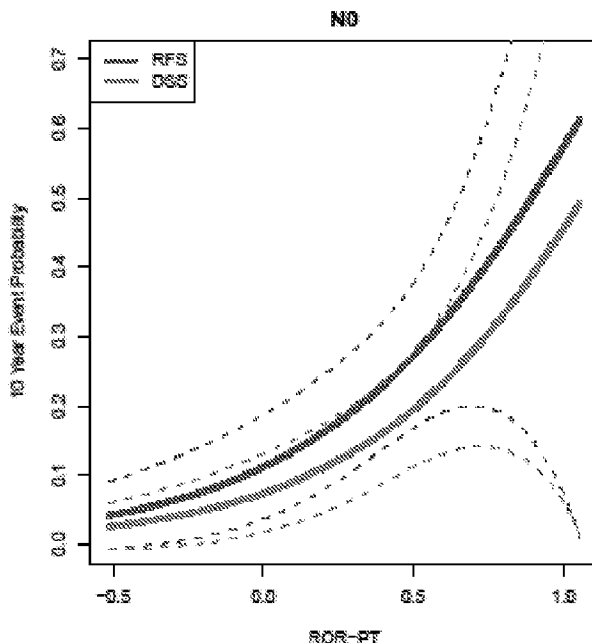




(86) Date de dépôt PCT/PCT Filing Date: 2013/05/22
(87) Date publication PCT/PCT Publication Date: 2013/11/28
(45) Date de délivrance/Issue Date: 2021/10/19
(85) Entrée phase nationale/National Entry: 2014/11/21
(86) N° demande PCT/PCT Application No.: US 2013/042157
(87) N° publication PCT/PCT Publication No.: 2013/177245
(30) Priorités/Priorities: 2012/05/22 (US61/650,209);
2013/01/17 (US61/753,673)

(51) Cl.Int./Int.Cl. *C40B 40/06* (2006.01),
C40B 30/04 (2006.01)
(72) Inventeurs/Inventors:
FERREE, SEAN M., US;
PARKER, JOEL S., US;
STORHOFF, JAMES JUSTIN, US;
PEROU, CHARLES M., US;
ELLIS, MATTHEW J., US;
...
(73) Propriétaires/Owners:
BRITISH COLUMBIA CANCER AGENCY BRANCH, CA;
NANOSTRING TECHNOLOGIES, INC., US;

(54) Titre : GENES NANO-46 ET PROCEDES DE PREDICTION DE L'EVOLUTION DU CANCER DU SEIN
(54) Title: NANO46 GENES AND METHODS TO PREDICT BREAST CANCER OUTCOME



(57) Abrégé/Abstract:

The present invention provides methods for classifying and for evaluating the prognosis of a subject having breast cancer are provided. The methods include prediction of breast cancer subtype using a supervised algorithm trained to stratify subjects on the

(72) **Inventeurs(suite)/Inventors(continued):** BERNARD, PHILIP S., US; NIELSEN, TORSTEN O., CA

(73) **Propriétaires(suite)/Owners(continued):** THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL, US;
UNIVERSITY OF UTAH RESEARCH FOUNDATION, US; WASHINGTON UNIVERSITY, US

(74) **Agent:** RIDOUT & MAYBEE LLP

(57) **Abrégé(suite)/Abstract(continued):**

basis of breast cancer intrinsic subtype. The prediction model is based on the gene expression profile of the intrinsic genes listed in Table 1. Further provided are compositions and methods for predicting outcome or response to therapy of a subject diagnosed with or suspected of having breast cancer. These methods are useful for guiding or determining treatment options for a subject afflicted with breast cancer. Methods of the invention further include means for evaluating gene expression profiles, including microarrays and quantitative polymerase chain reaction assays, as well as kits comprising reagents for practicing the methods of the invention.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
28 November 2013 (28.11.2013)



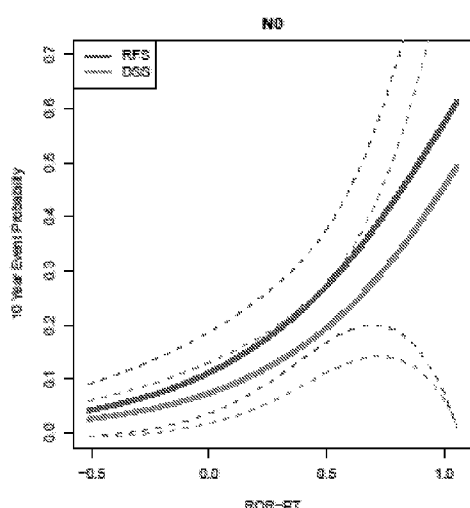
(10) International Publication Number
WO 2013/177245 A2

- (51) International Patent Classification:
C12Q 1/68 (2006.01)
- (21) International Application Number:
PCT/US2013/042157
- (22) International Filing Date:
22 May 2013 (22.05.2013)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
61/650,209 22 May 2012 (22.05.2012) US
61/753,673 17 January 2013 (17.01.2013) US
- (71) Applicant: NANOSTRING TECHNOLOGIES, INC.
[US/US]; 530 Fairview Avenue North, Suite 2000, Seattle
WA 98109 (US).
- (72) Inventors: FERREE, Sean M.; 7018 3rd Avenue NW,
Seattle, WA 98117 (US). PARKER, Joel S.; 1113
Churchwood Drive, Apex, NC 27502 (US). STORHOFF,
James Justin; 6520 36th Avenue NW, Seattle, WA 98117
(US).
- (74) Agents: ELRIFI, Ivor R. et al.; Mintz Levin Cohn Ferris
Glovsky and Popeo, P.C., One Financial Center, Boston,
MA 02111 (US).
- (81) Designated States (*unless otherwise indicated, for every
kind of national protection available*): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP,
KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD,
ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI,
NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU,
RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ,
TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA,
ZM, ZW.
- (84) Designated States (*unless otherwise indicated, for every
kind of regional protection available*): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ,
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: NANO46 GENES AND METHODS TO PREDICT BREAST CANCER OUTCOME

Figure 4



(57) Abstract: The present invention provides methods for classifying and for evaluating the prognosis of a subject having breast cancer are provided. The methods include prediction of breast cancer subtype using a supervised algorithm trained to stratify subjects on the basis of breast cancer intrinsic subtype. The prediction model is based on the gene expression profile of the intrinsic genes listed in Table 1. Further provided are compositions and methods for predicting outcome or response to therapy of a subject diagnosed with or suspected of having breast cancer. These methods are useful for guiding or determining treatment options for a subject afflicted with breast cancer. Methods of the invention further include means for evaluating gene expression profiles, including microarrays and quantitative polymerase chain reaction assays, as well as kits comprising reagents for practicing the methods of the invention.

WO 2013/177245 A2

WO 2013/177245 A2



Published:

- *without international search report and to be republished
upon receipt of that report (Rule 48.2(g))*

NANO46 GENES AND METHODS TO PREDICT BREAST CANCER OUTCOME

CROSS-REFERENCE TO RELATED APPLICATIONS

[01] This application claims the benefit of U.S. Provisional Application No. 61/650,209, filed May 22, 2012 and U.S. Provisional Application No. 61/753,673, filed January 17, 2013.

FIELD OF THE INVENTION

[02] This disclosure relates generally to the field of cancer biology, and specifically, to the fields of detection and identification of specific cancer cell phenotypes and correlation with appropriate therapies.

BACKGROUND OF THE INVENTION

[03] Current approaches to treating early breast cancer, including adjuvant therapy, have indeed improved survival and reduced recurrence. However, the risk of recurrence may be underestimated in some patients, but overestimated in others.

[04] While the risk of recurrence does diminish somewhat over time, ongoing risk has been observed in many studies, some of them involving tens of thousands of patients with breast cancer. In fact, some of the patients who experienced recurrence after five years in these studies had previously been considered "low risk" - for example, their cancer had not spread to the lymph nodes at the time of their initial diagnosis, or their estrogen receptor status was positive. In one of these studies, a substantial number of recurrences occurred more than five years post-treatment. Thus, there is a need in the art to determine risk of recurrence and determine therapies which reduce that risk and improve overall survival.

SUMMARY OF THE INVENTION

[05] The present invention provides a method of predicting outcome in a subject having breast cancer comprising: providing a tumor sample from the subject; determining the expression of the genes in the NANO46 intrinsic gene list of Table 1 in the tumor sample; measuring the similarity of the tumor sample to an intrinsic subtype based on the expression of the genes in the NANO46 intrinsic gene list, wherein the intrinsic subtype consists of at least Basal-like, Luminal A, Luminal B or HER2-enriched; determining a proliferation score based on the expression of a subset of proliferation genes in the NANO46 intrinsic gene list;

determining the size of the tumor, calculating a risk of recurrence score using a weighted sum of said intrinsic subtype, proliferation score and tumor size; and determining whether the subject has a low or high risk of recurrence based on the recurrence score. In one embodiment a low score indicates a more favorable outcome and high score indicates a less favorable outcome.

[06] The methods of the present invention can include determining the expression of at least one of, a combination of, or each of, the NANO46 intrinsic genes recited in Table 1. In some embodiments, the methods of the present invention can include determining the expression of at least one of, a combination of, or each of, the NANO46 intrinsic genes selected from ANLN, CCNE1, CDC20, CDC6, CDCA1, CENPF, CEP55, EXO1, KIF2C, KNTC2, MELK, MKI67, ORC6L, PTTG1, RRM2, TYMS, UBE2C and/or UBE2T. The expression of the members of the NANO46 intrinsic gene list can be determined using the nanoreporter code system (nCounter® Analysis system).

[07] The methods of the present invention can include determining at least one of, a combination of, or each of, the following: tumor size, tumor grade, nodal status, intrinsic subtype, estrogen receptor expression, progesterone receptor expression, and HER2/ERBB2 expression

[08] The sample can be a sampling of cells or tissues. The sample can be a tumor. The tissue can be obtained from a biopsy. The sample can be a sampling of bodily fluids. The bodily fluid can be blood, lymph, urine, saliva or nipple aspirate.

[09] While the disclosure has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the disclosure, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

[11] While this disclosure has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the

scope of the disclosure encompassed by the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

- [12] Figure 1 is a heatmap of the breast cancer intrinsic subtypes and the intrinsic genes of Table 1.
- [13] Figure 2 shows a Kaplan Meier survival curves from a cohort of untreated breast cancer patients.
- [14] Figure 3 shows a Kaplan Meier survival curves from a cohort of node-negative, ER+ Breast Cancer Patients treated with tamoxifen.
- [15] Figure 4 shows a 10 Year event probability as a function of ROR Score in ER+, Node-negative breast cancer patients treated with tamoxifen. The graph shows the sub-population subtyped as Luminal A or B within this population. RFS = Recurrence-free survival; DSS = disease-specific survival
- [16] Figure 5 is a schematic of the breast cancer intrinsic subtyping assay.
- [17] Figure 6 is a schematic of the algorithm process.
- [18] Figure 7 is an illustration showing the hybridization of the CodeSet to mRNA.
- [19] Figure 8 is an illustration showing the removal of excess reporters.
- [20] Figure 9 is an illustration showing the binding of the reporters to the surface of a cartridge.
- [21] Figure 10 is an illustration showing the immobilization and alignment of a reporter.
- [22] Figure 11 is an illustration of data collection.
- [23] Figure 12 is an illustration of the nCounter analysis system breast cancer test assay process.
- [24] Figure 13 is an illustration of the nCounter Prep Station.
- [25] Figure 14 is an illustration of nCounter Digital Analyzer.

DETAILED DESCRIPTION OF THE INVENTION

- [26] The disclosure presents a method of predicting outcome in a subject having breast cancer comprising: providing a tumor sample from the subject; determining the expression of the genes in the NANO46 intrinsic gene list of Table 1 in the tumor sample; determining the intrinsic subtype of the tumor sample based on the expression of the genes in the NANO46 intrinsic gene list, wherein the intrinsic subtype consists of at least Basal-like, Luminal A, Luminal B or HER2-enriched; determining a proliferation score based on the expression of a subset of proliferation genes in the NANO46 intrinsic gene list; determining the size of the

tumor, calculating a risk of recurrence score using a weighted sum of said intrinsic subtype, proliferation score and tumor size; and determining whether the subject has a low or high risk of recurrence based on the recurrence score. In one embodiment a low score indicates a more favorable outcome and high score indicates a less favorable outcome.

[27] Intrinsic genes are statistically selected to have low variation in expression between biological sample replicates from the same individual and high variation in expression across samples from different individuals. Thus, intrinsic genes are used as classifier genes for breast cancer classification. Although clinical information was not used to derive the breast cancer intrinsic subtypes, this classification has proved to have prognostic significance. Intrinsic gene screening can be used to classify breast cancers into five molecular distinct intrinsic subtypes, Luminal A (LumA), Luminal B (LumB), HER2-enriched (Her-2-E), Basal-like, and Normal-like (Perou *et al.* Nature, 406(6797):747-52 (2000); Sorlie *et al.* PNAS, 98(19):10869-74 (2001)).

[28] A NANO46 gene expression assay, as described herein, can identify intrinsic subtype from a biological sample, e.g., a standard formalin fixed paraffin embedded tumor tissue. The methods utilize a supervised algorithm to classify subject samples according to breast cancer intrinsic subtype. This algorithm, referred to herein as the NANO46 classification model, is based on the gene expression profile of a defined subset of intrinsic genes that has been identified herein as superior for classifying breast cancer intrinsic subtypes. The subset of genes, along with primers target-specific sequences utilized for their detection, is provided in Table 1. Table 1A provides the sequences of target specific probe sequences for detecting each gene utilized in Table 1. The sequences provided in Table 1A are merely representative and are not meant to limit the invention. The skilled artisan can utilize any target sequence-specific probe for detecting any of (or each of) the genes in Table 1.

[29] Table 1

GENE	REPRESENTATIVE GENBANK ACCESSION NUMBER	FORWARD PRIMER	SEQ ID NO:	REVERSE PRIMER	SEQ ID NO:
ACTR3B	NM_020445 NM_001040135	AAAGATTCCCTGGG ACCTGA	1	TGGGGCAGTTCTGTA TTACTTC	47
ANLN	NM_018685	ACAGCCACTTTCA GAAGCAAG	2	CGATGGTTTGTGACA AGATTTCTC	48
BAG1	NM_004323	CTGGAAGAGTTGA ATAAAGAGC	3	GCAAATCCTTGGGC AGA	49
BCL2	NM_000633	TACCTGAACCGGC ACCTG	4	GCCGTACAGTTCCAC AAAGG	50
BLVRA	BX647539	GCTGGCTGAGCAG AAAG	5	TTCCTCCATCAAGAG TTCAACA	51
CCNE1	BC035498	GGCCAAAATCGAC AGGAC	6	GGGTCTGCACAGAC TGCAT	52
CDC20	BG256659	CTGTCTGAGTGCC GTGGAT	7	TCCTTGTAATGGGA GACCA	53
CDC6	NM_001254	GTAAATCACCTTC TGAGCCT	8	ACTTGGGATATGTGA ATAAGACC	54
CDCA1	NM_031423	GGAGGCGGAAGA AACCAG	9	GGGGAAAGACAAAAG TTTCCA	55
CDH3	BC041846	GACAAGGAGAAT CAAAAGATCAGC	10	ACTGTCTGGGTCCAT GGCTA	56
CENPF	NM_016343	GTGGCAGCAGATC ACAA	11	GGATTCGTGGTGGG TTC	57
CEP55	AB091343	CCTCACGAATTGC TGAACCT	12	CCACAGTCTGTGATA AACGG	58
CXXC5	BC006428	CATGAAATAGTGC ATAGTTTGCC	13	CCATCAACATTCTCT TTATGAACG	59
EGFR	NM_005228	ACACAGAATCTAT ACCCACCAGAGT	14	ATCAACTCCCAAAC GGTCAC	60
ERBB2	NM_001005862	GCTGGCTCTCACA CTGATAG	15	GCCCTTACACATCGG AGAAC	61
ESR1	NM_001122742	GCAGGGAGAGGA GTTTGT	16	GACTTCAGGGTGCTG GAC	62
EXO1	NM_130398	CCCATCCATGTGA GGAAGTATAA	17	TGTGAAGCCAGCAA TATGTATC	63
FGFR4	AB209631	CTTCTTGGACCTT GGCG	18	TATTGGGAGGCAGG AGGTTTA	64
FOXA1	NM_004496	GCTACTACGCAGA CACG	19	CTGAGTTCATGTTGC TGACC	65
FOXC1	NM_001453	GATGTTTCGAGTCA CAGAGG	20	GACAGCTACTATTCC CGTT	66
GPR160	AJ249248	TTCGGCTGGAAGG AACC	21	TATGTGAGTAAGCTC GGAGAC	67
HSPC150 (UBE2T)	NM_014176	GGAGATCCGTCAA CTCCAAA	22	AGTGGACATGCGAG TGGAG	68
KIF2C	NM_006845	TGGGTCGTGTCAG GAAAC	23	CACCGCTGGAAACT GAAC	69

KNTC2	NM_006101	CGCAGTCATCCAG AGATGTG	24	CGTGCACATCCATGA CCTT	70
KRT14	BC042437	ACTCAGTACAAGA AAGAACCG	25	GAGGAGATGACCTT GCC	71
KRT17	AK095281	GTTGGACCAGTCA ACATCTCTG	26	GCCATAGCCACTGCC ACT	72
KRT5	M21389	TGTGGCTCATTAG GCAAC	27	CTTCGACTGGACTCT GT	73
MAPT	NM_001123066	GACTCCAAGCGCG AAAAC	28	CAGACATGTTGGTAT TGCACATT	74
MDM2	M92424	CCACAAAATATTC ATGGTCTTG	29	AGGCGATCCTGGGA AATTAT	75
MELK	NM_014791	CCAGTAGCATTGT CCGAG	30	CCCATTGTCTGTCT TCAC	76
MIA	BG765502	GTCTCTGGTAATG CACACT	31	CTGATGGTTGAGGCT GTT	77
MKI67	NM_002417	GTGGAATGCCTGC TGACC	32	CGCACTCCAGCACCT AGAC	78
MLPH	NM_024101	AGGGGTGCCCTCT GAGAT	33	TCACAGGGTCAAAC TTCCAGT	79
MMP11	NM_005940	CGAGATCGCCAAG ATGTT	34	GATGGTAGAGTTCC AGTGATT	80
MYC	NM_002467	AGCCTCGAACAAT TGAAGA	35	ACACAGATGATGGA GATGTC	81
NAT1	BC013732	ATCGACTGTGTAA ACAAC TAGAGAA GA	36	AGTAGCTACATCTCC AGGTCTCTG	82
ORC6L	NM_014321	TTTAAGAGGGCAA ATGGAAGG	37	CGGATTTTATCAACG ATGCAG	83
PGR	NM_000926	TGCCGCAAGACTC ACTTG	38	CATTGCCGTCCCTC ATCG	84
PHGDH	AK093306	CCTCAGATGATGC CTATCCA	39	GCAGGTCAAACTC TCAAAG	85
PTTG1	BE904476	CAGCAAGCGATGG CATAGT	40	AGCGGGCTTCTGTAA TCTGA	86
RRM2	AK123010	AATGCCACCGAAG CCTC	41	GCCTCAGATTTCAAC TCGT	87
SFRP1	BC036503	TCGAACTGAAGGC TATTTACGAG	42	CTGCTGAGAATCAA AGTGGGA	88
SLC39A6	NM_012319	GTCGAAGCCGCAA TTAGG	43	GGAACAAACTGCTC TGCCA	89
TMEM45B	AK098106	CAAACGTGTGTTT TGGAGG	44	ACAGCTCTTTAGCAT TTGTGGA	90
TYMS	BQ56428	TGCCCTGTATGAT GTCAGGA	45	GGGACTATCAATGTT GGGTTCTC	91
UBE2C	BC032677	GTGAGGGGTGTCA GCTCAGT	46	CACACAGTTCAGTGC TCCACA	92

Table 1a. Probes for detecting NANO46 genes

Gene Name	RefSeq Accession	Target Sequence	SEQ ID NO:
ACTR3B	NM_001040135.1	CCAGAAGAAGTTTGTATAGACGTTGGTTACG AAAGATTCCCTGGGACCTGAAATATTCTTTCAC CCGGAGTTTGCCAACCCAGACTTTATGGAGTCCATC	140
ANLN	NM_018685.2	CGTGCCAGGCGAGAGAATCTTCAGAGAAAAA TGGCTGAGAGGCCACAGCAGCTCCAAGGTC TATGACTCATGCTAAGCGAGCTAGACAGCCA CTTTCAG	141
BAG1	NM_004323.3	CTTCATGTTACCTCCCAGCAGGGCAGCAGTGA ACCAGTTGTCCAAGACCTGGCCCAGGTTGTTG AAGAGGTCATAGGGGTTCCACAGTCTTTTCAG AAAC	142
BCL2	NM_000633.2	CCAAGCACCGCTTCGTGTGGCTCCACCTGGAT GTTCTGTGCCTGTAAACATAGATTCTGCTTTCC ATGTTGTTGGCCGATCACCATCTGAAGAGCA GACG	143
BLVRA	NM_000712.3	TTCCTGAAAAAAGAAGTGGTGGGGAAGACC TGCTGAAAGGGTCGCTCCTCTTCACAGCTGGC CCGTTGGAAGAAGAGCGGTTTGGCTTCCCTGC ATTCA	144
CCNE1	NM_001238.1	GAGAACTGTGTCAAGTGGATGGTTCCATTTC CATGGTTATAAAGGAGACGGGGAGCTCAAAA CTGAAGCACTTCAGGGGCGTCGCTGATGAAG ATGCAC	145
CDC20	NM_001255.1	CCCGAGTGGGCTCCCTAAGCTGGAACAGCTA TATCCTGTCCAGTGGTTCACGTTCCTGGCCACA TCCACCACCATGATGTTTCGGGTAGCAGAACA CCATGT	146
CDC6	NM_001254.3	GGGGAAGTTATATGAAGCCTACAGTAAAGTC TGTCGCAAAACAGCAGGTGGCGGCTGTGGACC AGTCAGAGTGTTTGTCACTTTCAGGGCTCTTG GAAGCC	147
CDCA1	NM_145697.1	GCCTGGCGGTGTTTTCGTCTGTGCTCAGCGGTG GGAGGAGGCGGAAGAAACCAGAGCCTGGGA GATTAACAGGAAACTTCCAAGATGGAAACTT TGTCTTT	148
CDH3	NM_001793.3	CCCTCGACCGTGAGGATGAGCAGTTTGTGAG GAACAACATCTATGAAGTCATGGTCTTGGCCA TGGACAATGGAAGCCCTCCCACTGGCAC GGGAAC	149
CENPF	NM_016343.3	AGAAAATCTTGCAGAGTCCTCCAAACCAACA GCTGGTGGCAGCAGATCACAAAAGGTCAAAG TTGCTCAGCGAGCCCAGTAGATTCAAGGCAC CATCCTC	150
CEP55	NM_018131.3	GTACTACCGCATTGCTTGAACAGCTGGAAGA GACAACGAGAGAAGGAGAAAGGAGGGAGCA GGTGTGAAAGCCTTATCTGAAGAGAAAGAC GTATTGAA	151
CXXC5	NM_016463.5	AGCTGCCCTCTCCGTGCAATGTCAGTCTCGT GTGGTCTCCAGCAAGGGATTCTGGGCGAAGAC AAACGGATGCACCCGTCTTTAGAACCAAAAA TATTCT	152
EGFR	NM_005228.3	GCAGCCAGGAACGTACTGGTGAAACACCGC AGCATGTCAAGATCACAGATTTTGGGCTGGCC AAACTGCTGGGTGCGGAAGAGAAAGAATACC	153

		ATGCAG	
ERBB2	NM_004448.2	TGAAGGTGCTTGGATCTGGCGCTTTTGGCACA GTCTACAAGGGCATCTGGATCCCTGATGGGG AGAAATGTGAAAATTCCAGTGGCCATCAAAGT GTTGAG	154
ESR1	NM_000125.2	AGGAACCAGGGAAAAATGTGTAGAGGGCATGG TGGAGATCTTCGACATGCTGCTGGCTACATCA TCTCGGTTCCGCATGATGAATCTGCAGGGAGA GGAGT	155
EXO1	NM_006027.3	TGGCCACAAAGTAATTAAAGCTGCCCCGGTCT CAGGGGGTAGATTGCCTCGTGGCTCCCTATGA AGCTGATGCGCAGTTGGCCTATCTTAACAAAG CGGG	156
FGFR4	NM_002011.3	CCCACATCCAGTGGCTGAAGCACATCGTCATC AACGGCAGCAGCTTCGGAGCCGACGGTTTCC CCTATGTGCAAGTCCTAAAGACTGCAGACATC AATAG	157
FOXA1	NM_004496.2	TGGATGGTTGTATTGGGCAGGGTGGCTCCAG GATGTTAGGAACTGTGAAGATGGAAGGGCAT GAAACCAGCGACTGGAACAGCTACTACGCAG ACACGCA	158
FOXC1	NM_001453.1	TTCGAGTCACAGAGGATCGGCTTGAACAAC CTCCAGTGAACGGGAATAGTAGCTGTCAAAT GGCCTTCCCTTCCAGCCAGTCTCTGTACCGCA CGTCCG	159
GPR160	NM_014373.1	GGATTTCAGTCCTTGCTTATGTTTTGGGAGAC CCAGCCATCTACCAAAGCCTGAAGGCACAGA ATGCTTATTCTCGTCACTGTCCTTTCTATGTCA GCAT	160
UBE2T	NM_014176.1	GTGTCAGTCACTGCATCCCAGGCAGCTCTTA GTGTGGAGCAGTGAAGTGTGTGTGGTTCCTTC TACTTGGGGATCATGCAGAGAGCTTCACGTCT GAAG	161
KIF2C	NM_006845.2	GTTGTCTACAGGTTACAGCAAGGCCACTGGT ACAGACAATCTTTGAAGGTGGAAGCAACT TGTTTTGCATATGGCCAGACAGGAAGTGGCA AGACAC	162
KNTC2	NM_006101.1	AAAAGGTCATAAGCATGAAGCGCAGTTCAGT TTCCAGCGGTGGTGCTGGCCGCCTCTCCATGC AGGAGTTAAGATCCCAGGATGTAAATAAACA AGGCCT	163
KRT14	NM_000526.3	GCAGTCATCCAGAGATGTGACCTCCTCCAGCC GCCAAATCCGCACCAAGGTCATGGATGTGCA CGATGGCAAGGTGGTGTCCACCCACGAGCAG GTCCTT	164
KRT17	NM_000422.1	CTGACTCAGTACAAGAAAGAACC GG TGACCA CCCGTCAGGTGCGTACCATTGTGGAAGAGGT CCAGGATGGCAAGGTCATCTCCTCCCGCGAG CAGGTCC	165
KRT5	NM_000424.2	CTGGTTCTCTTGCTCCACCAGGAACAAGCCAC CATGTCTCGCCAGTCAAGTGTGTCTTCCGGA GCGGGGGCAGTCGTAGCTTCAGCACCGCCTCT GCCA	166
MAPT	NM_016835.3	GCCGGGTCCCCTCAACTCAAAGCTCGCATGGTC AGTAAAAGCAAAGACGGGACTGGAAGCGATG ACAAAAAAGCCAAGACATCCACACGTTCCCTC TGCTAA	167

MDM2	NM_006878.2	GGTGAGGAGCAGGCAAATGTGCAATACCAAC ATGTCTGTACCTACTGATGGTGTCTGTAACCAC CTCACAGATTCCAGCTTCGGAACAAGAGACC CTGGTT	168
MELK	NM_014791.2	AGAGACAGCCAACAAAAATATTCATGGTTCTT GAGTACTGCCCTGGAGGAGAGCTGTTTGACT ATATAATTTCCCAGGATCGCCTGTCAGAAGAG GAGACC	169
MIA	NM_006533.1	CCGGGGCCAAGTGGTGTATGTCTTCTCCAAGC TGAAGGGCCGTGGGCGGCTCTTCTGGGGAGG CAGCGTTCAGGGAGATTACTATGGAGATCTG GCTGCT	170
MKI67	NM_002417.2	GCTTCCAGCAGCAAATCTCAGACAGAGGTTT CTAAGAGAGGAGGAGAAAGAGTGGCAACCTG CCTTCAAAAAGAGAGTGTCTATCAGCCGAAGT CAACATG	171
MLPH	NM_024101.4	GAGGAAGTCAAACCTCCCGATATTTCTCCCTC GAGTGGCTGGGAAACTTGGCAAGAGACCAGA GGACCCAAATGCAGACCCTTCAAGTGAGGCC AAGGCA	172
MMP11	NM_005940.3	AGCAGCCAAGGCCCTGATGTCCGCTTCTACA CCTTTTCGTACCCACTGAGTCTCAGCCCAGAT GACTGCAGGGGCGTTCAACACCTATATGGCC AGCCC	173
MYC	NM_002467.3	CACCGAGGAGAATGTCAAGAGGCGAACACAC AACGTCTTGAGCGCCAGAGGAGGAACGAGC TAAAACGAGAGCTTTTTTGCCCTGCGTGACCAG ATCCCCG	174
NAT1	NM_000662.4	AGCACTTCCTCATAGACCTTGGATGTGGGAGG ATTGCATTTCAGTCTAGTTCCTGGTTGCCGGCT GAAATAACCTGAATTCAAGCCAGGAAGAAGC AGCAA	175
ORC6L	NM_014321.2	GACTGTGTAAACAACCTAGAGAAGATTGGACA GCAGGTCGACAGAGAACCTGGAGATGTAGCT ACTCCACCACGGAAGAGAAAAGAAGATAGTGG TTGAAGC	176
PGR	NM_000926.2	GGGATGAAGCATCAGGCTGTCTATTATGGTGT CTTACCTGTGGGAGCTGTAAGGTCTTCTTTAA GAGGGCAATGGAAGGGCAGCACAACTACTTA TGTGC	177
PHGDH	NM_006623.2	GCGACGGCTTCGATGAAGGACGGCAAATGGG AGCGGAAGAAGTTCATGGGAACAGAGCTGAA TGGAAGAGCCCTGGGAATTCTTGGCCTGGGC AGGATTG	178
PTTG1	NM_004219.2	CACCAGCCTTACCTAAAGCTACTAGAAAGGC TTTGGGAACGTGTCAACAGAGCTACAGAAAAG TCTGTAAAGACCAAGGGACCCCTCAAACAAA AACAGCC	179
RRM2	NM_001034.1	TTCTTTTGGACCGCCGAGGAGGTTGACCTCT CCAAGGACATTCAGCACTGGGAATCCCTGAA ACCCGAGGAGAGATATTTTATATCCCATGTTT TGGCT	180
SFRP1	NM_003012.3	GTGGGTACACACACGCACTGCGCCTGTCTAGT AGTGGACATTGTAATCCAGTCGGCTTGTTCCT GCAGCATTCCTCGCTCCCTTCCCTCCATAGCCA CGCT	181
SLC39A6	NM_012319.2	GATCGAACTGAAGGCTATTTACGAGCAGACT CACAAGAGCCCTCCCACTTTGATTCTCAGCAG CCTGCAGTCTTGGAAGAAGAAGAGGTCATGA TAGCTC	182

TMEM45B	NM_138788.3	CTGGCTGCCCTCAGCATTGTGGCCGTCAACTA TTCTCTTGTTTACTGCCTTTTGACTCGGATGAA GAGACACGGAAGGGGAGAAATCATTGGAATT CAGA	183
TYMS	NM_001071.1	TGCTAAAGAGCTGTCTTCCAAGGGAGTGA ATCTGGGATGCCAATGGATCCCGAGACTTTTT GGACAGCCTGGGATTCTCCACCAGAGAAGAA GGGGAC	184
UBE2C	NM_007019.2	GTCTGCCCTGTATGATGTCAGGACCATTCTGC TCTCCATCCAGAGCCTTCTAGGAGAACCCAAC ATTGATAGTCCCTTGAACACACATGCTGCCGA GCTC	185

[30] Table 2 provides select sequences for the NANO46 genes of Table 1.

[31] Table 2

GENBANK ACCESSION NUMBER	SEQUENCE	SEQ ID NO:
NM_020445	CAGCGGCGCTGCGGCGGCTCGCGGGAGACGCTGCGCGCGGGGCTAGCGGGCGGCGGAGCGGACGGCGACG GGGCGCTCTCGGGGTGCCGGCGGGGCCAGCGCCGCGCGTCCCGAGCATGGCAGGCTCCCTGCCTCCCTG CGTGGTGGACTGTGGCACC GGGTATACCAAGCTTGGCTACGCAGGCAACACTGAGCCCCAGTTTCATTATT CCTTCATGTATTGCCATCAGAGAGTCAGCAAAGGTAGTTGACCAAGCTCAAAGGAGAGTGTGTAGGGGAG TTGATGACCTTGACTTTTTCATAGGAGATGAAGCCATCGATAAACCTACATATGCTACAAAGTGGCCGAT ACGACATGGAATCATTGAAGACTGGGATCTTATGGAAAGGTTTCATGGAGCAAGTGGTTTTTAAATATCTT CGAGCTGAACCTGAGGACCATTATTTTTTAAATGACAGAACCTCCACTCAATACACCAGAAAACAGAGAGT ATCTTGCAGAAATATGTTTGAATCATTTAACGTACCAGGACTCTACATTGCAGTTCAGGCAGTGTCTGGC CTTGGCGGCATCTTGGACATCTCGACAAGTGGGTGAACGTACGTTAACGGGGATAGTCATTGACAGCGGA GATGGAGTCACCCATGTTATCCCACTGGCAGAAGGTTATGTAATTGGAAGCTGCATCAAACACATCCCCGA TTGCAGGTAGAGATATTACGTATTTTCATTCAACAGCTGCTAAGGGAGAGGGAGGTGGGAATCCCTCCTGA GCAGTCACTGGAGACCGCAAAAGCCATTAAAGGAGAAATACTGTTACATTTGCCCCGATATAGTCAAGGAA TTTGCCAAGTATGATGTGGATCCCCGGAAGTGGATCAAACAGTACACGGGTATCAATGCGATCAACCGA AGAAAGTTTTGTTATAGACGTTGGTTACGAAAGATTCTCTGGGACCTGAAATATTCITTCACCCGAGTTTGC CAACCCAGACTTTTATGGAGTCCATCTCAGATGTTGTTGATGAAGTAATACAGAACTGCCCATCGATGTG CGGCGCCCCGCTGTATAAGAAATGTCGTACTCTCAGGAGGCTCCACCATGTTTCAGGGATTTTCGACGCGGAC TGCAAGGGATTTGAAGAGAGTGTGGATGCTAGGCTCAGCGAGGAGCTCAGCGCGCGGAGGAGT CAAGCCGAAGCCTGTGGAGTCCAGGTGGTCAACGATCAGATGCAGCGCTACGCGCTGTGTTTCGGAGGC TCCATGCTGGCCTCGACTCCCGAGTTCTTTAGGTTCTGCCACACCAAGAAGGACTATGAAGAGTACGGGC CCAGCATCTGCCGCCACAAACCCGCTCTTTGGAGTCATGTCCTAGTGTCTGCCTGAACGCGTCTGTTCCGATG GTGTCACGTTGGGGAACAAGTGTCTCTCAGAACCCAGAGAAGGCGCGCTTCTGTAATAGCGAGTCGG TGTTGCTGCCCCAGCAGCGTGTGCTTGCATTGCGGTCATGAGGCGCGGCGCGGGCCCTTCAGTAAAAAGCCA TTTATCCGTGTGCGGACCGCTGTCTGCCAGCCTCCTCTCTCCGCCCCCTCACCCTCGCTCTCCCTC CTCCTCCTCCTCCGAGCTGCTAGCTGACAAATACAATTCGAAGGAATCCAAATGTGACTTTGAAAATTG TTAGAGAAAACAACATTAGAAAAATGGCGCAAAATCGTTAGGTCCCAGGAGAGAATGTGGGGCGCAAAAC CTTTTCTCCAGCCTATTTTTGTAAATAAAATGTTTAAACTTGAAATACAAATCGATGTTTATATTTCC TATCATTTTGTATTTTATGTTATTTGGTACAACTGGCTGATACTAAGCACGAATAGATATTGATGTTATG GAGTGTGTAATCCAAAGTTTTTAAATTGTGAGGCATGTTCTGATATGTTTATAGGCAAAACAAATAAAACA GCAAACTTTTTTGCCACATGTTTGTAGAAAATGATTATACTTTTATGGAGTGACATGAAGTTTGAACAC TAAACAGTAATGTATGAGAATTACTACAGATACATGTATCTTTTACTTTTTTTGTTTGAACCTTTCTGGA GCTGTTTTATAGAAGATGATGTTTGTGTTGCGGTGAGTGTGGAATGAAATACTTCTTGCACCATTGTAA TAAAGCTGTTAGAATATTGTAATAATC	93
NM_00104013 5	CAGCGGCGCTGCGGCGGCTCGCGGGAGACGCTGCGCGCGGGGCTAGCGGGCGGCGGAGCGGACGGCGACG GGGCGCTCTCGGGGTGCCGGCGGGGCCAGCGCCGCGCGTCCCGAGCATGGCAGGCTCCCTGCCTCCCTG CGTGGTGGACTGTGGCACC GGGTATACCAAGCTTGGCTACGCAGGCAACACTGAGCCCCAGTTTCATTATT CCTTCATGTATTGCCATCAGAGAGTCAGCAAAGGTAGTTGACCAAGCTCAAAGGAGAGTGTGTAGGGGAG TTGATGACCTTGACTTTTTCATAGGAGATGAAGCCATCGATAAACCTACATATGCTACAAAGTGGCCGAT ACGACATGGAATCATTGAAGACTGGGATCTTATGGAAAGGTTTCATGGAGCAAGTGGTTTTTAAATATCTT CGAGCTGAACCTGAGGACCATTATTTTTTAAATGACAGAACCTCCACTCAATACACCAGAAAACAGAGAGT ATCTTGCAGAAATATGTTTGAATCATTTAACGTACCAGGACTCTACATTGCAGTTCAGGCAGTGTCTGGC CTTGGCGGCATCTTGGACATCTCGACAAGTGGGTGAACGTACGTTAACGGGGATAGTCATTGACAGCGGA GATGGAGTCACCCATGTTATCCCACTGGCAGAAGGTTATGTAATTGGAAGCTGCATCAAACACATCCCCGA TTGCAGGTAGAGATATTACGTATTTTCATTCAACAGCTGCTAAGGGAGAGGGAGGTGGGAATCCCTCCTGA GCAGTCACTGGAGACCGCAAAAGCCATTAAAGGAGAAATACTGTTACATTTGCCCCGATATAGTCAAGGAA	94

	TTTGCCAAGTATGATGTGGATCCCCGGAAGTGGATCAAACAGTACACGGGTATCAATGCGATCAACCAGA AGAAGTTTGTATAGACGTTGGTTACGAAAGATTCTGGGACCTGAAATATCTTTACCCGGAGTTTGC CAACCCAGACTTTATGGAGTCCATCTCAGATGTTGTTGATGAAGTAATACAGAACTGCCCATCGATGTG CGGGCCCCGTGTATAAGCCCGAGTTCTTTAGGTCTGCCACCAAGAAGGACTATGAAGAGTACGGGC CCAGCATCTGCCGCCACAACCCCGTCTTTGGAGTCATGTCTTAGTGTCTGCCTGAACGCGTCGTTCCGATG GTGTACGTTGGGGAACAAGTGTCTTCAGAACCCAGAGAAGGCCGCCGTTCTGTAAATAGCGACGTGG TGTGCTGCCAGCAGCGTGTTCGATTGCCGGTGCATGAGGCGCGGGCGGGGCCCTTCAGTAAAGCCCA TTTATCCGTGTGCCGACCGCTGTCTGCCAGCCTCCTCTCTCCGCCCCCTCACCCCTCGCTCTCCCTC CTCCTCCTCCTCCGAGCTGTAGCTGACAAATACAATTCTGAAGGAATCCAAATGTGACTTTGAAAAATTG TTAGAGAAAACAACATTAGAAAATGGCGCAAAATCGTTAGGTCCCAGGAGAGAATGTGGGGCGCAAACC CTTTTCTCCAGCCTATTTTTGTAAATAAAATGTTAAACTTGAAATACAAATCGATGTTTATATTTCC TAICATTTTGTATTTATGGTATTTGGTACAACCTGGCTGATACTAAGCACGAATAGATATTGATGTTATG GAGTCTGTAAATCCAAAGTTTAAATTTGAGGCATGTTCTGATATGTTTATAGGCAACCAATAAACA GCAAACTTTTTTGCCACATGTTTGTAGAAAATGATTATACTTTATGGAGTGACATGAAGTTTGAACAC TAAACAGTAATGTATGAGAATTACTACAGATACATGTATCTTTTAGTTTTTTTGTGTTGAACTTTCTGGA GCTGTTTTATAGAAGATGATGGTTTGTTCGGTGAGTGTGGAAGAAATACTTCTTGCACCATTGTAA TAAAGCTGTAGAAATATTGTAATAATC	
NM_018685	CTCGGCGCTGAAATTCAAATTTGAACGGCTGCAGAGGCCGAGTCCGTCAGTGAAGCCGAGAGGAGAGGA CAGCTGGTTGTGGGAGAGTTCCCCCGCCCTCAGACTCCTGGTTTTTTTCCAGGAGACACACTGAGCTGAGAC TCACTTTTCTCTTCTGAAATTTGAACACCGGTTTCCATCTGCTCTGATGCTCGACGCGCTGGGGCGATGGAT CCGTTTACGGAGAACTGCTGGAGCGAACCCGTGCCAGGCGAGAGAATCTTCAGAGAAAATGCGTGAAGA GGCCACAGCAGCTCCAAGGTCTATGACTCATGTAAAGCAGCTAGACAGCCACTTTCAGAAAGCAAGTAA CCAGCAGCCCTCTCTGGTGGTGAAGAGAAATCTTGTACAAAACCATCGCCATCAAAAAACGCTGTTCT GACAACACTGAAGTAGAAGTTTCTAACTTGGAAAAATAACAACCAAGTTGAGTCGACATCTGCAAAATCTT GTTCTCCAAGTCTGTGTCTCTCAGGTGCGAGCCACAAGCAGCAGATACCATCAGTGTCTGTTGCTGT CCCCGATCACTGCTGGGATGAGGAGAGGGCTGAAGTCAAGATTGGAAGCACTGCAGCCTCTCAGTT AAAACACGTATGCAAAAACCTTGCAGAGCAACGGCGCCGTTGGGATAATGATGATATGACAGATGACATTC CTGAAAGCTCACTCTTCTCACCATGCCATCAGAGGAAAAGGCTGCTTCCCTCCAGACCTCTGCTTTC AAATGCTCGGCAACTCCAGTTGGCAGAAAGGGCCGCTCGGCCAATCTTGCTGCAACTATTGCTCCTGG GAAGATGATGTAATCACTCATTGCAAAAACAAACAGTGTACAAGAACAGCCTGGTACCCTGTTGTTTAT CCAAATTTTCTCTGCAAGTGGAGCATCTGCTAGGATCAATAGCAGCAGTGTAAAGCAGGAAGCTACATT CTGTTCCCAAAGGATGGCGATGCCTCTTTGAATAAAGCCCTATCTCAAGTGTCTGATGATGCGCTCTTG GTTAATGCCTCAATTTCCAGCTCTGTGAAGCTACTTCTCCAGTGAAATCTACTACATCTATCACTGATG CTAAAAGTTGTGAGGACAAAATCTGAGCTACTTCCAAAACCTCTATTAGTCTCTGAAAACGGGGGT ATCGAAACCAATTTGTAAGTCAACTTTATCCAGACAGTTCCATCCAAGGAGAAATTAAGTAGAGAAAT TGCTGCAATCTCAATCTAAAGACAAATCTACGACACCAGGAGGAACAGGAATTAAGCCCTTCTGGAAC GCTTTGGAGAGCGTTGTCAAGACATAGCAAAAGAAAGTCCAGCTCGTAGCACACCCCAAGAACCCCAT TATTACTCCAAATACAAAGGCCATCCAAGAAAGATTATTCAAGCAAGACACATCTTCACTACTACCCAT TTAGCACAAACAGCTCAAGCAGGAACGTCAAAAAGAACTAGCATGTCTTCCGTGGCCGATTTGACAAGGGCA ATATATGGAGTGCAGAAAAAGGCGGAACTCAAAAAGCAAACTAGAAACCAACAGGAAACTCACTG TCGAGCACTCCCTCAAAAACACCAAGGTGTTTCAAAAACCTCAGTCACTTCCAGTACAGAACGAAAGGTG ACCGAAAACAGATACCAGCCAAAAATCTAGTACAGAACCTAAAGGTTTCACTGAATGCGAAATGACGA AATCTAGCCCTTTGAAAATAACATTGTTTTTAGAAGAGGACAAATCTTAAAGTAACATCAGACCCAAA GGTTGAGCAGAAAATTGAAGTGATACGTGAAATTTGAGATGAGTGGGATGATGATGATATCAATAGTTTCG AAAGTAATTAATGACCTCTTCAGTGATGCTTAGAGGAAGGTGAAGTATGATGAGGAAAGGCAAGAGG AGATGGATCAAGCATTAGCAGAAAGCAGCGAAGAACAGGAAGATGCACTGAATATCTCTCAATGTCTTT ACTTGCACCATTTGCCACAAACAGTTGGTGTGGTAAGTCCAGAGAGTTTAGTGTCCACACCTAGACTGGAA TTGAAAGACACCAGCAGAAGTGAAGAAAGTCCAAAACAGGAAAAATCCAAAGAACTCGTGTCCCTCGAG CTGAATCTGGTGATAGCCTTGGTTCTGAAGATCGTGATCTCTTTACAGCATTGATGCATATAGATCTCA AAGATTCAAAGAAACAGAACGTCCATCAATAAAGCAGGTGATTGTTCCGGAAGGAAGATGTTACTTCAAAA CTGGATGAAAAAATAATGCCTTTCTTGTCAAGTTAATATCAACAGAAAAATGCAGGAACCTCAATAACG AAATAAATATGCAACAGACAGTGAICTATCAAGCTAGCCAGGCTCTTAAGTGTGTGTTGATGAAGAACA TGGAAAAGGGTCCCTAGAAAGAGCTGAAGCAGAAAAGACTTCTTCTAATTGCAACTGGGAAGAGAACACTT TTGATTGATGAATTGAATAAATTGAAGAACGAAGGACCTCAGAGGAAGAATAAGGCTAGTCCCCAAAGTG AATTTATGCCATCCAAAGGATCAGTTACTTTGTGCAAAATCCGCTTGCCTCTAAAAGCAGATTTTGTCTG CAGTACGGTTTCAGAAACAGATGCAGCAAAATTAATTACTTAATTATACTAAAAGCAGGAGCTGAAAAAT ATGGTAGCCACACCATTAGCAAGTACTTCAAACCTCTCTTAACGGTGATGCTCTGACATCTCACTACTACAT TTACTCTGCAAGATGTATCCAATGACTTTGAAATAAATAATTGAAGTTTACAGCTTGGTGCAAAAGAAAGA TCCCTCAGGCCTTGATAAGAAGAAAAAATCCAAAGTCCAAAGGCTATTACTCCAAAGCAGCTCCTCACA TCTATAACCACAAAAAGCAACATTCTTTCAGTCAATGCCAGTCCAGGAGGCTTAGTGTCTGTGCGAA CCAGCAACTTCGCCCTTGTGGATCTTACACATTATCATGCTCTCAGTAGGAAATTAAGTTTGTCTCT GGACAAGGTCCTTTTTATCTTCTTTGGAAGGTCATATTTATTTAAAAATAAAATGTCAAGTGAATTC AGTGTGGAAGAAAGAGGTTTTCTAACCATATTTGAAGATGTTAGTGGTTTTGGTGCCTGGCATCGAAGAT GGTGTGTTCTTCTGCAAACTGTATATCTTATTGGACTTATCCAGATGATGAGAAACGCAAGAATCCCAT AGGAAGGATAAATCTGGCTAATTGTATGTAACAGTCGTGATAGAACCAGCCAAACAGAGAATTTGTGCAAGA CGCAACACTTTTGAATTAATTACTGTCCGACCACAAAGAGAAGATGACCGAGAGACTCTTGTGAGCCAAAT GCAGGGACACACTCTGTGTACCAAGAACTGGCTGTCTGAGATACATAAGAAAGAGCGGGATCTCTGGAT GCAAAAACCTCAATCAAGTTCTTGTGTATATTCGCTCTGGCAACCTGATGCTTGTACAAACCTATGGA	95

	AAGCCTTAAACCGGGAATTTCCATGCTATCTAGAGGTTTTTGATGTCATCTTAAGAAACACACTTAAAGA GCATCAGATTACTGATTGCAATTTTATGCTTTAAGTACGAAAGGGTTTGIGCCAATATTCACCTACGTAAT ATGCGAGTATTTATATCTTTTGTATGTAAACTTTTAACTGATTTCTGTCTATTCATCAATGAGTAGAAGTAA ATACATTATAGTTGATTTTGCTAAATCTTAATTTAAAGCCTCAITTTCCTAGAAATCTAATTATTCAGT TATTCATGACAATATTTTTTAAAGTAAGAAATTCGTAGTTGTCTTCTTGGAGCTGTAGGCTCTTGAAGC AGCAACGTCCTTTCAGGGGTTGGAGACAGAAACCCATTCTCCAATCTCAGTAGTTTTTTCGAAAGGCTGTG ATCATTATTTATGATCGTGATATGACTTGTACTAGGGTACTGAAAAAATGTCTAAGGCCTTTACAGAAAC ATTTTTAGTAATGAGCATGAGAACTTTTICAAATAGCAATATATATTGGCTTAAAGCATGAGGCTGTCT TCAGAAAAGTGATGTGGACATAGGAGCAATGTGTGAGACTTGGGGGTTCAATATTTATATAGAAGAGT TAATAAGCAGATGGTTTACATTTACTCAGCTACTATATATGCAGTGTGGTGCACATTTTTCACAGAAATCT GGCTTCATTAAGATCATTATTTTGTCTGCGTAGCTTACAGACTTAGCATATTAGTTTTTCTACTCCTAC AAGTGTAAATTGAAAAATCTTTATATTAAGTAAGTAACTGTTATGAAGCTGCTATGTACTAATAATACT TTGCTTGGCAAAGTGTTTTGGGTTTTGTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG GTCTAGAAACCCATTTTGAAAGTGGAAAAATTTAAGTCACCTATCACCTTTAAACGCCCTTTTTTAAAA TTATAAAATATTGTAAAGCAGGGTCTCAACTTTTAAATACACTTTGAACCTCTTCTCTGAATTTATTAAG TTCTTTATGACCTCATTTATAAACACTAAATCTGTGCACCTCCTGTCTATTTATTTTTTATTCATTCAA TGATTTTTTCTTGTGCATATTATAAAATATATTTTATGAGCTCTTACTCAAAATAAATACCTGTAAATG TCTAAAGGAAAAAAAAAAAAAAAAA	
NM_004323	AGGCCGGGGCGGGGCTGGGAAGTAGTCGGGCGGGGTTGTGAGACGCCGCGCTCAGCTTCCATCGCTGGGC GGTCAACAAGTGCGGGCTGGCTCAGCGCGGGGGGGCGCGGAGACCGCGAGGCGACCGGGAGCGGCTGGG TTCCCGGCTGCGCGGCCCTTCGGCCAGCGCGGGAGCGCGCGCAGTCGGAGGCCCGCGGGCGGTGTCGG CCICCCCTCICGGCGTCCACCTGCCCGGAGTACTGCCAGCGGGCAIGACCGACCCACCAGGGGCGCGCGCG CCGGCGCTCGCAGGCCGCGGATGAAGAAGAAAACCGCGCGCGCTCGACCCGGAGCGGAGGAGTTGACCCG GAGCGAGGAGTTGACCCTGAGTGAGGAAGCGACCTGGAGTGAAGAGGCGACCCAGAGTGAGGAGGCGACC CAGGCGCAAGAGATGAATCGGAGCCAGGAGGTGACCCGGGACGAGGAGTCCGACCCGAGCGAGGCTGA CCAGGGAGGAAATGGCGGCAGCTGGGCTCACCGTACTGTCAACCCACAGCAATGAGAAGCAGCACCTTCA TGTTACCTCCAGCAGGGCAGCAGTGAACCAAGTGTCCAAAGACCTGGCCCAAGGTTGTTGAAGAGGTGATA GGGGTTCCACAGTCTTTTTCAGAACTCATATTTAAGGGAAATCTCTGAAGGAAATGGAACACCGTGTGT CAGCACTTGAATACAAAGATGGTTTGC CGGGT CATGTTAATTGGGAAAAAGAACAGTCCACAGGAAGAGGT TGAACATAAGAAGTTGAACATTTGGAGAAGTCTGTGGAGAAGATAGCTGACCAGCTGGAAGAGTTGAAT AAAGAGCTTACTGGAATCCAGCAGGGTTTTCTGCCCAAGGATTTGCAAGCTGAAGCTCTCTGCAAACTTGT ATAGGAGAGTAAAGCCACAATAGAGCAGTTTATGAAGATCTTGGAGGAGATTGACACACTGATCTCGCC AGAAAAATTTCAAAGACAGTAGATTGAAAAGGAAAGGCTTGGTAAAAAAGTTTCAAGGCATTCTAGCCGAG TGTCACACAGTGGAGCAGAACATCTGCCAGGAGACTGAGCGGCTGCAGTCTACAAACTTTGCCCTGGCCG AGTGAGGTGTAGCAGAAAAAGGCTGTGCTGCCCTGAAGAAATGGCGCCACCAGCTCTGCCCTCTCTGGAGC GGAATTTACCTGATTTCTTCAGGGCTGCTGGGGGCAACTGGCCATTTGCCAATTTTCTACTCTCACACT GGTTCTCAATGAAAAATAGTGTCTTTGTGATTTTGTAGTAAAGCTCTCTATCTGTTTTCTCTCTCTGCTCT GTGGTTGTACTGTCCAGCAATCCACCTTTTCTGGAGAGGGCCACCTCTGCCCAAATTTTCCAGCTGTTT GGACCTCTGGGTGCTTTCTTTGGGCTGGTGAGAGCTCTAATTTGCCTTGGGCCAGTTTCAGGTTTATAGG CCCCCTCAGTCTTCAGATACATGAGGGCTTCTTTGCTCTTGTGATCGTGTAGTCCCATAGCTGTAAAAAC AGAAATCACCAGGAGGTGTGACCTAGTCAAGAAATATGGGAATGGCTAGAACAAAGTGTGTTGGCAGATAA GTAGACCACTTATCCCTCATTTGTGACCTAATTCAGAGCATCTGGCTGGGTGTTGTTGGGTTCTAGACTTTG TCCTCACCTCCCAGTGACCTGACTAGCCACAGGCCATGAGATACCAGGGGGCCGTTCTCTTGGATGGAGC CTGTGGTTGATGCAAGGCTTCTTTGTCCCCAAGCAAGTCTTCAGAAGGTTAGAACCAGTGTGTACTGAG CTGTGCTTGAACACAGGCCAGAGCCATGGATTAGGAAGGGCAAGAGAGAAGGCACACAGATGAGTAAAGC AGGCGAGTGGTGAAGCCAACCATAAACTTCTCAGGAGTGACATGTGCTTCTTCAAAGGCATTTTTGTGTA ACCATATCCTTCTGAGTTCTATGTTTCTTTCACAGCTGTTCTATCCATTTTGTGGAGTGTCCCCACCCC CACCCCATCATTTGTTTTTAAAAAATTAAGGCCTGGCGCAGCAGCTATGCCATATAATCCCAGCACTTTGG GAGGCTGAGGCGGGCGGATCACTTGAGGCCAGGAGTTTGAGACCAGGCCAGGCACACATAGCAAAAACCCCA TTCTGCTTTAAAAAATAAAAAAAAAAATAGCTTGGCGTAGTGGCATGTGCCTATAATCCCAGCTACT GGGGAGGCTGAGGCACAAGAAATCATTTGAACCTGGGAGGTAGAGGTTGCTGTGAGCCGAGATTACGCCCC TGCACTCCAGCCTGGGTACAGAGTGAGACTCCATCTCAGAAAAAATAAATGAGTCAAGGTGAGTGTAGT CTCCTTCCGTGAGTCCAGCTACTTGGAGGCTGAGGCTAGAGGATCACTTGAGCCAGGAGTTTGAGTC TAGTCTGGGCAACATAGCAAGACCCCATCTCTAAATTTAAGTAAGTAAAGTAGATAAATAAAGAA AAAAAACTGTTTATGTGCTCATATAAAGTAGAAGAGTGGTTTGCTTTTTTTTTTTTTTTGGATTAAATG AGGAAATCATTTCTGTGGCTCTAGTCATAATTTATGCTTAATAACATTGATAGTAGCCCTTTGCGCTATAA CTCTACCTAAAGACTCACATCATTTGGCAGAGAGAGAGTCTTGAAGTCCCAGGAATTCAGGACTGGGCA GGTTAAGACCTCAGACAAGGTAGTAGAGGTAGACTTGTGGACAAGGCTCGGGTCCCAGGCCACCGCACCC CAACTTTAATCAGAGTGGTTCACTATTGATCTATTTTTGTGTGATAGCTGTGTGGCGTGGGCCACACAT TTAATGAGAAGTTACTGTGCACCAAACTGCCGAACACCAATTTCAAATATTATATATATTAGTCAITTTA ATTCTTACATAACTTGAGAGGTAGACAGATATCCTTATTTTAGAGATGAGGAAACCAAGAGAACCTTAGGT CATTAGCGCAAGGTTGTAGAGTAAGCGGCAAGGCAAGACACAAAGCTGGGTGGTTTGGTTTCAGAGCCA GTGCTTTTTCCCTCTACTGTACTGCTCTCAACCAACACAGGGTTGCACAGGCCCATCTCTGATTTTTT TCCTCTTGCTCTGCTCTGCTCTGCTCTAGCTCCACTTCTCTCTGCTCTAGTTTCTTTCTTTAGAGCAG CCCGAGTGATCATGAAGTGCAAATCTTGCCATGTGAGTCCCTGCTTAGAACCCCTCCAATGGCTCACTTT CTCTTTAGGCAAAAGTCTTTACCCCATGCTTCTCTCCATCTCATCTCAACCCCTCATTTGTGGCTGTC TGCTGTGAGCCACTCTTCTTTCAGGTCTTCAGATGCACCTCTCTGCTGGGGGCTTTGCTCC TGCTACTACCTCTGCTTGAACAGCTCTCACCTTCTCTCCCAACCCCTACCCCTTGATAGGTGACTTTT	96

