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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 May 2014 (15.05.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/075083 A1

(51) International Patent Classification:
C12Q 1/68 (2006.01) *G01N 33/48* (2006.01)
C12N 15/11 (2006.01) *G01N 33/53* (2006.01)

(21) International Application Number:
PCT/US20 13/0697 10

(22) International Filing Date:
12 November 2013 (12.1 1.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/724,528 9' November 2012 (09. 11.2012) US

(71) Applicants: **THE REGENTS OF THE UNIVERSITY OF CALIFORNIA** [US/US]; Office Of The President, 111 Franklin Street, 5th Floor, Oakland, CA 94607-5200 (US). **SAGE BIONETWORKS** [US/US]; 1100 Fairview Avenue N, MS: M1-C108, Seattle, WA 98109 (US).

(72) Inventors: **ZHANG, Kang;** 9415 Campus Point Drive, La Jolla, CA 92093 (US). **HANNUM, Gregory;** Pharmaceutical Science Building, Rm. 4244, 9500 Oilman Drive, La Jolla, CA 92093 (US). **D3EKER, Trey;** Pharmaceutical Science Building, Rm. 4230, 9500 Oilman Drive, La Jolla, CA 92093 (US). **GUINNEY, Justin;** Sage Bionetworks, 1100 Fairview Avenue N, Seattle, WA 98109 (US). **FRIEND, Stephen, H.;** Sage Bionetworks, 1100 Fairview Avenue N, Seattle, WA 98109 (US).

(74) Agent: **ADRIANO, Sarah, B.;** Adriano & Associates, 690 East Green Street, Suite 300, Pasadena, CA 91101 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available):* AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Continued on nextpage]

(54) Title: METHODS FOR PREDICTING AGE AND IDENTIFYING AGENTS THAT INDUCE OR INHIBIT PREMATURE AGING

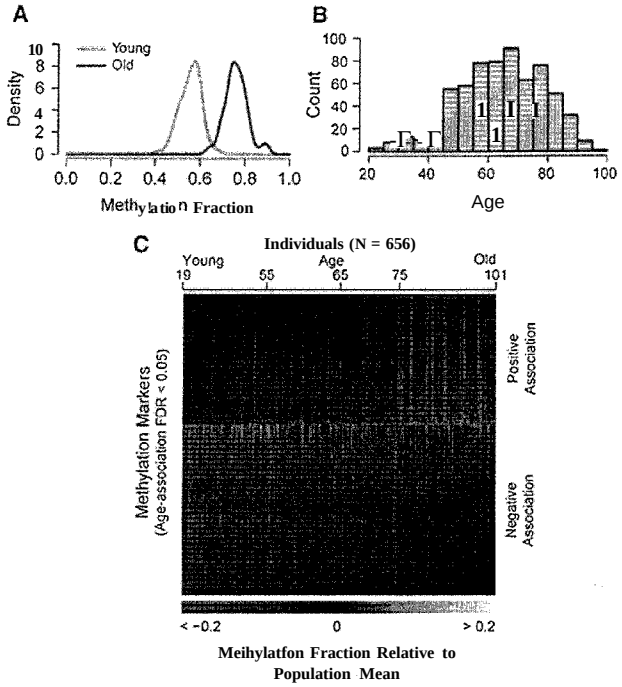


FIGURE 1

(57) Abstract: The invention provides for methods for predicting age of a subject based on the epigenome of the subject.

WO 2014/075083 A1

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available):* ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

METHODS FOR PREDICTING AGE AND IDENTIFYING AGENTS THAT INDUCE OR INHIBIT PREMATURE AGING

This invention was made with government support under Grant Nos. P50GM085764 and R01E501481 1 awarded by NIH and Grant Nos. EY014428, EY018660, EY019270 and EY021374 awarded by NEI/NIH. The government has certain rights in the invention.

Throughout this application various publications are referenced. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art to which this invention pertains,

BACKGROUND OF THE INVENTION

Not everyone ages in the same manner. It is well known that women tend to live longer than men, and lifestyle choices such as smoking and physical fitness can hasten or delay the aging process (Steven N., 2006; Blair *et al,* 1989). These observations have led to the search for molecular markers of age which can be used to predict, monitor, and provide insight into age-associated physiological decline and disease. One such marker is telomere length, a molecular trait strongly correlated with age (Harley *et al,* 1990) which has been shown to have an accelerated rate of decay under environmental stress (Epel *et al,* 2004; Valdes *et al).* Another marker is gene expression, especially for genes that function in metabolic and DNA repair pathways which are predictive of age across a range of different tissue types and organisms (Fraser *et al,* 2005; Zahn *et al,* 2007; de Magalhaes *et al,* 2009).

A growing body of research has reported associations between age and the state of the epigenome—the set of modifications to DNA other than changes in the primary nucleotide sequence (Fraga and Esteller, 2007). In particular, DNA methylation associates with chronological age over long time scales (Alisch *et al,* 2012; Christensen *et al,* 2009; BoUati *et al,* 2009; Boks *et al,* 2009; Rakyan *et al,* 2010; Bocklandt *et al,* 201 1; Bell *et al,* 2012) and changes in methylation have been linked to complex age-associated diseases such as metabolic disease (Barres and Zierath, 201 1) and cancer (Jones and Laird, 1999; Esteller, 2008). Studies

have also observed a phenomenon dubbed "epigenetic drift", whereby the DNA methylation marks in identical twins increasingly differ as a function of age (Fraga *et al.,* 2005; Boks *et al.,* 2009). Thus, the idea of the epigenome as a fixed imprint is giving way to the model of the epigenome as a dynamic landscape that reflects a variety of chronological changes. The current challenge is to determine whether these changes can be systematically described and modeled to detect different rates of human aging, and to tie these rates to related clinical or environmental variables.

The mechanisms that drive changes in the aging methylome are not well understood, although they have been attributed to at least two underlying factors (Vijg and Campisi, 2008; Fraga *et al.,* 2005). First, it is possible that environmental exposure will over time activate cellular programs associated with consistent and predictable changes in the epigenome. For example, stress has been shown to alter gene expression patterns through specific changes in DNA methylation (Murgatroyd *et al.,* 2009). Alternatively, spontaneous epigenetic changes may occur with or without environmental stress, leading to fundamentally unpredictable differences in the epigenome between aging individuals. Spontaneous changes may be caused by chemical agents that disrupt DNA methyl groups or through errors in copying methylation states during DNA replication. Both mechanisms lead to differences between the methylomes of aging individuals, suggesting that quantitative measurements of methylome states may identify factors involved with slowed or accelerated rates of aging.

To better understand how the methylome ages and to determine whether human aging rates can be quantified and compared, we initiated a project to perform genome-wide methylomic profiling of a large cohort of individuals spanning a wide age range. Based on these findings, we constructed a predictive model of aging rate which we show is influenced by gender and specific genetic variants. These data help explain epigenetic drift and suggest that age-associated changes in the methylome lead to changes in transcriptional patterns over time. These findings were replicated in a second large cohort.

The ability to measure human aging from molecular profiles has practical implications in many fields, including disease prevention and treatment, forensics, and extension of life. Although

chronological age has been linked to changes in DNA methylation, the methylome has not yet been used to measure and compare human aging rates. Here, we have created a quantitative model of aging using measurements at more than 450,000 CpG markers from the whole blood of 656 human individuals, aged 19 to 101. This model measures the rate at which an individual's methylome ages. Furthermore, we have discovered that differences in aging rates may explain epigenetic drift and are reflected in the transcriptome. Our discovery highlights specific components of the aging process and provides forensic methods, screening methods for agents retarding or accelerating aging, and methods for preventing and treating diseases.

SUMMARY OF THE INVENTION

The invention provides methods for predicting age of a subject based on the epigenome of the subject. In one embodiment, the method comprises (a) obtaining a biological sample of the subject; (b) determining the methylation status of a set of age-associated epigenetic marker(s) in the epigenome of the subject as shown in any of Figure 9, Tables S3, S4 and/or S5; and (c) comparing the methylation status of a set of age-associated epigenetic marker(s) of the subject with the methylation status of the same markers from an age correlated reference population so as to obtain a value or a range of values for the predicted age of the subject thereby predicting the age of a subject based on the epigenome of the subject.

The invention also provides for methods for identifying type of tissue for a biological sample from a subject with a known chronological age. In one embodiment, the method comprises (a) ascertaining the chronological age of a subject; (b) determining the AMAR of the subject from the biological sample by dividing the predicted age of a subject from the chronological age of the subject; (c) comparing to a reference standard relating AMAR to chronological age for various types of tissue; (d) determining which value from step (b) closely matches the AMAR in the reference standard for various types of tissue from step (c); and (e) based on the closest match in step (d), assigning the type of tissue for the biological sample, thereby identifying type of tissue for a biological sample from a subject with a known chronological age.

The invention further provides for methods for predicting age of a subject based on age-associated epigenetic modification affecting gene expression comprising: (a) obtaining a biological sample of the subject; (b) determining the expression of one or more gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age; (c) comparing the expression of one or more gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age with the expression of the same gene(s) from an age correlated reference population; and (d) obtaining a value or range of values for the predicted age of the subject; wherein comparing the expression of one or more gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age with the expression of the same gene(s) from an age correlated reference population comprises any statistical method, multivariate regression method, linear regression analysis, tabular method, or graphical method used to predict the age of a subject based on expression of gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age; thereby predicting age of a subject based on age-associated epigenetic modification affecting gene expression.

The invention also provides methods for predicting age of a tissue or organ of a subject based on the epigenome of the tissue or organ of the subject. In one embodiment, the method comprises (a) obtaining a biological sample of a tissue or organ from the subject; (b) determining the methylation status of a set of age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; and (c) comparing the methylation status of the set of age-associated epigenetic marker(s) of the subject with the methylation status of the same markers from an age-correlated reference population so as to obtain a value or a range of values for the predicted age of the tissue or organ, thereby predicting the age of a tissue or organ of a subject based on the epigenome of the tissue or organ of the subject.

The invention also provides for a kit for determining age of a subject based on epigenetic modification of subject's genetic material comprising any age-associated epigenetic marker or markers as listed in Figure 9, Table S3, Table S4 or Table S5.

The invention further provides for a kit for predicting age of a subject based on the epigenome of the subject utilizing the set of the age-associated epigenetic marker(s) provided in Figure 9, Table S3, S4 and/or S5.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1. Global Data on the Aging Methylome. (A) A density plot of methylation fraction values for the marker cgl 6867657, separated by young (green) and old (blue) individuals. (B) A histogram of the age distribution for all individuals. (C) A heatmap of the age-associated methylation markers, sorted by the magnitude of association (regression coefficient). The individuals are ordered youngest to oldest. See also Figure SI and Tables SI and S2 for a specific example of an age-associated region and for annotation coincidence tables, respectively.

Figure 2. Model Predictions and Clinical Variables. (A) A flow chart of the data (green boxes) and analyses (red ovals) used to generate aging predictions (blue boxes). (B) A comparison of predicted and actual ages for all individuals based on the aging model. (C) Out-of-sample predictions for individuals in the validation cohort. (D) Apparent methylomic aging rate (AMAR) for each individual, based on the aging model without clinical variables. The distribution of aging rates shows faster aging for men than women. A table of the markers used in the aging model is provided in Table S3. See also Figures S2 and S3, Table S3 and Figure 9.

Figure 3. Genetic Effects on Methylomic Aging. (A) We surveyed genomic variants for an association with age-associated methylation markers. Eight genetic variants, corresponding to 14 meQTLs, were chosen for validation. Of these, seven were significant in the validation cohort and two showed an association with AMAR. (B) A plot of the trend between the methylation marker cg27367526 (STEAP2) and age. The state of variant rs42663 (GTPBP10) causes an offset in this relationship. (C) A second example for cgl 8404041 and rs2230534 (ITIH1, NEK4). See also Table S4 for a table of confirmed genetic associations.

Figure 4. Multitissue Support. (A) Predictions of age made by the full aging model on the TCGA control samples. There is a high correlation between chronological and predicted age, but

each tissue has a different linear intercept and slope. (B) After adjusting the intercept and slope of each tissue, the error of the model is similar to that of the original whole-blood data. Age predictions made on cancer samples are presented in Figure S2. (C) Age predictions made on matched normal and tumor samples from TCGA. Predictions are adjusted for the linear offset of the parent tissue (breast, kidney, lung, or skin). (D) Tumor samples show a significant increase in AMAR. See also Figure S4 and Table S5.

Figure 5. Age Associations for Methylation. Fraction and Deviance (A) Methylation fraction values for are shown for the marker cg24724428. Over any subset of the cohort, we consider two group methylation statistics: the mean and variance. Marker variance is a measure of the mean methylation deviance, which is defined as the squared difference between an individual's methylation fraction and their expected methylation fraction. (B) A density plot showing the change in mean methylation with age for the marker cg24724428. Young and old groups are based on the top and bottom 10%. (C) A histogram of the significance of association between the methylation fraction of all markers and age. p values are signed such that positive values represent an increase of methylation with age. Markers that exceeded the FDR < 0.05 threshold are grouped into the most extreme bins. (D) A density plot showing the change in methylation deviance with age for the marker cg24724428. (E) A histogram, in the same form as (D), of the significance of association between the methylation deviance of all markers and age. Aging trends are mapped for CpG islands in Figure S3. See also Figure S5.

Figure 6. Methylome-wide Trends with Age. (A) Aggregate regression lines for all methylation markers that increased with age (red) and decreased with age (blue). The darkest color represents the median regression line and the bounds represent the 25% and 75% quantile. Both increasing and decreasing markers trend toward moderate methylation fraction values. (B) An entropy aging rate was calculated as the mean Shannon entropy of age-associated methylation markers divided by chronological age. This was strongly associated with AMAR.

Figure 7. Transcription Aging Model. (A) We built an aging model using mRNA expression data for genes that showed an aging trend in the methylome. Its standard error (RMSE = 7.22 years) is increased due to the rounding of ages to the nearest 5 year interval in the data set. (B)

Similar to the methylome, the transcriptome shows an increased aging rate for men as compared to women ($p < 10^{-4}$). See also Table S6 and Table S7.

Figure 8 shows the model of biological age.

Figure 9 is a collection of Figures 9A (9A through 9A-7), **9B** (9B through 9B-7), **9C** (9C through 9C-7), **9D** (9D through 9D-7) and 9E (9E through 9E-7) which are spreadsheets showing age-associated epigenetic markers, designated by "eg" prefix followed by a number (cg#), related to Table S3 in which each of the CpG dinucleotide so examined is embedded within the sequence shown in column entitled "Forward_Sequence" in the third subpanel of each series (i.e., Figs. 9x-2, where "x" is A-E). For example, the dinucleotide of interest is bounded by brackets. Additional information may be found in the Gene Expression Omnibus (GEO) database with GEO accession number GPL13534 and Bibikova *et al.* Genomics, 201 1, 98:288-95. The various subpanels of the spreadsheets should be assembled as shown below:

9A	9A-1	9A-2	9A-3	9A-4	9A-5	9A-6	9A-7
9B	9B-1	9B-2	9B-3	9B-4	9B-5	9B-6	9B-7
9C	9C-1	9C-2	9C-3	9C-4	9C-5	9C-6	9C-7
9D	9D-1	9D-2	9D-3	9D-4	9D-5	9D-6	9D-7
9E	9E-1	9E-2	9E-3	9E-4	9E-5	9E-6	9E-7

Figure SI. An example aging association map, related to Figure 1

Age association levels for the gene Four and a Half LIM Domains 2 (FHL2). A strong aging association is shown for several markers (red: -logio(p-value)) at a CpG island in the center of the gene, coincident with an internal promoter (black: average methylation fraction).

Figure S2. Apply the aging model to the Heyn *et al.* dataset, related to Figure 2

We obtained methylation profiles from the Heyn *et al.* dataset and applied the age prediction model. Our model successfully separated old and young samples (black circles). In addition, we applied the aging model to the three samples in the Heyn *et al.* dataset which were measured using bisulfite sequencing rather than the bead-chip technology used for our data. Despite the

differences in technology, the model successfully separated the young, middle-aged, and old samples (green dots).

Figure S3. Measuring the effects of batch-correlated variables, related to Figure 2

The model covariates of ethnicity and diabetes status were highly correlated with batch variables, such that their effect on the aging process could not be determined. Nonetheless, we built separate models for the subgroups (A) European, (B) Hispanic, (C) Non-diabetic, and (D) Diabetic. Each model was used to predict the age of its complementary cohort. The results show a strong predictive power despite the covariate and/or batch effects.

Figure S4. Normal and tumor aging model predictions, related to Figure 4

Aging models were built in matched normal and tumor samples using the model markers identified in the primary cohort. The aging rate (AMAR) of tumor samples predicted by normal tissue was found to be higher than expected (red, Wilcox test, $P < 10^{-21}$) and the aging rate of normal samples predicted by the tumor model was lower than expected (black, Wilcox test, $P < 10^{-17}$).The separation of the two aging rates was also highly significant (Wilcox test, $P < 10^{-25}$).

Figure S5. A map of aging trends in CpG Islands, related to Figure 5

(A) An aggregate genomic map of the methylation fraction for 27,176 CpG islands (black). The aging coefficient relating methylation fraction to age is shown in the same region (green). Color bars indicating the island and shore regions represent 75% confidence intervals. (B) A CpG island map showing methylation deviance (red) and the aging coefficient for deviance (green).

DETAILED DESCRIPTION OF THE FIGURES

DEFINITIONS

As used in this application, the biological age (bioage), chemical age, methylomic age and molecular age are equivalent or synonymous. The biological age is determined using a set of age-associated epigenetic markers of a subject or an organism. In the current invention, the

biological age is determined from an analysis of the modification status of specific CpG dinucleotide and, in particular, e.g., the methylation status at the C-5 position of cytosine.

Chronological age is the actual age of a subject or organism. For animals and humans, chronological age may be based on the age calculated from the moment of conception or based on the age calculated from the time and date of birth. The chronological age of the cell, tissue or organ may be determined from the chronological age of the subject or organism from which the cell, tissue or organ is obtained, plus the duration of the cell, tissue or organ is placed in culture. Alternatively, in the case of the cell or tissue culture, the chronological age may be related to the total or accumulative time in culture or passage number.

As used in this application, the term "tissue" may be replaced with "cell," or vice versa, for a biological sample.

The methylation marker as provided in Tables S3, S4 and S5 under the column "Marker" or "Methylation Marker," provided in Figure 9 under the column "ID" or "Name" in Figs. 9x where "x" is A-E and discussed in the text with a "cg#" designation are age-associated epigenetic markers. The specific CpG dinucleotide within each epigenetic marker probed in the invention is provided in Figure 9 under the heading "Forward_Sequence" and the specific CpG dinucleotide probed within brackets, i.e., [CG]. Additional sequence information for all "cg#" designation, such as in Tables S4 and S5 and in the text, may be obtained at the National Center for Biotechnology Information of the National Institutes of Health (Bethesda, MD) in the Gene Expression Omnibus (GEO) database with GEO accession number GPL13534.

The methylation markers as provided in Figure 9, Tables S3, S4 and S5 were used in an Illumina's Infinium Methylation Assay using the HumanMethylation450 BeadChip. However, these age-associated epigenetic markers may be used in other assays outside of the Infinium Methylation Assay system, based on the sequence, homology, or normal association to sequence for each cg# provided in the invention.

METHODS OF THE INVENTION

The invention provides for methods for predicting age of a subject based on the epigenome of the subject. The subject may be human, mammal, animal, plant, or any multicellular organism. Examples of suitable mammals include but are not limited human, monkey, ape, dog, cat, cow, horse, goat, pig, rabbit, mouse and rat. The age of a subject may be a chronological age or a molecular age, chemical age, methylomic age or biological age. The epigenome may be deoxyribonucleic acid (DNA) in which the DNA may be subjected to epigenetic modification. The epigenetic modification may be methylation of CpG residues. In one embodiment, the methylation is the covalent attachment of a methyl group at the carbon-5 (C-5) position of cytosine.

In one embodiment, the method comprises obtaining a biological sample of the subject. Additionally, the method comprises determining the methylation status of a set of age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5. Further, the method comprises comparing the methylation status of a set of age-associated epigenetic marker(s) of the subject with the methylation status of the same markers from an age correlated reference population so as to obtain a value or a range of values for the predicted age of the subject, thereby predicting the age of a subject based on the epigenome of the subject.

In one embodiment, the method comprises use of a statistical method to compare the methylation status of a set of age-associated epigenetic marker(s) of the subject with the methylation status of the same markers from an age correlated reference population. Examples of suitable statistical methods include but are not limited to multivariate regression method, linear regression analysis, tabular method or graphical method comprises Elastic Net, Lasso regression method, ridge regression method, least-squares fit, binomial test, Shapiro-Wilk test, Grubb's statistics, Benjamini-Hochberg FDR, variance analysis, entropy statistics, and/or Shannon entropy. In a preferred embodiment, the statistical method comprises a multivariate regression algorithm or linear regression algorithm.

In accordance with the practice of the invention, determining the methylation status may comprise isolating genomic DNA or nuclear DNA from the sample, reacting the isolated genomic DNA or nuclear DNA with one or more probe/agent (e.g., a chemical probe/agent) which differentially reacts with unmodified cytosine so that the cytosine is converted to uracil. The step may also comprise determining or analyzing the methylation status at the cytosine position (also referred to herein as the C position) of a CpG dinucleotide in the isolated genomic DNA or nuclear DNA of the sample by detecting the presence of a cytosine or uracil. The presence of cytosine or uracil indicates the presence of a 5-methylcytosine or unmodified cytosine, respectively, in the original CpG dinucleotide. Alternatively, resistance to cleavage by a restriction enzyme may indicate the presence of 5-methylcytosine at the original CpG dinucleotide. Sensitivity to cleavage by the restriction enzyme may indicate presence of unmodified cytosine at the original CpG dinucleotide. Further, the step may further comprise determining the proportion of 5-methylcytosine or unmodified cytosine initially present at each age-associated epigenetic marker; or alternatively, determining the ratio of 5-methyl-cytosine to unmodified cytosine or the ratio of unmodified cytosine to 5-methyl-cytosine cytosine initially present at each age-associated epigenetic marker based on characterizing outcome of probing the isolated genomic DNA or nuclear DNA.

In accordance with the practice of the invention, determining the methylation status may comprise isolating genomic DNA or nuclear DNA from the sample, incubating the isolated genomic DNA or nuclear DNA with one or more restriction enzyme which recognizes a specific DNA sequence, is affected by a CpG dinucleotide, within or adjacent to the restriction enzyme recognition or cleavage site, and differentially cleaves the DNA based on the presence or absence of a methyl group at C-5 position of cytosine of the CpG dinucleotide. The step may also comprise determining or analyzing the methylation status at the C position of a CpG dinucleotide in the isolated genomic DNA or nuclear DNA of the sample by its resistance to cleavage at a potential cleavage site by the restriction enzyme indicating presence of 5-methylcytosine at the original CpG dinucleotide within or adjacent to the restriction enzyme recognition or cleavage site. Sensitivity to cleavage by the restriction enzyme may indicate presence of unmodified cytosine. Further, the step may further comprise determining the proportion of 5-methylcytosine or unmodified cytosine initially present at each age-associated

epigenetic marker; or alternatively, determining the ratio of 5-methylcytosine to unmodified cytosine or the ratio of unmodified cytosine to 5-methylcytosine initially present at each age-associated epigenetic marker.

In accordance with the practice of the invention, the methylation status may be determined based on five or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; ten or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; fifteen or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; twenty or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; twenty-five or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; thirty or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; thirty-five or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; forty or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; forty-five or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; fifty or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; fifty-five or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; sixty or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; sixty-five or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; or seventy or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5.

Further, in a preferred embodiment, the methylation status may be determined based on five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; ten or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; fifteen or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; twenty or more age-associated epigenetic marker(s) in the

epigenome of the subject from Figure 9 or Table S3; twenty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; thirty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; thirty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; forty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; forty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; fifty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; fifty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; sixty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; sixty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; or seventy or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3. For example, the set of markers having individual CpG residues subject to methylation of C-5 position of cytosine in the genome of a subject may comprise any one or more of the following methylation marker cg05652533 of Table S4, cg27367526 of Table S4, cgl 8404041 of Table S4, cg23606718 of Figure 9, Tables S3 and S5, cgl 6867657 of Figure 9, Tables S3 and S5, cg04474832 on chromosome 3 at position 52008487, cg05442902 on chromosome 22 at position 21369010, cg06493994 on chromosome 6 at position 25652602, cg09809672 on chromosome 1 at position 236557682, cgl 9722847 on chromosome 12 at position 308491 14, cg22736354 on chromosome 6 at position 18122719, cg05652533 of Table S4, cg27367526 of Table S4, cgl 8404041 of Table S4, cg23606718 on chromosome 2 at position 131513927, and/or cgl6867657 of chromosome 6 at position 11044877.

In one embodiment, the set of markers having individual CpG residues subject to methylation at C-5 position of cytosine in the genome of a subject may comprise methylation marker cg04474832 on chromosome 3 at position 52008487, cg05442902 on chromosome 22 at position 21369010, cg06493994 on chromosome 6 at position 25652602, cg09809672 on chromosome 1 at position 236557682, cgl9722847 on chromosome 12 at position 308491 14, and cg22736354 on chromosome 6 at position 18122719.

In another embodiment, the set of markers having individual CpG residues subject to methylation at C-5 position of cytosine in the genome of a subject may be any one or more of methylation marker cg20822990 of Figure 9 or Table S3, cg04400972 of Figure 9 or Table S3, cgl6054275 of Figure 9 or Table S3, cg036071 17 of Figure 9 or Table S3, cg20052760 of Figure 9 or Table S3, cgl6867657 of Figure 9 or Table S3, cg06493994 of Figure 9 or Table S3, cg066851 1 1 of Figure 9 or Table S3, cg004861 13 of Figure 9 or Table S3, cg20426994 of Figure 9 or Table S3, cgl4361627 of Figure 9 or Table S3, cg08097417 of Figure 9 or Table S3, cg07955995 of Figure 9 or Table S3, cg22285878 of Figure 9 or Table S3 and/or cg08540945 of Figure 9 or Table S3.

In further embodiment, the set of age-associated epigenetic marker(s) may be any one or more of methylation marker cg23606718 of Figure 9, Tables S3 and S5 and/or cgl 6867657 of Figure 9, Tables S3 and S5.

In accordance with the practice of the invention, the methods of the invention may be automated.

