaatgtaa | 4800 |
| gaagtggcct | tgacagaaat | tacaaataatc | atcgtaaaaat | attttgtat | gatttcaat | 4860 |
| aaactgaaaa | gagaacgagg | gcttggaaagaa | gctcagaaac | aacttttgc | tgtaaaatgt | 4920 |
| ttatttgatg | aaaagaagaa | atgcaggatgg | atgttgcac | tctggggctc | ggcttaatct | 4980 |
| ggcaaggatc | ttcagacgc | tttggggacaa | aatggctgc | tgactactc | tggtgcgtt | 5040 |
| atttgcgtt | tgacatagg | tttccatttgc | ggcactgtt | ttttttagaa | ccaaggccaca | 5100 |
| tgacagcgtt | tttagaaagca | taccaaccct | tggtgcctgt | tgtatcacgt | ggggaccctt | 5160 |
| ctgtaaatag | agttgttgc | gttgcaccc | acagcccttgc | tgtttacaga | gaatacaagg | 5220 |
| ccagtaacgc | aatgttgc | ttgttacactt | agttccactt | taagcacaat | gatataaagt | 5280 |
| gttttgcgtt | aaaactacag | ctatgttgc | cttgcgttgc | actgcaccc | tgcattgtaa | 5340 |
| agggatactg | ccagtgctca | aaacaaaatg | tgaatgatgt | cattttggaaa | caagggtgggg | 5400 |
| gtgttagggc | aacatcgagg | atttgcagca | tttggaaactt | ccccagtagt | tcttggaaaa | 5460 |
| gtctgacccgca | gaattttgttgc | gtgttacactt | agcatattttgc | agtgtgtgt | tggtttttaaa | 5520 |
| caaaaaaacta | acagtgttgc | aaatcttgc | aaagggtctc | tgtttttcag | tggttcatcaa | 5580 |
| ctgcacttca | tcaaaatc | cttccattca | ccaaggact | cttaatgttgc | gagagtgggc | 5640 |
| tttattttaaa | tgaacagcat | tttaaccaga | tacttttgc | taatgtatgt | tcctttctt | 5700 |
| catctgtttt | ttcatactaa | atgttgc | tgatggatatt | tttgcattttt | atacaaaaaaa | 5760 |
| atcatttattt | catttgcgtt | ccaaaaccc | tgatcagaac | gatctgttgc | agagtaactc | 5820 |
| cattttttata | tgatgttgc | tctccgttgc | tttagatttgc | ggtgcaccc | tggtttatgt | 5880 |
| atacttgcgt | tgatgttgc | aaaaaaaaga | tacattttac | atttcatgc | attgtgttc | 5940 |
| acactggatgt | atttatata | aatatata | tttgaggccc | aaggcttgc | aaatattatgt | 6000 |
| atacaacttgc | gtatcttgc | tttactatgt | acttttttgc | agatatttcc | gcaggagaaa | 6060 |
| gaattttaaa | tacccatttt | attatgc | ttttttttaa | agaatttcc | atccaggat | 6120 |
| actgtgttgc | tttttttttgc | gattttttgc | tttcttgc | tttgcgttgc | atgaccat | 6180 |
| taatattacc | tgtgttttgc | tttcttgc | ccagaagaag | taaacatgc | atttgcac | 6240 |
| tcttgatgtt | tttttttttgc | aggccatcc | tacccttgc | gttccaaatc | ccccatcttgc | 6300 |
| gaaaagaaat | tttttttttgc | aaatataat | tgatcataca | ttcccttgc | tcatatcca | 6360 |
| acactgtatg | gaggcatttgc | attttaat | gaattttgc | taaaggcaat | tccttttgc | 6420 |
| ttctgttata | tctggaaaat | catggagag | gtgacaccc | aaacaaatgt | tcaagaaaa | 6480 |
| taaggcgttca | tacccctgtat | aggccatcc | ccaaatcttgc | tgtttttgc | cattgtctgt | 6540 |
| ccaaacggaca | caccttcaaa | aaacaaaact | accaatata | tgacagatca | gaataaagg | 6600 |
| gagaggctgc | gtccccatttgc | aaggcttgc | cagtcttca | agaggttgc | gagtttataa | 6660 |
| gagaacaaaca | gttggaaatgt | tgagggccaa | gggttaggg | gttgccttcc | agacttcccc | 6720 |
| tactacttgc | gggtacttgc | aacttata | atcgatc | tttttttgc | tttttttgc | 6780 |
| aaatgttaca | tgagggttgc | cctgtcc | tttttttgc | tttttttgc | tttttttgc | 6840 |
| taccccttgc | tccaggatgttgc | tttttttgc | tttttttgc | tttttttgc | tttttttgc | 6900 |
| gcctcaattt | ctgtataata | tttcttgc | tttttttgc | tttttttgc | tttttttgc | 6960 |
| atgtatgttgc | tttttttgc | tttttttgc | tttttttgc | tttttttgc | tttttttgc | 7020 |
| tcacatttgc | caaatttgc | tttttttgc | tttttttgc | tttttttgc | tttttttgc | 7080 |
| tgggttctata | ttaaacagca | tttttttgc | tttttttgc | tttttttgc | tttttttgc | 7140 |
| tctttgcata | tttttttgc | tttttttgc | tttttttgc | tttttttgc | tttttttgc | 7200 |
| tacagcgttca | tttttttgc | tttttttgc | tttttttgc | tttttttgc | tttttttgc | 7260 |
| ttatttgcata | cctgttgc | tttttttgc | tttttttgc | tttttttgc | tttttttgc | 7320 |
| aaaataaaaca | tttttttgc | tttttttgc | tttttttgc | tttttttgc | tttttttgc | 7346 |

SEQ ID NO: 21 moltype = DNA length = 2828
 FEATURE Location/Qualifiers
 source 1..2828
 mol_type = other DNA
 organism = Homo sapiens

-continued

SEQUENCE: 21

```

agaacagtga cagcgcgcgc gcagccgacc ccgcctcctc ggcggacagc gatgtcagc 60
tggctcgccc cgagtcatcg cctagegtcg gcagggccgc tgaccgaccc acggaggcgc 120
cgattggcgc attgtccact ggcagaagg agcagctgc ccgcgcggcc ccgcgcgcgc 180
ctgaggcgcg ggtccgcagg gccggggga agccgagggc tgccggagaa ccctgcaggt 240
gtcaactcggg acgcggaaatg ggcgttgcgg agggtttgttatacataccgc ttgagactcc 300
ccgacaagcg taatttggtc gagttcgacg ggaaagtact ctccccaccc cagcgcggcc 360
cgcgtagtcc gaggttactg tccccggcgc gtctctgtt gcccagttcc agaggctgc 420
cttgaacccg ggcgcgcacg agcgcagggc atccgaggcg acagccctg gcacggcccg 480
acctgtaccc accctcgccag gaagactgtta atcgtggaa tacactacc tacccaggca 540
atatgaatgtttatgttgc gaaactgcgc tttctttagtgcattttgttgcatttttgaa 600
ccggtccact gacaacgc当地 tatgtttatc ggagaatatg ggaagaaaact ggcactaca 660
cttttcatc tgatagcaat atttttgatgt gtaaaaaaa caaaaggcgc ccaattttg 720
cattccagga ggaagtccatg aaaaaaaaaatgttgcacgttatactgcagatg gacataatgt 780
gattaattcc tggcttagtgc tttatccatca tttttttgttgcattttgttgcatttttgaa 840
gaaaatttcc tatgttttgc tttccgttgc ttgtcttgc aaccaggcgtt tggtctgtt 900
tgcgttgc当地 ttttgcctt ccattccgc当地 ttttgcatttgc atcttacccctt attttgttgcatt 960
tttgc当地 tttatccatca ttttggggag ctgtcttgc当地 tttatccatgcattttgttgcatttttgaa 1020
aagaacacaaa aaaaaaaaaaaa attcgc当地 atttttatgc当地 tttatccatgcatttttgaa 1080
ctggactaac accgtgtca tctggcttatttgc当地 tttatccatgcatttttgaa 1140
tttcaatttat tgctgtgttgc当地 tttatccatgcatttttgaa 1200
atccagtggaa agagtgttca tctcagaatg ttactatgtc atgtatgttgc当地 ggcttccaaa 1260
acctatccatcc cggaaatccatc atgttttttgc当地 agaatgttcc ttttgcatttgc当地 1320
tctgttttgc当地 atttttatgc当地 gtaatccatc ttttgcatttgc当地 ggacttgc当地 1380
tttttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1440
gatcagttt gggtagtgc当地 ttttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1500
attgtatggaa agatgttca ttttgc当地 tttatccatgc当地 tttatccatgc当地 1560
ctatgaccgc gtttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1620
tcaactatgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1680
aacaaggatcc ttttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1740
tttcaatttat tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1800
tgcgttgc当地 ttttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1860
gctggaaatgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1920
acagagccttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1980
atatacatgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2040
aaagaatgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2100
ttctatatgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2160
ttaaaaaaatgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2220
tttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2280
ctcttagtcc tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2340
catttttcaat tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2400
ggaaaatgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2460
tgccttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2520
tccatgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2580
aaggcttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2640
ccccagttcc tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2700
tgcactgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2760
gattactcat tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2820
aaaaaaaaaaaaaaaatgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 2828

```

SEQ ID NO: 22

moltype = DNA length = 2840

FEATURE

Location/Qualifiers

source

1..2840

mol_type = other DNA

organism = Homo sapiens

SEQUENCE: 22

```

tttttgc当地 agaaacagcc cttccgc当地 ttgtcctggc ggtgaggcacc cagggcttgc 60
cttttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 120
cttcttgc当地 ttgttgc当地 tttatccatgc当地 tttatccatgc当地 180
tccaggccatc tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 240
gggcacccca cttccatgc当地 tttatccatgc当地 tttatccatgc当地 300
aggccccatc tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 360
gagagaaatgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 420
attttatccatgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 480
agcacatgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 540
catcccaatc tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 600
tttccccatc tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 660
gaacaccatc tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 720
ccctcttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 780
tggctcatgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 840
agagggttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 900
tgcactgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 960
ctcagttgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1020
tttgttgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1080
cctgc当地 tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1140
cccaggggatc tttatccatgc当地 tttatccatgc当地 tttatccatgc当地 1200

```

-continued

| | |
|--|------|
| aggccagccc ttgcgggggc aggccagagga cagagttggc tttgcgcggt ctctgcctgc | 1260 |
| tgacgcgttc caatttcctc catggggaga acaaggaggc agtcgcgtgt gcatgttcca | 1320 |
| gaaggtttac tggggaggag gaagcgagca gaggagctg tgggtgcgt tgaaggggtg | 1380 |
| ggcagggtgg gagggatgca cgcttatgtc agcatagcat gtgttagtac tacacacatc | 1440 |
| tccatgcaga agcacaactg ggcagccctg gcttccagct ctggggttca gcacaacaga | 1500 |
| caccagccgt tggtctctca gaagccaggc agaccacatc gggctcagga cggtttaccc | 1560 |
| aaagtccaga gttttatgc ctctccctgg cattctccat aagaagggg aggtcgatg | 1620 |
| accctttaga tcttgttcat ctgggattt ctttgggtgg gtttagacac gatgcctct | 1680 |
| tttctcagg atagcagata acctgtttt aaagagggtc taattctgtg ggtcttaaat | 1740 |
| tttcttcctt ctcttcctc ttctgtgtt gtgtgttggg aaaatggcaa gtttccaaa | 1800 |
| ccagctttgg aggaacgatt acgtttccc tccaaatttca agtccgaaag accagagccc | 1860 |
| tcatccaaa gccccccccc catggatt tttcgtttc atttgcatac cgtcccatgg | 1920 |
| gaggggccca tgcgttctca gaaccatcc tggaggcage aggtcggtt gagttagttt | 1980 |
| ggccgtctca tgacccccc ccctggatgtt gtaacaagg atgtctggg cgatgtcgag | 2040 |
| aatgttttg aagtcgttcc cagatgcgc tgatgtatc accaggatgg aatgtcgat | 2100 |
| cgccttgagt ccaacccctg cataaacqg gttctataaa acaaggctac tctaccctaa | 2160 |
| gctaagtctt tgcgttccaa cccacttcat ctttgcgttcc tggagacccgtt gttacactaa | 2220 |
| cctgtacttgc acctgttcat gtatgttgc tgggtgtt catgcgtgt ggaccaagca | 2280 |
| atggcatggg tgggtgtgtt gtgtgtgtt gtgtgtgtt gtgtatgtcg | 2340 |
| tcacacttgc gtgtgtatata tgcgtatgtc atgcgtatc aatgtttt gatgtcaag | 2400 |
| acaacacat tccattgttt taaatatttca attatgtaaa caatacgcag agggaccata | 2460 |
| tctacttgc tcatattttt tgcgtatgttca aacatgtcat ttgcataaaa ttaagcttcc | 2520 |
| tggaaaggca agcgttattg gacccaaacg actgtctcg aacatgtgtt gttatctcg | 2580 |
| gttccatatac aatccaaacg taaatggggcc tccccccccc tccagggggg aacaccaggaa | 2640 |
| ccccggagtt tttttttttt gctatattttt aaatgttgcatac tgacgtttt cttcccttcc | 2700 |
| tttgcgttca gttggaaatgta gcaatgttgc taaaatgtgtt ttgcgttgc tttgggttca | 2760 |
| ttatgtttt aaaaatgtat ctgtttgtt tttggaaataa aaaaatctca tttgggttca | 2820 |
| tgaaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa | 2840 |

SEQ ID NO: 23 moltype = DNA length = 2309
 FEATURE Location/Qualifiers
 source 1..2309
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 23

| | |
|--|------|
| gcggggggcg gccggggcg ggccaggccg gctagaggggg cgggtctagc ggcggccccc | 60 |
| ggcgaagttc actgcgttgc cgctgacaga cgaagatggc cggacatgtc ggaactaaag | 120 |
| caaatggta tgagccttag agtttctgaa ctccaaatgtc tgggtggctt cggccggaga | 180 |
| aacaagcgc cgcggccaa cgaacttcc acaaaagccc tgcattttgtt aaaggctggc | 240 |
| tgtgtctgtt ctgtgcggat gaaaatggaa gaacttccat ggcgggggtt cccacagaaa | 300 |
| atcatgcgc ctgcggactt gtccatcccc aacgttccat cttttttttt ggcggccat | 360 |
| ttgtctccat tccatattcc acaacttact tacgtgttcc accctgcattc atgcatttt | 420 |
| ctccctgtttt ctcttcgtggg accttccatc gaaacttgcac tcccaatctt tacatcgat | 480 |
| cttcaccccg tccatcccgaa taaaatggaa cttttttatgtt tttactggat | 540 |
| gaactgttac aaccaccccg tcttcgtatc gacaacatgtt cttttttttt gggatatttt | 600 |
| tttgcatttgc ctttcgttcc acacaaatgtt cttttttttt gggatattttt gggatatttt | 660 |
| ggggccatggat gtggacttcc acgtacgggtt cttttttttt gggatattttt gggatatttt | 720 |
| tgtccaccaat aagatccat cccacccat tttttttttt gggatattttt gggatatttt | 780 |
| agcccttcggc gtatcttcc accttccatc aatggcggtt aacccaaacg accccacccg | 840 |
| ccaaatataa tcacccatc tttttttttt gggatattttt gggatattttt gggatatttt | 900 |
| tcttcgttgc ctttcgttcc acacaaatgtt tttttttttt gggatattttt gggatatttt | 960 |
| tccttcacacg tttttttttt gggatattttt gggatattttt gggatattttt gggatatttt | 1020 |
| agacgtttttt tttttttttt gggatattttt gggatattttt gggatattttt gggatatttt | 1080 |
| ctaaagggtttt tttttttttt gggatattttt gggatattttt gggatattttt gggatatttt | 1140 |
| cttacatgtt tttttttttt gggatattttt gggatattttt gggatattttt gggatatttt | 1200 |
| aaacccaaatc gggatattttt gggatattttt gggatattttt gggatattttt gggatatttt | 1260 |
| gtatggctttt tttttttttt gggatattttt gggatattttt gggatattttt gggatatttt | 1320 |
| gaggatggca ctggggccacc gatggatca aaaaaggaaatg tacaggaaatg ttctgcctt | 1380 |
| tacaatggatg tttttttttt gggatattttt gggatattttt gggatattttt gggatatttt | 1440 |
| cgttccttc acatggatcc acatggatcc gatggatcc gatggatcc gatggatcc | 1500 |
| gtatggatcc acatggatcc gatggatcc gatggatcc gatggatcc gatggatcc | 1560 |
| ccactaaata ataaaggcat tttttttttt gggatattttt gggatattttt gggatatttt | 1620 |
| ccaaaggccctc ctgtgttgc acacaaatgtt tttttttttt gggatattttt gggatatttt | 1680 |
| catcccttc acatggatcc acatggatcc gatggatcc gatggatcc gatggatcc | 1740 |
| tttacatgtt tttttttttt gggatattttt gggatattttt gggatattttt gggatatttt | 1800 |
| tcagatgttgc acacaaatgtt tttttttttt gggatattttt gggatattttt gggatatttt | 1860 |
| tttcttcgttgc acatggatcc gatggatcc gatggatcc gatggatcc gatggatcc | 1920 |
| agttttttttt tttttttttt gggatattttt gggatattttt gggatattttt gggatatttt | 1980 |
| gtcacaacaa gggatggatcc gatggatcc gatggatcc gatggatcc gatggatcc | 2040 |
| tcatggatcc gatggatcc gatggatcc gatggatcc gatggatcc gatggatcc | 2100 |
| gcagaaagaa gggatggatcc gatggatcc gatggatcc gatggatcc gatggatcc | 2160 |
| cgtttttttt tttttttttt gggatattttt gggatattttt gggatattttt gggatatttt | 2220 |
| tatatttcgtt tttttttttt gggatattttt gggatattttt gggatattttt gggatatttt | 2280 |
| taaaaaaa aaaaaaaaaaaaaaaa aaaaaaaaaaaaaaaa | 2309 |

SEQ ID NO: 24 moltype = DNA length = 1740

-continued

| FEATURE | Location/Qualifiers |
|---|-----------------------------|
| source | 1..1740 |
| | mol_type = other DNA |
| | organism = Homo sapiens |
| SEQUENCE: 24 | |
| ccacgcgttag cccgcggcca ggcacccggg aggagcggc egcgcgtcga cctctcccg 60 | |
| cctgtctgtt cgcttcagg ctggatgg ccggcacatc ggggtctgt cgctcgct 120 | |
| gcagagccg aggctcgccg ccggctctgg cggccggc ectgacagcc cccacctcgc 180 | |
| aagagcggcc tatgcgaca aaaggatcaa ggtggegaag cccgtgtgg 240 | |
| agatggatgg tgatgatgt acccgtt tctggcaigt catcaaggag aagtcatcc 300 | |
| tgcggccatgt ggacatccag ctaaagtatt ttgacctgg gtcceaaac cgtgaccaga 360 | |
| ctgtatgacca ggttaccatt gacttgac tggccaccca gaagtagact gtggctgtca 420 | |
| agtgtgcac catcaccctt gatggggcc gtgtgaaaga gttcaagctg aagaagatgt 480 | |
| ggaaaatgtcc caatggact atccggaaacat ccttgggggg gactgtttc cgggagcc 540 | |
| tcatctgcaaa aacatccca cgcctagtgc ctggatggac caagccatc accattggca 600 | |
| ggcacgcggca tggcggaccatcataaaggccca cagactttt ggcagacccg gccggactt 660 | |
| tcaaaatgtt cttcacccca aaagatggca gtgtgtcaaa ggagtggaaa gtgtacaact 720 | |
| tcccccgagg cggcgtgggg atggcatgt acaacacccg ccagttccat tcagggttt 780 | |
| cgcacagctg ctteccatgt gccatccaga agaaatggcc gctgtacatg agcacaaga 840 | |
| acaccatatac gaaaggctac gatggggctt tcaaggacat cttccaggag atctttgaca 900 | |
| agoactataa gaccgcattc gacaagaata agatctggta tgagcaccgg ctcattgtat 960 | |
| acatgggtgc tcaaggcttc aagtcttgg gtggctttgt gtggccttgc aagaactatg 1020 | |
| acggagatgt cggatcagac atccggccccc agggctttgg cttcccttgc ctgtacatg 1080 | |
| cctgcctgtt ctgcctgtat gggaaagcga tggaggctga ggccgtctat gggaccgtca 1140 | |
| cccgccacta tcggggggcc cagaaggggc gggccaccag cacaacccc atcggcagca 1200 | |
| tcttgcgtt gacacgtggc ctggggccccc gggggaaact ggtatggaaac caagacctca 1260 | |
| tcagggttgc ccaaggatgtt gggatggatgt gggatggatgt ggagccatgt 1320 | |
| cacaggaccc ggcgggctgc attacggcc tcagcaatgt gaagatgttcc gggccatcc 1380 | |
| tgaacaccac ggacttccctc gacaccatca agaccaacctt ggacagagcc ctggccaggc 1440 | |
| agtaggggggaa ggccggccaccc atgggtgcacat tggggggcc ggggcttgc cggccgggtcc 1500 | |
| tccctggcgc ggcggggggatgtt gggccatctt gggccatctt ctggggatgt 1560 | |
| tttttttttata aaccatgtt tttttttttat atatgtgtt tttttttttat ggtgacgtga 1620 | |
| ggcaggagca gtgcgtttt cctcggccatc tcaatgttgc ttgcataactg taatttat 1680 | |
| tgcggccatgttggac acacatgttgc ctataaaaaa gcttttccaca aaaaaaaaaa 1740 | |
| SEQ ID NO: 25 | moltype = DNA length = 1552 |
| FEATURE | Location/Qualifiers |
| source | 1..1552 |
| | mol_type = other DNA |
| | organism = Homo sapiens |
| SEQUENCE: 25 | |
| ggtcgcgttta agaaaggagt agctgtatc tgaagcgtc tggacgttgg attagaaggc 60 | |
| agcaaaaaaa gctctgtgtt ggctggggcc ccctcgttgtt gcaggcttagt agggacttagg 120 | |
| ctgggtgtgg agctgcacgc tatccacagg ccccaggatg caggccctgg tgctactct 180 | |
| ctgcattttgg ggccttcgtc ggcacagcg ctggccagac cctggccaggcc ccccccggg 240 | |
| gggttcccca gacccggaca gcacaggggc gctgggtggag gaggaggatc tttttttcaa 300 | |
| agtcccccgtt aacaaggctt cggcggctgt ctccaaactt ggctatgtacc tgtaaccgggt 360 | |
| gcgcgtccatc acggggccca cggaccaatgtt gtcctgtctt cctcgttgtt ggccacccggc 420 | |
| cctctcgccc ctctcgctgg gageggggca gcaacacaaatc tccatcatc accgggtct 480 | |
| ctactatgac ttgatcagca gcccggacat ccattgttacc tataaggagc ttcttgacac 540 | |
| ggtcactggc ccccaagaaga acctcaagatg tgcttcccg atcgttttgc agaagaagct 600 | |
| gcccataaaa tccaggtttt tgccacccat gggaaatgtca tatggggacca ggcccaggat 660 | |
| cctgacggcc aaccctcgat tggacatcgca agagatcaac aactgggtgc aggcggat 720 | |
| gaaaggggaaatctggccaggat ccacaaaggaa aattcccgat gagatcggca ttcttctt 780 | |
| cggtgtggcc cacttcagg ggcggatggg aacaaatgtt gactccggaa agacttccct 840 | |
| cgaggatgtt tacttggatg aagaggaggac cgtgggggtt cccatgtatg cggaccctaa 900 | |
| ggctgttttta cgctatgtt tggatcgatc ttctcgttgc agatgtccc agctggccct 960 | |
| gaccggaaatc atgatgtatca ttcttcttccat gcccctgaaat gtaaccggaa atttggac 1020 | |
| gatagaggatcc acccttcaccc tccggatcatc tcatgtatca gaccggaaac tgaagacgt 1080 | |
| gcaggccggcc ctactgtcc ccaacgttgc gctggatgtt gggggggccatccatc 1140 | |
| cctgcggggat atgaatgttccat tggatccatc gactttggatc agatcaccagg 1200 | |
| caaacccatc aagatgtactc aggtggaaaca cccgggtggcc ttggatgttgc acggggatgg 1260 | |
| ggcggggaaacc acccccgatcc cagggtgtca gcctggccacat ctcacccatcc cgtggacta 1320 | |
| tccatcttacaa cggccatccatc tccatgttgc gggggggccatccatc 1380 | |
| cattggccatc atcttggggccatccatccatc agtttataat tccatccatc 1440 | |
| tagaagaaaa cccggggggatcc acggatgttcc acggggccatccatc 1500 | |
| tcaatgtatccatc aataaaaaaa gctttatccatc taaaaaaaat aaaaaaaaaaa 1552 | |
| SEQ ID NO: 26 | moltype = DNA length = 4816 |
| FEATURE | Location/Qualifiers |
| source | 1..4816 |
| | mol_type = other DNA |
| | organism = Homo sapiens |
| SEQUENCE: 26 | |
| gttcccgat ttttggggccatccatc ccctcgatccatc ccctgtgttgc tccatccatc 60 | |
| gaggccatgtt ggtttaaggatccatc agggggccatccatc 120 | |