	GTTCATCCTTCAGAATTCAACTCACAATGCTCTCTTGCATGGAGAACCCTCACCTACTGTTGAGACCCCTG TCCAGCCCCCAGGTGGGATCCTCTCTCGACTTCCCATACATTTCTTTACAGCATTACATAGTCCATGA TAGTTTACTTGTGGGATTATTTGGTTAATCTTTGCCTTTAACACCAGGGTTCCCTGGGGTGAAGGAGCTTC TTTATCTTGGTAACAGCATTATTTCAAGCATAACTTGAATATAGTTATATTACATATATAACATATATA TATATAACATAACATATATAACATATATAACAAGCATAACTTGTATATAGTCTTGTATATAGTAAGACC TCAATAAATATTTGGAGAACAAAAA	
NM_000633	TTTCTGTGAAGCAGAAGTCTGGGAATCGATCTGGAATCCCTCCTAATTTTTACTCCCTCTCCCGCGACT CCTGATTTCATTGGGAAGTTTCAAATCAGCTATAACTGGAGAGTGTGAAGATTGATGGGATCGTTGCCTT ATGCATTTTGTGTTTGGTTTTACAAAAAGGAACTTTGACAGAGGATCATGCTGTACTTAAAAAATACACAT CACAGAGGAAGTAGACTGATATTAACAATACTTACTAATAATAACGTGCCCTCATGAAATAAAGATCCGAA AGGAATTGGAATAAAAAATTTCTCGCATCTCATGCCAAGGGGGAAACACCAGAATCAAGTGTTCGCGGTGA TTGAAGACACCCCTCGTCCAAGAAATGCAAAGCACATCCAATAAAATAGCTGGATTATAACTCTCTTCT TTCTCTGGGGGCCGTGGGGTGGGAGCTGGGGCGAGAGGTGCCGTTGGCCCCCGTTGCTTTTCTCTGGGA AGGATGGCGCACGCTGGGAGAACAGGGTACGATAACCGGGAGATAGTGAAGAAGTACATCCATTATAAGC TGTGCGAGAGGGGTACGAGTGGGATGCGGGAGATGTGGGCGCCGCGCCCCGGGGGCGCCCCCGCACC GGGCATCTTCTCTCCAGCCCGGGCACACGCCCATCCAGCCGATCCCGGGACCCGGTGCACAGGACC TCGCCGCTGCAGACCCCGGCTGCCCGCGCGCGCGGGGCTGCGCTCAGCCCGTGCCTGTTGGTGG TCCACCTGACCCCTCCGCCAGGCGCGGCGAGACTTCTCCCGCGCTACCGCCGCGACTTCGCCGAGATGTC CAGCCAGCTGCACCTGACGCCCTTCACCGCGCGGGGACGCTTTGCCACGGTGGTGGAGGAGCTCTTCAGG GACGGGGTGAAGTGGGGAGGATTGTGGCTTCTTTGAGTTCCGGTGGGGTGCATGTGTGTGGAGAGCGTCA ACCGGGAGATGTGCCCCCTGGTGGGACACATCGCCCTGTGATGACTGAGTACCTGAACCGGACCTGCA CACCTGGAATCCAGGATAACGGAGGCTGGGATGCCCTTTGTGGAACGTACGGCCCCAGCATGCGGCCCTGT TTTGATTTCTCTGGCTGTCTCTGAAGACTCTGCTCAGTTTGGCCCTGGTGGGAGGTTGCATCACCCTGG GTGCCATCTCTGGGCCACAAGTGAAGTCAACATGCTGCCCCAAACAAATATGCAAAAGGTTCACTAAAGC AGTAGAAATAATATGCATTGTCAAGTGTACCATGAAACAAAGCTGCAGGCTGTTTGAAGAAAAAGTAAAC ACACATATAAACATCACACACACAGACAGACACACACACACAACAATTAACAGTCTTCAGGCAAAACG TCGAATCAGCTATTTACTGCCAAAGGGAAATATCATTATTTTACATTATTAAGAAAAAAGATTATAT TTATTTAAGACAGTCCCATCAAACTCCCTGTCTTTGGAATCCGACCACTAATTGCCAAGCACCGCTTCG TGTGGCTCCACCTGGATGTCTGTGGCTGTAAACATAGATTGCGCTTCCATGTTGTGTGGCCGATCACC TCTGAAGAGCAGACGGATGAAAAAGGACCTGATCATTGGGGAAGCTGGCTTCTGGCTGCTGGAGGCTG GGGAGAAGGTGTTCACTTGCATTTCTTTGCCCTGGGGGCTGTGATATTAACAGAGGGAGGGTTCTCT GTGGGGGAAGTCCATGCCCTCCCTGGCCCTGAAGAAGAGACTCTTTGCATATGACTCACATGATGCATACC TGGTGGGAGGAAAAGAGTTGGGAATTCAGATGGACCTAGTACCCACTGAGATTTCACCGCGAAGGACA GCGATGGGAAAAATGCCCTTAAATCATAGGAAAGTATTTTTTAAAGCTACCAATTTGTCCGAGAAAAAGCA TTTTAGCAATTTATACAATATCATCCAGTACCTTAAGCCCTGATTGTGTATATTCATATATTTTGGATAC GCACCCCCCAACTCCCAATCTGGCTCTGTCTGAGTAAGAAACAGAATCCTCTGGAATCTGAGGAAGTGA ACATTTCCGTGACTTCCGCAATCAGGAAGGCTAGAGTTACCCAGAGCATCAGGCCCGCCACAAGCTGCT TTTAGGAGACCGAAGTCCGCAAGCTGCTGTGTCCCAGCTTGGAGGCTGGTCTGGAATGAGCCGG GGCCCTCACTGGCTCCTCCAGGGATGATCAACAGGGCAGTGTGGTCTCCGAATGTCTGGAAGCTGATGG AGCTCAGAATTCCTCTGTCAAGAAAGAGCAGTAGAGGGGTGTGGCTGGGCTGTACCCCTGGGGCCCTCC AGGTAGGCCCGTTTTTACAGTGGAGCATGGGAGCCACGACCTTCTTAAGACATGTATCTGAGAGGA AGGAACAGAGGCCCTGGGCCCTTCTATCAGAAGGACATGGTGAAGGCTGGGAACGTGAGGAGAGCAAT GGCCACGGGCCATTTTGGCTGTAGCACATGGCACGTTGGCTGTGTGGCCTTGGCCACCTGTGAGTTTAA AGCAAGGCTTTAAATGACTTTGGAGAGGGTCAAAATCCTAAAGAAGCATTGAAGTGAAGGTGTATGGA TTAATTGACCCCTGTCTATGGAATTACATGTAAACATTAATCTGTCTGACTGTAGTTTGGTTTTATTTGAA AACCTGACAAAAAAAGTTCCAGGTGTGGAATATGGGGTTATCTGTACATCCTGGGGCATTAAAAAA AAATCAATGGTGGGAACATAAAGAAGTAACAAAGAAGTGACATCTTCAGCAATAAATAGGAAATTT TTTTTTCTTCCAGTTTGAATCAGCCTTGAAACATTGAATGAATAACTCTGTGGCATTATTGCATTATA TACCATTATCTGTATTAACTTTGGAATGTACTCTGTTCATGTTTAAATGCTGTGGTTGATATTTCGAAA GCTGCTTTAAAAAATACATGCACTCAGCGTTTTTTTGTTTTTAATTGTATTTAGTTATGGCCTATACA CTATTTGTGAGCAAAGGTGATCGTTTTCTGTTTGAGATTTTATCTCTGATCTTCAAAAGCATTTCTGA GAAGGTGAGATAAGCCCTGAGTCTCAGCTACCTAAGAAAAACCTGGATGTCTAGTGGCCACTGAGGAGCTT TGTTCACCAAGTCATGTGCATTTCCACGTCAACAGAAATGTTTTATTGTGACAGTTATATCTGTTGTCC CTTTGACCTTGTCTTGAAGGTTTCCCTGCTCCCTGGGCAATTCGCAATTTAATTCATGTTATTCAGGAT TACATGCATGTTTGGTTAAACCCATGAGATTCAATTCAGTTAAAAATCCAGATGGCAATGACCAGCAGAT TCAAACTATGGTGGTTTGACCTTTAGAGAGTTGCTTTACGTGGCTGTTTCAACACAGACCCACCCAGA GCCCTCCTGCCCTCCTTCCGCGGGGGCTTTCTCATGGCTGTCTTCAGGGTCTTCTGAAATGCAAGTGGT GCTTACGCTCCACCAAGAAAGCAGGAAACCTGTGGTATGAAGCCAGACCTCCCGCGGGGCTCAGGGAA CAGAATGATCAGACCTTTGAATGATTCTAATTTTTAAGCAAAATATATTTTATGAAGGTTTACATTGT CAAAGTGATGAATATGGAATATCCAATCCTGTGCTGCTATCCTGCCAAAAATCATTTTAAATGGAGTCAGTT TGCAGTATGCTCCAGTGGTAAGATCCTCCAGCTGCTTTAGAAGTAACAAATGAAGAAGCTGGACGTTTTT TAATATAAGCCTGTTTTGCTTTTTGTTGTTGTTCAAACGGGATTCACAGAGTATTTGAAAAATGTATAT ATATTAAGAGGTACGGGGGCTAATTTGCTGGCTGGCTGCTTTTTGCTGTGGGGTTTTGTTACCTGTTTTT AATAACAGTAAATGTCCCGACCTCTTGGCCCCAGAACTGTACAGTATTGTGGCTGCCTTGTCTAAGA GTAGTTGATGTGCTATTTTCTTTTATTTTAAAAACATGTTAGAAGCAATGAATGTATATAAAGCCCTCAA CTAGTCAATTTTTTCTCTCTCTTTTTTTTTTCAATATATCTAATTTATTTTGCAGTTGGGCAACAGAGAAC CATCCCTATTTTGTATTGAAGAGGGATTACATCTGCATCTTAAGTGTCTTTATGAATGAAAAACAGT CCTCTGTATGTACTCCTCTTTTACACTGGCCAGGGTCAGAGTTAAATAGAGTATATGCATTTTCCAAATTG	97

	GGGACAAGGGCTCTAAAAAAGCCCCAAAAGGAGAAGAACATCTGAGAACCTCCTCGGCCCTCCAGTCC CTCGCTGCACAAATACTCCGCAAGAGAGGGCCAGAAATGACAGCTGACAGGGTCTATGGCCATCGGGTCTGTC TCCGAAGATTTGGCAGGGGCAGAAAACCTCTGGCAGGCTTAAGATTTGGAATAAAGTCACAGAATTAAAGGA AGCACCTCAATTTAGTTCAAACAAGACGCCAACATTCTCTCCACAGCTCACTTACCTCTCTGTGTTTCA TGTGGCCTTCCATTTATATGTGATCTTTGTTTTATTTAGTAAATGCTTATCATCTAAAGATGTAGCTCTGG CCCAGTGGGAAAAATTAGGAAGTGATTATAAATCGAGAGGAGTTATAAATCAAGATTAATGTAAATA ATCAGGGCAATCCCAACACATGTCTAGCTTTCACTCCAGGATCTATTGAGTGAACAGAATTGCAAAATAG TCCTATTTGTAAITGAACCTTATCTTAAAAAATAGTTTTATAAATGTGAACCTTAACTCTAATTAATTC CAACTGTACTTTTTAAGGCAGTGGCTGTTTTTAGACTTTCTTATCACTTATAGTTAGTAATTTACACCTAC TCTATCAGAGAAAAACAGGAAAGGCTCGAAATACAAGCCATTCTAAGGAAATTAGGGAGTCAGTTGAAAT TCTATTCTGATCTTATTCTGTGGTGTCTTTTGCAGCCAGACAAATGTGGTTACACACTTTTTAAGAAAT ACAATTTCTACATTGTCAAGCTTATGAAGTTCCAATCAGATCTTTATTGTATTCAATTTGGATCTTTCA GGGATTTTTTTTTTAAATTTATTATGGGACAAAGGACATTTGTTGGAGGGGTGGGAGGAGGAATTTTT TAAATGTAAACATTTCCCAAGTTTGGATCAGGGAGTTGGAAGTTTTTCAAGATAACCAGAATAAGGGTAT GAAGGACCTGTATTGGGGTCGATGTGATGCCTCTGCGAAGAACCCTGTGTGACAAATGAGAAACATTTTG AAGTTTGTGGTACGACCTTTAGATTCCAGAGACATCAGCATGGCICAAAGTGCAGCTCCGTTTGGCAGTG CAATGGTATAAATTTCAAGCTGGATATGTCTAATGGGTATTTAAACAATAAATGTGCAATTTTACCTAAC AGGATATTTAATGACAACCTTCTGGTTGGTAGGGACATCTGTTTCTAAATGTTTATTATGTACAATACAG AAAAAATTTTATAAATTAAGCAATGTGAAACTGAATTGGAGAGTGATAATACAAGTCTTTTAGCTTTA CCCAGTGAATCATTTCTGTCCATGTCTTTGGACAAACCATGACCTTGGACAAATCATGAAATATGCATCTCA CTGATGCAAAAGAAAATCAGATGGAGCATGAATGGTACTGTACCGGTTTCACTCGGACTGCCCAAGAAA TAACTTCAAGCAACATCTATCAACAACAAGGTTGTTCTGCATACCAAGCTGAGCAGAGAAGATGGGAA CACTGGTGGAGGATGGAAAGGCTCGCTCAATCAAGAAAAATCTGAGACTATTAATAAATAAGACTGTAGT GTAGATACTGAGTAAATCCATGCACCTAAACCTTTTGGAAAAATCTGCCGIGGGCCCTCCAGATAGCTCAT TTCAATTAAGTTTTTCCCTTCCAAGTTTGGAAATTTGCAAGAGTGACAGTGGATTGCAATTTTGGGGAAGC TTTCTTTTGGTGGTTTTGTTTATTATACCTTCTTAAAGTTTTCAACCAAGGTTTGCTTTTGTTTTGTAGTTA CTGGGGTTATTTTTGTTTTAAATAAAAAATAAGTGTAATAAGTGTTTTTGTATTGAAAGCTTTTGTATT CAAGATTTTCACTTTTACCTTCCATGGCTCTTTTTAAGATTGATACTTTTAAAGAGTGGCTGATATTC TGCAAACACTGTACACATAAAAAATACGGTAAGGATACTTTTACATGGTTAAGGTAAGTCAAGTCCAGGTT GGCCACCATTAGCTATAATGGCACTTTGTTTGTGTTGTTGAAAAAGTCACTTGGCATTAAACTTTTCTT TGCTGTCTAGTTAATTTGTGAAGAAAAATAAAGTACAGTGTGAGATACTG	
BX647539	AATGAGGGTATTTATAAACTACTTAAATTTAAAAAGAATGAGACATCAGACTTACAGTTTTGGATACTA ATTTTTTTCACTTAACGTTTCATTATGTGATAGGAGTTTTTCCATCCTATTATACCGCTGTGCGATCTGATC TTGGGCACGTTAACCACCTCTTGTGCTTCGATTTTCTCACCTGTAAAAGTGGGGGTAATCATAATGCT TACTTAGTAGGATAGCCCTGAAGAATAAGTGACTTAGCGAACATAAATAGCTTACAAATAGGGTTTTGAGC ATGGGAAGGATTTCAGTAAATGTTAGCTGTCATCATCACCACCTACAAAGGAAGCAATCTGTGCTGAAAG TTTTTCCATCATTAAATGTAATTTCTATAGTACGATTCCCAAGAAGATATTAAAAATTAGAAATAAGGT ATTGGTATATTCCTAATTTATTTCTTAAAGATTGTATTGATAAATATGTCATCTCTTCCCTTAACGGGAT GCATTCAGAAAAACAAGTCAAAATGTAGACAAAGTATCAGAAGGGAAATCTGTAGCCAGAGAGCTAA AATTACAATAGGGTCTCTAATTATCTTCAACTTTTTTAGGAATAATTTTCAGTGTGTTTTCCACATTT CATATGTAATTTTTTTTTTTTTTTTTTTTGGAGACAGAGCCTCGCCCTGTCCACAGGCTGGAGTACAGTG GCGCGATCTCGGCTCACTGCAACTTCCACCTGCTGGGTTCAAGCAATTTCTTCTGACCTCAGGTGATCCAC CCGCCCTCGGCCCTCCCAAAGTGCTGGGATTATAACAGGCGTGGCATGAGTCAACGCCGCCGCCGATCTTT ACTTTTTTATCTTTGTACCCCCCTGCTATCCAGTTAGCATGTGATTAAAGTCAAGATTTTGGCCTTTG GGCCACATCTATTAATTTTTCATCTTTTGTATATTTGTATTTAGTTTGTATCTACACTGCTTATTACTCC CAGTCAATTTTTATAGAATGAAAACTGGTAAATACTCAAAATTGCACTGACTTCTATGTAGAGGCGA CACTCCATCAGAACCGTGGGCTGACAGGGGAATCCCACTGTGCAGGAGCTGCGCGCATTTTCAITTTCTGAT TCCTTTTGGCGTATCCAGGACTCTGATGACATGATCATATTTTATCAGTAGTAACAGGTTGGGCCATTT GTTTTTGTGGTAAATCATATATTTAAGATTTTAGAAATAAGTTGATAGCCATGTATTTTGAATTTGAA AAAGACATTGCATTACTCAGCTTCAAATTAAGCTTTAATCAAATAGTGAACCTTTCCATTAAATGGACAGT GTATACCTTTTTGTGTTATTTAAAAAATAAAGCACTGAATATAGTGCCTTTTGTGACAGGGGAGCTTGGTTC CTGACAATGTCTCTTGAGCCTTTTTTTTTTTTTTGTAGATGGAGTCTCACTGTGTACCCAGGCTGGAGT GCAGTGGCGCCATCTTGGCTCAGTGCACCTCCGCCCTGGGTTCAAGTGTATCTCATTTCTCAGCTTCT CTAAGTAGCTGGGATTACAGGCACGCACCCACCATGACCAGCTAATTTTTTATCTTTTGTAGAGACAGGG TTTTGCCATGTTGGCTAGGTTGGTCTCGAAGTCTGACCTCAAGTAATCCACCCACCATGGCCTCCCCAA AGTGCTGGGATTACAGGCGTGAGCCATTTACCCGGCCTCTCTTCCGCTCTTTGAGCTGTGAGGAAATAGC TACATTACATGAGCTGTGATCTGCCTTATGGTCAGAAATGAAGGTTGAACCTCTCAGGAACAGTGACAT ATATACACACTGATATTTCCAAAGTACAATGCCCAAAATGTATCCACAAAGGAATTAAGGTCAATTTGCAA CAAAATCAGAGAAATAGTAACAAATAAATAGAAGATAAATATGGCCAGGGATGCTGCAAACTGATATAGT CCAAGTTTATCAGTTGGGAATCCCAACAGTGAAAAGCATAAAAATGAAGGAATTTAAGGAGACTTTTT ATAGAAGAGTGGGAAGGATTGGAGGAGCCAACAGTGATGTTGAGGCACAGAGGGAAGCTTTAGTGGG CACCATCCCTCTCTGGTTTGAAGGGGTAGGGAGGGGACCAGAGCTGGGAGGAGGGGGCTGGAATACTGC TGGAGGAGCCACTCCCTTCCAGACTGTGTGGCCATCAGAAATGCAGCCACTGCCAGAGCAGCAGCCC GAGGAACCAAGGAGGGGGAGCACAAGTACCCTAGCCTCTCTCTTTCTGTTTCTTGGCTGCCGATCTCTCT CACTGGCTAAACCCAGCTGGATGCTAAGAGTACAGTCAAGCTGCCTGCTGAGGAGGGACCAACAGGGACC ACCATCAGCAAGGGATCCAATGTCTTTCTGCTCTGCAAGTGAAGGTTGGGGCGCGGGGGCGCTCTAC TTCTTAGGGATATTGTGGGAATAAAGGAAATAGGCAAAAAATGTTTTGAAAAACAAAGCACATAGTGC GCACCCGTGGGCCACTACTGCTTTTGAACCCCTGGCTCTGTTTCAAGTAATGTGCTGTCAATCTCTTT	98

	TTAGGTGCTACAGGATTTCTTTAGGTTTGTCTGTCCACCATATTTCAACTCATGTGTGCTGTTTGT GTGCTAAAACAAATATTTGCTGATGCCCTGAGTGAATAGTTGAATATTTATATAAGTCAAATTTATACGT AATGATTTTTCTTTGTAACCTTAGCCGTTTCTCTTTTACAACTCAGAAAACTCAGACTTTGAAAAGGCCCT TGAAGTTCCTCACCTGAAATCTGAGAACCTGGAGCGCCTTAAAAATCTAAAGGAAAAACAAACAGTGAA AGAACATGATATAGTCAGTGTAGAGAAATAAATATTTATGTAATTAATATTGAGGATGCAGATAACACA TTGTGAAATCTTGCTTGTAAAAATCTCGATCTGCTGAAGAAAGATGTTCTCTCTAGAGATCTTTGAAAG CATAATTATTGAGCTTTTAAAAATGTTAGAAACAAAAGTTAGACCCACACATATTTCTGGCGTGTGGAGAT TTGCATTCTTCCCTGCCGCCGCCGCCACACTTGTGAGTTGTGCCGTGTGACGAGTTCCTGTAGC ACTCGGCTGGGCAGAAATCATCTTTTCAGCACTAAGGGAACATAGTTATGATCTGGACCTCTGGGAGTGG TCAGTGCCCAAGAACAGGTATGGGACTCCAGAAAGTTCTGCTCTCAACCCTATTTTGAATAGAGTTACA CATTGTTCTACAATTATTTGAGTTAATAAGCAGCTCTTTTCAACCGTGATTATGCCCTTCCAAGTTTAAA TACACTAGACTTTAGTGAAGTAATTGACCTCATCTCATTTCTCTCTCTGTTATATTAGATCACTTTTACG TAAAGGTAGAGCTTTTGAAGTGGTGAGGAGGAGGTAGAGGAGGACATAGAGCAGATAGGGGCTGGAA AGTGGGTGAGGAAGAGAGTGGCTTCTCTTTGGCAGAGTACCAAGGAAAAGCCCTATCTGTACAGAACCT TTGTGCTCGGGAACCTTGATGGCTGCAACCTGAGCCTCAACCTAGTTTGTCTGCGGAGCCAGAGAGAGC TAAAAACCTTCAGTTAACCAAGCCAGACACCAAGAAAGTTAAACCGAAAGAGAACCCCCCACCCTGGAA AAAAAAGAAAGTAAAGTGGGTAAAGTGATATCATGTTAGCACAGAAAGAGAACATAAGGGTCACTTAAG TTCATCTGCCCCCTCTTCTATTTCAAGGTCAGAACTAAGGCACAAGGGACCCCGTGTCTCTCTTGA TCACATAGCTAGTGGGTGCCAAGCCAGGCTAGAACTCTGTTCTCTGGGGTACAGGCTGGCTCTTCATC CCCTAGAGAGATAGCTCATCTGTGTGCACCTGAGCCGTTGTGTTTCGGAGTCAAAGCAAAATAAGGCT CAAACCTCAAGACTGTTTTGCAAGCCGGCTGCACTAGATATGGGGGAGGAGAAACCTGCTTTAAATTTGC TTCAAGCAAGTTGTTTCTGCAAGGTGTGACTTTTTTCTTTCACTTTCTAGTGAGTCACTGCAGCCTG AGCTGTTATTTGTCTATTATGCAATAATTTCAGGAACTAACCAAGATTCTCTTTTTTAAATTTATTTGTTA TTTAGAGACAGAGTCTTGCTCTGTTGCCAGGCTGGAGTGCAGTGGTGTGATCTCGGCTCACTGCAGCCT CTGCCCTCGGGTTCAAGCAATTTCTCATGTCTCAGCCTCCCGAATAGCTGGTATTGCGAGCTCCCGCTGC CACCCCTGCTAATTTTTTGAATTTTAGIGGAGACACGGTTTCGCCATGTTGGCCGGCTCGTCTTGAGC TCCTGGCCTCAGGTGATCCGCCCGCTCGGCCCTCCCAAGTGCTGGGATTGCGAGCGTGAGCCTCCACAC CCGGCTATTTATTTATTTTAAATTTGGCTGCTCTTAGAAAGGCATACCATGTTTCTGGATGGGAAGGCT TATTAATTCACCTTAATTTAATGTATATAAATTTGATGCAATCATAGTCACAGTCCCGAGTGGATTTTTTA CTTGGTAAGATGTTCTAAAATTAATGAGAGAACTTGAATTACCAGGTATTGAAACACTGTAAAGCCACAA TCATGTAAACAGTATGTTATAACCATGGGAATAGAGGTCTGTGATACAGCAGAAAAAGTGAAAAAAGA ATAACTGTATTCATAAAAAATTTAAATGTGGAGTCACTGGGGGAAAGGATTAAATATTCGATAATGTAGAA ACAACTCAACTATTGGAGAAATGTAAATTTAGAGCCTTATCTCATGCCATATACCAAAATCACTATTAG ATTTGATTAAAAATAAAAAAAAAAAAAAAAAAAAAA	
BC035498	GCGGCCGCCAGCGCGGTGTAGGGGGCAGGCGCGGATCCCGCCACCGCCGCGCGCTCGGCCCGCCGACTCC CGGCGCCGCCGCCGCCACTGCGCTGCGCCGCCGCCGCTGCCGGGACTGGAGCGCGCGCTCCGCCCGGAC AAGACCTTGGCCTCAGGCGGGAGCAGCCCATCATGCGGAGGGAGCGCAGGGAGCGCGGTGCGAAGGAGC GGGACACCATGAAGGAGGACGGCGCGCGGAGTTCTCGGCTCGCTCCAGGAAGAGGAAGGCAAACGTGAC CGTTTTTTTTGCGAGATCCAGATGAAGAAATGGCCAAAATCGACAGGACCGCGAGGGACCACTGTGGGAGC CAGCCTTGGGACAATAATGCAGTCTGTGCAGACCCCTGCTCCCTGATCCCCACACCTGACAAAGAAGATG ATGACCGGGTTTTACCCAACTCAACGTGCAAGCCTCGGATTATTGCACCATCCAGAGGCTCCCGCTGCC TGTACTGAGCTGGGCAAATAGAGAGGAAGTCTGGAAAATCATGTTAAACAAGGAAAAGACATACTTAAGG GATCAGCACTTTCTTGAGCAACCCCTCTTCTGCGACCAAAAATGCGAGCAATCTTCTGGATTGGTTAA TGGAGGTGTGTGAAGTCTATAAACTTCACAGGGAGACCTTTTACTTGGCACAAGATTTCTTTGACCGGTA TATGGCGACACAAGAAAATGTTGTAAAAACTCTTTTACAGCTTATTGGGATTTCATCTTTTATTTATGCA GCCAACTTGAGGAAATCTATCTCCAAAGTTGCACCAGTTTGCATATGIGACAGATGGAGCTTGTTCAG GAGATGAATTTCTCACCATGGAATTAATGATTATGAAGGCCCTTAAGTGGCGTTTAAATCCCTGACTAT TGTTCTCTGGCTGAATGTATACATGCAGGTTCATATCTAAATGACTTACATGAAGTGCTACTGCCCGAG TATCCCCAGCAAATCTTTATACAGATTGCAGAGCTGTTGGATCTCTGTGCTCTGGATGTTGCTGCTTG AATTTCTTATGGTATACTTGCTGCTTCGGCCTTGATCATTTCTCGTCTCTGAATTGATGCAAAAGGT TTCAGGGTATCAGTGGTGCACATAGAGAACTGTGTCAAGTGGATGGTTCCATTTGCCATGTTTATAAGG GAGACGGGGAGCTCAAACTGAAGCACTTCAGGGCGCTGCTGATGAAGATGCACACAACATACAGACCC ACAGAGACAGCTTGATTTGCTGGACAAAGCCCGAGCAAAGAAAGCCATGTGTCTGAACAAAATAGGGC TTCTCTCTCCCCAGTGGGCTCTCTACCCCGCCACAGAGCGGTAAGAAGCAGAGCAGCGGGCCGGAATG GCGTGACCAACCCATCCTTCTCCACCAAGACAGTTGCGCGCCTGCTCCACGTTCTCTTCTGTCTGTTGC AGCGGAGGCGTGCGTTTGCTTTTACAGATATCTGAATGGAAGAGTGTCTTCCACAACAGAAGTATTTG TGTGGATGGCATCAACAGGGCAAGTGTTTTTTATTGAATGCTTATAGGTTTTTTTTTAAATAAGTGGGT CAAGTACACCAGCCACCTCCAGACACCACTGCGTGTCTCCGATGCTGCTATGGAAGGTGCTACTTGACCT AAGGGACTCCACAAACAAGCTTGAAGCTGTGGAGGGCCACGGTGGCGTGGCTCTCTCGCAGGTG TTCTGGGCTCCGTTGTACCAAGTGGAGCAGGTGGTTGCGGGCAAGCGTTGTGAGAGCCCATAGCCAGCT GGGAGGGGGCTGCGCTCTCCCAATATCAGTTGACAGTGTACAAATGCTTGTATGAATGTTTGTAAAG TGCTGCTATATCTATCCATTTTTTAAATAAGATAATACTGTTTTTGAATAAAAAAAAAAAAAAAAAAAAAA AA	99
BG256659	GAGGGCACGGGCTCCGTAGGCACCACTGCAAGGACCCCTCCCGCTGCGGGCGCTCCCATGGCACAGTTC GCGTTCGAGAGTACCTGCATCGCTGCTTCAGCTGGATGCACCATCCCAATGCAACCCCTGCGCGCT GGCAGCGCAAGCCAAAGGAGCCGAGGCCCGGCCCTCACCCTATGCGGGCCGCCAACCGATCCACAG CGCGGCGCAGGACTCCGGGCGCAACTCTGCAAAATCCAGTTCCAAGGTTCAGACCACTCTAGCAACCT GGCGGTGACCGCTATATCCCCATCGCAGTGTGCGCCAGATGAGGTTGGCCAGCTTCTCTCTGAGCAAGG	100

	AGAACCAGCCTGAAAACAGCCAGACGCCCAACGAAGGAACATCAGAAAGCCTGGGCTTTGAACCTGAA CGGTTTTGATGTAGAGGAAGCCAAAGATCCTTCGGCTCAGTGGAAAAACCACAAAAATGCGCCAGAGGGTT ATCAGCAACAGACTGAAAGTACTCTACAGCCAAAAGGCCACTCCTGGCTCCAGCCGGAAGACCTGCCGTT TACATTCTTCCCTGCCAAGACCGTATCCTGGATGCGCCTGAAATCGAATGACTATTAACGTAACCTGTG GGACTGGCAGTCCCGGGGAATGTCCGGGCCGGGCCACGGCCACGAGGTGTTCCGTGTGGAGTGCAAGCTGG GACACACCGTGCCGCTTGTGCACAGGGCCACGCGGGGAAATAATCCCGGGCGCGCAAAGCGGCACCTGGC GAGAGCCGCACGGGCCGCTGTGGGGTGTACAAACAGGCCAAAACAACACAAAGGCCAACAAAGACATA CGCGCGCTGACACACGGTGCAGGCGCTCAGACGAGTAGTAACCGGCACTGTGGTTGCTGCCCTCCAC CTCTCCCGCTCTCAGCGTAAGATAAAAGAAAGAGAGCAAAAGCAAAAGAAAGAGACAGCAGCAGACAC ACAGGAACGAACAGTAAAGCAAGCTAAAGCAAACGCAAGACCAGACAACAGAAATAGAAAGAACCAACAG AGAGGAGACAGAACAGGACGCCAGCAACATAGCAACAAACGAACAGAAAGAGAGCACTAAACAAAAGCAGC AGCAAGACGAGACAGGAGAGAAGGAGGAAGGAGGGCCGAGCGAGCAGGGAGCGGAGCAGCGAGGCGCAAG CAGCAGACAAGGGCAGGCGAAGGGCAAGAGAGGAGGCCACACAAAAAGGAGAGGAGCAGAGCAAGC AGCGAGAGAAGCGGAGGAGCAACAAGAGGAAGAAAAGGAGAGGGAGAGGAGGAGAGCGGAAGGAGGA AGAAACAGCAGGCGACGAAGGGGGGAGACGCGGGGGCAGGAAAGACACAGGAAGGCAGCGCGGAGG AGGAGAAGGGGAAGCAGGAAGGAGACGGAAGGAGAAGAGGAGAGGACAGCGCAAGAGAGCGCGCGCGC GACGCGAGGGACGGAGCGAGAGAGAGGAAGCGGAAGCGAGAGGGAAGAGGAGAGGAGGAGGAGCAAC CAACCGAAAACAGCAGAAAGAGAGAGAAGGACGCGCAAGAGGCAAGCGCAAGACGACAGGAAACGAAG CGAGAGACGAGAAAGCCGCTGACGAGCAGGAGAAAGGGAAGGCAGGAGACAGGACAGGCGGAAGAGAGACA CGCGAGACGCAAGAGTGAGCAGAACGAAGCGAAGAGCAACGACAGAGAGAAACGAC	
NM_001254	GAGCGCGCTGGAGTTTGTGCTGCCGCTGTGCAAGTTTGTTCAGGGGCTTGTGGTGTAGTCCGAGAGG CTGCGTGTGAGAGACGTGAGAAGGATCCTGCACTGAGGAGGTGGAAAGAAGAGGATTGCTCGAGGAGGCC TGGGGTCTGTGAGGCAGCGGAGCTGGGTGAAGGCTGCGGGTTCGGCGGAGGCTTGAAGCTGTGCTGTGCTG ATGCTCAAACCCGATCCAGGCACAGGCTACAATCAGTTTTCCAAAAGGAAGCTGTCTCGGGCATTTGA ACAAGCTAAAAACTCCAGTGTAGTGCACAACTAGAACCACAAATGTCCAAACCGTAACTGTCTCTCTCG TGTAAAAGCCCTGCCTCTCAGCCCCAGGAAACGCTGTGGCGATGACAACCTATGCAACACTCCCCATTTA CCTCCTTGTCTCCACCAAGCAAGGCAAGAAAGAGAATGCTCCCTCTACTACATACACTTAAGGGAC GAAGATTGGTATTTGACAATCAGCTGACAATTAAGTCTCCTAGCAAAAGAGAAGTACCAAGGTTTACCA AAACAAAATACTTTCTCAGTTAGAAAAGTCAAGAGATCACAACAAATTTCTGAGCAGAGATGTCCTACTG AAGAAAGAACTCTGATGTGTGAGACTATTCAAGCAAGAAGGCCTTGTCTACCAGCAAGCAAGCTGGTCC TGAACACAGCTGTCCAGATCGGCTGCTGCCAGGGAAGGGAGATGGATGTCATCAGGAATTTCTTGAG GGAACACATCTGTGGGAAAAAGCTGGAAGCCCTTACCTTTCTGGTGTCTCTGGAAGTGGAAAAACCTGCC TGTCTAAGCCGGATTCTGCAAGACCTCAAGAAGGAAGTGAAGGCTTTAAACCTATCATGCTGAATGCA TGTCTTGGAGCTGCCAGGCTGTATTTCCAGCTATTGCTCAGGAGATTGTGTCAGGAAGAGGTATCCAG GCCAGCTGGGAAGGACATGATGAGGAAATTTGAAAAACATATGACTGCAGAGAAGGGCCCCATGATGTG TTGGTATTGGACGAGATGGATCAACTGGACAGCAAGGCCAGGATGTATTGTACACGCTATTGTAATGGC CATGGCTAAGCAATTTCTCAGTTGGTGTGTTGGTATTGCTAATACCTCTGGATCTCAGATGAAATTTCT ACCTAGGCTTCAAGCTAGAGAAAAATGTAAGCCACAGCTGTTGAACCTTCCACCTTATACCAGAAATCAG ATAGTCACTATTTTGCAAGATCGACTTAATCAGGTATCTAGAGATCAGGTTCTGGACAATGCTGCAGTTC AATTCTGTGCCCGCAAAGTCTCTGCTGTTTCAGGAGATGTTGCAAAAGCACTGGATGTTTGACGAGAGC TATTGAAATTTAGAGTCAAGTGTCAAAAGCCAGACTATTCTCAAACCACTGTCTGAATGCAACCT TCTGAGCCTCTGATTCCCAAGAGGGTTGGTCTTATTCACATATCCCAAGTCATCTCAGAAGTTGATGGTA ACAGGATGACCTTTAGCCAAAGAGGAGCACAAGATTCCTTCCCTCTCAGCAGAAAGATCTTGGTTTGTCT TTTGATGCTCTTGTACAGGCAGTTGAAAATCAAGAGGTCCTCTGGGGAAGTTATATGAAGCCTACAGT AAAGTCTGTGCAAAACAGCAGGTGGCGGCTGTGGACCACTCAGAGTGTGTTGCTCACTTTCTGG AAGCCAGGGGCATTTTAGGATTAAAGAGAAACAAGGAAACCCGTTTGACAAAGGTGTTTTTCAAGATTGA AGAGAAAGAAATAGAACATGCTCTGAAAGATAAAGCTTTAATTTGAAATATCTTAGCTACTGGATTGCCT TAAATTTCTCTCTACACCCACCCGAAAGTATTACAGCTGGCATTTAGAGAGCTACAGCTCTTCAATTTAG TGCTTTACACATTCGGGCTGAAAACAAATATGACCTTTTTTACTTGAAGCCAATGAATTTTAATCTATA GATTCTTTAATATTAGCACAGAATAATACTTTGGGTCTTACTATTTTTTACCATAAAAGTGACCAGGTA GACCTTTTTTAATTACATTCACACTTCTACCACTTGTGTATCTCTAGCCAATGTCTTGCAGGTGTACA GATCTGTGTAGAGGAATGTGTATATTTACCTCTTCGTTTGGTCTCAAACATGAGTGGGTATTTTTTGT TGTTTTTTTTGT TTTCTGTCTACTACAGCACCCGCTTCCAGGTTGAAGTGTATCTCTTGCCTCAGCCTCCCGAGTAGCTG GGATTACAGGTGCCACCACCGCGCCAGCTAATTTTTTAATTTTTAGTAGAGACAGGGTTTTACCATGT TGGCCAGGCTGGTCTTGAACCTCTGACCTCAAGTGATCTGCCACCTTGGCCTCCCTAAGTGCTGGGAT TATAGGCGTGAGCCACCATGCTCAGCGCTTAAGGTATTTTGTAAAGAACTTAAGTTTAGGTAAGGAAGA ATGAAATGATCCAGAAAAATGCAAGCAAGTCCACATGGAGATTGGAGGACACTGGTTAAAGAATTTAT TTCTTTGTATAGTATACTATGTTTATGGTGCAGATACTACAACATTTGTGGCATTTTAGACTCGTTGAGTT TCTTGGGCACTCCCAAGGGCGTTGGGGTCAATAAGGAGACTATAACTCTACAGATTGTGAATATATTTATT TTCAAGTTGCATTTTGTCTTTTGAAGCAATCAGATTTCAAGAGAGCTCAAGCTTTCAGAGTCAATGT GAAAATTCCTTCTTAGGCTGTCCACAGCTCTTGTGCTCCCTTAGATGAAGCCACTTGTTTCAAGATGACT ACTTTGGGTTGGGTTTTTCTATCTAAACACATTTTTCCAGTCTTATTAGATAAATTAGTCCATATGTTTGG TTAATCAAGAGCCTTCTGGGTTTGGTTTGGTGGCATTAAATGG	101
NM_031423	GCGGAATGGGGCGGACTTCCAGTAGGAGGCGGCAAGTTTGAAAAGTGAIGACGGTTGACGTTTGTGAT TTTTACTTTGCTTGTAGCTGCTCCCCGAAGTCCGCTCTTCTGTGCGCGGCGGCACTGTAGATTAAAC AGGAAACTTCCAAGATGGAACTTTGTCTTTCCCGAGATATAATGTAGCTGAGATTGTGATTCATATTCG CAATAAGATCTTAACAGGAGCTGATGGTAAAAACCTCACCAGAAATGATCTTTATCCAAATCCAAAGCCT	102

	GAAGTCCTTGCACATGATCTACATGAGAGCCTTACAAATAGTATATGGAATTCGACTGGAACATTTTTTACATGATGCCAGTGAACCTCTGAAGTCATGTAICCCACATTTAATGGAAGGCTTCTTACCATTACAGCAATTTAGTTACTCATCTGGACTCATTITTTGGCTATCTGCCGGGTGAATGACTTTGAGACTGCTGATATTTCTATGTCCAAAGCAAAAACGGACAAGTCGGTTTTTAAGTGGCATTATCAACTTTATTCACTTCAGAGAAGCATGCCGTGAAACGTATATGGAATTTCTTTGGCAATATAAATCCTCTGCGGACAAAATGCAACAGTTAAACGCCGCACACAGGAGGCATTAATGAACTGGAGAGACTTGATTCTGTTCCAGTTGAAGAGCAAGAAGAGTTCAAGCAGCTTTCAGATGGAATTCAGGAGCTACAACATCACTAAATCAGGATTTTCATCAAAAAACGATAGTGTCTGAAGAGGAAATTCCTCAAAAAGAAGTCAAAATATTTTTCAGAGAAAACCAAGCGTTGAATGAACATAAAATTTGTCGGTGGTTTTCTTTGAAAGAAATACAGAGAGTTTGAACAAAATTTGTGGATTCTCCAGAGAAATTGAAGAAATTATAAAGAAAAAATGAAAGATACGGTCCAGAGCTTAAAAATGCCAGACAAGAAGTGGTGGAGAAATATGAAATCTATGGAGACTCAGTTGACTGCCTGCCTTCATGTCAGTTGGAAGTGCAGTTATATCAAAAGAAATACAGGACCTTTTCAGATAATAGGAAAAATTAGCCAGTATCTTAAAGGAGAGCCTGAACCTTGGAGGACCAATTTGAGAGTGATGAGTCAGAACTGAAGAAATGAAGACTGAAGAAAATTCGTTCAAAAGACTGATGATTGTGAAGAAGGAAAACTTGCCACAGCACAAATCAAAATAAATAAGAAGCATGAAGATGTTAAGCAATACAACGCACAGTAATTGAGGATTGCAATAAAGTTCAAGAAAAAGAGGTGCTGTCTATGAACAGGATAACACAATTAATCAAGAAATCCAAAAATTAACCTTGAATTTCAACAATAAAAGATGCTGCTGAAAGGGAGAATCTGAAGTCCCGAGGAAATTTCTTAACTTGAACCTTGAAGAAATACCAGACAGGATTTGAAAAAGGCAGCAGAGGACTCCTATGCTAAGATAGATGAGAAGACAGCTGAACCTGAAGAGGAAGATGTTCAAAATGTCAACCTGATTAAACAAAATTACATGTCTTTTGTAAATGGCTTGCCATCTTTAATTTTCTATTTAGAAAAGAAAGTTGAAGCGAATGAAGTATCAGAAGTACCAATAATGTGGCTTCATCAGTTTTTATACACTCTCATAAGTAGTTAATAAGATGAATTTAATGTAGGCTTTTATTAATTTATAATTAATAATTTGTCAGACTATTTCATGCTCTACTCTGCCCTTGTGTGIAAATAGTTTGAGTAAAAACAACTAGTTACCTTTGAAATATATATATTTTTTCTGTTACTATC	
BC041846	GGCTAGCGCGGGAGGTGGAGAAAAGGGCTTGGGCGGCCCCGCTGTAGCCGCGTGTGGGAGGACGCACGGGCTCTGCTTCAAAGCTTTGGGATAACAGCGCCTCCGGGGGATAAATGAATGCGGAGCCTCCGTTTTTCAGTTCGCTTCAGATGTGCTCCACTTTTTTCCGCTGTAGCCGCAAGGCAAGGAAACATTTCTTCCCGTACTGAGGAGGCTGAGGAGTGCAGTGGGTGTTCTTTTCTCCTCTAACCCAGAACTGCGAGACAGAGGCTGAGTCCCTGTAAAGAACAGCTCCAGAAAAGCCAGGAGAGCGCAGGAGGGCATCCGGGAGGCCAGGAGGGGTTCGCTGGGCCCTCAACCGCACCCACATCGGTCCCACTGCGAGGGGGCGGGACCTCGTGGCGTGAACCAATCAGCACCCACCTGCGCTCACCTGGCCTCTCCTCCGCTGGCTCCCGGGGGCTGCGGTGCTCAAAGGGGCAAGAGCTGAGCGGAACACCGGCCCGCGCTCGCGGCAGCTGCTTCACCCCTCTCTCTGAGCCATGGGGCTCCCTCGTGACCTCTCGCGCTCTCTCCTCTCTCCAGTTTGTGGCTGCAGTGCAGCGGCCCTCCGAGCCGTGCCGGGCGGTCTCAGGGAGGCTGAAGTGACCTTGGAGGCGGGAGGCGCGGAGCAGGAGCCCGGCCAGGCGCTGGGAAAGTATTTCATGGGCTGCCCTGGGCAAGAGCCAGCTCTGTTTAGCACTGATAATGATGACTTCACTGTGCGGAATGGCGAGACAGTCCAGGAAAGAGGTCACTGAAGGAAAGGAATCCATTGAAGATCTTCCCATCCAAACGTATCTTACGAAGACACAAGAGAGATTGGGTGGTTGCTCCAAATATCTGTCCCTGAAAATGGCAAGGGTCCCTTCCCCCAGAGACTGAATCAGCTCAAGTCTAATAAAGATAGAGACACCAAGATTTTCTACAGCATCACGGGGCCGGGGCAGACAGCCCCCTGAGGGTGTCTTCTGCTGTAGAGAAGGAGACAGGCTGGTTGTGTGTAATAAGCCACTGGACCGGGAGGAGATTGCCAAGTATGAGCTCTTTGGCCACGCTGTGTGAGAGAATGGTGCCTCAGTGGAGGACCCATGAACATCTCCATCATAGTGACCGACCAGAAATGACCACAAGCCCAAGTTTACCCAGGACACCTTCCGAGGAGTGCTTTAGAGGGAGTCCCTACCAGTACTTCTGTGATGAGATGACAGCCACAGATGAGGATGATGCCATCTACACCTACAATGGGGTGGTTGCTTACTCCATCCATAGCCAAGAACCAAGGACCCACACGACCTCATGTTTCACAATTCACCGGAGCACAGGCACCATCAGCGTCATCTCCAGTGGCCTGGACCGGAAAAAGTCCCTGAGTACACACTGACCATCCAGGCCACAGACATGGATGGGGACGGCTCCACACCCACGGCAGTGGCAGTAGTGAGTACCTTGTATGCCAATGACAATGCTCCCATGTTTGAACCCCAAGTACGAGGCCCCATGTGCTGAGAAATGAGTGGGCCATGAGGTGCAGAGGCTGACGGTCACTGATCTGGACGCCCCCAACTCACCGCGTGGCGTGCCACCTACCTTATCATGGGCGGTGACGACGGGGACCATTTTACCATCACCCACCCCTGAGAGCAACAGGGCATCTTGACAACCAGGAAGGGTTTGGATTTTGAAGCCAAAAACAGCACACCTGTACGTTGAAGTGACCAACGAGGCCCCCTTTGTGCTGAAGCTCCCAACCTCCACAGCCCAATAGTGGTCCACGTGGAGGATGGAATGAGGCACCTGTGTTTGTCCACCCCTCCAAAGTCGTTGAGGTCCAGGAGGCATCCCCACTGGGGAGCCGTGTGTGTCTACACTGCAGAAGACCCTGACAAGGAGAAATCAAAGATCAGCTACCGCATCTCTGAGAGACCCAGCAGGGTGGCTAGCCATGGACCCAGACAGTGGGCAGGTCAAGCTGTGGGCACCTCGACCGTGAGGATGAGCAGTTTGTGAGGAACAACATCTATGAAGTCATGTGCTTTGGCCATGGACAATGGAAGCCCTCCCAACCTGAGCAGGGAAACCTTCTGCTAACACTGATTGATGTCAACGACCATGGCCAGTCCCTGAGCCCCGTGAGATCACCATCTGCAACCAAGCCCTGTGCGCCAGGTGCTGACATCACGGACAAGGACCTGTCTCCCCACACCTCCCCCTTCCAGGCCAGCTCACAGATGACTCAGACATCTACTGGACGGCAGAGGTCAACGAGGAAGGTGACACAGTGGTCTTGTCCCTGAAGAAGTTCCTGAAGCAGGATACATATGACGTGCACCTTTCTCTGTCTGACCATGGCAACAAAGAGCAGCTGACGGTGTATCAGGGCCACTGTGTGCGACTGCCATGGCCATGTGCAAACTTCCCTGGACCCCTGAAAGGAGGTTTTCATCCTCCTGTGCTGGGGGCTGTCTGTGGTCTGTCTTCTCCTGCTGGTGTGCTTTTGTGGTGGAGAAAGAGCGGAAGATCAAGAGAGCCCCCTCTACTTCCGAGAAGTACACCCGTGACAACGCTCTTCTACTATGGCGAAGAGGGGGTGGCGAAGAGGACAGGACTATGACATCACCCAGCTCCACCGAGGTCTGGAGGCCAGGCCGGAGGTGGTTCTCCGCAATGACGTGGCACCAACCATCTCCCGACACCATGTACCGTCTTAGGCCAGCCAAACCCAGATGAAATCGGCAACTTTATAATTGAGAACCTGAAGGCGGCTAACACAGACCCCAAGCCCCGCTTACGACACCTCTTGTGTTTCGACTATGAGGCGAGCGCTCCGACGCGCGTCCCTGAGCTCCCTCAGCTCCCTCCGCTCCCGACCAAGACCAAGATTACGATTATCTGAACGAGTGGGGCAGCCGCTTCAAGAAGCTGGCAGACATGTACGGTGGCGGGGAGGACACTAGGCGGCCCTGCCTGCAGGGCTGGGGACCAACGTCAGGCCACAGAGCATCTCCAAAGGGGCTCAGTTCCCCCTCAGCTGAGGACTTCGGAGCTGTGCAGGAAGTGGCCGTAGCAACT	103

