

METHODS FOR COLON CANCER DETECTION AND TREATMENT MONITORING

Cross-Reference to Related Applications

[001] This application claims priority to, and the benefit of, U.S. Provisional Application No. 62/620,015, filed January 22, 2018, the contents of which is incorporated herein by reference in its entirety.

Sequence Listing

[002] The instant application contains a Sequence Listing which has been submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on January 21, 2019, is named "LBIO-004_001WO.txt" and is 111 KB in size.

Field of the Invention

[003] The present invention relates to colon cancer detection.

Background of the Invention

[004] Colorectal cancer (CRC) is one of the most commonly diagnosed cancers worldwide. In the US, CRC is the second leading cause of death as it is in Europe, after lung cancer. Worldwide, it is the fourth most common cause of cancer death. Although surgical resection followed by chemotherapy is the leading treatment option, approximately half eventually die due to distant metastases. Currently, the 5-year overall survival rate of patients with primary CRC can be up to 90%, but it will be reduced to ~50% in patients with advanced non-metastatic tumors, and can be further decreased to <10% in patients in whom the disease is resected at its earliest stages, owing to an incomplete understanding of the molecular mechanisms underpinning its pathogenesis.

[005] Overall survival is associated with the disease stage at the time of diagnosis, suggesting that early detection of disseminated disease is of considerable significance. Consequently, the development of new diagnostic methods that better define disease stage and can better monitor disease progression is critical.

[006] Surveillance remains a cornerstone approach to detect recurrence at an early stage and plan further therapeutic strategies. After potentially curative resection, monitoring can be undertaken through measurement of blood biomarkers and/or imaging like CT to detect asymptomatic

metastatic disease earlier. Pooled data from randomized trials published from 1995 to 2016, however, identifies that a benefit from surgical treatment resulting from earlier detection of metastases, does not occur. This likely reflects the poor sensitivity of current biomarkers.

[007] The current biomarker is carcinoembryonic antigen (CEA), a glycoprotein involved in cell adhesion that is not generally expressed in adult tissues except in heavy smokers. Its specialized sialofucosylated glycoforms serve as functional colon carcinoma L-selectin and E-selectin ligands, which may play a role in metastatic dissemination of colon carcinoma cells. CEA is principally used to monitor colorectal carcinoma treatment, to identify recurrences after surgical resection, for staging or to localize cancer spread through measurement of biological fluids. There are, however, significant limitations. While preoperative CEA levels have shown an association with (disease-free) survival, this was chiefly because it was a surrogate for metastatic presentation. Extrapolating the predictive value of preoperative CEA has, however, been shown to be of limited significance for predictions of long-term outcomes in individual cases. This has been independently supported by a prospective analysis, which identified that levels of CEA, and other biomarkers like CA19-9, does not indicate metastasis even at a time-point where clinical signs and imaging techniques has already demonstrated metastasis.

[008] While the molecular basis for the colorectal cancer disease has been well-characterized e.g., microsatellite instability, K-RAS mutations *etc.*, the development of diagnostic and prognostic markers e.g., in urine or stool or as circulating-free DNA that captures this information, remains nascent but have begun to be developed. Examples include measurements of methylation of septin 9, a tumor suppressor involved in cytokinesis during cellular division. This has been used to detect colon cancer; the metrics range between 60-70%. Assessment of circulating free DNA (Line 1 and Alu-based PCR) has a predictive value of 81% with a ROC of 0.86 as a diagnostic, while measurements of circulating tumor cells are also considered useful. TPS (tissue polypeptide specific antigen) can be used as a monitor of colon cancer as can TAG-72 (tumor-associated glycoprotein) but measurements of other single analytes, like CEA or CA19-9, are non-specific.

Summary of the Invention

[009] Among other things, disclosed herein is a 14-gene expression tool for colon cancer detection.

[010] In one aspect, the present disclosure provides a method for detecting a colon cancer in a subject in need thereof, comprising: (a) determining the expression level of at least 14 biomarkers from a test sample from the subject by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*; (c) inputting each normalized expression level into an algorithm to generate a score; (d) comparing the score with a predetermined cutoff value; and (e) identifying the presence of a colon cancer in the subject when the score is equal to or greater than the predetermined cutoff value or identifying the absence of a colon cancer in the subject when the score is less than the predetermined cutoff value.

[011] In one aspect, the present disclosure provides a method for detecting a colon cancer in a subject in need thereof, the method comprising: (a) determining the expression level of at least 14 biomarkers from a test sample from the subject by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*; (c) inputting each normalized expression level into an algorithm to generate a score; (d) comparing the score with a first predetermined cutoff value; and (e) producing a report, wherein the report identifies the presence of a colon cancer in the subject when the score is equal to or greater than the first predetermined cutoff value or identifies the absence of a colon cancer in the subject when the score is less than the first predetermined cutoff value, wherein the first predetermined cutoff value is 50% on a scale of 0-100%.

[012] In one aspect, the present disclosure provides a method for determining whether a colon cancer in a subject is stable or progressive, the method comprising: (a) determining the expression level of at least 14 biomarkers from a test sample from the subject by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *SI00P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *SI00P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *SI00P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*; (c) inputting each normalized expression level into an algorithm to generate a score; (d) comparing the score with a predetermined cutoff value; and (e) identifying that the colon cancer in the subject is progressive when the score is equal to or greater than the predetermined cutoff value or identifying that the colon cancer in the subject is stable when the score is less than the predetermined cutoff value.

[013] In one aspect, the present disclosure provides a method for determining whether a colon cancer in a subject is stable or progressive, the method comprising: (a) determining the expression level of at least 14 biomarkers from a test sample from the subject by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *SI00P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *SI00P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *SI00P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*; (c) inputting each normalized expression level into an algorithm to generate a score; (d) comparing the score with a second predetermined cutoff value; and (e) producing a report, wherein the report identifies that the colon cancer is progressive when the score is equal to or greater than the second predetermined cutoff value or identifies that the colon cancer is stable when the score is less than the second predetermined cutoff value, wherein the second predetermined cutoff value is 60% on a scale of 0 to 100%.

[014] In one aspect, a method for determining the completeness of surgery in a subject having a colon cancer, the method comprising: (a) determining the expression level of at least 14 biomarkers from a test sample from the subject after the surgery by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*; (c) inputting each normalized expression level into an algorithm to generate a score; (d) comparing the score with a predetermined cutoff value; and (e) identifying that the colon cancer in the subject is not completely removed when the score is equal to or greater than the predetermined cutoff value or identifying that the colon cancer in the subject is completely removed when the score is less than the predetermined cutoff value.

[015] In one aspect, the present disclosure provides a method for determining the completeness of surgery in a subject having a colon cancer, the method comprising: (a) determining the expression level of at least 14 biomarkers from a test sample from the subject after the surgery by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*; (c) inputting each normalized expression level into an algorithm to generate a score; (d) comparing the score with a first predetermined cutoff value; and (e) producing a report, wherein the report identifies that the colon cancer is not completely removed when the score is equal to or greater than the first predetermined cutoff value or identifies that the colon cancer is completely removed when the score is less than the first predetermined cutoff value, wherein the first predetermined cutoff value is 50% on a scale of 0-100%.

[016] In one aspect, the present disclosure provides a method comprising: (a) determining the expression level of at least 14 biomarkers from a test sample from a subject by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*; (c) inputting each normalized expression level into an algorithm to generate a score; (d) comparing the score with a predetermined cutoff value; and (e) administering a first therapy to the subject when the score is equal to or greater than the predetermined cutoff value.

[017] In one aspect, the present disclosure provides a method for evaluating the response of a subject having a colon cancer to a first therapy, the method comprising: (1) at a first time point: (a) determining the expression level of at least 14 biomarkers from a first test sample from the subject by contacting the first test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*; (c) inputting each normalized expression level into an algorithm to generate a first score; (2) at a second time point, wherein the second time point is after the first time point and after the administration of the therapy to the subject: (a) determining the expression level of at least 14 biomarkers from a second test sample from the subject by contacting the second test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and the housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*,

SI00P, SNRPA, SORD, STOML2, and UMPS to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, SI00P, SNRPA, SORD, STOML2, and UMPS*; (c) inputting each normalized expression level into the algorithm to generate a second score; (3) comparing the first score with the second score; and (4) identifying that the subject is responsive to the first therapy when the second score is significantly decreased as compared to the first score or identifying that the subject is not responsive to the first therapy when the second score is not significantly decreased as compared to the first score.

[018] In one aspect, the present disclosure provides a method for evaluating the response of a subject having a colon cancer to a therapy, the method comprising: (1) at a first time point, performing the following steps that include (a) determining the expression level of at least 14 biomarkers from a first test sample from the subject by contacting the first test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, SI00P, SNRPA, SORD, STOML2, UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, SI00P, SNRPA, SORD, STOML2, and UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, SI00P, SNRPA, SORD, STOML2, and UMPS*; (c) inputting each normalized expression level into an algorithm to generate a first score; and (2) at a second time point, performing the following steps that include (d) determining the expression level of at least 14 biomarkers from a second test sample from the subject by contacting the second test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, SI00P, SNRPA, SORD, STOML2, UMPS*, and the housekeeping gene; (e) normalizing the expression level of each of *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, SI00P, SNRPA, SORD, STOML2, and UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, SI00P, SNRPA, SORD, STOML2, and UMPS*; (f) inputting each normalized expression level into the algorithm to generate a second score, wherein the second time point is

after the first time point and after the administration of the therapy to the subject; (3) comparing the first score with the second score; and (4) producing a report, wherein the report identifies that the subject is responsive to the therapy when the second score is significantly decreased as compared to the first score or identifies that the subject is not responsive to the therapy when the second score is not significantly decreased as compared to the first score.

[019] In some aspects, a method of the present disclosure can further comprise continuing to administer a first therapy to a subject when a second score is significantly decreased as compared to a first score.

[020] In some aspects, a method of the present disclosure can further comprise discontinuing administration of a first therapy to a subject when a second score is not significantly decreased as compared to a first score.

[021] In some aspects, a method of the present disclosure can further comprise administering a second therapy to a subject when a second score is not significantly decreased as compared to a first score.

[022] In some aspects, a second score is significantly decreased as compared to a first score when the second score is at least 25% less than the first score.

[023] In some aspects, a predetermined cutoff value can be 50% on a scale of 0-100%. A predetermined cutoff value can be 60% on a scale of 0-100%.

[024] In some aspects of any one of the methods disclosed herein, a housekeeping gene can be selected from the group consisting of *MRPL19*, *PSMC4*, *SF3A1*, *PUM1*, *ACTB*, *GAPD*, *GUSB*, *RPLP0*, *TFRC*, *MORF4L1*, *18S*, *PPIA*, *PGK1*, *RPL13A*, *B2M*, *YWHAZ*, *SDHA*, and *HPRT1*. For example, the housekeeping gene can be *MORF4L1*.

[025] In some aspects, a method of the present disclosure can have a sensitivity greater than 85%. In some aspects, a method of the present disclosure can have a specificity of greater than 85%.

[026] In some aspects, a biomarker can comprise RNA, cDNA, protein or any combination thereof.

[027] In some aspects, wherein when the biomarker is RNA, the RNA can be reverse transcribed to produce cDNA, and the produced cDNA expression level can be detected.

[028] In some aspects, a biomarker or the expression of a biomarker can be detected by forming a complex between the biomarker and a labeled probe or primer.

[029] In some aspects, when a biomarker is protein, the protein can be detected by forming a complex between the protein and a labeled antibody.

[030] In some aspects, when a biomarker is RNA or cDNA, the RNA or cDNA can be detected by forming a complex between the RNA or cDNA and a labeled nucleic acid probe or primer. A complex between the RNA or cDNA and the labeled nucleic acid probe or primer can be a hybridization complex.

[031] In some aspects, a predetermined cutoff value can be derived from a plurality of reference samples obtained from subjects not having or not diagnosed with a neoplastic disease. The neoplastic disease can be colon cancer.

[032] In some aspects, an algorithm can be XGBoost (XGB), Random Forest (RF), glmnet, cforest, Classification and Regression Trees for Machine Learning (CART), treebag, K-Nearest Neighbors (kNN), neural network (nnet), Support Vector Machine radial (SVM-radial), Support Vector Machine linear (SVM-linear), Naïve Bayes (NB), multilayer perceptron (mlp) or any combination thereof.

[033] In some aspects, the methods of the present disclosure can further comprise administering to a subject a first therapy when a score is equal to or greater than a predetermined cutoff.

[034] In some aspects, a first time point can be prior to the administration of a first therapy to the subject. A first time point can be after the administration of the first therapy to the subject.

[035] In some aspects, a therapy can comprise anti-cancer therapy, surgery, chemotherapy, targeted drug therapy, radiation therapy, immunotherapy or any combination thereof.

[036] In some aspects, surgery can comprise removing a polyp during a colonoscopy, endoscopic mucosal resection, a partial colectomy, an ostomy, removing at least one cancerous lesion from the liver, or any combination thereof.

[037] In some aspects, chemotherapy can comprise FOLFOX, FOLFIRI, a combination of 5-FU and leucovorin, capecitabine, irinotecan, CapeOx or any combination thereof.

[038] In some aspects, targeted drug therapy can comprise bevacizumab, cetuximab, panitumumab, regorafenib, a combination of trifluridine and tipiracil, a EGFR TKI inhibitor or any combination thereof.

[039] In some aspects, anti-cancer therapy can comprise anti-colon cancer therapy.

[040] In some aspects, immunotherapy can comprise pembrolizumab, nivolumab or a combination of pembrolizumab and nivolumab.

[041] In some aspects, a test sample can be blood, serum, plasma, neoplastic tissue or any combination thereof. A reference sample can be blood, serum, plasma, non-neoplastic tissue or any combination thereof.

[042] Any of the above aspects can be combined with any other aspect.

[043] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. In the Specification, the singular forms also include the plural unless the context clearly dictates otherwise; as examples, the terms “a,” “an,” and “the” are understood to be singular or plural and the term “or” is understood to be inclusive. By way of example, “an element” means one or more element. 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. About can be understood as within 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, or 0.01% of the stated value. Unless otherwise clear from the context, all numerical values provided herein are modified by the term “about.”

[044] Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. The references cited herein are not admitted to be prior art to the claimed invention. In the case of conflict, the present Specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and are not intended to be limiting. Other features and advantages of the disclosure will be apparent from the following detailed description and claim.

Brief Description of the Drawings

[045] FIGs. 1A-1B are graphs showing normalized gene expression of the 13 gene signature in colon mucosa (FIG. 1A) and cell lines (FIG. 1B). Gene expression was significantly ($p < 0.0001$) elevated in colon cancer samples ($n=7$) compared to matched normal mucosa ($n=7$). Levels were ~20-fold elevated in colon cancer tumor tissue than in normal mucosa. All genes were expressed in three different colon cancer cell lines. Levels were ~1000x elevated compared to normal mucosa. Horizontal lines identify median normalized expression of the 13 genes.

[046] FIGs. 2A-2B are graphs showing receiver operator curve analysis of the test set (FIG. 2A) and independent set (FIG. 2B). Each cohort included 136 cancers and 60 controls. The AUROC in the test set was 0.9 and the Youden J index was 0.71. In the independent set the AUROC was 0.86 with a Youden index of 0.6. Z-statistics ranged 11.2-15.6 and were highly significant ($p<0.0001$). The sensitivity and specificity of the test ranged 85-87.5% and 75-83%, respectively.

[047] FIGs. 3A-3B are graphs showing that gene expression in the entire cohort (controls: $n=120$; colon cancer cases: $n=272$) identified levels were significantly ($p<0.0001$) elevated in cases ($62.7\pm14\%$) versus controls ($34.6\pm18\%$) (FIG. 3A). The AUROC was 0.88, $p<0.0001$ (FIG. 3B). Horizontal lines identify median expression of the normalized 13 gene signature (ColoTest).

[048] FIGs. 4A-4B are graphs showing decision curve analysis (FIG. 4A) and risk analysis (FIG. 4B) for the ColoTest. This exhibited $>50\%$ standardized predictive benefit up to a risk threshold of 80%. The probit risk assessment plot identified a ColoTest score $>50\%$ was 75% accurate for predicting colon cancer in a blood sample. This was increased to $>80\%$ at a ColoTest score $>60\%$.

[049] FIG. 5 is a graph showing the effect of surgery on the ColoTest. Levels prior to surgery are elevated ($84\pm6\%$). In those with no evidence of disease (NED) levels were reduced by surgery to $14\pm9\%$ ($*p=0.0001$). In those with disease (D) remaining after surgery, levels remained similar to pre-surgical values ($74\pm4\%$).

[050] FIGs. 6A-6C are graphs showing ColoTest scores in stable and progressive disease. Test scores were not significantly different between those identified as stable and those with progressive disease at the time of assessment (FIG. 6A). Of the 17 with stable disease, 12 exhibited disease progression in the 3 month follow-up. Levels in those who truly had demonstrable stable disease were low ($16\pm10\%$) (FIG. 6B). In those who did progress in the 3 months levels were not different to those that had progressive disease ($73\pm16\%$ vs. $68\pm25\%$). The AUROC for differentiating stable from progressing/progressive disease was 0.97, $p<0.0001$ (FIG. 6C).

[051] FIG. 7 is a graph showing comparison of AUROC between the ColoTest and CEA for differentiating stable from progressive disease. The ColoTest was significantly more sensitive than CEA (difference in AUC: 0.18, z-statistic: 2.1, $p=0.03$).

[052] FIG. 8 is a graph showing the effect of treatment on the ColoTest. Levels prior to treatment are elevated ($82\pm9\%$). In those who responded to therapy with disease stabilization, levels were reduced to $14\pm7\%$ ($*p<0.0001$). In those that exhibited disease progression because of treatment failure, levels were elevated ($69\pm21\%$).

Detailed Description of the Invention

[053] The details of the invention are set forth in the accompanying description below.

Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, illustrative methods and materials are now described. Other features, objects, and advantages of the invention will be apparent from the description and from the claims. In the specification and the appended claims, the singular forms also include the plural unless the context clearly dictates otherwise. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. All patents and publications cited in this specification are incorporated herein by reference in their entireties.

[054] Colon cancer is cancer of the large intestine (colon). Symptoms of colon cancer include, but are not limited to: (a) a change in bowel habits, (b) rectal bleeding or blood in the stool, (c) persistent abdominal discomfort, such as cramps, gas or pain, (d) a feeling that the bowel doesn't empty completely, (e) weakness or fatigue, and (f) unexplained weight loss.

[055] Described herein are methods to quantitate (score) the circulating colon cancer molecular signature with high sensitivity and specificity for purposes including, but not limited to, detecting a colon cancer, determining whether a colon cancer is stable or progressive, determining the completeness of surgery, and evaluating the response to a colon cancer therapy. Specifically, the present invention is based on the discovery that the expression levels of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*, normalized by the expression level of a housekeeping gene such as *MORF4LI*, are elevated in subjects having colon cancers as compared to healthy subjects.

[056] Accordingly, the present disclosure provides a method for detecting a colon cancer in a subject in need thereof, the method comprising: (a) determining the expression level of at least 14 biomarkers from a test sample from the subject by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*,

PDXK, POP7, S100P, SNRPA, SORD, STOML2, and UMPS; (c) inputting each normalized expression level into an algorithm to generate a score; (d) comparing the score with a first predetermined cutoff value; and (e) identifying the presence of a colon cancer in the subject when the score is equal to or greater than the predetermined cutoff value or identifying the absence of a colon cancer in the subject when the score is less than the predetermined cutoff value.

[057] In some aspects of the preceding method, step (e) can comprise producing a report, wherein the report identifies the presence of a colon cancer in the subject when the score is equal to or greater than the first predetermined cutoff value or identifies the absence of a colon cancer in the subject when the score is less than the first predetermined cutoff value.

[058] In some aspects, the preceding method can further comprise administering to the subject a first therapy. The preceding method can further comprise administering to the subject a first therapy when the score is equal to or greater than the predetermined cutoff.

[059] The present disclosure also provides a method for determining whether a colon cancer in a subject is stable or progressive, the method comprising: (a) determining the expression level of at least 14 biomarkers from a test sample from the subject by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, S100P, SNRPA, SORD, STOML2, UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, S100P, SNRPA, SORD, STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, S100P, SNRPA, SORD, STOML2*, and *UMPS*; (c) inputting each normalized expression level into an algorithm to generate a score; (d) comparing the score with a second predetermined cutoff value; and (e) identifying that the colon cancer in the subject is progressive when the score is equal to or greater than the predetermined cutoff value or identifying that the colon cancer in the subject is stable when the score is less than the predetermined cutoff value.

[060] In some aspects of the preceding method, step (e) can comprise producing a report, wherein the report identifies that the colon cancer is progressive when the score is equal to or

greater than the second predetermined cutoff value or identifies that the colon cancer is stable when the score is less than the second predetermined cutoff value.

[061] In some aspects, the preceding method can further comprise administering to the subject a first therapy. The preceding method can further comprise administering to the subject a first therapy when the score is equal to or greater than the predetermined cutoff.

[062] In some aspects, the method further comprises treating the subject with a progressive colon cancer with surgery, chemotherapy, targeted drug therapy, radiation therapy, immunotherapy, or a combination thereof.

[063] The present disclosure also provides a method for determining the completeness of surgery in a subject having a colon cancer, the method comprising: (a) determining the expression level of at least 14 biomarkers from a test sample from the subject after the surgery by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*; (c) inputting each normalized expression level into an algorithm to generate a score; (d) comparing the score with a first predetermined cutoff value; and (e) identifying that the colon cancer in the subject is not completely removed when the score is equal to or greater than the predetermined cutoff value or identifying that the colon cancer in the subject is completely removed when the score is less than the predetermined cutoff value.

[064] In some aspects of the preceding method, step (e) can comprise producing a report, wherein the report identifies that the colon cancer is not completely removed when the score is equal to or greater than the first predetermined cutoff value or identifies that the colon cancer is completely removed when the score is less than the first predetermined cutoff value.

[065] In some aspects, the preceding method can further comprise administering to the subject a first therapy. The preceding method can further comprise administering to the subject a first therapy when the score is equal to or greater than the predetermined cutoff.

[066] The present disclosure also provides a method comprising: (a) determining the expression level of at least 14 biomarkers from a test sample from a subject by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*; (c) inputting each normalized expression level into an algorithm to generate a score; (d) comparing the score with a predetermined cutoff value; and (e) administering a first therapy to the subject when the score is equal to or greater than the predetermined cutoff value.

[067] The response of a subject having a colon cancer to a therapy can also be evaluated by comparing the scores determined by the same algorithm at different time points of the therapy. For example, the first time point can be prior to or after the administration of the therapy to the subject; the second time point is after the first time point and after the administration of the therapy to the subject. A first score is generated at the first time point, and a second score is generated at the second time point. When the second score is significantly decreased as compared to the first score, the subject is considered to be responsive to the therapy.

[068] Accordingly, the present disclosure provides a method for evaluating the response of a subject having a colon cancer to a first therapy, the method comprising: (1) at a first time point: (a) determining the expression level of at least 14 biomarkers from a first test sample from the subject by contacting the first test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*; (c) inputting each normalized expression level into an algorithm to generate a first score; (2) at a second time point, wherein the second time point is after the first

time point and after the administration of the therapy to the subject: (a) determining the expression level of at least 14 biomarkers from a second test sample from the subject by contacting the second test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and the housekeeping gene; (b) normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*; (c) inputting each normalized expression level into the algorithm to generate a second score; (3) comparing the first score with the second score; and (4) identifying that the subject is responsive to the first therapy when the second score is significantly decreased as compared to the first score or identifying that the subject is not responsive to the first therapy when the second score is not significantly decreased as compared to the first score.

[069] In some aspects of the preceding method, step (4) can comprise producing a report, wherein the report identifies that the subject is responsive to the first therapy when the second score is significantly decreased as compared to the first score or identifies that the subject is not responsive to the first therapy when the second score is not significantly decreased as compared to the first score.

[070] In some aspects of the preceding method, the second score is significantly decreased as compared to the first score when the second score is at least about 10% less than the first score, or at least about 20% less than the first score, or at least about 25% less than the first score, at least about 40% less than the first score, at least about 50% less than the first score, or at least about 60% less than the first score, or at least about 70% less than the first score, or at least about 75% less than the first score, or at least about 80% less than the first score, or at least about 90% less than the first score, or at least about 95% less than the first score or at least about 95% less than the first score. In some aspects, when the second score is not significantly decreased as compared to the first score, the subject is considered to be not responsive to the therapy.

[071] In some aspects of the preceding method, a first time point can be prior to the administration of a first therapy to the subject. A first time point can be after the administration of a first therapy to the subject.

[072] In some aspects, the preceding method can further comprise continuing to administer the first therapy to the subject when the second score is significantly decreased as compared to the first score.

[073] In some aspects, the preceding method can further comprise discontinuing administration of the first therapy to the subject when the second score is not significantly decreased as compared to the first score.

[074] In some aspects, the preceding method can further comprise administering a second therapy to the subject when the second score is not significantly decreased as compared to the first score.