-continued

| | | | | | | |
|--------------|-------------|-------------|--------------|--------------|-------------|------|
| ttccatgatc | tttttgagt | cgcaattgaa | gtaccacctc | ccgagggtga | ttgctcccc | 180 |
| atgcgggta | gaacccatgt | tgtctgttc | accacttac | ctccatgtaca | gaatttggct | 240 |
| tatgcctact | caatgttgaag | atgtgagga | tgaaaaacctt | tgtgtatgc | cacttccact | 300 |
| taatgaatgg | tggcaaaagca | aagctatatt | caagaccaca | tgcaaagcta | ctccctgagc | 360 |
| aaagagtcac | agataaaaacg | ggggcaccag | tagaatggcc | aggacaaaacg | cagtgcagca | 420 |
| cagagactca | gaccctggca | gccatgcctg | cgcaggcgt | gatgagagtg | acatgtactg | 480 |
| tttgtggacat | gcacaaaaagt | gagttgtgcac | ccggccacagac | atgaagtcgc | ggctccctgc | 540 |
| cagtcccgag | acccacctgg | acatgtctgg | ccacacatc | cagggtgtcc | agggtgtgca | 600 |
| gggaaacactg | gaactcacct | acctgcccac | caatgccage | ctgtcccttc | tgccaggatat | 660 |
| ccaggagggt | cagggtgtac | tgctcatcg | tcacaaccaa | gtgaggcagg | tcccaactgca | 720 |
| gaggcgtgg | attgtgcgg | gcacccatgt | cttgaggagg | aactatgccc | tgccgtgtgt | 780 |
| agacaatgga | gaccggctga | acataaccac | ccctgtcaca | ggggctccc | caggaggct | 840 |
| gccccggact | cagcttcgaa | gcctcacaga | gatcttga | ggaggggtct | tgatccagcg | 900 |
| gaaccccccag | ctctgttacc | aggacacatg | tttgtggaa | gacatcttcc | acaagaacaa | 960 |
| ccagctgttgc | tcacacatgt | tagacaccaa | cegtctcg | gectgeccacc | ctgttcttcc | 1020 |
| gtatgttaaq | ggctcccgct | gctggggaga | qagttctgag | gattgttaca | qcctgacqcg | 1080 |
| cactgtctgt | gccgggtggct | gtgcccgtg | caaggggcca | ctgcccactg | actgtgttcca | 1140 |
| tgagcgtgt | gtgtggcggt | gcacggggcc | caagcactt | gactgtctgg | cctgtcttcc | 1200 |
| cttcaaccac | agtggcactt | gtgagttgtca | ctggccagac | ctggcttac | acaacacaga | 1260 |
| cacgttttag | tccatgcccc | atccccgggg | ccggtatata | ttccggccca | gctgtgttcc | 1320 |
| tgccctgtccc | tacaactacc | tttctacgga | cgtggatcc | tgccacccctg | tctgtccccct | 1380 |
| gcacaaccaa | gagggtgacag | cagaggatgg | aacacagcgg | tgtgagaatg | gcagcaagcc | 1440 |
| ctgtggcccg | gtgtgtatgt | gtgtggccgt | ggagcactt | cgaggggtga | gggcgttac | 1500 |
| cagtgcataat | atccaggagt | ttgtgtgtgt | caagaagatc | tttggggagcc | ttggcatttt | 1560 |
| gcggggagac | tttggatgggg | accaggctt | caacactgtc | ccgctccagc | cagagcagct | 1620 |
| ccaaatgtttt | gagactctgg | aagagatcc | agtttacca | tacatctcg | catggccgga | 1680 |
| cacccctgc | gacccatcg | tcttcagaa | cctgcaagta | atccggggac | gaatttgcgt | 1740 |
| caatggcc | tactcgtga | ccctgc | ccctgc | agctgggtgg | ggctgcgtc | 1800 |
| actgaggggaa | ctggggactgt | gactggccct | catccacat | aacacccccc | tctgttctgt | 1860 |
| gcacacgggt | ccctgggacc | agcttttgc | gaacccgcac | caagcttgc | tccacactgc | 1920 |
| caacccggcca | gaggagggat | gtgtggccgt | ggggctggcc | tgccacccagc | tgtgegcccc | 1980 |
| aggggactgc | ttgggttcccg | ggcccaacca | gtgtgtcaac | tgccggccat | tccttgcggg | 2040 |
| ccaggagatgc | gtggggaaat | gcccggact | gcaggggttc | cccaggggat | atgtgtatgc | 2100 |
| caggcactgt | ttggcggtcc | accctgatgt | tcagccccag | aatggctc | tgacctgttt | 2160 |
| tggaccggag | gtgtggccgt | gtgtggccgt | tgcccaactt | aaggccctc | ccttctgtgt | 2220 |
| ggggccgtgc | cccaagggtgt | tggaaatgt | cctctccat | atggccatct | ggaagtttcc | 2280 |
| agatgaggag | ggcgcgttcc | acgcttgc | catcaactgc | acccacttct | gtgtgtacat | 2340 |
| ggatgacaag | ggctgccccg | ccgagcagag | agccagccct | ctgacgttca | tcatctctgc | 2400 |
| gtgtgttggc | atttgcgtgg | tcgtgtgtt | gggggtgttc | tttgggatcc | tcatcaagcg | 2460 |
| acggcagcag | aaatgttgcgt | agtacacatg | cgccggactgt | tcgcaggaaa | cgggactgtgt | 2520 |
| ggagccgtgt | acacccatgc | gagccatgtc | caaccaggcc | catgttgcgt | tccttgcggaa | 2580 |
| gacggagctg | aggaagggtga | aggtgttgg | atctggcg | tttggc | atctacaaggg | 2640 |
| catctgttgc | ctgtgtgggg | agaatgttgc | aattccatgt | gcccataa | tgtgtgggaa | 2700 |
| aaacacatcc | cccaaaatc | aaaaaaat | cttagacaa | gcatatgt | tgtgtgtgt | 2760 |
| gggctcccca | tatgttcc | gccttctgg | catctgttgc | acatccacgt | tcgcgttgcgt | 2820 |
| gacacagctt | atgcctatgt | getgttctt | agaccatgtc | cgggaaaacc | ggggccgtt | 2880 |
| gggctcccca | gacccgtgt | actgtgtat | gcatgttgc | aagggtatga | gtctacttgg | 2940 |
| gtatgttgggg | tcgtgttgc | ggggacttgc | cgctcgaa | gtgtgttgc | agagtcccaa | 3000 |
| ccatgttcaaa | atttacagact | tcgggttgc | tcgggttgc | gatattgtc | agagcagatgt | 3060 |
| ccatgcagat | ggggcaagg | tgcccat | gtggatggcg | tcgggttca | tttccggcc | 3120 |
| gggggttacc | caccaggatgt | atgtgttgc | ttatgtgttgc | actgtgttgc | agctgtatgc | 3180 |
| ttttggggcc | aaacccatgc | atggatcc | atggatcc | atccctggac | tgctggaaaa | 3240 |
| ggggggagcc | ctggccccc | ccccatctg | caccattgt | gtctatcat | tcatgttca | 3300 |
| atgtgttgc | atgtacttgc | aatgttgc | aaatgttgc | gagtttgcgt | ctgtatttcc | 3360 |
| ccgcacatggcc | aggggcccccc | agcgcttgc | ggtcatcc | aatgttgcgt | ttggcccgac | 3420 |
| cagtccttgc | gacacgttcc | tetaccgtc | actgtgttgc | ggatgttgc | ttggggact | 3480 |
| gggtgtgtgt | gagggtgtatc | ttgttccca | gcagggttcc | ttctgttcc | acccttgc | 3540 |
| ggggcgttgg | ggcatgttcc | accacaggca | ccgcacgtca | tctaccagg | ttggccgtt | 3600 |
| ggacccatgtca | ctagggttgc | agccctgtca | agaggaggcc | cccggttcc | cactgttcc | 3660 |
| ctccggggat | gtgtgttgc | atgtgttgc | tttgggttgc | ggaatgggg | cggccaaagg | 3720 |
| gtgtgttgc | cccccacac | atggccat | ccctcttac | cggttacatgt | aggacccac | 3780 |
| agtacccctgt | cccttgcata | ctgtgttgc | cggttgc | ctgacgttca | gccccccagcc | 3840 |
| tgaatatgt | aaaccaggcc | atgttgc | ccagccccc | tcgcccc | gagggccctt | 3900 |
| ggctgttgc | ccacccatgt | gtgtgttgc | ggaaaggccc | aaatgttgc | ccccaggggaa | 3960 |
| gaatggggat | gttgggttgc | atgtgttgc | ttttgttgc | tttgggttgc | tttgggttgc | 4020 |
| gacacccccc | ggagggtatgt | cccttgc | ccacccat | tttgggttgc | tttgggttgc | 4080 |
| cgacaaatctc | tattacttgg | accaggatcc | accagatgttgc | tttgggttgc | tttgggttgc | 4140 |
| caaaggggaca | ccatccggca | agaacccaga | gtacccatgt | ctggacgttgc | cagtgttgc | 4200 |
| caaaaggcc | atgtccatgt | agccctgtat | tgcccttgc | gagcagggttgc | tttgggttgc | 4260 |
| ctgtgttgc | caatgttgc | gggggttgc | cgaccat | tttgggttgc | tttgggttgc | 4320 |
| ggaaacccatgc | tatgttgc | tttgggttgc | tttgggttgc | tttgggttgc | tttgggttgc | 4380 |
| agctgttgc | gaggaggatgt | tttgggttgc | tttgggttgc | tttgggttgc | tttgggttgc | 4440 |
| tgagacttca | gggttccatgt | gtgtgttgc | cccaatgttgc | ccatgttgc | tttgggttgc | 4500 |
| ggtactgaaa | gcctttaggg | agctggcc | tttgggttgc | tttgggttgc | tttgggttgc | 4560 |
| agaacacaaat | cgaccat | agagactgt | cctgttgc | tttgggttgc | tttgggttgc | 4620 |
| agaacacagca | atgttgc | tatccatgt | tttgggttgc | tttgggttgc | tttgggttgc | 4680 |

-continued

| |
|--|
| acttttttttgg ttttgttttt tttaaagatga aataaaagacc cagggggaga atgggttg 4740 |
| tatggggagg caagtgtggg gggccttct ccacaccac ttgtccatt tgcaaatata 4800 |
| ttttggaaaa cagcta 4816 |
| SEQ ID NO: 27 moltype = DNA length = 6831 |
| FEATURE Location/Qualifiers |
| source 1..6831 |
| mol_type = other DNA |
| organism = Homo sapiens |
| SEQUENCE: 27 |
| ccaggccccca ttgttcccggtttccagca tggctgcat tacctgacca ggcacacgc 60 |
| cgggtcttcgttgcaggccgg ggagaatgtgccaagacaat ttctgtttt cacaggccgg 120 |
| gtttctcaac ggtacttgttggcgttgc ttctgtcgag cgagtcatgg cccgaaggca 180 |
| gaactactgttgcgttgc ttctactc aggatgcacgc cgagggtggc ccaaggggcc 240 |
| acgatgtggc ttggatgttctt gtgcgttgc ttctgtcgatggcataaagaa 300 |
| aactcttca caataaacag ttgttgcacatgaaagacccgcggactgtggcggtaa 360 |
| gggaagaacc tgaccctgca gtgcttcgg gatgtcggca ccaccttca cgtcaaggct 420 |
| cagcaccaga tgctgttcta taaggatgac gtgtgtttaa acacatctc ctccatgaag 480 |
| agcacacaga gtatattttat tccctgttgcggatctatgactcaggagatataaaatgt 540 |
| actgtgttgc tgaacaacaa agaaagaaacc actgcagat accagggttggtggaa 600 |
| gtgcccgatccagggtgtac actggacaag aaagaggccca tccaaagggttggatgttgc 660 |
| gtcaactgtt ctgtcccaga ggaaaaggcccaatataacttcacatgttgc 720 |
| ctaaatggaaa aatgttcaatgttgc gatgttgc gagaagaaacttcggatgttgc 780 |
| atactggat tcccccgttgc ggaaacaggac cgcgttttat cttccatgttgc 840 |
| atcatttctgttgc gatgttgc gatgttgc gatgttgc gatgttgc 900 |
| acggaatcttcttccatacc caagtccatc atcagccccca cccggatgtatcatggaa 960 |
| gttcagatccatcaatgttgc caccatgttgc tggacttcgg tggcccgaggaa 1020 |
| atcataatttcggaaaggcccaatgttgc gcccacaaca gacatggccaa caaggctgttgc 1080 |
| tactcgttca tggccatgttgc ggacacatgttgc gatgttgc gatgttgc 1140 |
| cgcataatccatcaatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 1200 |
| cttggatcttcttccatacc tcttgcacatcgttgc gatgttgc gatgttgc 1260 |
| ccaggagccatcccgccatc tttccatccatc cagaaggaaatatgttgcgttgc 1320 |
| caagatttca ccaagatagtc ctcaagtcg gacagtggca cgtatctc cacttcgg 1380 |
| attgacaatggatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 1440 |
| cagcccgaggatccatccatc tttccatccatc tttccatccatc aaggacacatc 1500 |
| cgttgcgatccatcaatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 1560 |
| gttttggatgttgc gatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 1620 |
| gaagacgttgc aataccatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 1680 |
| gagggttgc ggggttcaatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 1740 |
| agtaaggatgttgc tggatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 1800 |
| ggtcccatcaatgttgc tttccatccatc tttccatccatc aaggacacatc 1860 |
| tcaaataatccatccatc tttccatccatc tttccatccatc aaggacacatc 1920 |
| tattactgttgc cagccatccatc tttccatccatc tttccatccatc aaggacacatc 1980 |
| ctgacatgttgc gatgttgc tttccatccatc tttccatccatc aaggacacatc 2040 |
| ggagtgtatccatccatc tttccatccatc tttccatccatc aaggacacatc 2100 |
| gccaaggcaga tggccatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 2160 |
| gagaaatgttgc gatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 2220 |
| agaaacatgttgc gatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 2280 |
| tacacggaaatccatccatc tttccatccatc tttccatccatc aaggacacatc 2340 |
| gagacatgttgc gatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 2400 |
| agaaacggaaatccatccatc tttccatccatc tttccatccatc aaggacacatc 2460 |
| gacatccatc tttccatccatc tttccatccatc aaggacacatc 2520 |
| ttagaaaccttgccttc catccatccatc tttccatccatc aaggacacatc 2580 |
| taaatccatc tttccatccatc tttccatccatc aaggacacatc 2640 |
| cccacataccatccatc tttccatccatc tttccatccatc aaggacacatc 2700 |
| agatccatccatc tttccatccatc tttccatccatc aaggacacatc 2760 |
| taaaagacccatccatc tttccatccatc tttccatccatc aaggacacatc 2820 |
| atttctatgttgc tttccatccatc tttccatccatc aaggacacatc 2880 |
| gggttgcgttgc gatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 2940 |
| ccccctgttgc tttccatccatc tttccatccatc aaggacacatc 3000 |
| acccgcgttgc gatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 3060 |
| atgaaggatgttgc gatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 3120 |
| cctcttccatc tttccatccatc tttccatccatc aaggacacatc 3180 |
| aatggaaatgttgc tttccatccatc tttccatccatc aaggacacatc 3240 |
| ggcaggcaggatccatccatc tttccatccatc tttccatccatc aaggacacatc 3300 |
| cacccatgttgc gatgttgc gatgttgc gatgttgc gatgttgc gatgttgc 3360 |
| gaatttagccatccatc tttccatccatc tttccatccatc aaggacacatc 3420 |
| gattatccatc tttccatccatc tttccatccatc aaggacacatc 3480 |
| ccaggccatccatc tttccatccatc tttccatccatc aaggacacatc 3540 |
| taaaaggatgttgc tttccatccatc tttccatccatc aaggacacatc 3600 |
| agaggctgttgc tttccatccatc tttccatccatc aaggacacatc 3660 |
| actttccatc tttccatccatc tttccatccatc aaggacacatc 3720 |
| aaatttgcgttgc tttccatccatc tttccatccatc aaggacacatc 3780 |
| tctataatgttgc tttccatccatc tttccatccatc aaggacacatc 3840 |
| gtttttgttgc tttccatccatc tttccatccatc aaggacacatc 3900 |
| ggcgcacacttgc tttccatccatc tttccatccatc aaggacacatc 3960 |

-continued

| | | | | | | |
|--------------|-------------|--------------|-------------|------------|-------------|------|
| gcctcccgag | tagctggat | tacaggcaca | tgtcaccacg | cctggtaat | ttttgtat | 4020 |
| tttagtagaga | tggggttca | ccacgttgc | caggctgatc | tgcatagtc | gacc | 4080 |
| gatctgcgg | ccttgcctc | atgtgtgc | cacaggcctt | tgggttggga | ttcaggcgt | 4140 |
| gagccacat | gcccaccta | gacttttg | acaatatgat | gaaagtgtt | ggttccttc | 4200 |
| cccaacacac | acacacccgag | ttgtatcag | aaaatgtcat | acaattcca | ggtttctga | 4260 |
| gttgtggct | cagattgagg | tcaaaggatc | agacgaccc | taacgaccc | catgtctcg | 4320 |
| tttgatgatct | ggggacagcc | agatcccc | tgtccaggga | gttccctagt | cccttgcac | 4380 |
| caccagagaa | gggcattgc | cacgggagct | gcaaaagacc | tattctact | cctggtgcc | 4440 |
| tacttatgca | gcacgactga | atttttgtt | ttgtttgtt | ttgttgagac | aggggctgc | 4500 |
| tctgttgccc | aggctggat | gcacgtggac | aacaatggc | caccgcaggc | tgcacccct | 4560 |
| gggctcaagc | gatcctcca | tctcagttt | ctgggtatgt | gggacagag | gcgtgaggcc | 4620 |
| ccatagctgg | ctaatttttta | atttttttt | tgcagagat | aggttttacc | atggtgccca | 4680 |
| ggctggctc | gacttctgg | gctcaagtga | tcctccctc | ttggccctgc | aaagtgtgg | 4740 |
| gattgcggc | atgaggccac | gccccggcc | tgtggagcac | acatgagttt | aaaattactt | 4800 |
| tccttcgtc | ctatatttcc | gaggagaaa | cttcatgcgc | agggatctt | cttagtgat | 4860 |
| ttaatggcta | aaaggctctgt | ctgaatccag | gacgctggct | ttagccttc | tcggcagctg | 4920 |
| ccgtaaaccc | gggtctaaa | cctgaagcat | cccaggagca | cccactccag | gagtttctc | 4980 |
| ggccgcggaa | ctcattttt | agagccccc | cttgggttct | catgtgttta | tcggtactg | 5040 |
| aaggacttaa | aatggccctt | agccaacaca | cagttttttt | tttccctttt | ctgaccccaa | 5100 |
| gagggtcaggc | accatttca | tgagcatata | ctggcggccc | catagcgtt | ctctgatgg | 5160 |
| ctaaactgaac | ccactccccg | accttagactc | aagacaggcg | aagtgcac | taggtcaaca | 5220 |
| ttcaactact | aaagcaacga | ctgtggggcg | atttgttctc | ccgtgtgtt | tggaatggtg | 5280 |
| tctggagaca | tttttgggtt | tcacagctgt | tgggggtgtc | tccgggaaat | tgggggttag | 5340 |
| aaaccaagca | tgctcttaaa | catecttacag | gcacagaaac | gtctcccaac | accaaggatg | 5400 |
| atcaagtc | aatgcataat | aatggccagg | ttgagaaact | ctgcacagaa | gcatccagtt | 5460 |
| atttgcgtgt | ttgtctcaaa | agctttgtgt | catcatgtc | tgtgttctg | acgctgtgt | 5520 |
| gggtgtggc | ggtggaaaga | ttacaagat | cacatggcgt | ctgtctct | ggaaggataca | 5580 |
| acccaggataga | gatgcagact | aacagagago | caatttacaaa | cgatgtgt | aagcgtcatg | 5640 |
| gtggaaaatt | aaaagctca | acaaggccac | atgggggggg | cttccaaac | agactttttgg | 5700 |
| ggatccaggaa | aggcttaaga | ggaaaatggg | tctcaccaaa | gccttgacca | taggcagagg | 5760 |
| gttaccaggcg | aaaaggatggg | gttggaaaca | ttggaggacaa | aaggaaagaa | tgcaggaaag | 5820 |
| ccttggggca | aggggatggg | gggtccctg | gaggggatggc | agcaggccg | tctgtcagac | 5880 |
| ccaaatggcc | tccaggcc | aaagccaatt | agtcccttc | aaaaagctgt | cactgtcccc | 5940 |
| taagaattgc | tgccaggctc | ccactggct | gactcagtct | ttgagatgt | taaggaggag | 6000 |
| gtctctgtt | ggtacacacc | aagaaccttc | cccaagcac | ctgtttttaa | gacttccac | 6060 |
| cagegtcatt | ggcgtgttgg | gaagaaaccc | tctgccacag | aggccaggct | cagccttgc | 6120 |
| ctaacacccg | aaggccaaat | ggaaaatggtaa | acggggaggaa | gatgtctccc | cagcaggct | 6180 |
| tttgaggaca | gtttccctg | cagaagatct | caacctgggg | tccacagat | ggaatgttta | 6240 |
| gagttgggg | ctaggccaaat | atggcggagg | cagggtgggg | ccccccac | aatgtcaggc | 6300 |
| taccatcagg | tgtatgttgc | gttgggttta | aactgttct | gtaaaataat | aatgggttgc | 6360 |
| agccagtc | aagcaaggac | agtctctaa | tagatataaa | acacccctgt | ctgggtatca | 6420 |
| gccgcttccc | gataagatct | caggagctgg | gcaagcggc | tggagcatgc | gcaccaagag | 6480 |
| gcaaaatggc | ggaatttaa | cgtatataaa | cctacccctc | tctggaaac | cgactgtgt | 6540 |
| aaggggaaa | atgcctca | tgagcatgc | cgcaacttca | gtatcacac | tgtgcatgc | 6600 |
| acccttc | atgtgtgc | ggtaccaca | tacggggaca | gcctgtgc | agggaaagaa | 6660 |
| caggggagat | gagacgtaaa | tcccagaact | atgccaata | cataaaaccc | caagtttaagg | 6720 |
| gttggcaggc | gcaacttagat | ctctcaagtt | gcctgcgtc | cccaagtgt | gtgtacttcc | 6780 |
| ttttgttct | gtctaaaaac | ttttataaa | acttcactc | ctgtcttaaa | a | 6831 |