	TGGCGGAGACAGGCTATGAGTCTGACGTTAGAGTGGTTGCTTCTTACGCTTTCAGGATGGAGGAATGTG GGCAGTTTACCTTCAGCACTGAAAACCTCTCCACCTGGGGCCAGGGTTGCCTCAGAGGCCAAGTTTCCAGA AGCCTCTTACCTGCGCTAAAATGCTCAACCTGTGTCTGGGCTGGGCTGCTGTGACTGACCTACAGT GGACTTTCTCTCTGGAATGGAACCTTCTTAGGCTCTCTGGTGCAACTTAATTTTTTTTTTAATGCTATC TTCAAAACGTTAGAGAAAGTTCTTCAAAAGTGACGCCAGAGCTGCTGGGCCCACTGGCCGCTCCTGCATT TCTGGTTTCCAGACCCCAATGCCCTCCCATTCGGATGGATCTCTGCGTTTTTATACTGAGTGTCCCTAGGT TGCCCTTATTTTTTATTTTCCCTGTGCGTTGCTATAGATGAAGGGTGAGGACAATCGTGTATATGTAC TAGAACTTTTTTATTAAGAAACTTTTCCCAAAAAAAAAAAAAA	
NM_016343	GAGACAGAAAGCGGCGAATTGGGCACCCGGTGGCGGCTGCGGGCAGTTTGAATTAGACTCTGGGCTCCAG CCCCCGAAGCCGCGCCAGAACTGTACTCTCCGAGAGGTCTGTTTTCCCGTCCCGAGAGCAAGTTTATTT ACAAATGTTGGAGTAATAAAGAAGGCAGAACAAAATGAGCTGGGCTTTGGAAGAAATGGAAGAAGGGCTG CCTACAAGAGCTCTTCAGAAAATTCAGAGCTTGAAGGACAGCTTGACAACTGAAGAAGGAAAAGCAGC AAAGGCAGTTTTCAGCTTGACAGTCTCGAGGCTGCGCTGCAGAAGCAAAAACAGAAGGTTGAAAATGAAAA AACCAGGGTACAAACCTGAAAAGGGAGAATCAAGATTGATGGAATAATGTGAAGTCTGGAGAAAACCT AAGCAGAAGATTTTCATGAACCTCAAGTCAAGGAGTCAAGAGTGAATTTCCAGGAAGGACAACCTGAATT CAGGCAAAAAACAAATAGAAAACTGGAACAGGAACCTTAAAGGGTGAATCTGAGCTTGAAGAAGCCCA ACAGCTGCGCAGTCTGCAGATGTCTCTGGAATCCATGCAATACACCACAAAAATTTTTACAACTCCA CTAACACCAAGTCAATATTATAGTGGTTCCAAGTATGAAGATCTAAAAGAAAAATATAATAAAGAGGTTG AAGAACGAAAAAGATTAGAGGCAGAGGTTAAAGCCTTGACAGGCTAAAAAAGCAAGCCAGACTCTTCCACA AGCCACCATGAATCACCAGCAGATTGCCCGGCATCAGGCTTCATCATCTGTGTTCTCATGGCAGCAAGAG AAGACCCCAAGTCATCTTTTCATCTAATTTCTCAAGAAGCTCCAATTAGGAGAGATTTCTCTGCTACTTACT TTTCTGGGGAACAAGAGGTGACTCCAAGTCATCAACTTTGCAAAATAGGGAAAAGAGATGCTAATAGCAG TTTCTTTGACAATTTCTAGCAGTCTCTCATCTTTTGATCAATTAAAGCGCAGAAATCAAGAGCTAAGAAAC AAGATTAAATGAGTTGGAACCTACGCTGCAAGGACATGAAAAAGAAATGAAAGGCCAAGTGAATAAGTTTC AAGAACTCCAACCTCCAACCTGGAGAAAGCAAAAGTGAATTAATGAAAAAGAGAAAGTTTGAACAAATG TAGGGATGAACCTAGTGAGAACACAGCACAATACGACCAGGCGTCAACCAAGTATACTGCATTGGAACAA AAACTGAAAAAATGACGGAAGATTGAGTTGTCAGCGACAAAATGCAGAAAGTGCAGATGTTCTCTGG AACAGAAAATTAAGGAAAAAGAAAGGAGTTTCAAGAGGAGCTCTCCCGTCAACAGCGTTCTTTCCAAAC ACTGGACAGGAGTGCATCCAGATGAAGGCCAGACTCACCAGGAGTTACAGCAAGCCAGAATATGCGAC AACGTCTCTGAGGCTGAACCTGGATAAACTCACATCAGTAAAGCAACAGCTAGAAAACAAATTTGGAAGAGT TTAAGCAAAAGTTGTGCAGAGCTGAACAGGCGTTCCAGGCGAGTCAGATCAAGGAGAATGAGCTGAGGAG AAGCATGGAGGAATGAAGAAGGAAAAACACCTCCTTAAGAGTCACTCTGAGCAAAAGGCCAGAGAAGTC TGCCACCTGGAGGCAGAACTCAAGAACATCAAAACAGTGTTTAAATCAGAGCCAGAATTTTCAGAAAGAAA TGAAAGCGAAGAATACCTCTCAGGAAACCATGTTAAGAGATCTTCAAGAAAAATAAATCAGCAAGAAAA CTCCTTGACTTTAGAAAACTGAAGCTTGCTGTGGCTGATCTGGAAGAGCAGCCAGATTGTTCTCAAGAC CTTTTGAGAAAAAGAGAACATCACATTGAACAACCTTAATGATAAGTTAAGCAAGACAGAGAAAGAGTCCA AAGCCTTGCTGAGTGTCTTAGAGTTAAAAAGAAAGAAATATGAAGAATGAAAGAGAGAAAGTCTGTT TTCTTGTTGGAAGTGAAGAACGAAAACTTTTAACTCAGATGGAATCAGAAAAGGAAAACTTGACAGT AAAATTAATCACTTGGAACCTTGCTGAAGACACAGCAAAATAAAAGTCAATGAATACAACGAGAGAGTAA GAACGCTGGAGATGGACAGAGAAAACTTAAGTGTGAGATCAGAAACCTTCACAACGTGTTAGACAGTAA GTCAAGTGGAGGTAGACAGCCAGAACTAGCTTATATGGAGCTACAGCAGAAAGCTGATCTCAGATCAG AAACATCAGAAGGAAATAGAAAATATGTGTTTGAAGACTTCTCAGCTTACTGGGCAAGTTGAAGATCTAG AACACAAGCTTCAGTTACTGTCAAATGAATAATGGACAAAGACCGGTGTACCAAGACTTGCATGCCGA ATATGAGAGCCTCAGGGATCTGCTAAAATCCAAGATGCTTCTCTGGTGACAAATGAAGATCATCAGAGA AGTCTTTTGGCTTTTGATCAGCAGCTGCCATGCATCATTTCTTTGCAAAATATAATTTGAGAGCAAGGAA GCATGCTTTCAGAGAGGAGTGAATGTGTTTGAAGCAGACCAAGTCCGAAAAATTTCTGCCATCTTACA AAATAGAGTTGATTCACTTGAATTTTCATTAGAGTCTCAAAAACAGATGAACCTCAGACCTGCAAAAGCAG TGIGAAGAGTTGGTGCAATCAAGGAGAAATAGAAGAAATCTCATGAAAGCAGAACAGATGCATCAAA GTTTTGTGGCTGAAACAAAGTCAGCGCATTAGTAAGTTACAGGAAGACACTTCTGCTCACCAGAATGTTGT TGCTGAAACCTTAAGTGCCCTTGAGAACAAGGAAAAAGAGCTGCAACTTTTTAAATGATAAGGTAGAACT GAGCAGGCAGAGATTCAAGAATTAAGAAAGAGCAACCATCTACTTGAAGACTCTCTAAAGGAGCTACAAC TTTTATCCGAAACCTTAAGCTTGGAGAAGAAAGAAATGAGTTCCATCATTTCTCTAAATAAAAGGGAAT TGAAGAGCTGACCCAGAGAAATGGGACTCTTAAGGAAATTAATGCATCCTTAAATCAAGAGAAGATGAAC TTAATCCAGAAAAGTGAGAGTTTTTGCAACTATATAGATGAAAGGGAGAAAAGCATTTCAGAGTTATCTG ATCAGTACAAGCAAGAAAACTTATTTTACTACAAAGATGTGAAGAAACCGGAAATGCATATGAGGATCT TAGTCAAAAAATACAAAGCAGCACAGGAAAAAGAAATCTAAATTAGAATGCTTGCTAAATGAATGCATAGT CTTTGTGAAAAATAGAAAAATGAGTTGGAACAGCTAAAGGAAGCATTTGCAAAAGGAACCAAGAAATCT TAACAAAATTAGCATTTGCTGAAGAAAGAAATCAGAACTGATGCTAGAGTTGGAGACAGTGCAGCAAGC TCTGAGATCTGAGATGACAGATAACCAAAACAATTTCTAAGAGCGAGGCTGGTGGTTTAAAGCAAGAAATC ATGACTTTTAAAGGAAGAACAAAAACAAATGCAAAAGGAAGTTAATGACTTATTACAAGAGAAATGAACAGC TGATGAAGTAAATGAAGCTAAACATGAATGTCAAAATCAGAATCAGAACCAATGGAACCTCTGTGAA AGAAAGAGAGAGTGAAGAGAAATCAATGTAATTTTAAACCTCAGATGGATCTTGAAGTTAAAGAAATTTCT CTAGATAGTTATAATGCGCAGTTGGTGCAATTAGAAGCTATGCTAAGAAATAAGGAATTAAGAACTTCAGG AAAGTGAAGAGGAGAGGAGTGCCTGCAGCATGAATTACAGACAATTAGAGGAGATCTTGAAGACAGCAA TTTGAAGACATGAGTCAAGAAATTAGTGGCTTAAAGACTGTGAAATAGATGCGGAAGAAAGTAT ATTTCAAGGCTCATGAGTTGTCAACAAGTCAAAACGACAATGCACACCTTCAGTGCTCTCTGCAAAACAA CAATGAACAAGCTGAATGAGCTAGAGAAAAATGTGAAATCTGAGGCTGAAAAGTATGAACTCGTAAC TGAGCTGAATGATTCAAGGTCAGAAATGTATCACAGCAACTAGGAAAAATGGCAGAAGAGGTAGGGAACCTA	104

CTAAATGAAGTTAAATATTTAAATCATGACAGTGGTCTTCTCCATGGTGAGTTAGTGAAGACATACCAG
 GAGGTGAATTTGGTGAACAACCAAAATGAACAGCACCTTGTGTCTTTGGCTCCATTGGACGAGAGTAATTC
 CTACGAGCACTTGACATTGTGACAGACAAAGAGTTCAAATGCACCTTGGCCGAATGCAAGAGAAATTCCTTA
 TCTTTACAAAGTGAACACAAAATTTTACATGATCAGCACTGTGATGAGCTCTAAAATGTCAGAGCTGC
 AGACCTATGTTGACTCATTTAAAGGCCGAAAATTTGGTCTTGTCAACGAATCTGAGAACTTTCAAGGTGA
 CTTGGTGAAGGAGATGCAGCTGGGCTTGGAGGAGGGGCTCGTTCCATCCCTGTATCCTCTTGTGTGCCCT
 GACAGCTCTAGTCTTAGCAGTTTGGGAGACTCCTCCTTTTACAGAGCTCTTTTAGAACAGACAGGAGATA
 TGICTCTTTTGTAGTAATTTAGAAGGGGCTGTTTTCAGCAAAACAGTGCAGTGTAGATGAAGTATTTTGCAG
 CAGCTCGCAGGAGGAGAATCTGACCAGGAAAGAAACCCCTTCGGCCCCAGCGAAGGGTGTGGAAGACTT
 GAGTCCCTCTGTGAGGTGTACCGGCAGTCCCTCGAGAAGCTAGAAGAGAAAAATGGAAGTCAAGGGATTA
 TGAAAAATAAGGAAATTCAGAGCTCGAGCAGTTATTAAGTTCTGAAAGGCAAGAGCTTGACTGCCTTAG
 GAAGCAGTATTTGTCAGAAAATGAACAGTGGCAACAGAGCTGACAAGCGTGACTCTGGAGATGGAGTCC
 AAGTTGGCGCAGAAAAAGAAACAGACGGAACACTGTCACTTGAGCTGGAAGTAGACACTCCAGCTAC
 AAGGTCIGGACTTAAGTTCTCGGCTTTTGGCTTGGCATCGACACAGAAGATGCTATTCAAGGCCGAAATGA
 GAGCTGTGACATATCAAAAAGACATACTTCAGAACTACAGAAAGAACACCAAGCATGATGTTCTACAG
 ATTTGTGATAAAGATGCTCAGCAGGACCICAATCTAGACATTGAGAAAAATACTGAGACTGGTGCAGTGA
 AACCCACAGGAGAGTGTCTTGGGGAACAGTCCCCAGATACCAATTATGAGCCTCCAGGGGAAGATAAAAC
 CCAGGGCTCTTCAGAAATGCATTTCTGAATGTCTATTTTCTGGTCTTAATGCTTTGGTACCTATGGATTTCT
 CTGGGGAATCAGGAAGATATCCATAATCTTCAACTGCGGGTAAAAAGAGACATCAAATGAGAAATTTGAGAT
 TACTTCATGTGATAGAGGACCGTGACAGAAAAGTTGAAAGTTTGTCTAAATGAAATGAAAGAAATTAGACTC
 AAAACTCCATTTACAGGAGGTACAACTAATGACCAAAAATGAAGCATGCATAGAATTTGAAAAAATAGTT
 GGGGAACTTAAGAAAAGAAAACCTCAGATTTAAGTGA AAAAATTTGGAATATTTTCTTGTGATCACCAGGAGT
 TACTCCAGAGAGTAGAAAACCTTCTGAAGGCCCTCAATCTGATTTAGAAATGCATGCAGATAAATCATCACG
 TGAAGATATTGGAGATAATGTGGCCAAGGTGAATGACAGCTGGAAGGAGAGATTTCTTGATGTGAAAAAT
 GAGCTGAGTAGGATCAGATCGGAGAAAAGCTAGCATTGAGCATGAAGCCCTCTACCTGAGGCTGACTTAG
 AGGTAGTTCAAACAGAGAAGCTATGTTTAAAAAGACAAATGAAAAAAGCAGAGGTTATTGTCTGCCCT
 TGAGAAGAACTCTCAGTGGTCACAAGTGAGAGAAACAGCTTCGTGGAGAATTAGATACTATGTCAAAA
 AAAACACGGCCTGGATCAGTTGTCTGAAAAAATGAAGGAGAAAAACACAAGAGCTTGAGTCTCATCAAA
 GTGAGTGTCTCCATTGCATTGAGTTGAGTGGCAGAGGAGGAGTGAAGGAAAAAGCAGGAACTTTCAGACTTT
 GTCCTCTGATGTGAGTGAGCTGTTAAAAGACAAAACCTCATCTCCAGGAAAAGCTGCAGAGTTTGGAAAAAG
 GACTCACAGGCCTGTCTTTGACAAAAATGTGAGCTGGA AAAACCAATTTGCACAACCTGAATAAGAGAAAAAG
 AATTGCTTTGTCAAGGAATCTGAAAGCCTGCAGGCCAGACTGAGTGAATCAGATTATGAAAAGCTGAATGT
 CTCCAAGGCCCTTGGAGGCCGCACTGGTGGAGAAAAGGTGAGTTTCGCAATTGAGGCTCAGGCTCAACACAGGAG
 GAAGTGCATCAGCTGAGAAGAGGCATCGAGAACTGAGAGTTTCGCAATTGAGGCCGATGAAAAGAAGCAGC
 TGCACATCGCAGAGAACTGAAAGAACCGCAGCGGAGAGATGATTCACTTAAGGATAAAGTTGAGAACCT
 TGAAGGGGAATTGCAGATGTCAGAAGAAAACAGGAGCTAGTGATTCTTGATGCCGAGAATTCCAAAGCA
 GAAGTAGAGACTCTAAAAACACAAATAGAAGAGATGGCCAGAAGCTGAAAGTTTTTGAATTAGACCTTGT
 TCAGCTTAAGGCTCTGAAAAAGAAAATCTGACAAAACAAATACAAGAAAAACAAGGTGAGTTGTGAGAACT
 AGACAAGTTACTCTCTTCAATTTAAAGTCTGTTAGAAGAAAAGGAGCAAGCAGAGATACAGATCAAGAGAA
 GAATCTAAAACTGCAGTGGAGATGCTTCAGAACTCAGTTAAAGGAGCTAAATGAGGCAGTAGCAGCCTTGT
 GTGGTGACCAAGAAATATGAAGGCCACAGAACAGAGTCTAGACCCACCAATAGAGAACTCAGCTCAGCT
 GAGAAATAGCATTGAAAAGCTGAGAGCCCGCTTAGAAGCTGATGAAAAGAAGCAGCTCTGTGTCTTACAA
 CAACTGAAGGAAAGTGAGCATCATGCAGATTTACTTAAGGTAGAGTGGAGAACCTTGAAAAGAGAGCTAG
 AGATAGCCAGGACAAACCAAGAGCATGCAGCTCTTGAGGCAGAGAAATCCAAAGGAGAGGTAGAGACCTT
 AAAAGCAAAAATAGAAGGGATGACCCAAAGTCTGAGAGGTCTGGAATTAGATGTTTGTACTATAAAGTCA
 GAAAAAGAAAATCTGACAAATGAATTACAAAAAGAGCAAGAGCGAATATCTGAATTAGAAATATAAATTT
 CATCATTTGAAAAATTTTGCAGAAAAAGAGCAAGAGAAAGTACAGATGAAAGAAAAATCAAGCACTGC
 CATGGAGATGCTTCAAACACAATTTAAAGAGCTCAATGAGAGAGTGGCAGCCCTGCATAATGACCAAGAA
 GCCTGTAAAGGCCAAAGAGCAGAACTTAGTAGTCAAGTAGAGTGTCTTGAACCTTGAGAAGGCTCAGTTGC
 TACAAGGCCCTTGATGAGGCCAAAAATAATTATATTGTTTTGCAACTTTCAGTGAATGGCCTCATTCAAGA
 AGTAGAAGATGGCAAGCAGAACTGGAGAGAAAGGATGAAGAAATCAGTAGACTGAAAAATCAAATTCAA
 GACCAAGAGCAGCTTGTCTCTAACTGTCCAGGTGGAAGGAGAGCACCAACTTTGGAAGGAGCAAAAACCT
 TAGAACTGAGAAATCTGACAGTGGAAATGGAGCAGAAAGATCCAAAGTGTCTACAATCCAAAAATGCCTCTTT
 GCAGGACACATTAGAAGTGTGTCAGAGTTCTTACAAGAACTAGAGAATGAGCTTGAATTGACAAAAATG
 GACAAAAATGTCCTTTGTTGAAAAAGTAAACAAAATGACTGCAAGGAAACTGAGCTGCAGAGGGAAATGC
 ATGAGATGGCACAGAAAACAGCAGAGCTGCAAGAGAACCTCAGTGGAGAGAAAAATAGGCTAGCTGGAGA
 GTTGCAGTTACTGTTGGAAGAAAATAAGAGAGCAGCAAGATCAATTTGAAGGAGCTCAGCACTAGAAAAATG
 GAATTGAAGAAGAGCCTAGATTGCATGCACAAAGACCAGGTGGAAGAGGAAGGGGAAAGTGAAGAGGAAA
 TAGCTGAATATCAGCTACGGCTTCATGAAGCTGAAAAGAAACACCAGGCTTTGCTTTTGGACACAACAA
 ACAGTATGAAGTAGAAATCCAGACATACCGAGAGAAATTTGACTTCTAAAGAAGAAATGCTCTCAGTTTACAG
 AAGCTGGAGATAGACCTTTTAAAGTCTAGTAAAGAGAGCTCAATTAATTCATTGAAGACTACTACTCAGA
 TTTTGGAGAATTGAAGAAAACCAAGATGGACAATCTAAAAATATGTAATCAGTTGAAGAAGGAAAAATGA
 ACCTGCCCCAGGGGAAAAATGAAGTTGTGATCAAAATCCTGTAAACAGCTGGAAGAGGAAAAAGGAGATACTG
 CAGAAAGAACTCTCTCAACTTCAAGCTGCACAGGAGAAGCAGAAAAACAGGTACTGTTATGGATACCAAGG
 TCCGATGAATTAACAACCTGAGATCAAGAACTGAAAGAACTCTTGAAGAAAAACCAAGGAGGAGATGA
 ATACTTGGATAAGTACTGTTCTTGTCTTATAAGCCATGAAAAGTTAGAGAAAGCTAAAGAGATGTTAGAG
 ACACAAGTGGCCCATCTGTGTTTACAGCAATCTAAACAAGATTCCTGAGGGTCTCTCTTTGCTAGGTTCCAG
 TTGTTCCAGGACCATCTCCAATCCCTTCTGTTTACTGAAAAGAGGTTATCATCTGGCCAAAATAAAGCTTC

	AGGCAAGAGGCAAGATCCAGTGGAAATATGGGAGAATGGTAGAGGACCAACACCTGCTACCCCAGAGAGC TTTTCTAAAAAAGCAAGAAAGCAGTCAIGAGTGGTATTACCCCTGCAGAAACACGGAAGGTACTGAGT TTGAGCCAGAGGGACTTCCAGAAAGTTGTAAAGAAAGGGTTTGGCTGACATCCCGACAGGAAAGACTAGCCC ATATATCCTGCGAAGAACAACCATGGCAACTCGGACCAGCCCCGCTGGCTGCACAGAAGTTAGCGCTA TCCCCACTGAGTCTCGGCAAGAAAAATCTTGACAGAGTCCCTCCAAACCAACAGCTGGTGGCAGCAGATCAC AAAAGGTCAAAGTTGCTCAGCGGAGCCAGTAGATTGAGGCACCATCCTCCGAGAACCCACCACGAAATC CGTCCCAGTCAATAATCTTCTGAGAGAAGTCCGACTGACAGCCCCAGAGAGGGGCTGAGGGTCAAGCGA GGCCGACTTGTCCCCAGCCCCAAGCTGGACTGGAGTCCAAACGGCAGTGAGAAGTGAAGGTTCCAGTGAA GGCAGCTTTGTGTGTCAGTACCCCTGGGAGGTGCCAGTCAATTGAATAGATAAGGCTGTGCGCTACAGGACTT CTCTTTAGTCAGGGCATGCTTTATTAGTGAGGAGAAAAACAATTCCTTAGAAGTCTTAAATATATTGTA CTTTAGATCTCCCATGTGTAGGTATTGAAAAAGTTTGAAGCAGTGTATCAGTGTAGCATTTGCCATTCC TCTACTGCAATGTAATAGTATATAAGCTATGTATATAAAGCTTTTTGGTAATATGTTACAATTAATAATGA CAAGCAGTATATACAAATCTCTGTTTGTATGTGGGTTTTTACACTAAAAAATGCAAAACACATTTTATTC TTCTAATTAACAGCTCCTAGGAAAAATGTAGACTTTTGTCTTATGATATTCTATCTGTAGTATGAGGCATG GAATAGTTTTGTATCGGGAATTTCTCAGAGCTGAGTAAAAATGAAGGAAAAAGCATGTTATGTGTTTTTAAG GAAAAATGTGCACACATATACATGTAGGAGTGTATTCTTTCTCTTACAATCTGTTTTAGACATCTTTTGTCT TATGAAACCTGTACATATGTGTGTGGTATGTGTTTTATTCCAGTGAGGGCTGAGGCTTCCAGGAGG TGIGCTATACCATGCGTCTGTGCTGTGTTTTCTGTTTTTAGACCAATTTTTACAGTCTTTTGGTA AGCATTGTGCTATCTGGTGTGATTAACATATAGCCTTTGTTTTCTAATAAAATAGTCGCTTCGTTTT CTGTAAAAAAGAAAAAAGAAAAA	
AB091343	GGCAGAGGGGCCGACGCGAGCGCGCTTCGCTTCAGCTGCTAGCTGGCCCCAAGGAGGGCAGCCGCGG AGGGTGGCGAGGGGCGGCCAGGACCCGAGCCCCGGGCGCGGCCGCTCCGGACCGCCAGGGAGGGCAGG TCAGTGGGCGAGATCGCGTCCGCGGGATTCAATCTCTGCCGCTCTGATAACAGTCCCTTTCCCTGGCGCT CACTTCGTGCTGGCAGCCCGGCTGGGCGCTCAAGACCGTTGTCTCTCGATCGCTTCTTTGGACTTGGC GACCATTTAGAGATGTCTTCCAGAAGTACCAAGATTTAATTAAGTAAGTGGGAGTGAAGGCTTAGT AACTCCAAATCCGAAACTACATTAGAAAAATTAAGGGAGAAATTCACACTTAAAGACATCAGTGGATG AAATCACAAAGTGGGAAAGGAAAGCTGACTGATAAAGAGAGACACAGACTTTTGGAGAAAAATCGAGTCC TGAGGCTGAGAAGGAGAAGAAATGCTTATCAACTCACAGAGAAGGACAAAGAAATACAGCGACTGAGAGAC CAACTGAAGGCCAGATATAGTACTACCGCATTTGCTTGAACAGCTGGAAGAGACAACGAGAGAGAGGAAAA GGAGGGAGCAGGTGTTGAAAGCCTTATCTGAAGAGAAAGCGTATTGAAACAACAGTTGTCTGCTGAAC CTCAGCAATTGCTGAACCTTGAAGCAAAACCAATACACTCCGTTTTATCAGAGACTGTGGCTCCAAACTGC TTCAACTCATCAATAAATAATATTCATGAAATGGAATACAGCTGAAAGATGCTCTGGAGAAAAATCAGC AGTGGCTCGTGATGATCAGCAGCGGGAAGTCTATGTAAAAGGACTTTTAGCAAAAGATCTTTGAGTTGGA AAAGAAAAACGGAACAGCTGCTCATTTCACTCCACAGCAGACAAAAAGCCTGAATCAGAAGGTTATCTT CAAGAAGAGAAGCAGAAATGTTACAACGATCTCTTGGCAAGTGCAGAAAAAGATCTTTGAGTTGAACGAC AAACCATAACTCAGCTGAGTTTTGAACTGAGTGAATTTGGAAGAAATATGAAGAAACCCAAAAAGAGT TCACCAATTTAAATCAGCTGTTGATTTACAAAGAAAGGGCAGATGTGCAACATCTGGAAGATGACGAGAT AAAAACAGAGAAGATACAAAAACTCAGGGAAGAGAAATGATATTGCTAGGGGAAAACTTGAAGAAGAGAAGA AGAGATCCGAAGAGCTCTTATCTCAGGTCCAGTTTTCTTTACACATCTCTGCTAAAGCAGCAGAAAGAAC AACAAAGGGTAGCTCTGTTGGAACAACAGATGCAGGCATGTACTTTAGACTTTGAAAAATGAAAACTCGAC CGTCAACATGTGCAAGCATCAATTCGATGTAATCTTAAGGAGCTCCGAAAAAGCAAGAAATCTTTGAGTTGGA AGTTGGAATCCTTGAAACAGCTTCATGAGTTTGCCATCACAGAGCCATTAGTCACTTTCCAAGGAGAGAC TGAAAAACAGAGAAAAAGTTGCCGCTCACCAAAAAGTCCCACTGCTGCACTCAATGAAAGCCTGGTGGAA TGTCCTCAAGTGCAATATACAGTATCCAGCCACTGAGCATCGCGAICTGCTTGTCCATGTGGAATACGTGT CAAGTAGCAAAATAAGTATTTGTTTGTATTTAAAGATTCAATACTGATTTTTCTGTAGCTTTGTGGG CATTTTGAATTATATATTTTACATTTTGCATAAACTGCCTATCTACCTTTGACACTCCAGCATGCTAGT GAATCATGTATCTTTTAGGCTGCTGTGCATTTCTCTTGGCAGTGATACCTCCCTGACATGTTTTCATCATC AGGCTGCAATGACAGAAATGTTGTGAGCAGCGTCTACTGAGACTACTAACATTTTGCAGTGTCAAAATACT TGGTGAGGAAAAAGATAGCTCAGGTTATGCTAATGGGTTAATGCACCAGCAAGCAAAATATTTTATGTTT TGGGGGTTTGAATAAATAAGATAAATAAACAAGGATCTTAACTGTGTTTCGCAATTTTTATCCAAGCACT TAGAAAACTTACAATCCTAATTTTGTATGTCATTTGTTAAGAGGTGGTGATAGATACTATTTTTTTTTTCA TATTGTATAGCGGTTATTAGAAAAGTTGGGGATTTTCTTGATCTTTATTGCTGCTTACCATTGAAACTTA ACCCAGCTGTGTTCCCAACTCTGTTCTGCGCACGAAACAGTATCTGTTTGAGGCATAATCTTAAAGTGGC CACACACAATGTTTTCTCTTATGTTATCTGGCAGTAACTGTAACCTGAATTACATTAGCACATTTCTGCTT AGCTAAAAATGTTAAAAATAAACTTTAATAAACCCATGTAGCCCTCTCATTGATTGACAGTATTTTAGTT ATTTTTGGCATTTCTAAAGCTGGCAATGTAATGATCAGATCTTTGTTTGTCTGAACAGGTATTTTTATA CATGCTTTTTGTAAACCAAAAACTTTTAAATTTCTTCAGGTTTTCTAACATGCTTTACCCTGGGCTACTG TAAATGAGAAAAAGATAAAATATTTTAAAGTTTTTAAAAAAGAAAAA	105
BC006428	GGCGGCTGAGCCTGAGCGGGGATGTAGAGGCGGCGGCAGCAGAGCGGCAGTGGCGGCAAGAGCAGACGC CCGAGCCGAGCGAGAAGAGCGGCAGAGCCTTATCCCTGAAGCGGGGCCCCGCTCCCGCCCTGCCAGC CCGCGCCAGCCATGCGCGCGGCTGCTGAGTCCGGGCGCGCAGCTGAGCCCTCGCGCCGCGCAGCGG CGCTCAGCTCGGGGGTATTAGTTGCTTTTTGTTGTTTTTAAATTTGGGCGCGGGGAGGGGAGGAGGG GCAGGTGCTGCAGGCTCCCCCCCCCTCCCGCCTCGGGCCAGCCGCGGCGGCGGCTCGGGCTCCCGACC CGGGCACTGCTGGCGGCTGGAGCGGAGCGCACCAGCGGCGGTTGGTGCCAGAGCGGAGCGCAGCTCCCTGC CCGCCCCCTCCCCCTCGCGGCTCGCGGCGAGCGGCGGCGGTTGGCGGCTTGGACGACTCGGAGAGCCAGTGA AGACATTTCCACCTGGACACCTGACCATGTGCCTGCCCTGAGCAGCGAGGCCCCACAGGCATCTCTGTTG TGGGCAGCAGGGCCAGGTCCTGTGCTGTGGACCCCTCGGCAGTTGGCAGGCTCCCTCTGAGTGGGGTCTG GGCCTCGGCCCCACCAITGCGAGCCTCGGCGGTGGCTCCAGGATGCGCGGCGCAGTAGCAGCAGCAGCA	106

21

	AGATCACAGATTTTGGGCTGGCCAAACTGCTGGGTGCGGAAGAGAAAAGAATACCATGCAGAAGGAGGCCAA AGTGCCATCAAGTGGATGGCATTTGGAATCAATTTTACACAGAATCTATACCCACCAGAGTGATGCTGG AGCTACGGGGTGACCGTTTGGGAGTTGATGACCTTTGGATCCCAAGCCATATGACGGAATCCCTGCCAGCG AGATCTCCTCCATCCTGGAGAAAGGAGAACGCTCCCTCAGCCACCCATATGTACCATCGATGCTACAT GATCATGGTCAAGTGTGGATGATAGACGCAGATAGTCGCCCAAAGTTCCGTGAGTTGATCATCGAATTC TCCAAATGGCCGAGACCCCCAGCGCTACCTTGTCAATTCAGGGGATGAAAGAATGCATTTGCCAAGTC CTACAGACTCCAACCTTCTACCGTGCCCTGATGGATGAAGAAGACATGGACGACGTGGTGGATGCCGACGA GTACCTCATCCCACAGCAGGGCTTCTCAGCAGCCCCCTCCACGTCACGGACTCCCTCTGTAGCTCTCTG AGTGCAACCAGCAACAATTCACACCGTGGCTTGCATTGATAGAAATGGGCTGCAAGAGCTGCCATCAAGG AAGACAGCTTCTTGCAGCGATACAGCTCAGACCCACAGGCGCTTGACTGAGGACAGCATAGACGACAC CTTCTCCAGTGCCTGAATACATAAACAGTCCGTTCCCAAAGGCCCGCTGGCTCTGTGCAGAATCCT GTCTATCAAAATCAGCCTCTGAACCCCGCGCCAGCAGAGACCCACACTACCAGGACCCACAGCAGCTG CAGTGGGCAACCCGAGTATCTCAACACTGTCCAGCCCACTGTGTCAACAGCACATTCGACAGCCCTGC CCACTGGGCCCAGAAAGGCAGCCACCAATTAGCCTGGACAACCCCTGACTACCAGCAGGACTTCTTTCCC AAGGAAGCCAAGCCAAATGGCATCTTAAAGGGCTCCACAGCTGAAATGCAGAATACCTAAGGGTCGCGC CACAAGCAGTGAATTTATTGGAGCATGACCAGGAGGATAGTATGAGCCCTAAAAATCCAGACTCTTTC GATCCCAAGGACCAAGCCACAGAGGTCCCTCCATCCCAACAGCCATGCCCGCATTTAGCTCTTACCAAC AGACTGGTTTGCACGTTTACACCGACTAGCCAGGAAGTACTTCCACCTCGGGCACATTTTGGGAAGTT GCATTCCTTTGTCTTCAAACCTGTGAAGCATTTACAGAAACGCATCCAGCAGAATATTGTCCCTTTGAGC AGAAATTTATCTTTCAAAGAGGTATATTTGAAAAAAGTATATGTGAGGATTTTATTGATTGG GGATCTTGGAGTTTTTCATTGTGCTATTTGATTTTACTTCAATGGGCTCTTCCAAACAGGAAGACTT GCTGGTAGCACTTGTACCTGAGTTCAICCAGGCCCCAAGTGTGAGCAAGGAGCACAAGCCACAAGTCTT CCAGAGGATGCTTATTCCAGTGGTTCTGCTTCAAGGCTTCCACTGCAAAACACTAAAGATCCAAAGAGG CCITCAITGGCCCCAGCAGGCGGATCGGTACTGTATCAAGTCAATGGCAGGTACAGTAGGATAAGCCACTC TGTCCTTCTGGGCAAAAGAAACAGGAGGAGTGAATTTCTTCTTAGACTTACTTTTGTAAAGATGT CCCCACGGTACTTACTCCCACTGATGGACAGTGGTTTCCAGTCATGAGCGTTAGACTGACTTGTTTGT CTTCCATTCCATTGTTTTGAAACTCAGTATGCTGCCCTGCTTGTGTGATGAATCAGCAAGAGAGGA TGACACATCAAAATAAATACTCGGATTCAGCCACATTGGATTTCATCAGCATTTGGACCAATAGCCACA GCTGAGAATGTGGAATACCTAAGGATAGCACCGCTTTTGTCTCGCAAAACGATCTCTGTAATTTGAGG CTCAGATGAAATGCATCAGGTCTTTGGGGCATAGATCAGAAGACTACAAAATGAAGTGTCTGTGAAT CTCCTTTAGCCATCACCCCAACCCCAAAATTAGTTTGTGTACTTATGGAAGATAGTTTTCTCCTTTT ACTTCACTTCAAAGCTTTTTTACICAAAGAGTATATGTTCCCTCCAGGTCAGTGGCCCCAAACCCCTC CTTACGCTTTGTACACAAAAAGTGTCTCTGCTTGAATCATCTTCAAGCACTTACAGCTCTGGCCAC AACAGGGCATTTTACAGGTGCGAATGACAGTAGCATTATGAGTAGTGTGGAATTCAGGTAGTAAATATGA AACTAGGGTTTGAATTTGATAATGCTTTCACAACATTTGCAGATGTTTTAGAAGGAAAAAGTTCTTCC TAAATAAATTTCTCTACAATTGGAAGATTGGAAGATTAGCTAGTTAGGAGCCACCTTTTTCTTAATC TGTGTGTGCCCTGTAACTGACTGTGTTAACAGCAGTCTTGTGTAACAGTGTTTTTAACTCTGCTACTCA ATATCCACCCCATCAATTTATCAAGGAAGAAATGGTTTCAAAAAATATTTTACGCTACAGTTATGTCA GTCACACACACATACAAATGTTCTTTTGTCTTTTAAAGTAATTTTGAATCCAGATCAGTCAGAGCCC CTACAGCATTTGTTAAGAAAGTATTTGATTTTGTCTCAATGAAAAATAAACTATATTCATTTCCACTCTA AAAAA	
NM_00100586 2	GTTCCCGGATTTTGTGGGCGCCTGCCCGCCCCCTCGTCCCCCTGCTGTGTCCATATATCGAGGCGATAG GGTTAAGGGAAGGCGGACGCTGATGGGTAAATGAGCAAACTGAAGTGTTTTCCATGATCTTTTGTGAGT CGCAATTTGAAGTACCACCTCCCGAGGGTGATTGCTTCCCCATGCGGGGTAGAACCCTTTGCTGTCTGTTC ACCACCTACCTCCAGCAGCAAGATTTGGCTTATGCTTACTCAATGTGAAGATGATGAGGATGAAAACTT TGTGATGATCCACTTCCACTTAATGAATGGTGGCAAAGCAAAGCTATATTTCAAGACCACATGCAAAAGCTA CTCCCTGAGCAAAAGAGTCAAGATAAAACGGGGGACCAGTAGAATGGCCAGGACAAACGAGTGCAGCA CAGAGACTCAGACCTTGGCAGCCATGCCCTGCGCAGGCAGTATGAGAGTGACATGTACTGTTGTGGACAT GCACAAAAGTGAAGTGTGCAACCGGCACAGCATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACCCACCTGG ACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGGAAACCTGGAACCTACCTACCTGCCAC CAATGCCAGCCTGTCTTCTGCAAGGATATCCAGGAGGTGCAGGGCTACGTGCTCATCGCTCACAACCAA GTGAGGCAGGTCCCACTGCAGAGGCTGCGGATTGTGCGAGGCACCCAGCTCTTTGAGGACAACATATGCCC TGGCCGTGTGATAGCAATGAGAGCCGCTGAACAATACCACCCCTGTACAGGGGGCTCCCGAGGAGGCTT GCGGGAGCTGCAGCTTCAAGGCTCAGAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAACCCCCAG CTCTGCTACCCAGGACACGATTTTGTGGAAGGACATCTTCCACAAGAAACACAGCTGGCTCTCAGACTGA TAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGAAGGGCTCCCGCTGCTGGGGAGA GAGTTCTGAGGATTTGTGAGAGCCTGACGCGCATGTCTGTGCGGTGGCTGTGCCGCTGCAAGGGGCCA CTGCCACTGACTGCTGCCATGAGCAGTGTGCTGCGGCTGCACGGGCCCAAGCACTCTGACTGCCTGG CCTGCTTCACTTCAACCACAGTGGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACAACACAGA CACGTTTGTAGTCCATGCCCAATCCCGAGGGCCGGTATACATTCGGCGCCAGCTGTGTGACTGCCTGTCCC TACAACCTACCTTTTACGAGCTGCGGATCTGCAACCTGCTGTGCCCTGCAACCAAGAGGTGACAG CAGAGGATGGAACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGTGCCGAGTGTGCTATGGTCTGGGCAT GGAGCACTTGCAGAGGTGAGGGCAGTTACCAAGTGCCTATCCAGGAGTTTGTGGCTGCAAGAAAGATC TTTGGGAGCTGGCATTTCTGCCGAGAGCTTTGATGGGGACCCAGCCCTCAACACTGCCCGCTCCAGC CAGAGCAGCTCCAAGTGTGTGAGACTCTGGAAGAGATCAGAGTTACCTATACATCTCAGCATGTGCCGGA CAGCCTGCTGACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGGACGAATTTCTGCACATGGCGCC TACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTGCGCTCACTGAGGGAACCTGGGCAGTG GACTGGCCCTCATCCACCATAACACCCACCTCTGCTTGTGTCACAGGTGCCCTGGGACAGCTCTTTCG	108

	GAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCAGAGGACGAGTGTGTGGGCGAGGGCCTGGCC TGCCACCAGCTGTGCGCCCGAGGGCAGTGTGGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCCAGT TCCTTCGGGGCCAGGAGTGCCTGGAGGAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGC CAGGCACTGTTTGGCCTGCCACCCTGAGTGTGAGCCCCAGAATGGCTCAGTGACCTGTTTTGGACCGGAG GCTGACCACTGTGTGGCCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCCCGCTGCCCGAGCGGTG TGAAACCTGACCTCTCCTACATGCCCATCTGGAAGTTTCCAGATGAGGAGGGCGCATGCCAGCCTTGCCC CATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAGGGCTGCCCGCCGAGCAGAGAGCCAGCCCT CTGACGTCATCACTCTGCGGTGGTTGGCATTCTGCTGGTCTGCTGGTCTTGGGGTGGTCTTTGGGATCC TCATCAAGCGACGGCAGCAGAAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAACGGGAGCTGGT GGAGCCGCTGACACCTAGCGGAGCGATGCCCAACCAGGCGCAGATGCGGATCCTGAAAGAGACGGAGCTG AGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCACAGTCTACAAGGGCATCTGGATCCTGATGGGG AGAATGTGAAAATTCAGTGGCCATCAAAGTGTGAGGGAAAACACATCCCCAAAGCCAAACAAAGAAAT CTTAGACGAAGCATACGTGATGGCTGGTGGGCTCCCCATATGTCTCCCGCCTTGTGGCATCTGCGCTG ACATCCACGGTGCAGCTGGTGACACAGCTATGCCCTATGGCTGCCCTCTTAGACCATGTCCGGGAAAAACC CGGACGCTCGGGCTCCCCAGGACCTGCTGAACCTGGTGTATGCAGATTGCCAAGGGGATGAGCTACCTGGA GGATGTGCGGCTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAAA ATTACAGACTTCCGGCTGGCTGGCTGGGACATTGACGAGACAGAGTACCATGAGCAACGGGAGCAAGG TGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGCCGGCGGTTACCCACCAGAGTGTGTGTGGAG TTATGGTGTGACTGTGTGGGAGCTGATGACTTTTGGGGCCAAACCTTACGATGGGATCCCCAGCCCGGAG ATCCCTGACCTGCTGGAAAAGGGGAGCGGCTGCCCGAGCCCCCATCTGCACCATGTGATGTCTACATGA TCATGGTCAAATGTGGATGATTGACTCTGAATGTGCGGCAAGATTCCGGGAGTTGGTGTCTGAATCTC CCGCTATGGCAGGACCCCCAGCGCTTTGTGGTCAATCCAGAATGAGGACTTGGGGCCAGCCAGTCCCTTG GACAGCACCTTCTACCGCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATC TGGTACCCAGCAGGGCTTCTTCTGTCCAGACCCTGCCCGGGCGCTGGGGGATGGTCCACCACAGGCA CCGCAGCTCATCTACAGGAGTGGCGGTGGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGGGCC CCCAGGTCTCCACTGGCACCCCTCCGAAGGGGCTGGCTCCGATGTATTTGATGGTGACCTGGGAATGGGG CAGCCAAGGGGCTGCAAAGCCTCCCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAAGACCCAC AGTACCCCTGCCCTCTGAGACTGATGGCTACGTTGCCCCCTGACCTGCAGCCCCAGCCTGAATATGTG AACCAGCCAGATGTTCCGGCCAGCCCTTCTGCCCGGAGAGGGCCCTCTGCTGTGTGGCATCTGCGCTG GTGCCACTCTGGAAGGCCCAAGACTCTCTCCCCAGGGAAGAATGGGGTCTGCAAGACGTTTTTGCCTT TGGGGGTGCGGTGGAGAACCCCGAGTACTTGACACCCAGGGAGGAGCTGCCCTCAGCCCCACCTTCCT CCTGCTTCAGCCAGCCTTCGACAACCCTATTACTGGGACAGGACCCACCAGAGCGGGGGGCTCCAC CCAGCACCTTCAAAGGGACACCTACGGCAGAGAACCCAGAGTACCTGGGTCTGAGCTGCCAGTGAAC CAGAAGGCCAAGTCCGAGAAGCCCTGAITGTCTCAGGGAGCAGGGAAGGCCTGACTTCTGCTGGCAT CAAGAGGTGGGAGGGCCCTCCGACCACTTCCAGGGGAACCTGCCATGCCAGGAACCTGTCTTAAGGAACC TTCTTCTCTGCTTGAATTTCCAGATGGCTGGAAGGGGCTCAGCCCTGTTGGAAGAGGAACAGCACTGGG AGTCTTTGTGGATTCTGAGGGCTGCCCAATGAGACTCTAGGGTCCAGTGGATGCCACAGCCCGCTTGG CCCTTTCTTCCAGATCCTGGGTACTGAAAGCCTTAGGGAAGCTGGCCTGAGAGGGGAAGCGGCCCTAAG GGAGTGTCTAAGAAACAAAAGCGACCCATTGAGAGACTGTCCCTGAAACCTAGTACTGCCCCCATGAGGA AGGAACAGCAATGGTGTGATGATCCAGGCTTTGTACAGAGTGTCTTTCTGTTTACTTTTTTTG TTTTGTTTTTTTAAAGATAAAGAACCCAGGGGAGAAATGGGTGTTGTATGGGAGGCAAGTGTGGG GGGTCTTCTCCACACCCACTTTGICCAITTTGCAATATATTTTGGAAAACAGCTA	
NM_00112274 2	ATGGTCATAACAGCCTCCTGTCTACCGACTCAGAACGGATTTTACCAAAACGTGAAAATGCAGGCTCCATG CTCAGAAGCTCTTTAACAGGCTCGAAAGGTCCATGCTCCTTTCTCCTGCCATTCTATAGCATAAGAAGA CAGTCTCTGAGTGATAATCTTCTCTTCAAGAAGAAGAAAACCTAGGAAGGAGTAAGACACAAAGATCTCTTC ACATTCTCCGGGACTGCGGTACCAAAATATCAGCACAGCATTCTTTGAAAAGGATGTAGATTTTAACTCTG AACTTTGAACCATCACTGAGGTGGCCCGCCGGTTTCTGAGCCTTCTGCCCTGCGGGGACACGGTCTGCAC CCTGCCCGCGGCCACGGACCATGACCATGACCTTCCACACCAAAGCATCTGGGATGGCCCTACTGCATCA GATCCAAGGGAACGAGCTGGAGCCCTGAACCGTCCGAGCTCAAGATCCCCCTGGAGCGGCCCTGGGC GAGGTGTACCTGGACAGCAGCAAGCCCGCGGTGTACAACATACCCGAGGGCGCCGCTACGAGTTCAACG CCGCGGCCCGCCGCAACGCGCAGGTCTACGGTCAAGCCGCTCCCTTACGGCCCGGGTCTGAGGCTGC GGCGTTCGGCTCCAACGGCCTGGGGGTTTTCCCCCACTCAACAGCGTGTCTCCGAGCCCGCTGATGCTA GTCACCCCGCGCGCAGCTGTGCGCTTTTCTGACGCCCCACGGCCAGCAGGTGCCCTACTACCTGGAGA ACGAGCCAGCGGTACACGGTGGCGGAGGCCCGCCCGGCACTTACAGGCCAAATTCAGATAATCG ACGCCAGGGTGGCAGAGAAAGATTGGCCAGTACCAATGACAAGGGAAGTATGGCTATGGAATCTGCCAAG GAGACTCGCTACTGTGAGTGTGAATGACTATGCTTCAGGCTACCATTTATGGAGTCTGGTCTGTGAGG GCTGCAAGGCCCTTCTCAAGAGAAGTATTCAAGGACATAACGACTATATGTGTCCAGCAACCAACCATGT CACCATTGATAAAAAACAGGAGGAAGAGCTGCCAGGCTGCCGGCTCCGCAATGCTACGAAGTGGGAATG ATGAAAGGTGGGATACGAAAAGACCGAAGAGGAGGAGAAATGTTGAAACACAAGCCGACAGAGATGATG GGGAGGGCAGGGGTGAAGTGGGGTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAAGCCCGCTCAT GATCAAACGCTCTAAGAAGAAGCAGCCTGGCCCTTGCCCTGACGGCCGACAGATGTCTGAGTCCCTTGTG GATGCTGAGCCCCCATACTCTATTCCGAGTATGATCCTACCAGACCCCTTCACTGAAGCTTCCATGATGG GCTTACTGACCAACCTGGCAGACAGGGAGCTGGTTACATGATCACTGGGCGAAGAGGGTGCCAGGCTT TGTGGATTGACCCCTCCATGATCAGGTCCACCTTCTAGAATGTGCTGGCTAGAGATCCTGATGATTGGT CTGCTTGGCGCTCCATGGAGCACCCAGGGAAGCTACTGTTGCTCCTAACTTGTCTTGGACAGGAACC AGGGAAAATGTGTAGAGGGCATGGTGGAGATCTTCGACATGCTGTGCTGCTACATCATCTCGGTTCCGCAT GATGAATCTGCAGGAGAGGAGTTGTGTGCGCTCAAATCTATTATTTGCTTAATCTGGAGTGTACACA TTTCTGTCCAGCACCCCTGAAGTCTCTGGAAGAGAAGGACCATATCCACCGAGTCTGACAAAGATACAG	109

	ACACTTTTGATCCACCTGATGGCCAAAGGCAGGCCTGACCTGACAGCAGCAGCACCAGCGGCTGGCCAGCT CCTCCTCATCCTCTCCCACATCAGGCACATGAGTAACAAAGGCATGGAGCATCTGTACAGCATGAAGTGC AAGAACGTGGTGCCCTCTATGACCTGCTGCTGGAGATGCTGGACGCCCCACCGCTACATGCGCCCACTA GCCGTGGAGGGGCATCCGTGGAGGAGACGGACCAAAGCCACTTGGCCACTGCGGGCTCTACTTCATCGCA TTCCCTTGCAAAAGTATTACATCACGGGGAGGCAGAGGGTTCCCTGCCACGGTCTGAGAGCTCCCTGGC TCCCACACGGTTTCAGATAATCCCTGCTGCATTTTACCCTCATCAIGCACCCTTTAGCCAAATTCGTCT CCTGCATACACTCCGGCATGCATCCAACACCAATGGCTTTCTAGATGAGTGGCCATTTCATTGCTTGCTC AGTTCCTAGTGGCACATCTTCTGCTTCTGTTGGGAACAGCCAAAGGGATTCCAAGGCTAAATCTTTGTA ACAGCTCTCTTTCCCTCTGCTATGTTACTAAGCGTGAGGATTCCCGTAGCTCTTTCACAGCTGCACTCAG TCTATGGGTGGGGCTCAGATAACTCTGIGCATTTAAGCTACTTGTAGAGACCCAGGCTGGAGAGTAGA CATTTTGCTCTGATAAGCACTTTTTAAATGGCTCTAAGAATAAGCCACAGCAAAGAATTTAAAGTGGCT CCTTTAAATGGTGACTTGGAGAAAGCTAGGTCAAGGGTTTATTATAGCACCTCTTGTATTCCTATGGCA ATGCATCCCTTTATGAAAGTGGTACACCTTAAAGCTTTTATATGACTGTAGCAGATCTTGGTGATTTGT CAATTCATTCCCCCTATAGGAATACAAGGGGCACACAGGGAAGGCAGATCCCTTAGTTGGCAAGACTATT TTAATTGATACACTGCAGATTCAGATGTGCTGAAAGCTCTGCCCTTGGCTTTCCGGTCATGGGTCCAG TTAATTCATGCCCTCCCATGGACCTATGGAGAGCAGCAAGTTGATCTTAGITAAGTCTCCCTATATGAGGG ATAAGTTCCCTGATTTTTGTTTTTATTTTGTGTTTACAAAGAAAGCCCTCCCTCCCTGCTGCACTGCACTAA GGICAGCTTCAGGACCTGTTCAGTGGGCACTGTACTTGGATCTTCCCGGCTGTGTGTGCCCTTACACAG GGGTGAACCTGTTCACTGTGGTGATGCATGATGAGGTAATGGTAGTTGAAAGGAGCAGGGGCCCTGGTG TTGCATTTAGCCCTGGGGCATGGAGCTGAACAGTACTTGTGCAGGATTGTGTGGCTACTAGAGAAACA AGGGAAGTAGGGCAGAACTGGTACAGTTCTGAGGCACAGCCAGCTTGGCTCAGGGTGGCCCTGCCAC AGGCTGCAGCTACCTAGGAACATTCCTTGCAGACCCCGCATTTGCCCTTTGGGGTGGCCCTGGGATCCCTG GGGTAGTCCAGCTCTTCTTCATTTCCACAGCGTGGCCCTGGTTGGAAGAAGCAGCTGTCACAGCTGCTGTA GACAGCTGTGTTCTTACAATTGGCCAGCACCTTGGGGCAGGGGACAAGGGTGGGGACCGTTGCTGTAC TACTCAGGCTGACTGGGGCTGGTCAGATTACGTATGCCCTTGGTGGTTTATAGAGATAATATATATTTTGGAA GTTTGGTTTGGGAAGAAAATCCCTCCCTTCCCTCCCGGCCCGTTCCCTACCGCTCCACTCCTGCCA GCTCATTTCCCTTCAATTTCCCTTTCAGCTATAGGCTAAAAAAGAAAGGCTCATTCCAGCCACAGGGCAGCC TTCCCTGGGCCTTTGCTTCTCTAGCACAAATATGGGTTACTTCCCTTTTCTTAACAAAAAAGAAATGTTTG ATTTCTCTGGGTGACCTTATTGTCTGTAATTGAAACCCCTATTGAGAGGTGATGCTGTGTTAGGCCAATG ACCCAGGTGAGCTGCTCGGGCTTCTCTTGGTATGCTTGTGTTGAAAAGIGGATTTTCATTTCATTTCTGAT TGTCAGTTAAGTGATCACCAGGACTGAGAATCTGGGAGGGCAAAAAAAGAAAGTTTATGTTG CACTTAAATTTGGGGACAATTTTATGTAICTGTGTTAAGGATATGTTTAAAGACATAATTCCTTTTGTG TGTTTGTGTTAAGAACACCTTAGTTTGTGTTTAAAGAACCTTATATAGTATAATATATATTTTGGAAA TTACATTGCTTGTATCAGACAATTGAATGTAGTAATTCGTCTGGATTAAATTTGACTGGGTAAACA TGCAAAAACCAAGGAAAAATATTTAGTTTTTTTTTTTTTTTTTGTATACCTTTCAAGCTACCTTGTGATG TATACAGTCATTTATGCCATAAGCCTGGIGATTATTCATTTAAATGAAGATCACATTTTCATATCAACTTT TGTTCCACAGTAGACAAAATAGCTAATCCAGATGCCCTATTGTTGGTACTGAATGACAGCACTCTT ATGTAGCAAAGATTATGCCGAAAAGGAAAATTTTTCAGGGCAGCTAATTTTGCTTTTACCAAAATATCA GTAGTAATATTTTTGGACAGTAGCTAATGGGTGAGTGGGTCTTTTTTAAATGTTTACTTAGATTTTCTT TTAAAAAATTAATAAAAAAATAAAAAAATTTCTAGGACTAGACGATGTAATACCAGCTAAAGCCAAAC AATTATACAGTGGAGGTTTACATTATTCATCCAAATGTTTCTATTCTATGTTAAGATATATATATTTTGGAAA TGAAGTGGGCAGAGAACATCAGATGATTGAAATGTTGCCCAGGGGTCTCCAGCAACTTTGGAATCTCT TTGTATTTTACTTGAAGTGCCACTAATGACAGCAGATATTTCTGGCTGATGTTGGTATTTGGGTGTAG GAACATGATTAAAAAATACTCTTGCCCTGCTTTCCCCACTCTGAGGCAAGTTAAATGTAAAGAT GTGATTTATCTGGGGGCTCAGGTATGTGGGGAAAGTGGATTGAGGAATCTGGGGAATGGCAATATATTT AAGAAGAGTATTGAAAGTATTGGAGGAAAATGGTTAATCTGGGTGTGACCAGGGTTTCAGTAGAGTCC ACTTCTGCCCTGGAGACCACAAATCAACTAGCTCCATTTACAGCCATTTCTAAAAATGGCAGCTTCAGTTCT TAGAGAAGAAAGAACACATCAGCAGTAAAGTCCATGGAATAGCTAGTGGTCTGTGTTTCTTTTCGCCAT TGCTAGCTTGCCGTAATGATTCTATAATGCCATCATGCAGCAATTATGAGAGGCTAGGTATCCAAAGA GAAGACCTATCAATGTAGGTTGCAAAATCTAACCCCTAAGGAAGTGCAGTCTTTGATTGATTTCCTTA GTAACTTGCAGATATGTTTAAACCAAGCCATAGCCCATGCCCTTTGAGGGCTGAACAAATAGGGGACTTA CTGATAATTTACTTTTGATCACATTAAGGTGTTCTCACCTTGAATCTTATACACTGAAATGGCCATTGA TTTAGGCCACTGGCTTAGAGTACTCTTCCCTGCTGATGACACTGATTACAAATACTTTCCATTCTACT TTCCAATTTAGATGGACTGTGGGTACITGGGAGTGATCACTAACACCATAGTAATGTCTAATATTCACA GGCAGATCTGCTTGGGGAAGCTAGTTATGTGAAAGGCAATAGAGTCATACAGTAGCTCAAAAGGCAACC ATAATCTCTTTGGTGCAGGTCTTGGGAGCGTGATCTAGATTACACTGCACCATTCCCAAGTTAATCCCC TGAAAACCTTACTCTCAACTGGAGCAAAATGAACCTTTGGTCCCAAAATATCCATCTTTTCAGTAGCGTTAAT ATGCTCTGTTTCCAACCTGCATTTCTTTCCAATTGAATTAAGTGTGGCTCGTTTTAGTCATTTAAAA TTGTTTTCTAAGTAATGTCTGCCCTATTTATGGCACTTCAATTTTGCAGTCTCTTTTGTAGATTCAAGAAA AAATTTCTATTCTTTTTTTTGGCATCAATTTGTGCCGAACCTTTAAAAATATGTAATGCTGCCATGTTTCCA AAGCCATCTCAGTGTGTGTTTGTAGAGTGTGACCCCTAGAAACACATATTGTCCATGAGAGGTGTC CTGAGACACAGACCCCTTTGCATTCACAGAGAGGTGATTGTTTATAGAGACTTGAATTAATAAGTGACAT TATGCCAGTTCTGTCTCTCACAGGTGATAAACAATGCTTTTGTGCACTACATACTCTTCAGTGTAGA GCTCTGTTTTATGGGAAAAGGCTCAAAATGCCAATTTGTGTTGATGGAATTAATATGCCCTTTTGGCGAT GCATACATTACTGATGTGACTCGGTTTTGTGCGAGCTTTGCTTTGTTTAAATGAAACACACTGTGTAACCC TCTTTTGCACCTTTGAAAAAGAAATCCAGCGGGATGCTCGAGCACCTGTAAACAATTTCTCAACCTATTTG ATGTTCAAAATAAAGAATTAACCTAAA	
NM_130398	AAATTGAAGGTGAGCCTTTTCGCGCGCTGTGTAGGCAAGTTACCCGTGTTCTGCGTTGCCGCGCTGGGT	110