[075] In some aspects of the methods of the present disclosure, a predetermined cutoff value can be about 50% on a scale of 0-100%. A predetermined cutoff value can be about 60% on a scale of 0-100%. A predetermined cutoff value can be about 10%, or about 20%, or about 30%, or about 40%, or about 70%, or about 80%, or about 90% on a scale of 0-100%.

[076] In some aspects of the methods of the present disclosure, a test sample can be any biological fluid obtained from the subject. A test sample can be blood, serum, plasma, neoplastic tissue or any combination thereof. In some aspects, the test sample is blood. In some aspects, the test sample is serum. In some aspects, the test sample is plasma.

[077] In some aspects of the methods of the present disclosure, a housekeeping gene can comprise, but is not limited to, *MRPL19*, *PSMC4*, *SF3A1*, *PUM1*, *ACTB*, *GAPD*, *GUSB*, *RPLP0*, *TFRC*, *MORF4L1*, *18S*, *PPIA*, *PGK1*, *RPL13A*, *B2M*, *YWHAZ*, *SDHA*, and *HPRT1*. In some aspects, the housekeeping gene is *MORF4L1*.

[078] The methods of the present disclosure can have a sensitivity of at least about 50%, or at least about 60%, or at least about 70%, or at least about 75%, or at least about 80%, or at least about 85%, or at least about 90%, or at least about 95%, or at least about 99%. The methods of the present disclosure can have a sensitivity of greater than about 50%, or greater than about 60%, or greater than about 70%, or greater than about 75%, or greater than about 80%, or greater than about 85%, or greater than about 90%, or greater than about 95%, or greater than about 99%.

[079] The methods of the present disclosure can have a specificity of at least about 50%, or at least about 60%, or at least about 70%, or at least about 75%, or at least about 80%, or at least about 85%, or at least about 90%, or at least about 95%, or at least about 99%. The methods of the present disclosure can have a specificity of greater than about 50%, or greater than about 60%, or greater than about 70%, or greater than about 75%, or greater than about 80%, or greater than about 85%, or greater than about 90%, or greater than about 95%, or greater than about 99%.

[080] The methods of the present disclosure can have an accuracy of at least about 50%, or at least about 60%, or at least about 70%, or at least about 75%, or at least about 80%, or at least about 85%, or at least about 90%, or at least about 95%, or at least about 99%. The methods of the present disclosure can have an accuracy of greater than about 50%, or greater than about 60%, or greater than about 70%, or greater than about 75%, or greater than about 80%, or greater than about 85%, or greater than about 90%, or greater than about 95%, or greater than about 99%.

[081] In some aspects of the methods of the present disclosure, a predetermined cutoff value can be derived from a plurality of reference samples obtained from subjects not having or not diagnosed with a neoplastic disease. In some aspects, the neoplastic disease can be colon cancer.

[082] The plurality of reference samples can comprise about 2-500, 2-200, 10-100, or 20-80 reference samples. Each reference sample produces a score using the algorithm, and the first predetermined cutoff value can be an arithmetic mean of these scores. Each reference sample can be blood, serum, plasma, or non-neoplastic tissue. In some aspects, each reference sample is blood. In some aspects, each reference sample is of the same type as the test sample.

[083] Each of the biomarkers disclosed herein may have one or more transcript variants. The methods disclosed herein can measure the expression level of any one of the transcript variants for each biomarker.

[084] The expression level can be measured in a number of ways, including, but not limited to: measuring the mRNA encoded by the selected genes; measuring the amount of protein encoded by the selected genes; and measuring the activity of the protein encoded by the selected genes.

[085] In some aspects of the methods of the present disclosure, a biomarker can be RNA, cDNA, protein or any combination thereof. When the biomarker is RNA, the RNA can be reverse transcribed to produce cDNA (such as by RT-PCR), and the produced cDNA expression

level can be detected. The expression level of a biomarker can be detected by forming a complex between the biomarker and a labeled probe or primer. When the biomarker is RNA or cDNA, the RNA or cDNA can be detected by forming a complex between the RNA or cDNA and a labeled nucleic acid probe or primer. The complex between the RNA or cDNA and the labeled nucleic acid probe or primer can be a hybridization complex.

[086] In some aspects of the methods of the present disclosure, gene expression can be detected by microarray analysis. Differential gene expression can also be identified, or confirmed using the microarray technique. Thus, the expression profile biomarkers can be measured in either fresh or fixed tissue, using microarray technology. In this method, polynucleotide sequences of interest (including cDNAs and oligonucleotides) are plated, or arrayed, on a microchip substrate. The arrayed sequences are then hybridized with specific DNA probes from cells or tissues of interest. The source of mRNA typically is total RNA isolated from a biological sample, and corresponding normal tissues or cell lines may be used to determine differential expression.

[087] In some aspects of microarray techniques, PCR amplified inserts of cDNA clones are applied to a substrate in a dense array. In some aspects, at least 10,000 nucleotide sequences are applied to the substrate. The microarrayed genes, immobilized on the microchip at 10,000 elements each, are suitable for hybridization under stringent conditions. Fluorescently labeled cDNA probes may be generated through incorporation of fluorescent nucleotides by reverse transcription of RNA extracted from tissues of interest. Labeled cDNA probes applied to the chip hybridize with specificity to each spot of DNA on the array. After stringent washing to remove non-specifically bound probes, the microarray chip is scanned by a device such as, confocal laser microscopy or by another detection method, such as a CCD camera. Quantitation of hybridization of each arrayed element allows for assessment of corresponding mRNA abundance. With dual color fluorescence, separately labeled cDNA probes generated from two sources of RNA are hybridized pair-wise to the array. The relative abundance of the transcripts from the two sources corresponding to each specified gene is thus determined simultaneously. Microarray analysis can be performed by commercially available equipment, following manufacturer's protocols.

[088] In some aspects of the methods of the present disclosure, the biomarkers can be detected in a biological sample using qRT-PCR. The first step in gene expression profiling by RT-PCR is extracting RNA from a biological sample followed by the reverse transcription of the RNA

template into cDNA and amplification by a PCR reaction. The reverse transcription reaction step is generally primed using specific primers, random hexamers, or oligo-dT primers, depending on the goal of expression profiling. The two commonly used reverse transcriptases are avian myeloblastosis virus reverse transcriptase (AMV-RT) and Moloney murine leukemia virus reverse transcriptase (MLV-RT).

[089] In some aspects of the methods of the present disclosure, when the biomarker is a protein, the protein can be detected by forming a complex between the protein and a labeled antibody. The label can be any label for example a fluorescent label, chemiluminescence label, radioactive label, etc. Exemplary methods for protein detection include, but are not limited to, enzyme immunoassay (EIA), radioimmunoassay (RIA), Western blot analysis and enzyme linked immunoabsorbant assay (ELISA). For example, the biomarker can be detected in an ELISA, in which the biomarker antibody is bound to a solid phase and an enzyme-antibody conjugate is employed to detect and/or quantify biomarker present in a sample. Alternatively, a western blot assay can be used in which solubilized and separated biomarker is bound to nitrocellulose paper. The combination of a highly specific, stable liquid conjugate with a sensitive chromogenic substrate allows rapid and accurate identification of samples.

[090] In some aspects of the methods of the present disclosure, the methods described herein can have a specificity, sensitivity, and/or accuracy of at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%.

[091] In some aspects of the methods of the present disclosure, a labeled probe, a labeled primer, a labeled antibody or a labeled nucleic acid can comprise a fluorescent label.

[092] Any algorithm that can generate a score for a sample by assessing where that sample value falls onto a prediction model generated using different techniques, e.g., decision trees, can be used in the methods disclosed herein. The algorithm analyzes the data (i.e., expression levels) and then assigns a score. In some aspects, the algorithm can be a machine-learning algorithm. Exemplary algorithms that can be used in the methods disclosed herein can include, but are not limited to, XGBoost (XGB), Random Forest (RF), glmnet, cforest, Classification and Regression Trees for Machine Learning (CART), treebag, K-Nearest Neighbors (kNN), neural network (nnnet), Support Vector Machine radial (SVM-radial), Support Vector Machine linear (SVM-linear), Naïve Bayes (NB), multilayer perceptron (mlp) or any combination thereof.

[093] In some aspects of the methods of the present disclosure, the algorithm can be XGB (also called XGBoost). XGB is an implementation of gradient boosted decision trees designed for speed and performance.

[094] In some aspects of the methods of the present disclosure, a therapy can comprise anti-cancer therapy, surgery, chemotherapy, targeted drug therapy, radiation therapy, immunotherapy, or any combination thereof.

[095] In some aspects of the methods of the present disclosure, surgery can comprise removing a polyp during a colonoscopy, endoscopic mucosal resection, a partial colectomy, an ostomy, removing at least one cancerous lesion from the liver, or any combination thereof.

[096] In some aspects of the methods of the present disclosure, anti-cancer therapy can comprise anti-colon cancer therapy.

[097] In some aspects of the methods of the present disclosure, chemotherapy can comprise FOLFOX, FOLFIRI, a combination of 5-FU and leucovorin, capecitabine, irinotecan, CapeOx or any combination thereof.

[098] In some aspects of the methods of the present disclosure, targeted drug therapy can comprise bevacizumab, cetuximab, panitumumab, regorafenib, a combination of trifluridine and tipiracil, an EGFR TKI inhibitor or any combination thereof.

[099] In some aspects of the methods of the present disclosure, immunotherapy can comprise pembrolizumab, nivolumab or a combination of pembrolizumab and nivolumab.

[0100] For early-stage colon cancer, a minimally invasive approach to surgery can be used to remove the cancer. For example, if the cancer is completely contained within a polyp, the polyp can be removed during a colonoscopy. Endoscopic mucosal resection can be performed to remove larger polyps. Polyps that cannot be removed during a colonoscopy may be removed using laparoscopic surgery.

[0101] If the cancer has grown into or through the colon, partial colectomy can be performed to remove the part of the colon that contains the cancer, along with a margin of normal tissue on either side of the cancer. When it's not possible to reconnect the healthy portions of the colon or rectum, an ostomy can be performed to create an opening in the wall of the abdomen from a portion of the remaining bowel for the elimination of stool into a bag that fits securely over the opening. Lymph node removal can also be performed.

[0102] For advanced colon cancer, an operation to relieve a blockage of the colon or other conditions can also be performed. In specific cases where the cancer has spread only to the liver, surgery to remove the cancerous lesion from the liver can be performed.

[0103] For chemotherapies, either the FOLFOX (5-FU, leucovorin, and oxaliplatin) or CapeOx (capecitabine and oxaliplatin) regimens are used most often, but some patients may get 5-FU with leucovorin or capecitabine alone based on their age and health needs. Irinotecan can also be used as a chemotherapeutic agent for treating a colon cancer.

[0104] Targeted drug therapies target specific malfunctions that allow cancer cells to grow. These therapies include, but are not limited to, bevacizumab, cetuximab, panitumumab, ramucirumab, regorafenib, ziv-aflibercept, a combination of trifluridine and tipiracil, and an EGFR TKI inhibitor.

[0105] Immunotherapies for colon cancer include, but are not limited to, pembrolizumab (Keytruda®) and nivolumab (Opdivo®).

[0106] The sequence information of the colon cancer biomarkers and housekeepers is shown in Table 1.

[0107] Table 1. Colon Cancer Biomarker/Housekeeper Sequence Information

Gene Name	RefSeq Accession	Sequence	SEQ ID NO:
<i>ADRM1</i>	NM_007002.3	gtagagccggctgcgcggcttacggggctcaatcggcgccgagagcggcaggcggggcggccgaacgcgggttccggcgggcccgccaggcgccgaggaggaaagacgagccggacggcgccctctcgaacgagtggtggcgccaggcaggatgacgacctcaggcgccctcttccaagcctgtgcccaggctctcggggcgccccaacaagtacttggtggagtctgggcgggaagatgtccctgaaggggaccaccgtgactccggataagcggaaaggcgtggtgtacattcagcagacggacgactcgttattcacttctgctggaaggacaggacgtccgggaacgtggaagacgactgacatcttccctgacgactgtgagttcaagcgggtccgcagtgccccagcgggagggtctacgtgctgaagtcaaggcagggtccaagcggctttctcttgatgcaggaaccaagacagaccaggatgaggagcattgccggaagtcaacgagtatctgaacaacccccgatgcctggggcgctggggccagcgggaagcagcggccacgaactctctgcgtaggcggtaggggtgcccctgcagagccctgctgggaacatgagccacagccagctcatgcagctcatcgaccagccggcctcgaggagactgggtggctggggccctgactggacctggcctggccagcttactgggagcagtgggcctccagggagcagctcctccaactcccgaggcagtcggcagcgggtaccccgatccaccacctctccaccgtgccaccccagcccctctgctccagcagctgctcagcaactagcccagccccgcgccagttccgggaatggagccagcacagcagccagcccgaaccagcccacccagctgagcgacctcc	1

		agagcatcctggccacgataaacgtaccagccggccagcagggccagcaagtggac cctggccagtgctgacgcccagataatggctccatcctcgaacgcggatgccag gagcgcctgcttccctacttggcatctgggagtgctgcccagaccgcggatgagatcc agaataccctgacctgccccagttccagcagggccctggcatgttcagcgcagccttggc ctcggggcagctggggccccctcatgtccagttcggctgctcctgagaggctgtgaggcc gccaacaaggcgatgtggaagcgtttgccaagccatgcagaacaacgccaagcccgga gcagaaaggaggcgacacgaaggacaagaaggacgaaggaggacatgagcctggga ctgagccacgcgcgtcctccgaggaaactggcgcttcagtgctgtgcacacctcacct cccaccacgtgattatataaagcttttcttttaccctgcccccccccccccccccccc	
CDK4	NM_000075. 3	cacctctgtccgccccctcagcgcctgggtggcggtcagtgcccagaacgtccggcgttc gccccgccctcccagtttccgcgcgcctctttggcagctgctcacatgggtgagggtggggg tgagggggcccctctagcttgcggcctgtgtctatgtcggggccctctgcgtccagctgctcc ggaccgagctcgggtgtatggggccgtaggaaaccgctccggggccccgataacgggc cgccccacagcacccccgggctggcgtgagggtctccctgaltgagaatggtctacctct cgataltgagccagtgctgaaattgtgtcgggtccatgggacagtgatcaaggcccgtg atccccacagtgccactttgtggccctcaagagtgtagagtgccccaatggaggaggagg tgaggaggcccttcccacagcacagttcgtgagggtgctttactgaggcgactgaggcctt ttgagcatcccaatgttgcggcgtgagtgacgtctgtgcccacatccgaactgaccgggag atcaaggtaaccctgggtttgagcatgtagaccaggacctaaggacataatctggacaaaggc acccccaccaggcttgcagccgaaacgataaaggatctgtagcgcagtttctaaggaggc ctagatttcttcatgccaattgcatcttaccgagatctgaagccagagaacattctgtgtga caagtgtgtgaacagtcaggctggctgactttggcctggccagaatctacagctaccagatg gcacttacaccgctgggtgttacactctgtaccgagctcccgaaagtcttcttgcagctccacat atgcaaacctgtggacatgtggagtgttggctgtatctttgcagagatgttctgtcgaagcc tctcttctgtggaaactctgaagccgaccagttgggcaaaatctttgacctgattgggtgctcct ccagaggatgactggccctcagatgtatccctgccccgtggagccttccccccagagggc cccgccagtgtagtgggtgggtacctgagatggaggagtcgggagcacagctgctgctgg aatgctgacttttaaccacacaagcgaaatctgctcctttcgagctctgcagcacttattcta cataaggatgaaggtaatccggagtgagcaatggagtggctgccatgggaaggaaagaaag ctgccatttcccttctggacactgagagggaatctttgccccttattctgaggctatggagggt cctcctccatcttctacagagattactttgctgcccctaatgacattccccccccacctctcctttg aggcttctccttctccttcccatttctacactaagggggtatgttccctcttgccttctccctacc tttatatttgggtccttttttatacaggaaaaacaaaacaaagaaataatggtttttttttttta atgttttctcctctgttggctttgcatgtgctgatttggaaaaaccacttggaaagaggactt tcctgcaaaaccttaagactggttaattacaggccctagggaagtcagtgaggcccttga ctgacaaagcttagaaaggaaactgaattgcttctttgaatatggaatttaggcggggcgtgtgt ggctcacgcctataatccagcacgttggaggccaacgcgggtggatcacctgagggtca ggagtctgagaccagcctgactaacatgggtgaaacctgtctctactaaaaatacaaaattag tcaggcgtgtgtgtgcacacctgtaatccagctacttggagactgaggcaggaggatcg ctgaacccggaggcagagggttgcgtgagccgagatcatgcatgtcactccagcctg ggcaacagagcaagactctgtgtcaaaaaaaaaaaaaagaatatagatttttaaatggcaaaa aaaaaaaaaaaaaa	2

COMT	NM_000754.3	cgccctgctccgccaccggaagcgccctcctaataccccgagcgccaccgccaftgccc ccatcgtcgtgggcttctggggcagctagggtgcccggcgctgcccgtgcccggacc ggggcggtccagtcggggggggcgctcggggagagaaataacatctgcttgcctg gagctcagaggagaccccagacccctcccgagccagagggtgagcctgctcagag gtgcttgaagatccggagggcccccctctgctgttggcagctgtgtgctgggctgtg ctgctgtgtgtgctgctgcttctgaggcactggggctggggcctgtgcttctatggctg gaacgagttatcctgagcccataccacaacctgctcatgggtgacaccaaggagcagcg catcctgaaccacgtgctgagcatgaggagccgggaacgcacagagcgtgctgagg ccattgacacctactgagcagcagaaggagtgggccatgaacgtggcgacaagaaaggc aagatcgtgacgcccgtattcaggagcaccagccctccgtgctgtagctggggcc tactgtggtactcagctgtgcgcatggcccgtgctgtcaccaggggcgaggctcatca ccatcgagatcaacccgactgtgcccataccaccagcgatgggtggttgcgtggcgt gaaggacaaggctacccctgtgtgtggagcgtccagacatcaccagctgaagaa gaagtatgattggacacactggacatggtcttctcagaccactggaaggacggctacgtc cggacacgcttctcttggaggatgtgctgctgggaaggggacagtgctactgctga caacgtgatctgccagggtgcgcagacttcttagcacacgtgcggggagcagctgctt gagtgcacacactaccaatcgttctggaatacaggaggtgtgtgacggcctggagaa gccaatcacaaggggccaggcagcgaagcaggggccctgactgccccccggccccctc tcgggctctctcaccagcctgtgtactgaagggtgccagacgtgctcctgctgaccttgcg gtccggggctgtgtcctaataatgcaagcacacctcggccgaggcctgcccctgacatgt aaccttctgaactgcacactggattgttcttttaagactcaatcatgacttcttaaacac tggctagctatattatctatataataatcatgttttaaaaataaaaatagaattaaatcta aatatttagatataactcgtactgtacatccttctcaactgccatccccctgctcccctgactg ggcaccacaacattcaaaagctccccgtgacggacgctaacgctaaggcgggggccctagc tggctgggttctgggtggcacgctggcccactggcctccagccacagtgtgacagagt cagcccctgtagctagccagggggacacctgttagccccatggggacgactccggcct gggaaacgaagaggagtcagccagcattcacaccttctgaccaagcaggcgctggggac agggtggaccccgagcagcaccagcccctctggggccccatgtggcacagagtgggaagca tctcttccccactccccactgggcttgcctacagaagaggcaatggctcagaccagctccc gcatccctgtagtggctccctggcccagtagtaggagtagcagtgctgttctgcccacct cacctagagctgtcccctctctccaaagggtcagactgtagccacctcagaggctcca aggggccagttcccagcccagacagggaatcaacctgtgtagctgagttcacctgcac cgagaccagcccctagccaagattctactcctgggctcaaggcctggtgtagccccagcca ggccactcctatgtagacagaccagttagcccaagtgacaagtgtggggccaccag ggaccagaacagagcctctgaggacacagcagatgggcacctgggaccacctccacc caggggcctgccccagacgcgcagaggccgacacaagggaagacagccactgtg ccagacctgagtgagcagaagcaaaaagtcttctgtgttatttttaatttttcaaaaa atttaggtgtttaccaatagcttatttggcttatttttaa	3
DHCR7	NM_001163817.1	aatcgtgacatcatccggggcgggcgcccctgcccctgagggtgactccgaccctggc tagagggtaggcggtggagcagcgcgcaagcgaggccagggggaagggtggcg aggacttagccgttgagaaggatcaagcaggcatttgagcacagggtgttagaaacttt aaggggccggttcaagaaggaaaagtccctctgctgtgaaactatttggcaagggtg	4

		agggcccaatggctgcaaaatcgcaaccaacattcccaagccaagagtctagatggcgt caccaatgacagaaccgcatctcaagggcagtgggccgctgctgggaggtggactgggt ttactggcgagcgtcatcttctactgctgttgcctccctcatcgtctactactcatatggc ttgtaccagtacagctgcgcctgactggccctgtgttgacatcgtaccggacatgctc ggctctcgacatctgggccaagactccacctataacgaggaaagccgcccagctctatac ctgtgggtaccttccaggtgcttctgtacacgtctctccctgacttctgccataaagtctacc cggctactgtaggaggcatccaggaggggggcctgactcctgcagggggtgtgaacaagta tcagatcaatggcctgcaagcctggctcctcacgcacctgctctgtgttgaaacgctcatct cctgtcctgtgttctgcccaccatcatctctgacaactggatcccactgctgtgtgtgccc catccttggctatgcgctctccaccttgcctatggcaagggtacttcttccccaccagcgc cagagactgcaaatcacaggaatttctttacaactacatgatgggcatcgagttaacctt cggatcgggaagtgggtgactcaagctgttcttcaatggcgcccccggatcgtgcctg gacctcatcaacctgtccttcgagcgaaagcagcgggagctccacagccatgtgaccaat gccatgtctgtgtaacgtcctgcaggccatctacgtgatgtacttcttggaaacgaaacc tggtaactgaaagaccatgacatctgcaatgaccacttgggtgtacctgggctggcgga ctgtgtctgtgctgtatctttacacgtgcagggtctgtacttgggtgtaccaccccgctcag ctgtccaccccgacgcctgtggcgctcctgtgtgtggcctgtgtgtgtactactatctccg ggtggccaaccaccagaaggacctgttccggcagcgatgggctgctcctcatctgggg caggaaagcccaaggtatcgagtgtcctacacatcccgatgggagaggaccaca gcaagctgctgtgtcgggcttctgggctgtggcccgccacttcaactacgtcgcgacct gatgggcagcctggcctactgctgtggcctgtggcgcgccacctgtgctcctacttctac atcatctacatggccatcctgtgacccaccgctgctcctggggacgagcaccgctgcgcca gcaagtacggccgggactgggagcgctacaccgcccagtgcttaccgctgtgctcctg gaattcttaaggcagcgccttagggagaaagcctgtgggctgtcaaggcgtgttctgc caggctcatggggctggcatcccagctccaactcagaggagcctcagtttctcatctgtaa actggagagagcccagcacttggcagggttccagtacctaatacgcctgttctgtctttg ccttcaagggaattccgagtgtccagcactgctgttattgccaagcacagcggatttctctaat cagtgtccctggggcaggaggatgaccagctaccttctactgtcttctggagacaatttacc tgtattaggagcccaggccacgctacacttgcctcacactggtagcaggagggtcttccca cgccctgtcattaggctgcatttacttctgttaataaaagtgggagtgggcgctgcgcttat ccatgtattgcttctcagctctagatccccctcccctgctgtcgtcagtcgtgggtggggcc cgtgcgcgttctccttggtagcgtgcacgggtgttgaaactgggacactggggagaaaggg gcttctatgtcttcttctctcctgtcgtcacagctgcccaggagtgctcctgagctgtg cagacctcagagagggtccagcaccggctgtggccttccagggttaggcagggtggctctg ctcccgattcctgtgagcgccaccctctcgaagaaatttctgttgcctatgactgtgca gactctggctcagcaaccggggaacttaccctcaggggcctcccacaccttctccagc gaggaggctcagtcaccagcctcgggaggggcacctccttctgtgttcttctcctgaggca ttcttccicatccctagggtgtgtgtagaactcttttaactctatgctccgagtagagtcatct ttatataaacttcccctgttcaataa	
HMOX2	NM_001127 204.1	catctctaggccccccccgcgtgcgtgcccacgttgcgcggcctcgcgccagtcctgct gggctgcagggactgcggcgctgaggagtgctgacgggcagctgactggaggct ggccggacaggcgacagcgacctgcggcagagctctgtcgcgacaccaggctgagtgct	5