In accordance with the practice of the invention, the biological sample may be any of blood, lymphocyte, monocyte, neutrophil, basophil, eosinophil, myeloid lineage cell, lymphoid lineage cell, bone marrow, saliva, buccal swab, nasal swab, urine, fecal material, hair, breast tissue, ovarian tissue, uterine tissue, cervical tissue, prostate tissue, testicular tissue, brain tissue, neuronal cell, astrocyte, liver tissue, kidney, thyroid tissue, stomach tissue, intestine tissue, pancreatic tissue, vascular tissue, skin, lung tissue, bone tissue, cartilage, ligament, tendon, fat cells, muscle cells, neurons, astrocytes, cultured cells with different passage number, cancer/tumor cells, cancer/tumor tissue, normal cells, normal tissue, any tissue(s) or cell(s) with a nucleus containing genetic material, or genetic material in the form of DNA of a known or unknown subject.

The tumor or cancer cells may be derived from blood, lymph node, liver, brain, esophagus, trachea, stomach, intestine, pancreas, throat, tongue, bone, ovary, uterus, cervix, peritoneum, prostate, testes, breast, kidney, lung, or skin. The biological sample with tumor or cancer cells

may be predicted to have an older predicted age of at least about 30% or 40% more than the biological sample without tumor or cancer cells.

In one embodiment, the age-associated epigenetic marker(s) comprises a CpG residue. The methylation at C-5 position of cytosine may vary with the chronological age of a species associated with the subject. For example, the species associated with the subject may be *Homo sapiens.*

In another embodiment, the set of age-associated epigenetic marker(s) may comprise individual CpG residues subject to age-dependent methylation at C-5 position of cytosine in the genome of a subject. The set of markers may comprise about 70 distinct CpG residue-containing age-associated epigenetic markers. Additionally, the set of markers may comprise any one or more of markers as shown in Figure 9, Table S3, Table S4 or Table S5.

For example, the set of age-associated markers may comprise five or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; ten or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; fifteen or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; twenty or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; twenty-five or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; thirty or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; thirty-five or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; forty or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; forty-five or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; fifty or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; fifty-five or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; sixty or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; sixty-five or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; or seventy or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3.

In another embodiment, the set of age-associated markers may comprise five or more age-associated epigenetic marker(s) as shown in Table S5; ten or more age-associated epigenetic

marker(s) as shown in Table S5; fifteen or more age-associated epigenetic marker(s) as shown in Table S5; twenty or more age-associated epigenetic marker(s) as shown in Table S5; twenty-five or more age-associated epigenetic marker(s) as shown in Table S5; thirty or more age-associated epigenetic marker(s) as shown in Table S5; thirty-five or more age-associated epigenetic marker(s) as shown in Table S5; forty or more age-associated epigenetic marker(s) as shown in Table S5; forty-five or more age-associated epigenetic marker(s) as shown in Table S5; or fifty or more age-associated epigenetic marker(s) as shown in Table S5.

Merely by way of example, the correlation between chronological age and predicted age may be at least about 80%, 90% or *91%* with an error of less than about 5 years.

In yet another embodiment, the set of age-associated epigenetic marker(s) may be any of methylation marker cg23606718 of Figure 9, Tables S3 and S5 and/or cgl 6867657 of Figure 9, Tables S3 and S5 and the biological sample with tumor or cancer cells may be predicted to have an older predicted age of at least about *30%* or 40% more than the biological sample without tumor or cancer cells.

In an embodiment, a majority of the age-associated epigenetic markers in the epigenome of the subject may predict an older age for a biological sample with tumor than biological sample of the same type without tumor. Similarly, pre-cancerous lesions may show an older biological age or predicted age than a normal tissue type without such a lesion.

In another embodiment, a majority of the age-associated epigenetic markers in the epigenome of the subject predicting an older age for a biological sample with tumor than biological sample of the same type without tumor may be more than about 70% of total age-associated epigenetic markers.

In an embodiment, one or more probes (e.g., chemical probes) may differentially react with an unmodified cytosine and 5-methyl-modified cytosine. The probe may be chosen from a set comprising a sodium bisulfite, sodium metabisulfite, and/or bisulfite salts.

In another embodiment, the outcome of reacting the isolated genomic DNA or nuclear DNA with one or more probes may be the deamination of unmodified cytosine to uracil and unaltered 5-methylcytosine. Characterizing the outcome of probing (or reacting) the isolated genomic DNA or nuclear DNA with one or more probe(s) or analyzing the methylation status may involve DNA amplification and nucleic acid sequence determination and detecting for the presence of either cytosine or thymine at the C position of the CpG dinucleotide within the age-associated epigenetic marker. Further, DNA amplification may be followed by phage RNA polymerase transcription, RNase cleavage and matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) of RNase cleavage products.

The nucleic acid sequence determination may involve one or more of the following procedures: nucleic acid fragmentation, restriction enzyme digestion, nucleic acid hybridization, primer extension, pyrosequencing, single nucleotide extension, single nucleotide extension with biotin-labelled ddNTP, single nucleotide extension with 2,4-dinitrophenol (DNP)-labelled ddNTP, radioactive isotope labeling, non-radioactive label incorporation, fluorescent label incorporation, biotin incorporation, antigen-antibody complex formation, antibody detection, colorimetric detection, fluorescence detection, detection with fluorescent dye-labelled antibody, detection with labeled avidin or streptavidin, bead analysis or detection method, signal amplification, polymerase chain reaction, DNA amplification with thermostable DNA polymerase, phi-29 DNA polymerase DNA amplification, RNA production, in vitro transcription, phage RNA polymerase transcription, T7 RNA polymerase transcription, SP6 RNA polymerase transcription, T3 RNA polymerase transcription, RNAse digestion, RNase A digestion, DNA cloning, bacterial transformation, gel electrophoresis, mass spectroscopy, MALDI-TOF mass spectroscopy, microarray analysis, fluorescence scanner analysis, automated digital image capture, automated digital image analysis, ratiometric analysis, and Infmium® HumanMethylation450 BeadChip analysis.

In another embodiment, the proportion of unmodified cytosine initially present at each age-associated epigenetic marker may be a fraction or percent of an age-associated epigenetic marker with thymine at pyrimidine position of the CpG dinucleotide.

The ratio of 5-methylcytosine to unmodified cytosine initially present at each age-associated epigenetic marker may be the ratio of cytosine to thymidine at pyrimidine position of the CpG dinucleotide after exposure to one or more probe and following analysis of products of nucleic acid amplification.

In yet another embodiment, the ratio of unmodified cytosine to 5-methylcytosine initially present at each age-associated epigenetic marker may be the ratio of thymidine to cytosine at pyrimidine position of the CpG dinucleotide after exposure to one or more probe and following analysis of products of nucleic acid amplification.

In an embodiment, one or more restriction enzyme probe which recognizes a specific DNA sequence, is affected by a CpG dinucleotide, within or adjacent to the restriction enzyme recognition or cleavage site, and differentially cleaves the DNA based on the presence or absence of a methyl group at C-5 position of cytosine of the dinucleotide may be selected. Examples of such restriction enzymes include but are not limited to Aatll, Acc65I, Accl, Acil, Acll, Afel, Agel, Ahdl, Alel, Apal, ApaLI, Ascl, AsiSI, Aval, Avail, Bael, Banl, , BbrPI, BbvCI, BceAI, Bcgl, BcoDI, BfuAI, BfuCI, Bgll, BmgBI, BsaAI, BsaBI, BsaHI, Bsal, BseYI, BsiEI, BsiWI, BslI, BsmAI, BsmBI, BsmFI, BspDI, BsrBI, BsrFI, BssHII, BssKI, BstAPI, BstBI, BstUI, BstZ17I, Cac8I, Clal, Dpnl, Dralll, Drdl, Eael, Eagl, Earl, Ecil, Eco53kI, EcoRI, EcoRV, Faul, Fnu4HI, Fokl, Fsel, Fspl, Haell, Haelll, Hgal, Hhal, Hindi, Hinfl, HinPlI, Hpal, Hpall, Hpyl66II, Hpyl88III, Hpy99I, HpyAV, HpyCH4IV, HpyCH4V, Kasl, Mbol, MluI, Mmel, MspAlI, Mwol, Nael, Narl, Neil, NgoMIV, Nhel, NlalV, Notl, Nrul, Nt. BbvCI, Nt. BsmAI, Nt. CviPII, PaeR71, Phol, Plel, PluTI, Pmel, Pmll, PshAI, PspOMI, PspXI, Pvul, Rsal, RsrII, Sacll, Sail, Sau3AI, Sau96I, ScrFI, SfaNI, Sfil, Sfol, SgrAI, Smal, SnaBI, StyD4I, Tfil, Tlil, Tsel, TspMI, Xhol, Xmal, and Zral.

In an embodiment, the outcome of reacting the isolated genomic DNA or nuclear DNA with one or more restriction enzyme probe may be the production of a double-stranded DNA break at a restriction enzyme cleavage site when cytosine at a CpG dinucleotide is not modified or no double-stranded DNA break at a restriction enzyme cleavage site when cytosine at a CpG dinucleotide is modified at its C-5 position with a methyl group.

In another embodiment, analyzing the methylation status may comprise DNA amplification and analysis of age-associated epigenetic marker for specific DNA end(s) or fragment(s) due to cleavage by the restriction enzyme(s) and for intact restriction enzyme cleavage site associated at a particular age-associated epigenetic marker.

In another embodiment, the proportion of 5-methylcytosine initially present at each age-associated epigenetic marker may be a fraction or percent of the age-associated epigenetic marker with an intact restriction enzyme cleavage site due to resistance to cleavage by the restriction enzyme.

In yet another embodiment, the proportion of unmodified cytosine initially present at each age-associated epigenetic marker may be a fraction or percent of the age-associated epigenetic marker cleaved by the restriction enzyme.

In another embodiment, the ratio of 5-methylcytosine to unmodified cytosine initially present at each age-associated epigenetic marker may be the ratio of number or concentration of intact restriction enzyme cleavage sites to the number or concentration of double-stranded DNA breaks produced by the restriction enzyme for the age-associated epigenetic marker.

In another embodiment, the ratio of unmodified cytosine to 5-methylcytosine initially present at each age-associated epigenetic marker may be the ratio of number or concentration of double-stranded DNA breaks produced by the restriction enzyme to number or concentration of intact restriction enzyme cleavage sites resistant to cleavage by the restriction enzyme due to presence of 5-methylcytosine for the age-associated epigenetic marker.

In one embodiment, determining the methylation status comprises isolating genomic DNA or nuclear DNA from the sample. Additionally, the step involves probing the isolated genomic DNA or nuclear DNA with one or more probes which differentially reacts with unmodified and 5-methyl-modified cytosine and amplifying the DNA. The step also involves digesting the amplified DNA with one or more restriction enzyme that recognizes a restriction enzyme site

that contains a CpG dinucleotide but fails to digest the restriction enzyme site mutated to TpG dinucleotide from a CpG dinucleotide. Further, the step involves determining the proportion of 5-methylcytosine or unmodified cytosine initially present at each age-associated epigenetic marker based on the fraction or percentage of restriction enzyme sites sensitive or resistant to digestion.

Alternatively, the methylation status may involve determining the ratio of 5-methylcytosine to unmodified cytosine initially present at each age-associated epigenetic marker based on the ratio of number or concentration of sensitive restriction enzyme sites to number or concentration of resistant restriction enzyme sites to digestion. The methylation status may also involve determining the ratio of unmodified cytosine to 5-methylcytosine initially present at each age-associated epigenetic marker based on the ratio of number or concentration of resistant restriction enzyme sites to number or concentration of sensitive restriction enzyme sites to digestion.

In another embodiment, determining the methylation status of the set of age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5 may comprise isolating genomic DNA or nuclear DNA and fragmenting the genomic DNA or nuclear DNA. Additionally, the step involves exposing the fragmented DNA to a 5-methylcytosine-binding protein and separating 5-methylcytosine-binding protein-bound DNA fragments from 5-methylcytosine-binding protein-free DNA fragments. The step further involves determining for each age-associated epigenetic marker, the proportion of 5-methylcytosine-containing DNA fragments or unmodified cytosine-containing DNA fragments by determining the fraction or percent of 5-methylcytosine-binding protein bound or free DNA fragments, respectively, for each age-associated epigenetic marker.

Alternatively, the methylation status may involve determining for each age-associated epigenetic marker, the ratio of 5-methylcytosine-containing DNA fragments to unmodified cytosine-containing DNA fragments by determining the ratio of number or concentration of 5-methylcytosine-binding protein-bound DNA fragments to the number or concentration of 5-methylcytosine-binding protein-free DNA fragments.

The methylation status step may also involve determining for each age-associated epigenetic marker, the ratio of unmodified cytosine-containing DNA fragments to 5-methylcytosine-containing DNA fragments by determining the ratio of number or concentration of 5-methylcytosine-binding protein-free DNA fragments to the number or concentration of 5-methylcytosine-binding protein-bound DNA fragments for each age-associated epigenetic marker.

In one embodiment, the 5-methylcytosine-binding protein may be an antibody for 5-methylcytosine, MeCP2, MBD2, MBD2/MBD3L1 complex, core MBD domain of MBD2, or poly-MBD protein, a naturally occurring 5-methylcytosine binding protein, genetically engineered 5-methylcytosine binding protein, or derivative or fragment thereof.

In one embodiment, separating 5-methylcytosine-binding protein-bound DNA fragments from 5-methylcytosine-binding protein-free DNA fragments may include immunoprecipitation, immunocapture, solid phase chromatography, liquid chromatography, and/or gel electrophoresis.

In accordance with the practice of the invention, the invention provides methods for determining apparent methylomic aging rate (AMAR) of a subject. The method comprises predicting age by the method of the invention and dividing the age predicted by the actual chronological age.

In one embodiment, the invention provides methods for diagnosing the presence of tumor in a subject. The method comprises obtaining biological sample suspected to contain tumor and a second biological sample of the same type but known not to contain tumor. Additionally, the method comprises predicting the age of the each biological sample by the method of the invention. The method further comprises comparing the ages predicted for the two samples, such that a biological sample with tumor will have an older predicted age than biological sample without tumor.

In an embodiment, the invention provides a forensic diagnosis of human actual age from a tissue from a human by predicting age of a subject based on the epigenome of the subject by the method of the invention.

In another embodiment, the invention provides methods for health assessment of a subject by predicting age of a subject based on the epigenome of the subject by the method of the invention.

In yet another embodiment, the invention provides methods for screening whether an agent of interest can retard or accelerate aging process. The method comprises obtaining a biological sample from a living organism, and optionally, culturing cells, tissue, or organ derived from a living organism and predicting age or AMAR of the organism, following the method of the invention, using organism appropriate age-associated epigenetic marker(s) such that the age-associated epigenetic marker(s) for a human subject may need to be substituted with age-associated epigenetic marker(s) for the organism being examined. Additionally, the method comprises exposing the living organism or cultured living cells, tissue, or organ from the living organism to an agent of interest in a single dose, multiple doses, or continuous doses and obtaining a biological sample from the living organism or the cultured living cells, tissue, or organ. The method further comprises predicting age or AMAR of the organism from the biological sample using organism appropriate age-associated epigenetic marker(s) such that the age-associated epigenetic marker(s) for a human subject may need to be substituted with age-associated epigenetic marker(s) for the organism being examined. The method also comprises performing the same steps on another individual from the same organism or a duplicate cultured living cells, tissue, or organ from the same individual or organism but not treated with any agent of interest or treated with a placebo and comparing, for biological sample of the same predicted age or AMAR, the predicted age or AMAR of the agent-of-interest-treated organism/individual or cultured cells, tissue, or organ with the predicted age or AMAR of the untreated or placebo-treated organism/individual or cultured cells, tissue, or organ, such that a lower value or range of values for the agent-of-interest-treated organism/individual or cultured cells, tissue, or organ indicates that an agent of interest can retard an aging process whereas a higher value or range of values indicates an agent of interest can accelerate aging process. The agent of interest may be an anti-oxidant, reducing agent, DNA damaging agent, vitamin, dietary supplement, food, food

additive, food coloring, salt, vegetable, vegetable extract, fruit, fruit extract, flower, flower extract, fragrance, seed, seed extract, herb, herb extract, plant extract, fiber, fat, fatty acid, oil, sugar, artificial sweetener, probiotics, alcohol, wine, fungus, mold, cream, lotion, powder, makeup, sun blocker, gas, pollutant, smoke, environmental pollutant, paint, solvent, organic solvent, plastic, plasticizers, bisphenol, phenolic compounds, tobacco, inhalant, drug, biologic, hormone, endocrine disruptor, environmental estrogen, hormone antagonist, hormone agonist, caffeine, phytoestrogen, metal, enzyme, chelator, yogurt, sulfur compound, physical barrier, electromagnetic barrier, and radiation barrier.

In an embodiment, the organism may be yeast, fruit fly, fish, worm, insect, zebra fish, nematode, plant, or mammal. Mammal includes, but is not limited to, human, murine, simian, feline, canine, equine, bovine, porcine, ovine, caprine, rabbit, mammalian farm animal, mammalian sport animal, and mammalian pet.

In one embodiment, the invention provides methods for identifying type of tissue for a biological sample from a subject with a known chronological age. The method comprises ascertaining the chronological age of a subject and determining the predicted age of the subject from the biological sample by the method of the invention. Additionally, the method comprises comparing to a reference standard relating the predicted age for various types of tissue to chronological age and determining which value closely matches the predicted age in the reference standard for various types of tissue. Further, the method comprises assigning the type of tissue for the biological sample based on the closest match.

The invention also provides methods for identifying type of tissue for a biological sample from a subject with a known chronological age. The method comprises ascertaining the chronological age of a subject and determining the AMAR of the subject from the biological sample by dividing the predicted age of a subject from the chronological age of the subject. Additionally, the method comprises comparing to a reference standard relating the AMAR to chronological age for various types of tissue and determining which value closely matches the AMAR in the reference standard for various types of tissue. The method further comprises assigning the type of tissue for the biological sample based on the closest match.

In one embodiment, the set of age-associated epigenetic marker(s) comprises any one or more of methylation marker cg23606718 of Figure 9, Tables S3 and S5 and/or cgl 6867657 of Figure 9, Tables S3 and S5.

The invention further provides methods for predicting age of a subject based on age-associated epigenetic modification affecting gene expression. The method comprises obtaining a biological sample of the subject and determining the expression of one or more gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age. Additionally, the method comprises comparing the expression of one or more gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age with the expression of the same gene(s) from an age-correlated reference population. The method further comprises obtaining a value or range of values for the predicted age of the subject. Comparing the expression of one or more gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age with the expression of the same gene(s) from an age-correlated reference population may comprise any statistical method, multivariate regression method, linear regression analysis, tabular method, or graphical method used to predict the age of a subject based on expression of gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age. In one embodiment, the statistical method may be a multivariate regression algorithm or linear regression algorithm.

In another embodiment, one or more gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age may comprise one or more of the genes listed in Table S6 or Table S7.

In another embodiment, the gene expression may be a transcription or translation. In another embodiment, the transcription results in the production of RNA transcripts and translation results in the production of proteins.

In accordance with the practice of the invention, the invention provides a method of screening a tissue sample from a subject in order to predict the age of the tissue sample based on the epigenome of the subject by the method the invention.

In one embodiment, the tissue sample may be exposed to at least one test agent in a high-throughput screening assay. In another embodiment, said process may be used for any one of diagnosis and/or high-throughput screening.

The invention also provides methods for predicting age of a tissue or organ of a subject based on the epigenome of the tissue or organ of the subject. The method comprises obtaining a biological sample of a tissue or organ from the subject and determining the methylation status of a set of age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5. The method further comprises comparing the methylation status of a set of age-associated epigenetic marker(s) of the subject with the methylation status of the same markers from an age-correlated reference population so as to obtain a value or a range of values for the predicted age of the tissue or orgaa

The methylation status of the same markers from an age-correlated reference population may be determined on a same or a different type of tissue or organ. The methylation status of the same markers from an age-correlated reference population may be determined on blood or fractionated blood.

In an embodiment, the methods of the invention provides for determining differential aging rates of tissues or organs of a subject. The method comprises obtaining biological samples from different tissue(s) or organ(s) from the subject and predicting the age of the tissue or organ using the methods of the invention. The method further comprises comparing the predicted ages where a difference in the predicted ages indicates a difference in the aging rate of the tissue(s) or organ(s) of the subject. The predicted age may be divided by the chronological age of the subject to obtain AMAR.

COMPOSITIONS OF THE INVENTION

The invention further provides compositions which comprise a set of epigenetic markers based on five or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; ten or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; fifteen or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; twenty or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; twenty-five or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; thirty or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; thirty-five or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; forty or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; forty-five or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; fifty or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; fifty-five or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; sixty or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; sixty-five or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; or seventy or more age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5.

Further, in a preferred embodiment, the composition may comprise a set of epigenetic markers based on five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; ten or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; fifteen or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; twenty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; twenty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; thirty

or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; thirty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; forty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; forty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; fifty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; fifty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; sixty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; sixty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; or seventy or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3. For example, the set of age-associated epigenetic marker(s) may comprise any one or more of the following methylation marker cg05652533 of Table S4, cg27367526 of Table S4, cgl8404041 of Table S4, cg23606718 of Figure 9, Tables S3 and S5, cgl6867657 of Figure 9, Tables S3 and S5, cg04474832 on chromosome 3 at position 52008487, cg05442902 on chromosome 22 at position 21369010, cg06493994 on chromosome 6 at position 25652602, cg09809672 on chromosome 1 at position 236557682, cgl9722847 on chromosome 12 at position 308491 14, cg22736354 on chromosome 6 at position 18122719, cg05652533 of Table S4, cg27367526 of Table S4, cgl8404041 of Table S4, cg23606718 on chromosome 2 at position 131513927, and/or cgl6867657 of chromosome 6 at position 11044877.

In yet another embodiment, the composition comprises a set of age-associated epigenetic marker(s) of methylation marker cg04474832 on chromosome 3 at position 52008487, cg05442902 on chromosome 22 at position 21369010, cg06493994 on chromosome 6 at position 25652602, cg09809672 on chromosome 1 at position 236557682, cgl9722847 on chromosome 12 at position 308491 14, and cg22736354 on chromosome 6 at position 18122719.

In another embodiment, the composition comprises a set of age-associated epigenetic marker(s) of any one or more of methylation marker cg20822990 of Figure 9 or Table S3, cg04400972 of Figure 9 or Table S3, cgl6054275 of Figure 9 or Table S3, cg036071 17 of Figure 9 or Table S3, cg20052760 of Figure 9 or Table S3, cgl6867657 of Figure 9 or Table S3, cg06493994 of Figure

9 or Table S3, cg06685111 of Figure 9 or Table S3, cg00486113 of Figure 9 or Table S3, cg20426994 of Figure 9 or Table S3, eg14361627 of Figure 9 or Table S3, cg08097417 of Figure 9 or Table S3, cg07955995 of Figure 9 or Table S3, cg22285878 of Figure 9 or Table S3 and/or cg08540945 of Figure 9 or Table S3.

In further embodiment, the composition comprises a set of age-associated epigenetic marker(s) of any one or more of methylation marker cg23606718 of Figure 9, Tables S3 and S5 and/or egl 6867657 of Figure 9, Tables S3 and S5.

KITS

According to another aspect of the invention, kits are provided. Kits according to the invention include package(s) comprising compounds or compositions of the invention.

The phrase "package" means any vessel containing compounds or compositions presented herein. In preferred embodiments, the package can be a box or wrapping. Packaging materials for use in packaging pharmaceutical products are well known to those of skill in the art. Examples of pharmaceutical packaging materials include, but are not limited to, blister packs, bottles, tubes, inhalers, pumps, bags, vials, containers, syringes, bottles, and any packaging material suitable for a selected formulation and intended mode of administration and treatment.

The kit can also contain items that are not contained within the package but are attached to the outside of the package, for example, pipettes.

Kits may optionally contain instructions for administering compounds or compositions of the present invention to a subject having a condition in need of treatment. Kits may also comprise instructions for approved uses of compounds herein by regulatory agencies, such as the United States Food and Drug Administration. Kits may optionally contain labeling or product inserts for the present compounds. The package(s) and/or any product insert(s) may themselves be approved by regulatory agencies. The kits can include compounds in the solid phase or in a liquid phase (such as buffers provided) in a package. The kits also can include buffers for

preparing solutions for conducting the methods, and pipettes for transferring liquids from one container to another.

The kit may optionally also contain one or more other compounds for use in combination therapies as described herein. In certain embodiments, the package(s) is a container for intravenous administration. In other embodiments, compounds are provided in an inhaler. In still other embodiments compounds are provided in a polymeric matrix or in the form of a liposome.

The invention provides for a kit for determining age of a subject based on epigenetic modification of subject's genetic material comprising the set of age-associated epigenetic marker or markers as listed in Figure 9, Table S3, Table S4 or Table S5 as described in the methods of the invention, supra.

The invention further provides for a kit for predicting age of a subject based on the epigenome of the subject utilizing the set of the age-associated epigenetic marker(s) provided in Figure 9, Table S3, S4 and/or S5 as described in the methods of the invention, supra.

In one embodiment, the age-associated epigenetic marker(s) may comprise a nucleic acid with a CpG dinucleotide. In another embodiment, the cytosine of the CpG dinucleotide may be subject to age-dependent changes in methylation at the C-5 position. In another embodiment, the CpG dinucleotide is at the chromosomal position as indicated in Figure 9, Table S3, S4, and/or S5.

In an embodiment, the age-associated epigenetic marker(s) may be a human marker and selected from cg04474832 on chromosome 3 at position 52008487, cg05442902 on chromosome 22 at position 21369010, cg06493994 on chromosome 6 at position 25652602, cg09809672 on chromosome 1 at position 236557682, cgl9722847 on chromosome 12 at position 308491 14, cg22736354 on chromosome 6 at position 18122719, cg05652533 of Table S4, cg27367526 of Table S4, cgl8404041 of Table S4, cg23606718 on chromosome 2 at position 131513927, and cgl 6867657 of chromosome 6 at position 11044877.

In yet another embodiment, the age-associated epigenetic marker(s) may have the sequence as provided in Figure 9 or as can be found at the National Center for Biotechnology Information of the National Institutes of Health (Bethesda, MD) in the Gene Expression Omnibus (GEO) database with GEO accession number GPL13534.

The following examples are provided to further illustrate aspects of the invention. These examples are non-limiting and should not be construed as limiting any aspect of the invention.