	GCTCTGGCCACAGTGAGTTAGGGGCGTCGGAGCGGGTTTCTCCAACCGCAATCGGCTCCGCTCAAGGGGA GGAGGAGAGTCCCTTCTCGGAAGGCCAAGGAAACGTTGTCGTCTGGAATGGGCTTGGGGGCCACGCTGC ACATCTCCGCGAGACAGAGGGATAAAGTGAAGATGGTGTGTTATTGTTACCTCGAGTGCCACATGCGAC CTCTGAGATATGTACACAGTCATTCTTACTATCGCACTCAGCCATTCTTACTACGCTAAAGAAGAAATAA TTATTCGAGGATATTTGCCCTGGCCAGAAAGAACTTATGTAAATTTTCATGAACATATATATCCGTTTCC TCGAGTGAGAGAAAACCTCTTTTAGATATCATCTGAGAGAACTAGTGAATCCAGTCACTGAGTGCAGT TGAGAGTCTAAGAACCTCTGAAATTTGAGAACTGCTGGACCAGAGCCTTTAGAGCTCTGATAAGGTGTCA ACAGGGTAGTTAAATTTGGCACCATGGGGATACAGGGATTGCTACAATTTATCAAAGAAGCACTTCAGAACCC ATCCATGTGAGGAAGTATAAAGGGCAGGTAGTAGCTGTGGATACATATTGCTGGCTTACAAAGGAGCTA TTGCTTTGCTGAAAACTAGCCAAAGGIGAACCTACTGATAGGTATGTAGGATTTTGTATGAAATTTGT AAATATGTTACTATCTCATGGGATCAAGCCTATTCTCGTATTGATGGATGTACTTTACCTTCTAAAAAG GAAGTAGAGATCTAGAAGAGAAAGACGACAAGCCAATCTTCTTAAGGGAAAGCAACTTCTTCGTGAGG GGAAAGTCTCGGAAGCTCGAGAGTGTTCACCCGGTCTATCAATATCACACATGCCATGGCCCAAGAGT AATTAAGCTGCCCGGTCTCAGGGGTAGATTGCTCTGCTGGCTCCCTATGAAGCTGATGCGCAGTTGGCC TATCTTAACAAAGCGGGAATTTGTCAAGCCATAATTACAGAGGACTCGGATCTCCTAGCTTTTGGCTGTA AAAAGGTAATTTAAAGATGGACCAGTTTGGAAATGGACTTGAATTTGATCAAGCTCGGCTAGGAATGTG CAGACAGCTTGGGGATGTATTTACGGAAGAGAAAGTTTCGTTACATGTGTATCTTTTCAGGTTGTGACTAC CTGTCAICTACTGCGTGGGATTGGATTAGCAAAGGCATGCAAAGTCTTAAGACTAGCCAATAATCCAGATA TAGTAAAGGTTATCAAGAAAATTGGACATTATCTCAAGATGAATATCACGGTACCAGAGGATTACATCAA CGGGTTTATTCGGGCCAACAAATACCTTCTCTATCAGCTAGTTTTTGTATCCCATCAAAGGAAACTTATT CCCTGTAACGCCCTATGAAGATGTATGTTGATCCTGAAACACTAAGCTACGCTGGGCAATATGTGATGATT CCATAGCTCTTCAAATAGCACTTGGAAATAAGATATAAATACTTTTGAACAGATCGATGACTACAATCC AGACACTGCTATGCTGCTGCCATTCAAGAAGTCAATAGTTGGGATGACAAAACATGTCAAAGTCAAGTAAAT GTTAGCAGCATTTGGCATAGGAATTACTCTCCAGACCAGAGTGGGTACTGTTTCAGATGCCCCACAAT TGAAGGAAAATCCAAGTACTGTGGGAGTGAACGAGTGATTAGTACTAAAGGGTTAAATCTCCCAAGGAA ATCATCCATTGTGAAAAGACCAAGAAGTGCAGAGCTGTGAGAAGATGACCTGTTGAGTCACTATTCTCTT TCATTTACGAAGAAGACCAAGAAAATAGCTCTGAAGGCAATAAATCATTGAGCTTTTCTGAAGTGTG TGCTGACCTGGTAAATGGACCTACTAACAAAAAGAGTGTAAAGCACTCCACCTAGGACGAGAAATAAAT TGCAACATTTTACAAAGGAAAAATGAAGAAAGTGGTGCAGTTGTGGTTCCAGGGCAAGACAGGTTT TTTTGCAGTTCAGATTCTACTGACTGTGTATCAAAACAAAGTGAGCATCCAGCCTCTGGATGAAACTGCTG TCACAGATAAAGAGAACAATCTGATGAATCAGAGTATGGAGACCAAGAAGGCAAGAGACTGGTTGACAC AGATGTAGCACGTAATTCAGTGATGACATTCGGAATAATCATATTCAGGTGATCATATTCAGACAAG GCAACAGTGTTTACAGATGAAGAGTCTACTCTTTTGGAGCAGCAAAATTTACAAGGACCATTTACCCAC CCACTTTGGGAACACTAAGAAGTTGTTTTAGTTGGTCTGGAGGTCTTGGAGATTTTTCAAGAACGCCGAG CCCCCTCTCCAAGCAGCATTGACGAGTTCGGAAGAAAGAGCGATTCCCCCACCTCTTTGCTGAGAAAT AATATGCTGATGTGTGCGAGTTAAAGAGCGAGGAGTCCAGTGACGATGAGTCTCATCCCTTACGAGAAG AGGCAATGTTCTTACAGTCCCAGGAAAGTGGAGAATTTCTCACTGAGAGTTCAAATGTCATCAAAGCTTTT TCAGTGCTCTAGTAAGGACTCTGATTACAGAGGAATCTGATTGCAATATTAAGTTACTTGACAGTCAAAGT GACCAGACCTCCAAGCTACGTTTATCTCATTTCTCAAAAAAGACACACCTCTAAGGAACAGGTTCTCTG GGCTATATAAGTCCAGTTCTGCAGACTCTCTTTCTACAACCAAGATCAAACCTCTAGGACCTGCCAGAGC CAGTGGGCTGAGCAAGAAGCGCGCAAGTCCAGAAAGAGAAAGCATATATGCCGAGAAACAGCGGGG TTACAGATCAAACCAATGAGCTCTGGAAAACTTTGGATTTAAAAAAGATTCTGAAAAGCTTCTCTCTT GTAAAGAAACCCCTGTCCCGAGTCAGAGATAACATCCAACCTAAGTCCAGAAAGCGGAAGAGGATATATTAA CAAACCTGAATGTGGCGGTGTTCAAAGAGCAATATTCAGTAAATGCAGACTGCTGCAAAGCTTTTGGCT GCAAGAGAAATCTGATCAATTTGAAGTCCCTGTTTTGGGAATGAGGCACTTATCAGCATGAAGAAATTTTCT TCATTCTGTGCCATTTTAAAAATAGAATACATTTTGTATATTAACCTTATAAATGGGTTGTGGTTTTTTT GCTCAGCTTTTATATTTTATAAGAAGCTAAATAGAAGAATAATGTATCTCTGACAGGTTTGTGAGG TTTTAGTGTTAATTTGGGAAAATCCCTCTGGAGTTTATAAAAGTCTACTCTAAATATTTCTGTAATGTGTG AAGTAGAAAGATAGTAAATGGAGAACTACAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	
AB209631	CCATGACCTGCCTTGAGAAGGGGCGAGGGGAAGCCAGATGGACTGGAAGTGGAGTGGCAGTGACCAAGGAG GAGGAGGTGTGATAGGCTTCCCACGCGAGGGTAGATCCAGAGACACCAGTGCCACCCATAGGCCCTTAGGA CTGCAAGTGGTACCCGATTCCCTTTTGTCCAGCTGAGACTCAGTTCTGAGTGTCTATTTTGGGGAACAGA GGCGTCCCTTGGTAGCATTTTGAAGAGGATAGCCAGCTGGGGTGTGTGTACATCACAGCCTGACAGTAACA GCATCCGAACCAGAGGTGACTGGCTAAGGGCAGACCCAGGGCAACAGGTTAACCCTTCTAGGGCCGGGCA CAGGGAGGAGAACATTCCAACACTCTGTGTGCCAGTGGCGACGACGCTTCTCTTTTATCCTCAAAC AGTCTATGAGGATATAAGCCAGAGAGAGACAGAGACAAGGAATTAAGTTGGTGGAGTCAAGGATTTG AACTTGGCTCTGGCAGATGGAATAATAGGGTCTGTATTCTTTACAAAACCGTGTGTGCTCAGATGGAGT TGGTGATAAACAAGCAGAGGTATCCAGGGTCGCGGTCTGCTTGGCCAGGAAGGGCCGCTTGTGAGTT GTGACCACCCAGCCCTGGAATGTGAGTAAATGCTGTAAGGAGTGGGGATCGGATCAGATGCCATCCAGAT GCTGAAGTTTGACCTTGTGTCATTTTCACTTTCTTTTTTGGCTCTTCTGCAATCAATTCATTTATTTAG CAAAAAAGAAATTTATGTGTGCTGAGAGATGCAAGAGATGTCTGCTGCTTCTGCTGCTTCCCTCCAAAAA GAATCCCAAACTGCTTTCTGTGAACGTGTGCCAGGGTCCAGCAGGACTCAGGGAGAGCAGGAAGCCCA GCCAGACCCCTTGCACAACCTACCGTGGGGAGGCTTAGGCTCTGGCTACTACAGAGCTGTTCCAGTC TGCACTGCCACAGCTGGCCAGGGACTTGGACACATCTGCTGGCCACTTCTGCTCAGCTTTCTTATCT GCAAAATAAGGGAAAAGCCCCACAAGGTGCACGTGTAGCAGGAGCTCTTTCCCTCCCTATTTTAGGA AGGCAAGTTGGTGGGAAGTCCAGCTTGGGTCCTGAGAGCTGTGAGAAGGAGATGCGGCTGCTGCTGGCCC TGTGGGGGCTCTGCTGAGTGTGCTGGGCTCCAGTCTGTCCCTGGAGGCTCTGAGGAAGTGGAGCT TGGTATGGCTTCTGAGGTGGGAGAGGGTGGCAGGGGTGGGAAGAGTGGGCACCAGGAGGGGGCTGCTGGG	111

26

27

	GAAGTTACCTATTCTCCACCTAAATCTCTGAAAAATGGAGAAACCCCTCTGACTAGTCCATGTCAAAATTTT ACTAAAAGTCTTTTTGTTTAGATTTATTTTCCTGCAGCACTCTTCGCAAAATGTACTATATAGTCAGCTT GCTTTGAGGCTAGTAAAAAGATATTTTCTAAACAGATTGGAGTTGGCATATAAACAAAATACGTTTTCTC ACTAATGACAGTCCATGATTTCGGAAATTTTAAGCCCATGAATCAGCCGCGGTCTTACCAGGTGATGCCT GTGTGCCGAGAGATGGGACTGTGCGGCCAGATATGCACAGATAAATATTTGGCTTTGTATTCATATAA AATTGCAGTGCATATTATACATCCCTGTGAGCCAGATGCTGAATAGATATTTTCTATTATTTTCAGTCTT TTATAAAAGGAAAAATAAACAGTTTTTAAATGTATGTATATAATTTCTCCCCATTACAAATCCTTCATG TATTACATAGAAGGATTGCTTTTTTAAAAATATACTGCGGGTTGGAAAGGGATTTAATCTTTGAGAAA CTATTTTAGAAAAATATGTTTGTAGAACAATTATTTTGA AAAAGATTTAAAGCAATAACAAGAAGGAAGG CGAGAGGAGCAGAACATTTTGGTCTAGGGTGGTTTTCTTTTAAACCATTTTTTCTTTGTTAATTTACAGTT AAACCTAGGGGACAAATCCGATTGGCCCTCCCCCTTTTGTAAATAACCCAGGAAATGTAATAAATTCATT ATCTTAGGGTGATCTGCCCTGCCAATCAGACTTTGGGGAGATGGCGATTGATTACAGACGTTTCGGGGG GTGGGGGGCTTGCAGTTTCTTTTGGAGATAATACAGTTTCTGCTATCTGCCGCTCTATCTAGGCGCAA CACTTAAGCAGTAATTGCTGTGTGTGTGTTCAAAATTTGATCATTGTTAAAGGATTGCTGCAATAAAT ACACTTTAATTTTCAGTCAAAA	
AJ249248	GTGGCCTCGAGGTGGTGGCAGGGCCGCCCTGTCAGTCCGGAGACGAACGCACGGACCGGGCCCTCCGGAG GCAGGTTTCGGCTGGAAGGAACCGCTCTCGCTTCGCTTACACTTGCAGCAATGCTCCGAGCTTACTCAC ATAGCATATTGGTATATCAAAATGAAATGCAAGGAACCAAAAATAACATAATTGAAGGCAGTAAAAGTGA AATTAATAGGAAGATCATCAGTCAAGGAAGACCCACTGGAGAGGACAGAAAATGAAGCAGTGTTTTATC ATGTGATTTTCAGCAGGTCTTCTTGAAATTTAACTAAAAATATGACTGCTCTCTCTTCAGAGAACTGCTC TTTTAGTACAGTACAGTCAAAACAAACAGCCCTAGACGTAACTATCTGCTATCTTGTATCATACTT GGGAAAATATTATTAATATCCTTACACTAGGAATGAGAAGAAAAACACCTGTCAAAATTTTATGGAAT ATTTTTGCAATTTCACTAGCATTCTGTTGATCTTTTACTTTTTGGTAAACATTTCCATTATATTGTAATTCAG GGATTTTGTACTTTTAAAGCATTAGGTTCACTAAATACCACATCTGCCTATTACTCAAATTTATTTCTTT ACTTATGGCTTTTTGCAATTATCCAGTTTTCTCTGACAGCTTGTATAGATTATGCGCTTATCTCTAAAA CAACCAAGCTTTTCAATTAAGTGTCAAAAATTTATTTTCTTTTACAGTAATTTTAAATTTGGAATTCAGT CCTTGCTTATGTTTTGGGAGACCCAGCCATCTACCAAAGCCTGAAGGCACAGAATGCTTATCTCGTCAAC TGICCTTTCTATGTCAGCATTGAGGTTACTGGCTGTCAATTTTCATGGTGATGATTTTATTTGTAGCTT TCATAACCTGTTGGGAAGAAGTTTACTACTTTGGTACAGGCTATCAGGATAACTTCTATATGAATGAAAC TATCTTATATTTTCTTTTTTATCCACACCCAGTTATACGTGAGATCTAAAAAAATATTCTTATCCAAG CTCATTTGCTGTTTTCTCAGTACCTGGTTACCATTGTACTACTTCAGGTAATCATTTGTTTTACTTTAAAG TTCAGATTCAGCATATATTGAGATGAATATTCCTGGTTATACCTTTGTCAATAGTTTTCTCATTGCTAC AGTGATTTGGTTTTAATTGTCAACAGCTTAAATTTAAAGACATTGGATTACCTTTGGATCCATTTGTCAAC TGGAAGTGCTGCTTCATTCCACTTACAATTCCTAATCTTGAGCAAATGAAAAGCCTATATCAATAATGA TTTGTTAATATTATTAATTAAGTTACAGCTGTCAAGATCATATTTTATGAACAGAAAGAACTCAG GACATATTA AAAAATAAACTGAACATAAAACAACTTTTGCCCCCTGACTGATAGCATTTCAGAATGTGTCT TTTTGAAGGCTATACCAAGTTATTAATAAGTGTTTTTATTTTAAAAACAAAAATAATTCAGAAAGTTTTTAT AGTTATTCAGGGACACTATATTACAAATATTACTTTGTTATTAACACAAAAGTGATAAGAGTTAACATT TGGCTATACTGATGTTTGTGTTACTCAAAAAAATACTGGATGCAAACTGTTATGTAATCTGAGATTTT ACTGACAACCTTTAAGATATCAACCTAAACATTTTTATTAATGTTCAAATGTAAGCAAGAAAAA	114
NM_014176	AGTCAGAGTTCGCGCAGGCGCTGGTACCCCGTTGGTCCGCGCGTTGCTGCGTTGTGAGGGGTGTCAGCTC AGTGATCCCAAGGCAGCTCTTAGTGTGGAGCAGTGAACCTGTGTGTTGTTCTTCTACTTTGGGATCATGC AGAGAGCTTCACGCTCGAAGAGAGAGCTGCACATGTTAGCCACAGAGCCACCCCCAGGCATCACATGTTG GCAAGATAAAGACCAATGGATGACCTGCGAGCTCAAATATTAGGTGGAGCCAACACACCTTATGAGAAA GGTGTTTTTAAGCTAGAAAGTTATCATTTCTGAGAGGTACCCATTGAACCTCCTCAGATCCGATTTCTCA CTCCAATTTATCATCCAACATTGATTCTGCTGGAAGGATTTGTCTGGATGTTCTCAAATTTGCCACCAAA AGGTGCTTTGGAGACCATCCCTCAACATCGCAACTGTGTTGACCTCTATTACAGCTGCTCATGTCAGAACCC AACCTGTATGACCCGCTCATGGCTGACATATCCTCAGAAATTAATAATAAGCCAGCCTTCTCAAGA ATGCCAGACAGTGGACAGAGAAGCATGCAAGACAGAAACAAAAGGCTGATGAGGAAGAGATGCTTGATAA TCTACCAGAGGCTGGTGAATCCAGAGTACACAACCTCAACACAGAAAAGGAAGGCCAGTCAGCTAGTAGGC ATAGAAAAGAAATTTTATCCTGATGTTTAGGGGACTTGTCTGTTTATCTTAGTTAATGTGTTCTTTGC CAAGGTGATCTAAGTTGCTTACCTTGAATTTTTTTTTAAATATATTGATGACATAAATTTTGTGTAGTT TATTTATCTTGTACATATGATTTTGAATCTTTTAAACCTGAAAAATAAATAGTCATTTAATGTTGAAA AAAAA	115
NM_006845	ACGCTTTCGCGCGGGATTTAAACTGCGGCGGTTTACGCGCGGTTAAGACTTCGTAGGGTTAGCGAAATG AGGTTTCTTGGTATTGCGCGTTTTCTCTTCTTGTGACTCTCCGAATGGCCATGGACICGTCGCTTCAGG CCGCGCTGTTTTCCCGGTCTCGCTATCAAGATCCAACGCAGTAATGGTTTTAATTCAGAGTGCCAAATGTAAG GACTGTGAACCTTGAGAAATCCTGTGTTTTCAGTGGAAATGGCAGAAAGGAGGTGCCACAAAGGGCAAAGAG ATTGATTTTGTATGATGTGGCTGCAATAAACCCAGAACTCTTACAGCTTCTTCCCTTACATCCGAAGGACA ATCTGCCCTTGCAGGAAATGTAACAATCCAGAAACAAAAACGGAGATCCGTCAACTCCAAAATTCCTGCT TCCAAAAGAAAGTCTTCGAAGCCGCTCCACTGCGATGTCCACTGCTCAGAGCTTCGCTACAGGCTCAG GAGAAATGACATGGAGGTGGAGCTGCTGACAGCTGCAAACTCCCGCAAGCAGTTTTTTCAGTTCTCTGCCC CCACTAGGCTTCTGCCCCTGCAGTGGCTGAAATACCATTGAGGATGGTCAGCGAGGAGATGGAAGAGCA AGTCCATTCCATCCGAGGAGCTCTTCTGCAACCCCTGTGAACCTCAGTTCCGGAGGAAATCATGCTTGTG AAGGAAGTGGAAAAATGAAGACAAGCGAGAGAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG GAGCTCAGGAGTATGACAGTAGTTTTTCCAACTGGGAATTTGCCCAATGATTAAGAATTTTCGGCTAC TTTGGAAATGTCATCCACTTACTATGACTGATCCTATCGAAGAGCAGAAATATGTGCTGTGTTAGGAAA CGCCCACTGAATAAGCAAGAAATTTGGCCAAGAAAGAAATTTGATGTGATTTCCATTCTTAGCAAGTGTCTC	116

	TCCTTGGTACATGAACCCAAAGTTGAAAGTGGAATTAACAAAGTATCTGGAGAACCAGCAATCTCTGCTTTGA CTTTGCATTTGATGAAACAGCTTCGAATGAAGTTGTCTACAGGTTACACGACAGGCCACTGGTACAGACA ATCTTTTGAAGGTGGAAAAGCAACTTGTCTTTTGCATATGGCCAGACAGGAAGTGGCAAGACACATACATATGG GCGGAGACCTCTCTGGGAAAGCCGAGAAATGCATCCAAAGGATCTATGCCATGGCCCTCCCGGGACGCTCTT CCTCCTGAAGCAATCAACCCCTGCTACCGGAAAGTTGGCCCTGGGAAGTCTATGTGACATCTCTTCGAGATCTAC AATGGGAAGCTGTTTGTACCTGCTCAACAAGAAGGCCAAGCTGCGCGTGTCTGGAGACGGCAAGCAACAGG TGCAAGTGGTGGGGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATCGACATGGG CAGCGCCTCGCAGAACCTCTGGGCAGACATTTGCCAACTCCAATTTCTCCCGCTCCCACGGCTGCTTCCAA ATTATTTCTCGAGCTTAAAGGGAGAATGCATGGCAAGTTCTCTTTGGTAGCTCTGGCAGGAATGAGCGAG GCGCGGACACTTCCAGTGTGACCGGCAGACCCGCATGGAGGGCGCAGAAATCAACAAGAGTCTCTTAGC CCTGAAGGAGTGCATTCAGGGCCCTGGGACAGAACAAGGCTCACACCCCGTTCCGTGAGAGCAAGCTGACA CAGGTGCTGAGGGACTCTTCATTTGGGGAGAATCTAGGACTTGATGATGATTGCCAGATCTCACCAGGCA TAAGTCTCTGTGAATATACCTTTAAACACCTGAGATATGCAGACAGGGTCAAGGAGTGAAGCCCCACAG TGGGCCCAAGTGGAGAGCAGTTGATTCAAATGAAACAGAAGAGATGGAAGCCTGCTCTAACGGGGCGCTG ATTCCAGGCAATTTATCCAAAGGAAGAGGAGAACTGTCTTCCAGATGTCCAGCTTTAACGAAGCCATGA CTCAGTCAAGGAGCTGGAGGAGGAGGCTATGGAAGAGCTCAAGGAGATCATACGACAGGACCCAGACTG GCTTGAGCTCTCTGAGATGACCGGACGCCAGACTATGACCTGGAGACCTTTGTGAACAAGCGGAATCT GCTCTGGCCAGCAAGCCAAAGCATTTCTCAGCCCTGCGAGATGTCTCAAGGCCCTGCGCCTGGCCATGC AGCTGGAAGAGCAGGCTAGCAGACAATAAGCAGCAAGAAACGGGCCAGTGCAGCATGCAAAATAAAAAAT CTGTTTGGTTTGACACCCAGCCTCTTCCCTGGCCCTCCCCAGAGAACTTTGGGTACCTGGTGGTCTAGG CAGGGTCTGAGCTGGGACAGGTTCTGGTAAATGCCAAGTATGGGGCATCTGGGCCAGGGCAGCTGGGG AGGGGGTCAGAGTGACATGGGACACTCCITTTCTGTTTCTCAGTTGTGCGCCTCAGCAGAGGAAGGAGCT CTTAGTTACCCCTTTTGTGTGCCCCCTCTTTCCATCAAGGGGAATGTCTCAGCATAGAGCTTTCTCCGCA GCATCGCTGCTGCGTGACTGGCTGCTAATGGAGAGCTCCCTGGGTTGTCTGCTGCTGGGGAGAGACA CGGAGCCTTTAGTACAGCTATCTGCTGGCTCTAAACCTTCTACGCCCTTTGGGCCGAGCACTGAATGTCTT GTACTTTAAAAAATGTTTCTGAGACCTCTTTCTACTTTACTGTCTCCCTAGAGATCCTAGAGGATCCCT ACTGTTTTCTGTTTTATGTGTTTTATACATTGTATGTAACAATAAAGAGAAAAAATAAATCAGCTGTTTTAA GTGTGTGAAAAAATAAAAAA	
NM_006101	ACTGCGCGCGTCTGTCGTAATGACCTCAGCGCCGCGGAGAATTTCAAATTCGAACGGCTTTGGCGGGCC GAGGAAGGACCTGGTGTTTTGTATGACCGCTGTCTCTAGCAGATACTTGCACGGTTTACAGAAATTCG GTCCTTGGGTGCTGTGCAGGAAACTGGAAAAAAGGTCATAAGCATGAAGCGCAGTTGAGTTTCCAGCGGTG GTGCTGGCCGCTCTCCATCGAGGAGTTAAGATCCAGGATGTAATAAACAAGGCCCTCTATACCCCTCA AACCAAAGAGAAACCAACCTTTGGAAGTTGAGTATAAACCAAACCGACATCTGAAAGAAAAAGTCTCGCTA TTTGGCAAAAGAACTAGTGACATGGATCCCGGAATAGTCAACTTGGTATATTTTCCAGTTCTGAGAAAA TCAAGGACCCGAGACCACTTAATGACAAAGCATTTCATTGACAGTGTTATTCGACAACTCTGTGAGTTTCT TACAGAAATGGTTATGTCACATAATGTGTCATGAAATCTCTACAGCTCCCTCTGTTAAAGACTCTCTG AAGATCTTACATTTCTTTATGGCTTCCGTGCCCCCTCATACGAACCTTCTGACACAAAGTTTGAAGAAG AGGTTCCAAGAATCTTTAAAGACCTTGGGTATCCTTTTGCACTATCCAAAGCTCCATGTACACAGTGGG GGCTCCTCATACATGGCCCTCACATTGTGGCAGCCTTAGTTTGGCTAATAGACTGCATCAAGATACATACT GCCATGAAGAAGCTCACCTTTTATTTGATATGGCAGCCTTGGGGAGAAGAACTGAAGATGGAATTA TGCATAATAAGTTGTTTTTGGACTACCCATAAAATGCTATGAGAGTTTTATGAGTGGTGCCGACAGCTT TGATGAGATGAATGCAGAGCTGCAGTCAAACTGAAGGATTTATTTAATGTGGATGCTTTTAAAGCTGGAA TCATTAGAAGCAAAAAACAGAGCAATTGAATGAACAGATTTGCAAGATGGAACAAGAAAGAGAAAAAGCA CGAATCTGCTAGAGTTGTTGAGAAAACTGAAGGCTTCCCTCAAGAGGAGTTTCAAAGATATCAGGCATA CATGAGCAATTGGAGTCTCATTGAGCCATTCTTGACCAGAAATTAATGGTCTCAATGAGGAAATGTCT AGAGTAGAAGTGAATGTGAACAATAAAACAGGAGAACACTCGACTACAGAATATCATTGACAACCCAGA AGTACTCAGTTGCAGACATTGAGCGAATAAATCATGAAGAAATGAATTGCAGCAGACTATTAATAAAT AACCAAGGACCTGGAAGCTTGAAGACAGAAGTTGTGAATGAGGAGTTAAAATATGCCAGACCAAGAA GCGATTGAAACACAATTAGCAGAGTATCACAATTTGGCTAGAAAAATTAATACTTATCTAAAGGTGCTG AGAATTCCAAAGGTTATGACTTTGAAATTAAGTTTAATCCCGAGGCTGGTGCCAACTGCCTTGTCAAATA CAGGGCTCAAGTTTTATGACCTCTTAAGGAACTCCTGAATGAAACTGAAGAAGAAATTAATAAGGCCCTA AATAAAAAATGGTTTTGGAGGTACTTTAGAACAAATGAAATGCAATGAAATGAACAAGAAAGCAAGAGTG TGAGAACTCTGAAAGAGAAGTTCAAAGCTGGATGATCTTTACCAACAAAAAATTAAGGAAGCAGAGGA AGAGGATGAAAAATGTGCCAGTGAGCTTGAGTCTTGGAGAAACACAAGCACTTCTAGAAAGTACTGTT AACCAGGGGCTCAGTGAAGCTATGAATGAATTAGATGCTGTTGACGGGGAATACCAACTAGTTGTGCAAA CCACGACTGAAGAAAGACGAAAGTTGGAAATAAATGCAACGCTTGCAGCTGTTAGAGATGGTTGCTACACATGT TGGGTCGTGAGAGAAACATCTTGAGGAGCAGATTGCTAAAGTTGATAGAGAATATGAAGAAATGCATGTCA GAAGATCTCTCGAAAAATATTAAGAGATATAGAGATAAGTATGAGAAGAAAGCTACTCTAATTAAGTCTT CTGAAGAATGAAGATAAAATGTTGATCAIGTATATATCCATAGTGAATAAAATTTGCTCAGTAAAGTG TAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	117
BC042437	CTCCCTCCTCTGCACCATGACTACCTGCAGCCGCCAGTTTACCTCCTCCAGCTCCATGAAGGGCTCCTGC GGCATCGGGGGCGGCATCGGGGGCGGCTCCAGCCGCATCTCCTCGTCTCGCCGGAGGGTCTCTGCCGCG CCCCCAGACCTACGGGGGCGGCCGTCTCTGCTCATCTCCCGCTTCTCCTCTGGGGGAGCCATATGGGTT GGGGGCGCGGTATGGCGTGGCTTTCAGCAGCAGCAGCAGCAGCTTTTGGTAGTGGCTTTGGGGGAGGATAT GGTGGTGGCTTGGTGTGGCTTTGGGTGGTGGCTTTGGTGGTGGCTTTGCTGGTGGTGGTGGTGGCTTCTGG TGGGAGTGAGAAAGGTGACCATGCAAGACCTCAACGACCGCCTGGCCTCCTACCTGGACAGGTCGCTGTC TCTGGAGGAGGCCAACCGCCGACCTTGGAAAGTGAAGATCCGTGACTGGTACACGAGGCAGGCCCTGCTGAG ATCAAAGACTACAGTCCCTACTCTCAAGACCATTGAGGACCTGAGGAACAGAGATTCTCAGAGCCAGATGG 	118

	ACAATGCCAATGTCTTCTGCAGATTGACAATGCCCGTCTGGCCGCGGATGACTTCCGCACCAAGTATGA GACAGAGTTGAACCTGCGCATGAGTGTGGAAGCCGACATCAATGGCCTGCGCAGGGTGCTGGACGAACTG ACCCTGGCCAGAGCTGACCTGGAGATGCAGATTGAGAGCCTGAAGGAGGAGCTGGCCTACCTGAAGAAGA ACCACGAGGAGGAGATGAATGCCCTGAGAGGCCAGGTGGGTGGAGATGTCAATGTGGAGATGGACGCTGC ACCTGGCGTGGACCTGAGCCGCATTCTGAACGAGATGCGTGACCAGTATGAGAAGATGGCAGAGAAGAAC CGCAAGSATGCCGAGGAATGGTTCTTACCAAGACAGAGGAGCTGAACCGCGAGGTGGCCACCAACAGCG AGCTGGTGACAGCGGCAAGAGCGAGATCTCGGAGCTCCGGCGCACCATGCAGAACCTGGAGATTGAGCT GCAGTCCCAGCTCAGCATGAAAGCATCCCTGGAGAACAGCCTGGAGGAGACCAAAGGTGCGTACTGCATG CAGCTGGCCAGATCCAGGAGATGATTGGCAGCGTGGAGGAGCAGCTGGCCAGCTCCGCTGCGAGATGG AGCAGCAGAACCAGGAGTACAAGATCCTGCTGGACGTGAAGACGCGGCTGGAGCAGGAGATCGCCACCTA CCGCCGCTGCTGGAGGGCGAGGACGCCACCTCTCCTCCTCCAGTTCTCCTCTGGATCGCAGTCATCC AGAGATGTGACCTCCTCCAGCGCCAAATCCGCACCAAGGTCATGGATGIGCAGCATGGCAAGGTGGTGT CCAGCCACGAGCAGGTCTCTCGCACCAAGAAGTGGGCTGCCAGCCCGCTCAGGCTAGGAGGCCCCC CGGTGGACACAGATCCCACTGGAAGATCCCTCTCCTGCCCAAGCACTTCAGCTGGACCTGCTTCA CCGTCACCCCCTCTGGCAATCAATACAGCTTCATTATCTGAGTTGCATAAAAAAAAAAAAAAAAAAAAA AA AA	
AK095281	CTCTTTTGCAGGGGCCGTTCTCGGGGCATGACGCTGGCTCCTGCACAGATCCTGCTCCTCTGTGGCCTT CCTGGGCTGCCCTCCCCTCCTCCGGGACTGCTCTGGACTGACACTGCTCAGGTTCGGATTCCCTCAAAGA CTTTGGGAGACAAGACTTGGTCCCCCTTTTACAACAAGGAACGGAGGCTCTAGAAGTGACTTCTCTGAA AGGCTTGGATCCAAAGCTCCCTCAGTTCAGCGGCCACGCTATTTTCCCTCAGACACAGGATCCTTGAAC CTGTGGGCTGTATCTCCCCGCGGACTTGGAGAATCCCAAGAGAGTGGGGCTCCACAGGCTGGAGTGCA ATGGTGTGATCTCGGCTCACTGCAACCTCCACCTCCAGGTTCAAGCTATCTCCTGCTCAGCCTCCTG AGTAGCTGGGATTACAGATCCTGGTGGCTGTGGTGGTAAATCCAGCTTCGTGCTGGCTACAGGTGGATG ATGCCACCTGGCTGCCGATGACCTCTGCACCAAGTGAAGGCTGGGTCTCTGGAGCTGCCAGGCGGTGG ACAAGCTGACCCCTGGCCGGGGCCAACTGGAGATGCAGATTGAGAACCTCAAGGAGGACCTGGTCTACCT GAAGAAGAACCCACAGCAGGAAATGAACGTCCTTTGAGGTGAGGTGGATGAGGATGTCAGTGTGAAGATG GACACTGTGCTGGAGTGAACCTGAGCTGCATCCTGAATGAGATGCGTGACCAGGACAAGACATTGGTGG AGAAGAGCTGCAAGGATGCGAGGGCTGGTCTTTCAGCATGGTGGGTGGCCGTGCGTAAGCAGGCTGTGA CACGTGTGGGCACATGTGCTGCATGCTGGTGCAGCTGGAGCACTGGCAGATCCACAGGCTGTCCAGTTG GAAGGACTTTTGGAAACAGTGGAGCAGCCCTCATGTTTTAGATGTAAACGTCAGGCTCAGAGAGGA CTCAGGCTCACACAGCCCTTCACGTGGCCTGCAAAATAGATCCAGGTCTCTACAAGTCTGGTCTTGGGT TTCCACCACAGCTGTTTACAGGATGTGCGTATTTGAATACATATGTATACCCTTGGCAAGCACAGGCTGA GTATCTCCGGTATCCTAGGGACAGCAACAGGGCGCAAAAGAATAACACCCAGTGCCTGTCTTTGAGGTGCT GCAGTTTCACTAGGAAAAAGAAATGCAAAATGACCGCAGAGCAGGCTGAATTCCTCCAGTTCCAATGTGGG TGCAGAGGCTCTCTGTGTGCAGAAAGAGGGGCTGAAGTGCAGGTTGGCCACCAACACAGAGGCCCTGCAG AGTGGCTGGATAGAGATATGGAGCTCAGCTCTCTGTGCAAGCTGAGCCGTCCAGTGAAGCAGGCTGA GCATCGCTGGAGGGCAGCCTGGTGGAGATGGAGGTGTGTTACAGGACCTGCGCGCCAGCTGCAGGGGC TTAACAGAAAGCATGGAGCAGCAGCTGTGCGAGCTCTGCTGCGACACGGAGCACCAGGACCAAGCACAG GTCTCTCTGGACGTGAAGACGTGGCTGGAGCAGGAGATGCCACCTACCGCCGCTTGCTGGAGGTTGAGG ACGCCACAGGTTGATTAAGATGACATGAGGCTGGAGTCTGGCTGAGGAGCCTTGAATGCCAAGTTAAGCG TCTGGACTAGATCAGTAGGCAATGGGAGCCATGGAGGGATTTGGAGCAGGAGAGTGAATGAACATCA AGAGATTTTAGAACATTCACTCTGGCTGCAGAGGGAGAAATGGATCAGAGGGGTACAGGGCGGGGCCAGAG AGATGTGTGAGGGGGCTGGAGCAGGGAGTCTGGCCAGAGAAGTCCCGTGGCGGTGGTGGGTAGTGGGGCAG GGAAGGAAAGGTGGTGCACGCGAGAGAGGTTATAGCTCAAAACAGCGGGACTGGATGCTGGATCTCG GGGTAAGCATGGCTCACAGTCAGGACTCAGTAAGTGTGGGAGAACACATGAAGGAGCAGGCATTGATGG CCGTGGGTTTCTGGTTCTGATGACTGTGTGAGTGGTGAAGAGCAAGGTGGGTGGTGGTGGTGGTGGTGGT TGGGAAGGGTGATCAGGCCCTCAGCTGAGAGTGTCCCGGAGTCTCCATGCTTAGTCACACGTGTGAGCTT TTTGCTCCCCGAAATGGTGAAGTCCATCTATAGTCTAACAACAGTCTCTCCTGCTTTAATTGGGTCTAT TTGTGGGCCCTCTGGGTTATGGAAAAACCACTTGCTCAGCTTCTCCTTGTAAATTCCTGGTGAGTAGCC ACAGAGTGGCCGACAGCTACTGCTGTGCTGTCTTTCTTTCTCTCTGCTGTGCTGAACCCCTGCCCT TTCATTCTTGGGCCCTGCGCTAATTTCTGTGCTATCCCACTGTGATTTTTCACCAATTTAGGGGAACCTC CTCTGCCAGGGCCTACTTCTCCCGCAGAGTGTGTCAGGTGCTTGCAGGTGCTGGGCTGGCTGGCATCCCTGGGCTGAT GGGTGTCTTCTCTCCCTGCAGGCTGGCCACTCAGTACTCCTTGTCTCTGGCTTCGACGCCACCCGGGAAG CCACAGTGACCAAGCCACAGGTGTGCCATCGTGGAGGAAGTCCAGGTTGGAGAGGTGGTCTTCTCTGTG AGCAGGTCCACTTCTCCACCCACTGAGACCCCTTCTGTCTGCGACAGCCCACTCGAGGGCCACGGCA CAGCCATCAGCTCCAGCTCCAGCATGCTACTGCCACGCCCGAGTGTCCGTCTGGGCCCGGTGATGG CCGTGTGCTTTCTGTATCTACTTTCTGCAGCCCCCTACTGAGGAGGCTCCTGGGTTTGTCCAGTGCCT ACTATTAAGCTTTGCTCCAAGTTC	119
M21389	GCATCCTTTTGGGCTGCTCACAGCCCCAGCCTCTATGGTGAAGACATACTTGCTAGCAGCGTCACCAA CTGTGCTGCCAAGAGATCAGTGTGCTGCAAGGCAAGGTATTTCTAACTGAGCAGAGCCTGCCAGGAAGAAAG CGTTTGCACCCACACCACTGTGCAAGGTGTGACCGGTGAGCTCACAGCTGCCCCCAGGCATGCCAGGCC CACTTAATCATTCACAGCTCGACAGCTCTCTCGCCAGCCAGTTCTGGAAGGGATAAAAAGGGGGCATC ACCGTTCTCTGGGTAAACAGAGCCACCTTCTGCGTCTCTGCTGAGCTCTGTTCTCTCCAGCCTCCCAACCC ACTAGTGCTGGTTCTCTTGTCTGCTGCAAGGAACAAAGCCACCATGCTCGCCAGTCAAGTGTGCTCTCCG GAGCGGGGGCAGTCGTAGCTTACGACCCGCTCTGCCATCACCCCGTCTGCTCTCCCGCACCAGCTTCACC TCCGTGTCCCGGTCCGGGGTGGCGGTGGTGGTGGCTTCGGCAGGCTCAGCCTTGGGGTGTCTGTGGAG TGGGTGGCTAAGGAGCCGAGCCTCTACAACCTGGGGGGCTCCAAGAGGATAICCAICAGCACTAGAGG	120

	AGGCAGCTTCAGGAACCGGTTTGGTGCTGGTGGTGGAGCGGCTATGGCTTTGGAGGTGGTGCCGGTAGT GGATTITGGTTTCGGCGGTGGAGCTGGTGGTGGCTTTGGGCTCGGTGGCGGAGCTGGCTTTGGAGGTGGCT TCGGTGGCCCTGGCTTTCTGTCTGCCCTCTGGAGGTATCCAAAGAGGTCACTGTCAACCAGAGTCTCCT GACTCCCTCAACCTGCAAAATCGACCCAGCATCCAGAGGTGAGGACCGAGGAGCGGAGCAGATCAAG ACCTCAACAATAAGTTTGGCTCCTTCATCGACAAGGTGCGGTTCTGGAGCAGCAGAACAGGTTCTGG ACACCAAGTGGACCTGCTGCAGGAGCAGGGCACCAAGACTGTGAGGCAGAACCTGGAGCCGTTGTTCGA GCAGTACATCAACAACCTCAGGAGGCAGCTGGACAGCATCGTGGGGGAACGGGGCCGCTGGACTCAGAG CTGAGAAACATGCAGGACCTGGTGAAGACTTCAAGAACAAGTATGAGGATGAAATCAACAACCGTACCA CTGCTGAGAATGAGTTTGTGATGCTGAAGAAGGATGTAGATGCTGCCATACATGAACAAGGTGGAGCTGGA GGCCAAGGTTGATGCACTGATGGATGAGATTAACCTTCATGAAGAATGTTCTTGTATCGGAGCTGTCCCAG ATGCAGACGCATGTCTCTGACACCTCAGTGGTCTCTCCATGGACAACAACCGCAACCTGGACCTGGATA GCATCATCGCTGAGGTCAAGGCCAGTATGAGGAGATTGCCAACCGCAGCCGACAGAACCGGAGTCTCTG GTATCAGACCAAGTATGAGGAGCTGCAGCAGACAGCTGGCCGGCATGGCGATGACCTCCGCAACCAAG CATGAGATCACAGAGATGAACCGGATGAICAGAGGCTGAGAGCCGAGATTGACAAATGTCAAGAAACAGT GCGCAATCTGCAGAACGCCATTGCGGATGCCGAGCAGCTGGGGAGCTGGCCCTCAAGGATGCCAGGAA CAAGCTGGCCGAGCTGGAGGAGGCCCTGCAGAAGGCCAAGCAGGACATGGCCCGCTGCTGCTGAGTAC CAGGAGCTCATGAACACCAAGCTGGCCCTGGACGTGGAGATCGCCACTTACCGCAAGCTGCGGAGGCG AGGAATGCAGACTCAGTGGAGAAGGAGTTGGACCAGTCAACATCTCTGTGTGTCACAAGCAGTGTTCCTC TGGATATGGCAGTGGCAGTGGCTATGGCGGTGGCTCGGTGGAGGTCTTGGCGCGGCCCTCGGTGGAGGT CTTGCCGGAGGTAGCAGTGAAGCTACTACTCCAGCAGCAGTGGGGGTGTCGGCTAGGTGGTGGGCTCA GTGTGGGGGCTCTGGCTTCAGTGCAAGCAGTGGCCGAGGCTGGGGGTGGGCTTTGGCAGTGGCGGGG TAGCAGCTCCAGCTCAAAATTTGCTCCACCACCTCTCTCCCGGAAGAGCTTCAAGAGCTAAGAACCT GCTGCAAGTCACTGCCCTTCAAGTGCAGCAACCCAGCCATGGAGATTGCCCTCTCTAGGCAGTTGCTCA AGCCATGTTTTATCCTTTTCTGGAGAGTGTCTAGACCAAGCCAAATGCAGAACACATCTTTGGTTCC CAGGAGAGCCCCATTCCCGAGCTGTGCTTCCCGTGGCGAGTTCTATATTCTGCTTCAAAATCAGCTTCT AGGTTTCCACAGCATGGCCCTGCTGACACGAGAACCCAAAGTTTCCCAAATCTAAATCATCAAAACA GAATCCCCACCCCAATCCCAAATTTGTTTTGGTTCTAATACCTCCAGAAATGTGTCAATAAAATGCTT TTATAATAT	
NM_00112306 6	GGACGGCCGAGCGGCAGGGCGCTCGCGCGCGCCCACTAGTGGCCGGAGGAGAAGGCTCCCCGGGAGGCCG CGCTGCCCCCCCCCTCCCCGAGGGCTCGCGTTCCCGCTGCTCGCGCTGCGCGCCCGCCGCGCTCA GGAACGCGCCCTCTTCGCCGGCGCGCGCCCTCGCAGTCAACGCCACCCACAGCTCCGGCACCAACAGCA CGCGCGCTGCCACCGCCACCTTCTGCGCGCGCCACCACAGCCACCTTCTCTCTCTCGCTGTCTCTCTC CGTCTCGCTCTGTGACTATCAGGTGAACCTTTGAACCAAGGATGGCTGAGCCCGCCAGGAGTTGGAAG TGATGGAAGATCACGCTGGGACGTACGGGTTGGGGGACAGGAAAGATCAGGGGGGTACACCATGCACCA AGACCAAGAGGGGTGACACGGACGCTGGCTGAAAGAATCTCCCTGCGAGACCCCACTGAGGACGGATCT GAGGAACCGGGCTCTGAAACCTCTGATGCTAAGAGCACTCCAACAGCGGAAGATGTGACAGCACCCCTAG TGGATGAGGAGCTTCCCGGCAAGCAGGCTGCCGCGCAGCCCAACAGGAGATCCCAAGAGGAACACAGC TGAAGAAGCAGGCATTGGAGACACCCCAAGCTGGAAGACGAAGCTGCTGGTACGCTGACCCAAGAGCCT GAAAGTGGTAAGGTGGTCCAGGAAGGCTTCTCTCCGAGAGCCAGGCCCCCAGGCTCTGAGCCACAGCTCA TGTCGGCATGCTGGGGTCCCCCTCTGCTGAGGGCCCCAGAGAGGCCACACGCCAACCTTCGGGGAC AGGACCTGAGGACACAGAGGGCGCGCCAGCCCTGAGCTGCTCAAGCACAGCTTCTAGGAGCCAGCTG CACCAGGAGGGGCGCGCTGAAGGGGCGAGGGGCAAGAGAGGCGGGGAGCAAGGAGGAGGTGGATG AAGACCGCGACGTGATGAGTCTCCCCCAAGACTCCCCCTCCCTCCCAAGGCTCCCCAGCCCAAGATGG CGGGCTCCCCAGACAGCGCCAGAGAAGCCACAGCATCCAGGCTTCCAGCGGAGGGTGCCATCCCC CTCCCTGTGATTTCTCTCCAAAGTTTCCACAGAGATCCAGCCTCAGAGCCCGAGGGCCAGGTGATG GGCGGGCCAAAGGGCAGGATGCCCCCTGGAGTTACGTTTACGTTGGAATACACCCCAACGTGCAGAA GGAGCAGGCGCACTCGGAGGAGCATTTGGGAAGGGCTGCATTTCCAGGGGCCCCCTGGAGAGGGGCCAGAG GCCCGGGGCCCCCTTTGGGAGAGGACACAAAAGAGGCTGACCTTCCAGAGCCCTCTGAAAAGCAGCCTG CTGCTGCTCCGCGGGGAAGCCCTGAGCCGGGTCCCTCAACTCAAAGCTCGCATGGTCAAGTAAAGCAA AGACGGGACTGGAAGCGATGACAAAAAAGCCAAGACATCCACACGTTCCTCTGCTAAAACCTTGAAAAAT AGGCCTTGCTTAGCCCCAAACACCCCACTCTCTGGTAGCTCAGACCCCTCTGATCCAACCTCCAGCCCTG CTGTGTGCCAGAGCCACCTTCTCTCTAAATACGTCTCTTCTGTCACTTCCGAACTGGCAGTTCTGG AGCAAAAGGAGATGAAACTCAAGGGGGCTGATGGTAAACGAAGATCGCCACACCGCGGGAGCCGCCCT CCAGGCCAGAAGGGCCAGGCCAACGCCACAGGATTCAGCAAAAAACCCCGCCGCTCCAAAGACACCAC CCAGCTCTGCGACTAAGCAAGTCCAGAGAAGACCACCCCTGCGAGGGCCAGATCTGAGAGAGGTGAACC TCCAAAATCAGGGGATCGCAGCGGTACAGCAGCCCGGCTCCCCAGGCACTCCCGGCAGCCGCTCCCCG ACCCGCTCCCTTCCAACCCCAACCCACCCGGGAGCCCAAGAAGGTGGCAGTGGTCCGTACTCCACCCAAGT CGCCGTCTTCCGCCAAGAGCCGCTGCGACAGCCCCGCTGCCATGCCAGACCTGAAGAATGTCAAGTC CAAGATCGGCTCCACTGAGAACCTGAAGCACAGCCGGGAGGGCGGAAGGTGCAGATAATTAATAAGAAAG CTGGATCTTAGCAACGTCCAGTCCAAGTGTGGCTCAAAGGATAATATCAAACACGTCCCGGGAGGCGGCA GTGTGCAAAATAGTCTACAAACCACTGACCTGAGCAAGGTGACCTCCAAGTGTGGCTCATGAGCAACAT CCATCATAAAACAGGAGGTGGCCAGGTGGAAGTAAATCTGAGAAGCTTGACTTCAAGGACAGAGTCCAG TCCAGATTTGGGTCCCTGGACAATATCACCCACGTCCCTGGCGGAGGAAATAAAAGATTGAAACCCACA AGCTGACCTTCCGCGAGAACGCCAAAGCCAAGACAGACCAGGGGGGAGATCGTGTACAAGTCCGCCAGT GGTGTGAGGACAGCTTCCACGGCATCTCAGCAATGTCTCTCCACCCGCGAGCATCGACATGGTATGAGC TCGCCCCAGCTCGCCACGCTAGCTGACGAGGTGTCTGCTCCCTGGCCAAGCAGGTTTGTGATCAGGCC CCTGGGGCGGTCAATAATTTGTGGAGAGGAGAGAATGAGAGAGTGTGGAAAAAAAAGAATAATGACCCGG CCCCCGCCCTCTGCCCCAGCTGCTCTCTCGCAGTTGGTTAATTGGTTAATCACTTAACCTGCTTTTGTCT	121

	ACTCGGCTTTGGCTCGGGACTTCAAATCAGTGATGGGAGTAAGAGCAAATTTTCATCTTTCCAAATTGAT GGGTGGGCTAGTAATAAAATATTTAAAAAAAACATTCAAAAACATGGCCACATCCAACATTTCCCTCAGG CAATTCCTTTTGATTCTTTTTTCTTCCCTCCATGTAGAGAGGGAGAAGGAGAGGCTCTGAAAGCTGC TTCTGGGGATTTCAAGGGACTGGGGGTGCCAACACCTCTGGCCCTGTGTGGGGGTGTACAGAGGCA GTGGCAGCAACAAAGGATTGAAACTTGGTGTGTTCTGTGGAGCCACAGGCAGACGATGTCAACCTTGTGT GAGTGTGACGGGGTTTGGGGTGGGGCGGGAGGCCACGGGGGAGGCCAGGCAGGGCTGGGCAGAGGGGA GAGGAAGCACAAGAGTGGGAGTGGGAGAGGAAGCCACGTGCTGGAGAGTAGACATCCCCCTCCTTGGCC CTGGGAGAGCCAAAGCCATGCCACCTGCAGCGTCTGAGCGGCCCTGTCTTGGTGGCGGGGTGGG GGCTGTCTGTGGGTGAGTGTGCCACCTCTGTCAGGGCAGCCTGTGGGAGAGGGACAGCGGTAAAAAGA GAAGGCAAGCTGGCAGGAGGGTGGCACTTCTGTGATGACCTCCTTAGAAAAAGACTGACCTTGATGTCTTG AGAGCGCTGGCCCTCTTCCCTCCCTGCAGGGTAGGGGGCTGAGTTGAGGGGCTTCCCTCTGTCTCCAC AGAAACCTCTGTTTTATTGAGTCTCAGAAGTTGGAAGTGTGCCAATGATTTTGGCCACTTTGCAGACCTGG GACTTTAGGGCTAACCAAGTCTCTTTGTAAGGACTTGTGCTCTTGGGAGACGTCCACCCCTTTCCCAAGC CTGGGCCACTGGCATCTCTGGAGTGTGTGGGGTCTGGGAGGCAGGTCCCGAGCCCCCTGTCTTCCCAC GGCCACTGCAGTCAACCCCTCTGCGCCGCTGTGTCTGTGTCTGCGCTGAGAGCCCAATCAGTGCCTATAC CCCTCATCACACGTACAAATGTCCCGAATTCAGCCCTCACCACCCCTTCTCAGTAATGACCCCTGGTGTG TTGTCAGGAGGTACCTACCTCCACTGAGGGTGAATTAAGGGAAGGCAAGGTCCAGGCACAGAGTGGGA CCCCAGCCTCTCACTCTCAGTTCCTCATCAACTGGGACCCTCACCACGAATCTCATGATCTGATTCTG GTTCCCTGTCTCCTCCTCCCTGCAGATGTGAGCCAGGGCAGTGTGAGCTGTGAGCCCTAGGTGTCTTCT GCCTTGTGACATGGAGAGAGCCCTTCCCTGAGAAGGCCTGGCCCTTCTGTGTGAGCCACAGCA GCAGGCTGGGTGTCTTGGTGTCTGAGTGGTGGCCACAGGATGGAAGGGCAAGGCACCCAGGCGCCAC AGTCCCGCTGTCCCCACTTGCACCTAGCTTGTAGCTGCCAACCTCCAGACAGCCAGCCCGCTGCTC AGTCCACATGCATAGTATCAGCCCTCCACACCCGACAAAGGGGAACACACCCCTTGGAAATGGTCTCT TCCCCCAGTCCCAGCTGGAAGCCATGCTGTCTGTCTGTGGAGCAGCTGAACATATACATAGATGTGT CCTGCCCCCTCCCATCTGCACCCCTGTGAGTTGATTTGGATTGTCTGTTTATGTTGGATTCCAGAGA GTGACTATGATAGTGAAGAAGAAAAAAGGACGCATGTATCTTGAATGCTTGTAAAGAG GTTCTTAACCCACCTCAGCAGGTGTCTCTACCCCCACACTGGGACTCGTGTGGCTGTGTGGTGGCCAC CCIGCTGGGGCTCCCAAGTTTTGAAAGGCTTCTCAGCACCTGGGACCCAAACAGAGACCAGCTTCTAG CAGCTAAGGAGGCGTTCAGCTGTGACGAAGGCTGAAGCACAGGATTAGGACTGAAGCGATGTATCTCC CTCTCCCTACTTCCCTTGGGGCTCCTGTGTGTCAGGGCACAGACTAGGTCTTGTGGCTGGTCTGGCTTGG GCGGAGGATGGTCTCTCTGGTCTATAGCCCAAGTCTCATGGCAGTCCCAAAGGAGGCTTACAACCTCT GCATCACAAAGAAAAGGAAGCCACTGCCAGCTGGGGGATCTGCAGCTCCCAAGAGCTCCGTGAGCCCTCA GCCACCCCTCAGACTGGGTCTCCTCTCCAAGCTCGCCCTCTGGAGGGGACAGCGAGCTCCCAAGAGGC CCTGCGACCACAGCAGGGATTGGGATGAATGCTGTCTGGATCTGCTCTAGAGGCCCAAGCTGCCTGC CTGAGGAAGGATGACTTGACAAGTCAGGAGACACTGTTCCAAAGCCTTGACCAGAGCACCTCAGCCCGC TGACCTTGCAAACTCCATCTGCTGCCATGAGAAAAGGGAAGCCGCTTTGCAAAACATTGCTGCCTAA AGAAACTCAGCAGCCTCAGGCCAATCTGCCCCCTCTGGTTTGGGTACAGTTAAAGGCAACCTCAGGG ACTTGGCAGTAGAAATCCAGGGCTCCTCTGGGGTGGCAGCTTCGTGTGCAGCTAGAGCTTTACCTGAA AGGAAGTCTCTGGGCCAGAACTCTCCACCAAGAGCCTCCTGCCGTTCGCTGAGTCCAGCAATCTCTCC TAAGTTGAAGGGATCTGAGAAGGAGAAGGAAATGTGGGGTAGATTTGGTGGTGGTTAGAGATATGCCCC CTCACTACTGCCAACAGTTTCGGCTGCATTTCTTACGCACCTCGGTTCCTCTCTCTCTCTCTTGTGTG CCTGTCTTACGACCATGGGCTTCTTATACGGAAGGCTCTGGGATCTCCCCCTTGTGGGGCAGGCTCT TGGGGCCAGCCTTAAGATCATGGTTTAGGGTGTATGCTGTGGCAGATAAATTGAAAGGCAAGCTGGCTT GTGATCTTAAATGAGGACAAATCCCCCAGGGCTGGGCACCTCCTCCCCCTCCTTCTCCCACCTGCA GAGCCAGTGTCTTGGGTGGGCTAGATAGGATATAGTGTATGTCGGGCTCCTTCAAGCTGTGACTCACTT TATCAATAGTTCCATTTAAATTGACTTTCAGTGGTGTGAGCTGTATCTGTGTGCTATTGCTTGTGTGCTA TGGGGGGAGGGGGGAGGAATGTGTAAAGATAGTTAACATGGGCAAGGGAGATCTTGGGGTGCAGCACTTA AACTGCTCTGTAACCTTTTTCATGATTTCAACCACATTTGCTAGAGGGAGGGAGCAGCCACGGAGTTAGA GGCCCTTGGGGTTTTCTTTTCCACTGACAGGCTTCCCAAGCAGCTGGCTAGTTCAATTCCTCCCCAGC CAGGTGCAGGCTAGGAATATGGACATCTGGTTGCTTTGGCTGTGCCCCCTTTTACGGGGTCTTAAGCC CACAATCATGCCTCCCTAAGACCTTGGCATCCTTCCCTCTAAGCCGTGTGGCACCTCTGTGCCACCTCTCA CACTGGCTCCAGACACAGCCTGTGCTTTTGGAGCTGAGATCACTCGCTTACCCCTCCTCATCTTTGTT CTCCAAGTAAAGCCACGAGGTGGGGCGAGGGCAGAGGTGATCACTGCGTGTCCATGTACAGACTGTC AGCTTCATAAACTTCTGATTTCTCTTACGCTTTGAAAAGGGTTACCTGGGCACCTGGCCTAGAGCCTCA CCTCCTAATAGACTTAGCCCATGAGTTTGCCATGTTGAGCAGGACTATTCTGGCACTTGCAAGTCCCA TGATTTCTTCGGTAATCTGAGGGTGGGGGAGGGACATGAAATCATCTTAGCTTAGCTTTCTGTCTGTG AATGTCTATATAGTGTATTGTGTGTTTAAACAAATGATTTACACTGACTGTGTGTGTAAGAGTGAATTTG GAAATAAAGTTATTACTCTGATTAAA	
M92424	GCACCGCGGAGCTTGGCTGCTTCTGGGGCTGTGTGGCCCTGTGTGTGCGAAAGATGGAGCAAGAAGCC GAGCCCGAGGGGGCGGCGGACCCCTCTGACCGAGATCCTGTGCTTTTCGAGCCAGGAGCAGCCGTCCCT CCCCGGATTAGTGTGCTACGAGCGGCCAGTGCCTTGGCCCGAGAGTGGAAATGATCCCCGAGCCAGGGC GTCGTGCTTCCGCACTAGTCACTCCCGTGAAGGAACTGGGGAGTCTTGGGGACCCCCGACTCCAAGC GCGAAAACCCCGGATGGTGGAGGACAGGCAAAATGTGCAATACCAACATGTCTGTACCTACTGATGTGTCT GTAACCACTCACAGATTCAGCTTCGGAACAAGAGCCCTGGTTAGACCAAGCCATTGCTTTTGAAGT TATTAAAGTCTGTTGGTGCACAAAAGACACTTATACTATGAAAGAGGTTCTTTTTTATCTTGGCCAGTA TATTATGACTAAACGATTATATGATGAGAAGCAACAACATATTGTATATTGTTCAAATGATCTTCTAGGA GATTGTTTGGCGTGCCAAAGCTTCTCTGTGAAAGAGCACAGGAAAATATATACCATGATCTACAGGAAC TGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACATCTGTGAGTGAAGACAGGTGTACCTTGA	122