		aatggcgctatctcggctcactgcaacctccgcttccggattcaagcgtattctcctgacctca gctccccagtagggggactacaggaccagaggagcgagagcagcaagaaccacacc cagcagcaatgtcagcggaaagtggaacctcagaggggtagacgagtcagaaaaaaag aactctggggccctagaaaaggagaaccaaatgagaatggctgacctctcggagctcctga aggaaaggaccaaaggaaagcacacgaccgggcagaaaacaccagttgtcaaggacttc ttgaaaggcaacattaagaaggagctgtttaagctggccaccacggcactttacttcacatac tcagccctcaggagggaatggagcgcaacaaggaccatccagcctttgcccccttgtactt ccccatggagctgcaccggaaaggaggcgctgaccaaggacatggagtatttcttgggtgaa aactggaggagcagggtgcagtgccccaaaggctgccagaaatagctggagcggatcca ctacatagggcagaaagcggagctactgtgtgcccagcatacaccgctacatggg ggatctctcggggggccagggtgctgaagaagggtggcccagcagcactgaaactccca gcacaggggaaaggaccagttctacctgtttgagaatgtgacaatgccagcagttcaa gcagctctaccggggccaggatgaacccctggacctgaacatgaagaccaaaggaggga tcgtggaggaggccaacaggcctttgagtataacatgcagatatcgaatgaactggaccag ggcggtctccacactggccaggagagaccttggaggatgggttccctgtacacgattgggaaa ggagacatgcgttaaatgccctttctacgtgctgaacaagacaagggtgccctggagggca gcagctgtcccttccgaaacagctatgctgtgctgaggaaagcccagcctccagttcctctg ggcgctgggtgtgcccctagctgtgacttctggcctgtgtactacatgtgaagcaccatcat ggcacaccggtaacctctcccgactgaccactggcctacccttttctccagccctgactaa actaccacctcagggtgactttttaaaaaatgtgggtttaagaaggcaaccaataaaagcca gatgttagagcctctgctgacagcatcctctctatggccatattccgactgggcacagg ccgtcacccctgggagcagtcggcacagtgacgaagcctggccccgaccagctctact ccaggcttccacacttctggcccctaggctgcttccgtagtccctgttttgcagttacatggg tgactatctccccctgttggagggtgagtgccctgtaagccaagctgtcggaggggccttctg ggatgctgctgtacaacttctgggctctcttggacctgggagtgagggtgggtgtgggtg gaagcctcagaggccttgggagctcatccctctaccagaaatccctctaacccttgggtg cgtttgtctagccccagcttatctctctcccgctgtgttaatgtccagcactcaataaa gtgggctttgcaagctaaaaaaaaaaaaaaaaaaaaa	
MCM2	NM_004526. 3	atgacgtcgcgtccgtagggtcttcccgggctttgtgggtcacgtgaaccacttttgcg cgaaccctgtgtgtctgtatgtgggagaggatcgtgtactgctatgcgggaatcatcg gaatctctacatggcatccagcccggcccagcgctgcggaggcaatgatcctctacct ccagcccctggccgaagctcccgcgctactgatccctcactccagccctggccgtgacct tccaccatttgaggatgagtcggagggtctctaggcacagaggggcccctggagggaaga aggaggatggaggagctcattggagatggcatggaaagggaactaccgcgcatccag agctggacgcctatgaggccgagggaactggctctggatgagaggacgtagaggagctga cggccagtcagagggaagcagcagagcgggccatgcggcagcgtgaccgggaggctg ggcggggcccggccatgcggctgggtctctgtatgacagcgtatgaggagacgag gagcgccctgcccgaagcggccgaggtggagcgggcccagggaggacggcgaggga ggacgaggagatgacgagagcatcgaacctggaggatctcaaggccactctgtgcg cgagtgggtgagcatggcgggccccggctggagatccaccaccgcttaagaacttct gcgcactcacgtgacagccagggccacaacgtcttcaaggagcgcatcagcgacatgtg caaagagaaccgtgagagcctgtgtgtgaactatgaggacttggcagccaggagacgt	6

		gctggcctacttctgctgagggaccggcgagctgctgagatctttagagggctgccc tggaggtggtactggtccatgtaccccaagtagcaccgcatcaccacacatccatgtccg catctcccacctgctctggtggaggagctgctgctgctgagggagctgcatctgaaccag ctgatccgacacagtggtgggtgaccagctgactggtgctgctgccccagctcagcatgg tcaagtacaactgcaacaagtgcatttcgtctggtctcttctgcccagtcacagaccagg aggtgaaaccaggctctgctgagtgccagtcggccggccccctttaggtcaacatggga ggagaccatctatagaactaccagctatccgaatccaggagagtcacaggcaaatgtggc ggctgcccggctgccccgctccaaggacgcccattctctcgcagatctggtgacagctgc aagccaggagacgagatagagctgactggcatctatcaacaactatgaggtccctca acactgccaatggcttccctgcttctgcccactgtatccctagccaaccacgtggccaagaagg acaacaagggtgctgtaggggaaactgaccgataagatgtgaagatgactactagcctctcc aaggatcagcagatcggagagaagatcttgcaggcattgctctccatctatgtgatga gacatcaagagaggcctgctgctgcccctgtcggaggggagcccaaaaccagggtgg caagcacaagggtacgtggatataacgtgctcttgcggagaccctggcagacggaag tcgcagtttctcaagtatatgagaaggtgtccagccgagccatctcaccactggccaggggg gctgctggctgtgggctcagcggtatgtccagcggcaccctgtcagcagggaagtgacc ttggaggctggggccccgtgtctggtgaccgaggagtggtctcatgtatgaattgacaag atgaatgaccaggacagaaccagcatccatgaggccatggagcaacagagcatctccatct cgagggtggcatcgtcacctccctgagggtcgtcgcacggctattgctgcccgaaccc cataggagggtgctacgacccctcgtgacttctctgagaacgtggacctcacagagccc atcatctacgctttagatctctgtgtgtggtagggacaccgtggaccagtcacaggacga gatgctggcccccttctgtgtggcagccacgtcagacaccacccagcaacaaggagg aggaggggctggccaatggcagcgtgctgagcccgcctatcccgaacagtgatgctgtg gagcccctgccccaggaggctctgaagaagtacatcatctacgccaaaggagagggtccac ccgaagctcaaccagatggaccaggacaagggtggccaagatgtacagtgacctgaggaa agaatctatggcgacaggcagcatccccattacgggtcgggcacatcgaagtcattgccgc atggcggaggcccacgctgcatccatctgctgggactatgtatcgaagacgacgtcaac atggccatccgctgctgtgtagagcttcatagacacagaagttcagcgtcatgctgca gcatgccaagactttgccccgtacctttcattccgggtgacaacaatgagctgttctctt catactgaagcagttagtggtcagagcaggtgacatatacgcgaaccgcttggggcccag caggacactattgaggtccctgagaaggactgtgtgataaggctcgtcagatcaacatcca caacctctctgcatcttatgacagtgcctctcaggatgaacaaagttcagccacgacctga aaggaaaatgacctgacagcagttctgaggccctatgccatccataaggattccttgggattc tggtttgggtggcagtgcccctgtgtcttatggacacaaaaccagagcactgtatgaact cgggttactagggtcagggttatagcaggatgtctgtgctgcacctggcatgactgtttt ctccaagcctgcttgtgtcttcacaccttgggtgggtagccttgccagtgtgtcttactgtgtg ctgaacatcttgcacacctccagtgcttgtctcactcagcttggatcagagctgctgag ttcaggatgctgctgtgtgttaggtgttagccttcttcatgtagtgcaggagagctgctgc ccctgtgctgagtgctgtatcagcgtcttctgtgcttctgcccagagagctgtgtgaag atgtttgtaatcgtttcagctctcgcagggttctgtgcccctgtgttggaagaggcagac agtgccagcgcagcgttctgggtcctcagtcgcagggtgtggatgtgagtcagcggatt atccactgccacagttatcagctgccattgctcctgtgttccccactctcttattgtgcat	
--	--	--	--

		tcgggttgggttctgtagttttaatttttaataaagttgaataaaatataaaaaaaaaaaaaaaaaa a	
PDXK	NM_003681. 4	cggaactcgcgggttcggagccgcccgtgaggcagaaaggaggcgtctgcgtgatcg gggtccgcccgcgcgcagagccagagtcgcagccgaggggaggccggccggagccc gagcccgagccgagccggagcccgagcggcggagaccgtgccccgcctcgggc ccccgcgcgcgcgcgcagcccgcatggaggaggagtgcgggggtctctccataca gagccacgtcatccgcggctacgtgggcaaccggcgccacgttcccgtcaggtttt gggattgagattgacgggtgaactctgtccagtttcaaacacacaggctatgcccactg gaaggggccaagtctgaattcagatgagctccaggagtgtacgaaggcctgaggctgaa caacatgaataaatatgactacgtctcacaggttatacagggaacaagtcgttccgtggccat gggtgtggacattgtcaggagctgaagcagcagaaccccaggctgtgtacgtgtgtat ccagtcctgggtgacaagtgaggacggcgaaggctcgtatgtacgtccggaggaccccttc ccgtctacaagaataaagtgtgtccgttcgagacattatcacgccaaccagtttgaggcc gagttactgagtgccgggaagatccacagccaggagggaagccttcgggtgtatgacatg ctgcactctatggccccgacaccgtgggtcaccagctccgacctgccccccgcagg gcagcaactacctgattgtctggggagtcagaggagggaatcccgtgctccgtgtgt gatggaaacgcacccggatggacattcgaaagtggacgcccgtcttggggcactggggac ctgtttgtctccatgtctctgtggtggacacacaagcaccccaataacctcaagggtgcccgtg tgagaagaccgtgtctactctgcaccacgttctgcaggaggaccatccagtgtgcaaaagccc aggcccggggaaggagtgaaggcccagcccattgcagctggagctgcggatgtgtcagag caaaagggaacatgcaggaccagagatcgtctccagccacgggtgtgtgaggggcccc gcccgttcggcgtgacacgcagcggcgttgggtctcctggtttgtccctgtgaaaacatgtaa cgtctgcccattagagccatgaccgaactgtatattttttcttcatgagtgtccggcactgtctg gtcttcattgtgaacgtgtccagtcgtgttggtaaaaataacaagtgacagaaatttgt gatctgaaaaccggctcccttccccacaaggctcctggccctccgggaagacgggcccc gtttgtccatctcgggggtgttccctgtgggagggtgagtggtgaggccgagcctgtctc gtgtggagccctcagtgaggccctgctgcccactaccgtacagaggccgtgtcgcgtggg ctgggcctgggtggcctctgtctttgcatctctgagaaaggagtgggtgtaacgggtgggg tcaggaaagaattctccaagtatcttactgtcattctgaccatagccctttgttcccgcattcg aacttttgggttactttgtctgcgttttagtccctggggatttcagatcttagcgtgtgtttcac cgtatggagggtgtatgtgagctgtcttgagacacacgggtgcagcatcaggacacctcc caggccccagcaaatcaagtcgggtctgcagacctctcagctacccgagggaaccttctgtaa cccacggcatcttcagggaatccggcagtgacttgagggaagatgctaacgcagtaaggt ctgtctgtggccaagagcagcttgaagctccagagaaccaccccgtcagggttccctgtgga agctccccctacccgtgtgtcagcaggctgagcactgcggtttgccaagtgctgcccgtg acagcacattgagccacagcattgttagacaggacagagggtgcccctgccccctgcccctg ctggcacatttaacctgttccctgacctcagttctgtgccccaccaaatgcccaggggcga gaggccaccctggaagctgccaatctccaagggtgggtgtggggcacgggtggggggcggg cagctcccaggcccttgggcaggctgggtgtacggcagaggccacagcaccagctctga caagtccatcactctgtctcagcagtgacctccctgccccacttggccagagtttgggt ccccccaggtatagctataggcggcagtgccctgtccctgcccgtgtgttgcagccacac ccctgcagccctgcacccagctctgggggtgtgcagagggttgtgtctccagggaacccac	7

		ggctggagagaaataggagagatgcaggagtgggggcccatggggccccaagagcg gactctcaagggggtaccccaccccgctacctccccacggacggggccctctggagc ccataccctcctgtgagccattccagtgcttctagaagactcgttgcaggagtgcgtt ctttgtgaaaaatgccctgaagcgaaaagatgcagggttatatggaacccccacccctccc ccactctcccactctgttctgtgaatgtcttcacgagcgtgcatcaggggcgctgtctcccc cacctcagccagtgatcagacacgggttctgcagccatgttctgtgctccgaggacacg gggtggcaggcccggtgcagcccagagccactgtccctacaggggcgccgacaccagc aggaaaggaggatgctgtgtccggagcctggcgggaggcgccctcccagtatgtgag tgcagggatctgcagaaccacctggccctctgtaggcggttaactggaataccctact gccaagtggagactggggcggtgtccacattgccagccaccaggaaagcttttctttcttt tttttttttaaacaccaagagcacgtatagcatgggggaaagaacctaaatgtctctgtcc tgtgagctgtgaaaaaccagcatgagaacgcagtgtaggtgtgggactccttctgtccc ctgcagtggtgtgtacggcggtgtgtccctggcgagcaagcttgaattctgttctgtgagct cgtttcagaggctgagtcacacatcagctttagttcttggacttccctgtatlaagcaagaatt aggagaatggctgtccctgcaggcgccctccgtaaatcctgagctctctggcgcaatctgaa actctctctgttttcttggctgtatcagccgaaccaggagaggccctggcgctgactaagg agaaagaaatcggggtttctgagagcagatgggtcctttgtgggtgagggttttctgga aatgtcagccctacgggagagtcgggcatccctcccagactgcctgtgtcaaacca cggagcagctggagcctgcccgttccacggcccgtttccacccggcgatgttctctcat gacttcggcagagcccccgtgtggccttcagttctcctcaggaaggtaacctgtg gcattggcagtggtttccctatggcttgatccagattagaattgactttgttttcactttccat agtaataacatgcaaaataatgagaagaaattatitaaaggtagactatagtgccaacat cgccgtctatgtcagggtacagaagtttaatacttttaataccagttttcaacttccctgt agaccgttaaggatgaattccacaataggatccttttaaaatcgatttaaatgtgtcctagtcc tgccaagggtatattatgtcaicgttattttccaatacatgtaaacagttgcagcatgatgttgg tttaattctctgttcttaagctcgttagagccagttttgaaacgtttggcttaccgtgaacggag gctggcttggcttagccacgctgatgagtaagtagggtatgtctccatcttgatcaccag gcaagagaggtgctgcaccaggtaagaggccaagccctggggtaacagtccccacc gctaccggaggtaaaacaataaaagctatgtgttgagctcagggcctctcgtgctgtgtctc agagaaggcagagcccacagtagggtgcacgggtgcaaggccctgggagggcactggcca gggaagggtgtatagatggccctcagattgcggggccccgagcagctccccactctgccc gtccaccttccctggctccagcctcattctctttagtttaactatgcaaaagagagggttga gagtgcttggcagctggagctcttttcttctcttccctccctccgaltggggccacctgtgtc ggggcagcagtgccatgtttatggagatcagagggttccccactgtgtgctggactgtact ctgctgcccgggtagccaggagctctctctctctccctgcccctgctgctcctgtggtg cctccttcacacacccctccctgttgatcgcctgctggggcccagagcaggggaaactggag tttgtgagtgagcagagcagggtatgtgcagacagggaacgagaactttggacctgtgtt ctgagtcagggtgagagctgtgtggccccccgatgccactctgcccggagggtatgtg cctgtgagcccttttcttccacggccctctcactgcccagccagcggcttccgtgagact cgtgtgagaggcggtctccgtgtccgtccaccgagcactcagatggatgtgtatcaccag ggccgagggggtcccaagagacccagccctggggagggtgtgtggaggcc aagtcactgcccgaagctgtcagccctaagccagggaagcctggagcgtggcctgtg cgggtctgggtggacacgtccccactccggactccagcacaggggaggatccctgag	
--	--	--	--

		cctgtatggccctgtagccctgggcagagctgggctgtcgtgtgttctcctggcaggtg cagggtgctggccatctgcaggtggaaaggagggtggaaatctggattttttttttttttttt ttttttgagatgaagtcctcctgacacccaggctggcgtgcagtggtgtgtaiccggtcac tgcaaacctcgtctcctgggtcaagtggttctcctgcccagcctcccaagtagctgggtt acaggcatgcccaccacgctcagctgattttgtatttttagtagagatgggtttcaccatgt tggccaagctggtcacaactcctgacctcaagtatctgcccgcctcggcctccagagt ctgggattacaggcatgagccagtgacccggcggaaatctggaaattttatagacagacc tcagttctgactccagccgacacacctcctgacctgcccagagggtgtgcccagaccag agccagggccagggtccctgctccatccccccggtaggagtgacgtgagccatcctcta ggggactttttcagtgctgactcgtctcgtttagggtgtaggagccagttgtgtgcccgtg ccacgctccacagtgctggtgctgctgtgtgtgcccgtgtccccgttccctgcaggac ccagcaggcatcgtggtgacagctgtgtccaagccactgcccgggcatccatcacc accagggtgacggtctcctcctgctggggctttctgctgcatgtgtgtcctcgtgactctg cagttgttctcagagcagaatgtttcctgtctcagtgcaaaagacactgggtttcactggc gtctaaaaccacgttccctgctttcattgcaacacgggtgtttcattgttaaacagttatga gtaagtttagatgactgggtcaatatcttaaaaatgtatattagtaagaagtcttccctggaattttc tttcgattctggcagaataaacaggtgttttagtttccactgtctgagccaagcaggaccct gtcccagagcaagagatgtccccttccatctgaccttgcctgggacaaggctttagtggtg ggggcccagcttaaggctgtgtgtgggaacagcacccccataatgcccagcctcctttctc ccatccaccagtatactgcggggccatttctggtctttgtccaacaggaaacccatttctgtg ggatatgcttccagtgccacagggccactcaccctcagcatctctgctcctgcccgtcagtg ctgggacggacagcaaggccaagcccagtgctgtggcgatagggtgggtgggaacagag aggggagaaatccgttctaaagcttctgtgtgggatccccacacgacctgggtactgcct gggaacctgtcctaagtaaaactatggacctcgcctcgcacccagcctgcgaagccag catctcgtgaaggtgtgaggaagcgcccttctcctcattttgagctgaagctgggtcagcg gctctgaagccctcgtgagtgacttttaacccaagaccagcccctggcaggagggagggtg gggtgcagggtgtgtgtggacaaaaggagccctcagcaggccctgggaagaccttccagtac atcccacagcgtgtcagcagctggggaagacctgtgtcaagctcagcggctcatagggtccc catgagggtgtcgaagccccttctgtgtgtgtggaggcagagggtgtgacgtctgagcat ggacgtgagtcctcagctgctccgctgggccccttggagggtgcccaggtgtgtgtgacc ttctggtgagcctttaacttcagtgctgctgcttctgatttagaactttaaccggagcttcatcta gtgattgcaaaactggaccaatggaggacggcggcgcagcccgtcctcctcgtggaatg gagctcagctcttggaggcatcaaaagacctgtcgcctcgtgtgtcccctgctgaggga gtgcggcctctgcaagggttcgggggtggcttctgttgcctggagtgccggcccctgtgtg ccatgtggtgtgtgagcctcggctctacagcactgtgtaggctgcatctgtttcgtgctggt cctgtgactgtatgatatccacaataaataattttatggcgggtgtgtgaaaaaaaaa	
POP7	NM_005837.2	ggaaagggtggggcgaacggaagccgggaagcgaattcatagctcgggggtacggg cgcgctgctgcactccgcagcccgttcaggaccccggcgcggggcagggtgcccacgag ctggctggctgtgtcaccacatccttcttctgtggacctgggtgcgggttacttgggt ggccggcgaacccctgagtggctgtgcccgggagcgggctcgcgcctgtgagggtccc tgtggaacgaagagaggcacacagcatggcagaaaaccgagagccccgcgtgtgtg gaggctgaactggatccagtggataacaccttaggaaaaggcttcccagccgcctgcccc	8

		ggagacccaatgacatttatgtaacatgaagacggactttaaggcccagctggcccgtg ccagaagctgctggacggaggggccgggggtcagAACGCGTgctctgagatctacattca cgcttgggctggccatcaaccgcccacatcaacatcgcctgcagctgcagggggca gcttcgggctcttgcagggtggctgccaatactccaccgtggagcttggatgagctggag ccagagaccgacacacggggagccactgactcggatccgcaacaactcagccatccacatc cagcttctcagggtcacaccaagtaattgaaaagacactctccacttatccccctcgat atggctcttcgcatgctgagctgacctcggaccagagccatgtaagaaaaggcctgttc cctggaagcccaaaggactctgcatgagggtgggggtaattgtctcttggggcccagti agtggccttctcgtggtgtgtatgctggtctgtaactatgcatataaataaaaaatcctgtt gcactagtgtctgcatcccaaaaaaaaaaaaaaaaa	
<i>S100P</i>	NM_005980. 2	tgaggctgcttataaagcaccagaggctgccagtgggacatttctcggccctgccagcc cccaggaggaaagggtggtctgaatctagcaccatgacggaaactagagacagccatgggc atgatcatagacgtctttccgatatctggcagcaggggcagcagcagaccctgacca agggggagctcaagggtgctgagggagaggactaccaggcttctcagagtggtgaaaa gacaaggatgcccgtggaataatgctcaaggacctggacgccaatggagatgccagggtg gacttcagtgagttcatcgtgttcgtgctgcaatcacgtctgctgtcacaagtactttgaga aggcagggactcaaatgagctccctggagatgtcacagattcctggcagagccatgggtccag gcttcccaaaagtgttggcaattatccctaggctgagcctgctcaltacctctgattaa taaatgcttatgaatga	9
<i>SNRPA</i>	NM_004596. 4	ggcggggccaggagagaaagcttgggttgggtctcagggaagtagcaggcggcgttg agagaactacggccctgtcggaaagtaacctcgggtgcaaacgaccatcggcggcagggc gagcgggtacgcttggcgtccgggcttctcgggcccgtctgaggaaactgtctcaggg ccaggctgcctaggacctgtcccttttctatctggtcctccacatccgggttttctcggg acggcccttcggatgcttgggccaatgggaatcgccatttaggggtcctccgccaccgggt cgcgtagagcatcctggaagtcgtagtaaatctctcagaggttctctccgcacggcgtgg agaaagcgggtcctacgcacgcttgggtcgcgcttgcctccgtcctgcccctactccgc cttacctgacttcttctggagggaagatccttgagcagccgacgttggacaaaaggatttg agaaaccaggggctaaagtcacgttttctcctttaaagacttacctcaacacttccatcagg cagttccgagaccgcccctaaccacactatttatatacaaacctcaatgagaagatcaag aaggatgagctaaaaagtcctgtacgccatcttctccagtttgccagatcctggatctc ctggtatcacggagcctgaaagtgaggggccaggccttctcatcttcaaggaggctcagca gcgccaccaacgcccgtcctccatgcagggttcccttctatgacaaacctatgcgtatcc agtatgccaaagaccgactcagatcatgccaagatgaagggcaccttcgttgagcggga ccgcaagcgggagaagagggaagcccaagagccaggagaccccgccaccaagaagg ctgtgcaaggcgggggagccaccccggtgtgggggctgtccaggggcctgtccggggc atccgcccgtgactcaggcggcccgcattatgaccacatgccggggcagcggccctac atgccgcccctgtatgatcccccgccaggccttgcacctggccagatcccaccagggg gccatgccccgcagcagcttatgccaggacagatgccccctgcccagccttcttctgaga atccaccgaatcacatcttcttctcaccacacccgagaggagaccaacgagctcagctg tccatgctttcaatcagttccctggcttcaaggagggtccgtctggtaccggggcggcatgac atcgccttctgtgagtttgacaatgagggtacaggcagggggcagctcgcgtgcccctgcagg gctttaagatcagcagaaacgccaatgaagatccttggccaagaagtagcacctttcc	10

		ccccatgccctgccccctccccctgttctggggccacccctttccccctggctcagccccctga aggttaagcccccttgggggcttcttggagccgtgtgtgagtgagtgctgcccacacag catgtaccagagctgttccccagacattgcacctggcgctgttagggccggaattaaagt gctttttgagggttgggttttcacaatcaaaaaaaaaaaaaaaaaa	
SORD	NM_003104. 5	ctccacgctagcggccgcccaggctgacacaaggaggaaagcctagtcccggcccctgcgt ggggcgcttctccagggcccaccttccatccagtgccctggaccctcggctgggtagcgc caccagagcgaccaaacgtcccgcgcttccaggccgcactccagagccaaaagagctc catggcgggcgcgccaaagcccacaacctttccctgggtgtgtgcacggaccgggggactt ggcgctggagaaatctatccctgaaccaggcccaatgaggctgtgtgaggtatgcatt ctgttggaatctgtgtcagatgtccactactggagtagtgcgaattgggaattttattgtg aaaaagcccatgtgtcgtggacatgaagcttcgggaacagtcgaaaaagtgggcatcgc gtaaagcacctaaaaccagggtgacgtgtgttgcacagccgtgtgtccccggagaaaatg atgaattctgcaagatgggcccatacaatctgtaccctccatcttctgtgtccacgcccc cgatgacgggaacctctgcccgttctataagcacaatgcagcctttgttacaagcttctgac aatgtcaccttggaggagggcgccctgacgagccacttctgtggggtacatgcttgcag ggaggcggtggttaccctgggacacaaggcttctgtgtgtggagctgggccaatcggtgat gggtcacttctgtgtggccaagcaatgggagcagctcaagtagtggtgactgtatctgtcgt ctacccgattgtccaaagccaaggagattggggctgatttagtctccagatctcaaggag agccctcagggaatcgcagggaagttagaaggctcagctgggtgcaagccggaaagtcac catcgagtgacggggcgagaggcctccatccaggcgggcatctacgccactcgtctgtg tgggaacctcgtgtgtgtgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt ccatccgggagggtggtatcaaggcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt tcgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt aagctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt acccagtgaccagaatccctgatgttaattgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt tcagggcacaatggctgtgacatgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt ataactaatacaattcatgtgaacagaagtccttaagcagaggaaatgtgtgtgtgtgtgtgt acaatctgggtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt agcaagtagagcagagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt catgtagggaaggaatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt atgagggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt atgaaaatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt tgtttctagagacataatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt cctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt tgtcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt cactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt accccccccccaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt ggaaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt cccccccccaagatttagcatccatactagactatactcaactcaactaggtgtgtgtgtgt aatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt accagtaaaggctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt gtttttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	11