EXAMPLES

EXAMPLE 1

Materials and Methods

Sample Collection and Test Procedures

This study was approved by the institutional review boards of the University of California, San Diego; the University of Southern California; and West China Hospital. All participants signed informed consent statements prior to participation. Blood was drawn from a vein in the patient's arm into blood collection tubes containing the anticoagulant acid citrate dextrose. Genomic DNA was extracted from the whole blood with a QIAGEN FlexiGene DNA Kit and stored at -20°C. Methylation fraction values for the autosomal chromosomes were measured with the Illumina Infinium HumanMethylation450 BeadChip (Bibikova *et al.,* 201 1). This procedure uses bisulfate-treated DNA and two site-specific probes for each marker, which bind to the associated methylated and unmethylated sequences. The intensity of the methylated probe relative to the total probe intensity for each site represents the fractional level of methylation at that site in the sample. These values were adjusted for internal controls with Illumina' s Genome Studio software. Methylation fraction values with a detection p value greater than 0.01 were set to "missing." One patient sample and 830 markers were removed as they had greater than 5% missing values. The remaining missing values were imputed with the KNN approach (ten nearest markers) using the R "impute" package (Troyanskaya *et al.,* 2001). We performed exome sequencing on 258 of these samples, using a solution hybrid selection method to capture DNA

followed by parallel sequencing on an Illumina HiSeq platform. Genotype calls were made with the SOAP program (Li *et ah,* 2008). Calls with a quality score less than twenty were set as missing. Only variants that had fewer than 10% missing calls, were within Hardy-Weinberg equilibrium ($p \leq 10^{-4}$), and were of a common frequency (>5%) were retained (10,694). Individuals with less than 20% missing calls (252) were retained. Additional genotyping was done with multiplex PCR followed by MALDI-TOF mass spectrometry analysis with the iPLEX/MassARRAY/Typer platform.

Methylation Quality Control

We used principal component (PC) analysis to identify and remove outlier samples. We converted each sample into a z score statistic, based on the squared distance of its 1st PC from the population mean. The z statistic was converted to a false-discovery rate with the Gaussian cumulative distribution and the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995). Samples falling below an FDR of 0.2 were designated at outliers and removed. This filtering procedure was performed iteratively until no samples were determined to be an outlier. A total of 24 samples were removed in this manner.

Association Testing

Association tests for trends in methylation fraction and deviance were performed with nested linear models and the F test. As methylation levels may be sensitive to a number of factors, we included several covariates, including gender, BMI, diabetes status, ethnicity, and batch. Tests for whole-methylome changes in deviance were computed with the binomial test, based on the number of markers with a positive rather than negative coefficient. Markers were annotated as having support from the TCGA data when the coefficient of aging was the same sign and the significance was better than $p < 0.05$.

Annotation Enrichment

Methylation marker annotations for CpG islands and GO terms were obtained from the IlluminaHumanMethylation450k.db database from Bioconductor (Gentleman *et ah,* 2004). Annotation enrichment tests were performed with the two-sided Fisher's exact test.

Aging Model

The diagnostic model of age was made with a multivariate linear model approach based on the Elastic Net algorithm implemented in the R package "glmnet" (Friedman *et al.*, 2010). This approach is a combination of traditional Lasso and ridge regression methods, emphasizing model sparsity while appropriately balancing the contributions of correlated variables. It is ideal for building linear models in situations where the number of variables (markers) greatly outweighs the number of samples. Optimal regularization parameters were estimated via 10-fold crossvalidation. We employed bootstrap analysis, sampling the data set with replacement 500 times and building a model for each bootstrap cohort. We included in the final model only markers that were present in more than half of all bootstraps. The covariates gender, BMI, diabetes status, ethnicity, and batch were included in the model and were exempted from penalization (regularization). p values are based on a least-squares model built with the same terms and drop-one F tests. As BMI was strongly associated with age, the term was first adjusted for age before computing significance in the model. AMAR was computed with the aging model, but without the variables of gender, BMI, and diabetes status. The coefficients were not changed. AMAR was then taken as an individual's predicted age divided by her or his actual age.

Genetic VariantAssociations

Each genetic variant was tested for association in an additive model with the top aging-associated methylation markers with nested linear models and the F test. We included covariates for gender, BMI, diabetes status, ethnicity, and batch. Variant positions were based on the human reference build GRCh37 and gene annotations were based on chromosomal proximity within 20 kbp.

Computing Methylation Deviance

Methylation deviance was computed via the following approach: First, we removed the methylation trends due to all given variables, including age, gender, and BMI by fitting a linear model for each marker and acting only on the residuals. Next, we identified and removed highly nonnormal markers on the basis of the Shapiro-Wilk test ($p < 10^{-5}$). To allow for naturally occurring extreme deviations in the normality test, we first estimated the outliers of each marker based on a Grubb's statistic, choosing the inclusion threshold based on the Benjamini-Hochberg

FDR (Benjamini and Hochberg, 1995). If any samples had an FDR less than 0.4,weignored them and repeated the outlier detection until no outliers were detected. Finally, the deviance of each remaining marker was computed as the square of its adjusted methylation value.

Entropy Analysis

Entropy statistics were computed on methylation data adjusted for covariates and filtered for normality (see Computing Methylation Deviance). We computed the normalized Shannon entropy (Shannon and Weaver, 1963) of an individual's methylome according to the formula

$$\textbf{Entropy} = \frac{1}{N * \log(\frac{1}{2})} \sum_{i} [MFi * \log(MFi) + (1 - MFi) * \log(1 - MFi)],$$

where *MFi* is the methylation fraction of the i^{th} methylation marker and N is the number of markers.

Mapping CpG Islands

Genomic positions and marker annotations for 27,176 CpG islands were obtained from the IlluminaHumanMethylation450k.db database from Bioconductor (Gentleman *et al.*, 2004). We obtained the positions for markers within each island with at least four markers (25,028), as well as the nearest 100 markers upstream and downstream. These positions were then combined with the marker value of interest (i.e., methylation fraction, aging coefficient, or deviance) to produce a genomic map for each island and the surrounding region. After normalizing each map to the center of the island, we averaged the values at each relative genomic point across all islands to produce a common map.

Results and Discussion

Global Methylomic Profiling over a Wide Age Range

We obtained methylome-wide profiles of two different cohorts (Nl = 482, N2 = 174) sampled from a mixed population of 426 Caucasian and 230 Hispanic individuals, aged 19 to 101. Samples were taken as whole blood and processed with the Illumina Infinium HumanMethylation450 BeadChip assay (Bibikova *et al*, 201 1), which measures the methylation states of 485,577 CpG markers. Methylation was recorded as a fraction between zero and one,

representing the frequency of methylation of a given CpG marker across the population of blood cells taken from a single individual. Conservative quality controls were applied to filter spurious markers and samples. For simplicity, we discarded values for markers on sex chromosomes. Association tests revealed that 70,387 (15%) of the markers had significant associations between methylation fraction and age (Figure 1, false discovery rate [FDR] < 0.05 by F test). We were able to verify at a $p < 0.05$ significance level 53,670 (76%) of these associations using 40 young and old samples recently published by Heyn *et al.* (2012). More detailed accounts of the individual aging markers and their genomic features are presented in the Figure SI and Tables SI and S2. The resulting data set represents the largest and highest-resolution collection of methylation data produced for the study of aging, providing an unprecedented opportunity to understand the role of epigenetics in the aging process. The complete methylation profiles are available at the Gene Expression Omnibus (GSE40279).

Table SI, Functional annotations for age-associated markers, related to Figure 1

Genes with nearby age-associated markers were enriched for many functions. A selection of these functions are shown here.

Over-enriched Terms	Under-enriched Terms
Cell communication (FDR ~ 10^{-76}) Locomotion (FDR < 10^{-33}) Cell proliferation (FDR < 10^{-17}) Growth (FDR < 10^{-7})	G-protein coupled receptor activity (FDR < 10^{-11}) Ribosome (FDR < 10^{-5}) RNA splicing (FDR < 0.05) M phase (FDR < 0.05)

Table S2, Age-associated marker properties, related to Figure 1

A table of age-associated markers and their coincidence with several genomic features. Each value represents the percentage of the age-associated markers of a particular type (columns) that are coincident with a particular annotation (rows).

	All Markers (%)	Increasing Mean (%)	Decreasing Mean (%)	Increasing Variance (%)
MF > 0.5	55	13	57	51
MF < 0.5	45	87	43	49

	All Markers (%)	Increasing Mean (%)	Decreasing Mean (%)	Increasing Variance (%)
CpG Island	31	68	8	24
CpG Shore	23	18	35	36
CpG Shelf	10	3	13	8
Enhancer	22	22	30	27
Promoter	20	7	13	8
DHS	12	30	14	16

A Predictive Model for the Aging Methylome

We built a predictive model of aging on the primary cohort using a penalized multivariate regression method known as Elastic Net (Zou and Hastie, 2005), combined with bootstrap approaches. The model included both methylomic and clinical parameters such as gender and body mass index (BMI) (Figure 2A). The optimal model selected a set of 71 methylation markers that were highly predictive of age (Figure 2A and Table S3). The accuracy of the model was high, with a correlation between age and predicted age of 96% and an error of 3.9 years (Figure 2B). Nearly all markers in the model lay within or near genes with known functions in aging-related conditions, including Alzheimer's disease, cancer, tissue degradation, DNA damage, and oxidative stress. By way of example, two markers lay within the gene somatostatin *(SST)*, a key regulator of endocrine and nervous system function (Yacubova and Komuro, 2002). SST is known to decline with age and has been linked to Alzheimer's disease (Saito *et al.*, 2005). As a second example, six model markers lay within the transcription factor KLF14, which has been called a "master regulator" of obesity and other metabolic traits (Small *et al.*, 2011). Given the links between aging, longevity, and metabolic activity (Lane *et al.*, 1996; Tatar *et al.*, 2003), it is not surprising that several of our model markers are implicated in obesity and metabolism.

Table S3, Aging model markers, related to Figure 2

A table of the methylation markers included in the primary aging model. The coefficient listed for each marker is its regression coefficient within the model. A second table is provided for the model based on all samples (primary and validation).

Marker	Chromosome	Position	Genes	CpG Island	Coefficient
cg20822990	1	17338766	ATP13A2,SDHB	No	-15.7

Marker	Chromosome	Position	Genes	CpG Island	Coefficient
cg22512670	1	26855765	RPS6KA1	No	1.05
cg25410668	1	28241577	RPA2,SMPDL3B	No	3.87
Cg04400972	1	117665053	TRIM45,TTF2	Yes	9.62
cgl6054275	1	169556022	F5,SELP	No	-11.1
cgl0501210	1	207997020	Clorfl32	No	-6.46
cg09809672	1	236557682	EDARADD	No	-0.74
ch.2.30415474F	2	30561970	LBH	No	5.79
Cg22 158769	2	39187539	ARHGEF33	Yes	-2.06
cg02085953	2	97202260	ARID5A	No	1.02
Cg06639320	2	106015739	FHL2	Yes	8.95
cg22454769	2	106015767	FHL2	Yes	4.85
cg24079702	2	106015771	FHL2	Yes	2.48
cg23606718	2	131513927	FAM123C	Yes	8.35
cg22016779	2	230452311	DNER,RN U7-9P	No	1.79
Cg04474832	3	52008487	ABHD14A,ABHD14B, ACYl, GPR62,PCBP4,RPL29	No	-7.1
cg03607117	3	53080440	SFM BT1	Yes	10.7
cg07553761	3	160167977	SM C4,TRIM 59	Yes	3.72
cg00481951	3	187387650	SST	No	-2.72
Cg25478614	3	187387866	SST	No	4.01
cg25428494	4	84255411	HPSE	No	-1.81
cgO2650266	4	147558239	POU4F2	Yes	10.2
Cg08234504	5	139013317	CTD-3224K15.2,CXXC5, UBE2 D2	No	-3.16
Cg23500537	5	140419819	PCDHB1	No	5.67
Cg20052760	6	10510789	GCNT2	No	-12.6
cgl6867657	6	11044877	ELOVL2	Yes	10.8
cg22736354	6	18122719	NHLRC1JPMT	Yes	4.42
cg06493994	6	25652602	SCG N	Yes	9.42
Cg06685111	6	30295466	XXbac-BPG283016.8	No	-13.1
cg00486113	6	31105711		No	-10.7
cgl3001142	6	147528521	STXBP5	No	-5.8
Cg20426994	7	130418324	KLF14	Yes	19.1
cgl4361627	7	130419116	KLF14	Yes	10.7
cg08097417	7	130419133	KLF14	Yes	27.3
cg07955995	7	130419159	KLF14	Yes	13.7
Cg22285878	7	130419173	KLF14	Yes	-20.7
cg03473532	7	131008743	M KLN1	No	-3.31
cg08540945	7	152591698		Yes	9.41
cg07927379	7	156433108	C7orfl3,RN F32	Yes	-1.42

Marker	Chromosome	Position	Genes	CpG Island	Coefficient
cgl6419235	8	57360613	PENK	Yes	-1.6
cg07583137	8	82644012	CHMP4C,ZFAND1	No	3.03
Cg22796704	10	49673534	ARHGAP22	No	-10.6
cgl9935065	10	98062687	DNTT	No	13.4
cg23091758	11	9025767	NRIP3,SCUBE2	Yes	-0.392
Cg23744638	11	10323902	ADM,AMPD3,SBF2	No	0.0859
Cg04940570	11	12696758	TEAD1	Yes	11.6
Cgll067179	11	66083541	CD248,RIN1,TMEM151A	No	14.7
Cg22213242	11	66083573	CD248,RIN1,TMEM151A	Yes	23.7
Cg06419846	11	66083697	CD248,RIN1,TMEM151A	Yes	13.4
cg02046143	11	133797911	IGSF9B	No	-10.2
Cg00748589	12	11653486		Yes	8.21
cgl9722847	12	30849114	CAPRIN2,IP08	No	-5.66
cgl8473521	12	54448265	HOXC4,HOXC5	No	8.85
cg01528542	12	81468232	ACSS3	No	-2.98
ch.l3.39564907R	13	40666907		No	-20.6
cg03032497	14	61108227	SIX1	No	8.4
Cg04875128	15	31775895	OTUD7A	Yes	-4.37
cg21296230	15	33010536	GREM1	Yes	8.39
cg09651136	15	72525012	PARP6,PKM2	No	-15.8
cg03399905	15	79576060	ANKRD34C	Yes	28
cg04416734	16	30075192	ALDOA,PPP4C	No	11.9
cg07082267	16	85429035		No	2.87
cgl4692377	17	28562685	BLMH,SLC6A4,SNORD63.3	Yes	19.1
cg06874016	17	40177415	DNAJC7,NKIRAS2,ZNF385C	No	-4.37
cg21139312	17	55663225	MSI2	No	17.1
Cg02867102	17	62398693		No	-12.5
cgl9283806	18	66389420	CCDC102BJMX3	No	-4.29
cgl4556683	19	15342982	BRD4,EPHX3	Yes	-4.04
cg07547549	20	44658225	MMP9,SLC12A5	Yes	3.11
cg05442902	22	21369010	LZTR1,MIR649,P2RX6, SLC7A4JHAP7	No	-22.7
Cg08415592	22	36648973	APOLl,APOL2,Z82215.1	No	-6.92
Cg24724428					

We validated this model on the secondary cohort, consisting of an additional 174 independent samples. These samples were processed in the same manner as the primary cohort and were then used to predict age based on the original model (i.e., as trained on the original cohort). The predictions were highly accurate, with a correlation between age and predicted age of 91% and

an error of 4.9 years (Figure 2C). The significance of the aging model was also confirmed by the data set presented in Heyn *et al,* verifying the age association of 70 of the 71 markers (Heyn *et al,* 2012). Furthermore, the model was able to fully separate old and young individuals in the Heyn *et al.* study, even for profiles obtained via bisulfate sequencing rather than the bead-chip technology used in this study (Figure S2).

Methylome Aging Rate and Its Associations

While the aging model is able to predict the age of most individuals with high accuracy, it is equally valuable as a tool for identifying individual outliers who do not follow the expectation. For example, Figure 2B highlights two individuals whose age is vastly over- or under-predicted on the basis of their methylation data. To examine whether these differences reflect true biological differences in the state of the individual (i.e., versus measurement error or intrinsic variability), we used the aging model to quantify each individual's *apparent methylomic aging rate* (AMAR), defined as the ratio of the predicted age, based on methylation data, to the chronological age. We then tested for associations between AMAR and possibly relevant clinical factors, including gender and BMI. Analysis of ethnicity and diabetes status was not possible due to correlations with batch variables (Figure S3). We found that gender, but not BMI had significant contributions to aging rate (F test, $p = 6 \times 10^{-6}$, $p > 0.05$). The methylome of men appeared to age approximately 4% faster than that of women (Figure 2D), even though the overall distributions of age were not significantly different between the men and women in the cohort ($p > 0.05$, KS test). Likewise, the validation cohort confirmed the increased aging rate for men ($p < 0.05$), but was inconclusive for BMI ($p > 0.05$). This complements a previous finding of an epigenetic signal for BMI that does not change with age (Feinberg *et al.,* 2010).

As genetic associations have been previously reported with human longevity and aging phenotypes (Atzmon *et al,* 2006; Suh *et al,* 2008; Willcox *et al,* 2008; Wheeler *et al,* 2009), we examined whether the model could distinguish aging rates for individuals with different genetic variants. For this purpose, we obtained whole-exome sequences for 252 of the individuals in our methylome study at 15x coverage. After sequence processing and quality control, these sequences yielded 10,694 common single-nucleotide variants across the population. As a negative control, we confirmed that none of the genetic variants were

significant predictors of age itself, which is to be expected since the genome sequence is considered to be relatively static over the course of a lifetime. On the other hand, one might expect to find genetic variants that modulate the methylation of age-associated markers, i.e., methylation quantitative-trait loci or meQTLs (Bell *et al,* 201 1). Testing each genetic variant for association with the top age-associated methylation markers, we identified 303 meQTLs (FDR < 0.05, Figure 3A). For validation, we selected eight genetic variants (corresponding to 14 meQTLs) to test in a validation cohort of 322 individuals from our methylation study. This analysis found that seven of eight genetic variants (corresponding to seven meQTLs) remained highly significant in the validation cohort (FDR < 0.05, Table S4). While all of these variants acted in cis with their meQTLs (within 150 kbp), we confirmed that none directly modified the CpG site or associated probe sequence of the associated methylation marker.

Table S4, Genetic variants influencing age-associated methylation, related to Figure 3

A table of the genetic variants which were found to influence age-associated methylation. Distance is the genomic distance from the genetic marker to the methylation marker. Association values are listed as p-values. AMAR association is the significance of the association between the genetic marker and AMAR.

Genetic Marker	Genetic Genes	Methylation Marker	Methylation Genes	Distance (bp)	Meth-Age Association	Meth-Geno Association	AMAR Association
rs2818384	JAKMIP3	cg05652533	DPYSL4, JAKMIP3	38793	5.86×10^{-9}	3.73×10^{-21}	0.00133
rs42663	GTPBP10	cg27367526	STEAP2	142116	1.44×10^{-18}	8.05×10^{-22}	0.00476
rs2230534	ITIH1, NEK4	cg18404041	ITIH1, ITIH3, NEK4	21881	6.78×10^{-14}	1.26×10^{-87}	0.02125
rs17152433	CTBP2, ZRANB1	cg07906193		70390	8.51×10^{-9}	7.54×10^{-56}	0.05273
rs1058405	ATF6	cg19735514	ATF6	10998	4.50×10^{-11}	3.87×10^{-63}	0.55546
rs57913893	ACSS1, C20orf3	cg26306437	C20orf3, CST7	59756	5.20×10^{-8}	5.50×10^{-16}	0.80327
rs6115003	ACSS1	cg26306437	C20orf3, CST7	70693	5.20×10^{-8}	4.17×10^{-17}	0.92096

The methylation marker cg27 193080 was one of those found to be significantly associated with age ($p < 1O^{17}$), and its methylation fraction was found to be influenced by the single-nucleotide polymorphism (SNP) variant rs140692 ($p < 10^{-21}$) (Figure 3B). This meQTL was particularly interesting as both the SNP and the methylation marker mapped to the gene *methyl-CpG binding*

domain protein 4 {MBD4, with the SNP in an intron and the methylation marker just upstream of the coding region), one of the few known genes encoding a protein that can bind to methylated DNA. This meQTL thus captures a *cis* relationship in which rs140692 influences the methylation state of *MBD4.* That MBD4 plays a role in human aging is supported by previous work linking *MBD4* to DNA repair, as well as work showing that mutations and knockdowns of MBD4 lead to increased genomic instability (Bellacosa *et al.,* 1999; Bertoni *et al,* 2009).

Of the seven validated meQTLs, three were identified that had a statistically significant association not only with age but also with aging rate (AMAR, FDR < 0.05, Figures 3B and 3C). One is the genetic marker rs2230534, which is a synonymous mutation in the gene *NEK4,* and has a cis association with the methylation marker cgl 8404041. The *NEK* family of kinases plays a key role in cell-cycje regulation and cancer (Moniz *et al.,* 201 1). The second variant is rs2818384, which is a synonymous mutation in the gene *JAKMIP3* and has a cis association with the methylation marker cg05652533. Copy-number variants in *JAKMIP3* have been previously associated with glioblastoma (Xiong *et al.,* 2010). The final variant found to influence AMAR is rs42663, which is a missense mutation in the gene *GTPBP10* and associates with cg27367526 in the gene *STEAP2. STEAP2* is known to play a role in maintaining homeostasis of iron and copper—metals that serve as essential components of the mitochondrial respiratory chain (Ohgami *et al.,* 2006). Studies have shown that perturbations of iron concentrations can induce DNA damage through oxidative stress in mammalian cells (Hartwig and Schlepegrell, 1995; Karthikeyan *et al.,* 2002). These meQTLs represent genetic variants that appear to broadly influence the aging methylome and may be good candidates for further age-associated disease and longevity research.

A Multitissue Diagnostic

Our aging model was derived from whole blood, which is advantageous in the design of practical diagnostics and for testing samples collected from other studies. To investigate whether our aging model was representative of other tissues, we obtained DNA methylation profiles for 368 individuals in the control category of The Cancer Genome Atlas (TCGA) (Collins and Barker, 2007), including 83 breast, 183 kidney, 60 lung, and 42 skin samples. An aging model based on both our primary and validation cohorts demonstrated strong predictive power for chronological

age in these samples (expected value R = 0.72), although each tissue had a clear linear offset (intercept and slope) from the expectation (Figure 4A). This offset was consistent within a tissue, even across different batches of the TCGA data. We adjusted for each tissue trend using a simple linear model, producing age predictions with an error comparable to that found in blood (Figure 4B). Furthermore, predicted AMARs in each tissue supported the effect of men appearing to age more quickly than women ($p < 0.05$). Thus, computation of aging rate (AMAR) from blood samples reflects trends that are not specific to blood and may be common throughout many tissues of the human body. Furthermore, this analysis provides evidence that the observed methylomic changes are intrinsic to the methylome and not due primarily to cell heterogeneity, i.e., changing cell-type composition of whole blood with age. In this regard, this study is consistent with a prior analysis of purified CD4+ T cells and CD14+ monocytes, in which the age-associated epigenetic modifications were found to be similar to the changes observed in whole blood (Rakyan *et al.*, 2010).

To investigate the similarities and differences between the tissues, we built age models de novo for breast, kidney, and lung tissues (Table S5; the skin cohort had too few samples to build a model). Most of the markers in the models differed, although all of these models and the primary model share the markers cg23606718 and cgl6867657. These markers are both annotated to the gene *ELOVL2*, which has been linked to the photoaging response in human skin (Kim *et ah,* 2010).

Table S5, Aging model markers for TCGA data, related to Figure 4

To investigate the similarities and differences between the tissues, we built an age model for breast, kidney, and lung tissues. The skin cohort did not have enough samples to build a model. The markers and coefficients of each model are listed here.

Marker	Chromosome	Position	Genes	CpG Island	Coefficient
cg23040782	1	6762215	DNAJC11	No	-7.45
cg11197101	1	33219998	KIAA1522	Yes	8.73
cg00252781	1	179334658	C1orf125,SOAT1	No	13.2
cg16909962	1	229406711	RAB4A,TMEM78	Yes	27.6
cg23606718	2	131513927	FAM123C	Yes	25.1

Marker	Chromosome	Position	Genes	CpG Island	Coefficient
Cg03545227	2	220173100	MIR153-1,PTPRN,RESP18	Yes	13.5
cg00702638	3	44803293	KiAA1143,KIF15	Yes	53.1
cg05555455	3	148804550	HLTF,Y_RNA.240	Yes	-18.5
Cg03844506	4	4109441		Yes	-7.72
cgl6558177	4	4109446		Yes	-2.25
Cgll299854	5	132083184	CCNI2,KIF3A,SEPT8	Yes	43.6
cg05708550	5	137688227	CDC25C,FAM53C,KDM3B	Yes	8.4
cgl6867657	6	11044877	ELOVL2	Yes	23.1
Cg22736354	6	18122719	NHLRC1JPMT	Yes	-15
cgl4848772	6	27099813	HIST1H2AG,HIST1H2AH, HIST1H2BJ,HIST1H2BK,HIST1H4I	No	-14.3
cgl5623062	6	31747133	Y_RNA.307	No	32.2
cgl6489193	6	33240059		Yes	25.5
cgl8468088	6	35490818	TULP1	No	5.96
Cg04911280	6	44281184	AARS2JCTE1,TMEM151B	Yes	-9.67
cgl9291355	6	44281188	AARS2JCTE1JMEM151B	Yes	-5.16
cg05917988	6	44281197	AARS2JCTE1,TMEM151B	Yes	2.77
cg20160885	7	5013524	MMD2,RBAK,RNF216L	Yes	49.6
Cgl9230755	7	65878503	GS1-124K5.2,GS1-124K5.6, NCRNA00174,SKP1P1,TPST1, U6.862	Yes	-2.26
Cg09941452	7	97557827		Yes	-14.2
cg26830108	7	100813299	APlSl,C7orf52,VGF	No	1.22
Cgl9273773	7	102790112	NAPEPLD	Yes	1.56
cgl4361627	7	130419116	KLF14	Yes	1.17
Cg08097417	7	130419133	KLF14	Yes	53.5
Cg02821342	7	130793551	MKLN1	No	-19.8
cg07392449	8	11324666	FAM167A	Yes	73
cg083 18076	8	62051812	CLVS1	Yes	13
Cg02560186	11	3602584	OR7E117P	Yes	27.8
Cg08715791	11	66189297	MRPLI 1,NPAS4,SNORA43.2	Yes	-30.6
cg23156348	11	124981869	TMEM218	No	-21.1
cgl0820926	14	30397408	PRKD1	Yes	19.2
Cg06121469	15	44956098	PATL2,SPG11	No	-0.156
Cg07477282	15	44956107	PATL2,SPG11	No	2.22
cg21801378	15	72612125	CELF6	Yes	-49.5
cg02331561	16	2391081	ABCA17P,ABCA3	Yes	-17
cg06144905	17	27369780	PIPOX	No	9.13
cgl4692377	17	28562685	BLMH,SLC6A4,SNORD63.3	Yes	13.7
cgl8569335	17	40171970	DNAJC7,NK!RAS2,ZNF385C	Yes	-30.6

Marker	Chromosome	Position	Genes	CpG Island	Coefficient
cg26147554	18	712733	ENOSFl,YESl	Yes	17.7
cg21927946	19	4769688	C19orf30,MIR7-3	No	73.4
cgl5789607	19	4769690	C19orf30,MIR7-3	No	-15.4
cgl2589298	19	50828905	KCNC3,MYH14,NAPSB	Yes	15.3
cg06458239	19	58038573	MIR1274B,ZNF549,ZNF550, ZNF773	Yes	22.4
cgl0729426	19	58038585	MIR1274B,ZNF549,ZNF550, ZNF773	Yes	13.5
Cg26734668	19	58111094	ZIK1,ZNF134,ZNF530	No	9.2
cg22888484	20	37075185	SNHG11,SNORA39,SNORA60, SNORA71.3,SNORA71A, SNORA71C,SNORA71D	No	201

The TCGA data set also contains methylome profiles representing a total of 319 tumors and matched normal tissue samples (breast, kidney, lung, and skin). Interestingly, use of our aging model indicated that tumors appear to have aged 40% more than matched normal tissue from the same individual (Wilcox test, $p < 10^{-41}$, Figures 4C and 4D). Accelerated tumor aging was apparent regardless of the primary tissue type. We investigated whether this was the result of broad shifts in global methylation levels by examining all 70,387 age-associated markers, of which 44% tend to increase and 56% tend to decrease with age. Methylation fraction values in matched tumor and normal samples supported the finding that tumors coincide with older values for 74% of the markers regardless of the trending direction (binomial p ~0). Furthermore, separate aging models built in the matched normal and tumor samples confirm the apparent aging effect (Figure S4).