SEQ ID NO: 28 moltype = DNA length = 2956

FEATURE Location/Qualifiers

source 1..2956

mol_type = other DNA

organism = Homo sapiens

SEQUENCE: 28

| | | | | | | |
|-------------|-------------|--------------|--------------|-------------|-------------|------|
| ggggctgtgt | gctcgtgttgc | gagggggag | ttgtgtccac | cttgcgcacg | tcgcgtacccg | 60 |
| tggggctgtc | ctgggaaaggc | ggacggcag | cggccgggt | ccgcacttgc | ccgccttgc | 120 |
| tgccctgtct | cgcccggtgc | atcctca | gggacgcagg | gaccgtttt | aaatcacagg | 180 |
| ggcgtgtgtc | ggccgtccct | aggcttcat | gtctatataat | ttcccccattc | actgcccga | 240 |
| ctatctgttgc | tccggcaaga | tgtactgttgc | gtatgttgc | acccagccca | tggggagat | 300 |
| cggcctcagc | cccccgcaagg | atggcccttc | ctaccatgc | tcccaagacc | cgtcaggatt | 360 |
| tgaccgtgtc | tgcaaaactga | aggaccgtct | gcctccata | tgggtggaa | ccacagaagg | 420 |
| ggaggtgtgg | agcggggagc | tccgggtggcc | ccctggggag | ttcttgggtcc | aggaggatga | 480 |
| gcacgataac | tgcgacggaga | caggcggaa | aaataaaagaa | cgtatgtgt | cctgtggact | 540 |
| cccatgggtc | ataccagcc | gcatactgttgc | ctgaactgttgc | tttttccat | catgcggaa | 600 |
| gaagagatgt | agccgcatt | gttctggaaa | tgtccaaacg | ggcttctgtt | ttgcacccgtc | 660 |
| agatcacoga | gttgggtttc | ttttttttt | ttgcctttt | tttttttttgc | aatttgcgcg | 720 |
| gcagtggagc | cctctgacaa | tttgcacgg | cctctgacaa | aggaagctgc | ttagagccag | 780 |
| gggggttagt | ggtgggggg | gcgactgttg | tttttggat | cattatctga | actcaggcag | 840 |
| ccttagtagag | gcaactgttgc | gatccaaat | ggtcttgggt | ggtggggagg | ggggcatgt | 900 |
| caaagcaacg | aaggacatt | tggggtaaga | aaaacaaat | gaggcggaa | aaaaaaataca | 960 |
| tgtttttaag | aaaacattga | gcagagaact | gcagccagg | tgcgtc | agacattc | 1020 |
| tctggctgt | gggacatcag | aaaacaaat | cttcatctc | ctctccagg | tcacccaccc | 1080 |
| caccctttgc | tttcatatca | ggtgggttgg | tctatatgac | agggaggaga | gtaaaggaga | 1140 |
| gcaggagacaa | ttgggtgcct | gcaaaaccc | ctggggatgc | atgcgggg | aggaaagg | 1200 |
| accccattct | actccatggc | ctctctgttgc | ccagctgttg | taggtctaca | tagccatgt | 1260 |

-continued

| | | | | | | |
|-------------|-------------|------------|----------|-------------|------------|------|
| gatcggttt | taagaggcag | tgcttttcag | cttttccc | tgatataatcc | attttgc | 1320 |
| cacagcactt | ttaggtagtag | tgagagact | tcctgc | gttggaagcc | ccagggtgga | 1380 |
| cactcagcac | gaaggctct | cccttaactg | ctgc | aagacttgt | cccgagatgg | 1440 |
| agtggggcgt | gtcttcagg | ctggcc | cctt | ccgcac | ccctgcccc | 1500 |
| gccccagcag | ccatgggtac | atgggtcccc | agtcac | tggattccc | ccagtctgc | 1560 |
| cacgtc | actca | catggggat | cttgg | ttttctgt | ggagct | 1620 |
| ggagagcaga | cgtgg | tttttt | ttatgtgt | gttggggagg | tgatgtcat | 1680 |
| aggctgtcg | ggcaac | cggat | tttttt | tttttttt | tttttttt | 1740 |
| cttctccca | tggtgtccc | ggacaagg | tttttt | tttttttt | tttttttt | 1800 |
| gtccttactc | aacaggtgtc | ctcatccc | ccca | tttttttt | tttttttt | 1860 |
| cagtgtatg | aaaaaaatgg | ccttttttt | tttttttt | tttttttt | tttttttt | 1920 |
| tgaagcaga | ggttgc | tttttttt | tttttttt | tttttttt | tttttttt | 1980 |
| gagggtggact | cagggc | tttgc | tttttttt | tttttttt | tttttttt | 2040 |
| cetcagg | cggat | tttttttt | tttttttt | tttttttt | tttttttt | 2100 |
| ttaaatacct | caagg | tttgc | tttttttt | tttttttt | tttttttt | 2160 |
| tgttgacaca | gcaat | tttgc | tttttttt | tttttttt | tttttttt | 2220 |
| gettc | tttttttt | tttttttt | tttttttt | tttttttt | tttttttt | 2280 |
| tcc | tttttttt | tttttttt | tttttttt | tttttttt | tttttttt | 2340 |
| cccc | tttttttt | tttttttt | tttttttt | tttttttt | tttttttt | 2400 |
| ct | tttttttt | tttttttt | tttttttt | tttttttt | tttttttt | 2460 |
| ctgttactc | aat | tttttttt | tttttttt | tttttttt | tttttttt | 2520 |
| cgttaaaaat | tttttttt | tttttttt | tttttttt | tttttttt | tttttttt | 2580 |
| tcctttagacc | tttttttt | tttttttt | tttttttt | tttttttt | tttttttt | 2640 |
| gtcagg | tttttttt | tttttttt | tttttttt | tttttttt | tttttttt | 2700 |
| gag | tttttttt | tttttttt | tttttttt | tttttttt | tttttttt | 2760 |
| gattgtattt | aaaaaaatcc | tttttttt | tttttttt | tttttttt | tttttttt | 2820 |
| tgttaactt | tttttttt | tttttttt | tttttttt | tttttttt | tttttttt | 2880 |
| aactactt | tttttttt | tttttttt | tttttttt | tttttttt | tttttttt | 2940 |
| aaaaaaaaaa | aaaaaa | aaaaaa | aaaaaa | aaaaaa | aaaaaa | 2956 |

SEQ ID NO: 29 moltype = DNA length = 2262
FEATURE Location/Qualifiers
source 1..2262
mol_type = other DNA
organism = Homo sapiens

SEQUENCE: 29
ctggggagta atagcatggg caaccattat cctgtctcgc cgccacccag gacatggctt 60
ctgttcaat gccaagttag tacacctatg taaaacttaga aagtgttgc tcgaggcctt 120
ccctgcaat gtaccccgca gctaaaggaa agatggaaag gcccaggctg ttataaaag 180
acatcctcaa atgtacattt ctgttgtttt gagggttgtt cttttatatc ctcaagttaa 240
attataactac tgaagaatgt gacatggaaa aaatgcattt tttttttttt gtttttgtt 300
agagagctca gaaaatgtct cagcaagtct tgcaagaggaa atgtcgccc aagtttgc 360
agacataat ggcgtgtt tttttttttt ggtagtgcggt ggacttactc ctttttgtc 420
agaaggcccc caaagacagt gaagttgttgc ccaagttgttgc tttttttttt ggggttttgg 480
agttctccag taaaatccatg acccttctgg aactcttgcg agagcacagc cttcccttgc 540
acttggaaagc caagacctgt cggcgctgtt tttttttttt tttttttttt tttttttttt 600
gatttagaaact gggccacacc ctgaaaggctt cttttttttt tttttttttt tttttttttt 660
cagtgttgggg attttttttt tttttttttt tttttttttt tttttttttt tttttttttt 720
aggggcgcacc actgttgcactt tttttttttt tttttttttt tttttttttt tttttttttt 780
ttaagagtgt tgatttcaac tttttttttt tttttttttt tttttttttt tttttttttt 840
gggttacactt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 900
tcaggattt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 960
agccctactt aaggatgttgg tttttttttt tttttttttt tttttttttt tttttttttt 1020
tttgtttagt acatcatttctg tttttttttt tttttttttt tttttttttt tttttttttt 1080
atcaaccccg aacaccccttgg tttttttttt tttttttttt tttttttttt tttttttttt 1140
agaccatgtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1200
ttgttggaaa tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1260
aaaatgcac tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1320
ctgcagatac tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1380
atttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1440
accatattttt aatgtttttt tttttttttt tttttttttt tttttttttt tttttttttt 1500
gtaacgtgt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1560
acactgcaca tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1620
ttgttggaaa tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1680
acttcatttctg tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1740
cactttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1800
cgtaggtttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1860
agaaatctgtt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1920
cactttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1980
aagtgtttaga tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2040
actttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2100
cacttggaaat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2160
gaatgcctac tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2220
aaaaactttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 2280

SEQ ID NO: 30 moltype = DNA length = 4909

-continued

| FEATURE | Location/Qualifiers |
|---|-------------------------|
| source | 1..4909 |
| | mol_type = other DNA |
| | organism = Homo sapiens |
| SEQUENCE: 30 | |
| tagacgcacc ctctgaagat ggtgactccc tccctgagaag ctggaccctt tggtaaaaga | 60 |
| caaggccttc tccaagaaga atatgaaaagt gttactcaga cttatgtt tcatacgct 120 | |
| actgattttct tctctggagg ctgtataatg caaggaactg gaagaaaaaa taattttagt 180 | |
| gtcatctgca aatgaaattg atgttgcctt ctgtccctt aacccaaatg aacacaagg 240 | |
| cactataact tggtaaaag atgacagacca gacacctgtt tctacacaaac aagcttcag 300 | |
| gattcatcaa cacaaagaga aacttgggt tggctctgtt aagggtggg attcaggaca 360 | |
| ttactattgc gtggtaaaaa attcattta ctgcctcaga attaaaaataa gtgc当地 420 | |
| tgtggagaat gaggcataact tatgttataa tgccacaaggcc atatttaaagc agaaaactacc 480 | |
| cgttgcggaa gacggaggac ttgtgtgccc ttatatggg ttttttaaaa atgaaaataa 540 | |
| tgagtttacca aatttacatg ggttaaaaaa ttgcaaaaccc ctatcttgc acaatataca 600 | |
| ctttagtggaa gtccaaagata ggctcatacgat gatgttgc gctgaaaaagc atagaggaa 660 | |
| ctatacttgt catgcattt acacatactt gggcaagcaaa tttttttttt cccggtaat 720 | |
| agaattttt actcttagagg aaaacaaac cacaaggccctt gtgtttttttt gccccagctaa 780 | |
| tgagacaatgg aagtagact ttggatccca gatataatccat cttttttttt ccacccggcc 840 | |
| gttgagtgtc attgtttttt ggaatgtggaa ttggctctgtt attgttgc gatgttgc 900 | |
| gctaggggaa gactattaca gtgtggaaaa ttctgc当地 acacaaaatgg gttttttttt 960 | |
| cacagtgtt aatatatccgg aaatggaaat tagatttttt aaacatccat ttacccctttt 1020 | |
| tgccaaagat acacatggta tagatgtttt atatatccat cttttttttt ccacccggcc 1080 | |
| ttttccaaatgg cccatgtttt ttatgttgc ttatatccat cttttttttt ccacccggcc 1140 | |
| tttcatcttat aaaaatccca agatgtttt ttgttgc当地 tacagggtt cttttttttt 1200 | |
| ttttctccca aaaaatggcc cccatggaa gacctatgtt gatataatccat tttttttttt 1260 | |
| gactgttgggaa gaaagggttca cctctgtttt ttgttgc当地 ttgtttttttt 1320 | |
| gggtttttttt aaacagtgtt gatataatccat ttgtttttttt ggtttttttt 1380 | |
| ggaaaggccattt gttttttttt tttttttttt ttgtttttttt 1440 | |
| tttagtgc当地 gaaacatccat cttttttttt ttgtttttttt 1500 | |
| catgttataat gttttttttt ttgtttttttt 1560 | |
| ccaaagactt gaaatgtt gatataatccat ttgtttttttt 1620 | |
| ccgctgttca gggggacttta cacagggtttt acactgttca aagacaaaggcc ttgtttttttt 1680 | |
| ttgtttttttt ttgtttttttt 1740 | |
| accaggccactt aagggtttttt ttgtttttttt 1800 | |
| agtttgc当地 ttgtttttttt 1860 | |
| catgttactt ttgtttttttt 1920 | |
| ttttttttttt ttgtttttttt 1980 | |
| ttttttttttt ttgtttttttt 2040 | |
| ataatcccaatgg cccatgtttt ttgtttttttt 2100 | |
| ccagcccccactt gaaatgtt gatataatccat ttgtttttttt 2160 | |
| ttgttgc当地 ttgtttttttt 2220 | |
| ccggggagactt gggggacttca ttgtttttttt 2280 | |
| gagcaagactt cccatgtttt ttgtttttttt 2340 | |
| aaaaaaaggccatgggactt gggggacttca ttgtttttttt 2400 | |
| acagatgttccatgggactt gggggacttca ttgtttttttt 2460 | |
| cctgttactt gggggacttca ttgtttttttt 2520 | |
| ttttttttttt ttgtttttttt 2580 | |
| ccatacacat cccatgtttt ttgtttttttt 2640 | |
| aaaaacttccatgggactt gggggacttca ttgtttttttt 2700 | |
| agtttccatgggactt gggggacttca ttgtttttttt 2760 | |
| ttttttttttt ttgtttttttt 2820 | |
| ttttttttttt ttgtttttttt 2880 | |
| ccatgttccatgggactt gggggacttca ttgtttttttt 2940 | |
| ttttttttttt ttgtttttttt 3000 | |
| ccagcccccactt gggggacttca ttgtttttttt 3060 | |
| ttttttttttt ttgtttttttt 3120 | |
| ccgacccttccatgggactt gggggacttca ttgtttttttt 3180 | |
| ttttttttttt ttgtttttttt 3240 | |
| cgacttccatgggactt gggggacttca ttgtttttttt 3300 | |
| atctggggactt gggggacttca ttgtttttttt 3360 | |
| gggtgttctt gggggacttca ttgtttttttt 3420 | |
| taatgtttttt ttgtttttttt 3480 | |
| ttttttttttt ttgtttttttt 3540 | |
| atctttttttt ttgtttttttt 3600 | |
| gggtttttttt ttgtttttttt 3660 | |
| gcttggggactt gggggacttca ttgtttttttt 3720 | |
| aatgtttttt ttgtttttttt 3780 | |
| ttttttttttt ttgtttttttt 3840 | |
| aaatttttttt ttgtttttttt 3900 | |
| agaggactttt ttgtttttttt 3960 | |
| ggtttttttt ttgtttttttt 4020 | |
| ttttttttttt ttgtttttttt 4080 | |
| ctgtttttttt ttgtttttttt 4140 | |
| catggggactt gggggacttca ttgtttttttt 4200 | |
| ccatgtttttt ttgtttttttt 4260 | |

-continued

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|------------|------|
| gccaattttgg | aaacttaggt | tagtgacaaa | attggccaga | gagtgggggt | gatgtgacc | 4320 |
| aagaattaca | agtagaatgg | cagcttggaa | ttaaggaggg | acaagaatca | atggataa | 4380 |
| gtgggtggag | gaagatccaa | acagaaaaatg | gcaaaagtta | tccccatctt | ccaagggtt | 4440 |
| aattctggag | gaagaagaca | cattcctagt | tcccccgtaa | cttccttga | tttattgtcc | 4500 |
| ccactaaac | aaaacaaaaaa | acttttaatg | cttccacat | taatttagatt | ttcttgca | 4560 |
| tttttatgg | catttttta | aagatgcctt | aagtgttgaa | gaagagttt | caaatgcac | 4620 |
| aaaatattta | attaccggtt | gttaaaactg | gtttagcaca | atttattttt | tccctctt | 4680 |
| gccttccta | tttgcataaa | aaggatattg | gccatTTTTT | aatgcacatt | tttgataat | 4740 |
| tatgtttgt | ctagttgt | aaggagttt | ttttaacctg | tttataataat | tttgagcag | 4800 |
| aagccaaatt | ttttgtat | taaagcacca | aattcatgt | cagcatgtat | cacggatcaa | 4860 |
| tagactgtac | ttatTTTCA | ataaaaattt | caaactttt | actgttaaa | | 4909 |

| | | | | | | |
|---------------|-------------------------|---------------|-------------|-------------|-------------|------|
| SEQ ID NO: 31 | moltype = DNA | length = 2210 | | | | |
| FEATURE | Location/Qualifiers | | | | | |
| source | 1..2210 | | | | | |
| | mol_type = other DNA | | | | | |
| | organism = Homo sapiens | | | | | |
| SEQUENCE: 31 | | | | | | |
| cgcgccccctc | cctcttcg | gacctggcg | tgcggcgccc | cggagtggcc | ctttaaaagg | 60 |
| cagcttatttgc | tccggagggg | gccccgggg | ggcgccgacc | ggggctcgag | gccccggccc | 120 |
| tccccctctcc | cctccctctgt | cccccgctcg | ctcgctggct | agctcgctgg | ctcgctcgcc | 180 |
| cgtccggcgc | acgtcccgcc | tccgcgtcgt | ggctccgcgt | tccgggtgcgc | ggcgtggagc | 240 |
| gcoagccgggt | ctggacgcgc | ggccgggggt | gggggtgggg | agcgcggcgc | gcaagatctc | 300 |
| ccccgcgcag | agccccccct | gcccacgggg | gaggcctcg | ccgcgtatgc | agagatgggc | 360 |
| agtaaagggg | tgacggcg | aaagatcgcc | agcaacgtgc | agaagaagct | cacccgcgc | 420 |
| caggagaaagg | ttctccagaa | gctggggaa | gcagatgaga | ccaaggatga | gcagtttgag | 480 |
| cagtgcgtc | agaatgttcaa | caaggactg | acggaggggca | cccggtgc | gaaggatctc | 540 |
| cggacactacc | tggccctccgt | caaaaggat | cacggggctt | ccaagaagct | gaatgagtgt | 600 |
| ctgcaggagg | tgtatgagcc | cgatggccc | ggcagggtat | aggcaaaacaa | gatcgagag | 660 |
| aacaacgacc | tgctgtggat | ggattaccac | cagaagctgg | tggaccagc | gctgctgacc | 720 |
| atggacacgt | acttcggccca | gttcccccac | atcaagtcac | gcatgtccaa | gccccggcgc | 780 |
| aaagctgtgg | atcacagac | tgccggc | cactacgat | cccttcaaac | tgccaaaaaag | 840 |
| aaggatgaa | ccaaaaattgc | caaggccag | gaggagctca | tcaaagccca | gaagggttt | 900 |
| gaggagatga | atgtggatct | gcaggaggag | ctggcgtccc | tgtggaaacag | ccgcgttagt | 960 |
| tcttacgtca | acacgttcca | gaggatcg | ggccctggagg | aaaacttcca | caaggatgt | 1020 |
| agcaagctca | accaaaacct | caatgtgt | ctggcgtggc | tggaaagaca | acacgggagc | 1080 |
| aacaccccca | cggtcaaggc | ccagccc | aqaaaaatg | aactgtttc | gcggctgcgc | 1140 |
| agaaagaaga | acagtgcac | cgccgtc | aaagggaaaca | agagccctc | gcctccagat | 1200 |
| ggctccctcg | ccggccaccc | cgatcaga | gtcaaccac | agccagagcc | ggccggccgg | 1260 |
| gccacccccc | ggggccaccc | cccaaaatgc | ccatctc | cacggggc | ctcggaggt | 1320 |
| gcgggtggaa | cccaacactgc | ggctggagcc | caggagcc | gggagacggc | ggcaagtgaa | 1380 |
| gcagccctca | gtctcttcc | tgctgtcg | gtggagac | tccctc | tgtgtatggc | 1440 |
| accgtggagg | ggggcgtgg | ggccggcc | ttggac | ccccagg | ttatgtcaag | 1500 |
| gtacagggcc | acacgacta | cgccggc | acggac | atgtcgat | caaggcttgt | 1560 |
| gtatgtggtc | tgtgtatccc | cttccagaa | cctgt | ggatgtt | ctggctat | 1620 |
| ggcgtgttgg | agagcgact | gaacc | aggatgtt | agaatgtt | tggcgtt | 1680 |
| cccgagact | tcactcgat | ggtccat | cgccggggcc | caggcag | ccgggggt | 1740 |
| gaagaacacc | tctcccaa | aaatgttgg | ttctttttt | tgtttgtt | tcgttttca | 1800 |
| tcttttga | agcaaaagg | aatcaagagg | agacccca | gcagagggc | gttcccca | 1860 |
| agatttagtc | gttttccaa | gagccgc | ccggca | ccggega | caccagtgtt | 1920 |
| cctgttgc | ctgtgttccc | tagtggat | tctggc | ctggctgt | ccgcatgt | 1980 |
| gcctggccgc | aggggccggc | ttgggg | cgagccacca | tgttgc | aaatctggc | 2040 |
| cgccgcaccc | gggcacgggt | ccttttcc | tggcag | tgttgg | gccccacac | 2100 |
| cagccttagcc | tggctgtcc | ccgcagac | tctgtgt | gtttgaaaat | aatcttagt | 2160 |
| gttcaaaaca | aatgaaaca | aaaaaaaat | gataaaaaat | ctcaaaaaaa | | 2210 |

| | | | | | | |
|---------------|-------------------------|---------------|------------|----------|------------|-----|
| SEQ ID NO: 32 | moltype = DNA | length = 4664 | | | | |
| FEATURE | Location/Qualifiers | | | | | |
| source | 1..4664 | | | | | |
| | mol_type = other DNA | | | | | |
| | organism = Homo sapiens | | | | | |
| SEQUENCE: 32 | | | | | | |
| gcagccctcc | ggcgctg | gctttcc | ccggccggc | ttagcc | ggacccggg | 60 |
| agaatttcc | cagaaaaat | gcccacggc | ggcgccggc | ggcgatgt | ccggagcc | 120 |
| tgcgcgttcc | atcagaaga | ctgttgc | acataactgc | caagatgc | gtccctcccc | 180 |
| ctccagcacc | ccccggccc | ccgcgtt | cactggccaa | tacagaa | cctaccttga | 240 |
| ataagacaga | gcaggctgg | agaaaatgc | tcccttctga | tatcago | aaaaac | 300 |
| taaagaagac | ggtcacca | gacaga | actgt | accaata | ggacaaac | 360 |
| gtgtgtgg | cggtgtgt | ggcttgg | gaggcggc | atttggc | ggaggtgt | 420 |
| ggggaggccgg | tggaaat | ggggggcc | gaccc | tctgg | ggaggcc | 480 |
| cttggatgtcc | gaagctgt | tccacggc | acagg | atgttgc | ggaagcc | 540 |
| caccattgtt | gccacccgg | gaaagat | caatgc | acccttca | ccccca | 600 |
| gcccaggcc | gtttccgt | ccttcc | gccac | aga | tccca | 660 |
| gaaaccgaat | gcccacccca | aggccc | gac | tttgc | atgc | 720 |
| cgtacccat | tactcca | caat | caatgc | caacccgg | tcccac | 780 |
| tgcccgagg | ccccaggcc | ccca | ccactcc | tccccc | cctggaa | 840 |