		aag aatg agg atata tag ccatg acat ag cttg ag ctat ag cttt aatt cctt act ttg gct at ggg tgg aggg tgg ag ttg aag aggt tctg att tctt gta acc tggg aaag ccatg acc ttgtg ccc gatt cttt cag att gct ttggg taataaat attggtggtggtatctg actc atgctgctgttta tgg tctgt tttag tgggg aatgg actc aggtt accc atttccc agagg aaagg atccc agga ttttg aagg ttac atatttctgtaccaaatataatttcattgacatgaattatctctaaccctcatg acaagccacatacacaaatcattttgtagataaagaagatataaatgccagaggagaccttaa gattgtcttacaacacaaccccttcagttaacgagagagg	
<i>STOML2</i>	NM_001287 031.1	tccggggg agcgg aactg caag agg aaagg ctcgggt aggcttctgg agcgg accgctc cgctcgtctcgttggttccgg aggtcgtg cggcggtggg aaatgctggcg cgcgcggcg cggggcactggggcccttttctg aggggctctctactgttcttggcgcgtccgcgcgc gcccctcctctgg attgccccgaaac accgtggtactgttctgtccgcagcagg agggcctg gggtgtgg agcgaatggggccgattccaccggatcctgg agcctgtttg aacatcctcatcc ctgtgttag accggatccg atagtgcag agtctcaaggaaattg tcatcaacgtgcctg agc agtccgtgtg actctcga aatgt aactctgcaaatcg atgg agtcccttacctg cgcacat ggaccttacaagg caagctacgggtgtgg aggacctg agtatg ccgtc acccagctagct caaacaaccatg agatc agagctcgg caaac tctctctgg acaagcttccgggtgg agg cag agcggcggg aaacggggccacagtctag agtctg agggg acccg agagtcggccatc aatgtggcag aagggaagaaac agggccc agatcctggccctcga agcagaaaaggctga acagataaatc aggcagcagg agaggccagtg cagttctggcgaaggccaaggctaaag ctgaagctattcg aatcctggctgcagctctgacacaacataatgg agatgcagcagcttca ctgactgtggccg agcagtatgtcagcgcgttctccaactggccaaggactccaactat cctactgcccctcaaccctggcgtatgc accagcatgggtggctc agggccatgggtgtatag gagccctcaccaaagcccagtgcc agggactccagactcacttccag tggg agcagca gagatgtccagggtacagatgc aagcttctg atg aggaacttgatc agtcaag atg agttag tgg agctgggcttggccagggg agtctggg aacaagg aagcagatttctctg attctgctct agtctccctgccaagatttgg tttttatttttatttgaactttag tctgtataaaactcaccagt ggcaaacagaaactgtcctcttggattggggaatgaagtggg aaagtcactagcatittct tggatccagtcctgtcagcatg atgctccatg aataag agtg aacttctgtaaagtgaaact	12
<i>UMPS</i>	NM_000373. 3	ctgcagacgaggcaggcgg aagaggcggg acttgcgggtg acgtcatcggggcgccg gaggcccggggcgccctggg aatttg aagcaaacaggcagcgcgcgacaatggcggctc ctcgtgcagctttggggccattggtagcgggtctgtacgacgtgcaggctttcaagttgggg acttctgtctgaag agcgggcttctcctccccaatcacaatcgtatcgtggggcatcgtgtctc gaccggctcttctg agtcagggttcagatattttattccaactgccc aaatgcaggcatcag ttttgacaccgtgtgtg agtgcccttatacagctttgccaattggctacagttatctgttcaacca tcaaatccaatgcttattagaaggaaagaaacaaggattatggaactaagcgtctgttaga agg aactattaatcaggaga aacctgtttaatcattg aagatgtgtgaccagtggtatctagt gttttggaaactgttg aggttcttcag aaggagggttg aagggtcactg atgccaatgctgt tggacag agcaggagg gcaaggacaagttgcaggcgcacgggatccg cctccactca gtgtgtacattgtccaaatgctggagattctcagcagcagaaaaaagttg atgctgagac agtgggag agtgaagggtttattcaggagaatgtcttggtcaggcgaatcataatggttct cccctttctataaagg aagcacc caaagaactcagcttctggtgacgtgcag agctgccc ggatccaccagtgatcgaagcttctcaggcttatgcaaaagaaggagaccaatctgtgt	13

	ctatctgctg atgttctactggccagagagctgttgcagctagcag atgctttaggacctagta tctgcatgctgaagactcatgtagatatttgaatgatttactctggatgtgatgaaggagtga taactctggcaaaatgccatgagttcttgatattgaagaccgg aagtttgcagatagga cacagtga aaaaagcagtatg aaggaggatctttaaatagcttccctggcag atctagtaaa tgctcacgtggtg ccaggctcaggagttgtg aaaggcctgcaagaagtgggcctgcttgg catcggggggtg cctccttattg cggaaatgagctccaccggctccctggccactggggact acactagagcagcggttagaatggctg agggagcacttgaattgttgggttttattctggc tcccgagtaagcatgaaccagaatttcttacttgaactccaggagttcagttggaagcaggga ggagataatcttggccaacagtacaatagcccacaagaagttattggcaaacgagggttccga tatcatcattgtaggtcggtggcataaatctcagcagctgaicgtctgg aagcagcagagatga cagaaaagctgcttgggaagcg tattttagtagacttgggtttg agtgcctcagatacatttt cagatacaatgtgaagacattgaagatatgtgctcctctgaaagtcactgctggaataat ccaattattcctgcttggatcttccacaggccctgtgtaagaatgggttctgg agttctcatgg tctttaggaatattgagtaattgttaatcaccgcattgatactataaagttcattcttaagcttgg ctttttttagactggtgttggtagacagccacagtcctgctgggttaggggtcttccacatttga ggatccttccatctctccatgggactagactgcttggttattctatttttttattttttttaga caggatctcactctgttggcccaggatggagtgacgtggtgagatcacggtcctattgcagcct cgacctccaggtgaccccccacctcagcttccagattagctgggtgctataggcctgaccca ccacgtccatcaaatcttattattttagagatgaggtcttggccatgtaccaggcgtggtct caactcctgggtcgaagcgatcctcctgctcagctctc aaagtgtggtggattacaggtgtg agccactgtgcccagcctaattgcagtaagacaaaattctaggggcaccagaaggctaaag tcagcacagcttttctgtgctcctgtattctctgtctaatgttggcccaataatacctaatgtta gccattccctccatctctggtcctaaaagttagatgccagggtatccacatgggctgttccca gaactgccattgctcactctccaaagagggggaagggtgggaagggtgaagggtgactatagc tcagctcctgagctagtatctgctgttattcaacaaccggagttgggttgggtcctattttt cccctagccagcaattatgaccagtagtaacacaagtgacagcttccctgtgactgacttcac aattaggagggtctaagattccatttgggtatttggcttaaggatcccacataattgtcccacgggt cattagtagaggggaggtaagccttcattataataaagagaagagcccacattcaagggtgtgt gtttgagcaggggcagggtgagggtgtgctccgggtgtcattgaccagcacactcacatc cttctcatttggggcccacctgcaggaaagtggcacaggatcagccatttcccaccttctca gctgattggcccactgtctttaatgactcagaggaaatgcctaggatttttttttttttagacag aatctcactgtcggccagggtggagttcagtggtgcagatctcggctcactgcaacttgcct ctctgttcaagcagttctcctgctcagcctccgagtagctgggactacaagcctaggatt tttaactcagggttttattatattccctcctgaagtttttactcaagagcttctgtctaaagtccaa tttgggtctcatgtcccagtgctgcacatccagggaatgctgtctgtgggagagaccaact ctcaagggaagaaagtggccacagaaaggagcagggaagggttggccctcagggtactct gggggaagccaaaagtcagaaaggggagaaagaaatttctgacaaaacttgcaggaaatctct agggtctctcagtggtgagtgatatgttagaggcctttggagtgatgtctgagggtctcagg cgcccacctccctggtgtcacttccatgtgtcagtggttctcccacttttagcaggtatcagag tcacctggagctgttcaaaaacaggtaccagccccaccgcagcggttctgactctgggttag ctctgggtgaggtgtgagaatttgcgtttccaaaagggtccagggtgagctgctgggtgct gcgcagggaactggactttagaaccacttactggttattcacatttctgctcctgagtaga cagcctttaggtgtgctcctcgtctaagagtcacatgctcctgtcctttagaaattgtgggtcct	
--	---	--

[illegible]

		gagaagcagaccctgggtggcctgagagtcataatcgtcaggttaaggacagtcagtgggaa gtggacgggcccgcacaaccaagggtctcatgaggacaaccatgtcttcgggggtgcccttg tgcacagacagctccatagtcctgctcccaatgtcccaactgcattgtctccctgcacttag cagccctgcagggtgagacttggggaggatcctgaaatgattgtatttaacaagacatgctgt cctgtttacctggaaacctagcaatgtgtttctgccacaacttgaaatgatacttgaaagcaga gatgatgtgagttaaaaaaatatatacataaaaatatgggtctttcaacctgaatgatgg cctaaaaattcaaa	
<i>MORF4L1</i>	NM_001265 603.1	cggcgtgccctggggcggcgcggggcgaggggcgctgcgcggcgggcgtgctgtggc tggagcagcgctgcgcgggtcggggtctgtgaggtctgcggcgctggcgaatccgg cccaggatgttagagctggcagtgctgacggcgctgtacgcggagttgggtgggttag agagtagggggcggtagtcgggggtggaggagagggaggcggaatcactata aatggcgccgaagcaggacccgaagcctaattccaggagggtggatgaatgggtccg gagagcagagtactcaatacgtggacaccaatttgcagaacagcgaagaactcaaaaag ccaatcaggagcagtatgcagaggggaagatgagagggtgctccccaggaaagaaac atctggtctgcacagaaaaatgtgaagtgaaaacgaaaaagacaacagaaaaacact ggaaatggagatgggtgcagtaccagtgagacccctcagcctccctgggaagaaaagggc ccgggttagatcctactgttgaatgaggaaacattcatgaacagagttgaagttaaagtaaa gatcctgaagagclaaaccgtggctgttgatgactgggacttaattaccaggcaaaaaca gctctttatcttctccaagaaagaaatgtggtatccattctgaggatfatgcaaaftacaaga atctcgtggaaacacagataataaggagtatgcgggttaatgaagttgtggcagggtataaag aatactcaacgtaatgttgggtaccagctactctataaatttgagagaccacagtatgctga aatcttgcagatcatccgatgcacccatgtccagggtgatggagcgccacatcctcagag attattgtacgaattggagcaatgttggcttatcacctcgtgatgagaagagccctgtcttatt actcaattatctcagatttctaaagtacctggcgaagaaattctgaacttgttcagtccag cgattatgaagtggctcctcctgagtaccatcggaaagctgtgtgagagcgactctcactcac ttatgtttggatccgtaaacacattttgttcttagctatcttctacaaaacgatgtcttga gatgttagtgtataacaattgatgtttgtttctgtttgatttaaacagagaaaaataaaagg ggtaatagctcctttttcttcttttttttttattcaaaattgctgccagtgtttcaatgatgg acaacagaggatagctgtagaggttttatgcttagtgacaagctgtctttgaatgctgg tggttctatcttggacactacgcactttataatacatgttaatgctatatgacaaaagctctga ttccttagtgccaaaagggtcaattcagtgatataactgaacacactcatcattgtgtctttgttt ttttatgggtgttaagtaagagcccatccttgcgaagtcacatggtgttacttaggcatttta tcttggctcaaatgttgaagaatgggtgttctatggttttgtattgtgtctaatgcacgttt taacatgatagacgcaatgcattgtgtagctagtgttctggaaaagtcattcttttaggaattgttt ttcagatcttcaataaatttttctttaaattcaagaacaaaaaaaaaaaaaa	14
<i>MRPL19</i>	NM_014763. 3	gtagtcttgacgtgagctagctggcatggcgccctgcattgcagcggggcactgggctgca atgggcttagccggagtttcaagccgcccaggactctgtccccccgcccctctatcg cctgcagggtccacggcgggcctgtccggcagcagagcactgggcttccgagcccggt gcgttccaaccgcccgaaccgggtcatctgtgacaagcaccgccccgtgaaaccgga acgcagggtcttgagtcctgaattcattcctgaagggggaagaacagatcctctgaatttcaa atagaagaaaagataatgttagaaaggagaaaagtactccacattccagagttctatgttga agtattctctgttactacagctgacccatagccagtggaataatcagccagttcttggga	15

	tttgatcagagagatcagggaagaggacttggagctactttcattccttaggaatgtatcgaagg acaagggtctcagatcttcttgaactttataatcctcgggtccaggagatcaggtgtctcaaa tttagagaacacggctggatgatatgcttacttactacgagatgcccttccgtaatatagcacttt tgatgtgataatagaagccagtagtacaagagccctaaccaaaaagtctctgttaattgagctgaa agtaaaaaatgaagccctaagccctggctctaaacgcgtgggaacgtccaaattttaataaagg aatcagatcttcttcttgaactgaacagcaaatgaagaaagcctcagaagtgggaatcagcc atggcttgaattgtatgatatgagggaatatgatacttcaaaaatgaagcgtcaatatggaag gaaattgaagcgtcgaagaggcttctgattctgagaaatgaatttggtagttgcagaagatacat tggctctaaggagatatattttagagaccaatttaatttcaattataagaacatagtaattaaagtga ctaagcattcattgttttataacttttttctaaaaataaaactgtacaccagtttatfacctataa aagagaaftacacatgccaaatggaccaatgtccatttgcctatttggagggaagcctacataa gaagtcagagcctaccagaaatgtctttaaagcctggaacctgagcaaaagggaatagg tgggagtaattttttttaaattgtgaacaattcataagcacataatgattacagaataataaac attcatgtaccactatcagggttaagaaatagaacatttataatagtaggaatgttaagaata aaacatttaataagatctcagaagactccagtaaatctgcaattgtatctctctctttttaaagt aaatatactctgacttgttaattattcccttgcatttcttttagtttacttgcacacataatcttc aacataatatttaatttgaaaaaacctgaaaaaaaaaacctgttagcaagtataaaggggcag attactattattgcatgaaggcttcaagggaacgttacagcttttgggtcctatagctgcttca gcttctctcagagtttacagaggccaattttgagcaaatcatggtcaagggtatgagtgagtt ctgctaaacagaaggctcaccacaagggtatctggcaggattatactgggtagctgagatgtg cagaatgtgggttagagggaagttaaactgtttttgatgctcagcatgatgaatcaaacctctgt atcttaggattagggttaaacataaccttgggtatgatatgagtggtgtgctgaiccatgacga tggattggaaagcgtgggtataagcacacatgctaaagaaaaacatgttaattgtgctcact cacctggatatactgttctcaggttaaaaaatacagttactatcctaatacttgaaggcaactct cagcctatccattgagttacctcagatctgcccctctgggtccctagctgtcttgggactaacctct ttctcgcgtcagctgttttctggatccatgttttccattttattgagtactaacctgttttctgca gcacatcccttggtagcttctagagggaagtttgtgtggagggtaaaattttgagacctgcatgt ctcatgtttgattgatactttatcgttttaggtaggaggttaatttccctcaggactttaaaaatatt gttgcctcatttcttcttctattgtttattgagaatccaatgccattttgatttccccatcaaa attcatgatgatgtcttgggtgtgggtctatattatccattgtattgggttttaggtgaacctt ccagatagtaactcatttctcagttctgggaacacatgacattgggtgatatttattctctgc tgtcttcttctcccaactatttatttggatgttggatattccagcactgggtatctatttcttacctcc ctcccttgaacctcagctctgttttttagctctttagctcaatctccaaacttcttctattgtatttta aaatcttaagaccttcttctgtattgtagaagtcttcttcttacaacaaaaagcccttattctatg gattgttccacagataaagggttcaatatagtgtatttttttctattttaaattgtttgcgcact atttctccaaatttcttctgtatttatttttgggtgtctatatttctagacttttccaggataatctgataa tcttttggctgttcttctatgtgtgaagagggtactaaaaagccttggaaagcccttgggtgttggg aaggggctgtctttaggattatctgaatgggttttttgggagctccctctccacatgaatttt tgggttttgcagattccctagaatagaggcttccaatctctctctggagggtgtgtccaggga aggagattgtctagggtgtgtcagacagcagcttctcagctacttctctgacttttctactaatg attatatagtatcctaactactgtcaacaagtaataagataatctatcttctcactgttttagattattg ctgagataacctctcaaaagaaacctctcaaaaataaaaggtaacaagagccctatcttatattt ttcttcttcttctgtttagaagatagctattttaaacctgttcttctgtctgtgataaacacactt
--	--

		aafcttgtagaagtgtagatgggacagtatatitaggacctaagcctgcgaatgtatgat cagcttgtaagtacaggtgctcaaaaacatgtaacaatcagctttttactctgtaggaatc tttaaaattctgtgaatttttccccagaaagtaagcaaatctccccagaataaaatlaaatg tgcataatctaaagcctttttttattgtgtaggatatatataaaacataatttgcattgtaa acattttaatttacagtcagaggcatttaattacatcacatgttggaattattactactattc caaaattttctcatccccaactgaaactctgtaactgttagcaataacctcattctgtatc tctcccaacccaggttaacctcaaatctttttatctttgagacaaggctcattctatcactc aggtaggagtgtagtggtgtagcatagctcattgcagcctcaaatcctggctcaagcaa tctccttgtagtagtaagactatagggcacacattaactgcgctggctgattttgtttgtag agatgtggtctgtatgttcccatgctggtctgagttcctggcctcaagcagctccttaagatt catccatgttgtagcatgttcagaaattcattgtttttatgactaataatattccattgtatgat atacatittgttcatccatctctgtagaacactgggatagtctaccctttgctattgtgaataat gctgcagtaaacattgacataacaagtatgtattgattgctgcttctaaagtctttgggtatc atctgagtagaattgctagataatgtcatgtttttttctctgtgatttcttctgacccctgggt gagtggttaatttctacatgtttatgaatttccactgtttttgtttattgatttccaaagtccatcca ttgtgattagagaaagatacttagtatgttttaattgttttgagaaattgggtgtgctgctgtagat ggctgtcctggagaaatttctcctacacattgagcaaaatattatcatgctattgtgactgtat gtttctatgtctcttaggtcaagggtgtttacaatgtgttaagggtctcttttttaaaaaatlttt gcacagagatcttttctatgtgttccatgtattgtgtctttggagctatagctctgttagacag catatcactatctgtttgtttttttctgtccattctgccaatttctgcttttgattggaaaatt aatccatttgcatttaagtaattaaaggaaaggacttcttctaccatttaacacttctctatgtc atactatttttggccctcatttctctttatggccttcttttctgtttttgttagtaactagctgtat tctcttccactcccttgtgtatattgttagatgtttattgtgtgtgtctatgggattatagtta acatcctacacttaaaacaaictaatttaaacgtataccaatttaccctcaatagcatacaaaatc tctacictgttaagctctgccccgcccccttattgtattgtatggcaaaattgcttaataaat aatttatagttattgtatgagttgtcttttaaatcatttaggaaataaaagtgtagttagaaaac agtagatagtaataactgacttttatattgtcaatatattatcttattttggatcccttattcattatat agatttgagttactgttagtgcccttccatttgcggccaaaggattcccttattgacttcttgc gggcaagctcaattgtataaaactccctcagctttgttttactgagaatgtctgtatttccctt attttgattggataatttggccagatacatgaatttttggtaacagtattttcttccagcatttaaat atgtcatccactacccttgcactcatggttctcatgagatattagatgtataaaatttgaggaa ttcctcattctgtatgagtcagttctgttattgtcttctcgattgtctcagctttgtctttgacag ttgtattataacgcggctcagtggtgtctgtgatttatccacttagagtgttgagttctgtg gagtcatagatttatgtctttatcaaatttggacatatttggctattatttctcaattttttcactg cttcttcttcttcttctgaataattcttaattgtatatgttggctgtttgatgtgtctcaccagttctt aggctgtgttcttcttgttctcagactgtattattgcagttgccccttcttttttttcaagttgt tgtatttctccctgttcagatcaactgtgaactccctcagtgaaattatttgcagttactgtactttt cagctccaagatttatcttgggttcttttataacgtctgtgtctttatgtatttctcattttgtcat atgtctcttcttcccttagttcttgtccatgttttctttagctctttgggtattttaagacaattgtt taaagctttgcatagtaagccaatgtctgtgttctcagggaagggtttcattatttttttcaat gagccatacttctgtgtctttgtatgtcttctttgtttgaaaactgtatgttgaaacata acgtggtggccctgaaaacagatattcccccttctgagagttagttttattttatttgaag attgtagcagttctattgtctacatgtgcagtcatttccaactattttgcaagactgtatttcttct	
--	--	---	--

		gtgtgtcatcactgaagtcctgttcccttagtttgtgttaaatgtttgacatagatttccctgaaag gagttaaaactagcagaaaaatctctcccagcttccagcttcttagattggtctgtctg ggcttttccattacttagccaggctgtactgagcctaacaatcagcccaaaagcgtag ggctttgagatctgtctgagcatgcttctgtctgtatgcacgtagtttctaaatcctctgt atgtgtctgtgaatattctaattcccaagaaactccttgcagcttttctcacagaacatagat ggtttttggatcttgacctagcttctgacctagggtttgctgtgttagttcaccttacct ttttcaagcattgcttactgtctacgatgagtgctctgtcaatccttaagtagccccagacag gctaccagagacttaacaagaattgttaagtctgtctcagcttctctagaaatggggtatcag gggtcaagacagaaatgcagttgctgattcaagactgtgtcaacaccaggagctgtgtgg ggaagggaagcagaatgtcacaagaacttcttgcattttaagttgctgttctgactca gcattgtcttctgtatataaacttttactgttttcagagttctgataaaatggctatgctgtt ctgtcttaaaaaatataatatttttagggattgggtctcactatactgacctaggctgtctt gaacttctggcctcaagccatcctctcatttgcagcttcccaagtgctgcaattacacgcgtga accaccacaccagccctgtctgttttcaatgtgcttactccaccatgtgtctcaagtatgta tatttctaaactacctgttagttgtgtatgggaataaaatccctgagccttttgataaactcag agagatcaaaaacttagtttatcctattcgaaggattagaaaaatgatatcttctacatttctcag ggataggctcctcattagaaggctcctatgtccgatgtctgtacaagacatttcttcttfaat gtttacaacaagctgtgtgccaaggctgacttgaactcctggcctcaaacgatcctccagct cagcttcacaaagtgtgggtgtgtctggccaactaatgactatcttaactctgtgtttcaatgtt atgcttcttttactgtactgtatgtatgactatgttcttagaacaatgtgaacagaaatgtgt agagcagacatcctgtcttataatttaccattatataatgtatgttaggtatagattttctcacag atgcttttatcagattgaggaaattatattcctacttgcgaaagggtttgttagtatgaggggg tgtctgaatttgtcaaacacttttctggtaataatgagatgtatgttctgcagtcacgagatgt ggattttctcttattctgttctgtgagtgattacactgggtgactaatgttaaaacaacttacttt ccaggaaataaacctattatctttttatata	
PSMC4	NM_153001. 2	tgcgggtacggacagcgcatgagcttatgttagggcgagcccagaccagccctctgtc ctatcctgcccctccagcactctcagccgtaactaaactacacttcccagaaagcctctcag ccagggaacttccgtgtgtctcagcggaaagcgggtgacagatcatcccaggccacacagagg ccggctgtgtcactatgtgaggagataggcactctgtgtggaagggtcaggatgagatccc agcactgtccgtgtcccgccccagaccggcctgtccttctgtggccctgagcctgaggac ctggaggacctgtacagccgtacaaggaggagggtgaagcgaaatccaaagcatcccctgt gtcatcggacaatttctgtaggctgtgtatcagaatacagccatcgtgggtcttaccacagg ctccaactattatgtgcgcatcctgagcaccatcgatcgggagctgtcgaagcccaacgcct cagtggcccctcccaagcacagcaatgcactgggtgacgtgtgtgtgtgtgtgtgtgtgtgtgt gcagcatcatggtgtcactcagaccagaagccagatgtgtgtgtgtgtgtgtgtgtgtgtgtgt gcatgtgacatccagaaagcaggagggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt agctctacaagcagatcggcatcgtatccccccgaggcgtcctcatgtatgtgtgtgtgtgtgt ctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt gtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt tgttccgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt ccaccaagagattcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt agctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	16