Different Asins Rates Lead to Divergent Methylomes

If individuals indeed age at different rates, it might be expected that their individual methylomes should diverge over time. This is based on the premise that the methylomes of the very young share certain similarities and that these similarities diminish as individuals accumulate changes over time. This effect, called epigenetic drift, has been observed in monozygotic twins (Fraga *et al.*, 2005), but few specific hypothesis have been put forth to account for it. To examine epigenetic drift in our samples, we computed the deviance of each methylation marker value as its squared distance from the expected population mean (Figure 5A). Then, in addition to testing

for markers whose methylation fraction changes with age (Figures 5B and 5C), we were able to test for markers whose deviance changes with age (Figures 5D and 5E) (Breusch and Pagan, 1979). Increasing deviance was a widespread phenomenon—we identified 27,800 markers for which the deviance was significantly associated with age (FDR < 0.05), of which 27,737 (99.8%) represented increased rather than decreased deviance (Figures 5E and S5). For any given individual, especially high or low methylome deviance was a strong predictor of aging rate (R = 0.47, p ~0), suggesting that differences in aging rates account for part of methylome heterogeneity and epigenetic drift.

Another way to examine epigenetic drift is in terms of Shannon entropy, or loss of information content in the methylome over time (Shannon and Weaver, 1963). An increase in entropy of a CpG marker means that its methylation state becomes less predictable across the population of cells, i.e., its methylation fraction tends toward 50%. Indeed, over all markers associated with a change in methylation fraction in the sample cohort, 70% tended toward a methylation fraction of 50% (Figure 6A, binomial p ~0, Table S2). Consequently, we observed a highly significant increase in methylome entropy over the sample cohort ($R = 0.21$, $p < 10^{-7}$). Furthermore, extreme methylome entropy for an individual was highly correlated with accelerated aging rate based on AMAR (R = 0.49, p ~0, Figure 6B).

Aging Rates and the Transcriptome

As changes in methylation have been directly linked to changes in gene expression (Sun *et al.*, 2011), we were interested in whether these changes in the aging methylome were mirrored on a functional level in the human transcriptome and reflected differences in aging rates. For this purpose, we obtained and analyzed publicly available gene expression profiles from the whole blood of 488 individuals spanning an age range of 20 to 75 (Emilsson *et al,* 2008). We found strong evidence for genes whose expression associates with age (326 genes, FDR < 0.05) and for genes with increasing expression deviance (binomial $p < 10^{-276}$). Strikingly, we found that genes with age-associated expression profiles were more likely to have nearby age-associated methylation markers in our data ($p < 0.01$, Table S6). We used this information to build a model of aging based on the expression of genes that were associated with age in the methylome (Figure 7A, Table S7). This model demonstrated a clear ability to measure aging rate using

expression data, reproducing our finding of increased aging rates for men as compared to women (Figure **7B, 11%** difference, $p < 10^{-4}$**).** The gender effect was not present in a model built using all available genes rather than those associated with age-related changes in the methylome ($p >$ **0.05).** Thus, age-associated changes to the methylome are indicative of functional changes in gene expression patterns.

Table S6, Genes associated with aging in both the methylome and the transcriptome, related to Figure 7

A list of genes which mapped to age-associated methylation markers and showed age-associated changes the transcriptome.

ABCA3	BFSP1	CDKN1C	DEPDC7	FCGBP	IL7R	MAN1C1
ABCB9	BHLHE40	CEBPG	DGKA	FG FBP2	INPP4B	M B21D2
ABLIM1	BLN K	CECR5	DLL1	FGFRL1	IRS1	MEOX1
ACAA2	BYSL	CENPE	DNASE1L3	FLN B	ITFG2	M EST
ACCN2	C10orfl28	CENPV	DNMT3A	FOXP1	ITGA6	M LF1
ACSF2	C12orf23	CH MP7	DNMT3B	FZD1	ITM2C	MPI
ACVR2A	C16orf45	CHSY3	DPH5	GAL3ST4	JAKM IP1	MS4A3
AEBP1	C17orf58	CIAPIN 1	DUSP4	GATA3	KAT2A	MS4A4A
AGBL2	Clorfl72	CISH	DYN LL1	GFI1	KATNAL1	MT1E
AGPAT4	Clorf21	CMC1	ECT2	GLT25D2	KCNM B4	MT1M
AK5	Clorf216	COBLL1	EDAR	GNG7	KIAA1841	MTSS1
ALDH5A1	ClorF51	COL5A3	EEF1G	GPC2	KLF4	MTUS1
ANKRD13B	C21orf63	CR2	EFNA1	GPR114	KLF6	MXRA8
ANKS6	C2orf40	CRIPl	EOM ES	GPR137B	KLH L14	MYC
ANXA1	C6orf97	CRTA M	EPHA1	GPR153	KLH L3	MY06
APBA2	CACHD1	CRTC3	EPHA2	GPR56	KLRG 1	MYOF
APBB1	CACNA2D2	CSF1R	EPHX2	GSC	LAMA5	NBEA
APOBEC3H	CALHM2	CST7	EPPK1	GTSF1	LBH	NCAPH
ARAP2	CAMK2 N1	CTLA4	EXPH5	GYG1	LDLRAP1	NEFH
ARHGEF4	CAPN2	CTN NA1	FAIM3	GZMH	LEF1	NELL2
ATP1B1	CCDC106	CTSL1	FAM 129C	HEXIM1	LGALS1	NHLRC1
B3GAT1	CCR10	CX3CR1	FAM 134B	HIST1H3D	LILRA4	NKG7
BACH2	CCR7	CYP2J2	FAM 13A	HOPX	LIMS2	NMT2
BATF3	CD200	CYP4F12	FAN1	IG FBP7	LM07	NM UR1
BCAS4	CD244	DCBLD2	FASLG	IGLL1	LPCAT1	NOB1
BCL7A	CD8B	DDB2	FBL	IL10RA	LRP11	NOP16
BCL9	CD9	DEFA4	FBLN2	IL16	LRRC32	NOSIP
BDH1	CDCA7L	DENND2D	FBX024	IL4I1	LTK	NPM3

NRCAM	PDGFRB	PRSS23	RHOC	SLC27A5	TCF7	VCAM 1
NSUN5	PELI3	PRSS35	RNASE2	SLC2A6	TCF7L2	VIT
NT5E	PHGDH	PTGDS	RNASE3	SLC45A3	TGFBR3	WARS
NTAN1	PHLDA3	PTGER2	RNF144A	SOCS2	TIGIT	ZBED3
NUAK1	PHYHD1	PTPRK	ROBOI	SORCS3	TM6SF1	ZFYVE28
OSBPL10	PI16	PTTG 1	RPL13	SOX15	TM EM 12 1	ZNF135
OXNAD1	PIK3IP1	PUS1	RUNX3	SPEG	TM EM8B	ZNF167
P2 RX5	PLAG1	PYROXD1	S100A10	SPIB	TM IGD2	ZNF177
PACSIN1	PLCG1	RAB15	S1PR5	SPINK2	TNFRSF17	ZNF263
PALLD	PLEKHA7	RAB27B	SATB1	SPN	TNFRSF25	ZNF285
PAQR4	PLXDC1	RAB6B	SCARB1	STOM	TPPP3	ZNF365
PCBP4	PM EPA1	RAB6C	SEC14L2	STX8	TRAF5	ZNF462
PCDH12	PM P22	RAD54B	SEMA3G	SUSD1	TRAP1	ZNF528
PCSK4	POMC	RAM P1	SFRP5	SYT11	TRIM2	ZNF544
PCSK5	POU2AF1	RAPG EF6	SFTPD	TARBP1	TSPAN13	ZNF551
PDE6B	PPAP2C	RASD1	SIRPG	TBX21	TSPAN2	ZSCAN18
PDE7A	PPM 1J	RASG EF1A	SLAM F7	TCAP	TWIST1	
PDE9A	PPP2R2 B	RGMA	SLC1A7	TCF3	TXNDC5	
PDG FD	PRR5L	RGS9	SLC23A1	TCF4	USP18	

Table S7, Transcriptome aging model, related to Figure 7

The list of genes and coefficients used for predicting age based on transcriptome data.

Gene	Coefficient
ABLIM1	-4.537363687
ACCN2	-4.021935755
ACVR2A	5.862922173
AK5	-10.29726151
ANXA1	6.307730249
ASNS	23.66865779
AUTS2	-13.3985662
C16orf45	4.553248948
CACHD1	-7.768187189
CDKN1C	-0.0105012
CENPV	-2.462314825
CMC1	8.866490009
CR2	-1.78645877
CRIP1	4.33558575
EFNA1	-5.741766145
EPHA2	2.917895917
FAIM3	1.019538625

Gene	Coefficient
FBLN2	-2.061520114
FLNB	-3.844863485
FZD1	1.375051746
GPC2	-7.431385678
GSC	4.904090057
GTSF1	12.58953522
HIST1H3D	-8.692565907
IGLL1	-4.235566899
KRT72	2.814127932
LRP11	5.664133584
MEOX1	13.8516364
MS4A3	-2.661573104
NEFH	-6.728594491
NMT2	-15.38338708
NOSIP	-13.61680769
NT5E	-1.994658678
PHLDA3	18.71229769

Gene	Coefficient
PHYHD1	-5.052538719
PLXDC1	-5.337661458
POMC	-6.100365433
PRSS35	-6.498528559
PTGER2	-8.407414661
PYROXD1	13.70056157
RGMA	5.458322024
ROBO1	-7.342718162
SEC14L2	-2.887682148
SFRP5	4.430923586
SLC45A3	8.799140451
SORCS3	9.998064269
SPEG	0.574659287
SPINK2	0.302316458
SYT11	7.093819787
TMEM8B	-13.0069907
TMIGD2	6.006761191

Gene	Coefficient
TNFRSF17	-6.154501401

Gene	Coefficient
TXNDC5	-0.265342977

Gene	Coefficient
ZNF285	-4.729710661

Conclusions

In this study, we have shown that genome-wide methylation patterns represent a strong and reproducible biomarker of biological aging rate. These patterns enable a quantitative model of the aging methylome that demonstrates high accuracy and an ability to discriminate relevant factors in aging, including gender and genetic variants. Moreover, our ability to apply this model in multiple tissues suggests the possibility of a common molecular clock, regulated in part by changes in the methylome. It remains to be seen whether these changes occur on an intracellular level uniformly across a population of cells, or reflect consistent changes in tissue composition over time.

The ability to predict age from whole blood may permit a wider analysis in longitudinal studies such as the Framingham Study, the Women's Health Initiative, blood samples collected on neonatal Guthrie cards, and other longitudinal studies with rich annotation of biometric and disease traits. Aging trends could emerge from such studies with many potential practical implications, from health assessment and prevention of disease to forensic analysis. Similar to the effect of gender in this study, the identification of additional biometric or environmental factors that influence AMAR, such as smoking, alcohol consumption, or diet, will permit quantitative assessments of their impacts on health and longevity. A useful example would be to periodically assess the rate of aging of an individual using AMAR and determine whether diet or environmental factors can accelerate or retard the aging process and diseases such as age related macular degeneration. As models of human aging improve, it is conceivable that biological age, as measured from molecular profiles, might one day supersede chronological age in the clinical evaluation and treatment of patients.

References for Example 1

Alisch, R.S., Barwick, B.G., Chopra, P., Myrick, L.K., Satten, G.A., Conneely, K.N., and Warren, S.T. (2012). Age-associated DNA methylation in pediatric populations. Genome Res. 22, 623-632.

Atzmon, G., Rincon, M., Schechter, C.B., Shuldiner, A.R., Lipton, R.B., Bergman, A., and Barzilai, N. (2006). Lipoprotein genotype and conserved pathway for exceptional longevity in humans. PLoS Biol. 4, e113.

Austad, S.N. (2006). Why women live longer than men: sex differences in longevity. Gend. Med. 3, 79-92.

Barres, R., and Zierath, J.R. (2011). DNA methylation in metabolic disorders. Am. J. Clin. Nutr. 93, 897S-900.

Bell, J.T., Pai, A.A., Pickrell, J.K., Gaffney, D.J., Pique-Regi, R., Degner, J.F., Gilad, Y., and Pritchard, J.K. (2011). DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol. 12, R10. http://www.ncbi.nlm.nih.gov/pubmed/21251332.

Bell, J.T., Tsai, P.-C, Yang, T.-P., Pidsley, R., Nisbet, J., Glass, D., Mangino, M., Zhai, G., Zhang, F., Valdes, A., *et al;* MuTHER Consortium. (2012). Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. PLoS Genet. 8, e1002629.

Bellacosa, A., Cicchillitti, L., Schepis, F., Riccio, A., Yeung, A.T., Matsumoto, Y., Golemis, E.A., Genuardi, M., and Neri, G. (1999). MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. Proc. Natl. Acad. Sci. USA 96, 3969-3974.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. B 57, 289-300.

Bertoni, C , Rustagi, A., and Rando, T.A. (2009). Enhanced gene repair mediated by methyl-CpG-modified single-stranded oligonucleotides. Nucleic Acids Res. 37, 7468-7482.

Bibikova, M., Barnes, B., Tsan, C , Ho, V., Klotzle, B., Le, J.M., Delano, D., Zhang, L., Schroth, G.P., Gunderson, K.L., *et al.* (2011). High density DNA methylation array with single CpG site resolution. Genomics 98, 288-295.

Blair, S.N., Kohl, H.W., 3rd, Paffenbarger, R.S.J., Jr., Clark, D.G., Cooper, K.H., and Gibbons, L.W. (1989). Physical fitness and all-cause mortality. A prospective study of healthy men and women. JAMA 262, 2395-2401.

Bocklandt, S., Lin, W., Sehl, M.E., Sanchez, F.J., Sinsheimer, J.S., Horvath, S., and Vilain, E. (2011). Epigenetic predictor of age. PLoS ONE 6, e14821.

Boks, M.P., Derks, E.M., Weisenberger, D.J., Strengman, E., Janson, E., Sommer, I.E., Kahn, R.S., and Ophoff, R.A. (2009). The relationship of DNA methylation with age, gender and genotype in twins and healthy controls. PLoS ONE 4, e6767.

Bollati, V., Schwartz, J., Wright, R., Litonjua, A., Tarantini, L., Suh, H., Sparrow, D., Vokonas, P., and Baccarelli, A. (2009). Decline in genomic DNA methylation through aging in a cohort of elderly subjects. Mech. Ageing Dev. 130, 234-239.

Breusch, T.S., and Pagan, A.R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. Econometrica 47, 1287.

Christensen, B.C., Houseman, E.A., Marsit, C.J., Zheng, S., Wrensch, M.R., Wiemels, J.L., Nelson, H.H., Karagas, M.R., Padbury, J.F., Bueno, R., *et al.* (2009). Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet. 5, e1000602.

Collins, F.S., and Barker, A.D. (2007). Mapping the cancer genome. Pinpointing the genes involved in cancer will help chart a new course across the complex landscape of human malignancies. Sci. Am. 296, 50-57.

de Magalhaes, J.P., Curado, J., and Church, G.M. (2009). Meta-analysis of age-related gene expression profiles identifies common signatures of aging. Bioinformatics 25, 875-881.

Emilsson, V., Thorleifsson, G., Zhang, B., Leonardson, A.S., Zink, F., Zhu, J., Carlson, S., Helgason, A., Walters, G.B., Gunnarsdottir, S., *et al.* (2008). Genetics of gene expression and its effect on disease. Nature 452, 423-428.

Epel, E.S., Blackburn, E.H., Lin, J., Dhabhar, F.S., Adler, N.E., Morrow, J.D., and Cawthon, R.M. (2004). Accelerated telomere shortening in response to life stress. Proc. Natl. Acad. Sci. USA 101, 17312-17315.

Esteller, M. (2008). Epigenetics in cancer. N. Engl. J. Med. 358, 1148-1159.

Feinberg, A.P., Irizarry, R.A., Fradin, D., Aryee, M.J., Murakami, P., Aspelund, T., Eiriksdottir, G., Harris, T.B., Launer, L., Gudnason, V., and Fallin, M.D. (2010). Personalized epigenomic signatures that are stable over time and covary with body mass index. Sci. Transl. Med. 2, 49ra67.

Fraga, M.F., and Esteller, M. (2007). Epigenetics and aging: the targets and the marks. Trends Genet. 23, 413-418.

Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudosa, J.C., Urioste, M., Benitez, J., *et al.* (2005). Epigenetic differences arise during the lifetime of monozygotic twins. Proc. Natl. Acad. Sci. USA 102, 10604-10609.

Fraser, H.B., Khaitovich, P., Plotkin, J.B., Paabo, S., and Eisen, M.B. (2005). Aging and gene expression in the primate brain. PLoS Biol. 3, e274.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. J. Stat. Softw. 33, 1-22.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., *et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 5, R80.

Harley, C.B., Futcher, A.B., and Greider, C.W. (1990). Telomeres shorten during ageing of human fibroblasts. Nature 345, 458-460.

Hartwig, A., and Schlepegrell, R. (1995). Induction of oxidative DNA damage by ferric iron in mammalian cells. Carcinogenesis 16, 3009-3013.

Heyn, H., Li, N., Ferreira, H.J., Moran, S., Pisano, D.G., Gomez, A., Diez, J., Sanchez-Mut, J.V., Setien, F., Carmona, F.X, *et al.* (2012). Distinct DNA methylomes of newborns and centenarians. Proc. Natl. Acad. Sci. USA 109, 10522-10527.

Jones, P.A., and Laird, P.W. (1999). Cancer epigenetics comes of age. Nat. Genet. 21, 163-167.

Karthikeyan, G., Lewis, L.K., and Resnick, M.A. (2002). The mitochondrial protein frataxin prevents nuclear damage. Hum. Mol. Genet. 11, 1351-1362.

Kim, E.J., Kim, M.-K., Jin, X.-J., Oh, J.-H., Kim, J.E., and Chung, J.H. (2010). Skin aging and photoaging alter fatty acids composition, including 11,14,17-eicosatrienoic acid, in the epidermis of human skin. J. Korean Med. Sci. 25, 980-983.

Lane, M.A., Baer, D.J., Rumpler, W.V., Weindruch, R., Ingram, D.K., Tilmont, E.M., Cutler, R.G., and Roth, G.S. (1996). Calorie restriction lowers body temperature in rhesus monkeys, consistent with a postulated anti-aging mechanism in rodents. Proc. Natl. Acad. Sci. USA 93, 4159-4164.

Li, R., Li, Y., Kristiansen, K., and Wang, J. (2008). SOAP: short oligonucleotide alignment program. Bioinformatics 24, 713-714.

Moniz, L., Dutt, P., Haider, N., and Stambolic, V. (201 1). Nek family of kinases in cell cycle, checkpoint control and cancer. Cell Div. 6, 18.

Murgatroyd, C , Patchev, A.V., Wu, Y., Micale, V., Bockmiihl, Y., Fischer, D., Holsboer, F., Wotjak, C.T., Almeida, O.F.X., and Spengler, D. (2009). Dynamic DNA methylation programs persistent adverse effects of early-life stress. Nat. Neurosci. 12, 1559-1566.

Ohgami, R.S., Campagna, D.R., McDonald, A., and Fleming, M.D. (2006). The Steap proteins are metalloreductases. Blood 108, 1388-1394.

Rakyan, V.K., Down, T.A., Maslau, S., Andrew, T., Yang, T.-P., Beyan, H., Whittaker, P., McCann, O.T., Finer, S., Valdes, A.M., *et al.* (2010). Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. Genome Res. 20, 434-439.

Saito, T., Iwata, N., Tsubuki, S., Takaki, Y., Takano, J., Huang, S.-M., Suemoto, T., Higuchi, M., and Saido, T.C. (2005). Somatostatin regulates brain amyloid beta peptide Abeta42 through modulation of proteolytic degradation. Nat. Med. 11, 434-439.

Shannon, C.E., and Weaver, W. (1963). The Mathematical Theory of Communication (Champaign, IL: University of Illinois Press).

Small, K.S., Hedman, A.K., Grundberg, E., Nica, A.C., Thorleifsson, G., Kong, A., Thorsteindottir, U., Shin, S.-Y., Richards, H.B., Soranzo, N., *et al.;* GIANT Consortium; MAGIC Investigators; DIAGRAM Consortium; MuTHER Consortium. (201 1). Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. Nat. Genet. 43, 561-564.

Suh, Y., Atzmon, G., Cho, M.-O., Hwang, D., Liu, B., Leahy, D.J., Barzilai, N., and Cohen, P. (2008). Functionally significant insulin-like growth factor I receptor mutations in centenarians. Proc. Natl. Acad. Sci. USA 105, 3438-3442.

Sun, Z., Asmann, Y.W., Kalari, K.R., Bot, B., Eckel-Passow, J.E., Baker, T.R., Carr, J.M., Khrebtukova, I., Luo, S., Zhang, L., *et al.* (201 1). Integrated analysis of gene expression, CpG island methylation, and gene copy number in breast cancer cells by deep sequencing. PLoS ONE 6, el7490.

Tatar, M., Bartke, A., and Antebi, A. (2003). The endocrine regulation of aging by insulin-like signals. Science 299, 1346-1351.

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., and Altman, R.B. (2001). Missing value estimation methods for DNA microarrays. Bioinformatics 17, 520-525.

Valdes, A.M., Andrew, T., Gardner, J.P., Kimura, M., Oelsner, E., Cherkas, L.F., Aviv, A., and Spector, T.D. (2005). Obesity, cigarette smoking, and telomere length in women. Lancet 366, 662-664.

Vijg, J., and Campisi, J. (2008). Puzzles, promises and a cure for ageing. Nature 454, 1065-1071.

Wheeler, H.E., Metter, E.J., Tanaka, T., Absher, D., Higgins, J., Zahn, J.M., Wilhelmy, J., Davis, R.W., Singleton, A., Myers, R.M., *et al.* (2009). Sequential use of transcriptional profiling, expression quantitative trait mapping, and gene association implicates MMP20 in human kidney aging. PLoS Genet. 5, el000685.

Willcox, B.J., Donlon, T.A., He, Q., Chen, R., Grove, J.S., Yano, K., Masaki, K.H., Willcox, D.C., Rodriguez, B., and Curb, J.D. (2008). FOX03A genotype is strongly associated with human longevity. Proc. Natl. Acad. Sci. USA 105, 13987-13992.

Xiong, M., Dong, H., Siu, H., Peng, G., Wang, Y., and Jin, L. (2010). Genome-Wide Association Studies of Copy Number Variation in Glioblastoma. Proceedings of the 4th International Conference on Bioinformatics and Biomedical Engineering (iCBBE), 1-4.

Yacubova, E., and Komuro, H. (2002). Stage-specific control of neuronal migration by somatostatin. Nature 415, 77-8 1.

Zahn, J.M., Poosala, S., Owen, A.B., Ingram, D.K., Lustig, A., Carter, A., Weeraratna, A.T., Taub, D.D., Gorospe, M., Mazan-Mamczarz, K., *et al.* (2007). AGEMAP: a gene expression database for aging in mice. PLoS Genet. 3, e201.

Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. J. R. Stat. Soc. Series B Stat. Methodol. 67, 301-320.

EXAMPLE 2

Building a Methylation Model of Aging

We measured the methylation states of 485,577 CpG markers in genomic DNA collected from whole blood samples of 302 Caucasian individuals. Of these, 80 individuals had been diagnosed with type-2 diabetes, 22 of which were also characterized with diabetic nephropathy. For further study and validation, the methylation states of a second cohort were obtained, consisting of 188 Hispanic individuals. Everyone in the second cohort was diagnosed with type-2 diabetes, and 89

individuals also had diabetic nephropathy. Careful filtering and normalization was performed to remove the effects of gender, batch, and other unknown covariates.

In general, we assume biological activity will track with chronological age, allowing us to infer a biological model from chronological age. We hypothesize that changes in molecular activity from a common baseline will reflect a deceleration or acceleration of the aging process, to which disease, environment, and genetics might contribute.

Formally, we define biological age (bioage) as:

$$Bioage = f(M_{sub}) = Age + \sum_{j=1}^{f} \alpha_j c_j + \epsilon$$

where M_{sub} is a small subset of the methylation data, $\propto_j$ is a numerical coefficient, c_j is the j-th trait, and e is model error. A critical point here is that by selecting model probes that are coordinately linked to chronological aging, alterations to the methyl states corresponding to these same probes are likely to reflect either attenuation or amplification of the aging process.