	AGGTGGGAGTGATCAAAAGGACCTTGTACAAAGAGCTTCAGGAAGACAAACCTTCATCTTCACATTTGGTT TCTAGACCATCTACCTCACTAGAGGAGAGCAATTAGTGAGACAGAAGAAAATTCAGATGAATTATCTG GTGAACGACAAAGAAAACGCCACAAATCTGATAGTATTTCCCTTTCCCTTTGATGAAAGCCTGGCTCTGTG TGTAATAAGGGAGATATGTTGTGAAAGAAGCAGTAGCAGTGAATCTACAGGGACGCCATCGAATCCGGAT CTTGATGCTGGTGTAAGTGAACATTACAGGTGATTGGTTGGATCAGGATTCAGTTTTAGATCAGTTTAGTG TAGAATTTGAAGTTGAATCTCTCGACTCAGAAGATTATAGCCTTAGTGAAGAAGGACAAGAACTCTCAGA TGAAGATGATGAGGTATATCAAGTTACTGTGTATCAGGCAGGGGAGAGTGATACAGATTCAATTTGAAGAA GATCCTGAAATTTCCCTTAGCTGACTATTGGAAATGCACTTCATGCAATGAAATGAATCCCCCCTTCCAT CACATTGCAACAGATGTTGGGCCCTTCGTGAGAATTGGCTTCCTGAAGATAAAGGGAAGAAATAAGGGGA AATCTCTGAGAAAGCCAACTGGAAAACACAACAAGCTGAAGAGGGCTTTGATGTTCCCTGATTGTAAA AAAACATAGTGAATGATTCCAGAGAGTCATGTGTTGAGGAAAATGATGATAAAATTACACAAGCTTCAC AATCACAAGAAAGTGAAGACTATTTCTCAGCCATCAACTTCTAGTAGCATTATTTATAGCAGCCAAAGA TGTGAAAGAGTTTTGAAAGGGAAGAAACCCAAAGACAAAGAGAGAGTGTGGAATCTAGTTTGCCCTTAAT GCCATTGAACCTTGTGTGATTGTGCAAGGTCGACCTAAAAATGGTTGCAITGTCCATGGCAAAACAGGAC ATCTTATGGCCTGCTTTACATGTGCAAGAAGCTAAAGAAAGGAATAAGCCCTGCCAGTATGTAGACA ACCAATTCAAATGATTGTGCTAATTTATTTCCCTAGTTGACCTGTCTATAAGAGAATTATATATTTCTA ACTATATAACCTAGGAATTTAGACAACCTGAAATTTATTACATATATAAAGTGAAGAAATAGCTCAA TTCACATAGATTTCTCTCTTTAGTATAATTGACCTACTTTGGTAGTGGAATAGTGAATACTTACTATAA TTTGACTTGAATATGTAGCTCATCTTTACACCAACTCCTAATTTTAAATAATTTCTACTCTGTCTTAAA TGAGAAGTACTTGGTTTTTTTTTCTTAAATATGTATATGACATTTAAATGTAACCTATTATTTTTTTTG AGACCGAGTCTTGCTCTGTTTACCCAGGCTGGAGTGCAGTGGGTGATCTTGGCTCACTGCGCTGCCCC TCCCCGGGTTCCGACCATTTCTCCTGCTCAGCCTCCCAATTAGCTTGGCTACAGTCATCTGCCACCACA CTTGGCTAATTTTTTGTACTTTTAGTAGAGACAGGGTTTACCCTGTTAGCCAGGATGGTCTCGATCTCC TGACCTCGTGATCCGCCACCTCGGCCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCG	
NM_014791	GAGATTTGATTTCCCTTGGCGGGCGGAAGCGGCCACAACCCGGCGATCGAAAAGATTCTTAGGAACGCCGT ACCAGCCGCGTCTCTCAGGACAGCAGGCCCCCTGTCTTCTGTGCGGCGCGCTCAGCCGTGCCCTCCGCC CCTCAGGTTCTTTTCTAATTTCCAAATAAACTTGAAGAGGACTATGAAAGATTATGATGAACCTCTCAA ATATTATGAATTACATGAACTATTGGGACAGGTGGCTTTGCAAGGTCAACTTGCCCTGCCATATCCTT ACTGGAGAGATGGTAGCTATAAAAAATCATGGATAAAAAACACACTAGGGAGTGATTGCCCCGGTCTCAAAA CGGAGATTGAGGCCTTGAAGAACCCTGAGACATCAGCATATATGTCAACTCTACCATGTGCTAGAGACAGC CAACAAAAATTCATGGTTCTTGAGTACTGCCCTGGAGGAGAGCTGTTTGACTATATAATTTCCACAGGAT CGCCTGTGAGAAGAGGAGACCCGGGTGTGCTTCCGTGAGATAGTATCTGCTGTTGCTTATGTGCACAGCC AGGCTATGCTCACAGGGACCTCAAGCCAGAAAATTTGCTGTTTATGATGAATATCATAAATTAAGCTGAT TGACTTTGGTCTCTGTGCAAAACCCAAGGGTAACAAGGATTACCATCTACAGACATGCTGTGGGAGTCTG GCTTATGACGACCTGAGTTAATACAAGGCAAAATCATATCTTGGATCAGAGGCAGATGTTTGGAGCATGG GCATCTGTTATATGTTCTTATGTGTGGATTCTTACCATTGATGATGATAATGTAATGGCTTTATACAA GAAGATTATGAGAGGAAAAATATGATGTTCCCAAGTGGCTCTCTCCAGTAGCATTTGCTTCTCAACAA ATGCTGCAAGTGGACCCAAAGAAACGGATTCTATGAAAAATCTATTGAACCATCCCTGGATCATGCAAG ATTACAACATATCCTGTTGAGTGGCAAGCAAGAAATCCTTTTATTACCTCGATGATGATTGCGTAACAGA ACTTTCTGTACATCACAGAAACCAAGGCAAAATGGAGGATTTAATTTTCACTGTGGCAGTATGATCAC CTCAGGGCTACCTATCTTCTGCTTCTAGCCAAAGAGGCTCGGGGAAAACAGTTGCTTCTTAAAGCTTCTT CTTTCTCCTGTGGACAAGCCAGTGTCTACCCCATTCACAGACATCAAGTCAAATAATTGGAGTCTGGAAGA TGTGACCGCAAGTGATAAAAAATATGTGGCGGGATTAATAGACTATGATTGGTGTGAAGATGATTTATCA ACAGGTGCTGCTACTCCCCGAACATCACAGTTTACCAAGTACTGGACAGAAATCAATGGGGTGGAACTCA AATCATTAACCTCCAGCCTTATGCGAAGACCTGCAAAATAAATAAAGAAACAAAGAAATATGATATCTCC TAAGTCTGCTGTAAAGAATGAAGAGTACTTTATGTTTCTGAGCCAAAGACTCCAGTTAATAAGAACCAG CATAGAGAGAAATACTCACTACGCCAAATCGTTACACTACACCTCAAAGCTAGAAACAGTGCCTGA AAGAACTCCAATTAAAAATACCAGTAAATTCACAGGAACAGACAAGTTAATGACAGGTGTCATTAGCCC TGAGAGGGGTGCGGCTCAGTGGAAATGGATCTCAACCAAGCACATATGGAGGAGACTCCAAAAAGAAAG GGAGCCAAAGTGTTTGGGAGCCTTGAAGGGGGTGGATAAGGTTATCACTGTGCTCACCAGGAGCAAAA GGAAGGGTCTGCGCAGAGACGGGCCAGAAAGCTAAAGCTTCACTATAACGTGACTACAACATAGATTAGT GAATCCAGATCAACTGTTGAATGAATAATGTCTATTCTTCCAAAGAAGCATGTTGACTTTGTACAAAAAG GGTTATACACTGAAGTGTCAAACACAGTCAGATTTTGGGAAAGTGACAATGCAATTTGAATTAGAAGTGT GCCAGCTTCAAAAACCCGATGTGGTGGGTATCAGGAGGCAGCGGCTTAAGGGCGATGCTGGGTTTTACAA AAGATTAGTGAAGACATCCTATCTAGCTGCAAGGTATAATTGATGGATTCTTCCATCCTGCCGGATGAG TGTTGGGTGTGATACAGCTACATAAAGACTGTTATGATCGCTTTGATTTTAAAGTTTCAATGGAACCTACCA ACTTGTTTTCTAAAGAGCTATCTTAAGACCAATATCTCTTTGTTTTTAAACAAAAGATATTATTTGTGTA TGAATCTAAATCAAGCCCATCTGTCAATTTATGTTACTGTCTTTTTTAAATCATGTGGTTTTGTATATTAATA ATTGTTGACTTTCTTAGATTCACTTCCATATGTGAATGTAAGCTCTTAACATATGTCTCTTTGTAATGTGT AATTTCTTTCTGAAATAAAACCATTGTGTAATATAG	123
BG765502	GCAGCGGAGGAGCCAGTCCACGATGGCCCGGTCCTGGTGTGCCTTGGTGTGTCATCACTTGTGTCTGC CTTCTCCGGACCTGGTGTGAGGGGGTGGTCTATGCCCAAGCTGGCTGACCGGAAGCTGTGTGCGGACCAG GAGTGCAGCCACCCTATCTCCATGGCTGTGGCCCTTCAGGACTACATGGCCCCGACTGCGGATTCCTGA CCATTACCGGGGCCAAGTGGTGTATGTCTTCTCCAAGCTGAAGGGCCGTGGGCGGCTCTTCTGGGGAGG CAGGTTCAAGGGAGATTACATGAGAGACTGGCTGTGCTGCGCTGGGCTATTITCCCCAGTAGCATTTGCCGA GAGGACCAGACCCTGAAACCTGGCAAGTCGATGTGAAGACAGACAAATGGGATTTCTACTGCCAGTGTGAG CTCAGCCTACCGCTGGCCCTGCGCTTTCCCTCCTTGGGTTTATGCAAAATACAATCAGCCAGTGCAAAA AAAAAAAAAAAAAAAAAACTTCGGAGAGAGATAGCAACAAAAGGCCGCTTGTGTGAAGGCGCCAAAA	124

	GTTTTGCCCCAAGAGACCTTCGGCCTCCCCAGGGCGCGCGCAAAGGCGCCTTGTITTGACAACCTCTTG GACAACCGGAGGGGCTACCGCCCGGAGACCCCTGTGGTGGACCCCCGGGCAACCGGGTGTGACAGGGTA CTACCCCCACGGCTTTGTGGGGGTCCCACCAAAGGCCCCAAAGAGGCTCTTTCAAGGCACTATTCCTT GTTGTAGACCTTGTGTGTGCCACAGGCGCCAAAGAAACCTCGGGGGGCTAACAACGCACGTGCTTGGCA GCTCCGAGAAGGCTCTCTCCACCCGAGGGGTGGACGCAACAGGGGGAATGGGCCATCATTTGTTGCC CCGGTGGGCACCAACTCTTTTTCCCCCATAGAGAGGCTTAGCACACTATGTGGGCACGTTATTGCCGC CTAGAGAAACCGAGCGCCAGAAATTTGGAAGGGGGGGCGCTTCTCATTTTTGCGCAAAACCCCTT GTGGGAGTATGCCCGAACTCCTCTGGAACACACAAGCGACACTTGGCGGGGGTCTGCAAAAACCTCCT GTTGGGAAGCCGGCTTACN	
NM_002417	TACCGGGCGGAGGTGAGCGCGGCGCGGCTCCTCTCGGCGGACTTTGGGTGCGACTTGACGAGCGGTG GTTCGACAAGTGGCCTTGC GGCCGGATCGTCCCAGTGAAGAGTTGTAATTTGCTTCTGSCCTTCCCC TACGGATTATACCTGGCCTTCCCTACGGATTATACTCAACTTACTGTTTAGAAAATGTGGCCACGAGA CGCCTGGTTACTATCAAAAGGAGCGGGTTCGACGGTCCCCACTTTCCCTGAGCCTCAGCACTGCTTGT TTGGAAGGGGTATTGAATGTGACATCCGIATCCAGCTTCTGTTGTGTAACAAACATTGCAAAATGA AATCCATGAGCAGGAGGCAATATTACATAATTTCACTTCCACAAATCCAACACAAGTAAATGGGTCTGTT ATTGATGAGCCTGTACGGCTAAAACATGGAGATGTAATACTATTATGATCGTTCCTTCAGGTATGAAA ATGAAAAGTCTTCGAATGGAAGGAAGTCAACTGAATTTCCAAGAAAAATACGTGAACAGGAGCCAGCAGC TCGTGTCTCAAGATCTAGCTTCTCTCTGACCCTGATGAGAAAGCTCAAGATTCCAAGGCCTATTCAAAA ATCACTGAAGGAAAAGTTTCAGGAAATCCTCAGGTACATATCAAGAATGTCAAAGAGACAGTACCCGAG ATGACTCAAAAGACAGTGTGTCTCAGGGAACAATAATGTTTATTCCTCAGAACATGCTGGACGTAATGG CAGAAATGACAGCTGATCCCATTTCTGGGGATTTTAAAGAAATTTCCAGCGTTAAATTAGTTCAGTTCAT GGAGAAITGAAGTCTGTTCCCACTACACAATGTCTGACAATAGCAAAAAAATGAATCTCCCTTTTGGGA AGCTTTATGAGTCAGTGAAGAAAAGAGTTGGATGTAAAATCACAAAAGAAAATGTCTACAGTATTGTAG AAAATCTGGATTACAAACTGATTACGCAACAGAGAAAGAAAGTGTGATGGTTTACAGGGGGAGACCCAA CTGTTGGTCTCGCGCTAAGTCAAGACCAAAATCTGGTGGGAGCGGCCACGCTGTGGGACAGCTTCTCAC CTGAACAAGAGCTTGACCAGAACAAGGGGAAGGGAAGAGAGCTGGAGTCTGTTTCACTCCCAGCAAGGC TGTGGGCGCCAGCTTCTCTCTATGAGCGGCTAAAATGAAGACCCCTGTACAATATTACAGCAACAA AATCTCCCAAAAACATAAGAACAAGACCTGTATACTACTGGTAGAAGAGAATCTGTGAATCTGGGTA AAAGTGAAGGCTTCAAGGCTGGTGATAAACTCTTACTCCAGGAAGCTTTCAACTGAGAAATCGAACACC AGCTAAAGTTGAAGATGCAGCTGACTCTGCCACTAAGCCAGAAAATCTCTCTTCCAAAACAGAGGAAGT ATTCTACAGATGTGGAAGTTCTGCCACGAAACTGAAATTCACAATGAGCCATTTTAACTCTGTGGC TCACTCAAGTTGAGAGGAAGATCCAAAAGGATTCCCTCAGCAAGCCTGAGAAATTTGGGCACTACAGCTGG ACAGATGTGCTCTGGGTTACCTGGTCTTAGTTTCACTTGTATATCAACAACCTTGGTGTATCCATTAATGAG AGTGAGGGAATACCTTTGAAAAGAGGCGTGTGTCCTTTGGTGGGCACCTAAGACCTGAACATTTTGTG AAAACTTGCTCTTAATACGCTCTCAAAAGGGGAGAAGCCCCAACCAAGAAAAGTCTCTGGTAATGCA CACTCCACCTGTCTGAAGAAAATCATCAAGGAACAGCCTCAACCATCAGGAAAACAGAGTCAGGTTCA GAAATCCATGTGGAAGTGAAGGCACAAAGCTTGGTTATAAGCCCTCCAGCTCTTAGTCTAGGAACTCTC CAGTTGCCAGTGATCAACGCCGTAGGTCTGCAAAAACAGCCCTGCTTCCAGCAGCAAAATCTCAGACAGA GGTTCCTAAGAGAGGAGGGAGAAAGAGTGGCAACCTGCCTTCAAGAGAGATGTCTATCAGCCGAAGTCAA CATGATATTTTACAGATGATATGTTCCAAAAGAAGAGTGGTGTTCGGAAGCAAAATCTGATTGTTGCAA AATCATGGGCAGATGTAGTAAAACCTTGGTGCAAAAACAAACACAACTAAAGTCTATAAACTGCTCTCA AAGGTCAATGAACAAAAGGCAAAGAAGACCTGCTACTCCAAAGAAGCCTGTGGGCGAAGTTTACAGTCAA TTTAGTACAGGCCACGCAAACTCTCTTGTACCATAATAATAGGGAAAGCTCATACTGAAAAGTACATG TGCTGTCTGACCCCTACAGAGTGTCAACAACCTTCAATTTCAACCAAAAAATGGACTTTAAGGAAGATCT TTCAGGAATAGCTGAAATGTTCAAGACCCAGTGAAGGAGCAACCGCAGTTGACAGACAGTGTACATC GCTATTTCAAATTCAGAGAAATTTGCTTGGAAAACAGTTTCAAGGAACGATTTCAGGAGAAGAACCTCTGC TCCCCACCTCAGAGAGTTTTGGAGGAAATGTGTTCTTCACTGACAGAAATGCAGCAAAACAGCCATCTGA TAAATGCTCTGCAAGCCCTCCCTTAAAGACGGCAGTGTATTAGAGAAAATGGAAACGTAGCAAAAACGCC AGGAACACCTACAAAATGACTTCTCTGGAGACAAAACCTTCAGATACTGAGACAGAGCCTTCAAAAACAG TATCCACTGCAAAACAGGTGAGGAAGGTCTACAGAGTTCAGGAATATACAGAAGCTACCTGTGGAAGTAA GAGTGAAGAAACAAATACAGAAATTTGTGAGTGCATCCTAAAAAGAGGTGAGAAGGCAACACTACTACAA CAAAGGAGAGAAGGAGAGATGAAGGAAATAGAAAGACCTTTTGAACATATAAGGAAAATATTGAATTAA AAGAAAACGATGAAAAGATGAAAGCAATGAAGAGATCAAGAACTTGGGGGCAGAAATGTGCACCAATGTC TGACCTGACAGACCTCAAGAGCTTGCCTGATACAGAACCTCATGAAAGACACGGCACGTGGCCAGAATCTC CTCCAAACCCAGATCATGCCAAGGCACCAAGAGTGAGAAAGGCAAAATCACTAAAATGCCCTGCCAGT CATTACAACCAGAACCAATAAACACCCCAACACACAAAAACAACAGTTGAAGGCATCCCTGGGGAAGT AGGTGTGAAAGAAGAGCTCCTAGCAGTCGGCAAGTTACACGGACGTGAGGGGAGCCACGCACACGCAC AGAGAGCCAGCAGGAGATGGCAAGAGCATCAGAACGTTTAAAGGAGTCTCCAAAGCAGATCCTGGACCCAG CAGCCCGTGTAACTGGAATGAAGAAGTGGCCAAAGACGCTAAGGAAGAGGCCAGTCACTAGAAGACCT GGCTGGCTTCAAAGAGCTCTTCCAGACACCAGGTCCCTCTGAGGAATCAATGACTGATGAGAAAACCTACC AAAATAGCCTGCAAACTCCACCACCAAGTCACTGGACACTCCAAACAGCACAAAGCAATGGCCTAAGA GAAGTCTCAGGAAAGCAGATGTAGAGGAAGAAATCTTAGCACTCAGGAACTAACACCATCAGCAGGGAA AGCCATGCTTACGCCCCAACCCAGCAGGAGGTGATGAGAAAGACATTAAAGCATTTATGGGAACTCCAGTG CAGAAACTGGACCTGGCAGGAACCTTACCTGGCAGCAAAAGACAGCTACAGACTCCTAAGGAAAAGGCC AGGCTCTAGAAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGACTCCTGGTACACCCAGGGAATTAGTGGC TGCTGGTAAAACCACTAAAATACCTTGGGACTCTCCACAGTCAGACCCAGTGGACACCCCAACAGCACAA AAGCAACGACCCAGAGAAATATCAGGAAAGCAGATGTAGAGGGAGAACCTTAGCGTGCAGGAATCTAA TGCCATCAGCAGGCAAGCCATGCACACGCCCTAAACCATCAGTAGGTGAAGAGAAAGACATCATATATT	125

TGTGGGAACCTCCAGTGCAGAAACTGGACCTGACAGAGAACTTAACCGGCAGCAAGAGACGGCCACAAACT CCTAAGGAAGAGGCCAGGCTCTGGAAGACCTGACTGGCTTTAAAGAGCTCTTCCAGACCCCTGGTCATA CTGAAGAAGCAGTGGCTGCTGGCAAAACTACTAAATGCCTGCGAATCTTCTCCACCAGAAATCAGCAGA CACCCCAACAAGCACAAGAAGGCAGCCCAAGACACCTTTGGAGAAAAGGGACGTACAGAAGGAGCTCTCA GCCCTGAAGAAGCTCACACAGACATCAGGGGAAACCACACACACAGATAAAGTACCAGGAGGTGAGGATA AAAGCATCAACGCTTTAGGGAAACTGCAAAACAGAAACTGGACCCAGCAGCAAGTGTAACTGGTAGCAA GAGGCACCCAAAACTAAGGAAAAGGCCCAACCCCTAGAAGACCTGGCTGGCTTGAAGAGCTCTTCCAG ACACCAAGTATGCACTGACAAGCCACGACTCACGAGAAAACCTACCAAAATAGCCTGCAGATCACAAACCAG ACCCAGTGGACACACCAACAAGCTCCAAGCCACAGTCCAAGAGAAGTCTCAGGAAGGTGGACGTAGAAGA AGAATTTCTCGACTCAGGAAACGAACACCATCAGCAGGCAAAGCCATGCACACACCCAAACCAGCAGTA AGTGGTGAGAAAAACATCTACGCATTTATGGGAACCTCCAGTGCAGAAACTGGACCTGACAGAGAACTTAA CTGGCAGCAAGAGACGGCTACAAACTCCTAAGGAAAAGGCCAGGCTCTAGAAGACCTGGCTGGCTTTAA AGAGCTCTTCCAGACACGAGGTACACTGAGGAATCAATGACTAACGATAAAACTGCCAAAGTGGACGTGC AAATCTTCACAACCAGACCCAGACAAAAACCCAGCAAGCTCCAAGCGACGGCTCAAGACATCCCTGGGGA AAGTGGGCGTGAAAGAAGAGCTCTTAGCAGTTGGCAAGCTCACACAGACATCAGGAGAGACTACACACAC ACACACAGAGCCAAACAGGAGATGGTAAGAGCATGAAAGCATTTATGGAGTCTCCAAAGCAGATCTTAGAG TCACGAGCAAGTCTAACTGGCAGCAAGAGGCAAGCTGAGAAGTCTTAAGGGAAGTCTGAAGTCCCTGAAG ACCTGGCCGGCTTCATCGAGCTCTTCCAGACACCAAGTCACTAAGGAATCAATGACTAACGAAAAAAC TACCAAAGTATCCTACAGAGCTTCACAGCCAGACCTAGTGGACACCCCAACAGCTCCAAGCCACAGCCC AAGAGAAGTCTCAGGAAAGCAGACACTGAAGAAGAATTTTATGATTTAGGAAACAAACGCCATCAGCAG GCAAAAGCCATGCACACACCCAAACCAGCAGTAGGTGAAGAGAAAGACATCAACACGTTTTTGGGAACCTCC AGTGCAGAAACTGGACCAAGCCAGGAAATTTACCTGGCAGCAATAGACGGCTACAAACTCGTAAGGAAAAG GCCCAGGCTCTAGAAGAACTGACTGGCTTCAGAGAGCTTTTCCAGACACCATGCACTGATAACCCACACGA CTGATGAGAAAACTACCAAAAAATACTCTGCAAACTCTCCGCAATCAGACCCAGCGGACACCCCAACAAA CACAAAGCAACGGCCCAAGAGAAGCTTCAAGAAAGCAGACGTAGAGGAAGAATTTTATGATTCAGGAAAG CTAACACCATCAGCAGGCAAAGCCATGCACACGCCCTAAAGCAGCAGTAGGTGAAGAGAAAGACATCAACA CATTTGTGGGGAAGTCCAGTGGAGAACTGGACCTGCTAGGAAATTTACCTGGCAGCAAGAGACGGCCACA AACTCTTAAAGAAAAGGCCAAGGCTCTAGAAGATCTGGCTGGCTTCAAGAGCTCTTCCAGACACCAAGGT CCAACTGAGGAATCAATGACCCATGACAGAAATCAGAGAAGTATCCTGCAAACTCCACAAACAGACCCAG TCAAAACCCCAACAAGCTCCAAGCAACGACTCAAGATATCCTTGGGGAAGTAGGTGTGAAGAAGAGGT CCTACCAGTCGGCAAGCTCACACAGACGTCAGGGAAGACCACACAGACACACAGAGAGACAGCAGGAGAT GGAAGAGCATCAAAGCGTTTAAAGGAATCTGCAAGCAGATGCTGGACCCAGCAAACTATGGAATGGGA TGGAGAGGTGGCCAGAACACCTAAGGAAGAGGCCCCAATCACTAGAAGACCTGGCCGGCTTCAAAGAGCT CTTCAGACACCAAGACCACTGAGGAATCAACAATGATGACAAAACCTACCAAAATAGCCTGCAAACTCT CCACCAACCAAGATCAATGGACACTCCAACAAGCACAAGGAGGCGGCCCAAAACACCTTTGGGGAAGAGGG ATATAGTGAAGAGCTCTCAGCCCTGAAGCAGCTCACACAGACCACACACAGACAAAGTACCAGGAGA TGAGATAAAGGCATCAACGTTTCAGGGAAGCTGCAAAACAGAAACTGGACCCAGCAGCAAGTGAACCT GGTAGCAAGAGGCAGCCAAAGACCTTAAGGGAAGAGCCCAACCCCTAGAAGACTTGGCTGGCTTGAAG AGCTCTTCCAGACACCAATATGCACTGACAAGCCACGACTCATGAGAAAACCTACCAAAATAGCCTGCAG ATCTCCACAACCAAGACCCAGTGGGTACCCCAACAATCTTCAAGCCACAGTCCAAGAGAAGTCTCAGGAAA GCAGAGTGAAGGAAGAATCTTACGACTCAGGAAACGAACACCATCAGTAGGGAAGCTATGACACAC CCAAAACAGCAGGAGGTGATGAGAAAGACATGAAAGCATTTATGGGAACCTCAGTGCAGAAATTTGACCT GCCAGGAAATTTACCTGGCAGCAAAAGATGGCCACAAACTCTTAAGGAAAAGGCCAGGCTCTAGAAGAC CTGGCTGGCTTCAAAGAGCTCTTCCAGACACCAAGGCACTGACAAGCCACGACTGATGAGAAAACCTACCA AAATAGCCTGCAAACTCTCCACAACCAAGCCAGTGGACACCCAGCAAGCACAAGAACCGGCCCAAGAG AAACCTCAGGAAAGCAGAGCTAGAGGAAGAATTTTATGCACTCAGGAAACGAACACCATCAGCAGGCAAA GCCATGGACACACCAAAACCAAGCAGTAAAGTATGAGAAAATATCAACACATTTGTGGAACCTCAGTGC AGAACTGGACCTGCTAGGAAATTTACCTGGCAGCAAGAGACAGCCACAGACTCTTAAGGAAAAGGCTGA GGCTCTAGAGGACCTGGTTGGCTTCAAAGAACTCTTCCAGACACCAAGGTCACTGAGGAATCAATGACT GATGACAAAATCAGAGAAGTATCCTGTAATCTCCACAGCCAGAGTCATTCAAACCTCAAGAAGCTCCA AGCAAAAGGCTCAAGATACCCCTGGTGAAAGTGGACATGAAGAAGAGCCCTAGCAGTCAGCAAGCTCAC ACGGACATCAGGGGAGACTACGCAAAACACACAGAGCCAAACAGGAGATAGTAAGAGCATCAAAGCGTTT AAGGAGTCTCCAAAGCAGATCTTGGACCCAGCAGCAAGTGAACCTGGTAGCAGGAGGCAAGTGAAGACTC GTAAGGAAAAGGCCCGTGCCTAGAGAACTGGTTGACTTCAAAGAGCTCTTCTCAGCACCAAGGTACAC TGAAGAGTCAATGACTATTGACAAAACACAAAAATTCCTGCAAACTCTCCCCACCAAGAACTAACAGAC ACTGCCACGAGCACAAGAGATGCCCAAGACACGTCCAGGAAAGAAAGTAAAGAGGAGCTCTCAGCAG TTGAGAGGCTCACGCAAAACATCAGGGCAAGCACACACACACAAAGAACCAAGCAAGCGGTGATGAGGG CATCAAAGTATTGAAGCAACGTGCAAGAGAAAGAAACCAAAACCCAGTGAAGAGGAACCCAGCAGGAGAAGG CCAAGAGCACCTAAGGAAAAGGCCCAACCCCTGGAAGACCTGGCCGGCTTCACAGAGCTCTCTGAAACAT CAGGTACACTCAGGAATCACTGACTGCTGGCAAGGCCACTAAAAATACCTTGCAGATCTCCCCACTAGA AGTGGTAGACCCACAGCAAGCACAAGAGGCACTCTCAGGACACGTGTGAGAGAAGTCAAGATTAAGAA GAGCTTCAGCAGTCAAGTTACACAAAACATCAGGGGAAACCCAGGATGCAGACAAAGAACCAAGCAGGTG AAGATAAAGGCATCAAAGCATTAAGGAATCTGCAAAACAGACACCGGCTCCAGCAGCAAGTGTAACTGG CAGCAGGAGACGGCCAAGAGCACCCAGGGAAAGTGCCCAAGCCATAGAAGACCTAGCTGGCTTCAAAGAC CCAGCAGAGGTACACTGAAGAAATCAATGACTGATGACAAAACCACTAAAAATACCTGCAAAATCATCAC CAGAACTAGAAGACACCGCAACAAGCTCAAAGAGACGGCCAGGACACGTGCCAGAAAGTGAAGTGAAG GGAGGAGCTGTAGCAGTTGGCAAGCTCACACAAACCTCAGGGGAGACACGACACCCAGCAAAAGAGCCG GTAGGTGAGGGCAAGGCAAGCAAGCATTTAAGCAACCTGCAAGCGGAAGCTGGACGCAGAAGATGTAA	
--	--

	TTGGCAGCAGGAGACAGCCAAAGAGCACCTAAGGAAAAGGCCAACCCCTGGAAGATCTGGCCAGCTTCCA AGAGCTCTCTCAAACACCAGGCCACACTGAGGAACCTGGCAAAATGGTGTCTGATAGCTTTACAAGCGCT CCAAAGCAAAACACCTGACAGTGGAAAACCTCTAAAAATAATCCAGAAAGAGTTCTTCGGGCCCTTAAAGTAG AACCCGTGGGAGACGTGGTAAGCACCAGAGACCCTGTAAATACAAAAGCAAAAGCAACACTTCCCTGGC CCCACTGCCCTTCAAGAGGGAGGTGGCAAGATGGAAGCGTACCGGGAACCAAGAGGCTGCGCTGCATG CCAGCACCAGAGGAAATTGTGGAGGAGCTGCCAGCCAGCAAGAAGCAGAGGGTTGTCTCCAGGGCAAGAG GCAAAATCATCCGAACCCGTGGTTCATCATGAAGAGAAGTTTGAGGACTTCTGCAAAAAGAAATGAACCTGC GGAAGAGCTGAACAGCAACGACATGAAAACCAACAAAGAGGAACACAAATTACAAGACTCGGTCCCTGAA AATAAGGGAATATCCCTGCGCTCCAGACGCCAAAATAAGACTGAGGCAGAACAGCAAAATAACTAGGTCT TTGTATTAGCAGAAAGAATAGAAATAAACAGAAATGAAAAGAAGCCCATGAAGACCTCCCCAGAGATGGA CATTGAGAAATCCAGATGATGGAGCCCGGAAACCCATACCTAGAGACAAAGTCACTGAGAACAAAAGGTGC TTGAGGCTGTCTAGACAGAATGAGAGCTCCAGCCTAAGTGGCAGAGGAGAGCGAGGGCAGAAGAGTG CGAAGGTTCTCATGCAGAAATCAGAAAGGGAAGGAGAAGCAGGAATTCAGACTCCATGTGCTGAGTATC AAGAAAGACAAAAAGCCAGCCTGCAGCAAGCACTTTGGAGAGCAAAATCTGTGCAGAGAGTAACGCGGAGT GTCAGAGAGGTGTGCAGAAATCCAAAGAAGGCTGAGGACAAATGTGTGTGTCAAGAAAATAAGAACAGAA GTCATAGGGACAGTGAAGATATTGACAGAAAAATCGAAGCTGGGAAAAATATAATAAGTTAGTTTTGTG ATAAGTTCTAGTGCAGTTTTTGTGCATAAATACAAGTGAATTTCTGTAAGTAAGGCTGTGCATCTGCTTAA GGGAAGAAAACCTTTGGATTGTCTGGGTCTGAATCGGCTTCATAAACTCCACTGGGAGCACTGCTGGGCTC CTGACTGAGAATAGTTGAACACCGGGGGCTTTGTGAAGGAGTCTGGGCCAAGGTTGCCCTCAGCTTTG CAGAATGAAGCCTTGAGGCTGTCAACCCACAGCCACCCTACAGCAGCCTTAAGTGTGACACTTGGCA CACTGTGTGCTCGTTTGTTCCTATGTCTCCAGGGCAGGTTGGCAGGAACAACATCTGCTGTCTCC CAACACTGAGCAGGCACTCGGTAAACACGAATGAATGGATGAGCGCAGGATGAATGGAGCTTACAAGAT CTGTCTTTCCAATGGCCGGGGGCTTTGGTCCCCAAATTAAGGCTATTGGACATCTGCACAGGACAGTCC TATTTTIGATGTCCTTTCTTTCTGAAAAATAAGTTTTGTGCTTTGGAGAAATGACTCGTGAGCACATCTT TAGGGACCAAGAGTGACTTTCTGTAAAGGAGTGACTCGTGGCTTGCCTTGGTCTCTTGGGAATCTTTCT AACTAGGGTTGCTCTCACCTGAGACATTTCTCCACCCGCGGAATCTCAGGGTCCCAGGCTGTGGGCCATCA CGACCTCAAACCTGGCTCCTAATCTCCAGCTTTCTGTCTATTGAAAGCTTCGGAAGTTTACTGGCTGTGCT CCCGCTGTTTTCTTTCTGACTCTATCTGGCAGCCGATGCCACCAGTACAGGAAGTGACACCAGTACT CTGTAAAGCATCATCATCTTTGGAGAGACTGAGCACTCAGCACCTCAGCCACGATTTGAGGATCGCTTC CTTGTGAGCCGCTGCCTCCGAAATCTCTTTGAAGCCAGACATCTTTCTCCAGCTTCAGACTTGTAGAT ATAACTCGTTTCATCTTACATTTTCCACTTTGCCCCCTGTCTCTCTGTGTTCGCCAAATCAGAGAAT AGCCCCCATCCCCAGGTCACCTGTCTGGATTCTCCCATTCACCCACCTTGCAGGTGCAGGTGAGG ATGGTGCACACAGACAGGAGTGTCTGCCCCAAATGTGCCCTGTGCGGGCAGTGCCCTGTCTCCAGCTTT GTTTCCCCAGTGTCTGGCGGGGAGCCAGGTGACATCATAAATACTTGTGAATGAATGCAGAAATCAGCG GTACTGACTTGTACTATATTGGCTGCCATGATAGGTTCTCACAGCGTTCATCCATGATCGTAAGGGAGAA TGACATTTCTGCTTGAGGGAGGGAATAGAAAGGGGCAGGGAGGGGACATCTGAGGGCTTCACAGGGCTGCA AAGGGTACAGGGATTGCAACAGGGGCAAGCAGGGGAGGGTGTTCAGGAAGAGTGCTCTTAGCAGAGGC ACTTTGGAAGGTGTGAGGCATAAATGCTTCTCTACGTAGGCCAACCTCAAACCTTTAGTAGGAATGT TGGTATGATCAAGTTGTTCTAACACTTTAGACTTAGTAGTAATATGAACCTCACATAGAAAAATTCAT CCAGCCATATGCCGTGGAGTGGAATATCTGTTTAGTAGAAAAATCCTTTAGAGTTTCAAGCTTCAACAG AAATCTTGTGAAGTATGTCAGCACCTTTTCTCACCCCTGGTAAGTACAGTATTTCAAGAGCACGTTAAGG GTGGTTTTTCAATTTACAGGGCTGTGATGATGGGTAAAAATGTTTCAATTTAAGGGCTACCCCCGTGTTTA ATAGATGAACACCACTTCTACACACCCCTCTTGGTACTGGGGGAGGGAGAGATCTGACAAATACTGCC ATTCCCCTAGGCTGACTGGATTGAGAACAAATACCCACCATTTCACCATGGTATGGTAACTTCTCTG AGCTTCAGTTTCCAAGTGAATTTCCATGTAATAGGACATTOCCATTAAATAAAGCTGTCTTTTACTTTTT CGCCTCCCAGGGCTGTGGGATCTGGTCCCCAGCCTCTCTTGGGCTTTCTTACACTAACTCTGTACCTA CCATCTCCTGCCTCCCTTAGGCAGGCACCTCCAACCCACACACACTCCCTGCTGTTTCCCTGCCTGGAA CTTTCCCTCCTGCCCCACCAAGATCATTTTATCCAGTCTTGAGCTCAGCTTAAGGGAGGCTTCTTGCCGT TGGGTTCCCTCACCCCATGCTGTCTCCAGGCTGGGGCAGGTTCTTAGTTTGCCTGGAATGTCTGT ACCTCTTTGTAGCACGTAGTGTGTGGAACTAAGCCACTAATTGAGTTTCTGGCTCCCCTCCTGGGGTT GTAAGTTTTGTTCATTCATGAGGGCCGACTGCATTTCTGGTTACTCTATCCAGTGACCAGCCACAGGA GATGTCCAATAAAGTATGTGATGAAATGGTCTTAAAAA	
NM_024101	GCGCCGGGACGTGGCCAGTTGCCCGCTGCCCCGAGAGCCAGGCGCTAACCAGCCGCTCTGCGCCCGC GCCCTGCTTGCCCCCATTATCCAGCCTTGCCCCGGCGCCCTGACCTGACGCCCTGGCCTGACGCCCTGCT TCGTGCGCTCCTTTCTCTCCAGGTGCTGGACAGGGACTGAGCGTCCCCCGGAGAGGGTCCGGTGTGAC CCCGACAAGAAGCAGAAATGGGGAAGAACTGGATCTTTCAAGCTCACTGATGAAGAGGCCAGCATGT CTTGAAGTTGTTCAACGAGATTTTGACCTCCGAAGGAAAGAGAGGAACGGCTAGAGGGCTTGAAGGGC AAGATTAGAAGGAAAGCTCCAAGAGGGAGCTGCTTTCCGACACTGCCATCTGAACGAGACCCACTGCG CCCGCTGCCTGCAGCCCTACCAGTGCTTTGTGAATAGCAAAAGGCAAGTGCTGGAATGTGGCCTCTTCAC CTGCAAAAGCTGTGGCCGCTCCACCCGGAGGAGCAGGGCTGGATCTGTGACCCCTGCCATCTGGCCAGA GTGCTGAAGATCGGCTCACTGGAGTGGTACTATGAGCATGTGAAAGCCCGCTTCAAGAGGTTGCGAAGTG CCAAGGTTCATCCGGTCCCTCCACGGGCGGCTGCAGGGTGGAGCTGGGCTGAACTGATATCTGAAGAGAG AAGTGGAGACAGCGACGACAGAGATGAGGATGGAGAACCCTGGCTCAGAGGCCAGGCCAGGCCAGGCC TTTGGCAGCAAAAAAAGCGCCTCCTCTCCGTCCACGACTTCGACTTCGAGGGAGACTCAGATGACTCCA CTCAGCCTCAAGGTCACTCCCTGCACCTGCTCTCAGTCCCTGAGGCCAGGACAGCCACAGTCCCTCAC AGATGAGTCTGTCTCAGAGAAGGCAGCCCTCACAAAGGCTGAGGGCTGGAGGAGGCTGATACTGGGGCC TCGGGTGCCACTCCCATCCGAAGAGCAGCCGACAGCATCTCAGCTTCCAGACAGGGCCCTGGCTG AGCTCTGCCCGCTGGAGGCTCCACAGGATGGCCCTGGGACTGCTGCTGCACTCGGGTCAATGTCAT	126

	CAGGAATGAGCAGCTGCCCTGCAGTACTTGGCCGATGTGGACACCTCTGATGAGGAAAGCATCCGGGCT CACGTGATGGCCTCCACCATTCCAAGCGGAGAGGCCGGGCGTCTTCTGAGAGTCAGATCTTTGAGCTGA ATAAGCATATTTAGCTGTGGAATGCCTGCTGACCTACCTGGAGAACACAGTTTGTGCTCCCTTGGCCAA GGGTCTAGGTGCTGGAGTGGCAGCGAGGCCGATGTAGAGGAGGAGGCCCTGAGGAGGAAGCTGGAGGAG CTGACCGACAAAGTCAGTGACCGAGGACCTCGTCCGAGGAGGAGGAAGCCAAAGGACGAAAGGCCAGAGC CCAACAGGGACAAATCAGTTGGGCTCTCCCCAGGCGGACCCGGAGGTGGGCACGGCTGCCATCAAAC CAACAGACAGGAAAAAGCCCCAGGACCTGGGGACCCCGTCCAGTACAAACAGGACCACAGATGAGGAG CTGTGAGAGCTGGAGCACAGAGTGGCAGTGACGGCTCAGAAGTCCAGCAGGACAGAGCGAGGTTCAG ACATTGAATCCAGGATTGACGCCCTGAGGGCCGAGGGCTCACGGTGAAGCCCTCGGGAAGCCCCGGAG GAAGTCAAACCTCCGATATTTTCCCCCTGAGTGGCTGGGAAACTTGGCAAGAGACCAGAGGACCCAAAT GCAGACCCCTTCAAGTGAGGCCAAGGCAATGGCTGTGCCCTATCTTCTGAGAAGAAAGTTCAGTAATTCCT TGAAAAGTCAAGGTAAAGATGATGATTCITTTGATCGGAAATCAGTGTACCGAGGCTCGCTGACACAGAG AAACCCCAACGCGAGGAAAGGAATGGCCAGCCACACCTTCGCGAAACCTGTGGTGGCCCCAGCTCTAA CGGGACAGGACAGAGAGACAGAGCAGCCCTGCACTGTTTTCCCTCCACCACAGCCATCTGTCCCTCATT GGCTCTGTGCTTTTCCACTATACAGAGTCACCGTCCCAATGAGAAACAAGAGGAGACCCCTCCACATGGA CTCCCCACCTGCAAGTGGACAGCGACATTCAGTCTGCACTGCTCACCTGGGTTTACTGATGACTCCTGGC TGCCCCACCATCTCTCTGATCTGTGAGAAACAGCTAAGCTGCTGTGACTTCCCTTAGGACAATGTTGT GTAAATCTTTGAAGGACACACCGAAGACCTTTTACTGTGATCTTTTACCCCTTTTACTCTTGGCTTTCT TATGTTGCTTTTCAATGGAATGGAAGAAAGATGACTCAGTTAAGGCACCCAGCCATATGTGTAATCTTG ATGGTCTATATCGGGGTGTGAGCAGATGTTTGGTATTTCTTGTGGGTGTGACTGGATATAGACATCCG GACAAGTGACTCAACTAATGATCTGCTGAATAATGAAGGAGGAATAGACACCCAGTCCCCACCTACCT GCACCCGCTCTGCAAGTTCCCATGTGATCTGTAGACCAGGGGAAATTACACTGCGGTCAAGGGCAGAGCC TGACATGACAGCAAGTGAGCATTGATAGATGCTCAGATGCTAGTGACAGAGAGCTGCTGGGAGACGAA GAGACAGCAGGCAGAGCTCCAGATGGGCAAGGAAGAGGCTTGGTTCAGCCTGGCTCTGCCCTCACTGC AGTGGATCCAGTGGGGCAGAGGACAGAGGGTCACAACCAATGAGGAGATGCTGCCAAGGATGGGGTGCA GAGGCCACAGGAGTCAGCTTGGCACTCGCCCATTTGGTTACATAGATGATCTCTCAGACAGGCTGGGACTC AGAGTTATTTCTAGTATCGGTGTGCCCATCCAGTTTTAAGTGGAGCCCTCCAAGACTCTCCAGAGCTG CCITTGAACATCCTAACAGTAATCACATCTCACCTCCCTGAGGTTTACITTAGACAGGACCCAAATGGCT GCAGTCCCTTTGTGAGAGGGGGTGTGAGAGGAGTGGCTTCTTTAGAATCAAAACAGTAGACACAGAGT CAAGCCTTGTGCTTCAAGCATTGACCAAGTTAAGTGTTCCTTCCCTCTCTCAATAAGACACTTCCAGG AGCTTTCCAATCTCTCACTTAAACTAAGGTTTGAATCTCAAAGTGTGTGCTGGGAGGCTGATACTCTGC AACTTCAGGAGACCTGTGAGCACACATTAGCAGCTGTTTTCTGACTCCTTGTGGCATCAGATAAAAAAG TGGGAGTTTTTCCATATAATTTCCAGCCTTACTTATAAATCTATTTCTTTGAAAAAATTATTCAGGTAG GTAAGGTGGCTCATACCTATAATCCCAGCCCTTTGAGAGGCCAAGGTGGGAGAATTGCTTGAGGCCAGGA GTTTGAGACCTCCTGGGCAACATAGTGAGATCCCATCTCTACAAAAAACAAAAACAAAAAATTACCCAAG CATGATGGTATATGCCTGTAGTCGTACCTTACTTACTTAGGAGGCTGAGGCAGGAGGATCACTTGAGCCCT GGAGGTTGGGGCTGCAGTGAGGCTGATCGCATCACTATACTCGAGCCTGGGCAACAGAGTGAGACCTTG TCCTTAAAAAATTATAATAAATAAATGAAATAATTCTTCAGAAAAAATAAAAAA	
NM_005940	AAGCCAGCAGCCCGGGGCGGATGGCTCCGGCCGCTGGCTCCGCAGCGCGGCCGCGCGCCCTCCTG CCCCGATGCTGCTGCTGCTCCAGCCCGCCGCTGCTGGCCCGGGCTCTGCCCGCGGACGCCACCC ACCTCCATGCCAGAGAGGAGGGGCCACAGCCCTGGCATGCAGCCCTGCCAGTAGCCCGCTGCCCGCC TGCCACGCAGGAAGCCCCCGGCCCTGCCAGCAGCCTCAGGCCCTCCCGCTGTGGCGTGCCCGACCCATCT GATGGGCTGAGTGCCCGCAACCGACAGAAAGAGGTTCTGTGCTTTCTGGCGGGCGCTGGGAGAGACCGACC TCACCTACAGGATCCTTCGGTTCCCATGGCAGTTGGTGCAGGAGCAGGTGCGGCAGACGATGGCAGAGGC CCTAAAGGTATGGAGCGATGTGACGCCCTACCTTTACTGAGGTGCACGAGGGCGCTGCTGACATCATG ATCGACTTCGCCAGGTACTGGCATGGGGACGACCTGCCGTTTGAITGGGCTGGGGGCATCCTGGCCCATG CCTTCTTCCCCAAGACTCACCGAGAGGGGATGTCCACTTCGACTATGATGAGACCTGGACTATCGGGGA TGACCAGGGCACAGACCTGCTGCAGGTGGCAGCCCATGAATTTGGCCACGTGCTGGGGCTGCAGCACACA ACAGCAGCCAAGGCCCTGATGTCCGCCCTTCTACACCTTTCTGCTACCCACTGAGTCTCAGAGTCAGATGACT GCAGGGGCGTTCAACACCTATATGGCCAGCCCTGGCCACTGTCACTCCAGGACCCAGCCCTGGGCCC CCAGGCTGGGATAGACACCAATGAGATTGCACCGCTGGAGCCAGACGCCCGCCAGATGCCGTGTGAGGCC TCCTTTGACGCGGTCTCCACCATCCGAGGCGAGCTCTTTTCTTCAAAGCGGGCTTTGTGTGGCGCTCC GTGGGGGCCAGCTGCAGCCCGGCTACCCAGCATTGGCCTCTCGCCACTGGCAGGGACTGCCAGCCCTGT GGACGCTGCCTTCGAGGATGCCAGGGCCACATTGGTTCTTCAAAGGTGCTCAGTACTGGGTGTACGAC GGTGAAAAGCCAGTCCCTGGGCCCGCACCCCTCACCGAGCTGGGCTGGTGAGGTTCCCGGTCCATGCTG CCTTGGTCTGGGGTCCCGAGAAGAACAGATCTACTTCTTCCGAGGCAGGACTACTGGCGTTTCCACCC CAGCACCCGCGGTGAGACAGTCCCGTGCCCGCAGGGCCACTGACTGGAGAGGGGTGCCCTCTGAGATC GACGCTGCCTTCCAGGATGCTGATGGCTATGCCTACTTCTGCGCGGCCGCTCTACTGGAAGTTTGACC CTGTGAAGGTGAAGGCTCTGGAAGGCTTCCCCCGTCTCGTGGGTCTGACTTCTTTGGCTGTGCCGAGCC TGCCAACTTTCTCTGACCATGGCTTGGATGCCCTCAGGGGTGCTGACCCCTGCCAGGCCACGAATAT CAGGCTAGAGACCCATGGCCATGTTGTGGCTGTGGGCACAGGATGGGACTGAGCCCTGCTCTCTCA GGGGGATGGGGTGGGGTACAACCACCATGACAACCTGCCGGGAGGGCCACGCAGGTCTGTGTCACCTGCCA GCCACTGTCTCAGACTGGGCAGGGAGGCTTTGGCATGACTTAAGAGGAAGGGCAGTCTTGGGCCCGCTAT GCAGGTCTCTGGCAAACCTGGCTGCCCTGTCTCCATCCCTGTCCCTCAGGGTAGCACCATGGCAGGACTGG GGGAACCTGGAGTGTCTTGTGTATCCCTGTGTGAGGTTCTTCCAGGGGCTGGCATGAAGCAAGGGT GCTGGGGCCCCATGGCCTTCCAGCCCTGGCTGAGCAACTGGGCTGTAGGGCAGGGCCACTTCTTGAGGTCA GGTCTTGGTAGGTGCCTGCATCTGTCTGCCCTCTGGCTGACAATCTTGGAATCTGTCTCCAGAAATCCA GGCCAAAAAGTTCACAGTCAAATGGGGAGGGGTATCTTTCATGACAGGAGACCCAGGCCCTGGAGGCTGC	127