		acagagcagacacccctggatccggccctgctacggccaggacggctggaccgtaaaattg aatttccacttccctgaccgcccgagagagattgatttctccactatcactagcaagatgaa cctctctgaggaggttgacttggaagactatgtgcccggccagataagatttcaggagctg atattaactccatctgtcaggagagtggaatgttgctgtccgtgaaaaccgctacattgtcct ggccaaggacttcgagaagacatacaagactgtatcaagaaggacgagcaggagcatg agttttacaagtgaaccttcccttccctccaccacaccactcaggggctgggcttctctcgc acccccagcaccctgttcccaaaacctcattcccttttttcttaccaggattggtttctcaata aatagataagatcgaaatccatttaatttcttcttagaagttaactcctttgagaatgtggcctt gaataggatccctgtgggtccctcttaatctgacagatgagcagacgaggtgcatggcctggg ttgcagcttgagagaaccaaaatattcaaacagatgacttccaaaatgtgggaaaggat ggaaaatgaacctgagatggagtccttaacacgggataaagccctgtgcatctccctcattt cctacaggtaaaagacagtaaaagaaatcaggtcacaggccttgggagttcataggaaagga gatgtccagtgctgtccagtagaacttt	
SF3A1	NM_005877. 5	gggtccgggaagtgcgcccagtcgtaccttgcggccgcaactcgtcggccgcccacatctt gcgagctcgtctactgaccgagcggggaggctgtcttgaggcggccaccgctcaccgaca ccgaggcggactggcagccctgagcgtcgcagtcatgcccggccgaccgtgcaggcgc gtgccccggccgcccgtgcccacggagcccacagcccacagaagaagaagcat cttcaaaaggaggattctgcaccttcttaagccagttgtgggattatttacccttccacagggt cagaataattgttgacaagactgcagcttgtgtggccagaacgggcccgaatttgaagcta ggatccgacagaacgagatcaacaaccccgaattcaactttctgaaccccaatgaccttac catgcctactaccgcccacaaggctcagcaggtcaagggaaggagcctcaggagccgtcc ggcccatcccccaaggctcagcagcagcagcagaccaccagcagcagctgccccca gaagggtccaagcccgaagtaatccaagagaccatcgtgcccgaaggcctcctcctgagttt gagttcatgtctgacctcctctatctcagccttcgacttgatgtgtgaagctgacggctc agttgtggccaggaaatggcgccagtttctgaccagctgafgcagaagaagcagcgcga ctaccagtttgactttcggcccacagcacagccttctcaactacttcagaaagctagtggaa cagtaacacaagatctgattccacccaagggttattttcaaagctcaagaagaaggctgaa aacccccgagaagttttggtacaggtgtgttaccgagtggaatggccaaattccaggaaac gtgagagggaagaaggaagaagaggagaaaggagaaaggagcgggtggcctatgtctcagat cgactggcatgattttgtgtgtgtgaaacagtggaacttccaaaccaatgagcaagggaact tccctccccccaccacgcccaggagcgtggggggcccgaatcctcattcaggagcgtatg aaaagtttggggagagtgagggaagttagatggagggtcagctctgagaggagatgaca aacaggagaaggcggaggagcctccttccagctggaccaggacacccaagtaacagat atggatgagggttcagatgaaagaagaaggcagaagtgccccacccccagagac acccatgccctccacctctgcccccaactccagaccaagtcattgtccgaaggattatgatcc caaagcctccaagcccttgcctccagcccctgtccagatgagtatctgtgtccccattact ggggagaagatccccgcccagcaaatgcaggaaacacatgcgcatttgacttcttaccctc gctggctggagcagcggggtcgtccatccgtgagaagcagagcagatgaggtgtacg caccaggtctggaattgagagcagctgaagcagttggctgagcggcgctactgacatcttc ggtgtagaggaaacagccatttgtaagaagatcggtaggaggagatccagaagccaga ggaaaaggtagctggatggccactcaggcagcatggcccgaccagcaggctgccc aggccaacatcacctccaggagcagattgaggccattcacaaggccaaggcctgtgtgc	17

		cag agg atg ac act aag ag aag at tgg ccc cag ca ag ccc aat ga aat ccc t ca ag c cacc gcc acc at ttc ag cc acc aac at ccc cag ct ggc tcc acc ca t ac ttc ag tgc cc cg acc acc caca at tgc cacc tcc ag ttc tgc t ac t ac ag ttc tgc cgc ag t acc cgc t at gccc cgg ccc cca at tgg cat ct g tgc tcc ggc tgc ccc cag gct ag tgc at g ccc ccc at g cgc cc at c at cc ac g cgc ccc ag aat ca ac g tgg tgc ccc at g cct ccc tgc gccc tcc t at t at g gcccc cgc ccc acc ccc ccc at g at tgc tgc aac ag cct tgc tgc tgc tcc acc tgc tgc cacc tgc tcc ag c tcc ag ccc ca at g ccc cct g tgc at ccc ccc acc tcc at g ga ag at g ag ccc acc tcc aaaaaa ac tga ag ac ag ag g ac ag cct c at g cc ag ag g ag g ag ttc tgc g cag a aac aagg g tcc ag tgc at caa ag tcc ag g tgc cca aac at g cag g ata ag ac g ga at g g aa act g aat ggg cag g tgc tgg tct tacc cct ccc act c ac g g acc ag g tct tgc t at taa g gtg aag at tca tga ag cc ac ag g cat g cct g cag g ga ac ag aag ct ac ag tat g ag ggt at ttc at caa ag at tcc aact c act ggc t t ac t aac at g cca at g cgc ag t ac tcc ac ct ggc cct caa g g ag ag ag g cgg ag g aag aag tag ac aag ag g aac ct gct g t ca ag t ccc tgc ca ttt tgc tct cct g tct ccc acc ccc tgc ccc cag acc cag g ag ccc ccc tgc ag g ct tgc cct g cct g cat at ttt tgc tct t ac t ag ttt ggg aat caaat tgc tct g cag ag gtt cat tccc tgc acc ctt tccc ac at tgc ta ag ag tag ct ggg ttt tca ag c cact tct g g aat ct ttt g t ttag g tgc tgc at tgc ag g ac at tct ttc tgc ag ccc at tgc aact g ag ag ccc ag g g at g tct ac ag g at ag ttt cat ag tgc ac ag g tgg c act tgc taa tag aat at ggc tg at at tgc taa t ac at ttt tgc ttt g ac at g ggt tgc taa taaa act cgg acc ctt tgc tga aat cag tta aat aag act tgc tgc gtc acc tgc ccc tgc cag act cag g cag tgg ta ac act g cag ag tgc at g tgc tct ttt g ag g at ttt g ggt ttt g taact aaat tct tgc tgc cct c at act ttt at g t at t ag aat ca t at tgc t at tgc cct ttt aaaa ac at tgg g at ccc caa ag g cct gcccc at g t at t aac ag ta at ac ag g aag cat ggc ag g cacc at g caa acc aag g at g ga tgg tgc cag tccc tgc tgc ag tgg cgc g tgc gtt cct gct ggc cct gga at cact c at cct g at tgc at tgc tgc tgc tgc tgc ggc ag g tgc cct c at ag g tgc tgc g at at g at g ac gtt tct taaa at g t at g t at tta aca aat ac tta at tgc taa t ag g tca tgc tacc aag g at tgc taa taa t aat t t ac tct c act ttt at c at ttt at c at ttt aact c at g taat cct c at g tgc at tct cgt tta ac act g ag taa act g ag g c ac ag ag aac ata ag tgc at g ccc at ag tgc ac ac act g t gaa tgc aaaa g ag aat g tgc tgc aaaa ac ag tgc ac ag tgc tgc tgc tgc ag taa ag g cag gct g t at ctt tga aat ca ag ct ac ac ag g g ag at ag g ca at g ct g tgc g tgc tgc ag g aag g tgc ag ag cct g tgc tgc aac at tct g g ttt aag cta ag gct g at ttt at tgc gga ag at c ac at g t tgc tgc ccc tgc ag ag ttc ccc ag tgc cct ttt at tgc cag tcc tcc at tgc g acc tcc tag ct g c ccc at cag g tgc tcc cag ggc tgc ag ag g g tgc ag acc at tcc ca ag g tgc ac ag aacc ag ct c tgc tgc tgc ccc tgc cct cct cct cct ccc cag ag tgc ag tacc ag ttt at g ctt at t ac aa act gct ggc tgc cct ccc at ttt ca act g at tgc at g g g at g tgc at cct t at cct g tgc tgc at tgc cct ggc cct g g tgc tgc aag ttt g ccc cag g g c aag ag tgc aat ttt g cct cct g ag tgc ggc ttt g tgc ttt g cgt ccc ta aag tgc ag ccc act act g tgc tgc tgc tgc ccc at g cca ac ac cag aat ccc cct g ag cact ac cct ggc tgc at tcc aag aag g aag ag g tgc ag g ag ac cct g gg g ag tgc at at tcc aag tct cct tct tct tgc g g ag cag tgc ggc ag tgc at g tgc tgc ggc act ccccc ac ag act ggc aaaa ccc tgc cag g act tgc g tgc g tgc ag g c tgc acc g g ag gccc ag g aat g cgc tgc g g tgc at tgc ag tgc aat g g g ccc tgc ag cgc cct cgc ag ag c aa at cag ttt cct g ag c tgc tgc aat g aat at c tgc tgc tgc cgc tgc ag tgc cag aag ctc at tgc acc at g ag g cct tgc tgc aac at cag tgc ccc cag gccc ag tgc tgc tgc tgc cct ggc act gca
--	--	---

		ctagttaggacctgcagcatgtaggttagcgctcctagtgttataatacaaaagctgctctgcac agctttctgattctctgcaatctcctgaggattatctgccccattttaaaacgagggtggaata cccaagggtcatgtagccagtgagtgctctggaaagccaaagcagctcatcccttctgggg accacactgctctgcccaccagaccacactatgaaataggaaatagtgctcctgtgtagg actgctgggaaaacagggtgggtgggacttaagtcaccataatttgaagactgtcatgcaga gggctccaggaaatgttagacattaaggaaattcactttagtctaccactacttaagtactgt catgtactcttagaggaggccagtaatgacagaaccattttactttaaaatataatattgtatt agagaataataaatgggtatattgggttatgttaggatatatactgaaatggaatacatgtact attagcaatcatatttcatctcctgtaattagacaagaaagcataatagctctactcatgg gtacacataccagtgtataagatttttagaagttactttttaaataaaaagcaaatgttaagat cttaaaaaaaaaaaaaaaaa	
PUMI	NM_001020 658.1	agtggccgcccattgttgcaggagtgaaaggtaagggggagcagagcggccagagagag aagatcggggggctgaaatccatcttcatcctaccgctccgcccgtgtgttggaatgagcg ttgcatgtgtcttgaagagaaaagcagtgctttggcaggactctttagccccacctgaac atcacctcaagaaccagctaatcccaacatgctgtgtgtttgacatctggaaacagggtcgc aagcgcagccacaaccagctgcaaatcaggctcttgcagctgggactcactccagcccgt cccaggatctataggagttgcaggccgttcccaggacgacgctatgttgactactctttca gaggcagcatgtgtgagcagcttggggaggaggaaatggaggaggcggctataataata gcaaacatcgtatggcctactggggataacatcagcagaacatcaggtgctgtccatggat gaactgaatcatgatttcaagcacttgccttggagggaagagcgtgggagagcagctctt gccaggtaaaaagtgttgggaaacagatgaatcagcaagatggaccanaaggaaatffc ctgggtgatcaatggcagacagtgcttgggaaacatcagatcattcagttcccagccaat catgggtcagagaaacgtggtcagagttccatgtgaacagtgaggtaattctgtactgtc cccacgaticggagagtgggggactaggcgttagcatgttgagtatgtgttgagctatcc ccggggcgtattcgtctaaagaaaaggaggatttggcccagggaatgcagacagtatgaaa acgacaagggtgaaaagaagaacaagggtacgtttgatggagataagctaggagatttga ggaggagggtgatgtatggacaagaccaatggtttaccagtgcaaatgggattgatga gacgtcaaaagatttagccgtacccttggttaattgccagaactctgctaatgaatggatctc tgggtccaaaccagaatgtcttagggcttagcccagctgaccagcaccatgtgtccaa gccgtgtggaggatttccaaacatggagtcccagagtgtccccttggaccccattggaacatg tgggcatggagccctctttagttgattttagggcacgagggtacctgtggactcagcagcag caactgtgggactttttagctacaatttcaacaacagctgttccaaagacctaatgcgttct gtccagcagttgacagctgctcagcagcagcagtgatgactggcagctgctatcagccgc acatcgggttagctccgctgctgttggccccaatccatacatcagcgtgctccccag ggacggacccctacacagctggattggctgcagcagcagcagcagcagcagcagcagcag ctcaccagtattatggagtactccctggggagtctaccctgcccagctcttccagcagcaagc tgcctgctggcgtgcagcaactaattcagctaatcaacagaccacccacaggctcagcaa ggacagcagcaggttctccgtggaggagccagccaacgtctttagccccaaccagaac cagcagggacagcaaacgatcccttggcagctgcagcagtgaaatctgcccctgcat tggacaaggctgtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc ccaaactggtgcccctgtatgaatgcagcggcagaaatggcttggagctcctgttgcact gttagctcctgccccagctatcatttagttcctcagctgcacaagcagctgtgcagcagccgc	18

	agcttcagcaaatggagcagctgggtcttctgg aacaacaatggaccatttcgccctt aggaaacacagcagcctcagccccagccccagcagcagcccaataacaacctggcacc gttcttctacggcaacaactctctgaacagcaattcacagcagcctcctcttcccagg ctctgccagcctgccaaacacatcctgggaftcggagtagcagttctctcggcgccacc tgggatccgcccctggagggtttggaaacagcagttgcaactccaacactggcagtgctc ccgctgtactccctgactggcagcagtgacctttataaggagacatcagcagcttgacc cccattggacacagttttataacggccttagcttttctcctctcctggaccggtggcatgcc tctccctagtcaggaccaggacattcacagacaccacctcctcctctcttccatgac ctctcaagcttaacctggaggactcacgaatggcagtggaagatacatctctctctcc aggcgtgaagccaagtaccgagtgcaagcagcgccctccagcctctcagcccaggca gactcttttctctcctctctgttgcgatatggaatgtctgctcagccttctggcagga ggcctttggagattttcgaacaaccgggtaccccaatttacaactgcgggagatgctggac atataatggaaattttccaagaccagcagtggtccagattcattcagctgaaactggagcgtg ccacaccagctgagcgccagctgtcttcaatgaatcctccaggctgacctaccaactcag gtgcatgtgtttggtaattacgtcattcagaagtctttgaatttggcagtcctgaacagaagctg gctttggcagaacggattcaggccacgtcctgtcattggcactacagatgtatggctgcgt gttatccagaagctcttgagtttattcttcagaccagcaggttaattatgagatggtctggg aactagatggccatgtctgaagtgtgaaagatcagaatggcaatcacgtgttcagaaat gcattgaatgttacagcccagctcttgcaattatcagcagctttaaaggagacaggtattt gccttatccacacatccttatggctgcccagtgatcagagaatcctggagcactgtctccctg accagacactccctatttttagaggagcttcaccagcacacagagcagctgttacaggatcaa tatggaaattatgaatccaacatgtactggagcagggctgctcctgaggataaaagcaaat gtagcagaataccgaggcaatgtactgtatgtcagcacaaatttgaagcaatgttgtg gagaagtgtgttactcacgctcagctacggagcgctgtgtctatcagatgaggtgtgac catgaacgacgggtcccccacagtgccctatcaccatgatgaaggaccagatgccaaactac gtgtgccagaagatgtacgtgctggagccagggcagcgaagatcgtcatgcataag atccggccccacatcgcaactctctgaagtacacatgtggcaagcacattctggccaagct ggagaagtactacatgaagaacgggtgtgacttagggcccactgtgtgccccctaatgtga tcactgtaggcaggtcaccggctgttccctcattcccgctgacctcactggcccactggc atccaaccagcaaccagaaatgttctagttagagctgagacgggcaagtgtgtctccag gattactccctcctcaaaaaaggaaatcaatccacgagtggaagcctttgtaaatttaatt tattacacataacatgtactatttttttaattgactaattgccctgtgtttactgtgtataggat actgttacatagtaaccaatgtacatgggaggccacataattgttctactgtgtatctatattc acatgtggaaactttcagggtgtgtgttttaacaaaaaaaaaagcttaaaaaaaaaagaaa aaaaggaaaagggttttagctcatttgcctggccggcaagtgttgaatagctcttccccacc tctcatttttagtaaaaaacaacaaaaacaaaaaacctgagaagtgtgaattgtatgaat gaacccaaactggcatttaacactgtttataaaaaatataatataatataatataatga aaaggtttcagagttgctaaggctcagttgtgacattaaagttaataatctaaaaaatgcct ttttggagactatattatgtgaagaaggctgtcgtgaggaggagatgcgagcaccagga acgtctttgagcgtggcggtgtgattgttactgcctactggattttttctattaacattgaa aggtaaaatcgtattatttagcatgagaaaaaaaatccaactctgctttgtcttctata aatatatagtgtatcttggtgtgactttgcataatacaaatgttagtatttctgtttgtatg taatctgtatctataatgtacccatgtagtcgaacatactttgtattgtacaattgtacattgtata	
--	--	--

		cctgtaatgtaaatgtggagaggtttgaatcaacataaacacgttttttggaagaaaagagaa ttagccagccctgtgcattcagtgatattctcaccftttatggctgtagcatatagtggtatatt gtaaatgttaattcaaccagaagtaaattttttctttgaaggaaataatgttctttatcacgcct agttaatgtttaaaaagaaaaaatagcttggtttatgttcaictagictcaagtatacgagga ttcttctaaatgtattcaagattgagttctactagtggttttaacctaataaagtaattttg attttgtacagtcaaaagacgtgcaaaagtctagccttgcctgagcttcttacaatcaga gccccctcaccctgtaaagtgtaacgccttccctttgtacagaagatgaactgtattttgc attttgtactgttaagtgtaacatactgtcaattttcctgtttgaatatagaattgtaacac tacacgggtgtacattccagagccttggttatattccaatgaactttttgcaagcacactgtaa ccatatgtgtataaatacaaacctgtgtatgtctatgcctgggcaactattttgttaactctgtg tagattgtcttaacaatgtgtatctttattttgaaaaatacagaactttggaaatctgaaaaaa aaaaaaaaaaaaaaaaaaaaaa	
ACTB	NM_001101. 4	gagtgagcggcgcgggggccaatcagcgtgcgcccgttcgaaagtgtcctttatggctcga gcggccgcgcgcgccctataaaacccagcggcgcgacgcgccaccaccgcggaga ccgcgtccgccccgcgagcacagagcctgcctttgcccgtccgcccgcctccacacc gcccacagctcaccatggatgatatcgcgcgctcgtcgtcgaacagcctccggcat gtgcaaggccggtctcgcggcgacgatcccccgggcggcttccccctccatgtggg gcgccccaggcaccaggcggtgaltggtggcatgggtcagaaggatcctatgtggcgga cgaggcccagagcaagagaggcctcctaccctgaagtacccatcagacagcgcctcgt caccaactgggacgacatggagaaaatctggcaccacaccttctacaatgagctgctgtg gctcccaggagcaccctgtctgtgacgagggccccctgaacccaaggccaaccg cgagaagatgaccagatcatgtttgagaccttaacaccccagccatgtactgtctatcca ggctgtgtatcccgtacgctctggtgggtaccactggcatgtgtgactccgggtgacg gggtcaccacactgtgcccactacgagggtgatccccccatgccatcctgctgtg gacctgtgtgcccggacctgactgactacctatgaagatcctaccgagcgcggttaca gcttcaccaccacggccgagcgggaaatcgtgctgtacattaaaggagaagctgtgtactgt cgccctggacttcgagcaagagatggccacggctgtctccagctcctccctggagaagag ctacgagctgctgacggccagggtatcaccattggcaatgagcgttccgctgcccgtgag gcactctccagccttctcctggtgcatggagctctgtggcatccacgaaactaccttaact ccatcatgaagtgtgacgtgacatccgcaaaagacctgtacgccaacacagtgctgtggtg cggcaccaccatgtacctgtgcatgtccgacagatgcagaaggagatcactgccctggc acccagcacaatgaagatcaagatcattgtctcctcctgagcgcaagtactcctgtgtgacg gggctccatcctggtcctgtgtccacctccagcagatgtggatcagcaagcaggagat gacgagtcggccccctccatgtccaccgcaaatgtcttagcggactatgacttagttgc gttacacctttcttgacaaaacctactgtcgcagaaaacaagatgagattggcatggcttta ttgtttttttgtttgtttttttttttttgtgctgactcaggatitaaaaactggaacgggtg aaggtgacagcagtcgggtggagcagcattccccaaagttcacaaatgtgcccaggactt tgattgcacattgtgttttttaatagtcattccaaatagatgagcgtgttacaggaaagtcctt gccatcctaaaagccacccactctcttaaggagaatggcccagtcctctccaaagcca cacaggggaggtgatgcatgtcttctgttaattatgtaatgcaaaatttttaactctgcct taatactttttttgtttttttgaatgtagccctcgtgccccccctccccctttttgtcccc caactgagatgtatgaaggctttgtgtccctgggagtggtggaggcagccagggtcta	19

		cctgtacactgacttgagaccagttgaataaaaagtcacaccttaaaatgaggaaaaaaa aaaaaaaaa	
<i>GAPD</i>	NM_002046. 6	gctctctgctcctcctgttcgacagtcagccgcatcttcttttgcgtcgcagccgagccacat cgctcagacaccatggggagggtgaaggctcggagtcacggatltggctgtattggcgcc tgggtaccaggctgcttttaactctggtaaagtggatattgtgccatcaatgaccccttcatt gacctcaactacatgggtttacatgttccaatatgattccaccatggcaattccatggcaccg tcaaggctgagaaagggaagcttgcacatgaatgaatcccatcaccatctccaggagcga gatccctcctcaaatcaagtggggcgatgctggcgctgagtacgtcgtggagtcactggcg tctcaccaccatggagaaggctggggctcatttgcaggggggagccaaaagggtcatcat ctctgccccctctgctgatcccccatgttcgtcatgggtgtgaaccatgagaagtatgacaa cagcctcaagatcagcaatgcctcctgcaccaccaactgcttagcacccttggccaag gtcatccatgacaactttggtaicgtggaaggactcatgaccacagtcacatccatcactgcc accagaaagactgtggatggccctccgggaactgtggcggtatggccgaggggctctc cagaacatcatccctgcctctactggcgctgccaaggctgtgggcaaggctacccctgagct gaacgggaagctcactggcatggccttccgtgtccccactgccaacgtgctagtggtgac ctgacctgcccgtctagaaaaacctgccaatatgtagacatcaagaagggtggtaagcagg cgctggaggggccccctcaagggtcctctgggtacactgagcaccagggtgtctcctctga cttaacagcgaacccactcctccacctttagcgtggggctggcattggccctcaacgacc actttgtcaagctcatttctgtgtatgacaacgaatttggctacagcaacagggtgtgtgacct catggcccacatggcctccaaggagtgaagaccttggaccaccagccccagcaagagca caaggaggaaagagagagacctcactgctggggagtccctgccaactcagtcccccacca cactgaatctccctcctcacagtggccatgtagaccttgaagaggggaggggcctagg gagccgcaccttgtcatgtaccatcaataaagtacctgtgtcaaccagttaaaaaaaaa aaaaaaaaa	20
<i>GUSB</i>	NM_000181. 3	gtcctcaaccaagatggcgcgatggctcaggcgcatcacgacaccggcgctcagcg accgcctctacgggcacctcccgctgtttcttagcggcgagacgggtggccgagcgggg gaccgggaagcatggcccggggctggcggttgcctggcgcgctcgggcccgtgtt tggggctgctggctggggctgcagggggggagctgttaccgccaggagagcccgtcgg ggagtgcaggagctggacggcctctggagcttccgcccgaacttctctgacaaccgacg ccggggcttcgaggagcagtggtagggcgccgctgtggagtcagccccaccgtgg acatggccagttccctccagcttcaatgacatcagccaggactggcgtctgcggcatttgcg gtcgggtgtgtgtacgaacggggagggtatcctgcgggagcgtatggaccaggacctgcgc acaagagtgggtgtgaggttggcagtgcccattcctatgccatgtgtgggtgaatgggt cgacacgctagagcatgagggggctacctccccctcaggccgacatcagcaacctgt ccagggtggggccccctggccctccggctccgaatcactatcgcacatcaacaacacactacc cccaccacctgcccaccagggaaccatccaatcctgactgacacctccaagtatccaaagg gttactttgtccagaacacatatttgaacttttcaactacgctggactgcagcggtctgtacttct gtacacgacacccaccacctacatcgtatgacatcaccgtcaccaccagcgtggagcaaga cagtggcgctgtgaattaccagatctctgtcaagggcagtaacctgttaagtgtgaagtgc gtcttttggatgcagaaaacaagtcgtggcgaaatgggactgggaccaggggccaacttaa gggtgccagggtcagcctctgtgtggcgctacctgtagcacgaacgcctgcctatctgtatc attggagggtcagctgactgcacagacgtcactggggcctgtgtctgacttctacactcc	21