-continued

| | | | | | | |
|---------------|----------------|--------------|-------------|-------------|-------------|------|
| ggggcactgc | tttgggagga | ggctcaataac | gtcaagtcccc | cttgagatccc | tcctcgccct | 900 |
| tctccaaaccg | gctccccctg | ccgccttaccc | ccagcagggc | cttggatgac | aaacccctc | 960 |
| caccacccctc | cccaactggc | aacaggccct | ccatccacag | ggaagegggtt | ccccctctc | 1020 |
| cttctcagaa | caacaagccct | ccagtcctt | ccactcccg | gccttcggcc | tcctcacagg | 1080 |
| ccccacccctc | gcccacccct | cccagcaggc | ccggggccgc | tccttcgtct | ccaagttcca | 1140 |
| ggggcaatga | cgaaaacccca | agactcccaac | agcggaaatcc | gtccctcaat | tcgtccacgc | 1200 |
| ccccgttacc | tccgtccagga | cgttcaggtc | cttttcctcc | ccggcccaat | gagagacccc | 1260 |
| cacttcctagg | gaggggaccccg | ccaggccat | caggccccct | ccacccactc | ccttcagtaa | 1320 |
| gcagaaaacgg | oagcacatct | cggggccctg | ctgtcaaaaa | tcagttgcca | tccaggagtg | 1380 |
| gagttagacg | tcccaagggt | ggggccaggc | ctccccctcc | tcctgtatagg | cccagtgtcg | 1440 |
| gggcacccctc | ccccacccctca | ccatcaacat | ctattagaaa | tggctccaa | gacttcctat | 1500 |
| gtgaagatga | gtggggaaagc | agatttctact | tccatccgat | ttcccgattt | ccaccttcag | 1560 |
| agccatcatgt | acaaacgacc | aaaagttatc | ccagcaaaact | ggcaagaaaac | gaaagccgga | 1620 |
| gtggatccaa | ccgaagagaaa | agggggtgtc | caccactccc | tcccatcccg | aggtgtatctt | 1680 |
| tgcctgtet | tctctatccca | agctcaagag | tcgtttctgt | tgttatctaa | gaactgcata | 1740 |
| ccttcctcc | tgccttc | ttgtgcctc | atgtatggc | aggagggaaag | gtggggagggg | 1800 |
| gagtggaaat | atgcgtgtgt | gggtggaaat | cggtaaagaaa | tgcacccatgc | ttttcatatt | 1860 |
| gtgtttatcc | tccaggctat | tgcttgctt | agctgcagcc | tgccctgtgt | ggctgctggg | 1920 |
| gtcgatagge | ttttgtcgta | ataggcgag | atgacttgcc | tcccaatccc | ccaccaacca | 1980 |
| aattcaaaa | tcaactgttt | atttgttaca | gactgttaat | attaaaggccc | ctgagagctg | 2040 |
| ttttctccgg | ttcttttttc | gcatgtttgg | cctccctct | gtttctatga | accacagacc | 2100 |
| acctaagcaa | gtgtgtcgat | aagggtctac | tggaaacttt | cagtccacagg | atgtccat | 2160 |
| tttggcgtc | cgagcttggc | tcttgcggac | agctgtccaa | ttttatata | atgtgagccc | 2220 |
| catatacaat | ttttatccat | ttatataat | ttaaacaatgt | gaagttataa | tgcatccaaa | 2280 |
| atatttcaac | ctgtatcaaa | cataaaaattt | taatgagata | ttttatata | tttttggta | 2340 |
| ctgaatcttc | aaaatccaga | gtgttattta | cacttacccg | acatctccat | tcgactactg | 2400 |
| cacattttta | agtgtccatgt | agccacatgt | gggtgggtgg | tactggata | gacacacgca | 2460 |
| gtctggaaaga | tggaaatctg | tgcagaaacc | tcttggttaa | aaaacaaaaa | aggcaatgt | 2520 |
| gggttggcgc | atccaagagg | caactaaaaaa | taaaatttag | acccagccacc | ttgtttgaca | 2580 |
| cacagtttgc | ccttcgattt | tcctccctta | actccctct | tcccttaata | tctgtatata | 2640 |
| agtgttgcgtt | caaagtaccca | agggtcagaaa | ttgatttcgt | acggtttact | aaagtcatgt | 2700 |
| ggaataaaacg | cattggaaac | aatggaaag | cctgtcgggg | ttctggggct | cagaaccacg | 2760 |
| tggttcacgc | actccacttg | tcagtcggac | tcttgcctt | tgaaatggaa | gcagccttt | 2820 |
| ttccctttctg | gtcgagcaag | ctccctgaggt | ttggagagac | taggaaggct | ttgttaggagg | 2880 |
| ggaaaaaaatg | caggaaaaaa | tatcaatca | gaaacatcgaa | agaagaaggg | aaccqatttg | 2940 |
| agttgggtgg | caaaaactca | aaaatctaa | tctgtatgtt | atgtagggt | tgagcgaat | 3000 |
| agggagatgt | ctagtgtggaa | ttggagggaa | tttggtttgc | atcattttgc | taggtatctat | 3060 |
| gcaaataatag | ctccactaaa | ggaccatagg | gaagagccag | ccttgcctt | tcttatata | 3120 |
| ttttgttac | aaaattttac | ttgggactttt | aaatcttagt | atagagtgg | aaaaaaat | 3180 |
| ttccacttag | atattttaca | ttggttttgtt | taaaatttac | attactgtt | ttttaaaaac | 3240 |
| acatgaccac | atatgtatata | gtatatactac | ctaaacatgg | tatcatgttt | tcagttatgt | 3300 |
| attcatgtat | tactgggaga | tgctaccaag | aaaccaaccc | aaagaaaatt | ctgaaaaata | 3360 |
| cattttatcc | tatataatata | atgtttatcc | tatataaaaa | aaaaaaact | taggttctat | 3420 |
| ataaaatggaa | tgtctatata | attatgtat | tactgttttt | aatatgtat | tttttataata | 3480 |
| cttccttcgc | aaagtctgt | tttatgtat | tagagaacct | gtgtttctc | tcttcctatc | 3540 |
| cattcatctc | tcttcatac | agtcattttg | gtttttact | caaagagaat | caagaataaa | 3600 |
| taaggatataa | caagcttggc | aaagtgttgg | ctttttaaa | aaaatttt | ttaatctctt | 3660 |
| gcagtttgcgtt | aattttagcag | catcatttt | ttgggattt | tttattctgt | ttcaacagtg | 3720 |
| aaaaacatcc | ctatgtat | gcctaatgc | ccatccatca | aaagatggaa | tttgccttc | 3780 |
| ctagaaaata | tgacggagaa | aagtctgact | cagagaaatgt | gagtctgtat | tttataaggg | 3840 |
| gttagttagaa | ttgggacattt | cctttgcata | tctgtactt | gcaggtaccc | ttctaaatct | 3900 |
| gaaacaggggt | gatagctaa | agttgcattt | catccagaat | agattgttt | agaatgtat | 3960 |
| gtttaatgttgc | ctgttttccat | ataacacca | caccctttt | ttgaaatgtt | gcaaccta | 4020 |
| tgcatctaa | actatgtata | agttctgtgg | tttttttttt | tttttttttt | tttttttttt | 4080 |
| atgaattttt | cttcctctg | atcagatgtt | acatgtgtgt | tttttttttt | tttttttttt | 4140 |
| taaatgttac | caacatgttac | tcagttttt | taatccatcc | tttttttttt | tttttttttt | 4200 |
| gttattacatgt | ataaaaactca | agttgttgcgt | tttttttttt | tttttttttt | tttttttttt | 4260 |
| aagccaggag | aatatgtttaa | tccatgttgc | ccttactttt | tttttttttt | tttttttttt | 4320 |
| ttttttatcc | tcttaatgttgcgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 4380 |
| tttttaatgttgc | taatataatgttgc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 4440 |
| ttcccaaaaa | ttgtggatgttgc | ataatataat | ttcccttc | atagtcagg | acccgttaggg | 4500 |
| aagaaaaaaact | tttttttccat | accatgcact | atgtaaacacg | acacattttt | ctatctgtgt | 4560 |
| catcaggata | gtgttaatgttgc | taggttagag | actaccctag | acatctgtat | ttttgtatgt | 4620 |
| tagccagaca | ataaaagaaaa | gcagaatgttgc | aaaaaaa | aaaaaaa | aaaaaaa | 4664 |

SEQ ID NO: 33 moltype = DNA length = 1166
 FEATURE Location/Qualifiers
 source 1..1166
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 33
 atccccaaact gccaggatgtc tctgaaggccg actctggggc tcccttttgc ctctaaacaga 60
 cagcagcgac tttaggtctgg ataatagtca aattcttacc tcgttcttcc actgttagta 120
 agatcagatt gcgtttcttt cagttactct tcaatcgcca gtttctgtat tcgttctaa 180
 aagaagaatgt agagaagata aatctgttca tcaataccctg gaaggaaaaaaa caaataacc 240
 tcataactccgt ttggaaaaaaa acatccatca agagattttt ctttagatgt 300

-continued

| | | | | | | |
|------------|-------------|-------------|------------|-------------|-------------|------|
| ttacacaatg | aagaaagtac | atgcactttg | ggcttctgt | tgccctgtc | ttaatcttc | 360 |
| ccctggccc | cttaatgtc | attctggaa | agatgaagaa | cacaactta | tcacagatac | 420 |
| ggagttgc | ccactgaaac | ttatgcatt | attttgtca | ttcaaggccc | atgtggccc | 480 |
| atgtaaagca | atcatgaaaa | gatttttctt | caatatttc | actcgacagt | gcgaagaatt | 540 |
| tatataatgg | ggatgtgaag | gaaatcagaa | tcgatttgaa | agtctggaa | agtgcaaaaaa | 600 |
| aatgtgtaca | agagataatg | caaacaggat | tataaagaca | acattgcaac | aagaaaagcc | 660 |
| agatttctgc | tttttggaa | aagatctgg | aatatgtca | gggtatata | ccaggtattt | 720 |
| ttataacaat | cagacaaaaac | agtgtgaacg | tttcaagtat | gggtggatgcc | ttggcaat | 780 |
| gaacaattt | gagacactgg | aagaatgca | gaacatttg | gaagatggtc | cgaatggttt | 840 |
| ccaggtggat | aattatggaa | cccaagctca | tgctgtgaat | aactccctga | ctccgcatac | 900 |
| aaccaagggt | cccaaggctt | ttgttacaaa | agaaggaa | aatgtgggtt | ggaagaatgc | 960 |
| ggotcatatt | taccaagtct | ttctgaaacgc | cttctgcatt | catgcatacc | tgttcttct | 1020 |
| aggattggat | agcatttcat | gcctatgtt | atatttgtc | tttggcatt | tccttaat | 1080 |
| ttatatgtat | acgtgtgcc | tttgatagca | tactgcta | aaagtttaa | tatttacatg | 1140 |
| catagtaaaa | aaaaaaaaaa | aaaaaaa | | | | 1166 |

| | | | | | | |
|---------------|-------------------------|---------------|--------------|-------------|-------------|------|
| SEQ ID NO: 34 | moltype = DNA | length = 8449 | | | | |
| FEATURE | Location/Qualifiers | | | | | |
| source | 1..8449 | | | | | |
| | mol_type = other DNA | | | | | |
| | organism = Homo sapiens | | | | | |
| SEQUENCE: 34 | | | | | | |
| gccccgcgcg | gctgtgtc | acagggggag | gagaggaaac | cccaggcg | agcgggaaga | 60 |
| ggggacatgc | agcccaact | tctctggcc | tctgcattcc | ttctgtcct | ccaccgtcc | 120 |
| ccttccccac | cctctggccc | ccacccctt | ggaggcgaca | accccccgg | ggcattagaa | 180 |
| gggattttcc | ccgcagggtt | cgaagggaag | caaacttgg | ggcaacttgc | ctccgggtc | 240 |
| ggggcgttct | ccccccacgt | ctcaacatgc | tttaggggtcc | ggggccgggg | ctgtgtc | 300 |
| tggccgttca | ttggctgggg | acagcgttgc | cctccacggg | agcctcgaag | agcaagaggc | 360 |
| aggctcagca | aatgtttcag | cccccaatgc | cggtggctgt | cagtcaaa | aagccgggtt | 420 |
| gttatgacaa | ttggaaaaacac | tatcagataa | atcaacatgt | ggagcggacc | tacctaggca | 480 |
| atgcgttgg | ttgttactgtt | tatggaggaa | gcggagggtt | taactcgag | agtaaacctg | 540 |
| aagctgaaga | gacttgc | gacaatcaca | ctggaaacac | ttaccggatg | ggtgacactt | 600 |
| atgagcgtcc | taaagactcc | atgtatctgg | actgtatcttgc | catcggttgc | gggcgaggga | 660 |
| gaataagctg | taccatcgca | aaccgctgc | atgaaggggg | tcaatgttac | aagattgtgt | 720 |
| acacactggg | gagaccatgttgc | gagatgtttgc | tttacatgtt | agatgtgttgc | tgttcttgc | 780 |
| atggaaaagg | aaatggacc | tgcaacccca | tagctgagaa | gtgttttgc | catgtgtc | 840 |
| ggacttcctt | tgtgttgc | gaaacgtggg | agaacccctt | ccaaggctgg | atgtgttag | 900 |
| attgtacttg | cctggggagaa | ggcagggac | gcatcaatgc | cacttcttgc | aatagatgc | 960 |
| acgatcgagg | cacaaggaa | tcctatagaa | ttggagacac | ctggagcaag | aaggataatc | 1020 |
| gaggaaacct | gttccatgtc | atctcgacag | gcaacccgg | aggagatgg | aatgtgtgaga | 1080 |
| ggcacaccc | tgtcgacacc | acatcgagg | gatctggccc | tttcacccat | tttgcgtc | 1140 |
| ctgttttacca | accgcgcct | caccccgac | ctctcccttca | ttggccactgt | gtcacagaca | 1200 |
| gttgtgttgt | ctactctgt | gggtgttgc | ggctgttgc | acaaggaaat | aagcaaatgc | 1260 |
| tttgcacatgc | cttggggac | ggagtgcgt | gccaagagac | agtcgttacc | cgacttacg | 1320 |
| gttgtcaactc | aaatggagag | ccatgtgtt | taccatttac | ctacaatggc | aggacgttct | 1380 |
| actccctgcac | cacagaaggg | cgacaggac | gacatctttgc | gtgcagcaca | acttcgaatt | 1440 |
| atgagcgagg | ccagaaatcc | tctttctgc | caagacccatc | tgttttgc | cagactcgag | 1500 |
| gaggaaatcc | caatgttgc | tttgttgc | ttcccttc | tttgcacatc | tttttttttt | 1560 |
| ctgttgcacat | ttcttgc | agaagagaca | acatgttgc | tttgttgc | tttttttttt | 1620 |
| atgtatgcgc | ccagaatgtt | gggttctgc | ccatgttgc | tttttttttt | tttttttttt | 1680 |
| ccaaatgggg | gggtatgttgc | ccatgttgc | tttttttttt | tttttttttt | tttttttttt | 1740 |
| acatgtatgc | gttgcacatgc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 1800 |
| cccgatgttgc | agatgttgc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 1860 |
| acaagcgtca | tgaaggggg | cacatgttgc | actgtatgttgc | tttcgttgc | ggtcggggca | 1920 |
| gttgttgcgt | tgatccgc | gaccaatgc | aggattcaga | gactggggac | tttttttttt | 1980 |
| tttgttgcgt | tttgttgcgt | tttgttgcgt | tttgttgcgt | tttgttgcgt | tttgttgcgt | 2040 |
| gttgttgcgt | tttgttgcgt | tttgttgcgt | tttgttgcgt | tttgttgcgt | tttgttgcgt | 2100 |
| tcgaagtatt | tatcaatgc | acttcgatgc | agcccaatc | ccacccatc | cagtggat | 2160 |
| caccacagcc | atctccatc | tccaaatgc | ttcttcgggtt | gagacccat | tttttttttt | 2220 |
| ggcggttggaa | ggaatgttcc | ataccaggcc | actttaatcc | cttacaccat | aaaggccgtca | 2280 |
| agccctgggtt | gttataatgc | ggccatgttgc | tttttttttt | tttttttttt | tttttttttt | 2340 |
| tgactcgat | tgacttcacc | accaccatgc | ccagcacacc | tgttgcgttgc | tttttttttt | 2400 |
| caggagagac | gactccctt | tcttccttgc | ttggccacttc | tgaatctgt | tttttttttt | 2460 |
| cagcccaatgc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 2520 |
| tggatataatgc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 2580 |
| ccacttcgttgc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 2640 |
| agatataatgc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 2700 |
| atgcccctcc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 2760 |
| gcagacccca | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 2820 |
| gcagccacaga | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 2880 |
| cttgcgttgc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 2940 |
| tttgttgcgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 3000 |
| acctgcgttgc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 3060 |
| gtgcgttgc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 3120 |
| agaggctgc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 3180 |
| cttattacttgc | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 3240 |

-continued

| | | | | | | |
|-------------|--------------|-------------|--------------|-------------|-------------|------|
| aacagacaac | caaactggat | gtctccacta | acctccagtt | tgtcaatgaa | actgattcta | 3300 |
| ctgtccctgt | gagatggact | ccacccctggg | cccagataac | aggataccga | ctgaccgtgg | 3360 |
| gccttaccgg | aaggagacag | cccaggcagt | acaatgtggg | tccctctgtc | tccaaagtacc | 3420 |
| cactgaggaa | tctcagccct | gcatctgagt | acaccgtatc | cctcgtagggc | ataaaagggca | 3480 |
| accaagagag | ccccaaagcc | actggagttct | ttaccacact | gcagcctggg | agctctattc | 3540 |
| cacccatccaa | caccgagggt | actgagacca | ccattgtgat | catatggacg | cctgtctcaa | 3600 |
| gaattgggtt | taaagctgggt | gtacgacca | gccaaggggagg | agagcacca | cgagaagtga | 3660 |
| cttcagactc | aggaacatc | gttgcgtccg | gttgcgtcc | aggagtagaa | tacgtctaca | 3720 |
| ccatccaagt | octgagagat | ggacaggaaa | gagatgcgc | aattgtaaac | aaagtggta | 3780 |
| caccatgttc | tccaaaca | aactgtcattc | tggaggcaaa | ccctgcacat | ggagtgcctc | 3840 |
| cagtctctg | ggagaggagc | accaccccg | acattactgg | ttatagaatt | accacaaccc | 3900 |
| ctacaaaacgg | ccagcaggga | aatttttg | agaagttgtt | ccatgtgtat | cagagctccct | 3960 |
| gcacttttga | taaacctgagt | cccgccctgg | agtacaatgt | cagtgttac | actgtcaagg | 4020 |
| atgacaagga | aagtgtccct | atctgtata | ccatcatccc | agctgttct | cctcccaactg | 4080 |
| acctgcgatt | caccaactt | ggtccagaca | ccatgcgtt | cacccgggt | ccacccccc | 4140 |
| ccattgtatt | accaacttc | ctgtgcgtt | acttcacccgt | aaaaaatgag | gaagatgtt | 4200 |
| cagagttgtc | aattttctct | tcagacaatg | cagtggctt | aaaaaatctc | ctgcctggta | 4260 |
| cagaatatgt | agttagtgc | tccagtgtt | acgaacaaaca | tgagagcaca | cctcttagag | 4320 |
| gaagacagaa | aacagggttt | gatttttttt | ctggcatttg | ctttctgtat | attactgc | 4380 |
| actcttttac | tgtgcactgg | attgtccctc | gagocccat | cactgcgtac | aggatccgccc | 4440 |
| atcatcccg | gcacttcagt | gggagaccc | gagaagatcg | ggtgcggcc | tctcgaaatt | 4500 |
| ccatcaccc | caccaaccc | actccaggca | cagatgtatg | ggtcagcata | gttgccttta | 4560 |
| atggcagaga | ggaaaatccc | tttattgttt | gccaacaaatc | aacgttttct | gatgttc | 4620 |
| ggggacccgg | agttgttgt | gcggccccc | ccagctact | gatgcgtt | gtgcgtct | 4680 |
| ctgtcacagt | gagatattac | aggatcatt | acggagagac | aggagggaaat | agccctgtcc | 4740 |
| aggaggttac | tgtgcctggg | agcaagtca | caagtaccat | cagcggcc | aaacctggag | 4800 |
| ttgattatac | catactgt | tatgtgtca | ctggccgtgg | agacagoccc | gcaagcagca | 4860 |
| agccaaatttc | cattaattac | cgacacaaa | ttgacaaacc | atcccaatgt | caagtgcacc | 4920 |
| atgttcagga | caacagcatt | agtgtcaat | ggctgccttc | aagtccct | gttactgtt | 4980 |
| acagagtaac | caccatccc | aaaaatggac | caggacccaa | aaaaactaaa | actgcagg | 5040 |
| caagatcaac | aaaaatgact | atttaggtt | tgccaggat | tgccaggat | tggttagt | 5100 |
| tctatgtct | gaatccaa | ggagagatc | agcctctgtt | tcaactgca | gtaaaccaaca | 5160 |
| ttgatcgccc | taaaggactg | gcattcact | atgtggatg | cgattccatc | aaaattgttt | 5220 |
| gggaaagccc | acagggggca | gtttccagg | acagggttgc | ctactcgac | cctgaggat | 5280 |
| gaatccatga | gtatccccc | gcacatgtt | gttgaagaag | cactgc | ctgaagg | 5340 |
| tcagaccggg | ttctgtgt | acagtgttgc | ttgttgcctt | gcacatgtat | atggagagcc | 5400 |
| agccccctgt | ttggacccat | tcacagctt | ttctgttgc | ttctgttgc | ttctgttgc | 5460 |
| aggtcacacc | cacaagctg | agcgtccat | ggacaccacc | caatgttca | ctcaactggat | 5520 |
| atcgagtgc | ggtagcccc | aaggagaaga | ccggaccaat | gaaagaaatc | aaccctgtc | 5580 |
| ctgacagctc | atccgtgtt | gtatccagg | ttatgggtgc | ccaccaat | gttgcgttcc | 5640 |
| tctatgtct | taaggacact | tttgcacaa | gaccgttca | gggagggttgc | accactctgg | 5700 |
| agaatgttag | cccaccaaga | agggtctgt | tgacagatgc | tactgagacc | accatcacca | 5760 |
| ttagctggag | aaccaactg | gagacatca | ctggcttcca | agttgtatgc | gttccaggcca | 5820 |
| atggccagac | tccaaatcc | agaaccatca | agccagatgt | cagaacgttac | accatcacag | 5880 |
| gtttacaa | aggactgtac | tacaatgtt | acccgttac | tttgcacat | aatgtcgga | 5940 |
| gtctccctgt | ggtcatcgac | gcctccat | ccatgtatc | accatccaa | ctgcgttcc | 6000 |
| tggccaccac | acccaaatcc | tttgcgttgc | catggcag | gccacgttgc | aggattacc | 6060 |
| getacatcat | caatgtatg | aaggcttgc | ttcttccat | agaatgttgc | cctcgcccc | 6120 |
| ggccctgggt | cacagaggct | attactgt | ggctggaa | ggggacccaa | tatacaatt | 6180 |
| atgttcattgc | cctgttgc | aatcagaaga | gcgcggcc | gattggaa | aaaaagacag | 6240 |
| acgagatcc | ccaaactgtt | acccttcc | accccaat | tcatggacca | gagattctgg | 6300 |
| atgttccttc | cacagtctt | aaggccctt | tcgttacca | ccctgggtat | gacactggaa | 6360 |
| atggatattca | gttctgttgc | actttgttgc | agcaacccat | tttttgc | tttttgc | 6420 |
| tttggagaaca | ttgtttttag | cggttccat | cgccacac | ggccaccc | ataaggcata | 6480 |
| ggccaaagacc | atacccccc | aatgttaggt | aaaggatgtt | ctctca | accatctcat | 6540 |
| ggggccccc | ccaggacact | tacatgttgc | tttgcacat | tttgcacat | tttgcacat | 6600 |
| aagaacccctt | acagttcagg | gttctgttgc | tttgcacat | tttgcacat | tttgcacat | 6660 |
| ccagagggtc | cacccataac | atcatgttgc | aggcactgaa | agaccagcag | aggcataagg | 6720 |
| ttcgaaaga | gttgcgttac | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 6780 |
| atgactctgt | cttgcaccc | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 6840 |
| gaatgttgc | atcaggctt | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 6900 |
| tcagatgttgc | tttgcacat | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 6960 |
| agtgggaccg | tttgcacat | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 7020 |
| aaggagaaat | caatgttgc | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 7080 |
| acgttagggc | acatgttgc | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 7140 |
| gaggccacgc | gggttgcgc | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 7200 |
| aaggcactac | tggccatgttgc | tacaaccgt | tttgcacat | tttgcacat | tttgcacat | 7260 |
| ctaatgttta | tttgcacat | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 7320 |
| attccccgaga | gtaaatcatc | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 7380 |
| ccatctaaac | tttgcacat | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 7440 |
| cctttgttct | ggggggat | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 7500 |
| ccttggggat | tttgcacat | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 7560 |
| gaagtattca | ataccgttca | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 7620 |
| aataggaaag | catatgttgc | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 7680 |
| tttaagttag | aaagtccacc | aaacttct | tttgcacat | tttgcacat | tttgcacat | 7740 |
| tgttagaaaca | agcatgtatc | tttgcacat | tttgcacat | tttgcacat | tttgcacat | 7800 |