We considered that the rate of biological aging is not constant, and that during different milestones of human development and senescence large shifts in biological aging will occur. We tested this hypothesis by first using a univariate association test, to identify the top age-associated methylation markers in the primary cohort (FDR < 0.05). We then measured the relative coherence of these markers between young and old individuals using an entropy metric (Figure **1).** We found that the associated markers were much more coherent in the young individuals than in the old individuals ($p < 0.05$). This suggests that methylation aging patterns are similar for young individuals, but diverge over time.

What is claimed is:

1. A method for predicting age of a subject based on the epigenome of the subject comprising:

(a) obtaining a biological sample of the subject;

(b) determining the methylation status of a set of age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4 and/or S5; and

(c) comparing the methylation status of a set of age-associated epigenetic marker(s) of the subject with the methylation status of the same markers from an age correlated reference population so as to obtain a value or a range of values for the predicted age of the subject,

thereby predicting the age of a subject based on the epigenome of the subject.

2. The method of claim 1, wherein in (b), determining the methylation status comprises the following steps:

(a) isolating genomic DNA or nuclear DNA from the sample;

(b) reacting the isolated genomic DNA or nuclear DNA with one or more probes which differentially reacts with unmodified cytosine so that the cytosine is converted to uracil;

(c) analyzing the methylation status at the C position of a CpG dinucleotide in the isolated genomic DNA or nuclear DNA of the sample by detecting the presence of a cytosine or uracil, the presence of cytosine or uracil indicating presence of a 5-methylcytosine or unmodified cytosine, respectively, in the original CpG dinucleotide; and

(d) determining proportion of 5-methylcytosine or unmodified cytosine initially present at each age-associated epigenetic marker from (c); or alternatively, determining the ratio of 5-methyl-cytosine to unmodified cytosine or the ratio of unmodified cytosine to 5-methyl-cytosine initially present at each age-associated epigenetic marker from (c).

3. The method of claim 1, wherein in (b), determining the methylation status comprises the following steps:

(a) isolating genomic DNA or nuclear DNA from the sample;

(b) incubating the isolated genomic DNA or nuclear DNA with one or more restriction enzyme which (i) recognizes a specific DNA sequence, (ii) is affected by a CpG dinucleotide, within or adjacent to the restriction enzyme recognition or cleavage site, and (iii) differentially cleaves the DNA based on the presence or absence of a methyl group at C-5 position of cytosine of the CpG dinucleotide;

(c) analyzing the methylation status at the C position of a CpG dinucleotide in the isolated genomic DNA or nuclear DNA of the sample by its resistance to cleavage at a potential cleavage site by the restriction enzyme indicating presence of 5-methylcytosine at the original CpG dinucleotide within or adjacent to the restriction enzyme recognition or cleavage site, while sensitivity to cleavage by the restriction enzyme indicating presence of unmodified cytosine; and

(d) determining proportion of 5-methylcytosine or unmodified cytosine initially present at each age-associated epigenetic marker from (c); or alternatively, determining the ratio of 5-methylcytosine to unmodified cytosine or the ratio of unmodified cytosine to 5-methylcytosine initially present at each age-associated epigenetic marker from (c).

4. The method of claim 1, wherein in (b), the methylation status is determined from five or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5; ten or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5; fifteen or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5; twenty or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5; twenty-five or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5; thirty or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5; thirty-five or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5; forty or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5; forty-five or more age-associated epigenetic marker(s) selected from Figure 9,

Tables S3, S4 and/or S5; fifty or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5; fifty-five or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5; sixty or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5; sixty-five or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5; or seventy or more age-associated epigenetic marker(s) selected from Figure 9, Tables S3, S4 and/or S5.

5. The method of claim 1, wherein in (b), the methylation status is determined from five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; ten or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; fifteen or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; twenty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; twenty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; thirty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; thirty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; forty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; forty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; fifty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; fifty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; sixty or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; sixty-five or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3; or seventy or more age-associated epigenetic marker(s) in the epigenome of the subject from Figure 9 or Table S3.

6. The method of claim 1, wherein the age of a subject is a chronological age or a biological age.

7. The method of claim 1, wherein the epigenome is deoxyribonucleic acid (DNA) in which the DNA is subjected to epigenetic modification.

8. The method of claim 7, wherein epigenetic modification is the methylation of CpG residues.

9. The method of, claims 1 or 8, wherein the methylation is the covalent attachment of a methyl group at the carbon-5 position of cytosine.

10. The method of claim 1, wherein the biological sample is blood, lymphocyte, monocyte, neutrophil, basophil, eosinophil, myeloid lineage cell, lymphoid lineage cell, bone marrow, saliva, buccal swab, nasal swab, urine, fecal material, hair, breast tissue, ovarian tissue, uterine tissue, cervical tissue, prostate tissue, testicular tissue, brain tissue, neuronal cell, astrocyte, liver tissue, kidney, thyroid tissue, stomach tissue, intestine tissue, pancreatic tissue, vascular tissue, skin, lung tissue, bone tissue, cartilage, ligament, tendon, fat cells, muscle cells, neurons, astrocytes, cultured cells with different passage number, cancer/tumor cells, cancer/tumor tissue, normal cells, normal tissue, any tissue(s) or cell(s) with a nucleus containing genetic material, or genetic material in the form of DNA.

11. The method of claim 1, wherein the age-associated epigenetic marker(s) comprises a CpG residue wherein methylation at C-5 position of cytosine varies with the chronological age of a species associated with the subject.

12. The method of claim 11, wherein the species associated with the subject is *Homo sapiens.*

13. The method of claim 1, wherein the set of age-associated epigenetic marker(s) comprises individual CpG residues subject to age-dependent methylation at C-5 position of cytosine in the genome of a subject.

14. The method of claim 13, wherein the set of markers comprise about 70 distinct CpG residue-containing age-associated epigenetic markers.

15. The method of claim 13, wherein the set of markers comprises any one or more of markers as shown in Figure 9, Table S3, Table S4 or Table S5.

16. The method of claim 15, wherein the set of markers comprises five or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; ten or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; fifteen or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; twenty or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; twenty-five or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; thirty or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; thirty-five or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; forty or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; forty-five or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; fifty or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; fifty-five or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; sixty or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; sixty-five or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3; or seventy or more age-associated epigenetic marker(s) as shown in Figure 9 or Table S3.

17. The method of claim 15, wherein the set of markers comprises five or more age-associated epigenetic marker(s) as shown in Table S5; ten or more age-associated epigenetic marker(s) as shown in Table S5; fifteen or more age-associated epigenetic marker(s) as shown in Table S5; twenty or more age-associated epigenetic marker(s) as shown in Table S5; twenty-five or more age-associated epigenetic marker(s) as shown in Table S5; thirty or more age-associated epigenetic marker(s) as shown in Table S5; thirty-five or more age-associated epigenetic marker(s) as shown in Table S5; forty or more age-associated epigenetic marker(s) as shown in Table S5; forty-five or more age-

associated epigenetic marker(s) as shown in Table S5; or fifty or more age-associated epigenetic marker(s) as shown in Table S5.

18. The method of claim 1, wherein correlation between chronological age and predicted age is at least 90%.

19. The method of claim 13, wherein the set of markers having individual CpG residues subject to methylation of C-5 position of cytosine in the genome of a subject comprises any one or more of methylation marker cg05652533 of Table S4, cg27367526 of Table S4, cgl 8404041 of Table S4, cg23606718 of Figure 9, Tables S3 and S5, cgl 6867657 of Figure 9, Tables S3 and S5, cg04474832 on chromosome 3 at position 52008487, cg05442902 on chromosome 22 at position 21369010, cg06493994 on chromosome 6 at position 25652602, cg09809672 on chromosome 1 at position 236557682, cgl 9722847 on chromosome 12 at position 308491 14, cg22736354 on chromosome 6 at position 18122719, cg05652533 of Table S4, cg27367526 of Table S4, cgl 8404041 of Table S4, cg23606718 on chromosome 2 at position 131513927, and/or cgl6867657 of chromosome 6 at position 11044877.

20. The method of claim 13, wherein the set of markers having individual CpG residues subject to methylation at C-5 position of cytosine in the genome of a subject comprises methylation marker cg05652533 of Table S4, cg27367526 of Table S4, cgl8404041 of Table S4, cg23606718 of Figure 9, Tables S3 and S5, cgl 6867657 of Figure 9, Tables S3 and S5, cg04474832 on chromosome 3 at position 52008487, cg05442902 on chromosome 22 at position 21369010, cg06493994 on chromosome 6 at position 25652602, cg09809672 on chromosome 1 at position 236557682, cgl 9722847 on chromosome 12 at position 308491 14, cg22736354 on chromosome 6 at position 18122719, cg05652533 of Table S4, cg27367526 of Table S4, cgl 8404041 of Table S4, cg23606718 on chromosome 2 at position 131513927, and cgl 6867657 of chromosome 6 at position 11044877.

21. The method of claim 13, wherein the set of markers having individual CpG residues subject to methylation at C-5 position of cytosine in the genome of a subject comprises methylation marker cg04474832 on chromosome 3 at position 52008487, cg05442902 on chromosome 22 at position 21369010, cg06493994 on chromosome 6 at position 25652602, cg09809672 on chromosome 1 at position 236557682, cgl9722847 on chromosome 12 at position 308491 14, and cg22736354 on chromosome 6 at position 18122719.

22. The method of claim 13, wherein the set of markers having individual CpG residues subject to methylation at C-5 position of cytosine in the genome of a subject is any of methylation marker cg20822990 of Figure 9 or Table S3, cg04400972 of Figure 9 or Table S3, cgl6054275 of Figure 9 or Table S3, cg036071 17 of Figure 9 or Table S3, cg20052760 of Figure 9 or Table S3, cgl 6867657 of Figure 9 or Table S3, cg06493994 of Figure 9 or Table S3, cg066851 11 of Figure 9 or Table S3, cg004861 13 of Figure 9 or Table S3, cg20426994 of Figure 9 or Table S3, cgl4361627 of Figure 9 or Table S3, cg08097417 of Figure 9 or Table S3, cg07955995 of Figure 9 or Table S3, cg22285878 of Figure 9 or Table S3 and/or cg08540945 of Figure 9 or Table S3.

23. A method for diagnosing the presence of tumor in a subject comprising:

(a) obtaining biological sample suspected to contain tumor and a second biological sample of the same type but known not to contain tumor;

(b) predicting age of the each biological sample by the method of claim 1;

(c) comparing the ages predicted for the two samples, such that a biological sample with tumor will have an older predicted age than biological sample without tumor;

thereby diagnosing the presence of tumor in a subject.

24. The method of claim 23, wherein the biological sample with tumor is derived from blood, lymph node, liver, brain, esophagus, trachea, stomach, intestine, pancreas, throat, tongue, bone, ovary, uterus, cervix, peritoneum, prostate, testes, breast, kidney, lung, or skin.

25. The method of claim 23, wherein the set of age-associated epigenetic marker(s) in the epigenome of the subject comprises any of methylation marker cg23606718 of Figure 9, Tables S3 and S5 and/or cgl 6867657 of Figure 9, Tables S3 and S5.

26. The method of claim 23, wherein the biological sample with tumor is predicted to have an older predicted age of at least 30% more than the biological sample without tumor.

27. The method of claim 23, wherein the set of age-associated epigenetic marker(s) in the epigenome of the subject comprises any of methylation marker cg23606718 of Figure 9, Tables S3 and S5 and/or cgl 6867657 of Figure 9, Tables S3 and S5 and the biological sample with tumor is predicted to have an older predicted age of at least 30% more than the biological sample without tumor.

28. The method of claim 22, wherein a majority of the age-associated epigenetic markers in the epigenome of the subject predicts an older age for a biological sample with tumor than biological sample of the same type without tumor.

29. The method of claim 28, wherein a majority of the age-associated epigenetic markers in the epigenome of the subject predicting an older age for a biological sample with tumor than biological sample of the same type without tumor is more than 70% of total age-associated epigenetic markers.

30. The method of claim 2, wherein one or more probes which differentially reacts with unmodified cytosine and 5-methyl-modified cytosine is chosen from a set comprising sodium bisulfite, sodium metabisulfite, and/or bisulfite salts.

31. The method of claim 2, wherein the outcome of reacting the isolated genomic DNA or nuclear DNA with one or more probes is deamination of unmodified cytosine to uracil and unaltered 5-methylcytosine.

32. The method of claim 2, wherein analyzing the methylation status comprises (a) DNA amplification and nucleic acid sequence determination for the presence of either cytosine or thymine at the C position of CpG dinucleotide within the age-associated epigenetic marker; or, (b) DNA amplification followed by phage RNA polymerase transcription, RNase cleavage and matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) of RNase cleavage products.

33. The method of claim 32, wherein the nucleic acid sequence determination involves one or more of the following procedures: nucleic acid fragmentation, restriction enzyme digestion, nucleic acid hybridization, primer extension, pyrosequencing, single nucleotide extension, single nucleotide extension with biotin-labelled ddNTP, single nucleotide extension with 2,4-dinitrophenol (DNP)-labelled ddNTP, radioactive isotope labeling, non-radioactive label incorporation, fluorescent label incorporation, biotin incorporation, antigen-antibody complex formation, antibody detection, colorimetric detection, fluorescence detection, detection with fluorescent dye-labelled antibody, detection with labeled avidin or streptavidin, bead analysis or detection method, signal amplification, polymerase chain reaction, DNA amplification with thermostable DNA polymerase, phi-29 DNA polymerase DNA amplification, RNA production, in vitro transcription, phage RNA polymerase transcription, T7 RNA polymerase transcription, SP6 RNA polymerase transcription, T3 RNA polymerase transcription, RNAse digestion, RNase A digestion, DNA cloning, bacterial transformation, gel electrophoresis, mass spectroscopy, MALDI-TOF mass spectroscopy, microarray analysis, fluorescence scanner analysis, automated digital image capture, automated digital image analysis, ratiometric analysis, and Infinium® HumanMethylation450 BeadChip analysis.

34. The method of claim 2, wherein the ratio of 5-methylcytosine to unmodified cytosine initially present at each age-associated epigenetic marker is the ratio of cytosine to thymidine at pyrimidine position of the CpG dinucleotide after exposure to one or more probe and following analysis of products of nucleic acid amplification.

35. The method of claim 2, wherein the ratio of unmodified cytosine to 5-methylcytosine initially present at each age-associated epigenetic marker is the ratio of thymidine to cytosine at pyrimidine position of the CpG dinucleotide after exposure to one or more probe and following analysis of products of nucleic acid amplification.

36. The method of claim 3, wherein one or more restriction enzyme probe which recognizes a specific DNA sequence comprising at least one CpG dinucleotide and differentially cleaves the DNA based on the presence or absence of a methyl group at C-5 position of cytosine of the dinucleotide is/are selected from restriction enzyme selected from the group consisting of Aatll, Acc65I, Accl, Acil, Acll, Afel, Agel, Ahdl, Alel, Apal, ApaLI, Ascl, AsiSI, Aval, Avail, Bael, Banl, , BbrPI, BbvCI, BceAI, Bcgl, BcoDI, BfuAI, BfuCI, Bgll, BmgBI, BsaAI, BsaBI, BsaHI, Bsal, BseYI, BsiEI, BsiWI, BslI, BsmAI, BsmBI, BsmFI, BspDI, BsrBI, BsrFI, BssHII, BssKI, BstAPI, BstBI, BstUI, BstZ17I, Cac8I, Clal, Dpnl, Dralll, Drdl, Eael, Eagl, Earl, Ecil, Eco53kI, EcoRI, EcoRV, Faul, Fnu4HI, Fokl, Fsel, Fspl, Haell, Haelll, Hgal, Hhal, Hindi, Hinfl, HinPlI, Hpal, Hpall, Hpyl66II, Hpyl88III, Hpy99I, HpyAV, HpyCH4IV, HpyCH4V, Kasl, Mbol, Mlul, Mmel, MspAlI, Mwol, Nael, Narl, Neil, NgoMIV, Nhel, NlalV, Notl, Nrul, Nt. BbvCI, Nt. BsmAI, Nt. CviPII, PaeR71, Phol, Plel, PluTI, Pmel, Pmll, PshAI, PspOMI, PspXI, Pvul, Rsal, RsrII, SacII, Sail, Sau3AI, Sau96I, ScrFI, SfaNI, Sfil, Sfol, SgrAI, Smal, SnaBI, StyD4I, Tfil, Tlil, Tsel, TspMI, Xhol, Xmal, and Zral.

37. The method of claim 3, wherein the outcome of reacting the isolated genomic DNA or nuclear DNA with one or more restriction enzyme probe is the production of a double-stranded DNA break at a restriction enzyme cleavage site when cytosine at a CpG dinucleotide is not modified or no double-stranded DNA break at a restriction enzyme cleavage site when cytosine at a CpG dinucleotide is modified at its C-5 position with a methyl group.

38. The method of claim 3, wherein analyzing the methylation status in step (c) comprises DNA amplification and analysis of age-associated epigenetic marker (a) for specific DNA end(s) or fragment(s) due to cleavage by the restriction enzyme(s) and (b) for intact

restriction enzyme cleavage site associated at a particular age-associated epigenetic marker.

39. The method of claim 3, wherein the proportion of 5-methylcytosine initially present at each age-associated epigenetic marker is a fraction or percent of the age-associated epigenetic marker with an intact restriction enzyme cleavage site due to resistance to cleavage by the restriction enzyme.

40. The method of claim 3, wherein the proportion of unmodified cytosine initially present at each age-associated epigenetic marker is a fraction or percent of the age-associated epigenetic marker cleaved by the restriction enzyme.

41. The method of claim 3, wherein the ratio of 5-methylcytosine to unmodified cytosine initially present at each age-associated epigenetic marker is the ratio of number or concentration of intact restriction enzyme cleavage sites to the number or concentration of double-stranded DNA breaks produced by the restriction enzyme for the age-associated epigenetic marker.

42. The method of claim 3, wherein the ratio of unmodified cytosine to 5-methylcytosine initially present at each age-associated epigenetic marker is the ratio of number or concentration of double-stranded DNA breaks produced by the restriction enzyme to number or concentration of intact restriction enzyme cleavage sites resistant to cleavage by the restriction enzyme due to presence of 5-methylcytosine for the age-associated epigenetic marker.

43. The method of claim 1, wherein in (b), determining the methylation status comprises the following steps:

(a) isolating genomic DNA or nuclear DNA from the sample;

(b) probing the isolated genomic DNA or nuclear DNA with one or more probes which differentially reacts with unmodified and 5-methyl-modified cytosine;

(c) amplifying the DNA of step (b);

(d) digesting the amplified DNA of step (c) with one or more restriction enzyme that recognizes a restriction enzyme site that contains a CpG dinucleotide but fails to digest the restriction enzyme site mutated to TpG dinucleotide from a CpG dinucleotide;

(e) determining the proportion of 5-methylcytosine or unmodified cytosine initially present at each age-associated epigenetic marker based on the fraction or percentage of restriction enzyme sites sensitive or resistant to digestion in step (d), respectively; or,

alternatively, determining the ratio of 5-methylcytosine to unmodified cytosine initially present at each age-associated epigenetic marker based on the ratio of number or concentration of sensitive restriction enzyme sites to number or concentration of resistant restriction enzyme sites in step (d); or

alternatively, determining the ratio of unmodified cytosine to 5-methylcytosine initially present at each age-associated epigenetic marker based on the ratio of number or concentration of resistant restriction enzyme sites to number or concentration of sensitive restriction enzyme sites to digestion in step (d).

44. The method of claim 1, wherein in (b), determining the methylation status of the set of age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5 comprises the following steps:

(a) isolating genomic DNA or nuclear DNA;

(b) fragmenting the genomic DNA or nuclear DNA;

(c) exposing the fragmented DNA of step (b) to a 5-methylcytosine-binding protein;

(d) separating 5-methylcytosine-binding protein-bound DNA fragments from 5-methylcytosine-binding protein-free DNA fragments of step (c);

(e) determining for each age-associated epigenetic marker, the proportion of 5-methyl-cytosine-containing DNA fragments or unmodified cytosine-containing DNA fragments by determining the fraction or percent of 5-methylcytosine-binding protein bound or free DNA fragments, respectively, for each age-associated epigenetic marker; or

alternatively, determining for each age-associated epigenetic marker, the ratio of 5-methylcytosine-containing DNA fragments to unmodified cytosine-containing DNA fragments by determining the ratio of number or concentration of 5-methylcytosine-binding protein-bound DNA fragments to the number or concentration of 5-methylcytosine-binding protein-free DNA fragments; or

alternatively, determining for each age-associated epigenetic marker, the ratio of unmodified cytosine-containing DNA fragments to 5-methylcytosine-containing DNA fragments by determining the ratio of number or concentration of 5-methylcytosine-binding protein-free DNA fragments to the number or concentration of 5-methylcytosine-binding protein-bound DNA fragments for each age-associated epigenetic marker,

thereby determining the methylation status.

45. The method of claim 44, wherein the 5-methylcytosine-binding protein is an antibody for 5-methylcytosine, MeCP2, MBD2, MBD2/MBD3L1 complex, core MBD domain of MBD2, or poly-MBD protein, a naturally occurring 5-methylcytosine binding protein, genetically engineered 5-methylcytosine binding protein, or derivative or fragment thereof.

46. The method of claim 45, wherein separating 5-methylcytosine-binding protein-bound DNA fragments from 5-methylcytosine-binding protein-free DNA fragments includes immunoprecipitation, immunocapture, solid phase chromatography, liquid chromatography, and/or gel electrophoresis.

47. A method for determining apparent methylomic aging rate (AMAR) of a subject comprising:

(a) predicting age by the method of claim 1; and

(b) dividing the age predicted in (a) by the actual chronological age;

thereby determining the apparent methylomic aging rate of a subject.

48. The method of claim 1, wherein comparing the methylation status of the set of age-associated epigenetic marker(s) of the subject with the methylation status of the same markers from an age correlated reference population comprises a statistical method selected from the group consisting of multivariate regression method, linear regression analysis, tabular method or graphical method comprises Elastic Net, Lasso regression method, ridge regression method, least-squares fit, binomial test, Shapiro-Wilk test, Grubb's statistics, Benjamini-Hochberg FDR, variance analysis, entropy statistics, and/or Shannon entropy.

49. A forensic diagnosis of human actual age from a tissue from a human by predicting age of a subject based on the epigenome of the subject by the method of claim 1,

50. A method for health assessment of a subject by predicting age of a subject based on the epigenome of the subject by the method of claim 1.

51. A method for screening whether an agent of interest can retard or accelerate aging process comprising:

(a) obtaining a biological sample from a living organism, and optionally, culturing cells, tissue, or organ derived from a living organism;

(b) predicting age or AMAR of the organism, following the method of claim 1 or 46, using organism appropriate age-associated epigenetic marker(s) such that the age-associated epigenetic marker(s) provided in claim 1 for a human subject may need to be substituted with age-associated epigenetic marker(s) for the organism being examined;

(c) exposing the living organism or cultured living cells, tissue, or organ from the living organism of step (a) to an agent of interest in a single dose, multiple doses, or continuous doses;

(d) obtaining a biological sample from the living organism or the cultured living cells, tissue, or organ of step (c);

(e) predicting age or AMAR of the organism from the biological sample of step (d), following the method of claim 1 or 47, using organism appropriate age-associated

epigenetic marker(s) such that the age-associated epigenetic marker(s) provided in claim 1 for a human subject may need to be substituted with age-associated epigenetic marker(s) for the organism being examined;

(f) performing the same steps (a)-(e) on another individual from the same organism or a duplicate cultured living cells, tissue, or organ from the same individual or organism but not treated with any agent of interest or treated with a placebo at step (c); and

g) comparing, for biological sample of the same predicted age or AMAR in step (b), the predicted age or AMAR from step (e) with the predicted age or AMAR from step (f), such that a lower value or range of values for step (e) indicates that an agent of interest can retard an aging process whereas a higher value or range of values indicates an agent of interest can accelerate aging process,

thereby screening whether an agent of interest can retard or accelerate aging process.

52. The method of claim 51, wherein the agent of interest includes an anti-oxidant, reducing agent, DNA damaging agent, vitamin, dietary supplement, food, food additive, food coloring, salt, vegetable, vegetable extract, fruit, fruit extract, flower, flower extract, fragrance, seed, seed extract, herb, herb extract, plant extract, fiber, fat, fatty acid, oil, sugar, artificial sweetener, probiotics, alcohol, wine, fungus, mold, cream, lotion, powder, makeup, sun blocker, gas, pollutant, smoke, environmental pollutant, paint, solvent, organic solvent, plastic, plasticizers, bisphenol, phenolic compounds, tobacco, inhalant, drug, biologic, hormone, endocrine disruptor, environmental estrogen, hormone antagonist, hormone agonist, caffeine, phytoestrogen, metal, enzyme, chelator, yogurt, sulfur compound, physical barrier, electromagnetic barrier, and radiation barrier..

53. The method of claim 51, wherein the organism is yeast, fruit fly, fish, worm, insect, zebra fish, nematode, plant, or mammal.

54. The method of claim 53, wherein the mammal is human, murine, simian, feline, canine, equine, bovine, porcine, ovine, caprine, rabbit, mammalian farm animal, mammalian sport animal, and mammalian pet.

55. A method for identifying type of tissue for a biological sample from a subject with a known chronological age comprising:

(a) ascertaining the chronological age of a subject;

(b) determining the predicted age of the subject from the biological sample by the method of claim 1;

(c) comparing to a reference standard relating the predicted age for various types of tissue to chronological age;

(d) determining which value from step (b) closely matches the predicted age in the reference standard for various types of tissue from step (c); and

(e) based on the closest match in step (d), assigning the type of tissue for the biological sample,

thereby identifying type of tissue for a biological sample from a subject with a known chronological age.

56. A method for identifying type of tissue for a biological sample from a subject with a known chronological age comprising:

(a) ascertaining the chronological age of a subject;

(b) determining the AMAR of the subject from the biological sample by dividing the predicted age of a subject from the chronological age of the subject;

(c) comparing to a reference standard relating the AMAR to chronological age for various types of tissue;

(d) determining which value from step (b) closely matches the AMAR in the reference standard for various types of tissue from step (c); and

(e) based on the closest match in step (d), assigning the type of tissue for the biological sample,

thereby identifying type of tissue for a biological sample from a subject with a known chronological age.

57. The method of claim 55 or 56, wherein the set of age-associated epigenetic marker(s) in the epigenome of the subject comprises any one or more of methylation marker

cg23606718 of Figure 9, Tables S3 and S5 and/or cgl6867657 of Figure 9, Tables S3 and S5.