	AACATACCTCAATCCTGTCCCAGGCCGGATCCTCCTGAAGCCCTTTTCGCAGCACTGCTATCCTCCAAAG CCATTGTAAATGTGTGTACAGTGTGTATAAACCTTCTTCTCTTTTTTTTTTTTAACTGAGGATTGTC	
NM_002467	GACCCCGAGCTGTGCTGCTCGCGGCCGCCACCGCGGGCCCCGGCGCTCCCTGGCTCCCTCCTGCCTC GAGAAGGGCAGGGCTTCTCAGAGGCTTGGCGGGAAGAACGGAGGGAGGATCGCGCTGAGTATAAAA GCCGGTTTTTCGGGGCTTTATCTAACTCGCTGTAGTAATTCAGCGAGAGGCAGAGGGAGCGAGCGGGCGG CCGGCTAGGGTGGAAAGCCGGCGCAGCAGAGCTGCGCTGCGGGCGTCTTGGGAAGGGAGATCCGGAGCG AATAGGGGGCTTCGCCTCTGGCCAGCCCTCCCGCTGATCCCCAGCCAGCGGTCCGCAACCCCTTGGCGC ATCCACGAAACTTTGCCCATAGCAGCGGGCGGGCACTTTGCACTGGAACITACAACACCCGAGCAAGGAC GCGACTCTCCCGACGCGGGGAGGCTATTCTGCCCATTTGGGGACACTTCCCCCGCGCTGCCAGACGCCG TTCTCTGAAAGGCCTCTCTTGACGCTGCTTAGACGCTGGATTTTTTCGGGTAGTGGAAAACAGCAGCC TCCCGCGCAGCATGCCCTCAACGTTAGCTTCACCAACAGGAACATGACCTCGACTACGACTCGGTGCAG CCGTTTTCTACTGCGACGAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCGAGCTGCAGCCCCCGG CGGCCAGCGAGGATATCTGGAAGAAATTCGAGCTGCTGCCACCCCGCCCTGTCCCTCCTGCTC CGGGCTCTGCTCGCCCTCTACGTTGCGGTACACCCCTTCTCCCTTGGGGAGACAACGACGGCGGTGGC GGGAGCTTCTCCACGGCCGACAGCTGGAGATGGTGACCGAGCTGCTGGGAGGAGACATGTTGAACAGA GTTTCATCTGCGACCCGGACGACGAGACCTTATCAAAAAACATCATCATCCAGGACTGTATGTGGAGCGG CTTCTCGGGCGCCCAAGCTCTGCTCAGAGAAGCTGGCTCTCTACAGGCTGCGCGCTGCGCGAGCGGC AGCCCGAACCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTGTACCTGCAGGATCTGAGCGCCG CCGCCTCAGAGTGATCGACCCCTCGGTGGTCTTCCCTACCCCTCTCAACGACAGCAGCTCGCCCAAGTC CTGCGCCTCGCAAGACTCCAGCGCTTCTCTCCGTCTCGGATTCTCTGCTCTCTCGACGAGTCTCTC CCGACGGGACGCCCGAGCCCTGGTGTCTCCATGAGGAGACACCGCCCAACACGACGAGCAGCTGAGG AGGAACAAGAAGATGAGGAAGAAATCGAIGTTGTTTCTGIGGAAAAGAGGCAGGCTCCTGGCAAAAGGTC AGAGTCTGGATCACCTTCTGCTGGAGGCCACAGCAAACTCTCTCAGCGCCACTGGTCTCTCAAGAGGTGC CACGTCTCCACACATCAGCAAACTACGCGAGCGCTCTCTCCACTCGGAAGGACTATCTGCTGCCAAGA GGGTCAAGTTGGACAGTGTACAGTCTGAGACAGATCAGCAACAACTGCAACAGCTTCTGCGGTAA CTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACAACGTCTTGGAGCGCCAGAGGAGGAACGAGCTA AAACGGAGCTTTTTGCCCCGCGTGACAGATCCCGGAGTTGGAAAACAATGAAAAGGCCCCCAAGGTAG TTATCCTTAAAAAAGCCACAGCATACATCTGTCCGTCCAAGCAGAGGAGCAAAAGCTCATTCTGGAAGA GGACTTGTGCGGAACGACGAGAAGCTTGAACACAACTTGAACAGCTACGGAACCTTCTGCGGTAA GGAAAAGTAAGGAAAACGATTCTTCTTAACAGAAATGCTCTGAGCAATCACCTATGAACCTGTTTCAAT GCATGATCAAAATGCAACCTCACAACCTTGGCTGAGTCTTGGAGCTGAAAGATTTAGCCATAATGTAACCT GCCCAAAATGGACTTTGGGCATAAAAGAACTTTTTATGCTTACCATCTTTTTTTTTCTTTAACAGAT TTGTATTTAAGAATTGTTTTTAAAAAATTTAAGATTTACACAATGTTTCTCTGTAATATTTGCCATTAA ATGTAATAACCTTAAATAAACGTTTATAGCAGTTACACAGAATTTCAATCCTAGTATATAGTACCTAGT ATTATAGGTACTATAAACCTTAATTTTTTTTATTTAAGTACATTTTGCTTTTAAAGTTGATTTTTTCT ATTGTTTTTAGAAAAATAAAATAACTGGCAATATATCATTGAGCCAAATCTTAAAAAATAAAAAA	128
BC013732	GTGGGAGGATTGCATTCAGTCTAGTTCCCTGGTTGCGGGCTGAAATAACCTGCTCTCCAAATGTCTCCAA AAGTGACTTAAGTCAGGTTCCCCCAAAACGACACCAAGACAAGAATCCATGTGTGTGTGACTGAAGGAA GTGCTGGGAGAGCCCCAGCTGCAGCCTGGATGTGAACCTGCAACTCCAAAGTGTGTCCAGACTCAAGGCCAA GGGCACTAGGCTTTCCAGACCTCTACTAAGTCATTGATCCAGCACTGCCCTGCCAGGACATAAATCCCT GGCACTCTGTCTCTGCAAAAGGAGGGCAAGCAGCTTCAGGAGCCCTTGGGAGTCTCTCCAAAGAGT CTAGGGTACAGGTCGGAAGTAGAAGAACACAGAAGGCAGGCCAGGGGCACTGTGAGATGGTAAAGAGA TCTGAAGGATCCAGAATTCAGGCCAGGAAGAAGCAGCAATCTGTCTTCTGGATTAAACTGAAGATCAA CCTACTTTCAACTTACTAAGAAAGGGGATCATGGACATTGAAGCATATCTTGAAGAATTTGGCTATAAGA AGICTAGGAACAAATGGACTTGGAAACATTAAGTATATCTTCAACACAGACTCCGAGCTTCTCCCTT TGAGAACCCTTAACATCCATTGTGGGATGCCATGGACTTAGGCTTAGAGGCCATTTTGTACAGTTGTG AGAAGAAATCGGGTGGATGGTGTCTCCAGGTCAATCATCTTCTGTACTGGGCTCTGACCACTATTGGTT TTGAGACCACGATGTGGGAGGGTATGTTTACAGCACTCCAGCCAAAAATACAGCACTGGCATGATTCA CCTTCTCCTGCAGGTGACCATTTGATGGCAGGAACATACATTGTGATGCTGGGTTTGGACGCTCATACCCAG ATGTGGCAGCCTCTGGAGTTAATTTCTGGGAAGGATCAGCCTCAGGTGCCTTGTGTCTTCCGTTTGACGG AAGAGAATGGATTCTGGTATCTAGACCAATCAGAAGGGAACAGTACATTCCAAATGAAGAATTTCTTCA TTCTGATCTCCTAGAAGACAGCAAAATACCGAAAAATCTACTCCTTTACTCTTAAGCCTCGAACAATTGAA GATTTTGAGTCTATGAATACATACTGTGAGACATCTCCATCATCTGTGTTTACTAGTAAATCATTTTGTT CCTTGACAGCCCCAGATGGGGTTCACTGTTTGGTGGGCTTCAACCTCACCCATAGGAGATTCAATTATAA GGACAATACAGATCTAATAGAGTTCAAGACTCTGAGTGAGGAAGAAATAGAAAAAGTGTCTGAAAAATATA TTTAATATTTCTTGACAGAGAAAGCTTGTGCCCAACATGGTGATAGATTTTTTACTATTTAGAATAAGG AGTAAAAACAATCTGTCTATTTGTATCCAGCTCACCAGTTATCAACTGACGACCTATCATGTATCTTCT GTACCCCTTACCTTATTTTGAAGAAATCTAGACATCAAAATCATTTCACCTATAAAAAATGTCATCATATA TAATTAACAGCTTTTTTAAAGAAACATAAACCAAACTTTTCAATAATAATAATAATAATAATAATAA ATGTCTTTTTAAAGATGGCTGTGGTTATCTTGGAAATTTGGTATTTATGCTAGAAAGCTTTTAAATGTTGG TTTATTTGTTGAATTCCTAGAAAAGTTTTATGGGTAGATGAGTAAATAAAATATTTGAAAAAACTTATTG TCTATAAAGTATATTAACATTTGTTGGCTAATATAAAAAAATAAAAAA	129
NM_014321	GCGCGCGGGTTTCGTTGACCCGCGCGGTTTACGGGAATTGTTGCTTTAGTGCCGGCGCCATGGGGTCGG AGCTGATCGGGCGGCTAGCCCCGCGCTTGGGCTCGCCGAGCCCGACATGCTGAGGAAAGCAGAGGAGTA CTTGGCCCTGTCCGGGTGAAGTGTGTGCGCTCTCCGACGCAACACGAGACCGAGTGCAGTGCATG TGCTGAGACCTTGCAGCTTCTGGATGAAGTGCCCTTGGACAGGCTTATTTAATTAACTTTCTGGTT TGAACAAGGAGACATATCAGAGCTGTCTTAAATCTTTTGGTGTGTTTACTGGGCTGAATTCAAATATTGG AATAAGAGACCTAGCTGTACAGTTTAGCTGTATAGAAGCAGTGAACATGGCTTCAAAGTACTAAAAAGC	130

	TATGAGTCCAGTCTTCCCCAGACACAGCAAGTGGATCTTGACTTATCCAGGCCACTTTTCACCTTCTGCTG CAGTCTTTTCAGCATGCAAGATTCTAAAGCTGAAAGTGGATAAAAAACAAAATGGTAGCCACATCCGGTGT AAAAAAGCTATATTTGATCGACTGTGTAAACAACATAGAGAAGATTGGACAGCAGGTCGACAGAGAACCT GGAGATGTAGCTACTCCACCACGGAAGAGAAAGATAGTGGTTGAAGCCCCAGCAAAGGAAATGGAGA AGGTAGAGGAGATGCCACATAAAACCACAGAAAGATGAAGATCTGACACAGGATTATGAAGAATGGAAAAG AAAAATTTTGGAAAATGCTGCCAGTGTCTAAAAGGCTACAGCAGAGTGATTTCAGCTTCCAACTGCTAT ACATTCCAAACTGATAGTACATTGCCATCTCCAGGAAGACTTGACGGCTTTGGGATTTTGTTTAAACTTT TATAATAAGGATCCTAAGACTGTTCCTTTAAATAGCAAAGCAGCCTACCTGGAGGCTAAGTCTGGGCAG TGGCTTGGCCCCCTGGTGTGAGCATTAGACCAGCCACAGTGCTGATTGGTATAGCCTTATGTGCTTTCCCT ACAAAATGGAATTGGAGGCCGGCGCAGTGGCTCACGCCCTGTAATCCAGCACTTTGGGAGGCCAAGGTG GGTGGATCACCTGAGGTCAGGAGCTCGAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAA ATACAAAAATTAGCCAGGTGTGATGGTGCATGCCCTGTAATCCAGCTCCICAGTAGGCTGAGACAGGAGC ATCACTTGAACGTGGGAGGCAGAGTTGCAGTGAGCCGAGATTGCACCACCGCACTCCAGCTGGGTGAC AGAGCGAGACTTATCTCATAAATAAATAGATAGATACTCCAGCCTGGGTGACAGAGCGAGACTTATAGAT AGATAGATAGATAGATGGATAGATAGATAGATAGATAGATAGATAAACCAGGATTTGGAGCCATTTTG CTTTAAGTGAATGGCAGTCCCTTGTCTTATTAGAAATATAAAATTCAGTCTGAATGGCATCTTACAGATT TTACTTCAATTTTGTGTACGGTATTTTTATTTGACTAAATCAATATATTGTACAGCCTAAGTTAATAA ATGTTATTTATATATGCAAAAAAAAAAAAAAAAAA	
NM_000926	AGTCCACAGCTGTCACTAATCGGGGTAAAGCCTTGTGTATTGTGCGTGTGGGTGGCATTCTCAATGAGA ACTAGCTTTCACCTTGTCACTTTGAGTGAATCTACACCCGAGGCGGCTAGTGTCTCCGCACTACTGGGATC TGAGATCTTCGGAGATGACTGTTCGCCCGCAGTACGGAGCCAGCAGAAGTCCGACCTTCTCGGAAATGGG CTGTACCGAGAGGTCGACTAGCCCCAGGGTTTTAGTGAGGGGGCAGTGGAACTCAGCGAGGGACTGAGA GCTTCACAGCATGCACGAGTTTGATGCCAGAGAAAAGTGGGAGATAAAGGAGCGCGTGTCACTAAAT TGCCGTTCGAGCCGCAGCCACTCAAGTGCCGGACTTGTGAGTACTTGCCTCTCCAGTCTCTCGACAGAA GTTGGAAACTCTCTTGGAGAACTCCCGAGTTAGGAGACGAGATCTCTTAACAATTTCTACTTTCTTCTT GCGCTCCCCACTTGCCGCTCGCTGGGACAAACGACAGCCACAGTTCCTTGACGACAGGATGGAGGCCAA GGGCGAGGAGCTGACCAGCGCCGCCCTCCCCCGCCCCGACCCAGGAGGTGGAGATCCCTCCGCTCCAGCC ACATTCAACACCCACTTTCTCTCCCTCTGCCCCATATTTCCCGAAACCCCTCTCTCTCTCTCTCTCTCTCT TCCCTCTGGAGACGGGGAGGAGAGAAAAGGGGAGTCCAGTCTGATGACTGAGCTGAGCTGAAGCAAGGTGCTC CGGGCTCCCCACGTGGCGGGGAGCCGAGCCTCGGACACCTTGCCCTGAAGTTTCGGCCATACCTATCTCTCTGGA CGGGCTACTCTTCCCTCGGCCCTGCCAGGGACAGGACCCCTCCGACGAAAAGACGCGAGGACCAGCAGTCTG CTGTGAGACGTGGAGGGGCGCATATTCAGAGCTGAAGCTACAAGGGGTGCTGGAGGACGAGCTTCTAGTC CCCCAGAAAAGGACAGCGGACTGCTGGACAGTGTCTTGGACACTCTGTTGGCGCCCTCAGGTCCCGGGCA GAGCCAAACCCAGCCCTCCCGCCTGCGAGGTCACCAAGCTCTTGGTGCCTGTTTGGCCCCGAACTTCCCGAA GATCCACCGGCTGCCCGCCGCCACCCAGCGGGTGTGTGCCCGCTCATGAGCCGTCGGGTGCAAGGTGTG GAGACAGCTCCGGGACGGGAGCTGCCATAAAGTGTGTCGCCCGGGGCTGTACCAAGCCGCGCTGCTGCT GCTCCCGGCTCTGAGAGCCCTCACTGGTCCGGGGCCCCAGTGAAGCCGCTCTCCGAGGCCGCTGCGGTG GAGGTTGAGGAGGAGGATGGCTCTGAGTCCGAGGAGTCTGCGGGTCCGCTTCTGAAGGGCAAACTCCGGG CTCTGGGTGGCGCGGGCGGTGGAGGAGGAGCCGCGGCTGTCCCGCGGGGGCGGCGAGCAGGAGGCGTCCG CCTGGTCCCCAAGGAAGATTCCCGCTTCTCAGCGCCAGGGTGCCTTGGTGGAGCAGCAGCGCCGATG GCGCCCGGGCGCTCCCGCTGGCCACCACGGTGTATGGATTTTATCCACGTGCTATCTCTGCTCTCAATC ACGCTTATTGGCAGCCCGCACTCGGCAGTGTCTGGAAGACGAAAGTTACGACGGCGGGGCGGGGCTGC CAGCGCTTTGCCCGCGCGGGAGTTACCCCTGTGCTCTGTCACCCCGGTGCTGTAGGCGCACTTCCCC GACTGCGCGTACCCGCGCGAGCGAGCCCAAGGACGAGCGGTACCCCTCTCTATAGCGACTTCCAGCGCG CCGCTCTAAAGATAAAGGAGGAGGAGGAAGGCGCGGAGGCGCTCCGCGCGCTCCCGCGCTTCTTACCTGT GGCGGTTGCCAACCCCGCAGCCTTCCCGGATTTCCCGTTGGGGCCACCGCCCGCGTCCCGCGCGGAGCG ACCCCATCCAGACCCGGGAAGCGCGGTGACGGCCGACCCGCCAGTGCCTCAGTCTGCTCTGCTGCTCTCT CCTCGGGTTCAGACCTGGAGTGCATCTGTACAAAGCGGAGGGCGCGCGCCCGCCAGCAGGCGCGCTCGC GCCCGCGCTTGAAGGCGCGGGCGGAGCGGCTGCCTGCTCCCGCGGAGCGCTGCCCTCCACCTCC GCCTCTGCCGCGCGCGCGGGGCGGCCCCGCGCTCTACCTGCACTCGGCTCAACGGGCTCCCGCAGC TCGGCTACCAAGGCGCGGTGCTCAAGGAGGGCCTGCCGAGGCTTACCCGCCCTATCTCAACTACCTGAG GCCGATTTCAGAGCCAGCCAGAGCCCAATAACAGCTTCGAGTCAATACCTCAGAAGATTGTGTTAATC TGTTGGGATGAAGCATCAGGCTGTCAATATGGTGTCTTACCTGTGGGAGCTGTAAAGTCTTCTTTAAGA GGGCAATGGAAGGCGAGCACAACACTTATGTGCTGGAAGAAATGACTGCATCGTTGATAAAATCCGCGAG AAAAAAGTCCAGCATGTGCTTGTAGAAAGTGTGTGAGGCTGGCATGGTCTTGGAGGTGCAAAATTT AAAAAGTTCAATAAAGTCAGAGTTGTGAGAGCACTGGATGCTGTTGCTCTCCACAGCCAGTGGGCGTTT CAATGAAGCCAAAGCCCTAAGCCAGAGATTCACTTTTTTACCAGGTCAAGACATACAGTTGATTCCACC ACTGATCAACCTGTTAATGAGCATTGAACAGATGTGATCTATGAGGACATGACAAACACAAAACCTGAC ACCTCCAGTTCTTTGCTGACAAGTCTTAATCAACTAGGCGAGAGGCAACTTCTTTTCTAGTCAAGTGGT CTAAATCATTGCCAGGTTTTGCAAACTTACATATTGATGACGAGTAACCTCACTCAGTATCTTGGAT GAGCTTAATGGTGTTTGGTCTAGGATGGAGATCTTACAAACACGTCAGTGGGCGAGTGTGATTTTGGCA CCTGATCTAATACTAAATGAACAGCGGATGAAAGAATCATCTATTCTATTATGCTTACCATGTGGC AGATCCACAGGAGTTTGTCAAGCTTCAAGTTAGCCAAGAAGAGTTTCTCTGTATGAAAGTATTGTTACT TCTTAATAACAATTTCTTTGGAAGGCTACGAAGTCAAAACCCAGTTTGGAGAGATGAGGTCAAGCTACATT AGAGAGCTCATCAAGGCAATTGGTTTGGGCAAAAAGGAGTTGTGTGAGCTCACAGCGTTTCTATCAAC TTACAAAACCTTCTTGATAACTTGCATGATCTTGTCAAACAACTTCATCTGTACTGCTGGAATACATTAT CCAGTCCCGGGCAGTGTGTTGAATTTCCAGAAATGATGCTGAAGTTATGTGTGACAAATTACCAAG	131

ATATTGGCAGGGATGGTGAACCCCTTCTCTTTCATAAAAAGTGAATGTCATCTTTTTTCTTTTAAAGCAAT TAAATTTTGGGTATGTCITTTTTGTTTTGGTTCAGGATTATGAGGICTTGAGTTTTTATAATGTTCTTCTG AAAGCCCTTACATTTATAACATCATAGTGTGTAAATTTAAAGAAAAAATTGTGAGGTCTAATTTATTTTCT TTTATAAAGTATAATTAGAATGTTTAACTGTTTTGTTTACCCATATTTTCTTGAAGAATTTACAAGATTG AAAAAGTACTAAATTTGTTAAAGTAAACTATCTTATCCATATTATTTTCATACCATGTAGGTGAGGATTTT TAACTTTTGCATCTAACAAATCATCGACTTAAGAGAAAAAATCTTACATGTAATAACACAAAGCTATTAT ATGTTATTTCTAGGTAACCTCTTTGTGTCAATTATATTTCCAAAAATGAACCTTTAAAAATGGTATGCAA AATTTTGTCTATATATTTTGTGTGAGGAGGAAATTCATAACTTTCCCTCAGATTTTCAAAAGTATTTTATA ATGCAAAAAATGTAGAAAGATTTTAAACCCTAAATATAGATTGATGTTCTTCAAACCTAGGCAAAACAAAC TCATATGTTAAGACCATTTTCCAGATTGGAAACACAAATCTCTTAGGAAGTAAATAAGTAGATTTCATATC ATTATGCAAAATAGTATTGTGGGTTTTGTAGGTTTTTAAATAACCTTTTTTGGGGAGAGAATTGTCTCT AATGAGGTATTGCGAGTGGACATAAGAAATCAGAAGATTATGGCCTAACGTACTCTTACCAACTGTGG CATGCTGAAAGTTAGTCACTCTTACTGATTCATCTCCTTCACTTTGAAAGTAGTAAATATCTTTTCC TGCCAAATGCTCTTTGGGTGAGAGCTTATTAACATCTTTTCAAATCAAAGGAAAGAAAGAGGAGAGG AGGAGGAGGAGGATCAATTCACATACCTTTCTCTCTTTATCTCCACTATCATGAATTCATATTTATG TTTAGCCATGCAAACTTTTTTACCATGAAATTTCTTCCAGAAATTTTCCCTTTGACACAAATTCATATG CATGTTTTCAACCTTCGAGACTCAGCCAAATGTCATTTCTGTAAATCTTCCCTGAGTCTTCCAAAGCAGTA ATTTGCCITCTCTTAGAGTTTACCTGCCATTTTGTGCACATTTGAGTTACAGTAGCATGTTATTTTACAA TTGTGACTCTCTGGGAGTCTGGGAGCCATATAAAGTGGTCAATAGTGTTTGTGACTGAGAGTTGAATG ACATTTTCTCTCTGTCTTGGTATTACTGTAGATTTTCGATCATTCTTTGGTTACATTTCTGCTATTTTCTG TACCCATGACTTTTATCATTCTTCTCCATGCTTTATCTCCATCAATTTATCTTCTATTTTAAATTTT TCCACCTTTGCTTCTTCTTCTTGTGAGATCTCTCCCTTTACTGACTATAACATAGAAGAATAGAAGTGTAT TTTATGTGCTTTAAGGACAATACCTTTAGATTCTTGTCTTAAGTTTTTAAACTGAATGAATGGAATATTA TTCTCTCCCTAAGCAAAATTCACAAAACAATTTTCTTATGTTTATGTAGCCTTAAATTTGTTTTGT CTGTAAACCTCAGCATAAAAACCTTTCTTCAATTTCTAATTTTCAATTCACAAATATTGATTAATCTGGT ATTAGCACAAGAAAAATGTGCTAATAAGCCTTATGAGAATTTGGAGCTGAAGAAAGACATATAACTCAGG AAAGTTACAGTCCAGTAGTAGGTATAAATTACAGTGCCTGATAAATAGGCATTTTAAATTTGTACACTC AACGTATACTAGGTAGGTGCAAAACATTTACATATAAATTTTACTGATACCCATGCAGCACAAAGGTACTA ACTTTAAATATTAAATAACACCTTTTCTGTGTCAGTAATTCATTTGCATTAAATCTTATTGAAAAGGCTTT CAATATATTTTCCCCACAAATGTCATCCCAAGAAAAAGTATTTTAAACATCTCCCAATATAATAGTTA CAGGAAATCTACCTCTGTGAGAGTGACACCTCTCAGAATGAACCTGTGTGACACAAGAAAAATGAATGTAGG TCTATCCAAAAAACCCTCAAGAAACAAAAACAATATTATTAGCCCTTTTATGCTTAAAGTATGAGTCTGAG GGAACAGTTGATGTTGTGATCATTTTATTATCTGATTCTTGTACTTTGAATTAACCAATATTTTGATG ATATAAATCATTTCCACCAGCATATATTTAATTTCCATAATAACTTTAAATTTTCTAATTTCACTCAAC TATGAGGGGAATAGAAATGTGGTGGCCACAGGTTTGGCTTTTGTAAATGTTTGAATCTTCTGATGTTGAT CTCTGTCTGCAATGTAGATGTCTAAACACTAGGATTTAATATTAAAGGCTAAGCTTTAAAAATAAAGTAC CTTTTTAAAAAGAAATATGGCTTCCCAAAATGGAAATACCTAATTTCTAATCTTTTTCTCAAAAGTC CTATCTACTAATGTCTCCATTACTATTTAGTTCATCATAACCATTATCTTCAATTTTACATGTCGTGTTCTT TCTGGTAGCTCTAAATGACACTAAATCATAAGAAAGACAGGTTACATATCAGGAAATACCTGAAGGTAC TGAATAGATTCTTGAGTTAATGAAATATTTTCTGTAAAAAGGTTTGAAGAGCCATTTGAGTCTTAAAGC ATTATACCTCCATTATCAGTAGTTTATGTGACAAATTTGTGTGTTTAAATGTTTAAAGATGTGGCCTTT TTAATAAGGCAATGCTATGCTATTTTTCCTCATTTAACATTAAGATAATTTATGCTATACAGATGATAT GGAATATGATGAACAATATTTTTTTTGGCCAAACATATGCCTTGTAAAGTAGCCATGGAATGTCAACCTGT AACTTAAATTTATCCACAGATAGTCATGTTTTGATGATGGGCACGTGGAGATACTGACATAGGACTGT GCCCCCTTCTCTGCCATTACTAGCTGGATGAGATTAAGCAAGTCATTTAACTGCTCTGATTTAAACCTG CCTTTCCCAAGTGCTTTGTAAATGAATAGAAATGGAACCAAAAAAACGATACAGGCCTTCAGAAATAG TAATGCTACTATTTTGTCTTTCATTAAAGCCATAGTTCTGGCTATAATTTTATCAAACCTACCAGCTATAT TCTACAGTGAAAGCAGGATTTCTAGAAAGTCTCAGTGTATTTATTTATGTCACCATGTGCTATGATATATTT GGTGAATTCATTTGAAATTAGGGCTGGAAGTATTCAAGTAATTTCTTCTGCTGAAAAAATACAGTGT TGAGTTTAGGGCTGTTTTATCAAAGTTCTAAAGAGCCTATCACICTTCCATTGTAGACATTTTAAATA ATGACACTGATTTTAAATTTTAAAGTGTCTTTTGAACAGAGAGCCTGACTAGAACACAGCCCCCTCCA AAAACCATGCTCAAATTTATTTTACTATGGCAGCAATTCACAAAAGGGAACAATGGGTTTGAATAATTA CAATGAAGTCATCAACCCAAAAACATCCCTATCCCTAAGAAAGTTATGATATAAAATGCCCAACAAGAAA TCTATGTCGTCTTAAATCTGTCTTTTATIGCTTTTGAAGGATGGCTATTACATTTTATGTTTTGCTGTG AATACCTGAGCAGTTTCTCTCATCCATCTTATCTTTCACACATCAGAAGTCAGGATAGAATATGAATCA TTTTAAAACTTTTACAACCTCCAGGCCATGTGCATAAGAAGCATTTCAAACTTGCCAAAACATACATTT TTTTCAAATTTAAAGATACTCTATTTTGTATTTCAATAGCTCAACAACGTGGTCCCCACTGATAAAGT GAAGTGGACAAGGAGACAAGTAATGGCATAAGTTTGTTTTTCCCAAAGTATGCCTGTTCAATAGCCATTG GATGTGGGAAATTTCTACATCTCTTAAATTTTACAGAAAATACATAGCCAGATAGTCTAGCAAAAGTTC ACCAAGTCTTAAATTTGCTTATCTTACTTACTTAAAGTCAATGAAATCATTTTAAATGAAAGAACATCACCT AGGTTTTGTGGTTTTCTTTTCTTCTTATGCTGAGTGAAACCAACAATCTCTGTTTTCTCCCTAGCAT CTGTGGACTATTTAATGTACCATTTATCCACACTCTATGGTCTTACTAAATACAAAATGAACAAAAAG CAGTAAAAACAACCTGACTCTTCACCCATATTTATAAATATAATCCAAGCCAGATTAGTCAACATCCATAAG ATGAATCCAAGCTGAACCTGGGCTAGATTATTGAGTTTCAAGTTGGATCACATCCCTATTTATTAATAAAC TTAGGAAAGAAGGCTTACAGACCATCAGTTAGCTGGAGCTAATAGAACCCTACACTCTAAAGTTCCGGCC TAGAATCAATGTGGCTTTAAAGCTGAAAAGAGCAGGAAAGAACAGTTTTCTTCAATAATTTGTCCACC CTGTCACTGGAGAAAATTTAAGAAATTTGGGGGTGTGGTAGTAAGTTAAACACAGCAGCTGTTTATGGCA GAAATTTATCAATACATACCTTCTCTGAATATCTTATAACCAAGCAAGAAAAACCAAGGGGTTTGT	
--	--

TCTCCTCCTTGGAGTTGACCTCATTCCAAGGCAGAGCTCAGGTCACAGGCACAGGGGCTGCGCCCAAGCT TGTCGCGCAGCCCTTATGCAGCTGTGGAGTCTGGAAGACTGTTCAGGACTGCTGGCCTAGTCCCAGAAATGT CAGCCTCATTTTTCGATTTACTGGCTCTTGTGTCTGTATGTCATGCTGACCTTATTTGTTAAACACAGGTTT GTTTGTCTTTTTCCTCATGGAGACATGGGAGAGGCATTATTTTAAAGCTGGTTGAAAGCTTTAACCG ATAAAGCATTTTATAGAGAAATGTGAATCAGGCAGCTAAGAAAGCATACTCTGTCCATTACGGTAAAGAAA ATGCACAGATTATTAACCTCTGCAGTGTGGCATTAGTGTCTGGTCAATATTCCGATAGATATGAATAAAA TATTTAAATGGTATTGTAATAGTTTTCAGGACATATGCTATAGCTTATTTTATATCTTTTGAAATTG CTCTTAATACATCAAACTCTGATGTATTTCAATTTATCAGATATAAATTAATCTAAATGAAGCCAGTTAA ATGTTTTTGTCTTGTGAGTTATATGTTAAGTTTCTGATCTCTTTGTCTATGACGTTTATGATTAGGATGCTATT TTTACTGTTATGAATTATTTAGACAGCAGTGGTTTCAAGCTTTTTGCCACTAAAAATACCTTTTATTTT CTCCTCCCCCAGAAAAGTCTATACCTTGAAGTATCTATCCACCAAACTGTACTTCTATTAAGAAATAGTT ATTGTGTTTTCTTAATGTTTTGTTATTCAAAGACATATCAATGAAAGCTGCTGAGCAGCATGAATAACAA TTATATCCACACAGATTGATATATTTTGTGAGCCTTAACTTGATAGTATAAAATGTCTATTGCTTTTAA AATAATAGTTAGTCAATGGACTTCTATCATAGCTTTCTTAACTAGGTTAAGATCCAGAGCTTTGGGGTC ATAATATATTACATACAATTAAAGTTATCTTTTCTAAGGGCTTTAAATTCATGAGAATAACCAAAAAAG GTATGTGGAGAGTTAATAACAACATACCATATTTCTGTTGAAACAGAGAIGTGGCTCTGCTTGTCTTCCA TAAGGTAGAAATCTTTCCAGAAATTTGCTTAACTAGTAAGCCCTGAATTTGCTATGATTAGGATAGGA AGAGATTTTACATGGCAGACTTTAGAATTTCTTCACTTTAGCCAGTAAAGTATCTCTTTTGATCTTAGT ATTCTGTGTATTTTAACTTTTCTGAGTTGTGCATGTTTATAAGAAAAATCAGCACAAAGGGTTTAAAGTTA AAGCCTTTTTTACTGAAATTTGAAAGAAACAGAAGAAATATCAAAGTTCTTTGTATTTTGAGAGGATTAA ATATGATTTACAAAAGTTTACATGGAGGGCTCTCTAAAACATTAAATTAATTTATTTTGTGAAAAGCTCT TACTTTAGGCATCATTTTATTTCCICAGCAACTAGCTGTGAAGCCITTACIGTGCTGTATGCCAGTCACTC TGCTAGATTGTGGAGATTACCAGTGTTCCTGCTCTCTCCGAGCTTAGAGTTGGATGGGGAATAAAGACAG GTAACACAGATAGCTACAATATTGTACTGTGAATGCTTATGCTGGAGGAAGTACAGGGAACATTGGAGCA CCTAAGAGGAGCACCTACCTTTGAATTTAGGGTTAGCAGAGGCATCCTGAAAAAAGTCAAAGCTTAAGCCCA CAATCTATAAGCAGTTTAGGAATTAGCAGAACGTGCGTGGTGAGGAGATGCCAAAGGCAAGAGAGAAGA GTATTTCCAAACAGGAGGGATTCCAAAGAGAGAGAGTATCCCAACAACATTTGCCACAAACCTGATGGGG AGAGAGAAATGTGGGGTGGGGATGGATGATGAGACTGAAGAAGAAAGCCAGGTCTAGATAATCAGTGGCCT TGATACCATGTTAAAGAGTGTAGACTTGATTCTGTTGTAACAGGAAAGCAGCACAAATTCATATGAATA TTTTAGAAGACTCCCACTGGAATATGGAGAATAAAGTTGGAGATGACTAATCCTGGAAGCAGGGAGAACA TTTTTGAGGAAGTTGCATATTTTGGTGAAAATGATGATCATAAACATGAAGAATTGTAGGTGATCATGA CCCTCTCTTAATTTTCCAGAAGGGTTTGGGAAGATATAACATAGGAACATTGACAGGACTGACGAAAGG AGATGAATACACCATATAAATTTGTCAAACACAAGGCCAGATGCTAATTTATTTGCTTATGTGTTGAAA TTACAAATTTTTCATCAGGAAACCAAAACTACAAAATTTAGTTTTCCCAAGTCCCAGAATTTCTATCTGT CCAAACAATCTGTACCCTCCACCTATATCCCTACCTTTGCTATGCTGTCCAACTCCAAAGTCCAGGTCT ATACACACGGGTAAAGACTAGAGCAGTTCAAGTTTTCAGAAAATGAGAAAGAGGAACCTGAGTTGTGCTGAAC CCATACAAAATAAACACATTTCTTTGTATAGATTCTTGGAACTTCGAGAGGAATTCACCTTCTATAGGT ATTTGATGGTATGAATCCATGGCTGGGCTCGGCTTTTAAAAAGCCTTATCTGGGATTCTTCTATGGAAC CAAGTTCCATCAAAGCCCATTTAAAGCCCTACATTAACCAACAAATTTCTTGTGCTGCTATGATACAAATAA TGATGTCTATGATCAATAATCAGATGCCATTATCAAGTGAATTTACAAAATGGTATACCCACTCCAAAAA AAAAAAAAGCTAAATTTCTCAGTAGAAGATTGTGACTTATGAGCCCTCCACAGCCTTCTGAGGATGAGGA GGGAGCACTGGTGAGCAGTAGGTTGAAGAGAAAATTTGGCGCTTAATAATCTATCCATGTTTTTTCATCT AAAAGAGCCTTCTTTTGGATTACCTTATCAATTTCCATCAAGGAAATTTGTTAGTTCCACTAACCCAGAC AGCAGCTGGGAAGGCAGAAAGCTTACTGTATGTACATGGTAGCTGIGGGAAGGAGGTTTCTTTCTCCAGGT CCTCAGCTGGCCATACCCAGTCCCTTGTAGTTATGCTTGTGCTGCTAGACCCCTGCTATCTATCTCATA TTTTAAGTCTTTGGCTTGTGAATTTATCTATTTCTTTAGCTTTCAGCACTGCAGAGTGTCTGGGACTTTTGCTA ACTTCCATTTCTTGTGCTTGTAGCATTCTCTATAGGCCAGCTCTTTTCTCATCTGGCCCTGCTGTGG AGTCACCTTTGCCCCCTCAGGAGAGCCATGGCTTACCACTGCCTGCTAAGCCTCCACTCAGCTGCCACCAC ACTAAATCCAAGCTTCTCTAAGATGTGTCAGACTTTACAGGCAAGCATAAAAGGCTTGTATTTCTGTGGAC TTCCCTTTTACTTGTCTGAATCTCACCTCCTTCAACTTTTCACTCTCAGAATGTAGGCAATTTGCTCTTTG CCCTACATCTTCTTCTTGAATCATGAAAGCCTCTCACTTCTCTTGTCTATGCTGAGGCTTCTGT CAGGTTTTTGAATGAGTTCTCATCTAGTCTTAGTAGCTTTTGTATGCTTAAAGTCCACCTTTTAAAGGATACC TTTGAGATTTAGACCATGTTTTTCTGCTTGAAGAAAGCCCTAATCTCCAGACTTGCCCTTCTGTGATTCTCA AAGACCAACTGAGGAAGTCAAAAGCTGAATGTTGACTTTCTTTGAACATTTCCGCTATAACAATTTCAAT TCTCCTCAGAGCAATATGCTGCTCCCACTGACCAGGAGAAAGGTCCAGTGCCAAAGAGAAAAACACAA AGATTAATTTATTTAGTTGAGCACATACCTTCAAGTGGTTTGGGTATTCATATGAGGTTTTCTGTCAAG AGGGTGAGACTCTTCACTATCTGCTGCTGCTGACAGTTCTCCTGGCACTGGCTGCTAACAGATGCAAAA ACTGTAAAAATTAAGTGATCATGTATTTAACGATATCATCACATACTTATTTTCTATGTAATGTTTTAA ATTTCCCCTAACATACTTTGACTGTTTTCGACATGGTAGATATTCACATTTTTTTGTGTTGAAGTTGATG CAATCTTCAAAGTTATCTACCCCTGCTTATTTAGTAAACTAGTGTAAATCTTGGCAAGAGATGACAGG GAATCTTTCTCATGACTCAGCCCTATTTAGTTATTAATGCTACTACCCCTATTTTGTGTTGAAGTTAGGT CCCTAAGTACATTGTCCAGAGTTATCTTTTAAAGATATTTAGCCCCATATCTTCTGAACTCAAAAGTC ATACACCTTGCTCCTCATTTCTGAGTGGGAAAGACATTTGAGAGTATGTTGACAATTTGTTCTGAAGGTTT TTGCCAAGAAGGTGAAACTGTCCTTTCACTGTGTATGCTGGGGCTGGGTCCCTGGCAGTATGGGGTG ACAAATGCAAAGCTGAAAACTAGGTGCTAGTGGGCACCTAATATCATCATCATATATCTTTTCAAGC TAATATGCAAAATCCCATCTCTGTTTTTAACTAAGTGTAGATTTTCAGAGAAAATATTTTGTGGTTTACA TAAGAAAACAGTCTACTCAGCTTGACAAGTGTTTTATGTTAAATTTGGCTGGTGGTTTGAATGAATCATC TTCACATAATGTTTTCTTTAAAAATATTGTGAATTTAACTCTAATTTCTTGTATTCTGTGTATAATAAA	
---	--

	GAATAAACTAATTTCTA	
AK093306	ATTCTATGCTGCAGCCTAAGCATCATTCCTCTTCTCTTCTTAGTGGAGATAAAAATTACCCACTGCTCTCC TTACATTTTACTTTTGTCCATATTTGCTCCTATGCTCTAGGCTCGTGCAACAACACAGTGTGGGCCCTT ACCCTAGAAGCCAACTTCTCATGACCTTCTCTATCTCCAGAATCCATGCAGTGGGAATGAAGGTAAAAG AAGGTTTTTCATGGGATCCAGCTGAGAGCTCTACGGGGAATAATGGATCTGAGGAGCCATGTGCTCCATCTC TTTTATTTTACAGGTAGAGACTAGGGGTATAGAGTGAGGTGAATTACCGCAGTGACCCACACATTGTGG CAGACCTAGGATTAGAACTCTGTCTTCCGTGGTCCAGCTTGGTGCTTTTGAAAGCATACTTGCTGCTTT CTTACCGGCTGGTGTCTGCCACTTTGGGACAGAGTGTGGACTTGTCTACCTGCCCATTTCTTAGGGAT TCTCATTTCTGTGTTTGAGCAAGAATATTTCTTATTCTGGAAGAACCACATACCACAGATTCTGGGTGAG CATAAGGAAGATTGTCTTGGGAICTGACTTAGCTCACGTATAGTGGCTATGATGAATTCAGTGTCTTAT TTTTGTGATATGATATTTTATGCTAATATTGCCGTGGGTGTCTGAGCAAGTCTAGATGAATTTAATTGC TCTCATTTTCCCTGCCCTCTTCTTTGGTCTCTCTTTTAGGAAATGTTTTCTTTCAACATTTCGTTT CATTCATTATTTACTCATTCGGCCAACCAACATTTATTAGTGGCTTCCCTGTATCAGGGACAGGGCTT ACAAAGTAGAATTTGATCCCACTCTGCCCTCAGTAGCTCAGTGICTAATGGAGGTAGTGATGTTTATT AGCGTCGCCAGATACTGTGCTAGGTGCTGTGCCTGTTCTCTCTCGCTTGTCTCTACACACTTGAGAAGG CCGAAGCTGATTCATAGCTTGAAGGCAGGGGCCCTGGATTTGAACCCAGGCCTGACCAATGGCAGAACC TATCAGATGTGTGGACAGATGACATTGCCCTTCTTTCTTTGGATATATCAAAATCAGGCAGCGGAGGA ACTCCCATTTTGAAGCAATGTGCAGGAATGATAGGGTATACAGAGAGGAACAGGAGATGGCCCCCTGA CTTCCAGCATGTGTCTGATGGACATCCAGGCTGCAGGCATCATGTTGCTGTCTAGAGAGATGAGCCAGGT GCCCAGAGCCCCATGGGCCAATGCTGCCCTTCTTGGCATGCCAAACAAAGCGGTTGGTGTGTAGAGGC ACAGTCTCCTCCACTCTAAGTAAAAATCAGCATGAGTCTTAGCCACATTTCCCTAGTGAGTACCAAAA GATATCTATGAACGGCAGTCTCAGTGACTTCTTAAGGTTCCGGAAATGCATCTCTTACTCAGGAGTAA GCAATGATGTGCTGCGGCTTTACGAGTTCTCAGAGAATGACTTCTGGACCCAAATGTTTTTCTGCTT CAGGACTGTGAAGGCTTATGTTCTGCTCTGCCACCAAGGTGACCGCTGATGTCTCAACGCAGCTGAGA AACTCCAGGTGGTGGGACAGGGCTGGCAGAGGTGTGGACAATGTGGATCTGGAGGCCGCAACAGGAGGG CATCTTGGTTATGAACACCCCCAATGGGAACAGCCTCAGTGCCGAGAACTCACTTGTGGAATGATCATG TGCCCTGGCCAGGCAGATTCCCCAGGCGACGGCTTCGATGAAGGACGGCAATGGGAGCGGAAGATTCA TGGGAACAGAGCTGAATGGAAAGACCCCTGGGAATTTCTTGGCTGGGCAGGATTGGGAGAGAGGTAGCTAC CCGCATGCAATCCTTTGGGATGAAGACTATAGGGTATGACCCCATCATTTCCCCAGAGTCTCGGCCCTC TTTGGTGTTCAGCAGCTGCCCTGGAGGAGATCTGGCTCTCTGIGATTTTCACTGTGACACCTCCTC TCCTGCCCTCCACGACAGCTTGTGTAATGACAACACCTTTGCCAGTGCAAGAAAGGGGTGCGTGTGGT GAACTGTGCCCTGGAGGGATCTGIGGACGAAGGCGCCCTGCTCCGGGCCCTGCAGTCTGGCCAGTGTGCC GGGCTGCACTGGACGTGTTACGGAAGAGCCGCCACGGGACCGGGCTTGGTGGACCATGAGAATGTCTA TCAGCTGTCCCCACCTGGGTGCCAGCACCAGGAGGCTCAGAGCCGCTGIGGGGAGGAAATGTCTGTCTA GTTCGTGGACATGTTGAAGGGGAAATCTCTCAGGGGGTGTGTAATGCCAGGCCCTTACAGTGCCTTC TCTCCACACCAAGCCTTGGATTGGTCTGGCAGAAAGCTCTGGGGACACTGATGCCAGCCTGGGCTGGGT CCCCCAAAGGACCATCCAGGTGATAACACAGGGGAACATCCCTGAAGAATGCTGGGAATCTCGGCCCTC CGCAGTCAATGTGCGCCTCCTGAAAGAGGCTTCCAAGCAGGCGGATGTGAATTTGGTGAACGCTAAGCTG CTGGTGAAGAGGCTGGCCTCAATGTCAACACCTCCACAGCCCTGCTGCACAGGGGGGCAAGGCTTCG GGGAAATGCTCCTGGCCGTGGCCCTGGCAGGCGCCCTTACCAGGCTGTGGGCTTGGTCCAAGGCACTAC ACCTGTACTGCAGGGGCTCAATGGAGCTGTCTTACGGCCAGAAGTGCTCTCCGACAGGACCTGCCCTG CTCCTATTCCGACTCAGACCTCTGACCTTGAATGCTGCTTACCATGATTGGCTCCTGGCAGAGGCAG GCGTGCAGGCTGCTGCTCCTACAGACTTCACTGGTGTGAGATGGGAGACCTGGCAGCTCATGGGCATCTC CTCCTTGCTGCCAGCCTGGAAGCGTGAAGCAGCATGTGACTGAAGCCTTCCAGTTCACCTTCTAACCT TGGAGCTCACTGGTCCCTGCTGCGGCTTTTCTGAAGAAACCCACCCACTGTGATCAATAGGGAGAGA AAATCCACATTCTTGGGCTGAACGCGAGCCTCTGACACTGCTTACACTGCACCTCTGACCTGTAGTACAG CAATAACCGTCTAATAAAGAGCCTACCCCC	132
BE904476	CAAACAAAAACAGCCAAGCTTTTCTGCCAAAAAGATGACTGAGAAGACTGTTAAAGCAAAAAGCTCTGTT CCTGCCCTCAGATGATGCCATCCAGAAAATAGAAAAATTTCTTCCCTTCAATCCTCTAGACTTTGAGAGTT TTGACCTGCTGAAGAGCACCAGATTGCGCACCTCCCCTTGAAGTGGAGTGCCTCTCATGATCCTTGACGA GGAGAGAGAGCTTTGAAAAGCTGTTTCAGCTGGGGCCCCCTTACCTGTGAAGATGCCCTCTCCACCATGG GAATCCAATCTGTTGCAGTCTCCTTCAAGCATTCTGTGACCCCTGGATGTTGAATTTGCCACCTGTTTGT GTGACATAGATATTTAAATTTCTTAGTCTTCTAGAGTCTGTGTGATTTTGTATTAATAAAGCATTTCTTA ACAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA GAATTTCCCAAGAGGGGGGCAACAGATAATCAGAGGATATCACACAAGATCTCTCGGCGCACCAACGACG GGGGCCCAATAAGGGAGAGACCCAGAATCACACAGCCAGACACGGTGGACACGACGAAACAAACA CACAGCCAGACACGGGGGCAACACGCGCGCACACCGCGGACACCATGGGACAAAGCAGACACACCCCA CAAAACAACACCGCGGAGGGGGAAGAACAACAACAAGTGCAGAAACAGAACACAACCACAGAAAGAGA AAAATTAACAGGCCCCCAAGACGGCGACAACAACAACAACAACAACCACTACAGAGCGCTCAACAGCCGAG TAAAAACAACAACCGGACAACCTAACACACAAGGAATGAACAAAGCGGGGCCACACACCGGACACCGGA AATCCGGCAACAACCTCACACCGAGCGAGGGTCCAGACACAACAATACAGACACAAAGCAAAACGAGAAAC AAGACCAGCAAGACGAGCAGGCAAAAGACAACAAGACAGAGGAGACGACGACGAAAGGACAAAGA GGACACAACGACGCGAGGAGCGAGAGCGAGAGGAAGAGACAACAAGAACACAAAGAACAAAGCAAGCA GCAGCGAAGAACGACACACAACACACGAGACAGCAGGAGCAGAGCGGAGAAAAACAACGAGCAGAGCC AAGACCAAGAGAGGAGAACAAAAAATAACGAGAGCAGGCGGACGAGAGCAGAGACGAGACGAACAGACA AACGGGAATCAGAAGCATAACGATCCGCGACGCGAACAACN	133
AK123010	GTGCACCCTGTCCAGCCGTCTGTCTGCTGCTGCTCTGCTTCTGCTGCGCTCCACTATGCTCTCCC TCCGTGTCCTGCTGCGGCCATCAGGACCCGAGCAGCTGCAGCTCTCGCCGTGAAGGGGCTCAGCTT	134

	GGTCGACAAGGAGAACACGCCGCCGCCCTTGAGCGGGACCCGCGTCTTGGCCAGCAAGACCCGCGAGGAGG ATCTTCCAGGAGAAAACCCCGCGCCTTTTGTCATCTTCCCATCGAGTACCATGATATCTGGCAGATGTA TAAGAAGGCAGAGGCTTCCCTTTTGACCCGCGAGGAGGTGGACCTCTCCAGGACATTACGACTGGGAA TCCTGAAAACCCGAGGAGAGATATTTATATCCCATGTTCTGGCTTTCTTTGCAGCAAGCGATGGCATAG TAAATGAAAACCTTGGTGGAGCGATTAGCCAAAGAAAGTTCAGATTACAGAAGCCCGCTGTTCTATGGCTT CCAAATTGCCATGAAAACATACATTCTGAAATGTATAGTCTTCTTATTGACACTTACATAAAAGATCCC AAAGAAAGGGAATTTCTCTTCAATGCCATTGAAACGATGCCTTGTGTCAAGAAAGAGGCAGACTGGCCCT TGCGCTGGATTGGGGACAAAGAGGCTACCTATGGTGAACGTGTTGTAGCCTTTGCTGCAGTGGAAAGGCAT TTCTTTTCCGGTTCTTTTGGCTCGATATTCTGGCTCAAGAAACGAGGACTGATGCCTGGCCTCACATTT TCTAATGAACCTTATTAGCAGAGATGAGGGTTTACACTGTGATTTTGTCTGCTGATGTTCAAACACCTGG TACACAAACCATCGGAGGAGAGAGTAAGAGAAATAATTATCAATGCTGTTTCGGATAGAACAGGAGTTCCT CACTGAGGCTTGGCTGTGAAGCTCATTGGGATGAATTGCACCTCTAATGAAGCAATACATTGAGTTTGTG GCAGACAGACTTATGCTGGAACCTGGGTTTATAGCAAGGTTTTAGAGTAGAGAACCATTGACTTTATGAG AGAATATTTCACTGGAAGGAAAGACTAATCTTTTGAAGAAGAGTAGGCGAGTATCAGAGGATGGGAGT GATGTCAAGTCCAACAGAGAATCTTTTACCTTGGATGCTGACTTCTAAATGAACCTGAAGATGTGCCCTT ACTTGGCTGATTTTTTTTTTCCATCTCATAAGAAAAATCAGCTGAAGTGTACCAACTAGCCACACCAT GAATGTCCGTAATGTTTCAATTAACAGCATCTTTAAACTGTGTAGCTACCTCACACAGCTGCTGTCTGT TTATAGTGTGGTAGTATCACCTTTTGGCAGAAGGCTGGCTGGCTGTGACTTACCATAGCAGTGACAAAT GGCAGTCTTGGCTTTAAAGTGAGGGGTGACCTTTAGTGAGCTTAGCACAGCGGGATTAACAGTCCCTTT AACCAGCACAGCCAGTTAAAGATGCAGCCTCACTGCTTCAACGCAGATTTTAAATGTTTACTTAAATATA AACCTGGCACTTTACAAACAAATAAACATTTGTTTGTACTCACAGGCGATAATAGCTTATTTTGGT TTCTACACCAATACATTTCTCTGACCCTAATGGGAGCCAAATTCACAATTCATAAGTGACTAAAGTAA GTTAAACTTGTGTAGACTAAGCATGTAATTTTAAAGTTTATTTTAAATGAATTAATAATTTGTTAAACCA ACITTAAGTCACTCCTGTGTATACCTAGATATTAGTCAAGTGGTGGCAGATAGAAGCAGGTTGTGTTT TTATCTGTGGCTTGTGTAGTGTCTGGGATTTCTGTGCCCTCTGAGTAGAGTGTGTGGGATAGGGA ATCTCTCAGGGCAAGGAGCTTCTTAAGTTAAATCACTAGAAATTTAGGGGTGATCTGGGCTTCATATGT GTGAGAAGCCGTTTCATTTTATTTCTCACTGTATTTCTCAACGCTCTGGTGTATGAGAAAAAATCTTG AAGAGTTTTATATGTGGGAGCTAAGGTAGTATTGTAAATTTCAAGTCACTCTTAAACAAAATGATCCA CCTAAGATCTTGGCCCTGTTAAGTGGTGAATCAACTAGAGGTGGTTCTTACAAGTGTGTTCTTCTAGTT TTGTTTGGTGAAGTAGGTTGTGTGAGTTAATTCATTTATTTTACTATGTCTGTAAATCAGAAATTTT TTATATCTATGTTCTTCTAGATTTTACCTGTAGTTTCACTTCACTTCACTCAGCCAGTGTCTTATCTGGCAT TGICTAAATCTGAGCATTGTCTAGGGGATCTTAACTTTTAGTAGGAAACCATGAGCTGTTAATACAGTT TCCATTCAAATATTAATTTTCAAGATGAACATAAATTTTTTTTTTTTTTTTGGAGTGGAGTCTCGCTCT GTTGCCAGGCTGGAGTGCAGTGGCGCGATTTTGGCTCACTGTAACTCCATCTCTGGGTTCAAGCAAT TCTCTGTCTCAGCTCCCTAGTAGCTGGGACTGCAGGTATGTGTACCACACCTGGCTAATTTTTGTAT TTTTAGTAGAGATGGAGTTTACCATATTGGTCAAGGCTGGTCTTGAACCTCTGACCTCAGGTGATCCACC CACCTCGGCTCCCAAAGCTTGGGATTTGCAGGCGTGATAAACAAATATTCTTAAATAGGCTTCTGTGAA TTAATCTGCCCTTTATGTTTGGGAGAAGAAAGCTGAGACATTGCATGAAAGATGATGAGAGATAAATGTTG ATCTTTTGGCCCCATTTGTAAATGTATTGATTTTGAACGTCGCTCTGTTTATGTTAGTTTCTTCA TCATTTATTTGTATAGACAATTTTTAAATCTCTGTAAATATGATACATTTTCTATCTTTTAAAGTTATTGTT ACCTAAAGTTAAATCCAGATTATATGGTCTTATATGTGTACAACATTAATAATGAAGGCTTGTCTGTGCA TTGTGAGGTACAGGCGGAAGTTGGAATCAGGTTTTAGGATTCTGCTCTCATTAGCTGAATAATGTGAGG ATTAACCTTCTGCCAGCTCAGACCATTTCTTAATCAGTTGAAAGGGAAACAGTATTTCACTCICAAAAT GAATAATGCACAAGTCTTAAAGTATTAATAAACTGTTCTTATGTCAGTTT	
BC036503	AGCGGGGCACTCCAGCCCTGCAGCCTCCGGAGTCACTGCCGCGCGCCCGCGCCCTTCCCTGCT CGCCGCACCTCCGGGAGCCGGGGCGCACCCAGCCCGCAGCGCCGCTTCCCGCCCGCGCCGCTCCGACC GCAGGCCGAGGGCCGCACTTGGCCGGGGGACCGGGCAGCAGCTTGCAGCCGCGGAGCCGGGCAACGCTG GGGACTGCGCCTTTTGTCCCGGAGGTCCCTGGAAGTTTTCGGGCAGGACGCGCGGGGAGGCGCGGAG GCAGCCCGACGTCGCGGAGAACAGGGCGCAGAGCCGGCATGGGCATCGGGCGCAGCGAGGGGGCGCC GCGGGGCAGCCCTGGCGTGTCTGTGGCGCTGGGCGCGCGCTTCTGGCCGTGGGCTCGGCCAGCGAGTA CGACTACGTGAGCTTCCAGTCCGACATCGGCCCGTACCAGAGCGGGCGCTTCTACACCAAGCCACCTCAG TGCTGGAGCACGAGACCATGGCGAGGTGAAGCAGCAGGCCAGCAGCTGGGTGCCCTGCTCAACAGAA CTGCCACGCGCGCACCCAGGTCTTCTCTGCTCGCTCTTTCGCGCCCGTCTGCTGGACCGGCCCCATCTAC CCGTGTGCTGGCTCTGCGAGGCGGTGCGCGACTCGTGCGAGCCGGTCTATGAGTTCTTTCGCTTCTACT GGCCCGAGATGCTTAAAGTGTGACAAGTTCCCGGAGGGGACGCTGTGATCGCCATGACGCCGCCCAATGC CACGAAGCCTCCAAAGCCCCAAGGCACAACGGTGTGTCTTCTGTCGACACAGAGTTGAAATCTGAGGCC ATCATTGAACATCTCTGTGCCAGCGAGTTTGCAGTGGATGAAAAATAAAGAAGTGAAGAAAGAAAAATG GCCACAAGAAGATTGTCCCCAAGAAGAAGAGCCCTGAAGTTGGGGCCCATCAAGAAGAAGGACCTGAA GAAGCTTGTGCTGTACCTGAAGAAATGGGGCTGACTGTCTCTGCCACAGCTGGACAACCTCAGCCACCAC TTCTCATCATGGCCCGCAGGTGAAGAGCCAGTACTTGTGACGGCCATCCACAAGTGGGACAAAGAAA ACAAGGAGTTCAAAAACCTCATGAAGAAAAATGAAAAACCATGAGTGGCCACCTTTCAGTCCGTGTTTAA GTGATTTCTCCCGGGGCGAGGTTGGGGAGGAGCCTCGGGTGGGGTGGGAGCGGGGGGACAGTGCCTCCG GAACCCGGTGGGTACACACACGCACTGCGCCTGTCTAGTGTGACATTTAATCCAGTCCGCTTGTCTT GCAGCATTCCTGCTCCCTTCCCTCCATAGCCACGCTCCAAACCCAGGGTAGCCATGGCCGGGTAAAGCA AGGGCCATTTAGATTAGGAAGGTTTTTAAAGATCCGCAATGTGGAGCAGCAGCCACTGCACAGGAGGAGGT GACAAACCATTTCCAACAGCAACACAGCCACTAAAACACAAAAAGGGGGATTTGGGCGGAAAGTGAGAGCC AGCAGCAAAAACATACATTTTGAACCTTGTGTGGTGTGATCTATTGGCTGATCTATGCCCTTCAACTAGAA	135