-continued

| | | | | | | |
|------------|-------------|-------------|-------------|-------------|------------|------|
| tccaaatgt | cctagaatt | gcctagaaaat | atctttctct | tacctgttat | ttatcaattt | 7860 |
| ttcccagtat | ttttatacgg | aaaaaattgt | attgaaaaaca | cttagtatgc | agtgtataag | 7920 |
| aggaatttgg | tataattatg | gtgggtgatt | atttttata | ctgtatgtgc | caaagcttta | 7980 |
| ctactgtgga | aagacaactg | tttaataaaa | agatttacat | tccacaacct | gaaggtcac | 8040 |
| tatttgatat | aagacacctt | cggggaaat | aatttcctgt | aatattctt | ttcaattcag | 8100 |
| caaacattt | aaaatctatg | atgtgcaagt | ctaattgtt | atttcagtagc | aagattttct | 8160 |
| aaatcagttg | ctacaaaaac | tgatgggtt | ttgtcactt | atcttcac | taatggagat | 8220 |
| agctttcac | tttctgttt | aatagat | ttgtggacccc | aatatttt | aaaattgtct | 8280 |
| gtttaccgtt | cagaagtata | atagaataaa | tcttttagtt | ctctttctt | accattgtaa | 8340 |
| ttttccctt | tttccctcca | ccttccctt | attgaataaa | cctctgttca | aagagatgtc | 8400 |
| ctgcaaggga | aataaaaaatg | actaagat | taaaaaaaaaa | aaaaaaaaaa | | 8449 |

SEQ ID NO: 35 moltype = DNA length = 4625
 FEATURE Location/Qualifiers
 source 1..4625
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 35

| | | | | | | |
|-------------|-------------|------------|-------------|-------------|-------------|------|
| agcgctccgc | agtacacgtga | cgctcgccg | caacctctgc | tgtccctccgc | ggcgccccct | 60 |
| tccgcctgac | gcccccccg | cgcgccgcg | gcagccctgg | ctccctcgccg | gctcgccgg | 120 |
| cggctgeggc | ggggctatgg | cgagggcgg | tgggggggtt | aacactggcg | cgggtgggg | 180 |
| gcggggatg | ggccctggcc | tggggctggg | tctgggttgc | agcctagggca | tgagtggggc | 240 |
| caccagttag | cgacaggagg | aggcgccac | ggccgaggggc | tgggacggcc | ttggccacggc | 300 |
| gtctgtggct | cggtcccg | gtggggggc | gggtctggcc | ggggcgccgc | gggtgtgg | 360 |
| gtggggagaag | ccgctgcaca | gcctggc | ggggccgcgc | ctcaacggcc | tctttgtt | 420 |
| gtctgttcc | ttgtccctcc | ggccctt | cctactcagc | gtctcaactt | ttggcttattt | 480 |
| tctgtgttgc | ttctggc | ctcg | ccctgacgtt | tcagcat | ccccagaggaa | 540 |
| gccacactt | gacagtgggg | gtgggggtt | aggcgcccg | ccgcacccgc | tgagtgtgcc | 600 |
| cgagttgtgc | agataactgg | ctggagat | gtccaccc | cagatttacc | tgcaaggagct | 660 |
| gtctgcgtac | aaggaggcaga | atccagctca | gttctgcgtt | cgagttgtct | ctggctgtgc | 720 |
| tgtgttggct | ttgtggggc | actatgttt | agggtttat | tttcttata | ttgttttgtt | 780 |
| gagtatctt | ctgtggggcc | tgggttta | tcatgagtt | atccagaggaa | tgtacactcg | 840 |
| cctggagcc | ctgctcatgc | agctggacta | cagcatga | gcagaagcca | atgcctgc | 900 |
| tcacaaaac | gacaagagga | agcgtcagg | gaagaatgc | cccccaggag | tgatgtggcc | 960 |
| actggcagag | acagagatgt | aaaggagg | agagctggct | ggcttctccc | cagtgtgg | 1020 |
| tgtgaagaaa | acagcattgg | ccttggccat | taagactca | gagctgtcag | atgaggaggc | 1080 |
| tcttatctt | gagagttgt | gtcttccgt | atccggggc | acaactccgc | agctgactt | 1140 |
| tgtctccag | gattttggacc | agcagagct | gccaagtga | ccagaggaga | ccctaaggcg | 1200 |
| ggacactagg | gggggggg | aggggagat | ggccctccc | gaagaccc | taggccttcc | 1260 |
| tcaagctctg | tcaggaa | ccctggact | ggaggaa | gaaggaggat | tggcagctaa | 1320 |
| ggaacaccc | tttgcgtct | catccccct | ccatcttgc | aacacgcact | tcaatgggc | 1380 |
| agggtcccc | ccagatggag | tgaaaatgtc | ccctggagg | ccagtgaggaa | cactgagccc | 1440 |
| cggagactgt | atgtggggc | tcactgtt | gccccggc | ctgtcaccc | cacttgc | 1500 |
| tgttggaaat | gccccggcc | cetccctt | catttccca | ctgttccc | aggacttacc | 1560 |
| ccagccctgt | ctggccctt | aggagaaga | ggactcacc | actgaggact | ttgatgtgt | 1620 |
| ggatcagggg | gagctggagc | agctgaatgc | agactgggc | ttggagccag | agacacggcc | 1680 |
| aaacacccct | gtgtccac | ccctggggcc | cgacatccat | tcttgcgtac | agttagacca | 1740 |
| agaagctcg | gccccggcag | agccatgac | cagccgttga | ggaaggagct | gcaggcacag | 1800 |
| tagggctcc | ttgttgggg | tgttgcgtt | tctccctt | cattacc | ttgggtgggg | 1860 |
| cagtgtgtgg | ggaagctggc | tgttggatgg | tagtattcc | accctcttgc | tgccctgtc | 1920 |
| cctgtcttcc | ttggggatgtt | cgacttcc | tgcttgcgtt | tgtttttaa | tttgcataa | 1980 |
| attttccatt | ttgggttgcgt | gtatgttgc | gggttgcgtt | tttgcataa | tttgcataa | 2040 |
| cttgccttcc | ttgggttgcgt | ctatccat | tccactatgc | cccaagcc | tttgcataa | 2100 |
| ccctttttt | ttccctctat | cctcagg | ctgtgcgt | ctgcctcat | tttgcataa | 2160 |
| gttggatgtt | ttggggactt | tataatttt | cttgcgtt | tttgcataa | tttgcataa | 2220 |
| tcctgtgtt | gttgcgtt | tcgttgcgt | acccatcc | tttgcataa | tttgcataa | 2280 |
| gttggactgt | gccaaggctt | aggaggctc | ctgtgcgtt | tttgcataa | tttgcataa | 2340 |
| gcagtggttgc | aggcaggat | tttgcataa | tttgcataa | tttgcataa | tttgcataa | 2400 |
| cattcgttcc | ccatccagcc | ccggcc | tgatgcagg | gagcagggt | cccgccat | 2460 |
| atggcgecag | cactggat | ggtgatgc | tgttgcgtt | tttgcataa | tttgcataa | 2520 |
| ctgtcttcc | gggtgtatca | atgttgcgtt | ggccccc | tttgcataa | tttgcataa | 2580 |
| taacatctgt | tcagggttgc | acttgcataa | tttgcataa | tttgcataa | tttgcataa | 2640 |
| ccctgactgt | gagctgcctt | ccacccctg | cactgcact | tttgcataa | tttgcataa | 2700 |
| cctgtcttcc | ttgggttgcgt | tattttttt | tttgcataa | tttgcataa | tttgcataa | 2760 |
| aaggatttacc | tttgcataa | tttgcataa | tttgcataa | tttgcataa | tttgcataa | 2820 |
| ttatgtgtat | tttgcataa | tttgcataa | tttgcataa | tttgcataa | tttgcataa | 2880 |
| aatcggtaga | aattcgggac | cataaaaaat | tttgcataa | tttgcataa | tttgcataa | 2940 |
| actgtttctt | taagaagtct | tcacccat | acatgtca | aactca | tttgcataa | 3000 |
| atcttactg | gggaagccaa | acaacgata | tttgcataa | tttgcataa | tttgcataa | 3060 |
| ggagccaa | aactctgt | tttgcataa | tttgcataa | tttgcataa | tttgcataa | 3120 |
| tggctctgt | ctggatgtt | tataatgtt | tttgcataa | tttgcataa | tttgcataa | 3180 |
| tgttactaac | ttcattttat | tttgcataa | tttgcataa | tttgcataa | tttgcataa | 3240 |
| tttgcataa | tttgcataa | tttgcataa | tttgcataa | tttgcataa | tttgcataa | 3300 |
| tttgcataa | tttgcataa | tttgcataa | tttgcataa | tttgcataa | tttgcataa | 3360 |
| aatctgtatc | tttgcataa | tttgcataa | tttgcataa | tttgcataa | tttgcataa | 3420 |
| aagcccgaa | ctgacat | tttgcataa | tttgcataa | tttgcataa | tttgcataa | 3480 |

-continued

| | | | | | | |
|--------------|--------------|-------------|--------------|-------------|-------------|------|
| ggtagccatcat | ggcgatgcata | cctgttotttc | tgcgtggaggt | accatatggt | aatgtctgcct | 3540 |
| ggctgtctgc | tggaggtaacc | atatggtaat | gctgcctgtc | ttctggaggt | tgacttttat | 3600 |
| gcccattgtct | tcctaaagtgt | gtaaaattt | ttctgttgc | ttcacatgg | actggagaata | 3660 |
| atcttaggtt | ttgtatggac | ccctgtctct | tgcactaaa | ggaactcgaa | cttttcatca | 3720 |
| cttagagatt | tcagaggggaa | atggaaaaac | agttctaatac | aataaagcaag | caattcaaga | 3780 |
| aaaatagaat | taatcaggca | atgactgcaa | catgtctat | cttaatcta | ttttttttat | 3840 |
| aagcttggaa | attgacaataa | gaaccggaaag | cttgcgtctgt | gatccaaaata | tttccccatag | 3900 |
| ggctggagggt | tcatggagggt | ctatctttt | ttgttgggtt | ttttgttttt | tttttttttt | 3960 |
| tgggtttttt | tttttttttt | tttggagacgg | agttttgttc | tttttttttt | tttttttttt | 4020 |
| aatgttgcac | tcttggttca | ctgcaactcte | tgcccttccag | gttcaaaacaa | tttttttttt | 4080 |
| tcagcgctc | aagttagctgg | gattacaggt | gtatggccacg | atgccttgcgt | attttttttt | 4140 |
| tttttttttt | agggtgggggtt | tcacccatgt | ggccaggctg | gttccaaact | cttgcacttca | 4200 |
| ggtgattcac | ccacccctggc | ctcccaaaagt | gctgggattt | cagggtgttag | ccacccggcc | 4260 |
| cagccctatcg | agggtttatc | ctttatccat | accatgggtt | gttgggttgt | acatgtttgt | 4320 |
| ctatgatttt | ttttttctat | tatcggtgt | tttttttttt | tcatggccca | cgatggaaagg | 4380 |
| ccggccgggtt | tttccatgtat | taaaaagaaa | aaaggccaaaat | tttttttttt | caatttttttt | 4440 |
| aaacaaaaat | caactggaaa | agtttttttt | tttttttttt | tttttttttt | tttttttttt | 4500 |
| tttaggggggg | aggggagtaca | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 4560 |
| tgctttataaa | cttaattttt | ggctgcattta | ataaaagttt | tttttttttt | tttttttttt | 4620 |
| aaaaaa | | | | | | 4625 |

-continued

| | | | | | | |
|--------------|-------------|--------------|-------------|-------------|-------------|------|
| cttgggtatg | gtccctgctg | tgcagagaag | cgcacccgtc | ttaccagtgt | ccgaagactt | 3000 |
| taaatcgag | ctgtttcac | ctcaggatt | ttctttctcc | aagaatggc | tgctaagtgc | 3060 |
| attgctaga | caaaaatcaag | atagttacct | ggcagatgt | tcagacaggaa | gtcacagaaa | 3120 |
| taatgaaatg | gaoacttctag | aatcaaagaa | tctttgcatt | gtccctaaag | aaaggaaatc | 3180 |
| ttatacttag | ccattagaaa | atccattaa | aaagatgaaa | aacaacatgt | ttgatgtgc | 3240 |
| aaacaatcac | agtgcggccag | aagtactgt | tgggtccctt | cttaaccagg | aagagctgaa | 3300 |
| atttagcaga | aatgatcttg | aattttaaata | tcttgcgtgt | catggctcag | ccagcggaaag | 3360 |
| tgaacacagg | agttggccca | gagagagcaa | aagctttat | gttctaaac | agctgtctt | 3420 |
| ctcagaaaaac | tgtgtgcag | atttgtcccc | gcacagaagg | aactctgtgg | ctgacagat | 3480 |
| aaagaaaaagg | cacaaaataa | atgtgaccaa | cagcaaaactt | gaatttagca | tttcttctt | 3540 |
| aatatggactg | atgtgactgt | ccactcagcc | cacagttgc | atggaaaca | ggacatttc | 3600 |
| ataccggatg | gtatgaaaaa | ctctctgttag | tcctactttc | cctgagact | tgggtgtgc | 3660 |
| aggggtctaga | ccagaatctg | ggcttttga | tgggtgtcc | atgcccagtg | agaaaaggacc | 3720 |
| catthaatgg | gttatctact | atgcggagaa | gaatgagtt | aaaaaaagact | ctccaagatt | 3780 |
| gaccaaaaacc | aacccaaatac | tatattatcat | gcttcacaaa | ggaggcaatt | ctgttacac | 3840 |
| tcgagaaaaac | caagacaagg | acatttggag | ggaggctca | tctgtgtaaa | gtgtctcaca | 3900 |
| ggtcacagcc | aaagaagagt | tacttctac | tgcagaaac | aaagcttctt | tctttat | 3960 |
| aaagaagccct | tacaatagcc | atatggaaa | taatgtttct | cgcccaacaca | gcfgaaatgg | 4020 |
| agaagtttat | ggacttctgg | gaagctgtgt | aacgataaaag | aaagaatcag | aataaaaatgt | 4080 |
| acctgcccatt | cagtttgg | tctttttaa | actaatgat | atgatgttg | gatctgtata | 4140 |
| aataagagca | tgatgtggaa | aaaagcatgg | tataattgaa | actttttca | ttttgaaaag | 4200 |
| tattggttac | tgggtatgtt | gaaatatgca | taactat | tgcttaacat | tagatgtcat | 4260 |
| gaggaaacta | ctgaaactgc | aattttgggt | ttaacacttc | tgtatgcatt | agataacaaac | 4320 |
| tgtgagtagc | ctatgatgt | aattttttta | taataat | gtatgttttt | aaatgtaaaa | 4380 |
| ctccattcat | agtggattaa | tgcat | tgc | tgc | agggtactt | 4440 |
| tcagaagtc | gctcatacaa | cacattttt | aagtctaaac | tgttacaaa | ctctttaaag | 4500 |
| gataattatc | caataaaaaaa | aaacccatgt | ctgattcaca | gotttattatc | caatccaaa | 4560 |
| ataaattatg | aaatataatg | cttacat | tcacttttgc | taaaaaagaaa | aaaaaaaggt | 4620 |
| gtttat | tttttttttt | aactcttgg | aggggtttt | tggttccca | tgtgtctgtc | 4680 |
| ccttttcaat | atatat | ttttaaccttgc | tgctacttgc | taaaaatgt | ttacaatgt | 4740 |
| ggaaagtttg | atagatctt | tttttttttt | ggcatttttt | ttttttgt | ttttactac | 4800 |
| tttacttaat | taatacttct | ctttttacag | tttttttttt | ttttttgt | ttttactac | 4860 |
| ggtttctgaa | aaagttgaat | atthaagaaa | tttttttttt | cagaagcaaa | atggctt | 4920 |
| tttggacagt | tttcaccatc | tcttgc | tttttttttt | tttttttttt | tttttttttt | 4980 |
| gagtgtttag | accaggattc | cttacccat | aaactcagacc | acttgcattt | tttttttttt | 5040 |
| gagcacttgt | ttttttatgc | ttttttatgc | tttttttttt | tttttttttt | tttttttttt | 5100 |
| tctgctaaat | ggacccggaa | tcttccat | tttttttttt | tttttttttt | tttttttttt | 5160 |
| gtgtttgaga | accatgactt | atgacttcc | tcagaacat | tttttttttt | tttttttttt | 5220 |
| aaaactatgt | gtgttttca | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 5280 |
| tattgttagt | catatattac | cgttatgtt | tttttttttt | tttttttttt | tttttttttt | 5340 |
| acaaaaattt | gttggagggg | ataacttca | tatctggc | tttttttttt | tttttttttt | 5400 |
| gattatttgt | gtgttctaat | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 5460 |
| ggcccaatctt | ccaaactcag | gtgttcc | tttttttttt | tttttttttt | tttttttttt | 5520 |
| cttaacccat | ataaaaacca | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 5580 |
| tttaaacccca | agaaggccac | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 5640 |
| atacgtgaca | gaattgtaca | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 5700 |
| aaagtatcat | ttttttatgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 5760 |
| tcagttggaa | aacaggatctg | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 5820 |
| gatattttgt | tcttaattgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 5880 |
| atcactacta | gttagat | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 5940 |
| ttatataat | aaatgtgtt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6000 |
| ctcagaatgt | ttttttatgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6060 |
| gcaacatcat | ggagaatgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6120 |
| tgaatataat | ttttttatgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6180 |
| agctttttaa | caactgttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6240 |
| tcagaaaaaa | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6300 |
| gtgaaat | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6360 |
| gtgggtttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6420 |
| gagaatatgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6480 |
| aatgttccca | attttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6540 |
| caagtttact | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6600 |
| actattttag | atatcttagac | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6660 |
| tttatttgcata | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6720 |
| agagtttact | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6780 |
| tatataatata | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6840 |
| taataatgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6900 |
| gccttttaat | ctctttgg | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 6960 |
| gtttaaaatgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 7020 |
| cttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 7080 |
| gtgcattttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 7140 |
| tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 7200 |
| agggtcttaa | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 7260 |
| tactttcatg | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 7320 |
| caataactgga | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 7380 |
| aattttatgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 7440 |
| catgtatgt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | tttttttttt | 7500 |

-continued

| | |
|--|------------------------------------|
| cttttagtgt ttacttcctg cagtgaattt gaataaatga gaaaaaatgc attgtc | 7556 |
| SEQ ID NO: 37 | moltype = DNA length = 1516 |
| FEATURE | Location/Qualifiers |
| source | 1..1516 |
| | mol_type = other DNA |
| | organism = Homo sapiens |
| SEQUENCE: 37 | |
| tgc(ccc)acca ccg(cgt)ctcc tcagcaggcg cctcaccaggc ctccacaccc cttgcggccc 60 | |
| caga(aac)cgcg cctggccctg agcgtcacc accgacactc tccaggotcc ggacacgatg 120 | |
| cagg(c)catc aagtgtgtgt ggtggagat gggccgtgg gcaagacgtg ctttcatc 180 | |
| agctacacca ccaacgcctt tccggagag tatcatcccc cctgttta caactatcca 240 | |
| gccaatgtga tgg(tgg)acag caagcagtg aacctggggc tggggacac tgctggcag 300 | |
| gaggactacg accgtctcg gccgccttcc tatccacaga cggacgtctt cctcatctc 360 | |
| tttccctcg tcagccccagc ctcttatgg aacgtcccg ccaagtggg cccagaatgt 420 | |
| cggcaccact gcccacgac acatcatc ctgtggggca ccaagctgga cttgcgggac 480 | |
| gacaaggaca ccategagaa actgaaggag aagaagctgg ctcccatcac ctaccgcag 540 | |
| ggccctggcac tggccaaaggg gattgactcg gtgaaatacc tggagtgc agcttcacc 600 | |
| cagagggccc tggaaaacggt gttcgcacgg gccatccggg cctgtgtgtg ccctcagccc 660 | |
| acgcggcagc agaagcgcgc ctgcgcctc ctctagggtt tgccacccag cgctccacc 720 | |
| tagatgggtc tgatcctcca ggatccccac ccaaagcctg atggcacccc ggctggccat 780 | |
| gtgtgtccct ccctgtggcg ttcttagca gatggctgca gagcttcgtt gatggcttt 840 | |
| tctgtactgg aggccctctg aggccaggaa cgtgcaattt tgccagggtt gcatcccaag 900 | |
| cccccctatgc tcctgcctt ctggggca gaggggagcc ccaggaccca ttaagccacc 960 | |
| ccctgttcc tggcgtcaat gccaactgtc gcatgtggaa gcatctaccc ttactctca 1020 | |
| gtccccacccc accgcgtact cccctcttga aactgcggc cagatgtgtg ctgcaccaac 1080 | |
| tttgtactcttccatggatgg ggcttctact ccctctcgag gccagctgtc ctaatatcga 1140 | |
| tggtcctgtc tgccagagag ttcttctacc cagcaaaaat gagggttca gaagtgtgt 1200 | |
| cctctggcct cagtttccctt cttttggaa acataaaaac aaattttatt ttctacgcct 1260 | |
| ctggggatata ctgctcagcc aatggaaaat ctgggttca ccagccctg ccatttctta 1320 | |
| agactttctg ctgcacttca aggttctgtt gctgcacttca cctgtgtggag tcttcaact 1380 | |
| ttaaacccctt gcacgttcagg acttttgcata tttgcaatag aaaacccaaac tcaacctgt 1440 | |
| taagcagaaa ataaattttat tgattcaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1500 | |
| aaaaaaaaaaa aaaaaaa 1516 | |
| SEQ ID NO: 38 | moltype = DNA length = 1641 |
| FEATURE | Location/Qualifiers |
| source | 1..1641 |
| | mol_type = other DNA |
| | organism = Homo sapiens |
| SEQUENCE: 38 | |
| ctccctgtgt tgg(tgg)agga tggctgcagc agcatttaaa ttctggggagg gcttgggtgt 60 | |
| cacgcacgcg aggaggaggc agagcacacgc atcgtcgccga ccagactcgt ctcaaggccag 120 | |
| ttgcagecctt cttagccaaa cggccgatcc gggaaaactca ctaccatgag aattgcgtg 180 | |
| atttgcctt gcttccttagg cataccctgt gccataccag ttaaacaggc tgattcttga 240 | |
| agttctgagg aaaagcagct ttacaacaaa tacccagatg ctgtggccac atggctaaac 300 | |
| cctgaccatc tcacaaaggc gaaatccca gccccacaga atgctgtgtc ctctgaagaa 360 | |
| accatagact taaaacaaaga gacccttca agtaagtcga acgaaagcca tgaccacatg 420 | |
| gatgtatgg atgtatgaa tgatgtatc catgtggaca gccaggactc cattgcactg 480 | |
| aacgactctg atgtatgta tgacactgtat gattctcacc agtctgtatg gtctcaccat 540 | |
| tctgtatgtt ctgtatgtt ggtctactt tttccacgg acctgcccgg aaccgaagt 600 | |
| tttactccat ttttcccccac agtagacaca tatgtatggc gaggtgtatg tttgtttat 660 | |
| ggactgtgggtt caaaatctaa gaatgttccg agacactgaca tccagttacc tgatgttaca 720 | |
| gacgaggaca tcacccatca catggaaago gaggaggatgaa atgggtcata caaggccatc 780 | |
| cccggttgcggc aggacccgtt cggcccttctt gattggggca gccgtggggaa ggacagtat 840 | |
| gaaacacggc acgtgtgtt ccagatgtt gaaacccaca gccacaagca gtccagatta 900 | |
| tataaggccgaa aagccaatgt tgagagacat gggccatccatc atgtgtatg tttgtttat 960 | |
| ctttccaaatg tcacccgttca attccacacgc catgaatttc acagccatgt agatatgt 1020 | |
| gttgtatggcc caaaatgtt ggaagaatgtt aaacacccgtt aatgtgtatg ttctcatgaa 1080 | |
| tttagatgtt catcttctgtt ggtcaatattt aaggagaaaa aatacaatattt ctcaatttgc 1140 | |
| atttagtcaaa aaaaaaaaaat gctttatgtt aaaaatgtt aagacatgtt atgcttctt 1200 | |
| ctcagtttat tgggttataat tggatctttt tggttcttca aataactatatgtt gttttgtata 1260 | |
| attagtttag ttgtgtggctt catggaaact cccctgttacac taaaagcttcc agggttatgt 1320 | |
| ctatgttcat tctatgtatgg aatgttcaatc tttttatattt ttttttttttgc 1380 | |
| tcatgtatgtt aatattttatgtt aagacatgtt aatgttcaatc aaaaatgtt aatgttcaatc 1440 | |
| ataacatattt atgttcaatc aatctttttt tttttatgtt aatgttcaatc aatgttcaatc 1500 | |
| tatcttttttgggtgtgtt aatgttcaatc tttttatgtt aatgttcaatc aatgttcaatc 1560 | |
| aattgtttat ttttttttttgc aatgttcaatc aatgttcaatc aatgttcaatc aatgttcaatc 1620 | |
| gcctaaaaaaa aaaaaaaaaa a 1641 | |
| SEQ ID NO: 39 | moltype = DNA length = 6463 |
| FEATURE | Location/Qualifiers |
| source | 1..6463 |
| | mol_type = other DNA |
| | organism = Homo sapiens |
| SEQUENCE: 39 | |