58. A method for predicting age of a subject based on age-associated epigenetic modification affecting gene expression comprising:

(a) obtaining a biological sample of the subject;

(b) determining the expression of one or more gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age;

(c) comparing the expression of one or more gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age with the expression of the same gene(s) from an age-correlated reference population; and

(d) obtaining a value or range of values for the predicted age of the subject; wherein comparing the expression of one or more gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age with the expression of the same gene(s) from an age-correlated reference population comprises any statistical method, multivariate regression method, linear regression analysis, tabular method, or graphical method used to predict the age of a subject based on expression of gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age;

thereby predicting age of a subject based on age-associated epigenetic modification affecting gene expression.

59. The method of claim 58, wherein one or more gene(s) associated with age-associated epigenetic marker(s) whose expression changes with age comprises one or more of the genes listed in Table S6 or Table S7.

60. The method of claim 58, wherein gene expression is transcription or translation.

61. The method of claim 60, wherein transcription results in the production of RNA transcripts and translation results in the production of proteins.

62. A method of screening a tissue sample from a subject in order to predict the age of the tissue sample based on the epigenome of the subject by the method of claim 1.

63. The method of claim 62, wherein the tissue sample is exposed to at least one test agent in a high-throughput screening assay.

64. The method of claim 62, wherein said process is used for any one of diagnosis and/or high-throughput screening.

65. The method of claim 48, wherein the statistical method comprises a multivariate regression algorithm or linear regression algorithm.

66. A method for predicting age of a tissue or organ of a subject based on the epigenome of the tissue or organ of the subject comprising:

(a) obtaining a biological sample of a tissue or organ from the subject;

(b) determining the methylation status of the set of age-associated epigenetic marker(s) in the epigenome of the subject selected from Figure 9, Tables S3, S4, and/or S5; and

(c) comparing the methylation status of the set of age-associated epigenetic marker(s) of the subject with the methylation status of the same markers from an age-correlated reference population so as to obtain a value or a range of values for the predicted age of the tissue or organ,

thereby predicting the age of a tissue or organ of a subject based on the epigenome of the tissue or organ of the subject.

67. The method of claim 66, wherein the methylation status of the same markers from an age-correlated reference population is determined on a same type of tissue or organ.

68. The method of claim 66, wherein the methylation status of the same markers from an age-correlated reference population is determined on a different type of tissue or organ.

69. The method of claim 66, wherein the methylation status of the same markers from an age-correlated reference population is determined on blood or fractionated blood.

70. A method for determining differential aging rates of tissues or organs of a subject comprising:

(a) obtaining biological samples from different tissue(s) or organ(s) from the subject;

(b) predicting the age of the tissue or organ using the method of claim 66;

(c) comparing the predicted ages determined in step (b) where a difference in the predicted ages indicates a difference in the aging rate of the tissue(s) or organ(s) of the subject, thereby determining differential aging rates of tissues or organs of a subject.

71. The method of claim 70, wherein the predicted age is divided by the chronological age of the subject to obtain AMAR.

72. The method of claim 1 which is automated.

73. The method of claim 1, wherein the subject is human, mammal, animal, plant, or any multicellular organism.

74. A kit for determining age of a subject based on epigenetic modification of subject's genetic material comprising any age-associated epigenetic marker or markers as listed in Figure 9, Table S3, Table S4 or Table S5.

75. A kit for predicting age of a subject based on the epigenome of the subject utilizing the set of the age-associated epigenetic marker(s) provided in Figure 9, Table S3, S4 and/or S5.

76. The kit of claim 75, wherein the age-associated epigenetic marker(s) comprises a nucleic acid with a CpG dinucleotide.

77. The kit of claim 75, wherein the cytosine of the CpG dinucleotide is subject to age-dependent changes in methylation at the C-5 position.

78. The kit of claim 75, wherein the CpG dinucleotide is at the chromosomal position as indicated in Figure 9, Table S3, S4, and/or S5.

79. The kit of claim 75, wherein the age-associated epigenetic marker(s) is a human marker and selected from the group consisting of cg04474832 on chromosome 3 at position 52008487, cg05442902 on chromosome 22 at position 21369010, cg06493994 on chromosome 6 at position 25652602, cg09809672 on chromosome 1 at position 236557682, cg-19722847 on chromosome 12 at position 308491 14, cg22736354 on chromosome 6 at position 18122719, cg05652533 of Table S4, cg27367526 of Table S4, cgl8404041 of Table S4, cg23606718 on chromosome 2 at position 131513927, and eg16867657 of chromosome 6 at position 11044877.

80. The kit of claim 75, wherein the age-associated epigenetic marker(s) has the sequence as provided in Figure 9.

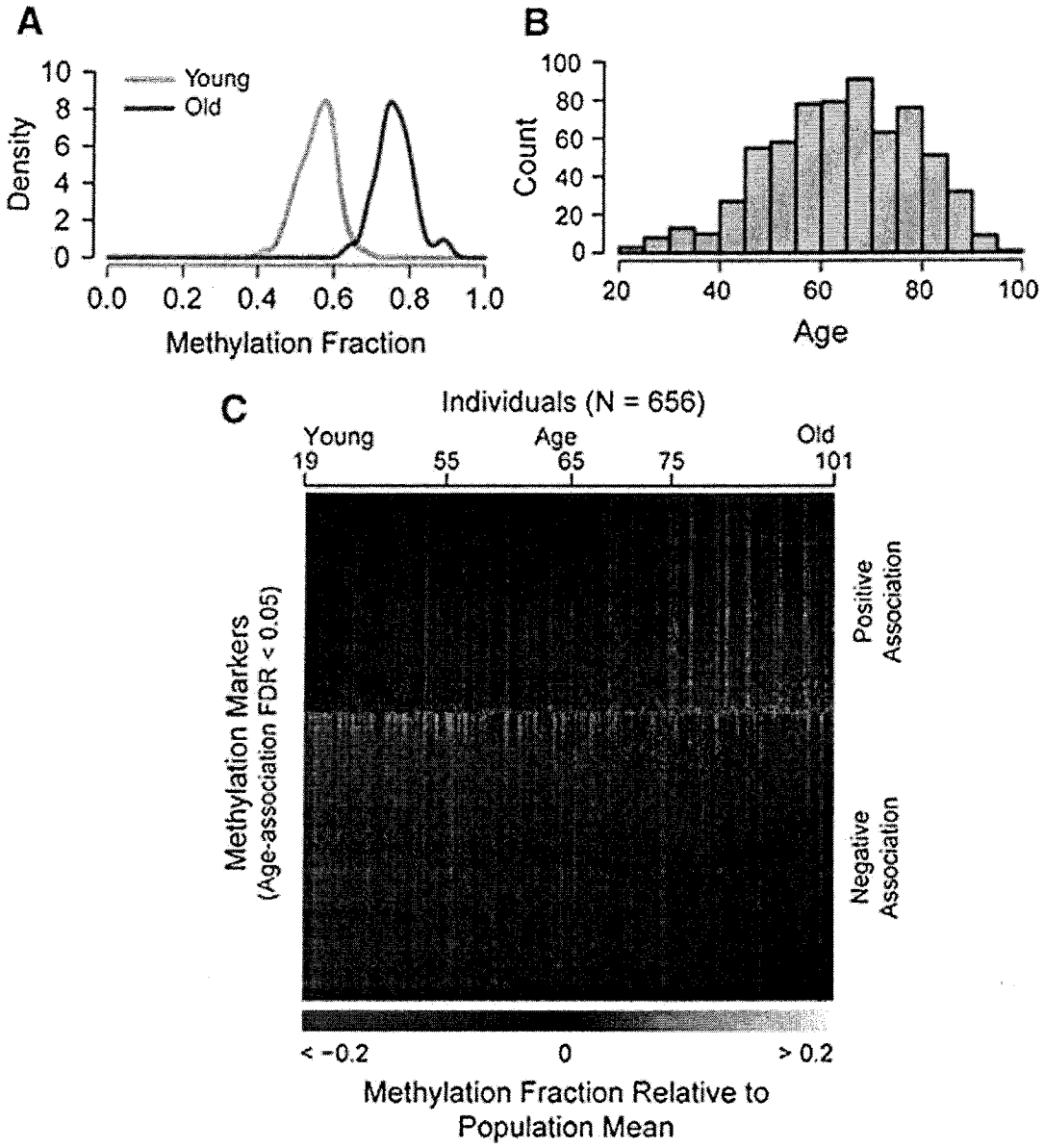

A
Density
Young
Old
Methylation Fraction
B
Count
Age
C
Individuals (N = 656)
Young
Age
Old
Methylation Markers
(Age-association FDR < 0.05)
Positive Association
Negative Association
< -0.2
0
> 0.2
Methylation Fraction Relative to Population Mean

FIGURE 1

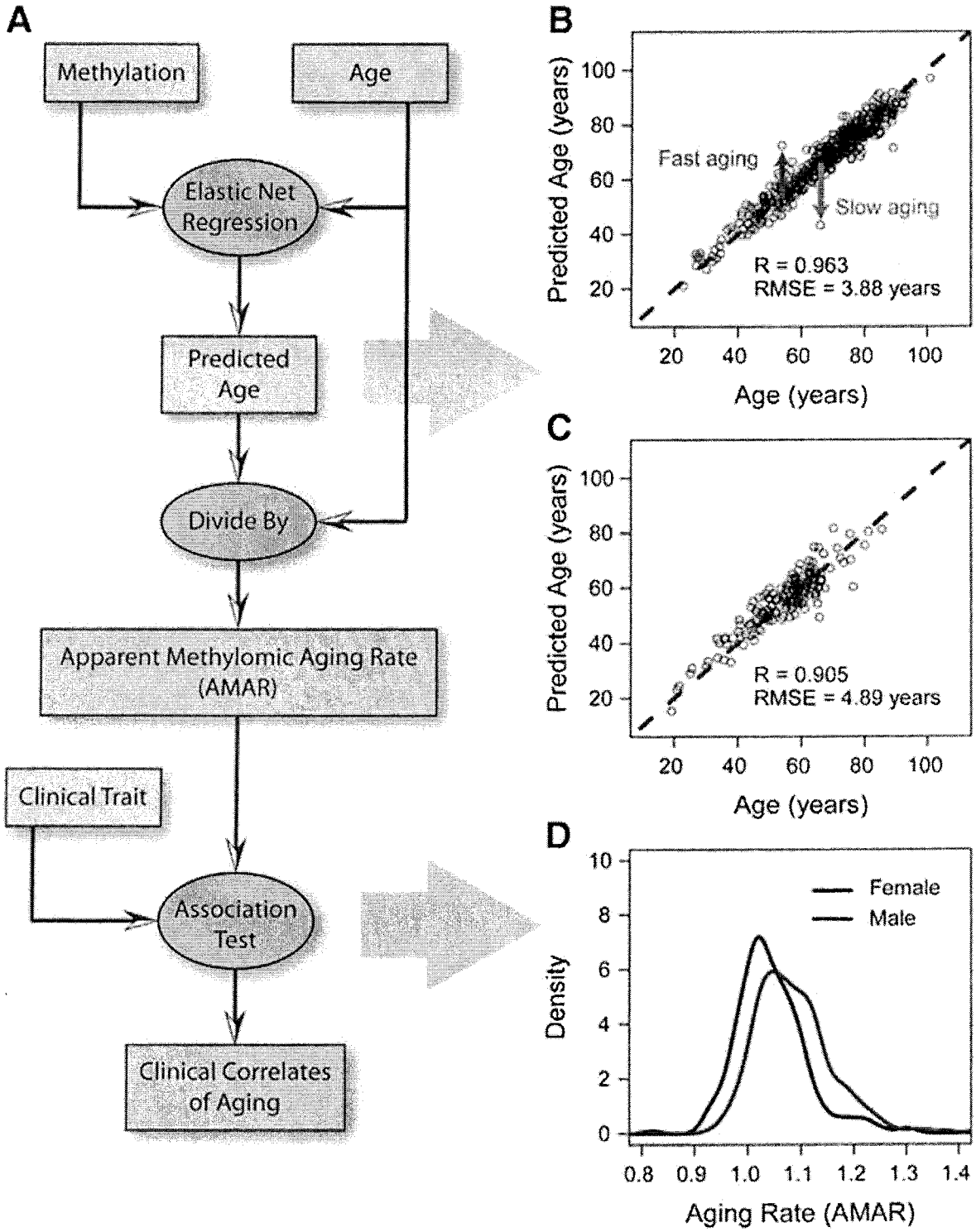

A
Methylation
Age
Elastic Net Regression
Predicted Age
Divide By
Apparent Methylomic Aging Rate (AMAR)
Clinical Trait
Association Test
Clinical Correlates of Aging
B
Predicted Age (years)
Fast aging
Slow aging
R = 0.963
RMSE = 3.88 years
Age (years)
C
Predicted Age (years)
R = 0.905
RMSE = 4.89 years
Age (years)
D
Density
Female
Male
Aging Rate (AMAR)

FIGURE 2

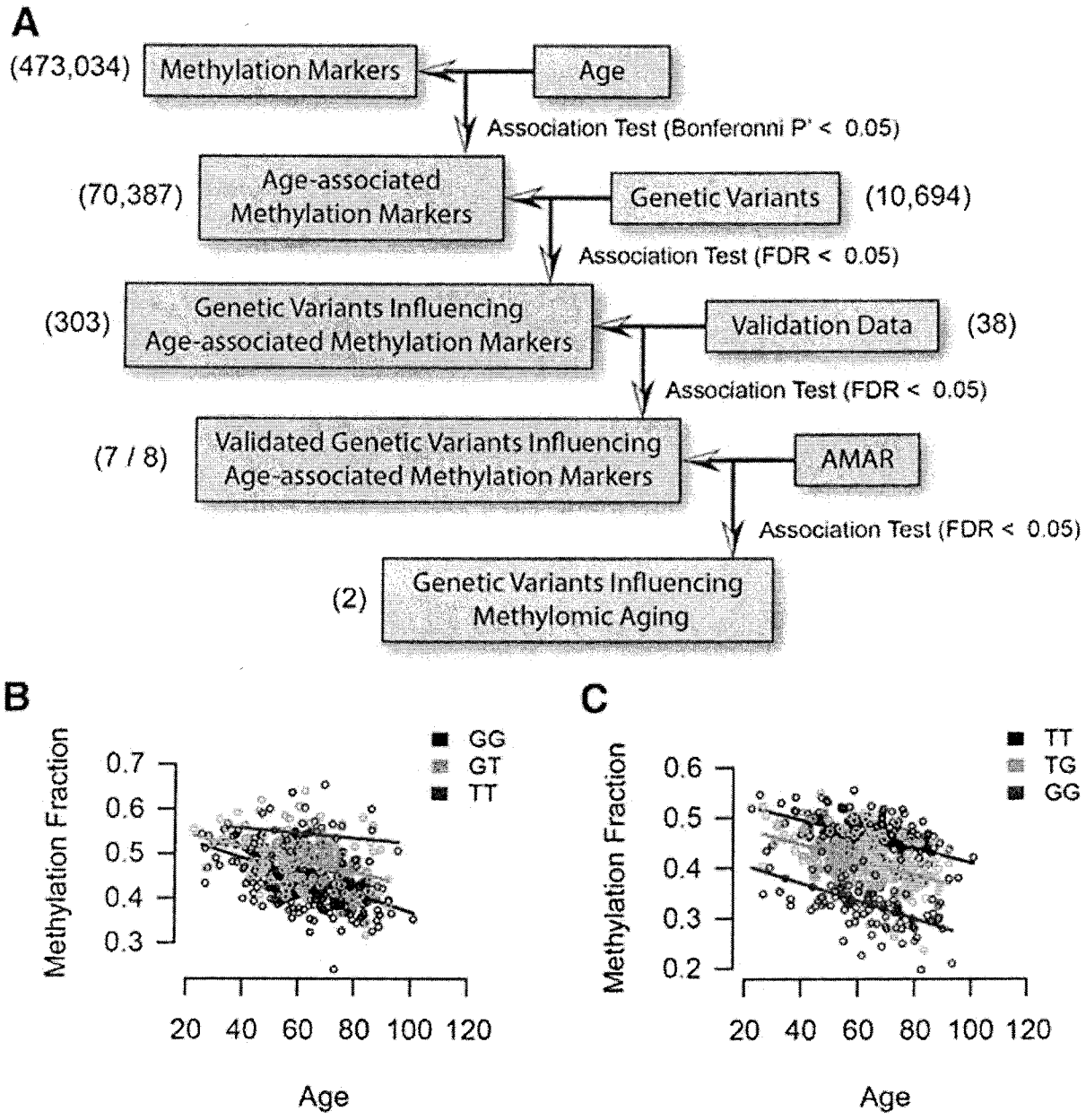

A
(473,034)
Methylation Markers
Age
Association Test (Bonferonni P' < 0.05)
(70,387)
Age-associated Methylation Markers
Genetic Variants
(10,694)
Association Test (FDR < 0.05)
(303)
Genetic Variants Influencing Age-associated Methylation Markers
Validation Data
(38)
Association Test (FDR < 0.05)
(7 / 8)
Validated Genetic Variants Influencing Age-associated Methylation Markers
AMAR
Association Test (FDR < 0.05)
(2)
Genetic Variants Influencing Methylomic Aging
B
Methylation Fraction
0.7
0.6
0.5
0.4
0.3
GG
GT
TT
20 40 60 80 100 120
Age
C
Methylation Fraction
0.6
0.5
0.4
0.3
0.2
TT
TG
GG
20 40 60 80 100 120
Age

FIGURE 3

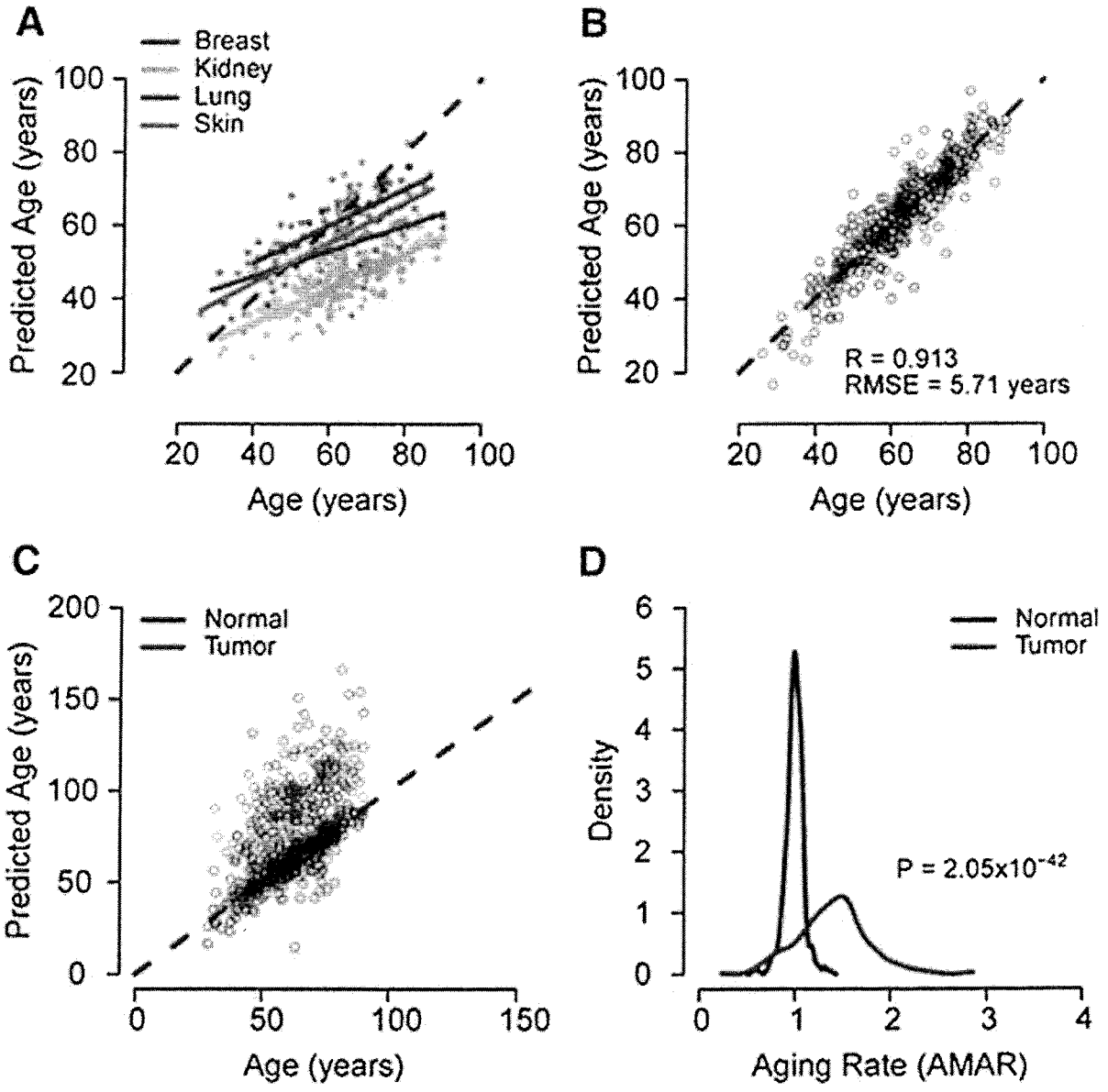

A
Breast
Kidney
Lung
Skin
Predicted Age (years)
Age (years)
B
R = 0.913
RMSE = 5.71 years
Predicted Age (years)
Age (years)
C
Normal
Tumor
Predicted Age (years)
Age (years)
D
Normal
Tumor
Density
P = 2.05x10^-42
Aging Rate (AMAR)

FIGURE 4

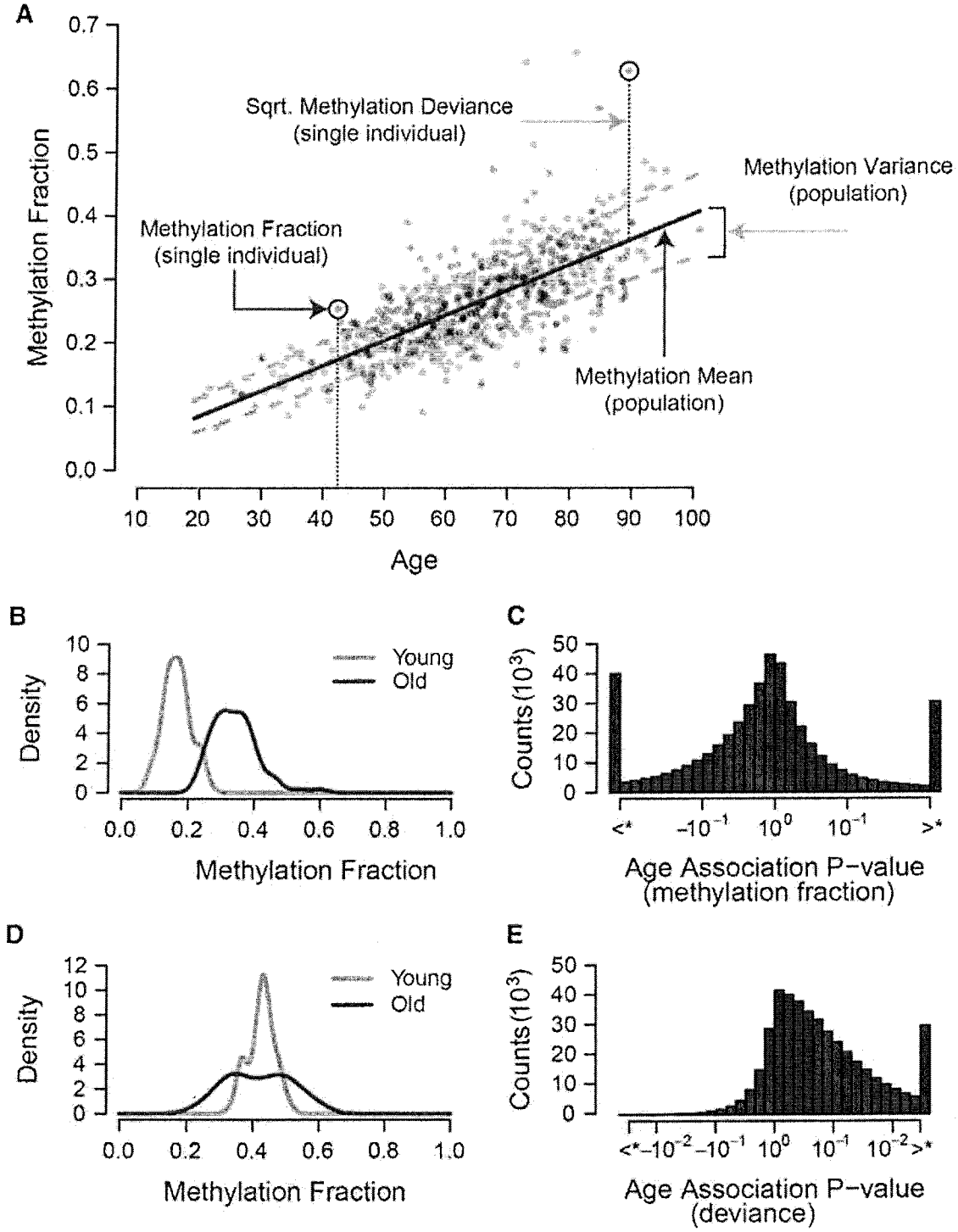

A
Methylation Fraction
Sqrt. Methylation Deviance
(single individual)
Methylation Variance
(population)
Methylation Fraction
(single individual)
Methylation Mean
(population)
Age
B
Density
Young
Old
Methylation Fraction
C
Counts (10^3)
Age Association P-value
(methylation fraction)
D
Density
Young
Old
Methylation Fraction
E
Counts (10^3)
Age Association P-value
(deviance)