	AATTCTAATGATTGGCAAGTCACGTTGTTTTTCAGGTCCAGAGTAGTTTCTTTCTGTCTGCTTTAAATGGA AACAGACTCATACCACACTTACAATTAAGGTCAAGCCCAGAAAGTGATAAGTGCAGGGAGGAAAAAGTGCA AGTCCATTATGTAATAGTGACAGCAAAGGGACAGGGGAGAGGCATTGCCTTCTCTGCCCCACAGTCTTTC CGTGTGATTGTCTTTGAATCTGAATCAGCCAGTCTCAGATGCCCAAAGTTTCGGTTCCTATGAGCCCGG GGCATGATCTGATCCCCAAGACATGTGGAGGGGCAGCCTGTGCCTGCCTTTGTGTCAAAAAAGGAAACC ACAGTGAGCCTGAGAGAGACGGCGATTTTCGGGCTGAGAAGGCAGTAGTTTTCAAACACATAGTTAAAA AAGAAACAAATGAAAAAATTTTGAACAGTCCAGCAAATTGCTAGTCAGGGTGAATTTGTGAAATTTGGGT GAAGAGCTTACGATTCTAATCTCATGTGTTTTTCCTTTTTCACATTTTAAAAAGAACATGACAAACACCCA CTTATTTTTCAAGGTTTTTAAACAGTCTACATTGAGCATTGAAAGGTGTGCTAGAACAAAGTCTCCTGA TCCGTCCGAGGCTGCTTCCCAGAGGAGCAGCTCTCCCCAGGCATTTGCCAAGGGAGGCGGATTTCCCTGG TAGTGTAGCTGTGTGGCTTTCCCTTCCCTGAAGAGTCCGTGGTTGCCCTAGAACCTAACACCCCTTAGCAAA ACICACAGAGCTTTCCGTTTTTCTTCTGTAAAGAAACATTTCCTTTGAACCTTGATTGCCATATGGAT CAAGAAATTCAGAACAGCCTGCCTGCCCCCGCACTTTTACATATATTGTTTCACTTCTGCAGATG GAAAGTTGACATGGGTGGGGTGTCCCAATCCAGCGAGAGAGTTTAAAAAGCAAACATCTCTGCAGTTTT TCCCAAGTGCCCTGAGATACTTCCCAAAGCCCTTATGTTTAAATCAGCGATGATATATAAGCCAGTTCACTT AGACAACCTTACCCTTCTTGTCCAATGTACAGGAAGTAGTTCTAAAAAATGCAATTAATTTCTTCCC CCAAAGCCGATTTCTAATCTCTGCAACACTTTGAGGACATTTATGATTGTCCCTCTGGGCTTCCGCTT ATACCCAGTGAGGATGCTGCAGTGAGGCTGTAAAGTGGCCCCCTGCGGCCCTAGCCTGACCCGGAGGAAA GGATGGTAGATTCTGTAACTCTTGAAGACTCCAGTATGAAATCAGCATGCCCGCTAGTTACCTACCG GAGAGTTATCCGATAAATTAACCTCTCACAGTTAGTGATCCTGICCTTTTAAACACCTTTTTTGTGGGGT TCTCTGACCTTTTCATCGTAAAGTGTGGGACCTTAAAGTGAATTTGCCGTGAATTTGGATTTAAAA AATGTGTATATATATTAGCTAATTAGAAATATTCTACTTCTCTGTGTCAAACCTGAAATTCAGAGCAAGT TCCTGAGTGGTGGATCTGGGTCTTAGTTCTGGTTGATTCACTCAAGAGTTCACTGCTCATACGTATCTGT CTCATTTTGACAAAGTGCTCATGCAACCGGGCCCTCTCTGCGGCAGAGTCCCTAGTGGAGGGGTTTA CCTGGAACATTAGTAGTTTACCACAGAAATCGGAAGAGCAGGTGACTGTGCTGTGCAAGTCTCTAAATGGG AATTCTCAGGTAGGAAGCAACAGCTTCAGAAAGAGCTCAAATAAATTTGAAATGTGAATCGCAGCTGTG GGTTTTTACCACCGTCTGTCTCAGAGTCCCAGGACCTTGAGTGTCAATAGTTACTTTATTGAGGGTTTAG ACCCATAGCAGCTTTGTCTCTGTACATCAGCAATTCAGAACCAAAAGGGAGGCTCTCTGTAGGCACAG AGCTGCACTATCAGAGGCTTTGTTTTCTCCACAAAGTATCTAACAAAACCAATGTGCAGCATGATTGG CCGTGGTCAATGGTCTCCGAGAGAGGAGGTTGCCTGTGATTTCTAATTATCGCTAGGGCCAAGGTGGGA TTTGTAAGCTTTTACAATAATCATCTTGGATAGAGTCCCTGGGAGGTCCCTTGGCAGAACTCAGTTAAATCT TTGAAGAAATTTGTAGTTATCTTAGAAGATAGCATGGGAGGTGAGGATTCAAAAACATTTTATTTTTTA AAATATCCTGTGTAAACACTTGGCTCTTGGTACCTGTGGGTAGCATCAAGTTCTCCCCAGGTAGAAATTC AATCAGAGCTCCAGTTTGCAATTTGGATGTGTAATTACAGTAATCCCATTTCCCAAACCTAAAATCTGTT TTTCTCATCAGACTCTGAGTAACCTGGTGTGTGTATCACTTCAATAGATGACAGGAGGCTCAGGTGATCT GTTTGAGCAGAGCACCCCTAGGCAGCCTGCAGGGAATAACATACTGGCCGTTCTGACCTGTTGCCAGCAGA TACACAGGACATGGATGAAATTCCTGTTTCTCTAGTTTCTTCTGTAGTACTCCTCTTTTAGATCCTAA GTCTCTTACAAAAGCTTTGAATACGTGAAAATGTTTTACATTCCATTTTCATTTGTGTGTTTTTTTAAAC TGCAATTTTACCAGATGTTTTGTATGTTATCGCTTATGTTAATAGTAATTCCTGACGTGTTCAITTTATTT TCATGCTTTTTTCAGCCATGTATCAATATTCACCTTGACTAAAAATCACTCAATTAATCAAAAAAAAAAAAA AA	
NM_012319	AGTCTTGGGCGAAGGGGGCGGTGGTTTCCCGCGGGCGCTGCGCGCGGCGGTAAATTAGTGATTGCTTCCAG CTTCGCGAAGGCTAGGGGGCGGGCTGCCGGGTGGGTGCGCGGGCGCTGCCCGCGGACCGAGGGGCAGCCAA CCCAATGAAACCACCGCGTGTTCGGCCCTGGTAGAGATTCTCGAAGACACCAGTGGGCGCGTTCGGAGC CCTCTGACCGCCCGTGTGGAACCAAACTGCGCGCGTGGCGGGCGCTGGGACAAACAGGCGCGGAGAGA CGAAGGCGCAATGGCGAGGAAGTTATCTGTAATCTTGATCCTGACCTTTGCCCTCTCTGTACAAAATCCC CTTCATGAACTAAAGCAGCTGCTTTCCCCAGACCCTGAGAAAATTAGTCCGAATTTGGGAATCTGGCA TTAATGTTGACTTTGGCAATTTCCACACGGCAATATCATCTACAACAGCTTTTCTACCGCTATGGAGAAAA TAATTTCTTTGTGAGTTGAAGGGTTTCAAAAATTACTTCAAAAATATAGGCATAGATAAGATTAAGAAGATC CATATACACCATGACCACGACCATCACTCAGACCACGAGCATCACTCAGACCATGAGCGTCACTCAGACC ATGAGCATCACTCAGACCACGAGCATCACTCTGACCATGATCATCACTCTCACCATAATCATGCTGCTTC TGGTAAAAATAAGCGAAAAGCTCTTTGCCAGACCATGACTCAGATAGTTCAGGTAAAGATCCTAGAAAA AGCCAGGGGAAAGGAGCTCACCAGCAGAACATGCCAGTGGTAGAAGGAATGTCAAGGACAGTGTGTAGTG CTAGTGAAGTGACCTCACTGTGTACAACACTGTCTCTGAAGGAACCTCACTTTCTAGAGACAATAGAGAC TCCAAGACCTGGAAGAACTCTTCCCCAAAGATGTAAGCAGCTCCACTCCACCCAGTGTCAATCAAGAGC CGGGTGAGCCGGCTGGTGGTAGGAAAACAAATGAATCTGTGAGTGAGCCCCGAAAAGGCTTTATGTATT CCAGAAAACAAATGAAAATCCTCAGGAGTGTTCATATGCATCAAAGCTACTGACATCTCATGCTAGGG CATCCAGGTTCCGCTGAATGCAACAGAGTTCAACTATCTCTGTCCAGCCATCATCAACCAAAATTGATGCT AGATCTTGTCTGATTTCATACAAGTGAAAAGAGGCTGAAATCCCTCCAAAGACCTATTTCATTACAAATAG CCTGGGTTGGTGGTTTTATAGCCATTTCCATCATCAGTTTCTGTCTCTGCTGGGGTTATCTTAGTGCC TCTCATGAATCGGGTGTTTTTCAAAATTTCTCTGAGTTTCTTGGGCACTGGCCGTGGGACTTTGAGT GGTGATGCTTTTTTACACCTTCTTCCACATTTCTCATGCAAGTCACCACCATAGTCATAGCCATGAAGAAC CAGCAATGGAATGAAAAGAGGACCACTTTTCAGTCATCTGTCTTCTCAAACATAGAAGAAAGTGCCCTA TTTTGATTCCACGTGGAAGGGTCTAACAGCTCTAGGAGGCTGTATTTTCATGTTTCTTGTGTAACATGTC CTCAGATTGATCAACAAATTTAAAGATAAGAAGAAAAGAAATCAGAAGAAACCTGAAAATGATGATGATG TGGAGATTAAAGAGCAGTTGTCCAAGTATGAATCTCAACTTTCAACAAATGAGGAGAAAGTAGATACAGA TGATCGAACTGAAGGCTATTACAGAGCAGACTCACAAGAGCCCTCCCACTTTGATTCTCAGCAGCCTGCA GTCTTGAAGAAGAAGAGGTCATGATAGCTCATGCTCATCCACAGGAAGTCTACAATGAATATGTACCCA	136

[illegible]

	TAATGCGCCTCCGTGGCATGACGCACGCGTTTGGTGTCCGCCGCCGTCTCATGTCCGCCGCGTGTGGACC CCCTTTTCTCTCGGGGCACATCCCCCTATTCCTTGCCCTTTGGGGGGCACCCTCTAGACCCGCGCT TCTCTTCTCGTCCGTGGGGGACATTGGTTTGCTGCGCGCGCGGGGCGNTAAAAATAAAACAGCCTG TTAGCCCGGCCAGTACCCCCCGCGCGGGCGCCTTNCGTITGCAITATACCCCAACCCATAAAG CCGCGCCCTTTAGCNCCNTAACTTTTGTGGTGTGGCCTCCCCCTTTTCCCGGGGAGCAGCAACGGAC ATCTGTACACTAATGCTGGCCCCGACCTTTCCCAAAAACCCCCGCGCGTGTCCCGTATAAATTTGGTGC CAANCCTGACGNGTCTCCCCCGCCTCGCCCCGTGGCCGCCCTTTAAAGCCCCCGGTGGTTGCGC CGCCCAACGAGTCCACCTATAGTTAANTCCACCAACACCCCACTTTTCTCCCGCGGCATCTTCCCC ACGTACCCCTTTTGTGCGGAGATGGCCACTCCCCCCCCCTGTTTGTAAAAACAACGAGATGGTGT GCCAACGCTGGTCTTTTCCCCCGCGACCGCGACCGCCAGGGGAATACGTACCATAAGCCCCCGCGCC CNCTTTTTCCTCCCTCCCGGCCAATCAAGATCCGCGTCCATTAGACGTATTATTTTCCCGCGATAC ACGAAAAAACAGGGCCGCCATTTATAACTAAATTCCTGCGCGCGCGCGGATATGTTTCCCAAAATA CCACCCCCCCCCCAATTTCTTTGCCCCCACTCCTGCGCACCGGTGTTACACGCTCGCGCGC	
BC032677	GGACGCGTGGGTGACCCACGCGTCCGGACCCACGCGTCCGGTCTGTCTTCTCCGAGTTCCTGTCTCTCTG CCAACGCCCGCCCGATGGCTTCCCAAAACCGCGACCCAGCGCCACTAGCGTCCGCCCGCGCCGTAAGG AGCTGAGCCGAGCGGGGGCGCCCGCGGGTCCGGTGGGCAAAAGGCTACAGCAGGAGCTGATGACCCTC ATGGTGAGTGATTAAAGTGCCAGAACCCAGCCTTCCATCCAATTTTCAGTAGCCTCTTTTTCGGTCA GCTTTTGTCTAGACATAGGGGTAATGTAATTTGCTCCCTCCTGGGAAAGAAGTTTATACACCCACCTA CACCATTTCTTCCAGCAGTCCCTCCTCCCAATTCATCCCCCACAGAGTTATCTCGAACACTTCCCT GAAGTCATACAAGACCTCCCTATCCAGTGTCTCCTACTTCTAGCCCCAACCAAGCTTTACCCACACC CAAGTCCCCGCGCTTCTTGGTATTTAGCCTATGAATTTGGTTGCTTTATTTTGGAGTCAGATGAG ATTAAGGGGAGGCTGGGCGCGGTAGCTCACACCTTATAATCCCAAGTGTGGGATTACAGGCGTGAGCC ACCGCGCCCGGCCAGCACTAATATTCTAATTGAAGTAAAGCACAGGATGCCAATTACAAATCCTTAGAC CAAAGAGTCACTGATGTCTCCACCAGATAAGAGGAAAGCATCAGGCTAGGCATAGTGGCTCACACCTGTA ATCTCAGCACTTTGGGAGGCTGAGGCAGGCAGATCACATGAGCCAGGAGTTTGAGACTGGCCTGGGCAA CATGGTGAACCCCTGTCTCTAAATAAAAACTAACTAAAAAACTTTTAAAAAGGCAGTGGGAGCAT CAGAACCAGCTCAACAGTTTGTCTACTGTCCGGTCCCAGAGAACTCAAGATTCTAGCAAGCCCTTGTG TGGGGCTTGGGTTGGGACATGAGGCTGCTGCTGGAGCTTACTCTGCAACTGTTTCTCCAAATGCCAGGTA TATGAAGACCTGAGGTATAAGCTCTGCTAGAGTTCCCCAGTGGCTACCTTACAAATCGGCCACAGTGA AGTTCTCTACGCCCTGCTATCACCCCAACGTGGACACCCAGGGTAACATATGCCTGGACATCCTGAAGGA AAAGTGGTCTGCCCTGTATGATGTGAGGACCATTTCTGCTCTCCATCCAGAGCCTTCTAGGAGAACCCAAC ATTGATAGTCCCTTGAACACACATGCTGCCGAGCTCTGGAAAAACCCACAGCTTTTAAAGTACCTGC AAGAAACCTACTCAAAGCAGGTACCCAGCCAGGAGCCCTGACCCAGGCTGCCAGCCTGTCTTGTGTG TCTTTTAAATTTTTCCTTAGATGGTCTGICCTTTTGTGATTTCGTATAGGACTCTTATCTTGAGCTG TGGTATTTTGTGTTTGTGTTTGTGTTTAAATTAAGCCTCGGTTGAGCCCTTGTATATTAATAAATGCA TTTTGTCTTTTAAAAAATAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA A	139

[32] At least 40, at least 41, at least 42, at least 43, at least 44, at least 46 or all 46 of the genes in Table 1 can be utilized in the methods of the present invention. Preferably, the expression of each of the 46 genes is determined in a biological sample. The prototypical gene expression profiles (*i.e.* centroid) of the four intrinsic subtypes were pre-defined from a training set of FFPE breast tumor samples using hierarchical clustering analysis of gene expression data. A heatmap of the prototypical gene expression profiles (*i.e.* centroids) of these four subtypes is shown in Figure 1, where the level of expression is illustrated by the heatmap. **Table 3** shows the actual values.

[33] **Table 3.**

Tumor Subtype Centroids for Comparison to a Sample				
Target Gene	Basal-like	Her2-enriched	Luminal A	Luminal B
ACTR3B	-0.2052	-0.7965	-0.2790	-0.4380
ANLN	1.0227	0.5006	-0.7289	0.1149
BAG1	-0.4676	-0.3132	0.4716	0.5879
BCL2	-0.7365	-0.7237	0.7234	0.6363
BLVRA	-0.8761	0.2270	0.1628	0.7138

CCNE1	1.3100	0.2201	-0.6231	-0.2729
CDC20	1.0995	0.1445	-1.0518	-0.1173
CDC6	0.5817	0.6601	-0.7032	0.3134
CDCA1	0.9367	0.1623	-0.4509	0.2692
CDH3	0.7639	0.0144	-0.0502	-1.0229
CENPF	1.0222	0.2944	-0.5657	0.2437
CEP55	1.0442	0.4881	-0.6365	0.2921
CXXC5	-0.9732	0.1866	0.5687	0.9463
EGFR	0.3352	-0.1326	-0.0011	-0.9755
ERBB2	-0.7045	1.4182	0.2420	0.1978
ESR1	-1.1847	-0.4926	0.7177	1.0101
EXO1	1.0546	0.4317	-0.7259	0.2559
FGFR4	-0.2073	1.4562	0.1707	-0.2223
FOXA1	-1.3590	0.5726	0.7131	0.7963
FOXC1	1.0666	-0.7362	-0.4078	-0.9877
GPR160	-1.0540	0.5524	0.6032	0.7305
KIF2C	0.9242	0.1104	-1.1001	-0.2771
KNTC2	1.1373	0.2266	-0.7593	0.1656
KRT14	0.4759	-0.5269	0.8187	-0.8879
KRT17	0.6863	-0.3777	0.6149	-1.1415
KRT5	0.7136	-0.4146	0.5832	-0.9462
MAPT	-1.1343	-0.2711	1.0957	0.8372
MDM2	-0.7498	-0.4855	-0.1788	0.2397
MELK	1.0209	0.2678	-0.8016	0.1012
MIA	1.2408	-0.5475	0.3289	-0.6320
MKI67	1.0446	0.4630	-0.6717	0.3161
MLPH	-1.4150	0.4842	0.8829	0.8194
MMP11	-0.1295	0.5220	0.3402	0.5653
MYC	0.5639	-0.9904	-0.3015	-0.2791
NAT1	-0.9711	-0.2708	1.2256	0.9576
ORC6L	1.0086	0.5152	-1.0385	-0.0336
PGR	-0.9216	-0.5755	1.2061	0.9278
PHGDH	0.9192	0.0322	-0.5194	-0.5371
PTTG1	0.9541	0.2079	-1.1207	0.1052
RRM2	0.7895	0.6336	-0.8099	0.3228
SFRP1	0.7694	-0.8271	0.2617	-1.0846
SLC39A6	-0.9992	-0.4573	0.6607	0.9222
TMEM45B	-1.0721	0.7926	0.3190	0.2016
TYMS	0.9823	-0.0960	-0.8593	0.1827
UBE2C	0.8294	0.3358	-1.0141	0.0608
UBE2T	0.6258	0.0617	-0.8652	-0.0487

[34] After performing the Breast Cancer Intrinsic Subtyping test with a test breast cancer tumor sample and the reference sample provided as part of the test kit, a computational algorithm based on a Pearson's correlation compares the normalized and scaled gene expression profile of the NANO46 intrinsic gene set of the test sample to the prototypical expression signatures of the four breast cancer intrinsic subtypes. The intrinsic subtype

analysis is determined by determining the expression of a NANO50 set of genes (which is determining the expression of the NANO46 set of genes and further includes determining the expression of MYBL2, BIRC5, GRB7 and CCNB1) and the risk of recurrence (“ROR”) is determined using the NANO46 set of genes). Specifically, the intrinsic subtype is identified by comparing the expression of the NANO50 set of genes in the biological sample with the expected expression profiles for the four intrinsic subtypes. The subtype with the most similar expression profile is assigned to the biological sample. The ROR score is an integer value on a 0-100 scale that is related to an individual patient’s probability of distant recurrence within 10 years for the defined intended use population. The ROR score is calculated by comparing the expression profiles of the NANO46 genes in the biological sample with the expected profiles for the four intrinsic subtypes, as described above, to calculate four different correlation values. These correlation values are then combined with a proliferation score (and optionally one or more clinicopathological variables, such as tumor size) to calculate the ROR score. Preferably, the ROR score is calculated by comparing only the expression profiles of the NANO46 genes.

[35] **Figure 6** provides a schematic of the specific algorithm transformations. The tumor sample is assigned the subtype with the largest positive correlation to the sample. Kaplan Meier survival curves generated from a training set of untreated breast cancer patients demonstrate that the intrinsic subtypes are a prognostic indicator of recurrence free survival (RFS) in this test population, which includes both estrogen receptor positive/negative and HER2 positive/negative patients, **Figure 2**.

[36] Independent testing on a cohort of node negative, estrogen receptor positive patients treated with tamoxifen shows predominantly Luminal A and B subtype patients with Luminal A patients exhibiting better outcome than Luminal B patients, **Figure 3**. The outcome of Luminal A patients is expected to improve even further using clinical trial specimens that use more modern treatment regimens (*i.e.* aromatase inhibitors) and have better adherence to therapy which will improve outcome

[37] The training set of FFPE breast tumor samples, which had well defined clinical characteristics and clinical outcome data, were used to establish a continuous Risk of Recurrence (ROR) score. The score is calculated using coefficients from a Cox model that includes correlation to each intrinsic subtype, a proliferation score (mean gene expression of a subset of 18 of the 46 genes), and tumor size, **Table 4**.

Table 4. Coefficients to calculate ROR-PT (equation 1)	
Test Variables	Coefficient
Basal-like Pearson's correlation (A)	- 0.0067
Her2-enriched Pearson's correlation (B)	0.4317
Luminal A Pearson's correlation (C)	- 0.3172
Luminal B Pearson's correlation (D)	0.4894
Proliferation Score (E)	0.1981
Tumor Size (F)	0.1133

[38] The test variables in **Table 4** are multiplied by the corresponding coefficients and summed to produce a risk score ("ROR-PT").

[39] **ROR-PT equation** = $-0.0067*A + 0.4317*B + -0.3172*C + 0.4894*D + 0.1981*E + 0.1133*F$

[40] In previous studies, the ROR score provided a continuous estimate of the risk of recurrence for ER-positive, node-negative patients who were treated with tamoxifen for 5 years (Nielsen *et al.* Clin. Cancer Res., 16(21):5222-5232 (2009)). This result was verified on ER-positive, node-negative patients from the same cohort, **Figure 4**. The ROR score also exhibited a statistically significant improvement over a clinical model based in determining RFS within this test population providing further evidence of the improved accuracy of this decision making tool when compared to traditional clinicopathological measures (Nielsen *et al.* Clin. Cancer Res., 16(21):5222-5232 (2009)).

[41] The gene set contains many genes that are known markers for proliferation. The methods of the present invention provide for the determination of subsets of genes that provide a proliferation signature. The methods of the present invention can include determining the expression of at least one of, a combination of, or each of, a 18-gene subset of the NANO46 intrinsic genes selected from ANLN, CCNE1, CDC20, CDC6, CDCA1, CENPF, CEP55, EXO1, KIF2C, KNTC2, MELK, MKI67, ORC6L, PTTG1, RRM2, TYMS, UBE2C and/or UBE2T. Preferably, the expression of each of the 18-gene subset of the NANO46 gene set is determined to provide a proliferation score. The expression of one or more of these genes may be determined and a proliferation signature index can be generated by averaging the normalized expression estimates of one or more of these genes in a sample. The sample can be assigned a high proliferation signature, a moderate/intermediate proliferation signature, a low proliferation signature or an ultra-low proliferation signature. Methods of determining a proliferation signature from a biological sample are as described in Nielsen *et al.* Clin. Cancer Res., 16(21):5222-5232 (2009).

[42] *Description of Intrinsic Subtype Biology*

[43] Luminal subtypes: The most common subtypes of breast cancer are the luminal subtypes, Luminal A and Luminal B. Prior studies suggest that luminal A comprises approximately 30% to 40% and luminal B approximately 20% of all breast cancers, but they represent over 90 % of hormone receptor positive breast cancers (Nielsen *et al.* Clin. Cancer Res., 16(21):5222-5232 (2009)). The gene expression pattern of these subtypes resembles the luminal epithelial component of the breast. These tumors are characterized by high expression of estrogen receptor (ER), progesterone receptor (PR), and genes associated with ER activation, such as LIV1, GATA3, and cyclin D1, as well as expression of luminal cytokeratins 8 and 18 (Lisa Carey & Charles Perou (2009). Gene Arrays, Prognosis, and Therapeutic Interventions. Jay R. Harris et al. (4th ed.), Diseases of the breast (pp. 458-472). Philadelphia, PA: Lippincott Williams & Wilkins).

[44] Luminal A: Luminal A (LumA) breast cancers exhibit low expression of genes associated with cell cycle activation and the ERBB2 cluster resulting in a better prognosis than Luminal B. The Luminal A subgroup has the most favorable prognosis of all subtypes and is enriched for endocrine therapy-responsive tumors.

[45] Luminal B: Luminal B (LumB) breast cancers also express ER and ER-associated genes. Genes associated with cell cycle activation are highly expressed and this tumor type can be HER2(+) (~20%) or HER2(-). The prognosis is unfavorable (despite ER expression) and endocrine therapy responsiveness is generally diminished relative to LumA.

[46] HER2-enriched: The HER2-enriched subtype is generally ER-negative and is HER2-positive in the majority of cases with high expression of the ERBB2 cluster, including ERBB2 and GRB7. Genes associated with cell cycle activation are highly expressed and these tumors have a poor outcome.

[47] Basal-like: The Basal-like subtype is generally ER-negative, is almost always clinically HER2-negative and expresses a suite of “basal” biomarkers including the basal epithelial cytokeratins (CK) and epidermal growth factor receptor (EGFR). Genes associated with cell cycle activation are highly expressed.

[48] *Clinical variables*

[49] The NANO46 classification model described herein may be further combined with information on clinical variables to generate a continuous risk of recurrence (ROR) predictor. As described herein, a number of clinical and prognostic breast cancer factors are known in the art and are used to predict treatment outcome and the likelihood of disease recurrence.

Such factors include, for example, lymph node involvement, tumor size, histologic grade, estrogen and progesterone hormone receptor status, HER-2 levels, and tumor ploidy. In one embodiment, risk of recurrence (ROR) score is provided for a subject diagnosed with or suspected of having breast cancer. This score uses the NANO46 classification model in combination with clinical factors of lymph node status (N) and tumor size (T). Assessment of clinical variables is based on the American Joint Committee on Cancer (AJCC) standardized system for breast cancer staging. In this system, primary tumor size is categorized on a scale of 0-4 (T0: no evidence of primary tumor; T1 : < 2 cm; T2: > 2 cm - < 5 cm; T3 : > 5 cm; T4: tumor of any size with direct spread to chest wall or skin). Lymph node status is classified as N0-N3 (N0: regional lymph nodes are free of metastasis; N1 : metastasis to movable, same-side axillary lymph node(s); N2: metastasis to same-side lymph node(s) fixed to one another or to other structures; N3: metastasis to same-side lymph nodes beneath the breastbone). Methods of identifying breast cancer patients and staging the disease are well known and may include manual examination, biopsy, review of patient's and/or family history, and imaging techniques, such as mammography, magnetic resonance imaging (MRI), and positron emission tomography (PET).

[50] *Sample Source*

[51] In one embodiment of the present disclosure, breast cancer subtype is assessed through the evaluation of expression patterns, or profiles, of the intrinsic genes listed in Table 1 in one or more subject samples. For the purpose of discussion, the term subject, or subject sample, refers to an individual regardless of health and/or disease status. A subject can be a subject, a study participant, a control subject, a screening subject, or any other class of individual from whom a sample is obtained and assessed in the context of the disclosure. Accordingly, a subject can be diagnosed with breast cancer, can present with one or more symptoms of breast cancer, or a predisposing factor, such as a family (genetic) or medical history (medical) factor, for breast cancer, can be undergoing treatment or therapy for breast cancer, or the like. Alternatively, a subject can be healthy with respect to any of the aforementioned factors or criteria. It will be appreciated that the term "healthy" as used herein, is relative to breast cancer status, as the term "healthy" cannot be defined to correspond to any absolute evaluation or status. Thus, an individual defined as healthy with reference to any specified disease or disease criterion, can in fact be diagnosed with any other one or more diseases, or exhibit any other one or more disease criterion, including one or more cancers other than breast cancer. However, the healthy controls are preferably free of any cancer.

[52] In particular embodiments, the methods for predicting breast cancer intrinsic subtypes include collecting a biological sample comprising a cancer cell or tissue, such as a breast tissue sample or a primary breast tumor tissue sample. By "biological sample" is intended any sampling of cells, tissues, or bodily fluids in which expression of an intrinsic gene can be detected. Examples of such biological samples include, but are not limited to, biopsies and smears. Bodily fluids useful in the present disclosure include blood, lymph, urine, saliva, nipple aspirates, gynecological fluids, or any other bodily secretion or derivative thereof. Blood can include whole blood, plasma, serum, or any derivative of blood. In some embodiments, the biological sample includes breast cells, particularly breast tissue from a biopsy, such as a breast tumor tissue sample. Biological samples may be obtained from a subject by a variety of techniques including, for example, by scraping or swabbing an area, by using a needle to aspirate cells or bodily fluids, or by removing a tissue sample (*i.e.*, biopsy). Methods for collecting various biological samples are well known in the art. In some embodiments, a breast tissue sample is obtained by, for example, fine needle aspiration biopsy, core needle biopsy, or excisional biopsy. Fixative and staining solutions may be applied to the cells or tissues for preserving the specimen and for facilitating examination. Biological samples, particularly breast tissue samples, may be transferred to a glass slide for viewing under magnification. In one embodiment, the biological sample is a formalin-fixed, paraffin-embedded breast tissue sample, particularly a primary breast tumor sample. In various embodiments, the tissue sample is obtained from a pathologist-guided tissue core sample.

[53] *Expression Profiling*

[54] In various embodiments, the present disclosure provides methods for classifying, prognosticating, or monitoring breast cancer in subjects. In this embodiment, data obtained from analysis of intrinsic gene expression is evaluated using one or more pattern recognition algorithms. Such analysis methods may be used to form a predictive model, which can be used to classify test data. For example, one convenient and particularly effective method of classification employs multivariate statistical analysis modeling, first to form a model (a "predictive mathematical model") using data ("modeling data") from samples of known subtype (*e.g.*, from subjects known to have a particular breast cancer intrinsic subtype: LumA, LumB, Basal-like, HER2-enriched, or normal-like), and second to classify an unknown sample (*e.g.*, "test sample") according to subtype. Pattern recognition methods have been used widely to characterize many different types of problems ranging, for example, over linguistics, fingerprinting, chemistry and psychology. In the context of the

methods described herein, pattern recognition is the use of multivariate statistics, both parametric and non-parametric, to analyze data, and hence to classify samples and to predict the value of some dependent variable based on a range of observed measurements. There are two main approaches. One set of methods is termed "unsupervised" and these simply reduce data complexity in a rational way and also produce display plots which can be interpreted by the human eye. However, this type of approach may not be suitable for developing a clinical assay that can be used to classify samples derived from subjects independent of the initial sample population used to train the prediction algorithm.

[55] The other approach is termed "supervised" whereby a training set of samples with known class or outcome is used to produce a mathematical model which is then evaluated with independent validation data sets. Here, a "training set" of intrinsic gene expression data is used to construct a statistical model that predicts correctly the "subtype" of each sample. This training set is then tested with independent data (referred to as a test or validation set) to determine the robustness of the computer-based model. These models are sometimes termed "expert systems," but may be based on a range of different mathematical procedures.

Supervised methods can use a data set with reduced dimensionality (for example, the first few principal components), but typically use unreduced data, with all dimensionality. In all cases the methods allow the quantitative description of the multivariate boundaries that characterize and separate each subtype in terms of its intrinsic gene expression profile. It is also possible to obtain confidence limits on any predictions, for example, a level of probability to be placed on the goodness of fit. The robustness of the predictive models can also be checked using cross-validation, by leaving out selected samples from the analysis.

[56] The NANO46 classification model described herein is based on the gene expression profile for a plurality of subject samples using the intrinsic genes listed in Table 1. The plurality of samples includes a sufficient number of samples derived from subjects belonging to each subtype class. By "sufficient samples" or "representative number" in this context is intended a quantity of samples derived from each subtype that is sufficient for building a classification model that can reliably distinguish each subtype from all others in the group. A supervised prediction algorithm is developed based on the profiles of objectively-selected prototype samples for "training" the algorithm. The samples are selected and subtyped using an expanded intrinsic gene set according to the methods disclosed in International Patent Publication WO 2007/061876 and US Patent Publication No. 2009/0299640. Alternatively, the samples can be subtyped according to any known assay for classifying breast cancer subtypes. After stratifying the training samples according to subtype, a centroid-based

prediction algorithm is used to construct centroids based on the expression profile of the intrinsic gene set described in Table 1.

[57] In one embodiment, the prediction algorithm is the nearest centroid methodology related to that described in Narashiman and Chu (2002) PNAS 99:6567-6572. In the present disclosure, the method computes a standardized centroid for each subtype. This centroid is the average gene expression for each gene in each subtype (or "class") divided by the within-class standard deviation for that gene. Nearest centroid classification takes the gene expression profile of a new sample, and compares it to each of these class centroids. Subtype prediction is done by calculating the Spearman's rank correlation of each test case to the five centroids, and assigning a sample to a subtype based on the nearest centroid.

[58] *Detection of intrinsic gene expression*

[59] Any methods available in the art for detecting expression of the intrinsic genes listed in Table 1 are encompassed herein. By "detecting expression" is intended determining the quantity or presence of an RNA transcript or its expression product of an intrinsic gene. Methods for detecting expression of the intrinsic genes of the disclosure, that is, gene expression profiling, include methods based on hybridization analysis of polynucleotides, methods based on sequencing of polynucleotides, immunohistochemistry methods, and proteomics-based methods. The methods generally detect expression products (*e.g.*, mRNA) of the intrinsic genes listed in Table 1. In preferred embodiments, PCR-based methods, such as reverse transcription PCR (RT-PCR) (Weis *et al.*, TIG 8:263- 64, 1992), and array-based methods such as microarray (Schena *et al.*, Science 270:467- 70, 1995) are used. By "microarray" is intended an ordered arrangement of hybridizable array elements, such as, for example, polynucleotide probes, on a substrate. The term "probe" refers to any molecule that is capable of selectively binding to a specifically intended target biomolecule, for example, a nucleotide transcript or a protein encoded by or corresponding to an intrinsic gene. Probes can be synthesized by one of skill in the art, or derived from appropriate biological preparations. Probes may be specifically designed to be labeled. Examples of molecules that can be utilized as probes include, but are not limited to, RNA, DNA, proteins, antibodies, and organic molecules.

[60] Many expression detection methods use isolated RNA. The starting material is typically total RNA isolated from a biological sample, such as a tumor or tumor cell line, and corresponding normal tissue or cell line, respectively. If the source of RNA is a primary

tumor, RNA (*e.g.*, mRNA) can be extracted, for example, from frozen or archived paraffin-embedded and fixed (*e.g.*, formalin-fixed) tissue samples (*e.g.*, pathologist-guided tissue core samples).

[61] General methods for RNA extraction are well known in the art and are disclosed in standard textbooks of molecular biology, including Ausubel *et al.*, ed., Current Protocols in Molecular Biology, John Wiley & Sons, New York 1987-1999. Methods for RNA extraction from paraffin embedded tissues are disclosed, for example, in Rupp and Locker, Lab Invest. 56:A67, (1987); and De Andres *et al.* Biotechniques 18:42-44, (1995). In particular, RNA isolation can be performed using a purification kit, a buffer set and protease from commercial manufacturers, such as Qiagen (Valencia, CA), according to the manufacturer's instructions. For example, total RNA from cells in culture can be isolated using Qiagen RNeasy mini-columns. Other commercially available RNA isolation kits include MASTERPURE™ Complete DNA and RNA Purification Kit (Epicentre, Madison, Wis.) and Paraffin Block RNA Isolation Kit (Ambion, Austin, TX). Total RNA from tissue samples can be isolated, for example, using RNA Stat-60 (Tel-Test, Friendswood, TX). Total RNA from FFPE can be isolated, for example, using High Pure FFPE RNA Microkit, Cat No. 04823125001 (Roche Applied Science, Indianapolis, IN). RNA prepared from a tumor can be isolated, for example, by cesium chloride density gradient centrifugation. Additionally, large numbers of tissue samples can readily be processed using techniques well known to those of skill in the art, such as, for example, the single-step RNA isolation process of Chomczynski (U.S. Pat. No. 4,843,155).

[62] Isolated RNA can be used in hybridization or amplification assays that include, but are not limited to, PCR analyses and probe arrays. One method for the detection of RNA levels involves contacting the isolated RNA with a nucleic acid molecule (probe) that can hybridize to the mRNA encoded by the gene being detected. The nucleic acid probe can be, for example, a full-length cDNA, or a portion thereof, such as an oligonucleotide of at least 7, 15, 30, 60, 100, 250, or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to an intrinsic gene of the present disclosure, or any derivative DNA or RNA. Hybridization of an mRNA with the probe indicates that the intrinsic gene in question is being expressed.

[63] In one embodiment, the mRNA is immobilized on a solid surface and contacted with a probe, for example by running the isolated mRNA on an agarose gel and transferring the mRNA from the gel to a membrane, such as nitrocellulose. In an alternative embodiment, the probes are immobilized on a solid surface and the mRNA is contacted with the probes, for

example, in an Agilent gene chip array. A skilled artisan can readily adapt known mRNA detection methods for use in detecting the level of expression of the intrinsic genes of the present disclosure.

[64] An alternative method for determining the level of intrinsic gene expression product in a sample involves the process of nucleic acid amplification, for example, by RT-PCR (U.S. Pat. No. 4,683,202), ligase chain reaction (Barany, PNAS USA 88: 189-93, (1991)), self sustained sequence replication (Guatelli *et al.*, Proc. Natl. Acad. Sci USA 87: 1874-78, (1990)), transcriptional amplification system (Kwoh *et al.*, Proc. Natl. Acad. Sci USA 86: 1173-77, (1989)), Q-Beta Replicase (Lizardi *et al.*, Bio/Technology 6:1197, (1988)), rolling circle replication (U.S. Pat. No. 5,854,033), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

[65] In particular aspects of the disclosure, intrinsic gene expression is assessed by quantitative RT-PCR. Numerous different PCR or QPCR protocols are known in the art and exemplified herein below and can be directly applied or adapted for use using the presently-described compositions for the detection and/or quantification of the intrinsic genes listed in Table 1. Generally, in PCR, a target polynucleotide sequence is amplified by reaction with at least one oligonucleotide primer or pair of oligonucleotide primers. The primer(s) hybridize to a complementary region of the target nucleic acid and a DNA polymerase extends the primer(s) to amplify the target sequence. Under conditions sufficient to provide polymerase-based nucleic acid amplification products, a nucleic acid fragment of one size dominates the reaction products (the target polynucleotide sequence which is the amplification product). The amplification cycle is repeated to increase the concentration of the single target polynucleotide sequence. The reaction can be performed in any thermocycler commonly used for PCR. However, preferred are cyclers with real time fluorescence measurement capabilities, for example, SMARTCYCLER® (Cepheid, Sunnyvale, CA), ABI PRISM 7700® (Applied Biosystems, Foster City, Calif.), ROTOR- GENE™ (Corbett Research, Sydney, Australia), LIGHTCYCLER® (Roche Diagnostics Corp, Indianapolis, Ind.), ICYCLER® (Biorad Laboratories, Hercules, Calif.) and MX4000® (Stratagene, La Jolla, Calif.).

[66] In another embodiment of the disclosure, microarrays are used for expression profiling. Microarrays are particularly well suited for this purpose because of the reproducibility between different experiments. DNA microarrays provide one method for the

simultaneous measurement of the expression levels of large numbers of genes. Each array consists of a reproducible pattern of capture probes attached to a solid support. Labeled RNA or DNA is hybridized to complementary probes on the array and then detected by laser scanning. Hybridization intensities for each probe on the array are determined and converted to a quantitative value representing relative gene expression levels. See, for example, U.S. Pat. Nos. 6,040,138, 5,800,992 and 6,020,135, 6,033,860, and 6,344,316. High-density oligonucleotide arrays are particularly useful for determining the gene expression profile for a large number of RNAs in a sample.

[67] In a preferred embodiment, the nCounter® Analysis system is used to detect intrinsic gene expression. The basis of the nCounter® Analysis system is the unique code assigned to each nucleic acid target to be assayed (International Patent Application Publication No. WO 08/124847, US Patent No. 8,415,102 and Geiss *et al.* Nature Biotechnology, 2008, 26(3): 317-325. The code is composed of an ordered series of colored fluorescent spots which create a unique barcode for each target to be assayed. A pair of probes is designed for each DNA or RNA target, a biotinylated capture probe and a reporter probe carrying the fluorescent barcode. This system is also referred to, herein, as the nanoreporter code system.

[68] Specific reporter and capture probes are synthesized for each target. Briefly, sequence-specific DNA oligonucleotide probes are attached to code-specific reporter molecules. Preferably, each sequence specific reporter probe comprises a target specific sequence capable of hybridizing to no more than one NANO46 gene of Table 1 and optionally comprises at least two, at least three, or at least four label attachment regions, said attachment regions comprising one or more label monomers that emit light. Capture probes are made by ligating a second sequence-specific DNA oligonucleotide for each target to a universal oligonucleotide containing biotin. Reporter and capture probes are all pooled into a single hybridization mixture, the “probe library”. Preferably, the probe library comprises a probe pair (a capture probe and reporter) for each of the NANO46 genes in Table 1.

[69] The relative abundance of each target is measured in a single multiplexed hybridization reaction. The method comprises contacting a biological sample with a probe library, the library comprising a probe pair for the NANO46 genes in Table 1, such that the presence of the target in the sample creates a probe pair – target complex. The complex is then purified. More specifically, the sample is combined with the probe library, and hybridization occurs in solution. After hybridization, the tripartite hybridized complexes (probe pairs and target) are purified in a two-step procedure using magnetic beads linked to

oligonucleotides complementary to universal sequences present on the capture and reporter probes. This dual purification process allows the hybridization reaction to be driven to completion with a large excess of target-specific probes, as they are ultimately removed, and, thus, do not interfere with binding and imaging of the sample. All post hybridization steps are handled robotically on a custom liquid-handling robot (Prep Station, NanoString Technologies).

[70] Purified reactions are deposited by the Prep Station into individual flow cells of a sample cartridge, bound to a streptavidin-coated surface via the capture probe, electrophoresed to elongate the reporter probes, and immobilized. After processing, the sample cartridge is transferred to a fully automated imaging and data collection device (Digital Analyzer, NanoString Technologies). The expression level of a target is measured by imaging each sample and counting the number of times the code for that target is detected. Data is output in simple spreadsheet format listing the number of counts per target, per sample.

[71] This system can be used along with nanoreporters. Additional disclosure regarding nanoreporters can be found in International Publication No. WO 07/076129 and WO 07/076132, and US Patent Publication No. 2010/0015607 and 2010/0261026. Further, the term nucleic acid probes and nanoreporters can include the rationally designed (*e.g.* synthetic sequences) described in International Publication No. WO 2010/019826 and US Patent Publication No. 2010/0047924.

[72] *Data processing*

[73] It is often useful to pre-process gene expression data, for example, by addressing missing data, translation, scaling, normalization, weighting, *etc.* Multivariate projection methods, such as principal component analysis (PCA) and partial least squares analysis (PLS), are so-called scaling sensitive methods. By using prior knowledge and experience about the type of data studied, the quality of the data prior to multivariate modeling can be enhanced by scaling and/or weighting. Adequate scaling and/or weighting can reveal important and interesting variation hidden within the data, and therefore make subsequent multivariate modeling more efficient. Scaling and weighting may be used to place the data in the correct metric, based on knowledge and experience of the studied system, and therefore reveal patterns already inherently present in the data.

[74] If possible, missing data, for example gaps in column values, should be avoided. However, if necessary, such missing data may be replaced or "filled" with, for example, the

mean value of a column ("mean fill"); a random value ("random fill"); or a value based on a principal component analysis ("principal component fill").

[75] "Translation" of the descriptor coordinate axes can be useful. Examples of such translation include normalization and mean centering. "Normalization" may be used to remove sample-to-sample variation. For microarray data, the process of normalization aims to remove systematic errors by balancing the fluorescence intensities of the two labeling dyes. The dye bias can come from various sources including differences in dye labeling efficiencies, heat and light sensitivities, as well as scanner settings for scanning two channels. Some commonly used methods for calculating normalization factor include: (i) global normalization that uses all genes on the array; (ii) housekeeping genes normalization that uses constantly expressed housekeeping/invariant genes; and (iii) internal controls normalization that uses known amount of exogenous control genes added during hybridization (Quackenbush Nat. Genet. 32 (Suppl.), 496-501 (2002)). In one embodiment, the intrinsic genes disclosed herein can be normalized to control housekeeping genes. For example, the housekeeping genes described in U.S. Patent Publication 2008/0032293 can be used for normalization. Exemplary housekeeping genes include MRPL19, PSMC4, SF3A1, PUM1, ACTB, GAPD, GUSB, RPLPO, and TFRC. It will be understood by one of skill in the art that the methods disclosed herein are not bound by normalization to any particular housekeeping genes, and that any suitable housekeeping gene(s) known in the art can be used.

[76] Many normalization approaches are possible, and they can often be applied at any of several points in the analysis. In one embodiment, microarray data is normalized using the LOWESS method, which is a global locally weighted scatter plot smoothing normalization function. In another embodiment, qPCR data is normalized to the geometric mean of set of multiple housekeeping genes.

[77] "Mean centering" may also be used to simplify interpretation. Usually, for each descriptor, the average value of that descriptor for all samples is subtracted. In this way, the mean of a descriptor coincides with the origin, and all descriptors are "centered" at zero. In "unit variance scaling," data can be scaled to equal variance. Usually, the value of each descriptor is scaled by $1/\text{StDev}$, where StDev is the standard deviation for that descriptor for all samples. "Pareto scaling" is, in some sense, intermediate between mean centering and unit variance scaling. In pareto scaling, the value of each descriptor is scaled by $1/\sqrt{\text{StDev}}$, where StDev is the standard deviation for that descriptor for all samples. In this way, each descriptor has a variance numerically equal to its initial standard deviation. The pareto

scaling may be performed, for example, on raw data or mean centered data.

[78] "Logarithmic scaling" may be used to assist interpretation when data have a positive skew and/or when data spans a large range, *e.g.*, several orders of magnitude. Usually, for each descriptor, the value is replaced by the logarithm of that value. In "equal range scaling," each descriptor is divided by the range of that descriptor for all samples. In this way, all descriptors have the same range, that is, 1. However, this method is sensitive to presence of outlier points. In "autoscaling," each data vector is mean centered and unit variance scaled. This technique is a very useful because each descriptor is then weighted equally, and large and small values are treated with equal emphasis. This can be important for genes expressed at very low, but still detectable, levels.

[79] In one embodiment, data is collected for one or more test samples and classified using the NANO46 classification model described herein. When comparing data from multiple analyses (*e.g.*, comparing expression profiles for one or more test samples to the centroids constructed from samples collected and analyzed in an independent study), it will be necessary to normalize data across these data sets. In one embodiment, Distance Weighted Discrimination (DWD) is used to combine these data sets together (Benito *et al.* (2004) *Bioinformatics* 20(1): 105-114. DWD is a multivariate analysis tool that is able to identify systematic biases present in separate data sets and then make a global adjustment to compensate for these biases; in essence, each separate data set is a multi-dimensional cloud of data points, and DWD takes two points clouds and shifts one such that it more optimally overlaps the other.

[80] The methods described herein may be implemented and/or the results recorded using any device capable of implementing the methods and/or recording the results. Examples of devices that may be used include but are not limited to electronic computational devices, including computers of all types. When the methods described herein are implemented and/or recorded in a computer, the computer program that may be used to configure the computer to carry out the steps of the methods may be contained in any computer readable medium capable of containing the computer program. Examples of computer readable medium that may be used include but are not limited to diskettes, CD-ROMs, DVDs, ROM, RAM, and other memory and computer storage devices. The computer program that may be used to configure the computer to carry out the steps of the methods and/or record the results may also be provided over an electronic network, for example, over the internet, an intranet, or other network.

[81] *Calculation of risk of recurrence*

[82] Provided herein are methods for predicting breast cancer outcome within the context of the intrinsic subtype and optionally other clinical variables. Outcome may refer to overall or disease-specific survival, event-free survival, or outcome in response to a particular treatment or therapy. In particular, the methods may be used to predict the likelihood of long-term, disease-free survival. "Predicting the likelihood of survival of a breast cancer patient" is intended to assess the risk that a patient will die as a result of the underlying breast cancer. "Long-term, disease-free survival" is intended to mean that the patient does not die from or suffer a recurrence of the underlying breast cancer within a period of at least five years, or at least ten or more years, following initial diagnosis or treatment.

[83] In one embodiment, outcome is predicted based on classification of a subject according to subtype. In addition to providing a subtype assignment, the NANO46 bioinformatics model provides a measurement of the similarity of a test sample to all four subtypes which is translated into a Risk of Recurrence (ROR) score that can be used in any patient population regardless of disease status and treatment options. The intrinsic subtypes and ROR also have value in the prediction of pathological complete response in women treated with, for example, neoadjuvant taxane and anthracycline chemotherapy (Rouzier *et al.*, J Clin Oncol 23:8331-9 (2005). Thus, in various embodiments of the present disclosure, a risk of recurrence (ROR) model is used to predict outcome. Using these risk models, subjects can be stratified into low, medium, and high risk of recurrence groups. Calculation of ROR can provide prognostic information to guide treatment decisions and/or monitor response to therapy.

[84] In some embodiments described herein, the prognostic performance of the NANO46-defined intrinsic subtypes and/or other clinical parameters is assessed utilizing a Cox Proportional Hazards Model Analysis, which is a regression method for survival data that provides an estimate of the hazard ratio and its confidence interval. The Cox model is a well-recognized statistical technique for exploring the relationship between the survival of a patient and particular variables. This statistical method permits estimation of the hazard (*i.e.*, risk) of individuals given their prognostic variables (*e.g.*, intrinsic gene expression profile with or without additional clinical factors, as described herein). The "hazard ratio" is the risk of death at any given time point for patients displaying particular prognostic variables. See generally Spruance *et al.*, Antimicrob. Agents & Chemo. 48:2787-92 (2004).

[85] The NANO46 classification model described herein can be trained for risk of recurrence using subtype distances (or correlations) alone, or using subtype distances with clinical variables as discussed *supra*. In one embodiment, the risk score for a test sample is

calculated using intrinsic subtype distances alone using the following equation:

[86] $ROR = 0.05*Basal + 0.11*Her2 + -0.25*LumA + 0.07*LumB + -0.11*Normal$, where the variables "Basal," "Her2," "LumA," "LumB," and "Normal" are the distances to the centroid for each respective classifier when the expression profile from a test sample is compared to centroids constructed using the gene expression data deposited with the Gene Expression Omnibus (GEO).

[87] Risk score can also be calculated using a combination of breast cancer subtype and the clinical variables tumor size (T) and lymph nodes status (N) using the following equation: $ROR (full) = 0.05*Basal + 0.11*Her2 + -0.19*LumA + 0.05*LumB + -0.09*Normal + 0.16*T + 0.08*N$, again when comparing test expression profiles to centroids constructed using the gene expression data deposited with GEO as accession number GSE2845.

[88] In yet another embodiment, risk score for a test sample is calculated using intrinsic subtype distances alone using the following equation:

[89] $ROR-S = 0.05*Basal + 0.12*Her2 + -0.34*LumA + 0.023*LumB$, where the variables "Basal," "Her2," "LumA," and "LumB" are as described *supra* and the test expression profiles are compared to centroids constructed using the gene expression data deposited with GEO as accession number GSE2845. In yet another embodiment, risk score can also be calculated using a combination of breast cancer subtype and the clinical variable tumor size (T) using the following equation (where the variables are as described *supra*): $ROR-C = 0.05*Basal + 0.11*Her2 + -0.23*LumA + 0.09*LumB + 0.17*T$.

[90] In yet another embodiment, risk score for a test sample is calculated using intrinsic subtype distances in combination with the proliferation signature ("Prolif") using the following equation:

[91] $ROR-P = -0.001*Basal + 0.7*Her2 + -0.95*LumA + 0.49*LumB + 0.34*Prolif$, where the variables "Basal," "Her2," "LumA," "LumB" and "Prolif" are as described *supra* and the test expression profiles are compared to centroids constructed using the gene expression data deposited with GEO as accession number GSE2845.

[92] In yet another embodiment, risk score can also be calculated using a combination of breast cancer subtype, proliferation signature and the clinical variable tumor size (T) using the ROR-PT described in conjunction with Table 3 *supra*.

[93] *Detection of Subtypes*

[94] Immunohistochemistry for estrogen (ER), progesterone (PgR), HER2, and Ki67 was performed concurrently on serial sections with the standard streptavidin-biotin complex method with 3,3'-diaminobenzidine as the chromogen. Staining for ER, PgR, and HER2

interpretation can be performed as described previously (Cheang *et al.*, Clin Cancer Res. 2008;14(5):1368–1376.), however any method known in the art may be used.

[95] For example, a Ki67 antibody (clone SP6; ThermoScientific, Fremont, CA) can be applied at a 1:200 dilution for 32 minutes, by following the Ventana Benchmark automated immunostainer (Ventana, Tucson AZ) standard Cell Conditioner 1 (CC1, a proprietary buffer) protocol at 98°C for 30 minutes. An ER antibody (clone SP1; ThermoFisher Scientific, Fremont CA) can be used at 1:250 dilution with 10-minute incubation, after an 8-minute microwave antigen retrieval in 10 mM sodium citrate (pH 6.0). Ready-to-use PR antibody (clone 1E2; Ventana) can be used by following the CC1 protocol as above. HER2 staining can be done with a SP3 antibody (ThermoFisher Scientific) at a 1:100 dilution after antigen retrieval in 0.05 M Tris buffer (pH 10.0) with heating to 95°C in a steamer for 30 minutes. For HER2 fluorescent *in situ* hybridization (FISH) assay, slides can be hybridized with probes to LSI (locus-specific identifier) HER2/neu and to centromere 17 by use of the PathVysion HER-2 DNA Probe kit (Abbott Molecular, Abbott Park, IL) according to manufacturer's instructions, with modifications to pretreatment and hybridization as previously described (Brown LA, Irving J, Parker R, *et al.* Amplification of EMSY, a novel oncogene on 11q13, in high grade ovarian surface epithelial carcinomas. Gynecol Oncol. 2006;100(2):264–270). Slides can then be counterstained with 4',6-diamidino-2-phenylindole, stained material was visualized on a Zeiss Axioplan epifluorescent microscope, and signals were analyzed with a Metafer image acquisition system (Metasystems, Altlusheim, Germany). Biomarker expression from immunohistochemistry assays can then be scored by two pathologists, who were blinded to the clinicopathological characteristics and outcome and who used previously established and published criteria for biomarker expression levels that had been developed on other breast cancer cohorts.

[96] Tumors were considered positive for ER or PR if immunostaining was observed in more than 1% of tumor nuclei, as described previously. Tumors were considered positive for HER2 if immunostaining was scored as 3+ according to HercepTest criteria, with an amplification ratio for fluorescent *in situ* hybridization of 2.0 or more being the cut point that was used to segregate immunohistochemistry equivocal tumors (scored as 2+) (Yaziji, *et al.*, JAMA, 291(16):1972–1977 (2004)). Ki67 was visually scored for percentage of tumor cell nuclei with positive immunostaining above the background level by two pathologists.

[97] Other methods can also be used to detect subtypes. These techniques include ELISA, Western blots, Northern blots, or FACS analysis.

[98] *Kits*

[99] The present disclosure also describes kits useful for classifying breast cancer intrinsic subtypes and/or providing prognostic information to identify risk of recurrence. These kits comprise a set of capture probes and/or primers specific for the intrinsic genes listed in Table 1. The kit may further comprise a computer readable medium.

[100] In one embodiment of the present disclosure, the capture probes are immobilized on an array. By "array" is intended a solid support or a substrate with peptide or nucleic acid probes attached to the support or substrate. Arrays typically comprise a plurality of different capture probes that are coupled to a surface of a substrate in different, known locations. The arrays of the disclosure comprise a substrate having a plurality of capture probes that can specifically bind an intrinsic gene expression product. The number of capture probes on the substrate varies with the purpose for which the array is intended. The arrays may be low-density arrays or high-density arrays and may contain 4 or more, 8 or more, 12 or more, 16 or more, 32 or more addresses, but will minimally comprise capture probes for the 46 intrinsic genes listed in Table 1.

[101] Techniques for the synthesis of these arrays using mechanical synthesis methods are described in, e.g., U.S. Patent No. 5,384,261. The array may be fabricated on a surface of virtually any shape or even a multiplicity of surfaces. Arrays may be probes (e.g., nucleic-acid binding probes) on beads, gels, polymeric surfaces, fibers such as fiber optics, glass or any other appropriate substrate, see U.S. Pat. Nos. 5,770,358, 5,789,162, 5,708,153, 6,040,193 and 5,800,992. Arrays may be packaged in such a manner as to allow for diagnostics or other manipulation on the device. See, for example, U.S. Pat. Nos. 5,856,174 and 5,922,591.