-continued

| | | | | | | |
|--------------|-------------|------------|-------------|-------------|----------------|------|
| ctctcttgc | cgtcgctcc | ctctcttc | tgctggctgc | ctgttctagg | aagccagcgc | 60 |
| ggagaggggg | gggatgcaca | gcacagggg | gagagattgc | gcatgttgtt | cagtctgttt | 120 |
| ttaaaagaga | cagtgcgggg | aggctgagag | gggcgcatgc | aacaacaact | tttggaaagga | 180 |
| tggaaagagaa | gaggcgaaaa | tactccatca | gcagtgcacaa | ctctgacacc | actgacagtc | 240 |
| atgcgcacatc | tacatccgca | tcaagatgt | ccaaactgc | cagcagcacc | aagtccggct | 300 |
| ggccccgaca | gaaacgaaaag | aageccctcg | agggtttcccg | gacagacttg | atcacagcca | 360 |
| tgaagatccc | ggactcatac | cagetcagcc | cggtgacta | ctacatctg | gcagaccat | 420 |
| ggcgacagga | atggggagaa | gggtgcagg | tgctctgggg | ggcagaggcc | atcccgagac | 480 |
| ccgtggtag | gatccctcca | ccactggaa | gccccctgc | ccaggeatcc | ccgagcagca | 540 |
| catgtcttgc | tgaggggctcc | cagectgtt | ggccagggggg | cagccgtat | gacttggacg | 600 |
| agattgtatc | ctactggctg | gagetcatca | aetccggat | taaggagatg | gagaggccgg | 660 |
| agctggacga | gtgtgacatta | gagctgtgc | tggaggagct | ggagacccctg | tgccaccaga | 720 |
| atatggccag | ggccattgag | acgcaggagg | ggctgggc | cgagatcgc | gaggatgtt | 780 |
| tctgcgacgt | gtgtcgctc | cctgagggg | aggatggca | cgagatgtgc | ttctgtgaca | 840 |
| agtgcacatc | ctgtgtgc | cgatgcgt | acgggatcc | caaggatgc | acgggcagct | 900 |
| ggctgtgcgc | gacgtgtgc | ctgggtgtcc | agccaaatgt | cctgcttc | cccaacgcqag | 960 |
| gaggagccctt | gaaagccact | agaagtggga | ccaaatgggtt | gcatgtc | tgtgccttat | 1020 |
| ggattccctga | ggtcagcata | ggctggccca | agaagatgg | gcccattacc | aagatctcgc | 1080 |
| atatccctcage | cageccatgg | gctctgttct | gcacatctg | caaggatgc | acaggcaacct | 1140 |
| gcatccatgt | ttccatgtcc | tcctcg | ccaggttcca | tgtcacatgc | gccttgcacc | 1200 |
| acggccttgg | aatcgggact | atattagcag | acaacatgt | ggtcaagt | tc aagtcatct | 1260 |
| gccaggagcga | cagtgcacgg | ggccacacgt | atgagccac | atctgacccc | acggaaaccca | 1320 |
| gcgcaggctgg | cgaggacgg | gaaaagggtt | ccctg | gcagcggctg | cagcagctag | 1380 |
| aggaggactt | ctacagactg | tggtggccgg | ctgagggtgc | tgagcggctg | gacccgtctg | 1440 |
| aggcactgtt | cgacttcata | taccagta | ggaagctgaa | gaggaaagcc | aatgcacacc | 1500 |
| agccgcgtct | gacccccaag | acccgacggg | tggacaaccc | ggcccaacg | gaggcaggacg | 1560 |
| tcctcttccac | cgcgttgcag | ctcttcac | atctgcggca | ggacat | tttgcggag | 1620 |
| atctgtgtca | catgtgtaca | atggcgcaga | gaaacgaaa | ccgcatttc | aaactccagg | 1680 |
| agcagatatt | ccacactgcag | atgaaactt | ttgaacagga | tctgtgtca | ggcctgtcca | 1740 |
| cctcatttcc | catcgatggc | accccttca | acagctgtt | ggcacatgc | gtgcagatca | 1800 |
| cagcagcatc | cattggccat | acgcgttgc | cactgaacaa | ttggcacccg | gaggaccctg | 1860 |
| ctccagggtt | gtgtgcag | gaactgtgc | aggacgagg | gacactgtc | agcttcatgc | 1920 |
| ggggacccctt | ggtgcaccc | gggtgaccc | ctaggaaggc | ccgaggc | ccgc accctgc | 1980 |
| ctgccaagaa | gaaaccacca | ccaccaccc | cgccaggacgg | gcctgg | ttca cggacactc | 2040 |
| cagacaaacg | ccccaaagaa | acccctggcc | aggatggcc | cagtgg | caag gggtcaag | 2100 |
| ggccacccatc | caggaaggca | ccacgttgc | catcttca | cttgcgttcc | agccctgcag | 2160 |
| ccggggactt | ccccatctca | ccacccctg | aaagcccc | ccactgt | ggcc | 2220 |
| cgacgcggc | gcctcttgc | gctgtgact | catgtgtca | atgcgc | tgc cctgc | 2280 |
| gcctctaacc | tttggggccgg | ctccggcc | cccgag | caaggtt | acccgc | 2340 |
| cgggtgcac | gcctgtatgc | ggggatggcc | cacccatgc | tgtgtgt | gg agggcc | 2400 |
| tcagcctgc | ttttgcacat | gagactgtat | gctacttctc | tgtatgggg | atgagcgt | 2460 |
| cagatgtaga | ggccgaggac | ggtgggg | tc | ccgc | ggggcaggag | 2520 |
| agggttgc | cattggcgt | ctggctt | aactcacc | cttcc | ccgcggcc | 2580 |
| ccctggccc | ccccacaggc | ctcagcc | tcacactgc | catttcc | ctctgtgt | 2640 |
| tgtcccacag | cctcgaggct | gcacatcc | ctgtgtt | ttttt | aaatagagatgt | 2700 |
| ttgaatttca | cactgtgtc | tttcttctgt | gctggc | ccat | ttcc | 2760 |
| ggctccggcc | gttaggcac | tgccagg | ccgc | gacccat | ccctgc | 2820 |
| attgtgttgc | tcctgggt | atgcacaa | ggtgg | ccat | ccgc | 2880 |
| tgcacacttgc | tgacacagac | tgtgttgg | gat | ttttt | tttgc | 2940 |
| gggccttacc | cgcccttgc | ccagatcc | ca | tttgggg | tttgc | 3000 |
| tcacccatgc | ccccaaagg | tataataaa | gat | tttgggg | tttgc | 3060 |
| aagcacaacg | atgtttatgt | agaaacact | gagg | tttgc | tttgc | 3120 |
| ggaaactgaa | agctgttcc | gat | ccctgc | tttgc | tttgc | 3180 |
| caccctcccc | tcccttgc | ctctgc | ccgc | tttgc | tttgc | 3240 |
| gggttccccc | atcccttcc | gatgtgt | ccat | tttgc | tttgc | 3300 |
| tgcacgttgc | ggtggggg | tgttaggg | ggc | tttgc | tttgc | 3360 |
| agaccaggccc | actgttagata | atcacttgc | ggc | tttgc | tttgc | 3420 |
| agtctcgcc | gaagtgcgtt | cacgtc | ca | tttgc | tttgc | 3480 |
| ttaaagtttc | ctatgttgc | tacaatgt | gtt | tttgc | tttgc | 3540 |
| ctcccttccaa | caagcgttgc | gatccc | gggg | tttgc | tttgc | 3600 |
| ctgtgttgc | ttttgtgttgc | tttgc | gggg | tttgc | tttgc | 3660 |
| cccccccccc | gtctgc | ccat | tttgc | tttgc | tttgc | 3720 |
| cccaaggatgt | tgctcttgc | agtttgc | tttgc | tttgc | tttgc | 3780 |
| aaaaaaaag | gaattccgt | aaaat | ttttt | tttgc | tttgc | 3840 |
| ttccaccaat | ggtaaaac | gaa | tttgc | tttgc | tttgc | 3900 |
| cagtgcacat | attctgttgc | tgccaa | tttgc | tttgc | tttgc | 3960 |
| cctgacaact | tgtcatttgc | actttttt | tttgc | tttgc | tttgc | 4020 |
| aacatgttca | ttgcacacac | ccaagg | tttgc | tttgc | tttgc | 4080 |
| ctgtgttgc | atgttgc | tttgc | gggg | tttgc | tttgc | 4140 |
| tttccatcaac | tgtgagactc | acag | tttgc | tttgc | tttgc | 4200 |
| acgtcttac | catgtgttgc | cct | tttgc | tttgc | tttgc | 4260 |
| aacagcttgc | ccttcatcca | ccc | tttgc | tttgc | tttgc | 4320 |
| tcatcgacac | tcaatttca | tttgc | tttgc | tttgc | tttgc | 4380 |
| ttcagaggat | gggacaaaat | tgt | tttgc | tttgc | tttgc | 4440 |
| caaggatttgc | gggcgttgg | atggc | tttgc | tttgc | tttgc | 4500 |
| ccttggatgt | agtgcacat | cgtgt | tttgc | tttgc | tttgc | 4560 |

-continued

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| tgcccttggaa | cagaacttct | tccttccccca | tgctttgggt | cacccgggc | tgcaaccctg | 4620 |
| tctgtcccg | atggcccggt | ctgacccctgc | aggaagcaaa | gagggtgact | taaagaacaa | 4680 |
| ccaaactctg | ccagggggtcc | cagaagccc | agggtccage | agtcctcaga | cttggccct | 4740 |
| tgcccttca | caccatctg | gggcaggggc | tgggcctccc | tggtggcagg | ggtgggttgg | 4800 |
| gaatttaggg | gagggtgca | cgagtctgc | cccttgcctc | gggtggctg | gtgttctcc | 4860 |
| aagagctct | gtcacattg | ttggcctctg | gattctggcc | cttcttcatt | ggctgttgt | 4920 |
| ttggactgga | ctgttgcgt | gectgtgtga | tgcaagaccc | agatgtctgt | taggttgt | 4980 |
| gggtgtcg | aggggagggg | gggtggctt | catttgggt | gcccttcac | tccaggcc | 5040 |
| agccctggag | caatttctt | caggcagctg | tetccacctc | caggatgtcc | agcaggctgc | 5100 |
| aaggagaaagg | atggcagcca | cccatccccc | ccaggatccc | ctgttggtca | 5160 | |
| cagccgttcc | tgtttttt | cggttactg | tecccagtgt | agagggtt | gtgttccgt | 5220 |
| agactgaggc | aggttcttt | tccaggctag | aggtggaggt | agatttttt | ctcaaccaca | 5280 |
| tctgcctcca | cacacagctc | ctccgcagg | aaggagaagg | tgctctgtaa | ctcattctgg | 5340 |
| ctatcgtccc | ccttctca | gacccgttcc | ccaccacact | ccttccccc | catcacatga | 5400 |
| caaaggataa | tgtcaagaa | aagtatttt | atgtatccat | aatgtatttt | aaaacaatg | 5460 |
| agaagaagaa | aggtgaaagg | gttattttt | ttaaatgac | ctgactttagt | gacagtgtgt | 5520 |
| gagcatttgc | aatgtaaagg | cctcagctt | cttggagaag | ccacccccc | tttccagaca | 5580 |
| tagatgttg | atgtttgt | gggggtgtgc | caggccacgt | ctcggtgtc | cgtatgcagg | 5640 |
| catgcctgt | tatactgtgt | atggcacac | ttggactatgc | ttggacaatt | cctagagatt | 5700 |
| caactgcaca | attctaacc | atccggcag | cggctgaact | ttgcatttcc | ttgctaactg | 5760 |
| ccagatgtgg | ccaaccttgc | tccatatgc | aaccactgaa | aatgtatctg | gatttctata | 5820 |
| gcaaggccct | ttggggagg | actctccat | gcccctggcc | tcgtggcc | cattggccaa | 5880 |
| tgagccagg | atggggatct | ttgttcttgc | aggcaccc | tgccacttcc | 5940 | |
| tccctca | agacaaacac | ttgttgcgt | acgtcaggat | gaggggacgg | ttggggatg | 6000 |
| atatgaatgt | cacaggagga | gacacccct | gtttttgtt | caaagaaagt | gtgtgcacat | 6060 |
| ttgttaat | acaagagaaa | tattgaaaat | atattgaaaa | gagcaattt | aaattat | 6120 |
| tggttatgt | tgcataattt | atttctgt | attagaaaaag | attcccttgt | agagaaaaaaa | 6180 |
| tgttattttc | attaacgca | agacattt | ctcccttttgc | tatctgtcc | atgtgcacaa | 6240 |
| cccttaacga | gcaatagaat | gtatggtac | ctgggtgtgg | ccagtgcccg | ctgtgcctg | 6300 |
| catgatttgc | ttttggcg | gctgcata | tcccaagcccc | atccgtctt | gctcaactcat | 6360 |
| ggggggctcc | agacccccc | ccccacggg | cttgcgtcat | agggagocct | ttgcactcc | 6420 |
| ctgtgttgg | caaaacgc | gtataaagg | gtgttttgc | tgc | | 6463 |

| | | | | | | |
|---------------|-------------------------|---------------|-------------|------------|-------------|------|
| SEQ ID NO: 40 | moltype = DNA | length = 2828 | | | | |
| FEATURE | Location/Qualifiers | | | | | |
| source | 1..2828 | | | | | |
| | mol_type = other DNA | | | | | |
| | organism = Homo sapiens | | | | | |
| SEQUENCE: 40 | | | | | | |
| catagatgaa | aatggcaagt | tccctggctt | tccttctgt | caactttcat | gtctccctcc | 60 |
| tcttggtca | gtgtctact | ccttgc | tcagttttc | tgtgttgg | ccctctggc | 120 |
| ccatcctggc | catgggtgg | gaagacgctg | atctgcctgc | tcacctgtc | ccgaccatga | 180 |
| gtgcagagac | catggagctg | aagtggtaa | gttccagact | aaggcaagg | gtgaacgtgt | 240 |
| atgcagatgg | aaaggatgt | gaagacagg | agatgcac | gtatcgagg | agaacttcga | 300 |
| tcttcgggaa | tggcatact | gcaggaaagg | ctgtctccg | aatacaaca | gtcacaggct | 360 |
| ctgacagatgg | aaagatcttgc | tgttatttcc | aaatgttgc | ttctatgaa | aaaggccctgg | 420 |
| tggagatggaa | gttgcagca | ctgggttca | atcttcacgt | cgaatgttgc | gtttatgtgg | 480 |
| atggaggat | ccatctgg | tgcagggtca | ccggctggta | cccccaaccc | caaatacagt | 540 |
| ggagcaacgc | caagggagag | aaacatccag | ctgtggaa | acctgtgtt | gcagatggag | 600 |
| tgggcctata | tgaagttag | gtatctgt | tcatgagg | cggtccggg | gagggtgtat | 660 |
| cctgcattat | caaaaaatcc | ctccctggcc | tggaaaagac | agccagcatt | tccatcgac | 720 |
| acccttctt | ccggacgtcc | cageccttgc | tcgcaccc | ggcaggacc | ctgcctatct | 780 |
| tgtgtgtgt | tctcggg | gcccgttact | tcttgcgttgc | acaacaa | gaaataactg | 840 |
| ctctgttc | ttagatagaa | agtgttgc | atgttgc | atgttgcac | 900 | |
| agcggggat | aaggcttca | gagacccctcc | aggaggact | caagggaaa | aaaatccat | 960 |
| acttgcatt | tggagatgg | tettgcgttgc | atccaaata | gtcgttgc | tgtctata | 1020 |
| gaaaaatgtc | ccttc | ttggggaaa | tggctgttgc | catggacacc | tcctcaact | 1080 |
| ctctgcac | gtatgtt | tgtatccaga | catggcaat | gcatactcc | ttgttttgc | 1140 |
| ggaccagg | atgttgcac | gtgttgc | gccccatgc | ctaccagaca | acccttgc | 1200 |
| atttgaatgg | ctgttgc | tgtgttgc | tggaaatgtt | atgttgc | gacactact | 1260 |
| ggaggtggaa | gtgggggac | aaaaaaatgt | gtatattgg | gtatgttgc | agaacgttgc | 1320 |
| gagaaaaaaa | gtttgggtca | aaatgacacc | ggagaacgg | tactggacta | ttggcctgc | 1380 |
| tgatggaaat | aagtatccgg | ctcttact | gttgcgttgc | gcccagaacc | aaccttgc | 1440 |
| tccttagggaa | gttggggtgc | tccttgcact | tgatgttgc | catatcttgc | tctacaatgc | 1500 |
| cacggatgg | tcttcatct | acacatttgc | gtcagccct | tccttgc | tcgttgc | 1560 |
| tgttattcaga | attttgcac | tggacccac | tccttgc | gtttgc | ccatggccaa | 1620 |
| agagatgttcc | cccgatccc | acctgttgc | tgatcatttgc | ctggagat | actgttgc | 1680 |
| aggcttgc | aatgaaatgt | ggggacgtca | ggctgttgc | acatcttgc | ttctccctgc | 1740 |
| ccagccctgg | gttaagggtc | tcacccttca | caacagcc | tcagacccat | aaagctacag | 1800 |
| gcacacactg | aaggacttca | ctgtatattca | ttaaatttgc | ccatggac | gttgcgttgc | 1860 |
| tttgggtgc | ccttatttgc | cccttata | agataggaa | actgggttgc | agaaaatgt | 1920 |
| attgacttca | caaagcagac | aggaatagt | aacaacag | ctgggttgc | aacaacaatg | 1980 |
| actaacat | atggagaatt | taaaaatgtt | tgatgttgc | gttgcgttgc | tttgcgttgc | 2040 |
| tcactc | aatcttcaca | acaccctgc | aggtgttgc | atttggca | tatggaa | 2100 |
| gaggcagg | aaatattatgt | agcttata | actcacacgg | taatttgc | atgttgc | 2160 |
| tttgcgttgc | cagtc | ccatgttgc | gttgcgttgc | acgcctgat | ttcctgc | 2220 |

-continued

| | | | | | |
|------------------------|-------------|-------------|-------------|-------------|------|
| ggaagagccc acatgttagcc | ctgagggttcc | cttcccaggaa | cagctccagg | atcgagatca | 2280 |
| ctgtgagtgg ttgtggaggat | aagcccccta | tggactcctt | cccagctgtat | tatcgagcc | 2340 |
| tttagaccagg cactcctgg | atggctctg | cagagtgtct | tggttgagag | aataacgttg | 2400 |
| cagttcccac agggcatgtg | actttgaaag | agactagagg | ccacactcag | ttaataatgg | 2460 |
| ggcacagatg tgttcccacc | caacaatgt | gataagtgtat | cgtgcagcca | gagccagcct | 2520 |
| tcottcagtc aagggttcca | ggcagagcaa | ataccctaga | gattctctgt | aatattggta | 2580 |
| atttggatga aggaagctag | aagaatata | gggatgtttt | taatcccact | atggactcag | 2640 |
| tctcctggaa aaggatctgt | ccactcctgg | tcattgggtt | atgttaaacc | catactcctt | 2700 |
| tcaactgctg cctgtctaggg | aaaactgctc | ctcattatca | tcactattat | tgctcaccac | 2760 |
| tgatccctt ctactggca | agtgttgc | aagttctagt | tgttcaataa | attttgttaat | 2820 |
| atagtctga | | | | | 2828 |