FIGURE 5

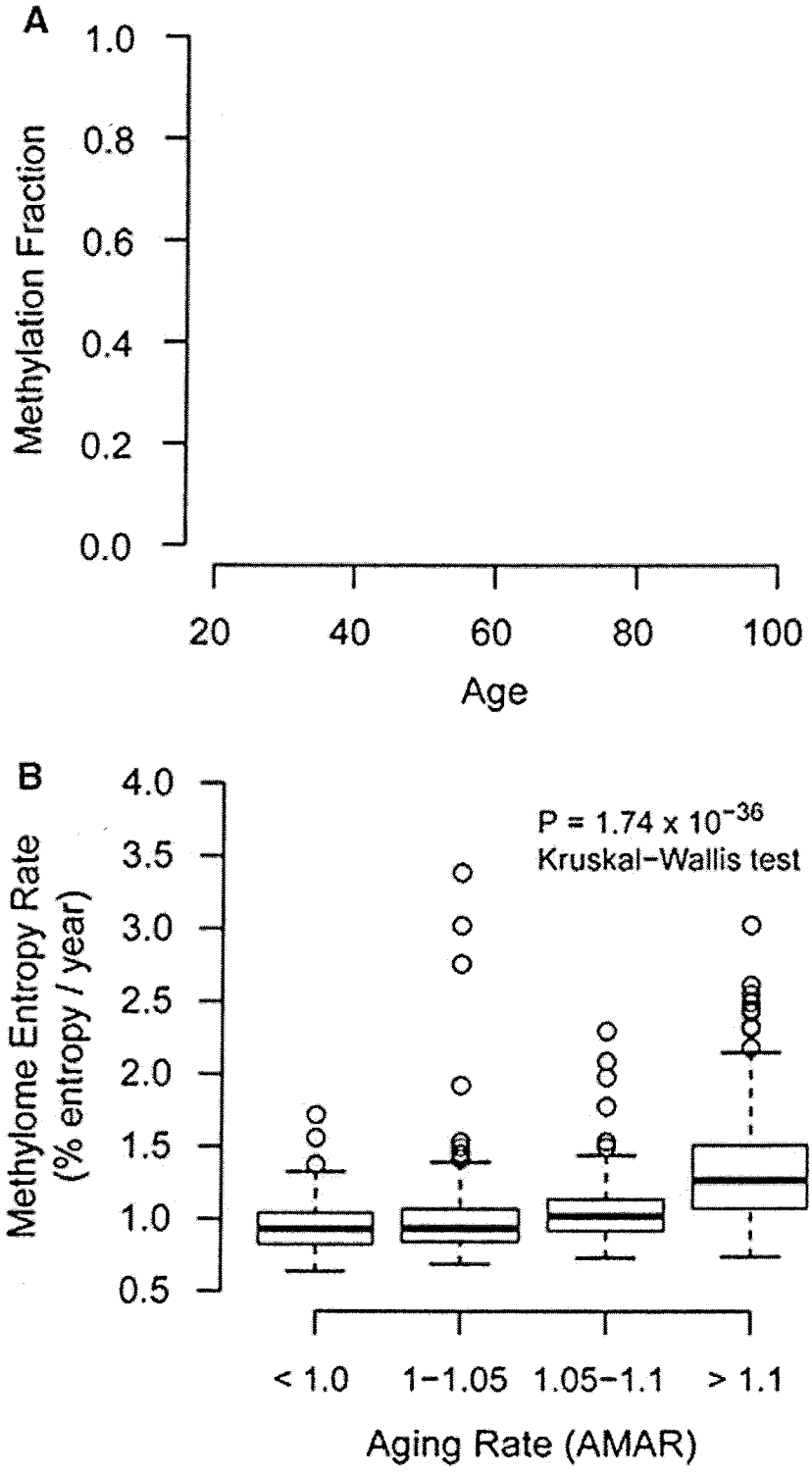

A
Methylation Fraction
1.0
0.8
0.6
0.4
0.2
0.0
20
40
60
80
100
Age
B
Methylome Entropy Rate
(% entropy / year)
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
P = 1.74 x 10^-36
Kruskal-Wallis test
< 1.0
1-1.05
1.05-1.1
> 1.1
Aging Rate (AMAR)

FIGURE 6

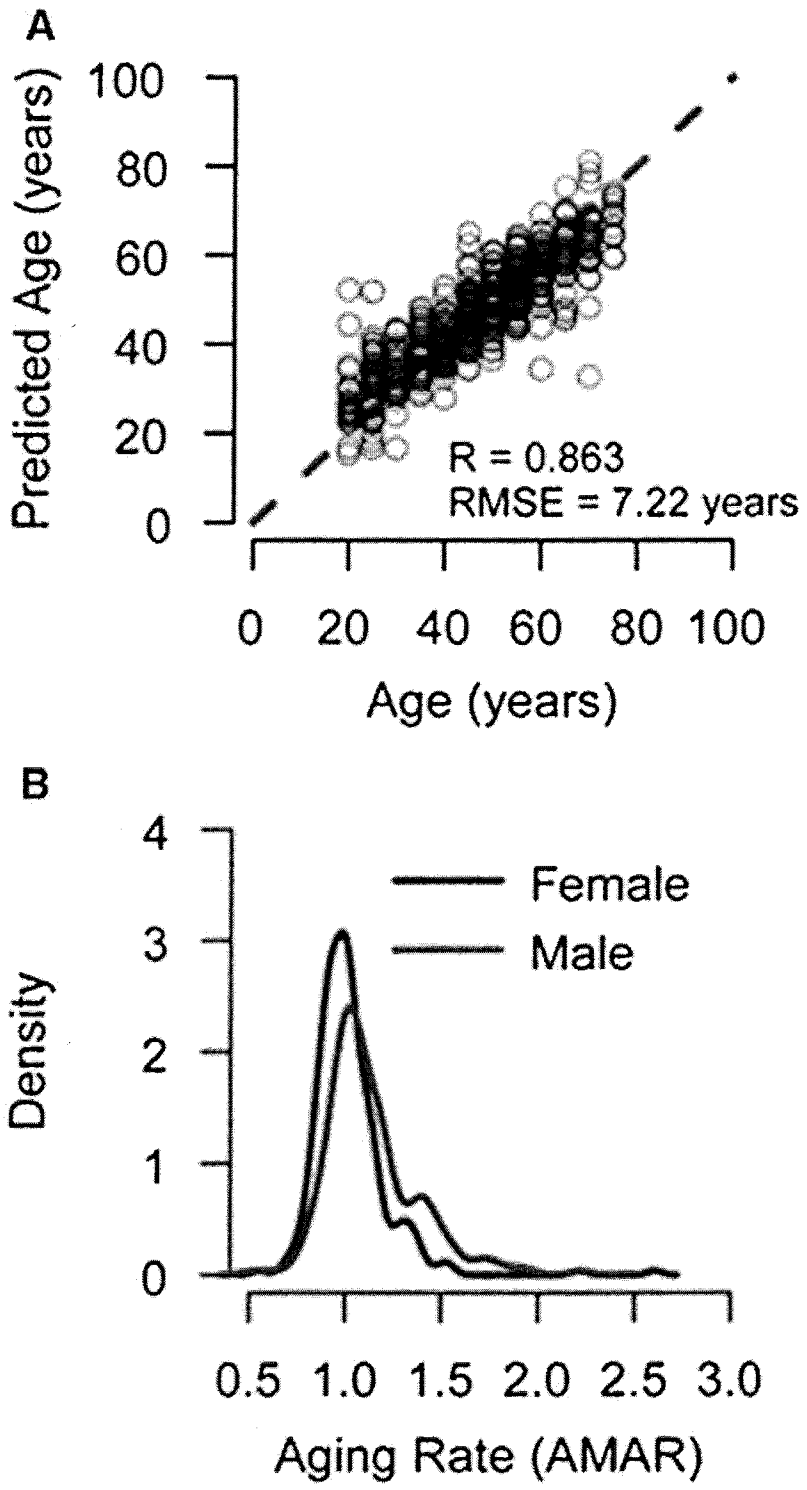
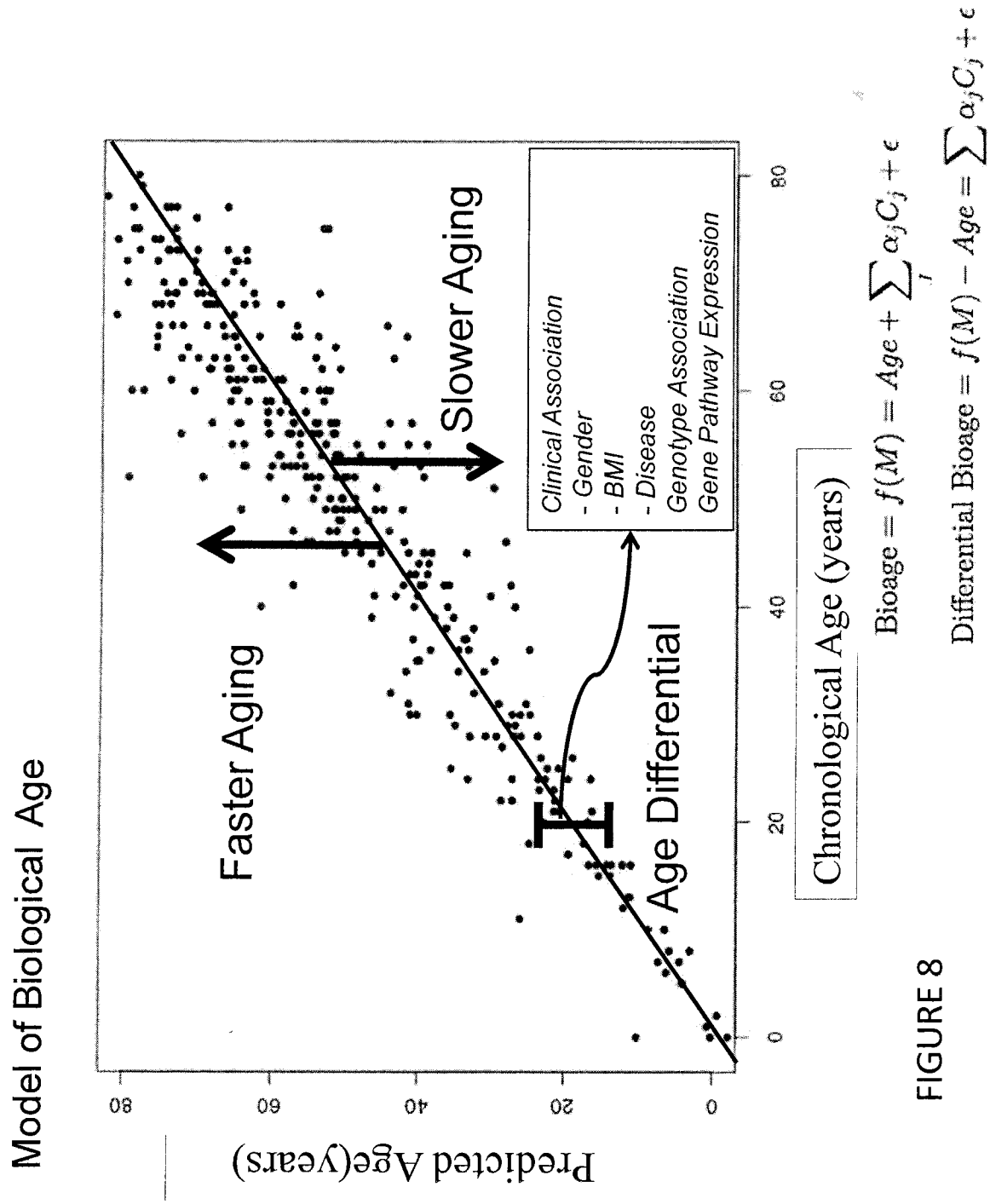

A
Predicted Age (years)
100
80
60
40
20
0
R = 0.863
RMSE = 7.22 years
0 20 40 60 80 100
Age (years)
B
Density
4
3
2
1
0
Female
Male
0.5 1.0 1.5 2.0 2.5 3.0
Aging Rate (AMAR)

FIGURE 7



$$\text{Bioage} = f(M) = Age + \sum_j \alpha_j C_j + \epsilon$$

$$\text{Differential Bioage} = f(M) - Age = \sum_j \alpha_j C_j + \epsilon$$

FIGURE 8

	ID	Name	AddressA_ID	AlleleA_ProbeSeq	AddressB_ID
19934	cg04400972	cg04400972	74745443	AAAATAAAAACATCTCAAACTCACATTACAAAAAC CAATTCAAAAAACCA	13601340
28973	cg09809672	cg09809672	12737501	TTTCATCTAAAAAATTTAACTCTAACCAAACAACCA ACRAACATCTTCTC	NA
30213	cg10501210	cg10501210	43774458	CAAAACTACAACACCTTACAACAAAACCAAAATTT AAATCTACACAAACA	66775341
39647	cg16054275	cg16054275	17771388	TCCAAATATAAACTTAAACCCAAACAAACCTCAAA TCAATAATAACRAAC	NA
47340	cg20822990	cg20822990	50699474	CACAAAAAACTACTACRAAAATCACACAAAAAAAT ACTTATCAACTAAAC	NA
49956	cg22512670	cg22512670	70754418	ATAATAACAACTACTAATACCCATCTTAACCTAAAA CACTAAAATCRCCC	NA
54990	cg25410668	cg25410668	26779422	ACACCRCRTAAAATTACTTATTCTTTTACATAAAAA ATCACATTCTCTTC	NA
61752	cg02085953	cg02085953	66719449	CTAAACTTTCAAACTAAAACCCCTAAAATAAACTCC TCACCTAAAAAAAC	NA
67888	cg06639320	cg06639320	13732383	CAAAACTCCTTTCTTCRTACCCTCCRAATCTTAAAA ACACAATAATTATC	NA
86462	cg22016779	cg22016779	10606390	RCAAACCCCTTAAAAAAAAACCAATACAACAAATT AATTAACCAAACCAC	NA
86644	cg22158769	cg22158769	17654495	ACACAAAAACTCTTTAAAAAAACAACTCAACAACT TAACCTTACCATACA	68643399
86991	cg22454769	cg22454769	66727363	CTTAAAAACACAATAATTATCAAAAACATCACCTC CAACATAAACTCTCA	23744359
88430	cg23606718	cg23606718	20628466	TCTCAAAACCTTAACAACTTACCACTAAAAACCCAC AATACAACAAAACA	58618322
88984	cg24079702	cg24079702	42601467	AAAACACAATAATTATCAAAAACATCACCTCCAAC ATAAACTCTCAAACA	50810372
93051	ch.2.30415474F	ch.2.30415474F	51667506	AAAACTAACTCTTCCTCTACCACATATAATCATCAA CAAATCCTATAAAT	NA

FIGURE 9A

	ID	Name	AddressA_ID	AlleleA_ProbeSeq	AddressB_ID
93891	cg00481951	cg00481951	14709443	CRTTTCAACACCTAAATCAACRCTTCCCAAAATCAA CACCAAAAATAAAC	NA
97218	cg03607117	cg03607117	70676435	TATCTACAACCCRAAATAAAACRCAAACTAAATCR AAACTAAACCAAAAC	NA
98073	cg04474832	cg04474832	31636354	TAACRTTTACTACTTATTACTACTCTACAAAAATTA ACAACATCTAACCC	NA
100937	cg07553761	cg07553761	43625429	ATATAACACCAAATAAAAAATAATTCCTCAAAAAC CATCAACCACCAACA	17688430
116505	cg25478614	cg25478614	60721480	AAAACCCAAAAAAACACCAAAACTCTTTAAAAAAA CTAAACATCCCTTAC	NA
120672	cg02650266	cg02650266	66683464	ACTATCCTCAAAAACCACCAAAATACTAAAAAAAA CAACAACAACAAACA	73675449
137234	cg25428494	cg25428494	68769372	AAATATATTACATTATAATAACATCTAAACTCTTAA TATAACTATCACCC	NA
146444	cg08234504	cg08234504	24749379	TAAAACCTAATTCACCTATTCACACACAAAAACAT AACTACCATAACAAC	NA
159652	cg23500537	cg23500537	48779348	TACAACCACACATCCAAAACTAACAAAACRAACAC TCTACCAAATCCTAC	NA
164052	cg00486113	cg00486113	37770376	AAACCTAAACCCACCCCRAAAAAAATACAAATAAT AAAATCCCCTCAACC	NA
172686	cg06493994	cg06493994	49740386	CAAAAAATACRATAAAAAAATCCTTCCCAAAATTA TCTAAATCCTTCCRC	NA
172936	cg06685111	cg06685111	59773369	TTACCAATCTAAATCCRTCCTAATACCTTACTATAC ATACAATTCTACTC	NA
181162	cg13001142	cg13001142	74697332	CACCRTCTCAAAAATTAACAAATTCAAATCTAACCT AATCAAATTTTATC	NA
186168	cg16867657	cg16867657	57761335	TCTAAACAACAAATAAATATTCCTAAAACTCCATA AACATTAAACCACCA	17685408
190156	cg20052760	cg20052760	32659462	ATTTTCACRCCATCATAACATTTTATTCCTCACCTAA CTAAAAACAACTC	NA

FIGURE 9B

	ID	Name	AddressA_ID	AlleleA_ProbeSeq	AddressB_ID
193374	cg22736354	cg22736354	26771442	ACTCRAATACAAAATATACTTTAAAAAATTTAACCA CCRACAACAACRAC	NA
204211	cg03473532	cg03473532	31780418	AATTCAAAATTTTCAAATAATAATAAACAACTACTA TCTCAAACACATAC	NA
209326	cg07927379	cg07927379	42797400	ACTAAAACACAAAAACATACAAACACCCAAACCTC ATATAATAAAAACCA	27785403
209350	cg07955995	cg07955995	49679509	ATCTCCACATTCTTTCTTCTACCAACAAACCAAATA ATAATAACAAAACA	57684340
209501	cg08097417	cg08097417	24706423	ATATTTAACAACCTCAAAAATTATCTTATCTCCRCR TTCTTTCTTCTACC	NA
209919	cg08540945	cg08540945	24789307	AATTCAAAACTTTAAACCCCCAACACTCTACAAAC AAAAAACAAAACACA	59640388
215998	cg14361627	cg14361627	42786323	RACCCCCCRACTAAATCATATTTAACAACCTCAAAA ATTATCTTATCTCC	NA
222324	cg20426994	cg20426994	55807502	ATAACACTTAACAACAAATATAACAAACCTCCTCC AAAACACCTAATCCA	12608480
224209	cg22285878	cg22285878	49781392	TCTTCTACCAACAAACCAAATAATAATAACAAAAC AAAACTCCCCAATCA	49684508
236164	cg07583137	cg07583137	23646337	ACAACCCCATTTAAAATTTTTAATACACTAAAAATC ATTCAAAAAACTTC	NA
242497	cg16419235	cg16419235	72778377	CAAAAAACTAATTTCTACAATCACTAAAACCTACA ACAACACTACTCCCA	26722332
279142	cg19935065	cg19935065	27739481	TATTATTTAAAACTTTTACTATCTCCAAAACACRTA AACACCTCACAACC	NA
280381	cg22796704	cg22796704	60625432	CRAAAAAAACCRCATTTCAAAACACTTACTCCTAA ATACCTAATATAATC	NA
286836	cg02046143	cg02046143	38797338	TTTCCACACATAACTCCTCACTACTTTATTACCTAA ACCAAAAAAAAACA	15671301
290175	cg04940570	cg04940570	46668492	ACACACCCTCAAACACCTTAAACAAAATACACTAA AAAACCAAAAATTCA	30738394

FIGURE 9C

	ID	Name	AddressA_ID	AlleleA_ProbeSeq	AddressB_ID
291937	cg06419846	cg06419846	40649319	TAATCACATATCCTACCACTACACTAAAAACTTCCA ACTAACAACAAACA	16728494
296834	cg11067179	cg11067179	41621317	CRAAAAATAATCCRCACCRCTATATAAACACAAAT AAATACCAAATTACC	NA
308206	cg22213242	cg22213242	25756359	ATCTACTAACACACACCAACAATCTAACACTCATCT ATATCCACACAACA	21781482
308942	cg23091758	cg23091758	55782376	AAAAACAACAACACTAATAAAAACTAACCCAACAA TCCAAAAATCCACCA	30707421
309640	cg23744638	cg23744638	35611335	CACAAAACTTTAAAAAACACTCTAACCCCTACTACT CACCCATACAAACA	14763344
314613	cg00748589	cg00748589	60757479	AACTCTACCTCAAAAAACTCAAAACCATCRTACTA AACCAACAAAAACTC	NA
315365	cg01528542	cg01528542	50719302	TCCCACTTTACAATCTTTACCAAATTTAATCATCAC TAACAAAAATTAAC	NA
330233	cg18473521	cg18473521	28617472	ATTACCCATTCTCRCTCRTAAATCCAATTCAATTAT ACTAACCCAAAATC	NA
331413	cg19722847	cg19722847	25621361	CAACAAATCTATCTTAAAACAAACCCAATTACRAT ACCCATACAATTCTC	NA
350626	ch.13.39564907R	ch.13.39564907R	61804317	GCTACTTCCCTAAATTAAACAAACCATTATTACCTC TCACCTAAACTACT	NA
352716	cg03032497	cg03032497	24810422	TATCTAACTCAACCCCTTTAAATATTCTTCCAAATA AAATTATTAAATTC	NA
367851	cg03399905	cg03399905	52669494	CCCATACTCRACCTTCTAAAAAATACCCACAAACAC TAACAATAATAAAC	NA
368628	cg04875128	cg04875128	66707375	AACTATAACTCTACAACAACAACAAACAATTCTAC TACAAATACATAACA	34724474
371123	cg09651136	cg09651136	69636474	CACTACACCCAAACAATTTAAAACTTTAAAATACA ATATAATCCAACATC	NA

FIGURE 9D

	ID	Name	AddressA_ID	AlleleA_ProbeSeq	AddressB_ID
377332	cg21296230	cg21296230	29625433	CRCRAAAATAAATTATAAAAAACCATCRCRAAATC CTTCCTACTAAAACC	NA
388382	cg04416734	cg04416734	71765505	TCAAAATCAAACTCCAATAAAAACCCAAAACCCCT ACCCTCTAAAAAACA	24794497
392814	cg07082267	cg07082267	49808346	AAAAAAAACCATAAAAATCTCCCRTTTCACAAATA AACACACCAAAACCC	NA
406094	cg02867102	cg02867102	24773488	AATAACTAAAAAATCTATCTAAATCCAAAACRCTA ACTTTAACCTTCCTC	NA
410441	cg06874016	cg06874016	71689361	AAAAACACAAAACRAACACAATTATAAATAATAAA CTTACTCTACAAACC	NA
418324	cg14692377	cg14692377	11791474	AACTACACAAAAAAACTAATCCCAAACTAAACAAA CAAACTAACCTCACA	74735323
424420	cg21139312	cg21139312	71738392	AAAACCACCCCAAACTCTATAATTTCCAAAACAAA TACAAAAACAACACA	36705403
434803	cg19283806	cg19283806	62709369	AATTTCTCCTTAAACAATCCCCACAAAAATAACAAC CAAAAAAAAATACA	12681323
449879	cg14556683	cg14556683	39707455	AAAAACACCAAACTCCACATAAAAACACACAACAA CTTCAACAACAAACA	23657395
465513	cg07547549	cg07547549	72801360	AACTCAACTCCATTAAAATACTCCRAACRCTATCCA AAATACTAAAATAC	NA
479076	cg05442902	cg05442902	45650355	CTCTCACTCTATACCTCTTAATTATCTTACATACTCT AATCTTTACATAC	NA
480022	cg08415592	cg08415592	18687383	AACAAATCTTTCTCCTTAAAACCACAACAAACCCCC AACCCTAAAAATAC	NA

FIGURE 9E

AlleleB_ProbeSeq	Infinium_Design_Type	Next_Base	Color_Channel
AAAATAAAAACATCTCAAACTCGCGTTACGAAA			
ACCGATTCGAAAAACCG	I	T	Red
	II		
CGAAACTACGACACCTTACGACGAAACCAAAAT			
TTAAATCTACGCAAACG	I	A	Red
	II		
	II		
	II		
	II		
	II		
	II		
	II		
ACGCGAAAACTCTTTAAAAAAACGACTCAACGA			
CTTAACCTTACCGTACG	I	C	Grn
CTTAAAAACACAATAATTATCGAAAACGTCGCCT			
CCGACGTAAACTCTCG	I	A	Red
TCTCGAAACCTTAACGACTTACCGCTAAAAACCC			
GCAATACAACAAAACG	I	C	Grn
AAAACACAATAATTATCGAAAACGTCGCCTCCG			
ACGTAAACTCTCGAACG	I	C	Grn
	II		

FIGURE 9A-1

AlleleB_ProbeSeq	Infinium_Design_Type	Next_Base	Color_Channel
	II		
	II		
	II		
GTATAACGCCAAATAAAAAATAATTCCTCAAAA ACCGTCGACCACCGACG	I	C	Grn
	II		
GCTATCCTCAAAAACCGCCAAAATACTAAAAAA AACGACAACAACGAACG	I	C	Grn
	II		
	II		
	II		
	II		
	II		
	II		
	II		
TCTAAACAACGAATAAATATTCCTAAAACTCCGT AAACGTTAAACCGCCG	I	C	Grn
	II		

FIGURE 9B-1

AlleleB_ProbeSeq	Infinium_Design_Type	Next_Base	Color_Channel
	II		
	II		
ACTAAAACACGAAAACGTACAAACGCCCAAACC			
TCGTATAATAAAAACCG	I	A	Red
ATCTCCGCGTTCTTTCTTCTACCGACGAACCAAA			
TAATAATAACAAAACG	I	A	Red
	II		
AATTCAAAACTTTAAACCCCCAACGCTCTACAAA			
CAAAAAACGAAACGCG	I	A	Red
	II		
ATAACGCTTAACAACAAATATAACAAACCTCCTC			
CGAAACGCCTAATCCG	I	C	Grn
TCTTCTACCGACGAACCAAATAATAATAACAAAA			
CGAAACTCCCCCAATCG	I	A	Red
	II		
CAAAAAACTAATTTCTACAATCGCTAAAACCTAC			
AACGACGCTACTCCCG	I	C	Grn
	II		
	II		
TTTCCACACGTAACTCCTCGCTACTTTATTACCTA			
AACCAAAAAAAAACG	I	A	Red
ACACACCCTCGAACGCCTTAAACGAAATACGCT			
AAAAAACCAAAAATTCG	I	A	Red

FIGURE 9C-1

AlleleB_ProbeSeq	Infinium_Design_Type	Next_Base	Color_Channel
TAATCACGTATCCTACCGCTACACTAAAAAACTTC			
CGACTAACAACAAACG	I	A	Red
	II		
ATCTACTAACACACACCGACAATCTAACACTCAT			
CTATATCCACACAACG	I	A	Red
AAAAACGACGACGCTAATAAAAACTAACCCGAC			
AATCCGAAAATCCACCG	I	C	Grn
CACGAAACTTTAAAAAACACTCTAACCCCTACTA			
CTCACCCATACAAACG	I	A	Red
	II		
	II		
	II		
	II		
	II		
	II		
	II		
AACTATAACTCTACGACGACAACGAACGATTCT			
ACTACGAATACGTAACG	I	C	Grn
	II		