		ctgtggggatccgactgtggctgtcaccaagagccagttcctcatcaatgggaacctttct atftccacggtgtcaacaagcatgaggatgcggacatccgagggaaggcgcttcgactggcc gctgctggtgaaggacticaacctgcttcgctggcttggccaacgctttccgtaccagcca ctacccctatgcagaggaaagtgtatgcagatgtgtaccgctatgggattgtgctatcgtatg agtgtccggcggtggcctggcgctggcgagttctcaacaacgtttctctcatcaccaca tgcaggtgatggaaagtgtgtcgtagggaacaagaaccaccccgggctgtgatgtgtgt ctgtggccaacgagcctgcgtcccacctagaatcgtcgtgctactactgaagatgtgtatc gctcacaccaaatccttggacccctccggcctgtgaccttgtgagcaactctaactatgca gcagacaagggggctccgtatgtgtgtatcgtttgaacagctactactcttggatcac gactacgggcacctggagtgtatcagctgcagctggccaccagtttgaactgtgtataa gaagtatcagaagcccattatcagagcgagtatggagcagaaaacgattgcagggtttacc aggatccacctctgtatgttctacgaaggtaccagaaaagtctgtatgagcaggtaccatctg gggtctggtatcaaaaacgcagaaaatcgtgtgtggagagctcatttggaaatttgcggtttca tgcactgaacagtcaccgacgagagtgtcggggataaaaaggggatcttctactggcaga gacaacaaaaagtgacgcttcccttttgcgagagagatactggaaagtccaatgaaacc aggatccccactcagtagccaagtcacantgtttgnaaacagcctgtttacttgcagcaaga ctgataccacctgcgtgtcccttctcccgagtcaggggcacttccacagcagcagaaca agtgcctcctgactgttcacggcagaccagaacgtttctgcccgggttttctgtctatctatt ctagcagggaacactaaagggtggaaataaaagatttctattatggaaataaagagttggcat gaaagtggctactgaaaaaaaaaaaaaaaaaaaaaaaaaaaa	
RPLP0	NM_001002. 3	gtctgacggcgatggcgagccaatagacaggagcgctatccgggttctgtattggcta ctttgttcgcattataaaagcagcgcgggcgaggcccttctcgcgaaggcgctcctg tggaaagtacatcgtctttaaacctgcgtggcaatccctgacgcaccgcccgtatgccag ggaaagacaggcgacctggaaagtccaactacttcttaagatcatccaactattggtatgta tccgaatgtttcattgtgtggagcagacaatgtgggtcccaagcagatgcagcagatccgca tgtcccttcgcggaaggctgtgtgtctgtatgggcaagaacaccatgatgcgaaggccat ccgaggggcacctggaaaacaaccagctctggagaaactgctgcctcatatccgggggaa tgtgggctttgtgtcacaaggaggacctcactgagatcagggacatgtgtctggccaataa gggtgccagctgtcccgtgtgtgtgctgctgctgctgctgctgctgctgctgctgctgctg aacactggctcggggccgagaagaccctcttttccaggcttttaggtatcaccactaaaaatc ccaggggcaccattgaaatcctgagtgatgtgcagctgtatcaagactggagacaaagtggg agccagcggaagccacgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt ccagcagggtgttcgacaatggcagcatctacaacctgaagtgtgtgtgtgtgtgtgtgtgtgt ctctgcattctcgtctcctggagggtgtccgcaatgttccagtgctgtgtgtgtgtgtgtgtgt ccaactgtgtcatcagtagcccatctatcatcaacgggtacaacagagtcctgtgctgtgtgt tggagacggattacaccttcccacttgcgtgaaaaggtaaggccttctgtgtgtgtgtgtgtgt cctttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt cagctaagggtgaagccaagggaagagtcggaggagtcggagcaggatattgggtgtgtgt tctttgactaatcaccaaaagcaaccaacttagccagttttatttgcacaaacagggaataaa ggcttacttctttaaagtaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	22

TFRC	NM_003234. 3	agagcgtcgggatafcgggtggcggctcgggacggaggacggcgtagtgtagtcggg cttctagaactacaccgacctcgtgtcctccctcctcgtcggggcgtgctggagcggcc gctccggtgctgtccagcagccatagggagccgcacggggagcgggaaagcggtcg gccccaggcggggcggccggatggagcggggccgcgagcctgtgggaaagggcgt gtggcggcgcctcagcggcgtcaggttcttctgtgtggcagttcagaatgatgaagc tagatcagcattcttaactgtttgtggagaaccattgtcalatacccggttcagcctggctc ggcaagtagatggcgataacagtcagtggagatgaaacttgcgttagatgaagaaaa tgctgacaataacacaaaggccaatgtacaaaacaaaaaggtagtggaagtatctgct atgggactatgtgtgtatcgtcttttctgtatggattatgtatggcctactgggctattgtaa ggggtagaaacaaaactgagtgtgagagactggcaggaaaccgagtcctcagtagggga ggagccaggagaggacttccctgcagcacgtcgttataatgggatgacctaaggaag ttgtcggagaacactggacagcacagacttcaccggcaccatcaagctgctgaatgaaaatc atatgtccctcgtgaggctggtatcctcaaaaagatgaaaatcttgcgtgtatgtgaaaatcaat ttcgtgaatttaactcagcaagctgtggcgtgtatcaacatttgaagattcaggtaaacac agcgtctaaaactcgggtgatacatagttgataagaacggtagactgtttacctgggtgagaat cctgggggttatgtggtgtatagtaaggctgcaacagttactggtaaacctggctccatgctaatt ttggtactaaaaagattttgaggattatacactcctgtgaatggatctatagtgattgtcagag cagggaaaaacacatttgcagaaaagggtgcaaatgtgaaagcttaaatgcaattgtgtgt tgatatacatggaccagactaaatttccattgttaacgcagaactttcattctttggacatgctc atcggggacaggtgaccttacacacctggatttcccttccctcaatcacactcagtttccacc atctcggtcacaggttgcttaataacacacagcaaatccagagctgctgcagaaaa gctgtttgggaatatggaaggagactgtccctcgtactggaacacagacttacaatgtagg ggaaccacagaaagcaagaatgtgaagctcactgtgagcaatgtgtgaaagagataaaa attcttaacatctttgaggtattaaaggctttgtagaaccagatcactatgtgtgtggggcc cagagagatgcatggggccctggagctgcaaaacccgggtgtaggcacagctcctattga aacttcccagatgttctcagataggtcttaaaagatgggttcagcccagcagaagcattat ctttgccagttggagtgctggagacttggatcgggttggtccactgaatggctagagggata ccttctgcctcgtcatttaaaaggcttctacttataatctgtgataaagcgggtctgtgtaccag caacttcaagggttctgcagcccactgtgtatagccttattgagaaaacaatgcataaatgtg aagcatccgggttactgggcaatttctatatacaggacagcaactggggcagcaaatgtgaa actcactttagacaatgctgcttcccttctcgtatattctggaaatccagcagtttcttctgttt ttgcgaggacacagattatccttatttgggtaccaccatggacacctataaggaaactgattga gaggatctcgtagtgaacaaagtggcacgagcagctgcagaggctcgtggtcagttcgtg attaaactaacccatgatgtgaattgaacctggactatgagaggtaaacagccaactgcttt catttggaggatcgaaccaatacagagcagacataaaggaaatggcctgagtttacag tggctgtattctgtcgtggagacttctccgtgtacttccagactaacacagatttgggaa tgctgagaaaacagacagattgtcatgaagaaactcaatgacgtgtcatgagagtgaggt atcacttctctctccctacgtatctcaaaaggatctccttccgacatgtcttctgggctccg gctctcacacgctgcccagctttactggagaacttgaactgcgttaacaaaataacgggtgctt ttaatgaacgctgttcagaaaccagttggctctagctacttggactattcagggagctgcaa atgccctctctgtgacgtttgggacatgacaatgagtttaaatgtgataccatagcttcca tgagaacagcagggttagtctgtttctagactgtgtgctgacgtgctaattttagtagggct acaaaaccgtatgttaaaattccatcccatcttctgtactactagatgtcttttaggcagcagc	23
------	-----------------	---	----

		tttaatacaggtagataacctgtacttcaagttaagtgaataaccacttaaaaaatgtccat gatggaaatattccccctatcttagaatftaagtgccttgaatgggaactgcctcttccctgtgtt gftaatgaaaatgtcagaaccagttatgtgaatgatctctctgaatcctaagggtctgtctctg ctgaagggtgtgaagtgtcgtcttactttgagtgatccccaacttcatttgatgctaataaggag ataccagggtgaaagaccttctccaaatgagatctaagccttccataaggaaatgtagctgggt tctctattcctgaaagaacagttacttcagaagagatgggtgttttcttccaatgagggt ctgaaatggaggtcctctgtctgataaaatgaggtcaactgtgtatgtaggaaataggcc ttaatatgttaacctagtgatattatgaaaaggaggaccagaagccaaagacttagtatat ttcttttctctgtcccttccccataagcctccatttagtcttgttattttgttcttccaaagca cattgaaagagaaccagtttcagggttttagttgcagactcagttgtcagactttaagaataa tatgctgcccattttgtccaaagtgttaatcttagggagagccttctgtcttctgtgcactga gataattatgtttattatcagtgacagagttcactataaatgggtgttttttaataagaataaattat cgggaagcagtgcttccataattatgacagttactgtcgttttttttaataaaagcagcatc tgcataataaaacccaacagatactggaaagtgttgcatttatggtcaacacttaagggttttagaa aacagccgtcagccaaatgtaattgaataaagtgaagctaagatttagagatgaattaaattt aattagggtgtgtcagaagcagcagcctgaccagataagaatgctgtgttctctaaatgcagt gaattgtgaccaagtataaatcaatgtcacttaaggcgtgtgtagtactcctgcaaaatttat agctcagtttatccaagggtgaactctaattcccatttgcataaatttccagttacctttgtcacaat cctaacacattatcgggagcagtgcttccataatgtataaagaacaagggtgttttaccctacc acagtgctgtatcggagacagtgatctccatattgtacactaagggtgtgaagtaattatcggg aacagtggttccataatttcttcatgcaatgacatcttcaagcctgaagatcgttagtatttaa catgtatcccaactcctataaattccctatctttagttttagttgcagaaacattttgtgtcatlaa gcattgggtgtgttaattcaaccactgtaaaatgaattactacaaaatttgaatttagcttgg gttttgttacctttatgttttctccagggtcctctacttaagagatagtagcatataattataatgtt gctatgtacaagtcattttaactttatcacattatttgcagtttacctcctataaaacttagtgcggac aagttttaatccagaattgacctttgacttaagcagagggactttgtatagaagggttggggg ctgtgggggaaggagagtcacctgaagggtctgacacgtctgcctaccattctgtgtgatacaa ttaaatgtaggtatgaataagttcgaagctccgtgagtgaaacatcattataaacgtgtatgac agctgtttgtcatagggtcagttggaaacggcctcctagggaaggttcatagggtctcttcag gttcttagtctcacttacctagatttacagcctcacttgaaatgtgtcactacacagctctttaa ctctcagttttatctttaaactcctcttttattcttggactgacatttagcgtagctaagtgaaaaggct atagctgagattcctgtgtcgggtgttacgcacacgtacttaaatgaaagcatgtgcattgttc atcgtataacacaatatgaatacagggtcatgtatttgcagcagtgagctcttcagaaaacc cttttctacagttagggttgagttacttctatcaagccagttacgtctaacagggtcaaatattcc tgaatgaaatatcagactagtgacaagctcctggtcttgaatgtcttctcgttaaggagatgg gccttttggagggtaaaggataaaatgaatgagttctgtcatgattcactattctagaacttgcat gaacctttacgtgttagctcttgaatgtcttgaatgttagactttcttgaataacaaatgatatgt ccttatcattgtataaaagctgttatgtgcaacagtggtgagattcctgtctgtatttaataaata cttaaacactgaaaaaaaaa	
18S	X03205.1	tacctggtgatcctgccagtagcatatgctgtctcaagattaagccatgcatgtctaagtac gcacggccgttacagtgaaactgcgaatggctcataaatacagttatgggtccttgggtcctc gctccttcccacttggataactgtgtgaattctagagctaatacatgccgacgggctgtgac	24

		ccccctgcgggggggatgcgtgcatctatcagatcaaaaccaaccgggtcagccccctcc ggccccggccggggggcgccggcgccggcgccgttggtagctctagataaacctcgggccga tcgcacgcccccggtgcggcgacgacccattcgaacgtctgccctatcaactttcgatggt agtcgccgtgctaccatggtgaccacgggtgacggggaatcagggttcgattccggaga gggagcctgagaacggctaccacatccaagggaaggcagcaggcgcgcaaatccac tcccgacccggggaggtagtacgaaaaataacaatacaggactctttcaggccctgtaa ttggaatgagtcactttaaatcttaacgaggatccattggagggaagtcgtgtagcagc agccgggtaattccagctccaatagcgtatattaaagtgtcgtcagttaaaagctcgtagt ggatcttgggagcggcgccgggtccggcgaggcgagccaccgcccgtcccccgc cttgcctctcgccgccccctcgatgctcttagctgagtgtcccgggggccgaagcgtta cttgaaaaaatagagtgtcaagcaggcccgagccgctgataccgcagctaggat aatggaataggaccgcggttctatittgtgttttcggaactgaggccatgattagaggga ggccgggggcatctgtattgcgcgctagagggtgaatcttggaccggcgcaagacgga ccagagcgaagcatttgcgaagatgtttcattaatcaagaacgaagtcggagggtcga agcagatcagataccgctgtagtccgaccataaacgagtcgaccggcgatgcggcgcc gttattccatgaccccgccggcgagcttcgggaaaccaagcttgggttcggggggga gtatggttgaagctgaacttaaggaattgacgggaaggcaccaccaggagtggagc ctgcggcttaattgactcaacacgggaaacctcaccggcccgacacggacaggattga cagattgtagctcttctcgttccgtgggtgggtgcatgcccgttcttagttgtgagc gattgtctgttaattccgataacgaacgagactctgcatgctaaactgtacgcgaccccc gagcgtcgcgtcccccaacttcttagaggacaaagtggcgttcagccaccgagattga gcaataacaggctctgtatgaccttagatgtccggggcgacagcgcgctacactgctgg ctcagcgtgtgctacctacgcccagcgcgcggttaacccgtgaacccattcgtgat ggggatcggggtgcaattattcccatgaacgaggaaatccagtaagtgcgggtcataa gcttgcgtgttaagtccttgccttctacacaccgcccgtcgtactaccgattggatggt tagtgaggccctcgatcgcccccgccgggtcggcccacggccctcgcgagcgcgta gaagacggtcgaactgactatctagagggaagtaaaagtcgtaacagggttccgtagggtga acctgcggaaggatcatta	
PPIA	NM_021130. 4	ggggccgaacgtgggtataaaaggggcggggagggccaggtcgtgcccgttttcagacgcc accgcccaggaaaaccgtgtactatagccatgtgaaccccaccgtgttcttcacatggcc gtcgacggcgagcccttggcgccgtctccttgagctgttcagacaaggicccaaga cagcagaaaatttctgtctcgtgagcactggagagaaggaattgtgtataagggttcctgctt tcacagaattatccagggtttatgtgtcagggtgtgtgacttcacacgccataatggcactggt ggcaagtcctatctatgggagaaatttgaaagtgaagaacttcactcfaaagcatacgggtcct ggcatctgtccatggcgaatgtggaaccaacacaatggttcccagttttcactcgtcactg ccaagactgagtggttgatggcaagcattgtgtgtttggcaagtgaaagaagcagtaaa tattgtgaggccatggagcgttttgggtccagggaatggcaagaccagcaagaagatcacc attgctgactgtggacaactcgaataagttgactgtgttttatcttaaccaccagatcattcctt ctgtagctcaggagagcaccctccacccatttgcctgcagtatcctagaatcttgtgctct cgtcgcagttcccttgggttccatgttttctgttccctccatgcctagctggattgcagagt taagttatgattatgaaaaaaactaataacaattgtcctcgttgaagtaagttgtatgt aggctttatitaaagcagtaaatgggttactctgaacatcactgtttgcttaattctacagta	25

50

		atgagggtgggaaagccactcttaggggctgcatcaccatcatagggtgggagacactgc cacttgctgtccaaatggaaacacggaggataaagtcagccatgtgagcactgggggtgtg gccagtttgagctcctggaaaggtaaagtccttctgggggtgagctctcagcaatatttagt acttctgctcttttagttctgtgcacagcccctaagtcacttagcattttctgcatctccactt ggcattagctaaaaccttccatgtcaagattcagctagtggccaaagatgcagtgccaggga accttaaacagttgcacagcatctcagctcacttccactgacccctggatttgcatacattctt caagatcccatttgaatttttagtgaactaaaccattgtgcatcttagagtgcataatattatttgg cctgttaaaaagaaagttagcagtggttagcttagttctctttgtagtggttattatgattagcttt gtcactgtttcactactcagcatggaaacaagatgaattccattttaggttagtgagacaaaa ttgtagtccattaaagtaacaataaaaagtgtccattgaaccgtgatttttttttctctgata ctttgttaggaagggtgagaaatagaatcttgaggaaacgcatcagatgtctatattgtgaatgc aagaagtgggagcagcagcagtgaggagatgggacaattagataaatgtccattcttatacaa gggacctactttatggcagacattgtgctagtgttttatttaacttttattttatcagttacacatg atcataatttaaaaagtcaggcttataacaaaaagccccagcccatttctccattcaagat tccacctccccagagggtgaccactttcaactcttgagttttcaggtatatacctccatgtttcta agtaatatgttatattgttcaactcttttttttttttttaaaagaaatctatttcataccatggaggga aggctctgttccacatataattccacttcttctcgttatagttttgtcacaattatagattag atcaaaaagctacataactaatacagctgagctatgtagtatgctatgattaaatttactatgta aaaaaaaaaaaaaaaa	
<i>RPL13A</i>	NM_012423. 3	cacttctgcccctctgtttcaaggataagaaccctgcgacaaaacctctctcttttccaa gcggctgcccgaagatggcggagggtgcaggctcctggtgctgtaggtgcaggccatctcctg ggccgcttgccgcccactgtggttaaacaggctactgtggtggccggaaaggtggtggtcgtga cgtgtggaaggcatcaacatttctggcaatttctacagaacaagttgaagtacgtggtcttcc tccgcaagcggatgaacaccaaccttcccgaggccccctaccacttccgggccccagcc gcattcttctgcccgcaccgtgcgaggatgctgccccacaaaaccagcgaggccaggccg ctctggaccgtctcaagggtttgacggcatcccaccgcccctacgacaagaaaaagcggat gggtgttctctgtgcccacaaggctgtgctgtcgaagcctacaagaaagtgttccctatctggg gcgccttggtcagcaggtgtggtgaggaagtaaccagcagtgacagccacctggaggaga agaggaaagagaagccaagatccactaccggagaagaagaacagctcagaggctacgg aaacaggccgagaagaacgtggagaaagaaattgacaataacacagaggctccaagac ccaggactcctggtctgagcccataaagactgttaattctcctatgctgttccctgcccctct ccattgttgcctggaaatgtacgggaccagggggcagcagcagtcagggtgcccacaggga gccctgggacataggaaagctgggagcaaggaaagggtctttagtactgctcccgaaagt gcttgaagcactcggagaaattgtgcagggtgcaattatctatgaccaatagggaaggaacc agttactatgagtgaaaggagccagaaagactgattggagggccctatcttgtgagtgagg catctgttggacttccacctgtgcatatactctgcagctgttagaattgtcaagcacttgggg acagcatgagctgtgtgtacacagggtatttctagaaagcagaatagactgggaagatg cacaaccaaggggttacaggcatgcccctgctcctcacctgtattttgtaatcagaataaat tgcttttaagaaaaaaaaaaaaaaaa	27
<i>B2M</i>	NM_004048. 2	aatataagtgaggcgctgcgtggcgggcattcctgaagctgacagcattcgggcccag atgtctgcctcgtggccttagctgtgtcgcgtactctcttcttggcctggaggctatcca gcgtactccaaagattcaggtttactcagctatccagcagagaatggaaagtc aaatttct	28

		g a a t t g c t a t g t c t g g g t t c a t c c a c a t t g a a g t t g a c t t a c t g a a g a a t g g a g a g a g a a t t g a a a a g t g g a g c a t t c a g a c t t g t c t t c a g c a a g g a c t g g t c t t a t c t t g t a c t a c a c t g a a t t c a c c c c a c t g a a a a g a t g a g t a t g c c t g c c g t g a a c c a t g t a c t t t g t c a c a g c c c a a g a t a g t t a a g t g g a t c g a g a c a t g t a a g c a g c a t c a t g g a g g t t g a a g a t g c c g c a t t t g g a t t g g a t g a a t t c c a a a t c t g c t t g c t t t t a a t a t t g a t a t c t t a t a c a c t t a c a c t t a t g c a c a a a t g t a g g g t t a t a a t a a t g t a a c a t g a c a t g a t c t t t t a t a a t t c a c t t t g a g t g c t g t c c a t g t t g a t g t a t c g a g c a g g t t g c c a c a g g t a g c t c t a g g a g g g c t g g c a a c t t a g a g g t g g g a g c a g a a a t t c t t a t c c a a c a t c a a c a t c t t g g t c a g a t t g a a c t c t t a a t c t t g c a c t c a a a g c t t g t a a g a t a g t t a a g c t g c a t a a g t t a a c t t c c a a t t a c a t a c t c t g c t t a g a a t t g g g g a a a a t t t a g a a t a f a a t t g a c a g g a t t a t t g g a a a t t t g t a t a a t g a a t g a a c a a t t t g t c a t a a g a t t c a t a t t a c t t c t t a c a t t t g a t a a a g t a a g g c a t g g t t g g t t a a t c t g t t a t t t t g t c c a c a a g t t a a t a a a t c a f a a a a c t t g a t g t t a t c t t a	
YWHAZ	NM_003406. 3	c t t t c t c t c c c c t t c t t c c g g g c t c c c g t c c c g g c t a t c a c c c g g c t g t g g c c c a c t c c c a c c g c c a g c t g g a a c c c t g g g a c t a c g a c g t c c c t c a a a c c t t g c t t c t a g g a t a a a a a g a a c a t c c a g t c a t g g a t a a a a t g a g c t g g t c a g a a g c c a a a c t g c c g a g c a g g c t g a g c g a t a t g a t g a c a t g g c a g c c t g c a t g a a g t c t g t a a c t g a g c a a g g a g c t g a a t t a t c c a a t g a g g a g a g g a a t c t t c t c a g t t g c t t a a a a a t g t t g t a g g a g c c c g t a g g t c a t c t t g g a g g t c g t c t c a a g t a t t g a a c a a a a g a c g g a a g g t g c t g a g a a a a c a g c a g a t g g c t c g a g a a t a c a g a g a g a a a t t g a g a c g g a g c t a a g a g a t a t c t g c a a t g a t g t a c t g t c t c t t t g g a a a g t c t t g a t c c c a a t g c t t c a a a g c a g a g a g c a a a g t c t t a t t t g a a a a t g a a a g a g a t t a c t a c c g t a c t t g g c t g a g g t g c c g c t g t g a t g a c a a g a a a g g a t t g t c g a t c a g t c a c a a a g c a t a c c a a g a a g c t t t g a a a t c a g c a a a a g g a a a t g c a a c c a a c a t c c t a t c a g a c t g g g t c t g c c c t t a a c t t c t g t g t c t a t t a t g a g a t t c g a a c t c c c c a g a g a a g c c t g c t c t t t g c a a a g a c a g c t t t g a t g a a g c c a t t g c t g a a c t g a t a c a t t a a g t a a g a g t c a t a c a a a g a c a g c a c g t a a t a a t g c a a t t a c t g a g a g a c a a c t g a c a t t g t g a c a t c g g a t a c c c a a g g a g a c g a a g c t g a a g c a g g a g a a g g a a a a t t a a c c g g c c t c c a a c t t t g t c t g c c t a t t c t a a a t t a c a c a g t a g a c c a t t g t c a t c c a t g c t g t c c c a c a a t a g t t t t t g t t a c g a t t a t g a c a g g t t a t g t t a c t t c a t t t g a a t t c t a t a t t t c c a t g t g g t t t t a t g t t a a t a t t a g g g a g t a g a c c a g t t a a c a t t a g g g a g t t a c t g t t t c a t c t g a g g t g g c c a a t a t g g g a t g t g a a t t t t a t a c a a g t t a a g t t t g g c a t a g t a c t t t g t a c a t t g t g c t t c a a a g g g c c a g t g t a a a c t g c t t c c a t g t c t a a g c a a g a a a a c t g c c t a c a t a c t g t t t g t c c t g c g g g a a t a a a g g g a t c a t t g g t c c a g t c a c a g g t g t a g t a a t t g g g t a c t t t a a g t t t g g a g c a c t t a c a a g g c t g t g t a g a a t c a t a c c c a t g g a t a c c a c a t a t t a a c c a t g t a t a t c t g t g g a a t a c t a a t g t g t a c a c c t t g a c t a c a g c t g c a g a a g t t c c t t a g a c a a a g t t g a c c c a t t t a c t c t g g a t a a g g c a g a a a c g g t t c a c a t t c c a t t a t t g t a a a g t a c c t g c t g t a g c t t c a t t a t t t t g c t a c a c t c a t t t a t t g t a t t t a a a t g t t t a g g c a a c c t a a g a a c a a t g t a a a g t a a a g a t g c a g g a a a a t g a a t t g c t t g t a t t c a t t a c t c a t g t a t a t c a a g c a c a g c a g t a a a c a a a a c c c a t g t a t t a c t t t t t t t a g g a t t t t g c t t t g t g a t t t t t t t t t g a t a c t t g c c t a a c a t g c a t g t c t g t a a a a t a g t t a a c a g g g a a a a a c t t g a g a t g a t g g c t a g c t t t g t a a t g t c t t a g a a a t t t c a t g a a c a a t c c a a g c a t a a t t g t a a g a a c a c g t g t a t t a a t t c a t g t a a g t g g a a t a a a g t t t a t g a a t g	29