SEQ ID NO: 41 moltype = DNA length = 2698
 FEATURE Location/Qualifiers
 source 1..2698
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 41

| | | | | | |
|-------------------------|--------------|---------------|---------------|--------------|------|
| ggcggtccgtg ctgactgagg | cgctgcagcc | aggagccgcg | gcccgtgcc | cagcgctcgc | 60 |
| cgccctcccg cgccgcgc | cgtccccgcg | ccgacatgcg | cttggccggcc | gcccgcgaacg | 120 |
| aggcgatcac ggcccccttg | gccccgtcgg | ggctgtctgg | ctgcaagcag | tgccggcgcccc | 180 |
| gccccgcacca ggacaggaa | cttggcatta | gaattccctcg | accactagga | cagggaccaa | 240 |
| gcagatccat ccaggaaaaa | gagatccctc | aagtggggag | tgaagacgc | cagatgcatt | 300 |
| ctttatttgc agatccctt | gtctttgg | gcccgttgg | taacattacg | ttatgtatgg | 360 |
| ttttccaccc acaatattta | gaaagttct | taaaaactca | gcactatcta | ctgcaatgg | 420 |
| atgggcgtt accccatcac | tatcgact | acatggaaat | aatggctgcg | gcaagacatc | 480 |
| agtgtctcta cttatgttgc | ctgactgtaa | atgatttccat | tcatgttgg | ggggacccca | 540 |
| agtggctcaa tgggtttagag | aatgttccct | aaaaactaca | gaattttaga | gaacttaaca | 600 |
| aagtgttagc ccatagac | tggcttatta | ccaaagaaca | cattggggaa | ctttaaaag | 660 |
| ctgaagagca cagctggtcc | cttgcggaa | ttgtatcatgc | agtagtttta | ctcacacact | 720 |
| atcattctct tcgccttccat | acattcggt | gtgaaatcat | tccagaaatt | cattgtatgg | 780 |
| gtggccacac attcagac | ccttctgtt | gcaactactg | catctgtgac | attacaatg | 840 |
| gcaatcacag tggatgttgc | atgcgggtca | actcagcaga | aaatgtttt | gtaatgtatt | 900 |
| ctttcttgc ggttgaagcc | ctcatggaaa | agatggggca | gttacaggaa | tgtcgagat | 960 |
| aagaagggc aagtccggaa | gagatggcc | cacgttttg | aatagaaaaaa | agagagagta | 1020 |
| tgtttgttct ctcttcat | gatgttgc | ttacaccagg | aagatgtgt | tctcgatcatt | 1080 |
| ttgaggatata tagttatggc | tataaaggat | tctcttagaca | ttggatgtat | gttccaaat | 1140 |
| ttcgtgttca ggactattgc | ttggaaagatc | atggttattc | tttggtaaat | cgcccttata | 1200 |
| caagatgtgg acagttgtt | gatgtaaaaaa | ttcacattgc | ttacaatctt | acttataata | 1260 |
| caatggcaat ggacaaat | gttgcaccc | caatgttttg | acggggcaatt | tggaaactata | 1320 |
| ttcaactgcat gtttggata | agatgtatgt | attatgtact | ttgtgaaatt | aaccagctat | 1380 |
| tggatgttag ctttaaaggat | tatataaaaa | ctgttgcgg | cactctgaa | aagggtacca | 1440 |
| aaagaatgtt tgatajcttc | ttggggcgt | tcaagcactt | tgagaagggt | catgttata | 1500 |
| tgcttcttat agaaggttgc | atgcacagg | aactccctta | tgctctgaga | gccattaccc | 1560 |
| gtctatgc ctgtgcctt | tccttcattt | aaatgttgc | ttggatgtatc | agcagatata | 1620 |
| gtctacaagg gggaaaggta | taagccccag | gaccaatgtt | agacaaaata | attcagaat | 1680 |
| ccattgtgc atgatccctt | tagtttgc | tattttctgt | ttggaaaacca | ctgtggcac | 1740 |
| aacgtgtac ttgtttggc | tttcaagg | agagctgttgc | agacggctg | ccattcacag | 1800 |
| tattttggtt ttgttgc | tttcaagg | tttcaactgttgc | tttcaactgttgc | aaatagtaac | 1860 |
| tctccaaatc ctgtgtctt | tatgttaat | aaagataaca | aaatggggac | atgcaagaa | 1920 |
| tgggacttgc ataatgactt | aaggtttata | ttgttgc | tttgcgttgc | tttgcgttgc | 1980 |
| ctattctgc tggaggatgt | aatagatgttgc | tttttcgttgc | tttgcgttgc | tttgcgttgc | 2040 |
| gaacatgtt actatgtat | cctacatata | taacttgc | tttgcgttgc | tttgcgttgc | 2100 |
| ttatgtgttta attttcttgc | atctaaaaaa | gcatagaatt | tttgcgttgc | tttgcgttgc | 2160 |
| acaaccattt tgatgttgc | tgatgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2220 |
| gatattaaat taaaatggca | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2280 |
| tgtgttaattt ataaagaaat | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2340 |
| ggaggggcgc tttttttttt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2400 |
| ctgttagactt ctgcgcgtt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2460 |
| tccccactttt aaagggttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2520 |
| ggtttatcaaa ttgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2580 |
| atatgttata agtactttgt | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2640 |
| actggatattt taaaataaaag | tttgcgttgc | tttgcgttgc | tttgcgttgc | tttgcgttgc | 2698 |

SEQ ID NO: 42 moltype = DNA length = 5215
 FEATURE Location/Qualifiers
 source 1..5215
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 42

| | | | | | |
|------------------------|------------|------------|------------|-------------|-----|
| cgagcgcggc gcccggat | tgcacccgg | cgcagggttt | cgagccact | tgtcagccgg | 60 |
| ccaaagaaaaa gaagctccgt | ccctcccg | tcacccggct | tcccccaccc | ttgtactcta | 120 |
| aactctgcgc | agggcgacgc | gcccggccac | ggaggcgccg | aggaggagcg | 180 |
| ggccggccgc | gtggccctgg | ggggaggggg | gcccggaggg | aggccggccgc | 240 |
| ggccggccgc | cgccatggca | gtgttttgc | ccggccggcc | cgagctgttgc | 300 |
| gtggccagg | cggtgcggct | gggggggggg | cgccggccgc | gttgcgtcc | 360 |

-continued

| | | | | | | |
|------------|-------------|------------|------------|-------------|------------|------|
| gatgagcacg | gaggcggacg | aggcatca | tttctgtg | ccaccctcg | ccccctcg | 420 |
| cttctgcacc | atccccggg | gcccacatcg | caggagggg | ggagcggccg | cggtggcg | 480 |
| ggggcggagg | caccatgc | caccggccg | ggggggcage | ttctggaa | tgagagcgc | 540 |
| cgtgcgcct | ggcattcg | gtccggccg | accctctcg | agcagtgc | cccgaggccg | 600 |
| gggcagct | gttgggggg | gcagccgacg | gaccacgg | gcatatgt | tcaacgaagc | 660 |
| gagccaagg | caactgg | tggccgagag | cgaggccct | cagagcttc | ggggggcg | 720 |
| cgagacagt | ggggccaccc | tggaaaccct | gcattttgg | aaactcgact | ttggagaaac | 780 |
| caccgtgt | gaccgtttt | acaatcg | tattgcgg | gtggagatg | gcatgcct | 840 |
| ccggcagccg | tcctgtttt | accacctgg | ggtgagagaa | agtttcag | tggccaacaa | 900 |
| catcatcctc | tactgtata | ctaactcg | ctctctcg | taactgaagg | aaataattt | 960 |
| ccagaagaat | atatgtca | ctgggg | cacctttgt | ctttacatg | taactccaca | 1020 |
| taacaaagg | tactgtgt | acagcagtt | catgaagg | ttgacagagc | tcatgcacc | 1080 |
| gaacttcgag | ctgcttct | gaccatcg | cttacctt | gtggatcg | ttattcaact | 1140 |
| tttgaaggt | gcacaacgg | gttctagca | gtacttccgg | aatctat | tcaatgcac | 1200 |
| caggaaagct | ctgatattat | acactgtt | agaatttgg | gtcgatgtt | caagaattcg | 1260 |
| gcacgcgat | gataatatcg | aagtcttgc | agcagatatt | gtcataatc | tggtactt | 1320 |
| ctacagagat | atccaggact | atgattctat | tgtgaagct | gttagagactt | tagaaaaact | 1380 |
| gccaaccc | tttgcattt | ccatccat | tgtgaagtt | cattatgc | ttgactgaa | 1440 |
| taggagaaat | tccectgg | acagacaaa | agcttgc | attatgtt | ccatgg | 1500 |
| aagcgaagg | caagttgtt | catatgt | ttgcctgtt | gttcgaatct | acaaagat | 1560 |
| gtttttggac | tctatttca | cgggact | aa | catggag | cttgg | 1620 |
| aaaggcattt | gaatctg | caactact | gtcaggaa | attatgc | ccctcc | 1680 |
| ggoagctg | caccatgtt | aatcttcc | tgagctt | aaagtgg | tgaagct | 1740 |
| tagtctttt | gtttaaaaagg | gaaacttgg | aaaactcc | agctactgg | aagtgg | 1800 |
| ttttctgggg | gocagegt | tagecaat | ccacatg | gtcattca | catctgaa | 1860 |
| gctttttaaa | ctgaagacac | cagcatgt | cctcaagt | atttgc | aaat | 1920 |
| atataagoat | tttgcgaa | tgccacac | acagctgt | gccaagoa | aacttgg | 1980 |
| ctttttggat | ttttccgt | tccgggg | aaagacat | tttactgt | tttaggtt | 2040 |
| agtatttaat | tttagacc | ccaaatct | tcaacctt | tatttgc | tcaacat | 2100 |
| agttgaggaa | aagcaatct | ctat | tttgc | gtgacaa | aaggta | 2160 |
| tgagtggaa | tttgcgtt | tttgcgtt | ggggat | atttctaa | tttgc | 2220 |
| atgctgtt | tttgcgt | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 2280 |
| acttcattt | aaaaat | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 2340 |
| cacagaggaa | ggagactgt | aaagtgc | gttgcgtt | tttgcgtt | tttgcgtt | 2400 |
| tggtgcac | gtcgtttt | gaaaagg | ttatgg | gtctacgg | gtcgggact | 2460 |
| gagcaaccaa | gtcagaattt | cttata | atccc | agagacac | gatact | 2520 |
| gcccctgc | gaagaaat | cattgc | acac | tttgcgtt | tttgcgtt | 2580 |
| tctgggtct | ttcagt | atgg | tttgcgtt | tttgcgtt | tttgcgtt | 2640 |
| aagtctttt | gtctcc | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 2700 |
| ttgtttttt | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 2760 |
| taaccgggac | ataaaagg | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 2820 |
| ctctgactt | ggaat | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 2880 |
| tggtaccc | cagat | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 2940 |
| agcagcagac | atctgtt | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3000 |
| attttatgaa | cttgg | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3060 |
| ccctgagatc | ccagat | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3120 |
| accagatctt | gcaagag | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3180 |
| ttaaagg | aaaagg | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3240 |
| atatctcagg | agtat | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3300 |
| tgagtacgg | tcagtt | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3360 |
| aagagcc | ttctcg | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3420 |
| tccagatgt | aaat | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3480 |
| atttttcat | cttgg | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3540 |
| agaccaagac | aaaat | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3600 |
| gaaaat | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3660 |
| cactgac | aaaat | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3720 |
| cagccatgc | atagcc | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3780 |
| agttcttgc | aatcata | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3840 |
| gaaggcgtt | cagac | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3900 |
| tgcacat | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 3960 |
| acagccttca | aatcata | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 4020 |
| ctcaggcgtt | tgca | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 4080 |
| actgaatgt | cagttgg | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 4140 |
| tccggaaag | aaagaatt | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 4200 |
| attaaacac | cttgg | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 4260 |
| tcttaatct | tctgg | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 4320 |
| tggagctgt | gaagacact | taagocgg | tttgcgtt | tttgcgtt | tttgcgtt | 4380 |
| tctctactat | gttacacgt | atgactt | tttgcgtt | tttgcgtt | tttgcgtt | 4440 |
| cacactgtt | aggcttac | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 4500 |
| aatcttcgat | ggaaattct | aaaatata | tttgcgtt | tttgcgtt | tttgcgtt | 4560 |
| cgaagggaga | ggagaaagg | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 4620 |
| aaaaaaaaaa | cagacaaaac | acactgaa | tttgcgtt | tttgcgtt | tttgcgtt | 4680 |
| aggactt | cataaggact | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 4740 |
| caaacc | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 4800 |
| agat | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 4860 |
| caaacc | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | tttgcgtt | 4920 |

-continued

| | | | |
|-------------------------|------------------------------------|--------------------------|------|
| atctgagtag aaatgcgtaa | tttcaattac tgtataaagt | ttatgtttt ttaagtgtgc | 4980 |
| agaatctcg agacaatggt | tttatcttc tggttaat | gtatatttg ctcttatttg | 5040 |
| taacttaagt ttctgacctg | tctgacatt gtttgagtcg | tttatgtact actgaactgt | 5100 |
| accagttgca catgttcaa | ctgtatgtat gttagctgt | tcttaaagcta tccattgtgt | 5160 |
| catatttact ctaaaaatta | aagagactct caacaaaaaa | aaaaaaaaaa aaaaaa | 5215 |
| SEQ ID NO: 43 | moltype = DNA length = 4655 | | |
| FEATURE | Location/Qualifiers | | |
| source | 1..4655 | | |
| | mol_type = other DNA | | |
| | organism = Homo sapiens | | |
| SEQUENCE: 43 | | | |
| tctgtgcacc ttgcgggggg | cggcgaacgg cagccggggc | agcagctagg gggcttgtgc | 60 |
| acacagcgag ggagacttag | ggactggcag acggacggac | ggacggcgag gaccctaccc | 120 |
| gagccccccg gccatggccg | agagaaagca atccgggaag | gcccggaggg acgaagaaagt | 180 |
| cctctgtttt tttaaaacc | tggctccgg cggcccaaa | ccccggcaga aattctgtgg | 240 |
| catgttctgc ccgggttggaa | ggtctcggaa gaacaagacc | atcgacttcg actegctgtc | 300 |
| gggtgggggg ggctggggc | aggtgttggc tcagcagcgg | gacgtcgccc acttggggcc | 360 |
| ggaccccgag cgccgtact | cgcggggggg cggccgcgc | ggggagagc catctgttga | 420 |
| atcggggccgg aagggtggaa | tccggggggc ctcggggaaa | aaagggctgc agaacatcaa | 480 |
| cgaccagagc gatgttctc | tgtatcaaagg aggtaaaatt | gttaatgtg accagtcttt | 540 |
| ctatgcacat atatacatgg | aaagatgggtt gatcaagcaa | ataggagaaa atctgtattgt | 600 |
| gccaggagga gtgaagacca | tcgaggccc ctccggat | gtgatccccg gagaatattga | 660 |
| cgttccacact cgtttccaga | tgcgtatca gggaatgacg | tctgtgtatc attttttcca | 720 |
| aggaaccaaag gggggctgg | ctggggggaa cactatgtac | attgaccacg ttgttccgtg | 780 |
| gcctgggaca accctgtctg | ctgccttta ccagtggagg | gaatggggcg acagcaagtc | 840 |
| ctgtgtgtac tactctctc | atgtgtacat cagcgagtg | cataaggcga tccaggagga | 900 |
| gtatggaaagcg ctgtgttgg | atcacggggt aaattcccttc | ctcggtatca tggcttccaa | 960 |
| agatcgcttc cagtaacgg | attgcacat ttatgtatgt | ctgagtgtg tccgggatat | 1020 |
| tggccatccat gccccatgtcc | acgcacggaaa tggcgcacat | attgcacagg agcagcagag | 1080 |
| gatccctggat ctgggcata | cggggccccc gggacatgtg | ctgacggac ctgaggaggt | 1140 |
| cgaggccggaa gccgtgaatc | gtgcatac catcgccaa | cagacaaact gcccgtgt | 1200 |
| tatcaccaaag gtgtatggca | aaaggtctgc tgaggatcat | gcccaggcac ggaagaagggg | 1260 |
| aactgtgtt tatggcgagc | ccatcaactgo cagcttggaa | acggacggct cccattactg | 1320 |
| gagcaaaacat tggggccaa | ctgtgcctt tgcgcaccc | ccacccatga gcccgtatcc | 1380 |
| aaccacttca gacttttca | actcttctgc tgcctgtgg | gaccctccgg tcacggccag | 1440 |
| tgcccatgtc acgttttaca | ctggcccgaa ggtgtatgg | aaggacaact tcaccctgtat | 1500 |
| tccggggggc accaatggca | ctgaggagcg gatgtccgtc | atctggggaca aggctgtgg | 1560 |
| cactgtggat atggatggaa | accagtgtt ggtgtgtacc | agccacaaatg cagccaaatg | 1620 |
| cttcaacattt tacccccgg | acggccatgtt tgcgcaccc | ccacccatcc acctgggtcat | 1680 |
| ctggggccccc gacagcggtt | aaaccatctc tgccaaagaca | cacaacagct ctctcgagta | 1740 |
| caacatctt gaaaggatgg | agtgcgcggg ctccccactg | gtggatcatca gccaggggaa | 1800 |
| gatgtctgtt gaggagccg | ccctgtcatg cccgggggg | tctggacgtc acatccccgg | 1860 |
| gaagcccttc ctgtatttt | tttataaaggc tataaaggcg | aggacgggc tggctgagct | 1920 |
| gagagggggtt ctctgtggcc | tgtatgcggg acctgtgtgt | gaagtgtctg tgacggccaa | 1980 |
| gacagtcaat ccacgttcc | cggccaaagac gtctctgtcc | aaggcggggg ccccacctgt | 2040 |
| ccggaaacctg caccgtctg | gattcaatgtt gtcgtgtgt | cagatgtatc acaacatcc | 2100 |
| cccgccgacc accacggcc | tegtggccgccccgggggg | cgtggccaaaca tcaccagct | 2160 |
| gggcttagagc ttctgggtg | tgccgtccac tggggactgg | ggatggggaca cctggggaca | 2220 |
| ttctcgagact tttttttttt | ttcctttttt tttttttttt | ttttttttt agacgtctgt | 2280 |
| atagtatctg tggggatgg | gtcccccttc gggccccaca | ccccgtcttc | 2340 |
| caccaagatgt tactgtttt | gtctatccacat ttccctacac | atctatgggt atcacaccca | 2400 |
| agacttacca ccaagtcat | acacgggaaac acacccaaaca | tttagacatg cgaacaagca | 2460 |
| gccccccagcg aggggttcct | tcgccttcaa ctccttagt | tctgttagca tcttccttt | 2520 |
| catggggggaa gggaaatata | agtatgttc ccagacgtgc | ctttttttt tctttttttt | 2580 |
| aattttttaaga agttttttt | ttggggctgg ggagggggcc | gggttggggaa gagttttttt | 2640 |
| ttttttttttt tttttttttt | ttttttttt atttttttt | ttttttttt aaggggtttt | 2700 |
| aactggacat ctaatgtat | caattacgtc atcacccagg | tgatccggg tggttggcaa | 2760 |
| acttcatcggt tctgtctgt | gagggtccac aatgcccacc | ccgcateccca ttctgtgtc | 2820 |
| ttcagggttca gctgtgtata | aaaggcggagg ctggcttagt | tggcttagat ttgtgtgcag | 2880 |
| atttaatctt ttgaggatc | tcttccttc taccatttt | tcgcgtgc tcactcttc | 2940 |
| tttctctctc tagtttttta | attcatgtat ttttctgt | ctgtctctc tcttcctgt | 3000 |
| gtttccatcca gcccgggtt | cgggagacggt gttttccccc | tttgccccc tatctttca | 3060 |
| ccctccggat ctaccatgtt | atggggggcc cttaaaggat | tgatgggtt tgatgggtt | 3120 |
| ccattgtcaag catgtgtctg | tgtcatctgt gtcgtatgt | gactgtgtct aaccacccgt | 3180 |
| catcatgggg agtgtgtcta | gagtgtggaa ccttggccaa | gtgcggaaat gggccatgcc | 3240 |
| gtctcaccca cagtatcaca | cgtggaaaccg cagacaggggc | ccagaagctt tagaggtat | 3300 |
| aggctgcaga accggagaga | ttttccctgtc tgcaatgt | tctggataaa gtcacgggtca | 3360 |
| aacttaaaca cccggatct | ttaacccaaatg tgaaaccaacc | aaagtccacca gttcagaatgt | 3420 |
| gctaagatcaa tagggatctg | acccggggcc ctgtgtcttc | ctgggtttaatg atcttttgag | 3480 |
| attcttagaac acatggggc | ttttttttt ccggggaaaaaa | ccgttattttt ttcttgccttca | 3540 |
| attattttcta aagacacact | acatagaaaag agggccctata | aactcaaaaa gtcattggga | 3600 |
| aacttaaagt ctattctact | tttgcaggggg agaaatgtgt | tttatgtatgtc atagatcaca | 3660 |
| tcagaactcc tggggggagg | aaacccatata aattaaacac | atggccccc tagagaccac | 3720 |
| aggtgtatgtc ttgtccatc | cttccctctc cttttctgtc | acctttcccc ctatgtgtct | 3780 |
| cctttggacc taccctgtc | cttgcgtact tttgtgttgc | actgtttttt acgtgtttac | 3840 |

-continued

| | | | | | | |
|--------------|-------------|--------------|--------------|-------------|-------------|------|
| aggttcttt | aagcaatgtt | gtatttgcag | gcttttctga | ataccaaatc | tgcttttgt | 3900 |
| aaaggctaaa | aacatcacaa | agtaggtcat | tccatcacca | cccttgcctc | tctacacatt | 3960 |
| ttgccttgg | ggatctgggtt | gggggtttttgg | gtttttttgtt | gttgggtttt | atttgttatt | 4020 |
| ttaaaggtaa | atgcacttt | aaaaaaaata | atgggttgc | ttaatatatt | tgctttttt | 4080 |
| ctcacctgc | cttagaggaa | attnaaca | gttggaaaaaa | aacaattttt | gttcaattc | 4140 |
| taagaaacac | ttgcagctct | agtttact | tgagtcttc | tggttttct | gtacgggtc | 4200 |
| atggtaattt | ttgggttgg | ttgggttgg | cttaaaaaaaaa | aagttaaaaac | ctgacgattt | 4260 |
| ctgcaggotg | tgtaaqcatg | tttaactgtt | gggttgc | tggtgtttgt | taatgtatg | 4320 |
| tcatatgtaa | atgcataaaat | aaatcgacag | tgtctcagaa | ctgataact | gcagtgcatt | 4380 |
| gtatgcctaa | aacatgttag | gattnaaga | tagatgtttt | taatcttgg | aaattgtgtat | 4440 |
| tgtgaccat | gagtggagga | actttcgtt | ctaaagctgt | taaagtgtgt | agccagaaga | 4500 |
| gtactttttt | ttttgttaacc | actgtcttg | tggcaaaaata | attatgttaa | aaaacaagtc | 4560 |
| tcgtgtttat | tattcttaa | gaactctgt | ttatattacc | atgaaacgcc | taataaagca | 4620 |
| aatatgtgggtt | gtttcaggaa | aaaaaaaaaa | aaaaaa | | | 4655 |