[102] In another embodiment, the kit comprises a set of oligonucleotide primers sufficient for the detection and/or quantitation of each of the intrinsic genes listed in Table 1. The oligonucleotide primers may be provided in a lyophilized or reconstituted form, or may be provided as a set of nucleotide sequences. In one embodiment, the primers are provided in a microplate format, where each primer set occupies a well (or multiple wells, as in the case of replicates) in the microplate. The microplate may further comprise primers sufficient for the detection of one or more housekeeping genes as discussed *infra*. The kit may further comprise reagents and instructions sufficient for the amplification of expression products from the genes listed in Table 1.

[103] In order to facilitate ready access, e.g., for comparison, review, recovery, and/or modification, the molecular signatures/expression profiles are typically recorded in a

database. Most typically, the database is a relational database accessible by a computational device, although other formats, *e.g.*, manually accessible indexed files of expression profiles as photographs, analogue or digital imaging readouts, spreadsheets, *etc.* can be used.

Regardless of whether the expression patterns initially recorded are analog or digital in nature, the expression patterns, expression profiles (collective expression patterns), and molecular signatures (correlated expression patterns) are stored digitally and accessed via a database. Typically, the database is compiled and maintained at a central facility, with access being available locally and/or remotely.

[104] *Devices and Tests*

[105] General - The NanoString nCounter Analysis System delivers direct, multiplexed measurements of gene expression through digital readouts of the relative abundance of hundreds of mRNA transcripts. The nCounter Analysis System uses gene-specific probe pairs (**Figure 7**) that are mixed together to form a single reagent called a CodeSet. The probe pairs hybridize directly to the mRNA sample in solution eliminating any enzymatic reactions that might introduce bias in the results.

[106] After hybridization, all of the sample processing steps are automated on the nCounter Prep Station. First, excess capture and reporter probes are removed (**Figure 8**) followed by binding of the probe-target complexes to random locations on the surface of the nCounter cartridge via a streptavidin-biotin linkage (**Figure 9**).

[107] Finally, probe/target complexes are aligned and immobilized (**Figure 10**) in the nCounter Cartridge. The Reporter Probe carries the fluorescent signal; the Capture Probe allows the complex to be immobilized for data collection. Up to 800 pairs of probes, each specific to a particular gene, can be combined with a series of internal controls to form a CodeSet.

[108] After sample processing has completed, cartridges are placed in the nCounter Digital Analyzer for data collection. Each target molecule of interest is identified by the “color code” generated by six ordered fluorescent spots present on the reporter probe. The Reporter Probes on the surface of the cartridge are then counted and tabulated for each target molecule (**Figure 11**).

[109] Reagents and Test Components - The Breast Cancer test will simultaneously measure the expression levels of NANO46 plus eight housekeeping genes in a single hybridization reaction using an nCounter CodeSet designed specifically to those genes. Each assay also includes positive assay controls comprised of a linear titration of *in vitro* transcribed RNA transcripts and corresponding probes, and a set of probes with no sequence homology to

human RNA sequences which are used as negative controls. Each assay run includes a reference sample consisting of *in vitro* transcribed RNA's of the targets and housekeeping genes for normalization purposes. The normalized gene expression profile of a breast tumor sample is correlated to prototypical gene expression profiles of the four breast cancer intrinsic subtypes (Luminal A, Luminal B, HER2-enriched, or Basal-like) that were identified from a training set of breast tumors. The gene expression profile, in combination with selected clinical variables, is used as part of a trained algorithm as a prognostic indicator of risk of distant recurrence of breast cancer.

[110] **Figure 12** outlines the assay processes associated with the nCounter Analysis System Breast Cancer Test.

[111] FFPE Tissue Extraction - The Breast Cancer Test will use RNA extracted from Formalin-fixed, Paraffin-embedded (FFPE) tissue that has been diagnosed as invasive carcinoma of the breast. A pathologist first performs an H & E stain of a tumor section mounted onto a slide to identify the region of viable invasive breast carcinoma containing tumor content above a minimum threshold. The pathologist circles the region on the H & E slide. The pathologist then mounts unstained tissue sections onto slides and marks the area of the slides containing invasive tumor. For larger tumors ($>100\text{mm}^2$ of viable invasive carcinoma on the H&E slide), the test requires only a single $10\mu\text{m}$ section. For smaller tumors ($<100\text{mm}^2$), the test requires 3 sections. The identified region of viable invasive breast carcinoma containing sufficient tumor content on the slides is macro-dissected prior to RNA extraction. Procedures for shipping FFPE tissue slides from the collection site to a testing site will be defined as part of the procedure.

[112] Following extraction of total RNA and removal of genomic DNA, the optical density is measured at wavelengths of 260 nm and 280 nm to determine both yield and purity. The assay procedure requires an input range of 125-500ng of total RNA for the subsequent hybridization step. NanoString plans to validate that this input range of RNA is sufficient to reproducibly perform the assay on the nCounter Analysis System. Additionally, the RNA quality will be measured using an OD 260/280 reading, with a target ratio of no less than 1.7 with an upper limit of 2.5. Procedures for storing RNA will be provided to the user so that downstream processing can be performed at a later point in time if desired.

[113] Requirements for Spectrophotometer to measure yield and purity post RNA extraction - RNA isolations from the FFPE sample result in a final sample volume of $30\mu\text{L}$. This volume is too low for the quantitation of nucleic acid abundance using absorbance

measurements in a cuvette-type UV-Vis spectrophotometer; therefore, NanoString's protocol includes a step for quantitating total RNA using a low volume spectrophotometer such as the NanoDrop™ spectrophotometer. NanoString will define performance specifications for the spectrophotometer so that the range of RNA input recommended for the test is above the limit of detection of the low volume spectrophotometer and is reproducibly measurable.

[114] Hybridization - For each set of up to 10 RNA samples, the user will pipette the specified amount of RNA into separate tubes within a 12 reaction strip tube and add the CodeSet and hybridization buffer. A reference sample is pipetted into the remaining two tubes with CodeSet and hybridization buffer. The CodeSet consists of probes for each gene that is targeted, additional probes for endogenous "housekeeping" normalization genes and positive and negative controls. The probes within the CodeSet pertaining to each of these genes within the four groups (target genes, housekeeping genes, and positive and negative controls) are each assigned a unique code and are therefore individually identifiable within each run. The reference sample consists of *in vitro* transcribed RNA for the targeted genes and housekeeping genes. Once the hybridization reagents are added to the respective tubes, the user transfers the strip tube into a heated-lid heatblock for a specified period of time at a set temperature.

[115] Requirement for Heat block with heated lid for hybridization step - The nCounter assay includes an overnight hybridization under isothermal conditions. Because the overnight hybridization is performed in a small volume at elevated temperature, care must be taken to avoid evaporation. Many commercial PCR thermocyclers are equipped with heated lids that will prevent the evaporation of small volumes of liquid. Because the assay does not require any fine control of temperature ramping, any heat block with a programmable heated lid and a block with dimensions that fit the NanoString tubes will work with the NanoString assay. NanoString plans to provide specifications for heat blocks that meet the assay requirements.

[116] Purification and Binding on the Prep Station - Upon completing hybridization, the user will then transfer the strip tube containing the set of 10 assays and 2 reference samples into the nCounter Prep Station along with the required prepackaged reagents and disposables described in **Table 1**. The Prep Plates contain the necessary reagents for purification of excess probes and binding to the cartridge (see section IIIC below for detailed description of purification process). The prep plates are centrifuged in a swinging bucket centrifuge prior to placement on the deck of the Prep Station. An automated purification process then removes excess capture and reporter probe through two successive hybridization-driven magnetic bead

capture steps. The nCounter Prep Station then transfers the purified target/probe complexes into an nCounter cartridge for capture to a glass slide. Following completion of the run, the user removes the cartridge from the Prep Station and seals it with an adhesive film.

[117] Imaging and Analysis on the Digital Analyzer - The sealed cartridge is then inserted into the nCounter Digital Analyzer which counts the number of probes captured on the slide for each gene, which corresponds to the amount of target in solution. Automated software then checks thresholds for the housekeeping genes, reference sample, and positive and negative controls to qualify each assay and ensure that the procedure was performed correctly. The housekeeping genes provide a measure of RNA integrity, and the thresholds indicate when a tested RNA sample is too degraded to be analyzed by the test due to improper handling or storage of tissue or RNA (e.g. improper tumor fixation, FFPE block storage, RNA storage, RNA handling introducing RNase). The positive and negative assay controls indicate a failure of the assay process (e.g. error in assay setup such as sample mixing with CodeSet, or sample processing such as temperature). The signals of each sample are next normalized using the housekeeping genes to control for input sample quality. The signals are then normalized to the reference sample within each run to control for run-to-run variations. The resulting normalized data is entered in the Breast Cancer Intrinsic Subtyping algorithm to determine tumor intrinsic subtype, risk of relapse score, and risk classification.

[118] Instrumentation- The nCounter Analysis System is comprised of two instruments, the nCounter Prep Station used for post-hybridization processing, and the Digital Analyzer used for data collection and analysis.

[119] nCounter Prep Station - The nCounter Prep Station (**Figure 13**) is an automated fluid handling robot that processes samples post-hybridization to prepare them for data collection on the nCounter Digital Analyzer. Prior to processing on the Prep Station, total RNA extracted from FFPE (Formalin-Fixed, Paraffin-Embedded) tissue samples is hybridized with the NanoString Reporter Probes and Capture Probes according to the nCounter protocol described above.

[120] Hybridization to the target RNA is driven by excess NanoString probes. To accurately analyze these hybridized molecules they are first purified from the remaining excess probes in the hybridization reaction. The Prep Station isolates the hybridized mRNA molecules from the excess Reporter and Capture probes using two sequential magnetic bead purification steps. These affinity purifications utilize custom oligonucleotide-modified magnetic beads that retain only the tripartite complexes of mRNA molecules that are bound to both a Capture probe and a Reporter probe.

[121] Next, this solution of tripartite complexes is washed through a flow cell in the NanoString sample cartridge. One surface of this flow cell is coated with a polyethylene glycol (PEG) hydrogel that is densely impregnated with covalently bound streptavidin. As the solution passes through the flow cell, the tripartite complexes are bound to the streptavidin in the hydrogel through biotin molecules that are incorporated into each Capture probe. The PEG hydrogel acts not only to provide a streptavidin-dense surface onto which the tripartite complexes can be specifically bound, but also inhibits the non-specific binding of any remaining excess reporter probes.

[122] After the complexes are bound to the flow cell surface, an electric field is applied along the length of each sample cartridge flow cell to facilitate the optical identification and order of the fluorescent spots that make up each reporter probe. Because the reporter probes are charged nucleic acids, the applied voltage imparts a force on them that uniformly stretches and orients them along the electric field. While the voltage is applied, the Prep Station adds an immobilization reagent that locks the reporters in the elongated configuration after the field is removed. Once the reporters are immobilized the cartridge can be transferred to the nCounter Digital Analyzer for data collection. All consumable components and reagents required for sample processing on the Prep Station are provided in the nCounter Master Kit. These reagents are ready to load on the deck of the nCounter Prep Station which can process up to 10 samples and 2 reference samples per run in approximately 2.5 hours.

[123] nCounter Digital Analyzer - The nCounter Digital Analyzer (**Figure 14**) collects data by taking images of the immobilized fluorescent reporters in the sample cartridge with a CCD camera through a microscope objective lens. Because the fluorescent Reporter Probes are small, single molecule barcodes with features of smaller than the wavelength of visible light, the Digital Analyzer uses high magnification, diffraction limited imaging to resolve the sequence of the spots in the fluorescent barcodes.

[124] The Digital Analyzer captures hundreds of consecutive fields-of-view (FOV) that can each contain hundreds or thousands of discrete Reporter Probes. Each FOV is a combination of four monochrome images captured at different wavelengths. The resulting overlay can be thought of as a four-color image in blue, green, yellow, and red. Each 4-color FOV is processed in real time to provide a “count” for each fluorescent barcode in the sample. Because each barcode specifically identifies a single mRNA molecule, the resultant data from the Digital Analyzer is a precise measure of the relative abundance of each mRNA of interest in a biological sample.

[125] Software - The Prep Station and the Digital Analyzer are stand-alone units that do not require connection to an external PC, but must be networked to one another using a Local Area Network (LAN). The nCounter System software securely manages operations through user accounts and permissions. Both instruments use setup and process wizards on an embedded touch screen user interface to guide the user through the sample processing and data collection steps of the assay. The user is led through the procedure by step-by-step instructions on the Prep Station and Digital Analyzer. The instrument touch screen uses a pressure sensitive method for controlling operations and enables the user to interact with the system by touching a selection on the screen. Because the touchscreen provides a limited human interface for data entry, the system also hosts a web-based application for user accounts management, sample batch definition, and sample status tracking.

[126] When samples are processed, the system software tracks the user account and reagent lots for each sample in a centralized data repository. After expression data for a sample is acquired by the Digital Analyzer, it is first analyzed to ensure that all pre-specified quality control metrics are met. The qualified data are then processed through a locked PAM50 algorithm to generate a report containing intrinsic subtype and risk of recurrence (ROR) score. The sample report is transferred to the central repository where it can be securely accessed for download by a user with the correct permissions.

[127] The Breast Cancer Intrinsic Subtyping Algorithm - The nCounter system will be used to identify the intrinsic subtype of an excised invasive carcinoma of the breast using a 50 gene classifier algorithm originally named the PAM50 (Parker J.S., et al. Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *Journal of Clinical Oncology*, 27: 1160-1167 (2009)). The gene expression profile will assign a breast cancer to one of four molecular classes or intrinsic subtypes: Basal-like, Luminal A, Luminal B, and HER2 enriched. A brief description of each subtype is provided below.

[128] Luminal subtypes: The most common subtypes of breast cancer are the luminal subtypes in the hormone-receptor positive population, Luminal A and Luminal B. Prior studies suggest that luminal A comprises approximately 30% to 40% and luminal B approximately 20% of breast cancers² and over 90 % of hormone receptor-positive breast cancers. The gene expression pattern of these subtypes resembles the luminal epithelial component of the breast (Nielsen, TO et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor positive breast cancer. *Clinical Cancer Research*, 16:5222-5232 (2010)). These tumors are characterized by high expression of estrogen receptor (ER), progesterone receptor (PR), and

genes associated with ER activation² such as LIV1, GATA3, and cyclin D1, as well as expression of luminal cytokeratins 8 and 18.

[129] Luminal A: Luminal A (LumA) breast cancers exhibit low expression of genes associated with cell cycle activation and the ERBB2 cluster resulting in a better prognosis than luminal B. The Luminal A subgroup has the most favorable prognosis of all subtypes and is enriched for endocrine therapy-responsive tumors.

[130] Luminal B: Luminal B (LumB) breast cancers express ER and ER-associated genes, but to a lower extent than LumA. Genes associated with cell cycle activation are highly expressed and this tumor type can be HER2(+) or HER2(-). The prognosis is unfavorable (despite ER expression) and endocrine therapy responsiveness is generally diminished relative to LumA.

[131] Basal-like: The Basal-like subtype is generally ER-negative, is almost always clinically HER2-negative and expresses a suite of “basal” biomarkers including the basal epithelial cytokeratins (CK) and epidermal growth factor receptor (EGFR). Genes associated with cell cycle activation are highly expressed.

[132] HER2-enriched: The HER2-enriched subtype is generally ER-negative and is HER2-positive in the majority of cases with high expression of the ERBB2 cluster, including ERBB2 and GRB7. Genes associated with cell cycle activation are highly expressed and these tumors have a poor outcome.

[133] Cutoffs for the intrinsic subtyping algorithm are pre-defined from training sets that defined the following: 1) intrinsic subtype centroids (i.e. the prototypical gene expression profile of each subtype), 2) coefficients for Risk of Recurrence (ROR) score, and 3) risk classification (Low/Intermediate/High). The intrinsic subtype centroids (Luminal A, Luminal B, Her2-enriched, Basal-like) were trained using a clinically representative set of archived FFPE breast tumor specimens collected from multiple sites. Hierarchical clustering analysis of gene expression data from the FFPE breast tumor samples was combined with breast tumor biology (i.e. gene expression of previously defined intrinsic subtypes) to define the prototypical expression profile (i.e. centroid) of each subtype. A computational algorithm correlates the normalized 50 gene expression profile of an unknown breast cancer tumor sample to each of the prototypical expression signatures of the four breast cancer intrinsic subtypes. The tumor sample is assigned the subtype with the largest positive correlation to the sample.

[134] 304 unique tumor samples with well-defined clinical characteristics and clinical outcome data were used to establish the ROR score. The ROR score is calculated using

coefficients from a Cox model that includes the Pearson correlation (R) to each intrinsic subtype, a proliferation score (P), and tumor size (T), as shown in the equation below.

$$\text{ROR} = aR_{\text{LumA}} + bR_{\text{LumB}} + cR_{\text{Her2e}} + dR_{\text{basal}} + eP + fT$$

[135] To classify tumor samples into specific risk groups (Low Risk/Intermediate Risk/High Risk) based on their calculated ROR score, cutoffs were set based on probability of recurrence free survival in a patient population consisting of hormone receptor positive, post-menopausal patients treated with endocrine therapy alone.

[136] Anticipated Use of NanoString Breast Cancer Test in Clinical Practice - Oncologists currently use a series of tests to develop a treatment protocol for breast cancer patients. Included in these are the IHC / FISH tests such as ER/PR IHC and HER2 IHC / FISH, and the Agendia MammaPrint® assay and the Genomic Health Oncotype Dx® test. These tests offer the oncologist additional information regarding the patient's prognosis and recommended treatment regimens.

[137] These tests, however, have limitations. ER, PgR, and Her2 testing is done locally by pathologists and reference labs, but the challenges with widespread standardization of IHC and FISH testing is well documented (Lester, J et al. Assessment of Tissue Estrogen and Progesterone Receptor Levels: A Survey of Current Practice, Techniques, and Quantitation Methods. The Breast Journal, 6:189-196 (2000); Wolff, A et al. American Society of Clinical Oncology / College of American Pathologists Guideline Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer. Archives of Pathology and Laboratory Medicine, 131:18-43 (2007)). The MammaPrint test is FDA cleared for use only with frozen or fresh-preserved tissue samples, yet most of the tumor samples collected in the United States are FFPE rather than fresh-frozen. This test is also not distributed and is only available through the Agendia reference labs. The Oncotype Dx test can be used to predict the risk of relapse for stage I/II, node negative, estrogen receptor-positive patients receiving adjuvant Tamoxifen therapy as well as response to cyclophosphamide/methotrexate/5-fluorouracil (CMF) chemotherapy. However this test is only offered as a lab-developed test (LDT) through Genomic Health's CLIA laboratory and is not FDA cleared for prognostic use, or FDA approved for predicting chemotherapy response.

[138] NanoString envisions a model that would have the Breast Cancer test used in conjunction with other sources of clinical data currently available to oncologists for breast cancer prognosis in selected patient segments. The Breast Cancer Test would be an additional source of prognostic information adding significant value to established clinical parameters (i.e tumor size, nodal status) used by oncologists in managing a patient with

breast cancer.

[139] *Methods, Assays and Kits*

[140] The methods, assays and kits of the present invention include a series of quality control metrics that are automatically applied to each sample during analysis. These metrics evaluate the performance of the assay to determine whether the results fall within expected values. Upon successful analysis of these quality control metrics, the Assay gives the following results:

Result	Output Values
The Intrinsic Subtype of the Breast Cancer Specimen	Luminal A Luminal B HER2-Enriched Basal-Like
Individual Estimate of the Probability of Distant Recurrence within 10 years	0-100%
Risk of Recurrence (ROR) Score	Integer value on a 0-100 scale
Risk Category	Low, Intermediate, High

[141] Intrinsic Subtypes

[142] The Intrinsic Subtype of a breast cancer tumor has been shown to be related to prognosis in Early Stage Breast Cancer. On average, patients with a Luminal A tumor have significantly better outcomes than patients with Luminal B, HER2-Enriched, or Basal-like tumors.

[143] The Intrinsic Subtype is identified by comparing the gene expression profile of 50 genes in an unknown sample with the expected expression profiles for the four intrinsic subtypes. The subtype with the most similar profile is assigned to the unknown sample.

[144] The most common subtypes of breast cancer are the luminal subtypes, Luminal A (LumA) and Luminal B (LumB). Prior studies suggest that Luminal A comprises approximately 30% to 40% and Luminal B approximately 20% of breast cancers. However, greater than 90% of hormone-receptor positive patients have luminal tumors. The gene expression pattern of these subtypes resembles the luminal epithelial component of the breast tissue. These tumors are characterized by high expression of estrogen receptor (ER), progesterone receptor (PR), and genes associated with ER activation, such as LIV1, GATA3,

and cyclin D1, as well as expression of luminal cytokeratins 8 and 18. Luminal A breast cancers exhibit lower expression of genes associated with cell cycle activation when compared to Luminal B breast cancers resulting in a better prognosis.

[145] Prior studies suggest that the HER2-Enriched subtype (Her2E) comprises approximately 20% of breast cancers. However, HER2-Enriched tumors are generally ER-negative, so only 5% of the tested ER-positive patient population was found to have HER2-Enriched breast cancer. Regardless of ER-status, HER2-Enriched tumors are HER2-positive in the majority of cases with high expression of the ERBB2 cluster, including ERBB2 and GRB7. Genes associated with cell cycle activation are also highly expressed.

[146] Published data suggest that the Basal-like subtype comprises approximately 20% of breast cancers. However, Basal-like tumors are generally ER-negative, so only 1% of hormone receptor-positive patients have Basal-like breast cancer. The Basal-like subtype is almost always clinically HER2-negative and expresses a suite of “basal” biomarkers including the basal epithelial cytokeratins (CK) and epidermal growth factor receptor (EGFR). Genes associated with cell cycle activation are highly expressed.

[147] ROR Score

[148] The ROR score is an integer value on a 0-100 scale that is related to an individual patient’s probability of distant recurrence within 10 years for the defined intended use population. The ROR score is calculated by comparing the expression profiles of 46 genes in an unknown sample with the expected profiles for the four intrinsic subtypes, as described above, to calculate four different correlation values. These correlation values are then combined with a proliferation score and the tumor size to calculate the ROR score.

[149] Probability of 10-Year Distant Recurrence

[150] The ROR scores for a cohort of post-menopausal women with hormone receptor-positive early stage breast cancer were compared to distant recurrence-free survival following surgery and treatment with 5 years of adjuvant endocrine therapy followed by 5 years of observation. This study resulted in a model relating the ROR score to the probability of distant recurrence in this tested patient population including a 95% confidence interval.

[151] Risk Classification

[152] Risk classification is also provided to allow interpretation of the ROR score by using cutoffs related to clinical outcome in tested patient populations.

[153] **Risk classification by ROR range and nodal status**

Nodal Status	ROR Range	Risk Classification
Node-Negative	0-40	Low
	41-60	Intermediate
	61-100	High
Node-Positive (1-3 nodes)	0-15	Low
	16-40	Intermediate
	41-100	High

[154] Quality Control

[155] Each lot of the Assay components is tested using predetermined specifications. All kit-level items are lot tracked, and the critical components contained within each kit are tested together and released as a Master Lot.

[156] The assay kit includes a series of internal controls that are used to assess the quality of each run set as a whole and each sample individually. These controls are listed below.

[157] Batch Control Set: *In vitro* transcribed RNA Reference Sample

[158] A synthetic RNA Reference Sample is included as a control within the Assay kit. The reference sample is comprised of in-vitro transcribed RNA targets from the 50 algorithm and 8 housekeeping genes. The Reference Sample is processed in duplicate in each assay run along with a set of up to 10 unknown breast tumor RNA samples in a 12 reaction strip tube. The signal from the Reference Sample is analyzed against pre-defined thresholds to qualify the run.

[159] The signal from each of the 50 algorithm genes of the breast tumor RNA sample is normalized to the corresponding genes of the Reference Sample.

[160] Positive Control Set: *in vitro* transcribed RNA targets and corresponding Capture and Reporter Probes

[161] Synthetic RNA targets are used as positive controls (PCs) for the assay. The PC target sequences are derived from the External RNA Control Consortium (ERCC) DNA sequence library. The RNA targets are in-vitro transcribed from DNA plasmids. Six RNA targets are included within the assay kit in a 4-fold titration series (128 – 0.125 fM final concentration in hybridization reaction) along with the corresponding Capture and Reporter Probes. The PCs are added to each breast tumor RNA sample and Reference RNA Sample

tested with the Prosigna Assay. A sample will be disqualified from further analysis if the signal intensities from the PCs do not meet pre-defined thresholds.

[162] Negative control set: exogenous probes without targets

[163] Negative control (NC) target sequences are derived from the ERCC DNA sequence library. The probes designed to detect these target sequences are included as part of the assay kit without the corresponding target sequence. The negative controls (NCs) are added to each breast tumor RNA sample and Reference Sample tested with the Prosigna Assay as a quality control measure. The sample will be disqualified from further analysis if the signal intensities from the NCs do not meet pre-defined thresholds.

[164] RNA Integrity Control Set: Housekeeping genes

[165] Capture and Reporter Probes designed to detect 8 housekeeping genes and 50 algorithm genes are included as part of the kit. The expression levels of the 8 housekeeping genes are analyzed to determine the quality of RNA extracted from the FFPE tissue sample and input into the assay. The sample will be disqualified from further analysis if the expression level of the housekeeping genes falls below pre-defined thresholds.

[166] The housekeeping genes are also used to normalize for any differences in the intact RNA amount in a sample prior to Reference Sample normalization.

[167] *Definitions*

[168] For the purposes of the present disclosure, "breast cancer" includes, for example, those conditions classified by biopsy or histology as malignant pathology. The clinical delineation of breast cancer diagnoses is well known in the medical arts. One of skill in the art will appreciate that breast cancer refers to any malignancy of the breast tissue, including, for example, carcinomas and sarcomas. Particular embodiments of breast cancer include ductal carcinoma *in situ* (DCIS), lobular carcinoma *in situ* (LCIS), or mucinous carcinoma. Breast cancer also refers to infiltrating ductal (IDC) or infiltrating lobular carcinoma (ILC). In most embodiments of the disclosure, the subject of interest is a human patient suspected of or actually diagnosed with breast cancer.

[169] The article "a" and "an" are used herein to refer to one or more than one (*i.e.*, to at least one) of the grammatical object of the article. By way of example, "an element" means one or more element.

[170] Throughout the specification the word "comprising," or variations such as "comprises" or "comprising," will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other

element, integer or step, or group of elements, integers or steps.

EXAMPLES

Example 1. NANO46 Subtyping Test

[171] **Figure 5** outlines the assay processes associated with the Breast Cancer Intrinsic Subtyping test. Following RNA isolation, the test will simultaneously measure the expression levels of 46 target genes plus eight housekeeping genes in a single hybridization reaction using an nCounter CodeSet designed specifically to those genes. For example, the housekeeping genes described in U.S. Patent Publication 2008/0032293 can be used for normalization. Exemplary housekeeping genes include MRPL19, PSMC4, SF3A1, PUM1, ACTB, GAPD, GUSB, RPLP0, and TFRC. The housekeeping genes are used to normalize the expression of the tumor sample. Each assay run also includes a reference sample consisting of *in vitro* transcribed RNA's of the 58 targets for normalization purposes.

[172] FFPE Tissue Review/Procurement and RNA Extraction: The Breast Cancer Intrinsic Subtyping Test will use RNA extracted from Formalin-fixed, Paraffin-embedded (FFPE) tissue that has been diagnosed as invasive carcinoma of the breast. A Pathologist reviews an H & E stained slide to identify the tissue area containing sufficient tumor tissue content for the test. Unstained slide mounted tissue sections are processed by macro-dissecting the identified tumor area on each slide to remove any adjacent normal tissue. RNA is then isolated from the tumor tissue, and DNA is removed from the sample.

[173] Assay Setup and Initiation of Hybridization: For each batch of up to 10 RNA samples isolated from a breast tumor, the user will set up a run using the nCounter Analysis x5 system software, which tracks sample processing, reagent lots, and results for each sample. To initiate the assay, the user will pipette the specified amount of RNA into separate tubes within a 12 reaction strip tube and add the CodeSet and hybridization buffer. A reference sample is pipetted into the remaining two tubes with CodeSet and hybridization buffer. The CodeSet consists of probes for each gene that is targeted, additional probes for endogenous "housekeeping" normalization genes and positive and negative controls that are spiked into the assay. The reference sample consists of *in vitro* transcribed RNA for the targeted genes and housekeeping genes. Once the hybridization reagents are added to the respective tubes, the user transfers the strip tube into a heated-lid heatblock for a specified period of time at a set temperature.

[174] Purification and Binding on the Prep Station: Upon completing hybridization, the user will transfer the strip tube containing the set of 10 assays and 2 reference samples onto the

nCounter Prep Station along with the required prepackaged reagents and disposables. An automated purification process then removes excess capture and reporter probe through two successive hybridization-driven magnetic bead capture steps. The nCounter Prep Station then transfers the purified target/probe complexes into an nCounter cartridge for capture to a glass slide. Following completion of the run, the user removes the cartridge from the Prep Station and seals it with an adhesive film.

[175] Imaging and Analysis on the Digital Analyzer: The cartridge is then sealed and inserted into the nCounter Digital Analyzer which counts the number of probes captured on the slide for each gene, which corresponds to the amount of target in solution. Automated software will then check thresholds for the housekeeping genes, reference sample, and positive and negative controls to qualify each assay and ensure that the procedure was performed correctly. The signals of each sample are next normalized using the housekeeping genes to control for input sample quality. The signals are then normalized to the reference sample within each run to control for run-to-run variations. The resulting normalized data is entered in the Breast Cancer Intrinsic Subtyping algorithm to determine tumor intrinsic subtype and risk of recurrence score.

[176] Example 2: Clinical Validation of the NANO46 risk of recurrence (ROR) score for predicting residual risk of distant-recurrence (DR) after endocrine therapy in postmenopausal women with HR+ early breast cancer (EBC): An ABCSG Study.

[177] The aim of the study is to assess the performance of the ROR score in predicting distal recurrence for postmenopausal patients with hormone receptor positive early breast cancer (HR+ EBC) treated with tamoxifen or tamoxifen followed by anastrozole when the NANO46 test is performed in a routine hospital pathology lab. Does the ROR score add prognostic information (Distant RFS) beyond the Clinical Treatment Score in all patients (CTS includes: nodes, grade, tumor size, age, treatment)? Do the ROR-based risk groups at prognostic information (Distant RFS) beyond the Clinical Treatment Score in all patients?

[178] Study Overview: 3,714 patients were enrolled in a ABCSG8. Patients were postmenopausal women with HR+ EBC (node negative and node positive), grade one or two, with no prior treatment. 1,671 patients re-consented for long-term follow-up or are deceased. The median follow-up was 11 years. 1,620 FFPE blocks were collected. 25 had insufficient cancer in the block on path review, 73 had insufficient RNA included, 44 failed QC specs for the NanoString device. 1,478 patients (91.2%) passed the NANO46 analysis.

[179] Methods: Three unstained 10 micron sections and 1 H&E slide for each patient was sent to an independent academic pathology laboratory at BCCA where tissue review, manual

micro-dissection and RNA extraction were performed. NANO46 analysis was then conducted on 250 ng of the extracted RNA using the NanoString nCounter Analysis System; both intrinsic subtype and ROR score were calculated.

[180] Results: The ROR Score adds statistically significant prognostic information (Distant RFS) beyond CTS in all patients (Likelihood ratio test $\Delta LR\chi^2 = 53.5$, $p < 0.0001$). The ROR-based risk groups add statistically significant prognostic information (Distant RFS) beyond CTS in all patients (Likelihood ratio test $\Delta LR\chi^2 = 34.1$, $p < 0.0001$). Differentiation between Luminal A and Luminal B adds statistically significant prognostic information (Distant RFS) beyond CTS in all patients (Luminal B vs. A: HR=2.38, 95% CI; 1.69-3.35, $p < 0.0001$). Results in the node-negative and node-positive subgroups are similar to the results for all patients that are reported in the study.

[181] Conclusions: The results show that both the ROR score and the ROR-based risk groups add statistically significant prognostic information beyond the Clinical Treatment Score. The results demonstrate that a complex, multi-gene-expression test can be performed in a hospital pathology laboratory and meet the same quality metrics as a central reference laboratory. The results of the TransATAC and ABCSG8 studies together provide Level 1 evidence for the clinical validity of the NANO46 test for predicting the risk of distant recurrence in postmenopausal women with HR+ EBC treated with endocrine therapy alone. The results also show that Luminal A subtypes have better outcomes than Luminal B subtypes in postmenopausal women with HR+ EBC treated with endocrine therapy alone.

What is claimed is:

1. A method of determining low or high risk of recurrence (ROR) in a subject having breast cancer comprising:

determining the size of a breast cancer tumor from the subject;

determining a Pearson correlation value of the breast cancer tumor for each of at least four intrinsic subtypes including Basal-like, Luminal A, Luminal B, or HER-2 enriched, by measuring the expression of at least 40 of the genes in a NANO46 intrinsic gene list of Table 1 in the breast cancer tumor;

determining a proliferation score by measuring the RNA expression of each of the 18-gene subset of proliferation genes in the NANO46 intrinsic gene list consisting of ANLN, CCNE1, CDC20, CDC6, CDCA1, CENPF, CEP55, EXO1, KIF2C, KNTC2, MELK, MK167, ORC6L, PTTG1, RRM2, TYMS, UBE2C, and UBE2T;

calculating a risk of recurrence (ROR) score using the following equation: $ROR-PT = -0.0067 * Basal-like + 0.4317 * HER-2\ enriched + -0.3172 * Luminal\ A + 0.4894 * Luminal\ B + 0.1981 * ProliferationScore + 0.1133 * TumorSize$,

wherein "*Basal-like*", "*HER-2 enriched*", "*Luminal A*" and "*Luminal B*" are the Pearson correlation values of the breast cancer tumor for each of the at least four intrinsic subtypes Basal-like, HER-2 enriched, Luminal A and Luminal B respectively, *ProliferationScore* is the proliferation score, and *TumorSize* is the size of the breast cancer tumor,

thereby determining whether the subject has a low or high risk of recurrence based on the risk of recurrence score.

2. The method of claim 1, further comprising determining at least one of the following: tumor grade, tumor ploidy, nodal status, estrogen receptor expression, progesterone receptor expression, and HER2/ERBB2 expression.

3. The method of claim 1, further comprising determining each of the following: tumor grade, tumor ploidy, nodal status, estrogen receptor expression, progesterone receptor expression, and HER2/ERBB2 expression.

4. The method of claim 1, wherein the risk is breast cancer specific survival, event-free survival or response to therapy.
5. The method of claim 1, wherein the expression of the members of the NANO46 intrinsic gene list is determined using a nanoreporter code system.
6. The method of claim 5, wherein the nanoreporter code system is the nCounter® Analysis system.
7. The method of claim 1, wherein the measuring the expression of at least 40 of the genes in the NANO46 intrinsic gene list comprises detecting gene expression products of the genes.
8. The method of claim 7, wherein the detecting the gene expression products comprises determining a quantity or presence of the gene expression products.
9. The method of claim 7, wherein the gene expression products comprise RNA.
10. The method of claim 9, wherein the RNA comprises mRNA.
11. The method of claim 1, further comprising assigning an intrinsic subtype to the breast cancer tumor.
12. The method of claim 11, wherein the intrinsic subtype is assigned to the breast cancer tumor using an algorithm.
13. The method of claim 12, wherein the algorithm is a supervised algorithm.

Figure 1

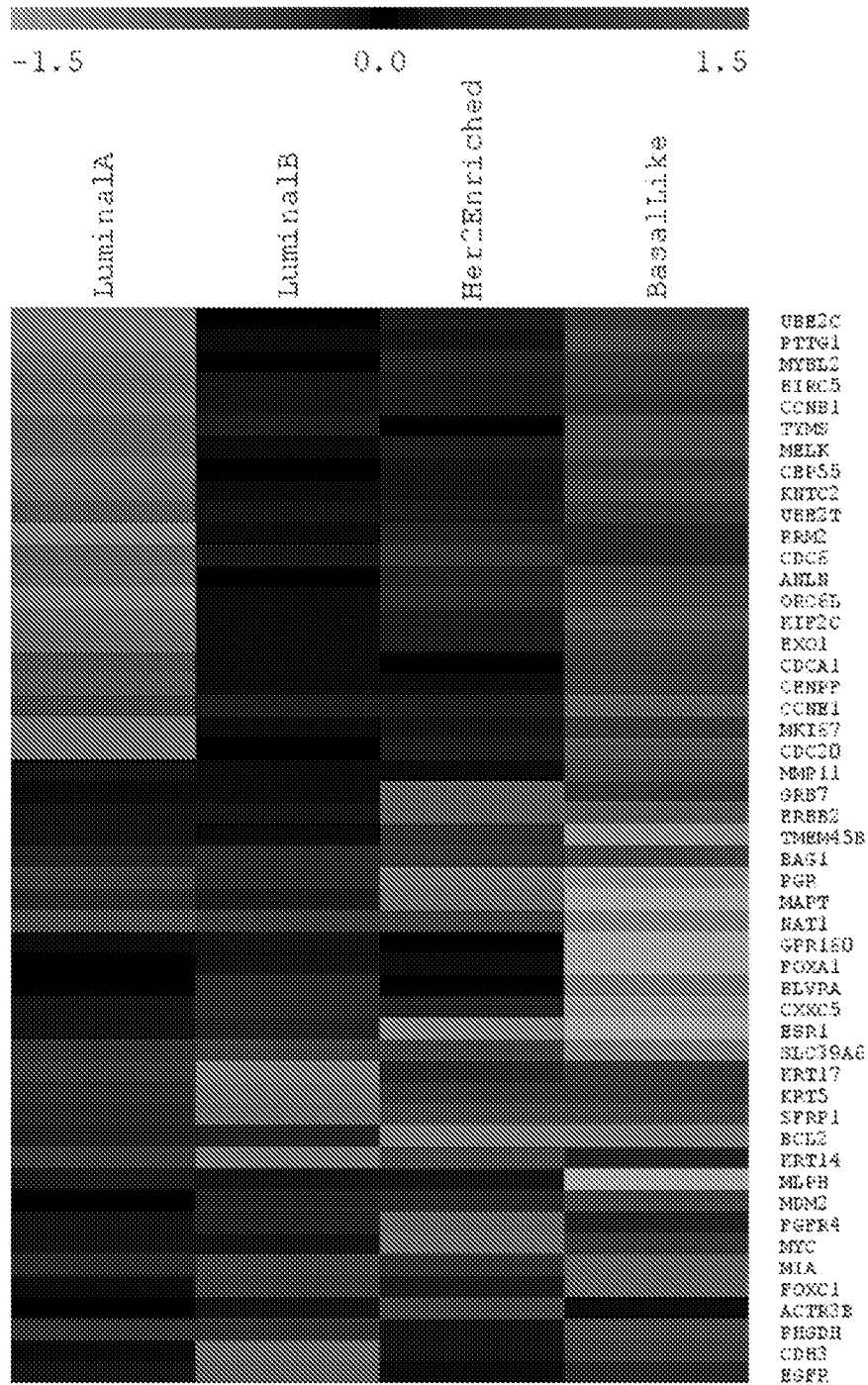


Figure 2

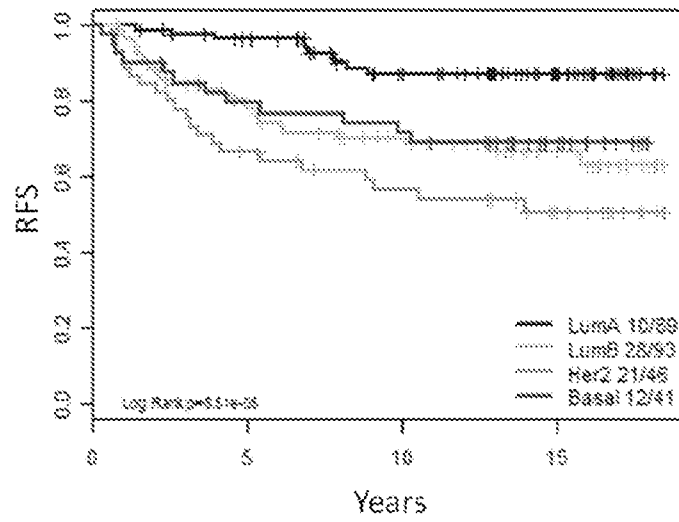


Figure 3

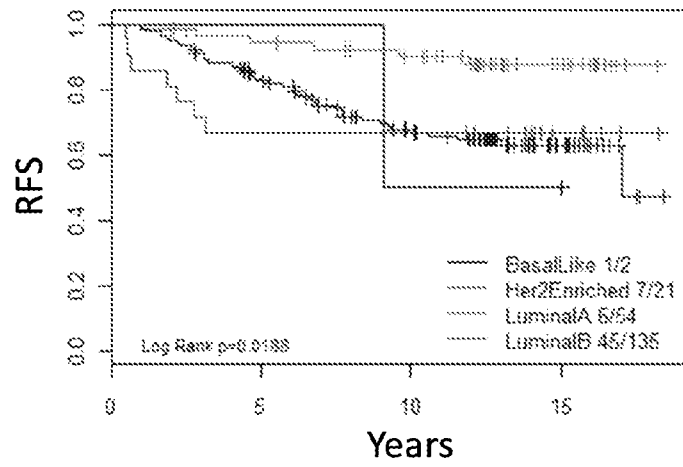


Figure 4

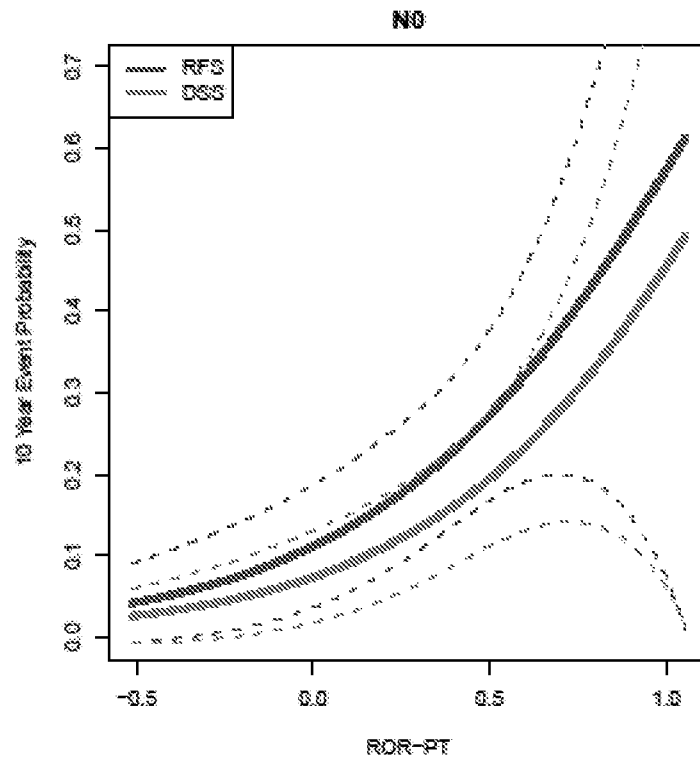


Figure 5

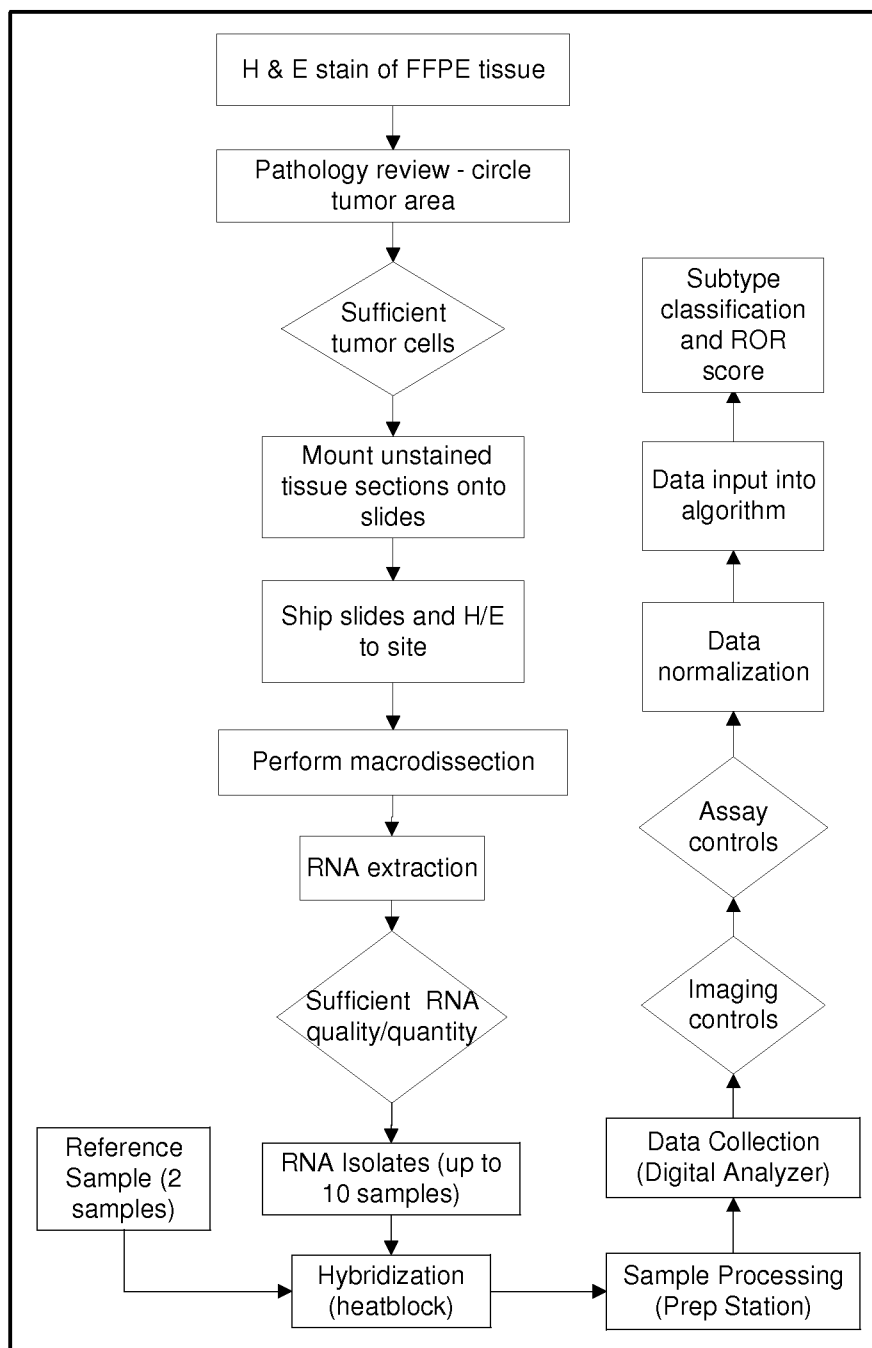


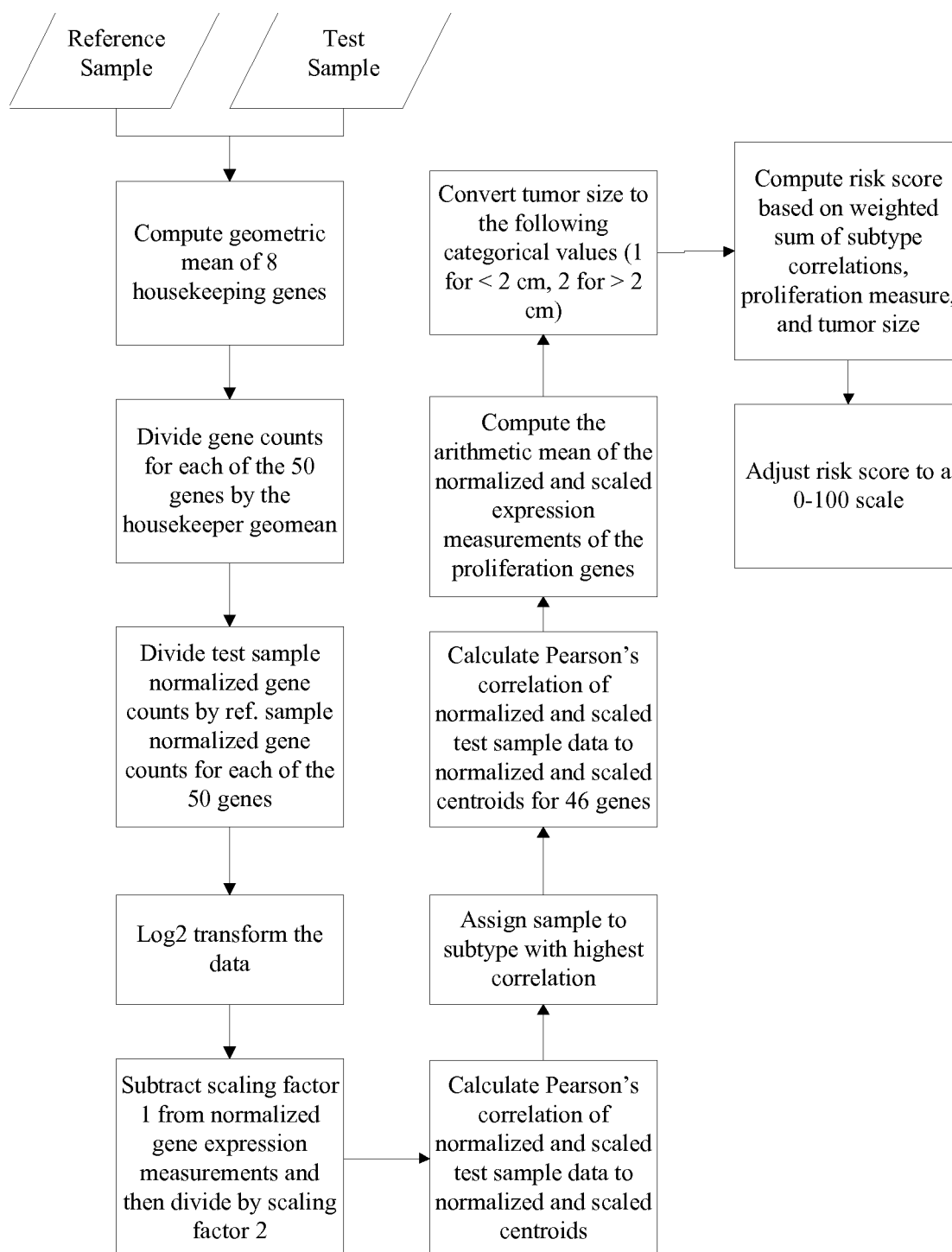
Figure 6

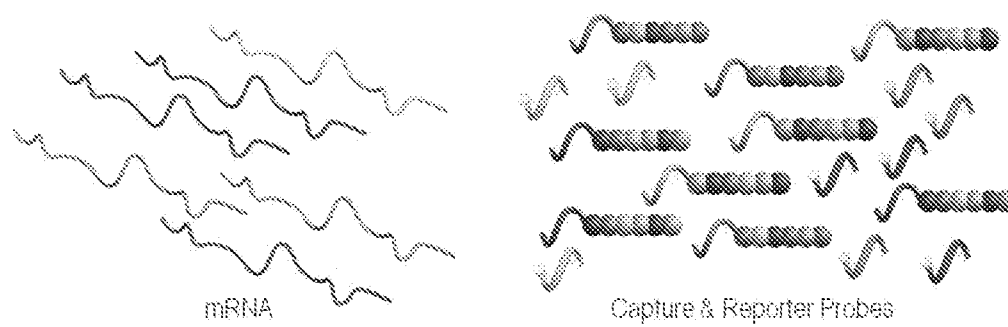
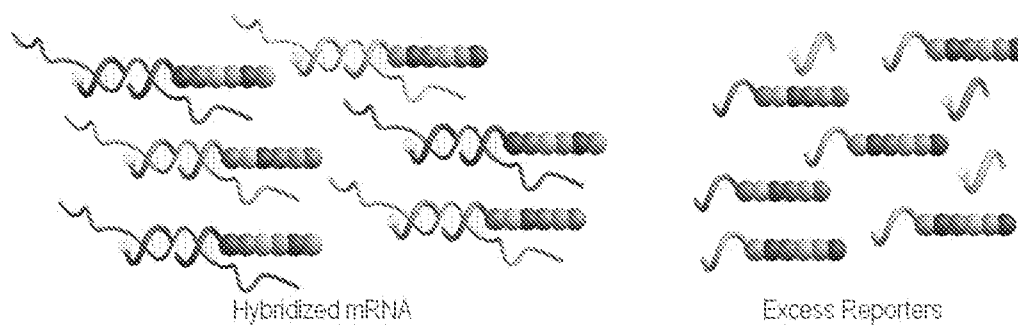
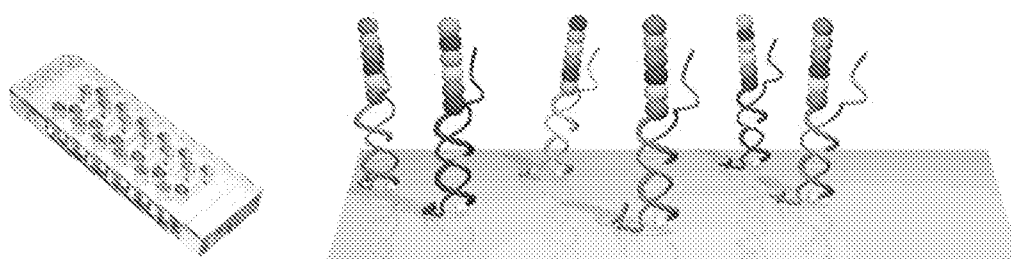
FIGURE 7**FIGURE 8****FIGURE 9**

FIGURE 10

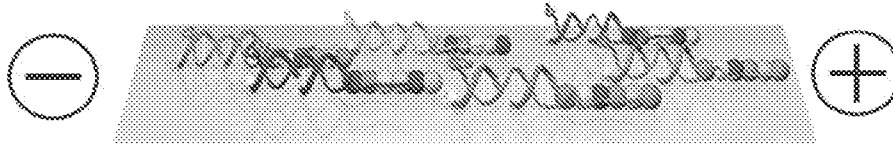


FIGURE 11

Code	Gene	Count
	x	3
	y	1
	z	2

FIGURE 12

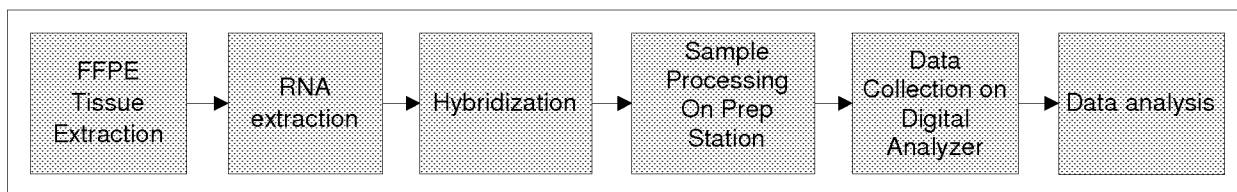


FIGURE 13

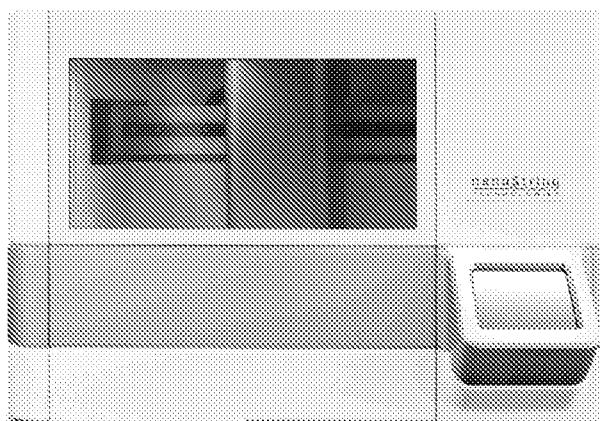
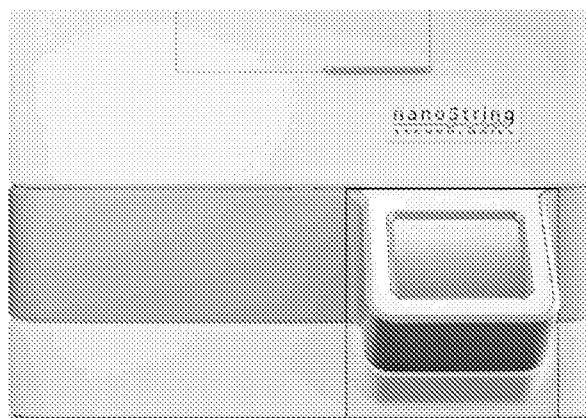


FIGURE 14



N0