		gactttcaactactttctacagctttcatgtaaatagttcttgggtctgaacttctctaaagg aaattgtacatttttgaatttattccttattccctcttggcagctaattgggtcttaccagtttaa acacaaaattatcataacaaaaatactactaatataactactgtttccatgtcccatgatccct ctctctctcccaccctgaaaaaatgagttcctatttttctgggagagggggggattgatta gaaaaaatgtagtgttccattttaaatttggcatatggcatttttaacttaggaagccaca atgttcttggcccatcatgacattgggtagcatttaactgttaagtgttcttccaaatcatttt gggtttlaagaatttctgatactcttatagccttgccttcaattttagctttatttcttatttga gggtgcacaagattaccttctgttttagccttctgtctgtcacaaccattcttacttgggtggc atgtacttggaaaaaggccgcatgtcttctggctccactcagtgctaaaggcaccctgtctc ctttgtctgcatcccacagactatttccctcatcctatttactgcagcaaatctctcttagttag gagactgtgtttatctccctttaaaccctacctaactgaatgggtctgtcattgtctgtctttaa atccttctcttcttctctcttatttcttaataatgtaggggctaagttatacccaagctcac ttacaaaaatatttctcagtttgcagaaaacacaaaacaaaatgccattttaaagggt gtatttttctttagaatgtaagctcccaagagcagggaacaatgtttctgtatgttctattgtc ctagtacactgtaaatgtcaataaatatgtatgtgggaggcagtgagcttgatgataagg gtgagaaactgaaatcccaaacactgtttgtgtctgttttattatgacctcagattaaattggg aaatattggcccttttgataattgtcccaaatattacattcaataaaagtgcantggagaaaa aaaaaaa	
SDHA	NM_004168. 3	actgcagcccgcctcgaactccggcggtggcgcagggcggtatccccctccccgccca gctcgaccctgggtgtgtgtcgcagggcgagtcgtcgcaggggactggcgggactgcggg cggcaacagcagacatgtcgggggtccggggcctgtcgggctgtcgtgagcgtcggcgc ctggcgctggccaaggcggtggccaacagtgttgcacaacaggaaaccgagggttttca ctgttgatgggaacaaggaggcatctgctaagttagattcatttctcagtatccagta gtgtgacatgaattttagcagtggtgtgtgtagcgcgtggaggggcaggcttgcagctgcatt tggcctttctgaggcagggtttaatacagcatgtgttaccaggctgttctaccaggctcac actgttcagcacaggagggaatcattgtctgtcgtgggaacatggaggaggacaactgg agggtggcatttctacgacaccgtgaagggtccgactggctgggggaccaggatgccatc cactacatgacggagcaggcccccccgccggtgtgtcagctagaaaattatggcatgccg tttagcagaactgaagatgggaagattatcagcgtgcatgttgggtggacagagcccaagtt ggaaaggcgggcaggcccatcgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt ttgtcacaccttatatggaagggtctcgtcgtatgataccagctatttgtgtgagtatttgccttg gatctcctgattggagaatggggagtgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt ccatccatgcataagagcaagaacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt cttcagctgcacgtctgtcccacaccagcactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt aggccttcttgcaggacctagagtttgttcagtccacctacaggcatatattgtgtgtgtgt tgtctcattacggaaggatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt gtttatggagcgatagccccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt ccatgactctggagatccgagaaaggagggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt gcagctgcaccacctacctccagagcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt agccatgatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt cattataacatggggcattcccaactacaaggggcagggtcgtgtgtgtgtgtgtgtgtgtgtgtgt gccaggatcagattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	30

		tgggtgccaaccgctcggggcaaacctgccttggacctgggtgtcttggctgggcatgtgc cctgagcatcgaaagtcagcagccctggagataaagtcctccaattaaaccaaacgct ggggaagaaatctgcatgaatctgacaaatgagatttgcctgagcagaataaagacatcg gaactgcgactcagcatgcagaagtcataatgcaaatcatgctccgtgtccgtgtgggaag cgtgtgcaagaaggtgtgggaaaatcagcaagctctatggagacctaaagcacctgaag acgttcgaccggggaatggtctggaaacacggacctggtggagacctggagctgcagAAC ctgatctgtgtgcgtgcagaccatctacggagcagagggcacggaaggagtcacgggg cgcgcatgccagggaagactacaagggtcggattgagtagcattacccaagcccac caggggcaacagaagaagcccttggaggagcactggagggaagcacacctgtctatgtg gacgttggcactgggaaggctactctggaaatagacctgtatgcacaaaactttgaacga ggctgactgtccaccgtcccgccagccattcgtctactatgagacaagatgtgtgagat gacagaatcagctttgttaattatgtataatagctcatgcatgtccatgtcataactgtctcat acgtcttcgactctggggaagaaggagtagattgaaggagattggcacctagtggctgg gagcttgcaggaacccagtgccaggggagcgtgacacttaccttgtccctgtctatctt gtgagatgataaaactgggcacagctcttaataaataataaagaacaaactttctttattcc aaatccatttgaataattttactgtgtgacttttagtcattttgtgacctaaaaatcaaatgta cttgtattgtgtacatcaaaatccagatatttgtatagttcttttcttttctttttttttt agacaggatcgggtgcagtatgacacacagctcactgcagcccaaacctcctggcagct cagggtgacttctcagcctctctgagtagttggggctacagggtgtgaccaccatgccc agctcatttatttgaattgtaggggacagggtctcactgtgtgacctaggctgtctcaagtgt cctccctccttggcctcccaagggtgtggaattatagggtgtgaacaaccaaaaaaaaaa aa	
HPRT1	NM_000194.2	ggccgggcccgtcttctcctcagcttcaggcggcgtgcgacgagccctcaggcgaacctctgc gctttcccgccggcgccgctctgtctgcgctccgctcctcctctgtctccgccaccggc ttctcctcctgagcagtcagcccgccgcccggcctccgttatggcagcccgagccc tggcgtcgtgattagtgatgatgaaccagggtatgaccttgatttatttgcatacctaataat gctgaggatttggaaagggtgtttattcctcatggactaattatggacaggactgaacgtctg ctgcgagatgtgaaaggagatgggaggccatcacattgtagccctctgtgtcacaagg gggctataaattcttgcctgacctggtgattacatcaagcactgaatagaataatgtaga tccatcctatgactgtatgtttatcagactgaagagctatgtaatgaccagtcacagggg acataaaagtaattggaggatgatctcctcaacttgaacgggaagaaatgctgtatgttggaa gatataattgacactggcacaacatgcagactttgcttcttggcaggcagtagataatccaa agatgtgcaaggctgcagcctgtgtgaaaaggaccocaggaagtgttggatataagcc agacttgggtgttgaattccagacaagttgtgttaggatatgcccctgactataatgaata ctgcagggtattgaatcatgtttgtctcattaggaactggaaaagcaaaatacaaacctaa gatgagagttcaagttgatttggaaacatctgagtcctattgacatgccagtaaaatfatc aatgttctagtctgtgccatctgcttagtagagcttttgcattgtatcttctaagaattttatctgt tttgtactttagaatgtcagttgctgcatcctaaactgtttatttgcactatgagcctatagacta tcagttcccttggcgggattgtgttgaactgttaaatgaaaaaattcttaaacacagcact attgagtgaacattgaactatctgttaagaaataagagaagatatattgttttaattgg tattttaattttatataatgcaggaaagaatagaagtgatgaataattgttaattataccaccgtgtg ttagaaaagtaagaagcagtcattttcacatcaagacagcatcaagaagtgttgtctgtcc	31

		tggaaattatitttagtagtggttcagtaaatgttgactgtatittccaactgttcaaattattaccagtg aatctttgtcagcagttcccttttaaatgcaaatcaataaattcccaaaaattaaaaaaaaaaaa aaaaaaaaaa	
--	--	--	--

Definitions

[0108] The articles “a” and “an” are used in this disclosure 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 element or more than one element.

[0109] The term “and/or” is used in this disclosure to mean either “and” or “or” unless indicated otherwise.

[0110] As used herein, the terms “polynucleotide” and “nucleic acid molecule” are used interchangeably to mean a polymeric form of nucleotides of at least 10 bases or base pairs in length, either ribonucleotides or deoxynucleotides or a modified form of either type of nucleotide, and is meant to include single and double stranded forms of DNA. As used herein, a nucleic acid molecule or nucleic acid sequence that serves as a probe in a microarray analysis preferably comprises a chain of nucleotides, more preferably DNA and/or RNA. In other aspects, a nucleic acid molecule or nucleic acid sequence comprises other kinds of nucleic acid structures such as for instance a DNA/RNA helix, peptide nucleic acid (PNA), locked nucleic acid (LNA) and/or a ribozyme. Hence, as used herein the term “nucleic acid molecule” also encompasses a chain comprising non-natural nucleotides, modified nucleotides and/or non-nucleotide building blocks which exhibit the same function as natural nucleotides.

[0111] As used herein, the terms “hybridize,” “hybridizing,” “hybridizes,” and the like, used in the context of polynucleotides, are meant to refer to conventional hybridization conditions, such as hybridization in 50% formamide/6XSSC/0.1% SDS/100 µg/ml ssDNA, in which temperatures for hybridization are above 37 degrees centigrade and temperatures for washing in 0.1 XSSC/0.1% SDS are above 55 degrees C, and preferably to stringent hybridization conditions.

[0112] As used herein, the term “normalization” or “normalizer” refers to the expression of a differential value in terms of a standard value to adjust for effects which arise from technical variation due to sample handling, sample preparation, and measurement methods rather than biological variation of biomarker concentration in a sample. For example, when measuring the expression of a differentially expressed protein, the absolute value for the expression of the

protein can be expressed in terms of an absolute value for the expression of a standard protein that is substantially constant in expression.

[0113] The terms "diagnosis" and "diagnostics" also encompass the terms "prognosis" and "prognostics", respectively, as well as the applications of such procedures over two or more time points to monitor the diagnosis and/or prognosis over time, and statistical modeling based thereupon. Furthermore, the term diagnosis includes: a. prediction (determining if a patient will likely develop aggressive disease (hyperproliferative/invasive)), b. prognosis (predicting whether a patient will likely have a better or worse outcome at a pre-selected time in the future), c. therapy selection, d. therapeutic drug monitoring, and e. relapse monitoring.

[0114] "Accuracy" refers to the degree of conformity of a measured or calculated quantity (a test reported value) to its actual (or true) value. Clinical accuracy relates to the proportion of true outcomes (true positives (TP) or true negatives (TN)) versus misclassified outcomes (false positives (FP) or false negatives (FN)), and may be stated as a sensitivity, specificity, positive predictive values (PPV) or negative predictive values (NPV), or as a likelihood, odds ratio, among other measures.

[0115] The term "biological sample" as used herein refers to any sample of biological origin potentially containing one or more biomarkers. Examples of biological samples include tissue, organs, or bodily fluids such as whole blood, plasma, serum, tissue, lavage or any other specimen used for detection of disease.

[0116] The term "subject" as used herein refers to a mammal, preferably a human. In some aspects, a subject can have at least one colon cancer symptom. In some aspects, a subject can have a predisposition or familial history for developing a colon cancer. A subject can also have been previously diagnosed with a colon cancer and is tested for cancer recurrence.

[0117] "Treating" or "treatment" as used herein with regard to a condition may refer to preventing the condition, slowing the onset or rate of development of the condition, reducing the risk of developing the condition, preventing or delaying the development of symptoms associated with the condition, reducing or ending symptoms associated with the condition, generating a complete or partial regression of the condition, or some combination thereof.

[0118] Biomarker levels may change due to treatment of the disease. The changes in biomarker levels may be measured by the present disclosure. Changes in biomarker levels may be used to monitor the progression of disease or therapy.

[0119] "Altered", "changed" or "significantly different" refer to a detectable change or difference from a reasonably comparable state, profile, measurement, or the like. Such changes may be all or none. They may be incremental and need not be linear. They may be by orders of magnitude. A change may be an increase or decrease by 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 99%, 100%, or more, or any value in between 0% and 100%. Alternatively, the change may be 1-fold, 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold or more, or any values in between 1-fold and five-fold. The change may be statistically significant with a p value of 0.1, 0.05, 0.001, or 0.0001.

[0120] The term "stable disease" refers to a diagnosis for the presence of a colon cancer, however the colon cancer has been treated and remains in a stable condition, i.e. one that is not progressive, as determined by imaging data and/or best clinical judgment.

[0121] The term "progressive disease" refers to a diagnosis for the presence of a highly active state of a colon cancer, i.e. one has not been treated and is not stable or has been treated and has not responded to therapy, or has been treated and active disease remains, as determined by imaging data and/or best clinical judgment.

[0122] The term "neoplastic disease" refers to any abnormal growth of cells or tissues being either benign (non-cancerous) or malignant (cancerous). For example, the neoplastic disease can be a colon cancer.

[0123] The term "neoplastic tissue" refers to a mass of cells that grow abnormally.

[0124] The term "non-neoplastic tissue" refers to a mass of cells that grow normally.

[0125] The term "immunotherapy" can refer to activating immunotherapy or suppressing immunotherapy. As will be appreciated by those in the art, activating immunotherapy refers to the use of a therapeutic agent that induces, enhances, or promotes an immune response, including, e.g., a T cell response while suppressing immunotherapy refers to the use of a therapeutic agent that interferes with, suppresses, or inhibits an immune response, including, e.g., a T cell response. Activating immunotherapy may comprise the use of checkpoint inhibitors. Activating immunotherapy may comprise administering to a subject a therapeutic agent that activates a stimulatory checkpoint molecule. Stimulatory checkpoint molecules include, but are not limited to, CD27, CD28, CD40, CD122, CD137, OX40, GITR and ICOS. Therapeutic agents that activate a stimulatory checkpoint molecule include, but are not limited to, MEDI0562, TGN1412, CDX-1127, lipocalin.

[0126] The term "antibody" herein is used in the broadest sense and encompasses various antibody structures, including but not limited to monoclonal antibodies, polyclonal antibodies, multispecific antibodies (e.g., bispecific antibodies), and antibody fragments so long as they exhibit the desired antigen-binding activity. An antibody that binds to a target refers to an antibody that is capable of binding the target with sufficient affinity such that the antibody is useful as a diagnostic and/or therapeutic agent in targeting the target. In one embodiment, the extent of binding of an anti-target antibody to an unrelated, non-target protein is less than about 10% of the binding of the antibody to target as measured, e.g., by a radioimmunoassay (RIA) or biacore assay. In certain embodiments, an antibody that binds to a target has a dissociation constant (K_d) of $< 1 \mu\text{M}$, $< 100 \text{ nM}$, $< 10 \text{ nM}$, $< 1 \text{ nM}$, $< 0.1 \text{ nM}$, $< 0.01 \text{ nM}$, or $< 0.001 \text{ nM}$ (e.g. 10^{-8} M or less, e.g. from 10^{-8} M to 10^{-13} M , e.g., from 10^{-9} M to 10^{-13} M). In certain embodiments, an anti-target antibody binds to an epitope of a target that is conserved among different species.

[0127] A "blocking antibody" or an "antagonist antibody" is one that partially or fully blocks, inhibits, interferes, or neutralizes a normal biological activity of the antigen it binds. For example, an antagonist antibody may block signaling through an immune cell receptor (e.g., a T cell receptor) so as to restore a functional response by T cells (e.g., proliferation, cytokine production, target cell killing) from a dysfunctional state to antigen stimulation.

[0128] An "agonist antibody" or "activating antibody" is one that mimics, promotes, stimulates, or enhances a normal biological activity of the antigen it binds. Agonist antibodies can also enhance or initiate signaling by the antigen to which it binds. In some embodiments, agonist antibodies cause or activate signaling without the presence of the natural ligand. For example, an agonist antibody may increase memory T cell proliferation, increase cytokine production by memory T cells, inhibit regulatory T cell function, and/or inhibit regulatory T cell suppression of effector T cell function, such as effector T cell proliferation and/or cytokine production.

[0129] An "antibody fragment" refers to a molecule other than an intact antibody that comprises a portion of an intact antibody that binds the antigen to which the intact antibody binds. Examples of antibody fragments include but are not limited to Fv, Fab, Fab', Fab'-SH, F(ab')₂; diabodies; linear antibodies; single-chain antibody molecules (e.g. scFv); and multispecific antibodies formed from antibody fragments.

[0130] Administering chemotherapy to a subject can comprise administering a therapeutically effective dose of at least one chemotherapeutic agent. Chemotherapeutic agents include, but are not limited to, 13-cis-Retinoic Acid, 2-CdA, 2-Chlorodeoxyadenosine, 5-Azacitidine, 5-Fluorouracil, 5-FU, 6-Mercaptopurine, 6-MP, 6-TG, 6-Thioguanine, Abemaciclib, Abiraterone acetate, Abraxane, Accutane, Actinomycin-D, Adcetris, Ado-Trastuzumab Emtansine, Adriamycin, Adrucil, Afatinib, Afinitor, Agrylin, Ala-Cort, Aldesleukin, Alemtuzumab, Alecensa, Alectinib, Alimta, Alitretinoin, Alkaban-AQ, Alkeran, All-transretinoic Acid, Alpha Interferon, Altretamine, Alunbrig, Amethopterin, Amifostine, Aminoglutethimide, Anagrelide, Anandron, Anastrozole, Apalutamide, Arabinosylcytosine, Ara-C, Aranesp, Aredia, Arimidex, Aromasin, Arranon, Arsenic Trioxide, Arzerra, Asparaginase, Atezolizumab, Atra, Avastin, Avelumab, Axicabtagene Ciloleucel, Axitinib, Azacitidine, Bavencio, Bcg, Beleodaq, Belinostat, Bendamustine, Bendeka, Besponsa, Bevacizumab, Bexarotene, Bexxar, Bicalutamide, Bicnu, Blenoxane, Bleomycin, Blinatumomab, Blincyto, Bortezomib, Bosulif, Bosutinib, Brentuximab Vedotin, Brigatinib, Busulfan, Busulfex, C225, Cabazitaxel, Cabozantinib, Calcium Leucovorin, Campath, Camptosar, Camptothecin-11, Capecitabine, Caprelsa, Carac, Carboplatin, Carfilzomib, Carmustine, Carmustine Wafer, Casodex, CCI-779, Ccnu, Cddp, Ceenu, Ceritinib, Cerubidine, Cetuximab, Chlorambucil, Cisplatin, Citrovorium Factor, Cladribine, Clofarabine, Clolar, Cobimetinib, Cometriq, Cortisone, Cosmegen, Cotellic, Cpt-11, Crizotinib, Cyclophosphamide, Cyramza, Cytadren, Cytarabine, Cytarabine Liposomal, Cytosar-U, Cytoxan, Dabrafenib, Dacarbazine, Dacogen, Dactinomycin, Daratumumab, Darbepoetin Alfa, Darzalex, Dasatinib, Daunomycin, Daunorubicin, Daunorubicin Cytarabine (Liposomal), daunorubicin-hydrochloride, Daunorubicin Liposomal, DaunoXome, Decadron, Decitabine, Degarelix, Delta-Cortef, Deltasone, Denileukin Diftitox, Denosumab, DepoCyt, Dexamethasone, Dexamethasone Acetate, Dexamethasone Sodium Phosphate, Dexasone, Dexrazoxane, Dhad, Dic, Diodex, Docetaxel, Doxil, Doxorubicin, Doxorubicin Liposomal, Droxia, DTIC, Dtic-Dome, Duralone, Durvalumab, Eculizumab, Efudex, Ellence, Elotuzumab, Eloxatin, Elspar, Eltrombopag, Emcyt, Empliciti, Enasidenib, Enzalutamide, Epirubicin, Epoetin Alfa, Erbitux, Eribulin, Erivedge, Erleada, Erlotinib, Erwinia L-asparaginase, Estramustine, Ethyol, Etopophos, Etoposide, Etoposide Phosphate, Eulexin, Everolimus, Evista, Exemestane, Fareston, Farydak, Faslodex, Femara, Filgrastim, Firmagon, Floxuridine, Fludara, Fludarabine, Fluoroplex, Fluorouracil, Fluorouracil (cream), Fluoxymesterone, Flutamide, Folinic Acid, Folutyn, Fudr,

Fulvestrant, G-CSF, Gazyva, Gefitinib, Gemcitabine, Gemtuzumab ozogamicin, Gemzar, Gilotrif, Gleevec, Gleostine, Gliadel Wafer, GM-CSF, Goserelin, Granix, Granulocyte - Colony Stimulating Factor, Granulocyte Macrophage Colony Stimulating Factor, Halaven, Halotestin, Herceptin, Hexadrol, Hexalen, Hexamethylmelamine, HMM, Hycamtin, Hydrea, Hydrocort Acetate, Hydrocortisone, Hydrocortisone Sodium Phosphate, Hydrocortisone Sodium Succinate, Hydrocortone Phosphate, Hydroxyurea, Ibrance, Ibritumomab, Ibritumomab Tiuxetan, Ibrutinib, Iclusig, Idamycin, Idarubicin, Idelalisib, Idhifa, Ifex, IFN-alpha, Ifosfamide, IL-11, IL-2, Imbruvica, Imatinib Mesylate, Imfinzi, Imidazole Carboxamide, Imlygic, Inlyta, Inotuzumab Ozogamicin, Interferon-Alfa, Interferon Alfa-2b (PEG Conjugate), Interleukin-2, Interleukin-11, Intron A (interferon alfa-2b), Ipilimumab, Iressa, Irinotecan, Irinotecan (Liposomal), Isotretinoin, Istodax, Ixabepilone, Ixazomib, Ixempra, Jakafi, Jevtana, Kadcyla, Keytruda, Kidrolase, Kisqali, Kymriah, Kyprolis, Lanacort, Lanreotide, Lapatinib, Lartruvo, L-Asparaginase, Lbrance, Lcr, Lenalidomide, Lenvatinib, Lenvima, Letrozole, Leucovorin, Leukeran, Leukine, Leuprolide, Leurocristine, Leustatin, Liposomal Ara-C, Liquid Pred, Lomustine, Lonsurf, L-PAM, L-Sarcosine, Lupron, Lupron Depot, Lynparza, Marqibo, Matulane, Maxidex, Mechlorethamine, Mechlorethamine Hydrochloride, Medrolone, Medrol, Megace, Megestrol, Megestrol Acetate, Mekinist, Mercaptopurine, Mesna, Mesnex, Methotrexate, Methotrexate Sodium, Methylprednisolone, Metcorten, Midostaurin, Mitomycin, Mitomycin-C, Mitoxantrone, M-Prednisol, MTC, MTX, Mustargen, Mustine, Mutamycin, Myleran, Mylocel, Mylotarg, Navelbine, Necitumumab, Nelarabine, Neosar, Neratinib, Nerlynx, Neulasta, Neumega, Neupogen, Nexavar, Nilandron, Nilotinib, Nilutamide, Ninlaro, Nipent, Niraparib, Nitrogen Mustard, Nivolumab, Nolvadex, Novantrone, Nplate, Obinutuzumab, Octreotide, Octreotide Acetate, Odomzo, Ofatumumab, Olaparib, Olaratumab, Omacetaxine, Oncospar, Oncovin, Onivyde, Ontak, Onxal, Opdivo, Oprelvekin, Orapred, Orasone, Osimertinib, Otrexup, Oxaliplatin, Paclitaxel, Paclitaxel Protein-bound, Palbociclib, Pamidronate, Panitumumab, Panobinostat, Panretin, Paraplatin, Pazopanib, Pediapred, Peg Interferon, Pegaspargase, Pegfilgrastim, Peg-Intron, PEG-L-asparaginase, Pembrolizumab, Pemetrexed, Pentostatin, Perjeta, Pertuzumab, Phenylalanine Mustard, Platinol, Platinol-AQ, Pomalidomide, Pomalyst, Ponatinib, Portrazza, Pralatrexate, Prednisolone, Prednisone, Prelone, Procarbazine, Procrit, Proleukin, Prolia, Prolifeprospan 20 with Carmustine Implant, Promacta, Provenge, Purinethol, Radium 223 Dichloride, Raloxifene, Ramucirumab, Rasuvo, Regorafenib,

Revlimid, Rheumatrex, Ribociclib, Rituxan, Rituxan Hycela, Rituximab, Rituximab Hyalurodinase, Roferon-A (Interferon Alfa-2a), Romidepsin, Romiplostim, Rubex, Rubidomycin Hydrochloride, Rubraca, Rucaparib, Ruxolitinib, Rydapt, Sandostatin, Sandostatin LAR, Sargramostim, Siltuximab, Sipuleucel-T, Soliris, Solu-Cortef, Solu-Medrol, Somatuline, Sonidegib, Sorafenib, Sprycel, Sti-571, Stivarga, Streptozocin, SU11248, Sunitinib, Sutent, Sylvant, Synribo, Tafinlar, Tagrisso, Talimogene Laherparepvec, Tamoxifen, Tarceva, Targretin, Tasigna, Taxol, Taxotere, Tecentriq, Temodar, Temozolomide, Temsirolimus, Teniposide, Tespa, Thalidomide, Thalomid, TheraCys, Thioguanine, Thioguanine Tabloid, Thiophosphoamide, Thioplex, Thiotepa, Tice, Tisagenlecleucel, Toposar, Topotecan, Toremifene, Torisel, Tositumomab, Trabectedin, Trametinib, Trastuzumab, Treanda, Trelstar, Tretinoin, Trexall, Trifluridine/Tipiricil, Triptorelin pamoate, Trisenox, Tspa, T-VEC, Tykerb, Valrubicin, Valstar, Vandetanib, VCR, Vectibix, Velban, Velcade, Vemurafenib, Venclexta, Venetoclax, VePesid, Verzenio, Vesanoid, Viadur, Vidaza, Vinblastine, Vinblastine Sulfate, Vincasar Pfs, Vincristine, Vincristine Liposomal, Vinorelbine, Vinorelbine Tartrate, Vismodegib, Vlb, VM-26, Vorinostat, Votrient, VP-16, Vumon, Vyxeos, Xalkori Capsules, Xeloda, Xgeva, Xofigo, Xtandi, Yervoy, Yescarta, Yondelis, Zaltrap, Zanosar, Zarxio, Zejula, Zelboraf, Zevalin, Zinecard, Ziv-aflibercept, Zoladex, Zoledronic Acid, Zolinza, Zometa, Zydelig, Zykadia, Zytiga, or any combination thereof.