SEQ ID NO: 44 moltype = DNA length = 4417
 FEATURE Location/Qualifiers
 source 1..4417
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 44

| | | | | | | |
|--------------|--------------|-------------|-------------|--------------|-------------|------|
| gtgtgtacac | tcacccgtgaa | ggtgcgago | ttcactccgg | agccagcgag | accacgaacc | 60 |
| caccagaagg | agaagaactct | gaacacatct | gaacatcaga | agggacagac | tccagacgcg | 120 |
| ccaccactct | gtaaacacca | gatagtgaa | agaacacccat | tgctgaaatg | tttgcacaca | 180 |
| ctgatggttt | gactctgtca | actggaaatgg | cttattgtgc | aagaaagtac | acctgggtcg | 240 |
| gtctctggggc | tcatctctag | caccagaaa | gatttctgaa | gacgttcttc | tagaaatgac | 300 |
| tggaaagttt | caagaggcat | aaagatacago | atttcttctg | aggccctgaa | gaagtatcaa | 360 |
| gtggggcttgg | acattgcgg | gggtgagagcg | acccccttc | accttggagaa | ctggggaaatg | 420 |
| tggattctca | gggacccgc | tgttcacgag | ctccaggctg | tgctgtggc | cctggctctg | 480 |
| ggggcgcgtag | cogcatctgc | aatagcacac | ttggccggcc | acctgtgtcc | gtgagccctt | 540 |
| gtctgcgtac | ccccctgggtt | cgccatctacc | tgatgaggat | gtgcaccccc | attagggggc | 600 |
| tgtctatgtc | cottgcgt | atgttggga | cagcgatggc | atttgcaccc | atacccccga | 660 |
| tacacccgttgg | gcacagagag | gtgcacccctg | tgagtttca | tgagccagac | atctacaact | 720 |
| acttcagcc | gtctgtgtgc | gaggacaagg | acacccttgc | cataggtgcc | cggggaggcgg | 780 |
| tcttcgtgttgc | gaacgcactc | aaacatctcc | agaagcagac | tgagggtgtat | ttggaaaggct | 840 |
| cagaagacaa | aaaagacaaa | tgtgcagaaa | agggaaaatc | aaaacagaca | gagtgcctca | 900 |
| actacatccg | ggtgcgtcg | ccacttcacg | ccacttcctt | ttacgtgtgt | gggaccaacg | 960 |
| cattccgc | ggcctgtgac | cacctgaact | taacatcctt | taagtttctg | gggaaaaatg | 1020 |
| aaagatggca | aggaqaatgt | cccttcgc | cagcacacac | ctacacatcc | gtcatggtt | 1080 |
| atggagaact | ttattcgggg | acgtcgatata | atttttgggg | aaatgttgc | atcatctccc | 1140 |
| gaaaattcttcc | ccacagtcct | ctgaggacag | aatatgtca | cccttgcgtc | aacgaccta | 1200 |
| gtttcgtgtt | tgctgtgtgc | atccgaaaaa | gcccagacag | ccccgacggc | gaggatgaca | 1260 |
| gggttctactt | tttttcacg | gagggtgtctg | tgaggatgtat | gtttgtgttc | agggtgtctg | 1320 |
| tcccacggat | ageaagagt | tgcaaggggg | accagggggg | cctgaggacc | ttgcagaaga | 1380 |
| aatgggcctt | tttccgtaaa | ggccgactca | tctgctccgg | ccacagacgc | ggcttggct | 1440 |
| tcaatgtgt | ggggatgtc | ttcgtgtca | ggtccccggg | cctgagggtg | cctgtgtct | 1500 |
| atgcacttc | caccggccac | ctgaaaca | ttgggtgtgtc | ggcgtgtgc | gcctacaacc | 1560 |
| tgtccacacg | cgaggagggt | tttcccaacg | ggaaggtacat | gcagcgacc | acagtggagc | 1620 |
| agtccccacac | caagttgggt | cgctataatg | ggccgggtacc | caagccgggg | cctggaggcgt | 1680 |
| gcatecgac | cgagggacacgg | ggccccaact | acaccagctc | tttgcatttg | ccagacaaga | 1740 |
| cgctgcgtt | cgtaaaagac | cacccttgc | ttggatgtac | ggtaacccca | atagacaaca | 1800 |
| ggcccgaggat | atacaagaaa | gatgttgc | acacccaggat | ctgggtggac | cggacccagg | 1860 |
| ccttggatgg | gactgtctat | gatgtcatgt | ttgtcagcac | aaacccgggg | gctctgcaca | 1920 |
| aaggccatcg | cttcgacac | gttgcacaca | tcatcgagg | gaccacgtc | tttccaggact | 1980 |
| tttgcggcgtt | ccagggcttc | ctgtgttgc | caaaaagggg | caacagggtt | gtctatgtc | 2040 |
| getctacttc | gggcgtgttc | caggccccgg | tttgccttc | ttggggccccc | ttggccaggct | 2100 |
| aggactgtgt | gttggcgccg | gacccctact | ggcccttggag | ccggcccccaca | ggcacctgc | 2160 |
| tggctgtca | ccagaccgg | agccccagca | ggggtttgc | tcaggatgt | agccggcgat | 2220 |
| tttctgttgt | ccggcccttc | tcttcataa | ccctcccttc | ttctgggtcc | tcttcctgt | 2280 |
| cctgtctggg | ccatgtgggg | gacaggaggc | tttcccttc | cttgcacccccc | ttggccaggct | 2340 |
| cggggtgcggg | ggccgacacgc | agctcgagg | tctccctgt | ggcccccctc | ctgagtgtac | 2400 |
| aggcacacga | ctgtgcacgc | ctggggaaact | tctaccttc | ctggcaggccc | acaggtctcg | 2460 |
| cagacattcg | ttttgtctgg | gagaagaatg | ggcgagctt | ggagacccgt | gttccctgtc | 2520 |
| agacccatgc | actggccatgc | ggcagggggg | atgcacttc | ctggctgcac | gacgcctatca | 2580 |
| ggggaaaggc | tgatgtatgc | tgctgtgttc | tctccatc | agggaaaca | acttcgaaagg | 2640 |
| tgcagggtgc | tgtgtatgaa | cctgaatgt | ccccaccagg | gagggtggacc | agagagctt | 2700 |
| ctgcctggag | ggctgtggct | ggggggacacg | accggatgt | gcagagotgg | aggaaggcgt | 2760 |
| ggggaaagctg | tagcaaggac | accctgtac | caccaggaa | gagtccctga | caccgactc | 2820 |
| aaccccaaca | agacccctgc | gcccactgac | acagccaccc | ccggagaagg | cctggccccc | 2880 |
| cacaactgt | aactgtcttgc | cccaactgc | ctctgaacac | agccattggg | ccaccacact | 2940 |
| atgggcacag | ggggacacgt | ggaaaggtt | ggaaacccaa | ttggccctgt | acaggaacta | 3000 |
| agacttaaaa | aatttaggtgc | tttactggaa | cagtaatgtt | tgtctggc | aagcaggtaa | 3060 |
| ccaggatggc | taacaggcctt | tgtatgtc | tctgtacta | aaacagcagg | gttgcgtc | 3120 |
| tttcccttc | tacggcgtc | cagcaggctc | tgaaggctga | ttctacaccc | tcccagtgac | 3180 |
| ttcccttgac | agagtgc | cccccttaa | tagccaa | ggttagcatg | ggccacacag | 3240 |
| atcgctgttt | ttattgtatgc | aatcaagcc | tgtgtttct | cctccctgc | gacttagcc | 3300 |

-continued

| | | | | | | |
|-------------|--------------|-------------|------------|------------|-------------|------|
| aggaactcca | agatgcata | ctgggacaag | aaaaggtga | actccacatg | gaaatgcctt | 3360 |
| gccttaaac | ttgaatgact | gtgagatgg | atctgggagt | gcatctgtca | agtcttttgt | 3420 |
| tttcttcac | taacctcaga | atactgggt | ctattttac | aagcgctgca | gtttatgcct | 3480 |
| ctgtccgc | aatgcctcgc | ttctgcaaca | ggacaccaa | cttgatgcag | aaaggccaaat | 3540 |
| aggtcaatta | tgcaaatctc | ctgggtccat | attaaatttc | ttgacgatgg | aatgagtctc | 3600 |
| atgagtgttt | tgttctacct | gcttcaagt | ctctaattat | taaagctgt | tctctgaaga | 3660 |
| ctgtgtca | tggtgtgt | acttgecta | aagctacta | gcctttaatc | ttacacacac | 3720 |
| gtttcttctt | gtctgttga | tgacagttt | catgtctatc | ataaaaacaa | agcctctgtt | 3780 |
| aaaagtcaag | ccgcacccct | ctgggtatcc | tagcaaatac | tgagtgctt | cccagcagt | 3840 |
| tgacaatgac | ctgttgcata | tccecttctt | ctggagctgg | acaaattctc | taccagcct | 3900 |
| tgtgtgggat | cagcatacat | cgcgttata | ttctttaggg | atccatcaca | acaggtgtcc | 3960 |
| tgaagatgt | ggagacaccc | tgggtgtctc | cacacgttcc | ccctccgcac | cccaagtgcg | 4020 |
| gaggcccage | tgccgtgt | gtgtgtgtt | gccccatccg | ccaaggatgc | cagtttgc | 4080 |
| cacggaaacca | tcacatactc | ataacacttca | gttttcttgc | aaaatatcca | tcagctca | 4140 |
| gtgtgttctt | ctttgggtgt | ggcttcaacc | actacaaact | gatgagtgg | atgtatggg | 4200 |
| ctttaggtt | atatttttgc | tgtgttttgc | tggttcttct | cctgaaatgc | ggatccaag | 4260 |
| cacttcaca | cttaacaaaa | taattacata | cttgaagttt | tcgtatgt | gagtgttcta | 4320 |
| ctgggaaatg | gagttatgag | gatgaatcc | tgagtcttcc | tttgctctgc | tggaaaaaaat | 4380 |
| aaaaatagag | ttgtacatttgc | aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | 4417 |

SEQ ID NO: 45 moltype = DNA length = 3108
 FEATURE Location/Qualifiers
 source 1..3108
 mol_type = other DNA
 organism = Homo sapiens

SEQUENCE: 45

| | | | | | | |
|-------------|--------------|--------------|--------------|--------------|-------------|------|
| gcctctgtct | cctcaggcga | ttccoggcg | ctccgggtt | ggcaacgcgg | acggggggagt | 60 |
| gcgactcggt | ctcgccgc | atggccgaga | agcggcacac | acgggactcc | gaagccgc | 120 |
| ggctcccccga | ctccctcaag | gacagccca | gttaaggcc | tggacettgc | ggatggattt | 180 |
| tggtggcgtt | ttcattttca | ttcaccgtt | taatcttcc | aatctcaata | tggatgtgc | 240 |
| taaagaggat | aaaagaggat | gaaagagcc | tcatcttttgc | atttgcatttgc | attttacaaag | 300 |
| gaggagccaa | aggacccgtt | tggtttttgc | tctcgccat | cactgcacgc | ttcatcaaaag | 360 |
| tggacatgt | aactatttca | tttgatatttgc | ctctctcaga | gatcctcaca | aaggatttgc | 420 |
| tgacaatttgc | ctgttgcgtt | gtgttctatt | accgcgttca | gaatgcaccc | ctggctgtgg | 480 |
| caaatatccac | caacgcgtac | tcagcaaccc | gtcttttggc | acaaactact | ctgaggat | 540 |
| ttctgggcac | caagaatctt | tctcgtatcc | tctctgcac | agaagaaaatt | gcacacaaaca | 600 |
| tgcagtcac | tctggatgt | gccactgt | cctggggaaat | aaaggtggag | cgtgtggaaa | 660 |
| ttaaggatgt | gaaactaccc | gtgcagctcc | agagagctat | ggctgcagaa | gcagaagcgt | 720 |
| cccgccggc | cccgccggc | gtttatgtcc | cccgaggaga | aatgaatgca | tccaggggctc | 780 |
| tgaagaagac | ctccatggt | actactgt | gttgcataat | cctctgcac | cgataccctgc | 840 |
| agacactgc | caccatgt | gctgagaaaa | actcaacaat | tgttcttcc | ctgccccatag | 900 |
| atatgcgtca | aggaatcata | ggggcaaaac | acagccatct | aggctatgt | agagatgagc | 960 |
| gtctggcc | caagcatgaa | gtcgccggacc | aaatttgcct | ttaactcata | aagagagggt | 1020 |
| agggtttttc | tttttttgc | tgtcaatttgc | gggtttttca | gaatgtatag | cagtattaaa | 1080 |
| aatagggttgc | aaatgttgc | gcttgcataat | actggagat | ttgtgttata | tataaggtaa | 1140 |
| tctgttagtc | ttaaaaatgt | taaaaggtttgc | tatttttgc | tttattatgt | gtagggttgc | 1200 |
| tcctcttgc | tttgcattcc | actgtatct | tctgcacccc | ctaagcaccc | aggccagagg | 1260 |
| caagaacccgt | ggctgttaact | gcccacccgt | accgcgtact | ggcttaatgc | tttgcagaaaa | 1320 |
| gtgtatgtt | tacaccacaa | ccagtttcc | cggttcatat | gtgccttacc | tccagaggt | 1380 |
| ctttttttt | ttttttctga | gatggattttgc | cacttttgc | gcccaggctg | gagtgcata | 1440 |
| gcgtatgtt | ggcttactgc | aacccccc | tcttgggtt | aagaggatct | cctgcctc | 1500 |
| cctcccccgt | agctgggtt | acaggctcat | gcccacccgt | ccagcttaat | tttgcattat | 1560 |
| tattttttgtt | ttttgttgc | gacgggggtt | cacccatgttgc | ccaggctatgt | tcacgact | 1620 |
| ctaaccttc | gtgatccacc | cacccatgttgc | tcccaaagtgc | ctgggattac | aggcatgagc | 1680 |
| taccacacgt | gtttttggaga | gttcaatttgc | aggaaatttgc | cataatgttgc | atttattttc | 1740 |
| taaatccac | ccgtgttca | gataatacttgc | tacttgc | ttggatgttgc | tttgcctgt | 1800 |
| cttgcacaa | cttagggaaa | tagccaaatgc | taatgttgc | tttgcctgt | tttgcctgt | 1860 |
| gactttaaa | ttcaactacta | gagtggat | tttgcacccat | tttgcacccat | tttgcacccat | 1920 |
| tgagccttct | ttatgttgc | atgtatcc | cttgcacccat | tttgcacccat | tttgcacccat | 1980 |
| ttgttcacaa | agtgttttttgc | tctcactgttgc | gcttgcataat | tttgcacccat | tttgcacccat | 2040 |
| ttttaattttt | tttataaaaat | tttgcacccat | tttgcacccat | tttgcacccat | tttgcacccat | 2100 |
| aatagctgaa | ggactaaaaat | acttttttgc | tttgcacccat | tttgcacccat | tttgcacccat | 2160 |
| ctatctgtat | gtgttgcact | gtggat | tttgcacccat | tttgcacccat | tttgcacccat | 2220 |
| taatgtatgt | aaaggatcc | cgtatataatgc | actatataatgc | gtctatgtat | cttggccaga | 2280 |
| tactgtatgt | tttataatgc | ctatataatgc | gttaggttgc | tttataatgc | ctatgtatgt | 2340 |
| tttgcacccat | ttatgtatgt | cttgcacccat | tttgcacccat | tttgcacccat | tttgcacccat | 2400 |
| tttagtgcata | ttatgtatgt | cttgcacccat | tttgcacccat | tttgcacccat | tttgcacccat | 2460 |
| ttggcatttgc | tttactggcc | agaaaatgc | tttgcacccat | tttgcacccat | tttgcacccat | 2520 |
| gactctgtat | tcacccat | aaatccat | tttgcacccat | tttgcacccat | tttgcacccat | 2580 |
| tatagcttc | tttaggttgc | ccactctgtc | catccacccat | tttgcacccat | tttgcacccat | 2640 |
| accccaattt | caagtttgc | tttgcacccat | tttgcacccat | tttgcacccat | tttgcacccat | 2700 |
| gagtttataa | gagtttataa | tttgcacccat | tttgcacccat | tttgcacccat | tttgcacccat | 2760 |
| tttaaaaatgt | tttttttttgc | tttgcacccat | tttgcacccat | tttgcacccat | tttgcacccat | 2820 |
| catatctgtat | ctgaaaatcc | tttgcacccat | tttgcacccat | tttgcacccat | tttgcacccat | 2880 |
| ttctctgtc | tccagatacc | aaatgcacccat | tttgcacccat | tttgcacccat | tttgcacccat | 2940 |
| atcagctgaa | attaaatcc | tttgcacccat | tttgcacccat | tttgcacccat | tttgcacccat | 3000 |

-continued

| | | | | | |
|-----------------------|-----------------------------|-------------|-------------|-------------|------|
| cagtgcagaa gtttctccag | gtggcctccc | tttccaacaa | tgtacataat | aaagtgtatg | 3060 |
| cacttcact | aataaaaaaa | aaaaaaaaaa | aaaaaaaaaa | aaaaaaaaaa | 3108 |
| <hr/> | | | | | |
| SEQ ID NO: 46 | moltype = DNA length = 4090 | | | | |
| FEATURE | Location/Qualifiers | | | | |
| source | 1..4090 | | | | |
| | mol_type = other DNA | | | | |
| | organism = Homo sapiens | | | | |
| SEQUENCE: 46 | | | | | |
| ggcgcgcccc ggaggtatcg | caaagggtt | cggcccgcccc | acagtgcgg | gtcccccccg | 60 |
| gttatcaaaa gaaggatcg | ctccgcccc | gggtcccc | ggggagttga | tagaagggtc | 120 |
| cttccccacc | tttgcgtcc | ccactcctgt | gcctacgacc | caggagcgtg | 180 |
| catggagaat | caagagaagg | cgagtatcg | gggccacatg | tgcgaegttag | 240 |
| aggtggcatt | tcaggactat | ctggtccaa | actcttgact | aatatggcg | 300 |
| ggttttagaa | gtcgggaca | gggtggagg | aagaacatata | actataaagg | 360 |
| tgattagtgta | gtatgggtg | gagttatgt | gggaccaacaa | aaaacagaa | 420 |
| gtctaaggag | ctggcatag | agactacaa | agtgaatgtc | agtggcg | 480 |
| tgtcaagggg | aaaacatata | catttcgggg | cgcccttcca | ccagttatgg | 540 |
| atattttgat | tacaataatc | tgtggaggac | aatagataac | atggggagg | 600 |
| tgtatgcacc | ttgggaggtc | aacatgtca | caaatggac | aaaatgacca | 660 |
| cattgacaaa | atctgtgg | caaagactg | taggcgtt | gettacttt | 720 |
| caatgtgacc | tctgacgtc | acgaagtgc | tgcctgtgg | ttcttgcgt | 780 |
| gtgcggggcc | accactcgga | tattctctgt | caccaatgtt | ggccaggaa | 840 |
| aggtggatct | ggtcaagtga | gcaacggat | aatggac | ctcggagacc | 900 |
| gaaccatct | gtcactc | ttggacgtt | aatgtacaa | atcatatcg | 960 |
| ccatgacat | tatgtgtt | aatacgtat | taatgcgt | cctccgact | 1020 |
| gattcacttc | agaccaggc | ttccagcaga | gagaaaccagg | ttaattcagc | 1080 |
| gggagctgtc | attaatgtca | tgatgttata | caaggaggcc | ttctggaa | 1140 |
| ctgtggctgc | atgatcattt | aagatgaaga | tgctccaatt | tcaataaac | 1200 |
| caagccagat | gggtcaactgc | ctgcccattat | gggcttcattt | cttgcgggaa | 1260 |
| acttgcgtt | stacataagg | aataaggaa | gaagaaatgt | tgtgagctt | 1320 |
| gtctggatcc | caagaactt | tatcatcg | gcattatgaa | gagaagaact | 1380 |
| gcagactct | gggggtgtc | acacggccta | cttcccttc | gggatcatga | 1440 |
| aagggtatt | ctgtcaacc | tggcaggat | tttcttgcg | ggcacagaga | 1500 |
| gtggagcggc | tacatggaa | ggcagttgt | ggctggaa | cgagcgtca | 1560 |
| aatatgttcc | gggaagggt | ccggaaa | tatctgggt | caagaacctg | 1620 |
| cgttccagcc | tgatggatca | cccacactt | ctggggaa | aacctgcct | 1680 |
| cctgctgaa | atcatggat | tttccacato | agtaactgc | ctggggtttgc | 1740 |
| atacaagctc | ctgcccacgt | tttgaagtt | tgttctttag | ctctctgtc | 1800 |
| ataaccacca | agaggaaaat | attgacaat | ttaaaggctg | tgttcttgg | 1860 |
| gtgtactgga | ttttaactacc | tttgctta | tttcaatcat | tgtttaa | 1920 |
| aaagaatcac | cttaaattt | ttagat | caagctccat | tttatttgc | 1980 |
| aactcatgtt | aattgtat | ataaaac | gtgtactt | tctgttttgc | 2040 |
| acgtgtatgt | ctcatagaa | acaatttctg | tgttctgttt | ttatttcc | 2100 |
| tacatgtga | tttcagaaac | aaacatgtt | acttctgtc | tgttgggg | 2160 |
| aggcccaac | tgtaactgt | ctttttctt | cctttagggaa | tggtaactgt | 2220 |
| gccttagggcc | tcacagcc | tctgggg | cagccctcc | tttggatc | 2280 |
| aggaaaccc | tggtgggg | agccgtc | agaccctcg | accagaatgg | 2340 |
| tggtctgt | cctcaggaat | ctcctgacca | tttgcgttcc | ctccgactt | 2400 |
| tagtctcgt | gtatgtt | tttgcgtt | ccttcttcc | ttttatgtt | 2460 |
| taactcggt | agaatgtt | aaactgtt | ctgtactt | tgttgcgtt | 2520 |
| cagagggaaa | tagtggc | actgtcccc | aggaggactt | cctgttttgc | 2580 |
| accaactac | actggcatct | tcttcttcc | tttggatc | tctgtgggg | 2640 |
| tgtcagtgt | aaacattctat | ttcaacttca | ggaaaggcc | agttttctt | 2700 |
| tatgaccata | aaatttagaca | tacctctaa | tttacatat | tcttcaat | 2760 |
| gcataaaat | tagcaagac | tgccaaat | tcttaatgt | tgaaat | 2820 |
| tgactgtt | tttgtgg | acttata | attaacctt | aaaagaaat | 2880 |
| gcaagctac | ttgttttagt | tgtttagt | tataaaac | tttgcgtt | 2940 |
| atatgaaac | taagttctt | atctttag | taatgtcgg | ccacaagggt | 3000 |
| tgttttcccc | ttttaaaaac | tcagtttt | aaaaggcc | tccaaagggt | 3060 |
| aataacttctt | ttttaatgt | atcaat | ttttatgtt | tcaactgtt | 3120 |
| ggaaaccac | caaccaat | tttataat | attaacctt | tttgcgtt | 3180 |
| tgttggaccc | ttgtgtt | acataatct | tccaaatgtt | tgcctat | 3240 |
| tttttctact | gtctttacc | tttgcatttt | gcttatttgc | tttttttgc | 3300 |
| atacgctcca | gtacataagg | tttataat | tcactgtt | tttttttgc | 3360 |
| tcccttagtct | attgcgtt | tttactt | tttgcgtt | tttttttgc | 3420 |
| tttggggaaat | gtgtt | acttcttac | tttgcgtt | tttttttgc | 3480 |
| taacatattt | aaagtctt | tttgcgtt | tttgcgtt | tttttttgc | 3540 |
| ggaaaacccat | ttaattttt | tttgcgtt | tttgcgtt | tttttttgc | 3600 |
| aaatgaatgc | cgttccagat | tttgcgtt | tttgcgtt | tttttttgc | 3660 |
| tgcaagtca | tccaaacgt | tttgcgtt | tttgcgtt | tttttttgc | 3720 |
| tcagtttaag | aggctgtat | tttgcgtt | tttgcgtt | tttttttgc | 3780 |
| gtcgatgt | ccccacttca | cagttaggcc | aagaattt | ttttaaact | 3840 |

-continued

| | | | | | | |
|--------------|------------|-------------|-------------|------------|------------|------|
| gtccccaggta | actgttaaat | aaggcctcat | cctccactga | agagtatgga | ttgaaggatt | 3900 |
| gtgaactatg | ttagtgtga | tttgtgaactt | ggtgccataat | gttccatgtc | tgaagtttgc | 3960 |
| cccaactgtcta | cacgtggag | tatacctatg | tgtgtgcctt | gccactgaag | taagattttg | 4020 |
| cctgttatggt | actgttttgt | ttgttaataaa | agtgcactgc | cacccccaat | gcaaaaaaaa | 4080 |
| aaaaaaaaaa | | | | | | 4090 |

What is claimed is:

1. A method 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 agent.
2. The method of claim 1, wherein the breast cancer cells were previously administered a first chemotherapeutic more than two weeks prior to the combination of anti-cancer compounds.
3. The method of claim 2, wherein the breast cancer cells that were previously administered a first chemotherapeutic are chemo-resistant.
4. The method of claim 3, wherein the breast cancer cells are determined not to be chemo-sensitive or are determined to be chemo-resistant.
5. The method of claim 1, wherein the breast cancer cells are glucocorticoid receptor-positive (GR+).
6. The method of claim 1, wherein the breast cancer cells do not express detectable levels of estrogen receptor alpha by immunohistochemistry (IHC).
7. The method of claim 1, wherein the glucocorticoid receptor antagonist is administered prior to the chemotherapeutic.
8. The method of claim 1, wherein the glucocorticoid receptor antagonist is administered after the chemotherapeutic is administered.
9. The method of claim 1, wherein the glucocorticoid receptor antagonist is administered prior to and after administration of the chemotherapeutic.
10. The method of claim 1, wherein the glucocorticoid receptor antagonist has undetectable or a lower level of activity as a progesterone receptor antagonist.
11. The method of claim 10, wherein the glucocorticoid receptor antagonist does not have detectable progesterone receptor antagonist activity.

12. The method of claim 1, wherein the breast cancer is an unresectable breast cancer.

13. A method of treating chemotherapy-insensitive breast cancer cells comprising administering to a breast cancer patient an effective amount of a glucocorticoid receptor antagonist followed by chemotherapy.

14. The method of claim 13, wherein the glucocorticoid receptor antagonist is not specific for progesterone receptor.

15. The method of claim 13, wherein the patient was previously treated with chemotherapy.

16. A method of treating breast cancer cells comprising administering to a breast cancer patient an effective amount of a glucocorticoid receptor antagonist followed by at least one apoptosis-inducing agent.

17. The method of claim 16, wherein at least one apoptosis inducing agent is radiation, a chemotherapeutic, or an immunotherapy.

18. The method of claim 17, wherein the breast cancer cells are determined to be resistant to apoptosis.

19. A method 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.

20. The method of claim 19, wherein said glucocorticoid receptor antagonist is selected from beclometasone, betamethasone, budesonide, ciclesonide, flunisolide, fluticasone, mifepristone, mometasone, triamcinolone, CORT 0113083, or CORT 00112716.

* * * * *