[0131] The terms “effective amount” and “therapeutically effective amount” of an agent or compound are used in the broadest sense to refer to a nontoxic but sufficient amount of an active agent or compound to provide the desired effect or benefit.

[0132] The term “benefit” is used in the broadest sense and refers to any desirable effect and specifically includes clinical benefit as defined herein. Clinical benefit can be measured by assessing various endpoints, e.g., inhibition, to some extent, of disease progression, including slowing down and complete arrest; reduction in the number of disease episodes and/or symptoms; reduction in lesion size; inhibition (i.e., reduction, slowing down or complete stopping) of disease cell infiltration into adjacent peripheral organs and/or tissues; inhibition (i.e., reduction, slowing down or complete stopping) of disease spread; decrease of auto-immune response, which may, but does not have to, result in the regression or ablation of the disease lesion; relief, to some extent, of one or more symptoms associated with the disorder; increase in the length of disease-free presentation following treatment, e.g., progression-free survival;

increased overall survival; higher response rate; and/or decreased mortality at a given point of time following treatment.

[0133] The terms "cancer" and "cancerous" refer to or describe the physiological condition in mammals that is typically characterized by unregulated cell growth. Included in this definition are benign and malignant cancers. Examples of cancer include but are not limited to, carcinoma, lymphoma, blastoma, sarcoma, and leukemia. More particular examples of such cancers include adrenocortical carcinoma, bladder urothelial carcinoma, breast invasive carcinoma, cervical squamous cell carcinoma, endocervical adenocarcinoma, cholangiocarcinoma, colon adenocarcinoma, lymphoid neoplasm diffuse large B-cell lymphoma, esophageal carcinoma, glioblastoma multiforme, head and neck squamous cell carcinoma, kidney chromophobe, kidney renal clear cell carcinoma, kidney renal papillary cell carcinoma, acute myeloid leukemia, brain lower grade glioma, liver hepatocellular carcinoma, lung adenocarcinoma, lung squamous cell carcinoma, mesothelioma, ovarian serous cystadenocarcinoma, pancreatic adenocarcinoma, pheochromocytoma, paraganglioma, prostate adenocarcinoma, rectum adenocarcinoma, sarcoma, skin cutaneous melanoma, stomach adenocarcinoma, testicular germ cell tumors, thyroid carcinoma, thymoma, uterine carcinosarcoma, uveal melanoma. Other examples include breast cancer, lung cancer, lymphoma, melanoma, liver cancer, colorectal cancer, ovarian cancer, bladder cancer, renal cancer or gastric cancer. Further examples of cancer include neuroendocrine cancer, non-small cell lung cancer (NSCLC), small cell lung cancer, thyroid cancer, endometrial cancer, biliary cancer, esophageal cancer, anal cancer, salivary, cancer, vulvar cancer or cervical cancer.

[0134] The term "tumor" refers to all neoplastic cell growth and proliferation, whether malignant or benign, and all pre-cancerous and cancerous cells and tissues. The terms "cancer," "cancerous," "cell proliferative disorder," "proliferative disorder" and "tumor" are not mutually exclusive as referred to herein.

Examples

[0135] The disclosure is further illustrated by the following examples, which are not to be construed as limiting this disclosure in scope or spirit to the specific procedures herein described. It is to be understood that the examples are provided to illustrate certain aspects and that no limitation to the scope of the disclosure is intended thereby. It is to be further understood that resort

may be had to various other aspects, embodiments, modifications, and equivalents thereof which may suggest themselves to those skilled in the art without departing from the spirit of the present disclosure and/or scope of the appended claims.

EXAMPLE 1. Derivation of a 13-marker gene panel

[0136] Raw probe intensities from $n = 24$ colon cancer tumor tissue samples were compared to $n = 22$ control colon mucosa to identify genes that best discriminated between disease using the transcriptional profile of E-MTAB-57. Gene co-expression networks were generated to identify temporal patterns of gene regulation associated with colon cancer. A total of 513 nodes with 53,786 links were identified. Differential expression analysis identified 103 genes were upregulated in tumor tissue compared to blood. To identify blood-specific colon cancer gene biomarkers, we evaluated expression of the 103 genes in peripheral blood transcriptomes ($n = 7$). Thirty-three (32%) of the 103 genes were below the level of detection in blood identifying these as candidate genes. Evaluation of transcripts in a preliminary dataset of blood samples from colon cancer ($n = 20$) and matched normal blood ($n = 20$) identified thirteen genes and one house-keeping gene as markers of colon cancer (**Table 2**). These genes were demonstrated to be highly expressed in colon cancer tumor tissue compared to normal mucosa and in three different colon cancer cell lines, LOVO (metastatic, hyperdiploid, MSI unstable cell line), LS-180 (derived from a Duke's B, colorectal adenocarcinoma) and Colo 320DM (derived from a Duke's C, colorectal adenocarcinoma). These data demonstrate target transcripts are produced by neoplastically transformed colon mucosal cells (**FIGs. 1A-1B**).

[0137] An artificial intelligence model of colon cancer disease was built using normalized gene expression of these 13 markers in whole blood from Controls ($n=120$) and Colon Cancers ($n=272$) samples. The dataset was randomly split into training and testing partitions for model creation and validation respectively. Twelve algorithms were evaluated (XGB, RF, glmnet, cforest, CART, treebag, knn, nnet, SVM-radial, SVM-linear, NB and mlp). The top performing algorithm (XGB – “gradient boosting”) best predicted the training data. In the test set, XGB produced probability scores that predicted the sample. Each probability score reflects the “certainty” of an algorithm that an unknown sample belongs to either “Control” or “Colon Cancer” class. For example, an unknown sample S1 can have the following probability vector [Control = 20%, Colon Cancer = 80%]. This sample would be considered a colon cancer sample.

Example 2. Clinical Utility

[0138] The data (receiver operator curve analysis and metrics) for the utility of the test to differentiate patients with colon cancer ($n=136$) from controls ($n=60$) in the training and test sets are included in **FIGs. 2A-2B**. The score exhibited an area under the curve (AUC) of 0.90 (training) and 0.86 (test set). The metrics are: sensitivity: 85.3-87.5% and specificity: 75-83.3%.

[0139] Overall, ColoTest scores were significantly elevated in cancers ($63\pm1\%$) and controls ($34\pm2\%$) (**FIGs. 3A-3B**). The overall accuracy (training and test cohort) is 84%, with an AUC: 0.88. The z-statistic for differentiating controls was 18.5.

[0140] A decision curve analysis was used to quantify the clinical benefit of the diagnostic test (**FIGs. 4A-4B**). The ColoTest exhibited >50% standardized predictive benefit up to a risk threshold of 80%. The probit risk assessment plot identified a ColoTest score >50% was 75% accurate for predicting colon cancer in a blood sample. This was increased to >80% at a ColoTest score $\geq 60\%$. The tool can therefore accurately differentiate between controls and colon cancer disease.

[0141] Specific evaluation of a colon cancer cohort before and after surgery identified that complete removal of a tumor and no evidence of disease was associated with a significant decrease ($p<0.0001$) in the ColoTest (**FIG. 5**). Levels were not significantly different in those with evidence of residual disease.

[0142] Examination of a separate colon cancer cohort by disease status (clinical evaluation at time of blood-draw) identified that the ColoTest was not significantly different between stable ($n=17$: $56\pm7\%$) and progressive disease ($n=32$: $68\pm4\%$) (**FIGs. 6A-6C**). However, 12 of the 17 patients progressed with 3 months of blood collection. Those that did progress exhibited elevated ColoTest scores at time of blood draw ($n=12$: $73\pm4\%$) that were not different to those with progressive disease at time of blood draw ($n=32$: $68\pm4\%$) (**FIGs. 6A-6C**). Levels in patients with stable disease were significantly lower ($n=5$: $16\pm4\%$, $p<0.0001$). A direct comparison between the ColoTest and CEA in these samples identified that the gene expression assay was significantly more sensitive ($p<0.05$) than CEA for predicting disease progression (**FIG. 7**). The ColoTest tool can therefore accurately predict progressive colon cancer disease.

[0143] ROC analysis identified the ColoTest had an AUC: 0.97 for differentiating stable from progressive disease. The z-statistic for differentiating controls was 20.6. Further evaluation of this

cohort identified that patients who exhibited disease progression despite therapy exhibited higher scores than those responding to therapy (**FIG. 8**). Therapies included bevacizumab, chemotherapy and EGFR TKI inhibitors. The tool can therefore accurately identify treatment failure in colon cancer disease.

Table 2.

Colon Cancer Biomarker or Housekeeping Genes		NCBI Chromosome location	UniGene ID	RefSeq	Amplicon length	Exon Boundary	Assay Location
Symbol	Name						
<i>ADRM1</i>	adhesion regulating molecule 1	Chr.20: 62302056 – 62308862	<i>Hs.90107</i>	NM_007002.3	60	3-4	486
<i>CDK4</i>	cyclin dependent kinase 4	Chr.12: 57747727 – 57752447	<i>Hs.95577</i>	NM_000075.3	65	5-6	928
<i>COMT</i>	catechol-O-methyltransferase	Chr.22: 19941740 – 19969975	<i>Hs.370408</i>	NM_000754.3	118	5-6	864
<i>DHCR7</i>	7-dehydrocholesterol reductase	Chr.11: 71434411 – 71448431	<i>Hs.503134</i>	NM_001163817.1	74	3-4	351
<i>HMOX2</i>	heme oxygenase 2	Chr.16: 4474697 – 4510347	<i>Hs.284279</i>	NM_001127204.1	81	5-6	1002
<i>MCM2</i>	minichromosome maintenance complex component 2	Chr.3: 127598357 – 127622436	<i>Hs.477481</i>	NM_004526.3	67	13-14	2374
<i>MORF4LI</i> (housekeeping gene)	mortality factor 4 like 1	Chr.15: 78872781 – 78897739	<i>Hs.374503</i>	NM_001265603.1	62	1	116
<i>PDXK</i>	pyridoxal (pyridoxine, vitamin B6) kinase	Chr.21: 43719097 – 43762307	<i>Hs.284491</i>	NM_003681.4	103	9-10	959
<i>POP7</i>	POP7 homolog, ribonuclease P/MRP subunit	Chr.7: 100706053 – 100707500	<i>Hs.416994</i>	NM_005837.2	136	2	828
<i>S100P</i>	S100 calcium binding protein P	Chr.4: 6693839 – 6697170	<i>Hs.2962</i>	NM_005980.2	73	1-2	234

<i>SNRPA</i>	small nuclear ribonucleoprotein polypeptide A	Chr.19: 40750854 -- 40765392	<i>Hs.466775</i>	NM_004596.4	123	3-4	986
<i>SORD</i>	sorbitol dehydrogenase	Chr.15: 45023104 -- 45075089	<i>Hs.878</i>	NM_003104.5	72	4-5	601
<i>STOML2</i>	stomatin like 2	Chr.9: 35099776 -- 35103195	<i>Hs.3439</i>	NM_001287031.1	68	2-3	290
<i>UMPS</i>	uridine monophosphate synthetase	Chr.3: 124730366 -- 124749273	<i>Hs.2057</i>	NM_000373.3	85	3-4	1082

[0144] References:

[0145] 1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. *CA Cancer J Clin.* 2017; 67: 7-30. doi: 10.3322/caac.21387. Epub 2017 Jan 5.

[0146] 2. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, Forman D, Bray F. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer.* 2013; 49: 1374-403.

[0147] 3. Fritzmann J, Morkel M, Besser D, Budczies J, Kosel F, Brembeck FH, Stein U, Fichtner I, Schlag PM, Birchmeier W. A colorectal cancer expression profile that includes transforming growth factor beta inhibitor BAMBI predicts metastatic potential. *Gastroenterology.* 2009; 137: 165-75.

[0148] 4. Chen VW, Hsieh MC, Charlton ME, Ruiz BA, Karlitz J, Altekruse SF, Ries LA, Jessup JM. Analysis of stage and clinical/prognostic factors for colon and rectal cancer from SEER registries: AJCC and collaborative stage data collection system. *Cancer.* 2014; 120: 3793-806.

[0149] 5. Heald RJ, Lockhart-Mummery HE. The lesion of the second cancer of the large bowel. *Br J Surg.* 1972; 59: 16-9.

[0150] 6. Mokhles S, Macbeth F, Farewell V, Fiorentino F, Williams NR, Younes RN, Takkenberg JJ, Treasure T. Meta-analysis of colorectal cancer follow-up after potentially curative resection. *Br J Surg.* 2016; 103: 1259-68.

[0151] 7. Thomas SN, Zhu F, Schnaar RL, Alves CS, Konstantopoulos K. Carcinoembryonic antigen and CD44 variant isoforms cooperate to mediate colon carcinoma cell adhesion to E- and L-selectin in shear flow. *J Biol Chem.* 2008; 283: 15647-55.

- [0152] 8. Amri R, Bordeianou LG, Sylla P, Berger DL. Preoperative carcinoembryonic antigen as an outcome predictor in colon cancer. *J Surg Oncol.* 2013; 108: 14-8.
- [0153] 9. Jansen N, Coy JF. Diagnostic use of epitope detection in monocytes blood test for early detection of colon cancer metastasis. *Future Oncol.* 2013; 9: 605-9.
- [0154] 10. Locker GY, Hamilton S, Harris J, Jessup JM, Kemeny N, Macdonald JS, Somerfield MR, Hayes DF, Bast RC, Jr. ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J Clin Oncol.* 2006; 24: 5313-27.
- [0155] 11. Warren JD, Xiong W, Bunker AM, Vaughn CP, Furtado LV, Roberts WL, Fang JC, Samowitz WS, Heichman KA. Septin 9 methylated DNA is a sensitive and specific blood test for colorectal cancer. *BMC Med.* 2011; 9:133.: 10.1186/741-7015-9-133.
- [0156] 12. Mead R, Duku M, Bhandari P, Cree IA. Circulating tumour markers can define patients with normal colons, benign polyps, and cancers. *Br J Cancer.* 2011; 105: 239-45.
- [0157] 13. Molnar B, Floro L, Sipos F, Toth B, Sreter L, Tulassay Z. Elevation in peripheral blood circulating tumor cell number correlates with macroscopic progression in UICC stage IV colorectal cancer patients. *Dis Markers.* 2008; 24: 141-50. doi:
- [0158] 14. Mishaeli M, Klein B, Sadikov E, Bayer I, Koren R, Gal R, Rakowsky E, Levin I, Kfir B, Schachter J, Klein T. Initial TPS serum level as an indicator of relapse and survival in colorectal cancer. *Anticancer Res.* 1998; 18: 2101-5.
- [0159] 15. Piepoli A, Cotugno R, Merla G, Gentile A, Augello B, Quitadamo M, Merla A, Panza A, Carella M, Maglietta R, D'Addabbo A, Ancona N, Fusilli S, et al. Promoter methylation correlates with reduced NDRG2 expression in advanced colon tumour. *BMC Med Genomics.* 2009; 2:11.: 10.1186/755-8794-2-11.

Equivalents

[0160] While the present invention has been described in conjunction with the specific aspects set forth above, many alternatives, modifications and other variations thereof will be apparent to those of ordinary skill in the art. All such alternatives, modifications and variations are intended to fall within the spirit and scope of the present invention.

Claims:

1. A method for detecting a colon cancer in a subject in need thereof, comprising:
 - determining the expression level of at least 14 biomarkers from a test sample from the subject by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene;
 - normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*;
 - inputting each normalized expression level into an algorithm to generate a score;
 - comparing the score with a predetermined cutoff value; and
 - identifying the presence of a colon cancer in the subject when the score is equal to or greater than the predetermined cutoff value or identifying the absence of a colon cancer in the subject when the score is less than the predetermined cutoff value.
2. A method for determining whether a colon cancer in a subject is stable or progressive, the method comprising:
 - determining the expression level of at least 14 biomarkers from a test sample from the subject by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene;
 - normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*;
 - inputting each normalized expression level into an algorithm to generate a score;

comparing the score with a predetermined cutoff value; and
identifying that the colon cancer in the subject is progressive when the score is equal to or greater than the predetermined cutoff value or identifying that the colon cancer in the subject is stable when the score is less than the predetermined cutoff value.

3. A method for determining the completeness of surgery in a subject having a colon cancer, the method comprising:

determining the expression level of at least 14 biomarkers from a test sample from the subject after the surgery by contacting the test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, *UMPS*, and a housekeeping gene;

normalizing the expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*, *PDXK*, *POP7*, *S100P*, *SNRPA*, *SORD*, *STOML2*, and *UMPS*;

inputting each normalized expression level into an algorithm to generate a score;

comparing the score with a predetermined cutoff value; and

identifying that the colon cancer in the subject is not completely removed when the score is equal to or greater than the predetermined cutoff value or identifying that the colon cancer in the subject is completely removed when the score is less than the predetermined cutoff value.

4. A method for evaluating the response of a subject having a colon cancer to a first therapy, the method comprising:

(1) at a first time point:

(a) determining the expression level of at least 14 biomarkers from a first test sample from the subject by contacting the first test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1*, *CDK4*, *COMT*, *DHCR7*, *HMOX2*, *MCM2*,

PDXK, POP7, S100P, SNRPA, SORD, STOML2, UMPS, and a housekeeping gene;

(b) normalizing the expression level of each of *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, S100P, SNRPA, SORD, STOML2, and UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, S100P, SNRPA, SORD, STOML2, and UMPS*;

(c) inputting each normalized expression level into an algorithm to generate a first score;

(2) at a second time point, wherein the second time point is after the first time point and after the administration of the therapy to the subject:

(a) determining the expression level of at least 14 biomarkers from a second test sample from the subject by contacting the second test sample with a plurality of agents specific to detect the expression of the at least 14 biomarkers, wherein the 14 biomarkers comprise *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, S100P, SNRPA, SORD, STOML2, UMPS*, and the housekeeping gene;

(b) normalizing the expression level of each of *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, S100P, SNRPA, SORD, STOML2, and UMPS* to the expression level of the housekeeping gene, thereby obtaining a normalized expression level of each of *ADRM1, CDK4, COMT, DHCR7, HMOX2, MCM2, PDXK, POP7, S100P, SNRPA, SORD, STOML2, and UMPS*;

(c) inputting each normalized expression level into the algorithm to generate a second score;

(3) comparing the first score with the second score; and

(4) identifying that the subject is responsive to the first therapy when the second score is significantly decreased as compared to the first score or identifying that the subject is not responsive to the first therapy when the second score is not significantly decreased as compared to the first score.

5. The method of claim 4, wherein the subject has not received the first therapy before the first time point.
6. The method of claim 4, wherein the first time point is after the administration of the first therapy to the subject.
7. The method of any one of claims 4-6, wherein the second score is significantly decreased as compared to the first score when the second score is at least 25% less than the first score.
8. The method of any one of claims 4-7, wherein the first therapy comprises anti-cancer therapy, surgery, chemotherapy, targeted drug therapy, radiation therapy, immunotherapy or any combination thereof.
9. The method of claim 8, wherein when the first therapy comprises surgery, the surgery comprises removing a polyp during a colonoscopy, endoscopic mucosal resection, a partial colectomy, an ostomy, removing at least one cancerous lesion from the liver, or any combination thereof.
10. The method of claim 8, wherein when the first therapy comprises chemotherapy, the chemotherapy comprises FOLFOX, FOLFIRI, a combination of 5-FU and leucovorin, capecitabine, irinotecan, CapeOx or any combination thereof.
11. The method of claim 8, wherein when the first therapy comprises targeted drug therapy, the targeted drug therapy comprises bevacizumab, cetuximab, panitumumab, regorafenib, a combination of trifluridine and tipiracil, an EGFR TKI inhibitor or any combination thereof.
12. The method of claim 8, wherein when the first therapy comprises anti-cancer therapy, the anticancer therapy comprises anti-colon cancer therapy.

13. The method of claim 8, wherein when the first therapy comprises immunotherapy, the immunotherapy comprises pembrolizumab, nivolumab or a combination of pembrolizumab and nivolumab.
14. The method of any one of claims 1-3, wherein the predetermined cutoff value is at least 50% on a scale of 0-100%.
15. The method of any one of claims 1-3, wherein the predetermined cutoff value is at least 60% on a scale of 0-100%.
16. The method of any one of claims 1-15, wherein the housekeeping gene is selected from the group consisting of *MRPL19*, *PSMC4*, *SF3A1*, *PUM1*, *ACTB*, *GAPD*, *GUSB*, *RPLP0*, *TFRC*, *MORF4L1*, *18S*, *PPIA*, *PGK1*, *RPL13A*, *B2M*, *YWHAZ*, *SDHA*, and *HPRT1*.
17. The method of any one of claims 1-16, wherein the housekeeping gene is *MORF4L1*.
18. The method of any one of claims 1-17, having a sensitivity greater than 85%.
19. The method of any one of claims 1-18, having a specificity greater than 85%.
20. The method of any one of claims 1-19, wherein at least one of the at least 14 biomarkers is RNA, cDNA or protein.
21. The method of claim 20, wherein when the biomarker is RNA, the RNA is reverse transcribed to produce cDNA, and the produced cDNA expression level is detected.
22. The method of claim 20, wherein the expression level of the biomarker is detected by forming a complex between the biomarker and a labeled probe or primer.

23. The method of any one of claims 1-3, wherein the predetermined cutoff value is derived from a plurality of reference samples obtained from subjects not having or not diagnosed with a neoplastic disease.
24. The method of claim 23, wherein the neoplastic disease is colon cancer.
25. The method of any one of claims 1-24, wherein the algorithm is XGBoost (XGB), Random Forest (RF), glmnet, cforest, Classification and Regression Trees for Machine Learning (CART), treebag, K-Nearest Neighbors (kNN), neural network (nnet), Support Vector Machine radial (SVM-radial), Support Vector Machine linear (SVM-linear), Naïve Bayes (NB), or multilayer perceptron (mlp).
26. The method of any one of claims 1-25, wherein the algorithm is XGBoost.