FIGURE 9D-1

AlleleB_ProbeSeq	Infinium_Design_Type	Next_Base	Color_Channel
	II		
TCGAAATCAAACTCCGATAAAAACCCAAAACCC CTACCCTCTAAAAAACG	I	A	Red
	II		
	II		
	II		
AACTACGCGAAAAAACTAATCCCGAACTAAACA AACGAACTAACCTCGCG	I	C	Grn
AAAACCGCCCCGAACTCTATAATTTCCAAAACAA ATACGAAAACGACACG	I	T	Red
AATTTCTCCTTAAACAATCCCCGCAAAAATAACA ACCAAAAAAAAATACG	I	A	Red
GAAAACACCAAACTCCACATAAAAACGCGCAAC AACTTCAACGACAAACG	I	C	Grn
	II		
	II		
	II		

FIGURE 9E-1

Forward_Sequence	Genome_Build	CHR	MAPINFO
GGGTATTTTTCGGGGGTGAGGGCATCTCAGGCTCGCGTTACGGGGACCGGTTCGGGAGAC[CG]TGG AGCCGAGGTGTCGAATGGAGGGCTTACTTCGGCCGGGCCAGGGAGCGCCACCCTCCT	37	1	117665053
CCCCAGAGAGCTTTCATCTAGAAGGTTTGACTCTGGCCAGACAACCAGCGAGCATCTTCT[CG]CAATC TGTTGCTTCTTCCATGGCAAACTCCAGAGAATTAAGAAGCCAAACTCAACATCGC	37	1	236557682
ACGTGGGGGAAGAAGGGGGTTACGCCATCAAGTCCTGAAGCCCGTCGGACCACCCATCGC[CG]CCT GCGCAGACCCAAATCTTGGTCCCGCCGTAAGGTGCCGCAGTCCCGAATGTTCCAGAA	37	1	207997020
TTCATGAGCCGGACAAAGCTGTATCCCTCCATTTCCACCTGCCAACACCACGGAAGCAGT[CG]TCCGT TACCACTGACCTGAGGCCTGCCTGGGTCCAAGCTCACACTTGGAGAACCTTCTGT	37	1	169556022
AATGCCTGCTTCACAGAGAACTGCTGCGAGGATCACACAAGAAAATGCTTGTCAACTGGG[CG]TGGT GGCGCATGCCTGTAATCCCAGCTACTCGGAGACTAAGCCAGGAGAATCGCTTGAAC	37	1	17338766
GGGTCCTGATGGTGGTAACAACTGCTGATGCCCATCTTGGCCTGGGGCACTGAGATCGCC[CG]GAGA TCACAGTGTTAGCTTCAGGGCGGGGTAGAAATTAGAGGATAGGGGATCTCTAGGGC	37	1	26855765
CTAGCCTCACAGCACCGCGTGGAGTTGCTTGTTCTTTTACATAGGAGGTCACATTCTCTT[CG]TGTAAT GCCACCAATGGTGCCGATTCTCCCCAGTGGGGCTGTGAGAAACCTACGCCCTCT	37	1	28241577
AGTTTTGCCTCCAGGGAAACTGAGGCACAAGGCAGCAATGATTACTGAGGGTCCTGCCTC[CG]CTCC TCTAGGTGAGGAGCCTATTCCAGGGGCTCCAGTCTGAAAGCCTAGAGGCGAGGGGC	37	2	97202260
CCTTTGTTTGCCAGGGCTCCTTTCTTCGTGCCCTCCGGGTCTTGGGAGCACAGTAGTTAT[CG]GGAGC GTCGCCTCCGGCGTGGGCTCTCGGGCGCGAGTTTCGGACGAGGCCTGGGCGCGGT	37	2	106015739
CTACAGTGCCCGCAGGCCCCTTAAAAGGGAGCCAGTGCAACAGATTGGTTAACCAGGCCA[CG]TGTT CCCAGTCTTCTAAAAATCCCCAAAGACTGGACAGCAAATGCCCCTTAGGTTGCATG	37	2	230452311
AGAGCGCTACGTCGCCGGCGGGCAGCAGCAGCGCCTACAAACTGGAGGCGGCGGCGCAGG[CG]CA CGGCAAGGCCAAGCCGCTGAGCCGCTCTCTCAAAGAGTTCCCGCGTGCGCCGCCAGCC	37	2	39187539
TGCCCTCCGGGTCTTGGGAGCACAGTAGTTATCGGGAGCGTCGCCTCCGGCGTGGGCTCT[CG]GGCG CGAGTTTCGGACGAGGCCTGGGCGCGGTGGCAGGGGTCTGCCCACGCCGGGATCTC	37	2	106015767
CTGACCGTGGTGCTGAGCGCGGCTCGCGCTCCGACGCGGTGCCCGAGCCTGTCGCGGCCG[CG]CCCT GCTGCACTGCGGGCCCCCAGCGGTAAGTCGCCAAGGCCCCGAGAGGCTGCGTTGGT	37	2	131513927
CTCCGGGTCTTGGGAGCACAGTAGTTATCGGGAGCGTCGCCTCCGGCGTGGGCTCTCGGG[CG]CGA GTTTCGGACGAGGCCTGGGCGCGGTGGCAGGGGTCTGCCCACGCCGGGATCTCTGCC	37	2	106015771
ACCCGCTATCTGGTGGAGATAGAGAAAGGCAGATGGCTTGAAGATCCAATTTGGAGATTA[CA]TCCA TAGGACTTGCTGATGATCACATGTGGCAGAGGAAGAGCCAGTCCTCAAATATGGCC	36	2	30561970

FIGURE 9A-2

Forward_Sequence	Genome_Build	CHR	MAPINFO
TCAGAGGTCTGCGTTTCAGCACCTGGGTCAGCGCTTCCCAGGGTCAGCACCAGGGATAGA[CG]CCAT TGTCGCTTCGTGCCCAGACAACTACCCAGATTTTGTAATTTTCTTGAATGCCAAGT	37	3	187387650
CGCTGTGGCCCCGAGCGGAACGGCCCGGAAGAGGAGACGCGTCCCCGGGAACCCAGTGCC[CG]CCC TGGCCCAGCCCCGATCCAGCCTGCGCCTCACCTCGGGTTGTAGACAGAGCGGCGGGG	37	3	53080440
CCAGCCAAGTGGCCTTGATCGTTTTCCCAATGCCCCCGAGCCTGTTTCCTGCCAGTAGAG[CG]GGTCA GATGTTGCCAACCTCTGCAGAGTAGCAATAAGCAGTAAACGCCACGCTCTGCACA	37	3	52008487
AATCCGCATGGCACCGGTGGTCTGGGGGAGAGGCTGGGCCTGGCGCGGGACGAGGCGAAG[CG]CC GGTGGCCGACGGCTTCTGAGGAATTATCTTTTACTTGGCGCCACACGGGGCGGGGCCT	37	3	160167977
TCGGGAGCTGAGGGACCCAGAAAAGCACCAAAACTCTTTAGAAGGACTGAGCATCCCTTA[CG]TCCA AACCAATGGGGCAGGAGCAAGGCTTAGGGAGGGCTGGAGAATCCGGGAGACGTCGA	37	3	187387866
GCCCGAGAGGATCCAGGGAAAGCAGAAGGGGGTTAAGGACCATGGACAGAGCCCGTCGCG[CG]CT CGTTGCTGCCGCCTTCCCCAGCACTCTGGCGGCTCCTGAGGACAGCGGTCCCATCTTG	37	4	147558239
GGGAGGGTGGGAGAGGGATGAGGGATGAGAAATTACCTATTAGATACAATGTACAACATT[CG]GGT GACAGTTACACTAAGAGCCCAGATGCCACCACAATGCAATATACCCATGGAACAGAA	37	4	84255411
TTGTATTTCAGCCCAAAGCCTACTGGAAGTGTCAAGCTGCCAGCTCCCCTCTGCCCTCCC[CG]TTGCTA TGGCAGCCATGTCTCTGTGTGTGAATAGGTGAACCAGGCTCCAGGTTAGGACCT	37	5	139013317
CAGGAGTGCGGTGCAGCCACACATCCAAGGCTGACAGGGCGGGCACTCTGCCAAGTCCTG[CG]CGC TGCTCGCCTTCCACAACACCTTCCTCAGCTTCGTCTGTATTTGAAGAGCTTAGTAAA	37	5	140419819
ATTCACGGGGAATCAGAGGGTGGAGAGGGCGTGGGTGCCTGGAGATGCCTGGGAACAGAA[CG]GC TGAGGGGACTCCATTATCTGTACTCTTCCCGGGGTGGGTCTAGGTCTGGCTCCTCCTG	37	6	31105711
GGAGAGCAAGTCAAGAAATACGGTGAAGGAGTCCTTCCCAAAGTTGTCTAGGTCCTTCCG[CG]CCGG TGCCTGGTCTTCGTCGTCAACACCATGGACAGCTCCCGGGAACCGACTCTGGGGCG	37	6	25652602
TCACCACTTCTTTGCCAGTCTAGATCCGTCCTGGTGCCTTACTGTGCATACAGTTCTACT[CG]TCTCAG GTGAGGAGGCCACTTAATTTGTAAAAGACTGAGGAAGGGGTAGGATCACCACAA	37	6	30295466
TGTGTTACTAAGTGAAGTCTACTATACATAGTAAATACTACGCAGTCATACCTTTATAAA[CG]ACAAA ATCTGACCAGGCTAGATCTGAATCTGTTAATTTCTGAGACGGTGTGTCTGAAGTG	37	6	147528521
CCGCGGCGTCCCCTGCCGGCCGGGCGGCGATTTGCAGGTCCAGCCGGCGCCGGTTTCGCG[CG]GCG GCTCAACGTCCACGGAGCCCCAGGAATACCCACCCGCTGCCCAGATCGGCAGCCGCT	37	6	11044877
CTTGCGCCTCGAATGCCACGTTGAATACTCCTCATGTCTTTGGAGACATGTCCTTCCCTT[CG]AGCTGC TCCCAGTCAGGTGAGGAATAAAATGCTATGATGGCGTGAAAATTCTCCCTTGGT	37	6	10510789

FIGURE 9B-2

Forward_Sequence	Genome_Build	CHR	MAPINFO
TGCGCCAGGGCGGCCACGCAGGCCAGGCAGACCACGTGGCCGCAGGACAGGTTGCGCGGG[CG]CC			
GCTGCTGCCGGTGGCCAAACTTCTCAAAGCACACCTTGCACTCGAGCAGGCTGATCTC	37	6	18122719
AATTAAAGACTAATTCAGAATTTTCAAGTGATAGTAAACAACTGCTATCTCAAACACATA[CG]ATATA			
AAATGAAACCACTGGTGCCTAACTGCCAGTTCTTTCACTCAAACCTCTGCTGTGA	37	7	131008743
CGTGGCTGCGCCCACAAAGCCGCCGGGGGCTGCGGACTACAGCGAAGCCGGCGCGGGGCT[CG]GC			
CCTCACTACACGAGGCCTGGGCGCCTGCACGCCCCCGTGCTTCAGCCCGCGGCTCCCG	37	7	156433108
AGAAATTATCTTGTCTCCGCGTTCTTTCTTCTGCCGGCGAGCCAGGTAATGGTAACAGAG[CG]AAACT			
CCCCAGTCGGAACTTCTGGGTTGCAGCAGCCTCGCCCCCCCTCCCCGCAGCCCGC	37	7	130419159
CCGGCTAAGTCATGTTTAACAGCCTCAGAAATTATCTTGTCTCCGCGTTCTTTCTTCTGC[CG]GCGAGC			
CAGGTAATGGTAACAGAGCGAAACTCCCCAGTCGGAACTTCTGGGTTGCAGCAG	37	7	130419133
CCCCGAGGCGGACGCCAGAGGGCGCGCGCCCCCACTCCTGCCCGCGTCGGGGCCGCAGC[CG]CGC			
TCCGCCCTTTGCCTGCAGAGCGCTGGGGGTTTAAAGTCCTGAACCCATGCACGGCTG	37	7	152591698
TCGGAACGCGCGGCCCCCCGGCTAAGTCATGTTTAACAGCCTCAGAAATTATCTTGTCTC[CG]CGTTC			
TTTCTTCTGCCGGCGAGCCAGGTAATGGTAACAGAGCGAAACTCCCCAGTCGGAA	37	7	130419116
GAAGGGGCATTGGTGGCGCTTGGCAGCAGGTGTGACAGACCTCCTCCGGGGCGCCTGATC[CG]CGG			
CGGGGGCGGGGCCTGCCCCTAGGGCCCCTCCAGAGAACCCACCAGAGGCTGCTGGTG	37	7	130418324
CTCCGCGTTCTTTCTTCTGCCGGCGAGCCAGGTAATGGTAACAGAGCGAAACTCCCCAGT[CG]GAACT			
TCTGGGTTGCAGCAGCCTCGCCCCCCCTCCCCGCAGCCCGCCCGCTGGTCCG	37	7	130419173
CAAACACCAGGGCAGCCCCCATTTAAGGTTTTTGATACACTGAGGATCATTCAGAAAACTT[CG]GATTC			
CTAGTTATAGAGTTGAATCCAACCACCAACACACTCCAGAAGTCCTGACATTAGG	37	8	82644012
CTGCGCCCTCTGCAAAGGGCTGATTTCTACAGTCGCTAGGACCTGCAGCGGCGCTGCTCCC[CG]CGGG			
GCTCCGGCCGCGCTGCATGTCCCATTATAGTCGCTAGAGGGCAGCGCTCTCCTGCG	37	8	57360613
GCAACACAGAGTGCTTTACCCATGCCTTAAGGCAGCTGTCACATGACTCTTCCCCTAATA[CG]GTTGT			
GAGGTGCTCACGTGTTTTGGAGATAGCAAAAGTCTCAAATAATATGGCACAGACG	37	10	98062687
TCCTAAGCCTCTCTGAGCTGGGCTTGGCCACCTTCCGGGGTGTGAGCGTCCACGGGAGAT[CG]ACCA			
CACCAGGCACCCAGGAGCAAGTGCTTTGAAATGCGGCTTTCTCCGGACCTTGCAGG	37	10	49673534
GAGTCAGTGCACAGAGTTAGAGGCAGGCAGAAAGAAAGGAGACCTCCCCAACCTCGGCCC[CG]CCC			
TTCCCTGGCCCAGGCAACAAAGCAGCGAGGAGCCACGTGTGGAAAAGCAGTGCAGAT	37	11	133797911
GCGCACACACGCACACACCCTCGGGCGCCTTGGACGGGGTGCGCTGGGGAGCCAGAAGTT[CG]GAG			
CGAGCGCGGGCGGGCAGAGCCGCCTCGGAGCCCGGAGCCGGCCTGCACCCCCCT	37	11	12696758

FIGURE 9C-2

Forward_Sequence	Genome_Build	CHR	MAPINFO
CACCGGGCTCACACTGCTGCTCGCACGGAGCCTGGGCACAGGGGTCCTCGCAACTGCGCC[CG]TCTG CTGCCAGCCGGAAGCCCTCAGTGCAGCGGCAGGACACGTGACCATCCACCTCCTCC	37	11	66083697
CCTCGCTACAATAACACTCGAAGCCACCAACGTAGTTGACACACATCTGCTGGCACACAC[CG]GCAAT CTGGCACTCATCTGTGTCCACACAGCGGTGCGGATCATCCTCCGCTGGCCGGAAA	37	11	66083541
TAGTTGACACACATCTGCTGGCACACACCGGCAATCTGGCACTCATCTGTGTCCACACAG[CG]GTGCG GATCATCCTCCGCTGGCCGGAAACCCAGGCGACAGTGGCAGCTGTAGCCTTGTGG	37	11	66083573
CAAAGCCGGCGAGGAGGCGGCGGCGCTGGTGGGGACTGACCCGGCAGTCCGAGAATCCAC[CG]CG GCCTTTTCACCCAACCGCCCCCTCCTGCGTGGGGGCCCCGCATCCCCTGGACTGGCGT	37	11	9025767
TCGACTTAAACCCACGAAGCTTTGGGGAGCACTCTAGCCCCTGCTACTCACCCATGCAAG[CG]GGGT GCGCGCTCGCGCACACACTCACTCACTCCAAGATAGGGGCTTTCTAGGAAAATACT	37	11	10323902
CCGGTGCGCCGGGCTCTACCTCAAGGAGCTCAGGGCCATCGTGCTGAACCAACAGAGGCT[CG]TCCG CACCCAGCGCCAGAGCATCGACGAGCTGGAGCGGCGGCTGAACGAGCTGAGCGCCT	37	12	11653486
TGTTACAATTTAACTACTTTCTCTTTCTCTTTCTCTCTCTCTCTCTCTCTCTGGTAAAAA[CG]TTAACCTCT GCTAGTGATGACCAAACCTGGTAAAGATTGTAAAGTGGGAAAAATTGGATT	37	12	81468232
CTCCTCGGCCCCCTCTGGCCCCCGTCCTCCTTTCTCTCCTTCCCCCTCTCTCTCCAGGAG[CG]ACTCTGG GTTAGCACAATTGAACTGGATTTACGAGCGAGAATGGGTAATTACATCCCCCA	37	12	54448265
TCTGCTTACAGCTGCTTCCAAATTAAGCATATCTGGATGGTGTGACACTTTTTGTTAGTC[CG]AGAACT GTATGGGCATCGCAACTGGGCCTGTTCCAAGATAGACTTGTTGGGACCTTCAAA	37	12	30849114
GTGAATGGATTTGGGGAGGGAGGATCAGACTAGAGGCCAGAAGTCCGGTGAGCAATATTA[CA]GTA GTCCAGGTGAGAGGTAATAATGGCCTGCCTAACCCAGGGAAGCAGCCACGGGGATTA	36	13	40666907
CACACCACTCGTATCTAACTCAACCCCTTTAGATATTCTTCCAGGTGGAATTATTGGATT[CG]GTCAGA ATGGGGGAGGGGCCACTATGCCCTTAAGAGGCTCAGAAGTGCCTACCTGGCTAA	37	14	61108227
TTCACCTGACCCCCATGCTCGGCCTTCTGGAAGATGCCCACAGACACTGGCAATAATGGA[CG]CTGGC ACACTCTGCCGGCGCGGCCGGAGCCCCGGAGTTCAGCACCTCGGACAGAGCCCGA	37	15	79576060
CGGCGCGCGCCGGGCTGTAGCTCTGCGACGACAGCGAGCGGTTCTGCTGCGGGTACGTGG[CG]CAC GGCCGCAGCGCCCCCACGGCCGGCGCGCACGCCTCGTCCCGCGCGCCCGACGCCTGC	37	15	31775895
CCTGAGCTGAGGACAAACTAAGAACAGTGGTGGACAACACTGGGATGGCAAGGATCTGGC[CG]ATG CTGGACTATATTGTATTTTAAAGTTCTAAATTGTCTGGGTGCAGTGGCTCACGCCTG	37	15	72525012

FIGURE 9D-2

Forward_Sequence	Genome_Build	CHR	MAPINFO
GGTGCGTTGTTCGCGGGGGTGAATTGTGAAGAACCATCGCGGGGTCCTTCCTGCTGAGGC[CG]CGG ACACCGTGACCTCGCTGCTCTGGGTCTGCAGGGAAACGTAGGAAAAAAAGTTGTCAG	37	15	33010536
ACTGTCTCTGCTTCGAGATCAAGCTCCGATGAGGACCCAGGGCCCCTGCCCTCTGGGGAG[CG]GCCA GCCCCCAGGGCCCATGTGCCCTCCTCCCTGAAGAGCCTTTCCCCACGCCACTGGAA	37	16	30075192
GCTCCTCATGTGAGAAGGACCATAGGAATCTCCCGTTTCACAGGTGGGCACACCAAGGCC[CG]ACAA TGGGTCCAGGCTGCCAAGGGTGGAGCCGAGATGCAAAGGGGCACCTCAGAGCCTGC	37	16	85429035
CTCCTGGAGTGGGTGCTCCTGGGATGCTTCAGGTTTAGACACCGGGGTTACGGCAGCTGC[CG]AGGA AGGCTAAAGCCAGCGTCCTGGATTCAGACAGACCTTTTAGCCATTAAATCCACTAA	37	17	62398693
CAGCCTCTCAGGAGCTGACAGGTCCTCTTTCGGGGCTCAGGAGGGTGGGCACACACCCAG[CG]GCCT GCAGAGTAAGCTTATTACCCACAACTGTGCCCGCTTTGTGCTTCTAAGGTGCACAC	37	17	40177415
ACCCACCTGCCAGGCTGCGCGGGGAGGCTGGTCCCGGGCTGGGCAGGCGGGCTGGCCTCG[CG]CCC TCGAGGCACCCGGCGGCGCTGGCTGTGCGGAGGGGCGCCGGCGCGGCCGTATTTGTA	37	17	28562685
TTGGAGCTCGCCAAAGCCGCCCCGGGCTCTGTGGTTTCCAGAGCAGATGCGGAGGCGGCA[CG]TCCT CGTGCCCTTGCTCCAGCTGCGCACACGACCTCAGCCTCCTCTGCCCCGCTGGGCGC	37	17	55663225
TCCGTAGTATTGTCTCTGGCTTTGAACGCTGTTGAGGGAGGGGAATGTTTGCACTCATCC[CG]CATCC TTTTTTGGCTGCTATCTTTGCGGGGATTGTTCAAGGAGAAATCCATCCTGACTGG	37	18	66389420
ACCAGCGCCACCGAGAACACCAGGCTCCACATGAAGGCGCGCAGCAGCTTCAGCGACAGG[CG]CGA CGGCGCCAGCAGCGCGGTCACCACCAGCTCCGGCATGTCGCCGCGCTCCGGGACCAC	37	19	15342982
TTGCAGCCTGGAGCTCAGCTCCATTGGAATGCTCCGGGCGCTGTCCAAGGTGCTGGAATG[CG]CCGC GCCCGGGGGCAGAGCTGCGGGCCGGGGGATTATCGCTGCCCACGGCTTCGGGCTGA	37	20	44658225
GCCAGGTCACCCTCTCACTCTGTGCCTCTTAGTTATCTTGCATGCTCTGGTCTTTGCATA[CG]CTGCTCC CTGCACCAGGAACCTCCATCCCCATCTTTGTCTGCTTGTCGAACTTCAGAAAT	37	22	21369010
AGTATGTCAGTGGCAGGTCTTTCTCCTTGAGACCACAGCAGACCCCCAGCCCTGAGGATG[CG]AGGC AGGTGGGTTGGATGAGAGGGATCTGGATGTCTGGTCTCAGGCTGCTCCTCTAAGGG	37	22	36648973

FIGURE 9E-2

SourceSeq	Chromosome_36	Coordinate_36	Strand	Probe_SNPs	Probe_SNPs_10
CGGTCTCCCGAACCGGTCCCCGTAACGCGAGCCTGAGATGCCCTCACCCC	1	117466576	R		rs78180333
TTCATCTAGAAGGTTTGACTCTGGCCAGACAACCAGCGAGCATCTTCTCG	1	234624305	R		
CGGGACTGCGGCACCTTACGGCGGGACCAAGATTTGGGTCTGCGCAGGCG	1	206063643	F	rs61821292	
CGTCCGTTACCACTGACCTGAGGCCTGCCTGGGTCCAAGCTCACACTTGG	1	167822646	F	rs9332483	
CGCCCAGTTGACAAGCATTTTCTTGTGTGATCCTCGCAGCAGTTCTCTGT	1	17211353	R		
CGGGCGATCTCAGTGCCCCAGGCCAAGATGGGCATCAGCAGTTGTTACCA	1	26728352	R	rs59827434	
CACCGCGTGGAGTTGCTTGTTCTTTTACATAGGAGGTCACATTCTCTTCG	1	28114164	R		
TAGGCTTTCAGACTGGAGCCCCTGGAATAGGCTCCTCACCTAGAGGAGCG	2	96565987	F		rs13020129
AGGGCTCCTTTCTTCGTGCCCTCCGGGTCTTGGGAGCACAGTAGTTATCG	2	105382171	R		
CGTGGCCTGGTTAACCAATCTGTTGCACTGGCTCCCTTTTAAGGGGCCTG	2	230160555	R		
ACGCGGGAACTCTTTGAGAGAGCGGCTCAGCGGCTTGGCCTTGCCGTGCG	2	39041043	F		
CTTGGGAGCACAGTAGTTATCGGGAGCGTCGCCTCCGGCGTGGGCTCTCG	2	105382199	R		
TCTCGGGGCCTTGGCGACTTACCGCTGGGGGCCCGCAGTGCAGCAGGGCG	2	131230397	F		
CGCCCGAGAGCCCACGCCGGAGGCGACGCTCCCGATAACTACTGTGCTCC	2	105382203	R		
CATCCATAGGACTTGCTGATGATCACATGTGGCAGAGGAAGAGCCAGTCC	2	30415474	F		

FIGURE 9A-3

SourceSeq	Chromosome_36	Coordinate_36	Strand	Probe_SNPs	Probe_SNPs_10
GTTTCAGCACCTGGGTCAGCGCTTCCCAGGGTCAGCACCAGGGATAGACG	3	188870344	R		
CGCCCTGGCCCAGCCCCGATCCAGCCTGCGCCTCACCTCGGGTTGTAGAC	3	53055480	F		
CGGGTCAGATGTTGCCAACCTCTGCAGAGTAGCAATAAGCAGTAAACGCC	3	51983527	F		
CGCCGGTGGCCGACGGCTTCTGAGGAATTATCTTTTACTTGGCGCCACAC	3	161650671	F		
GGACCCAGAAAAGCACCAAAACTCTTTAGAAGGACTGAGCATCCCTTACG	3	188870560	R		
GCTGTCCTCAGGAGCCGCCAGAGTGCTGGGGAAGGCGGCAGCAACGAGCG	4	147777689	F		rs76832102
CGGGTGACAGTTACACTAAGAGCCCAGATGCCACCACAATGCAATATACC	4	84474435	F		
GGAGCCTGGTTCACCTATTCACACACAGAGACATGGCTGCCATAGCAACG	5	138993501	F		
GCAGCCACACATCCAAGGCTGACAGGGCGGGCACTCTGCCAAGTCCTGCG	5	140400003	R		
CGGCTGAGGGGACTCCATTATCTGTACTCTTCCCGGGGTGGGTCTAGGTC	6	31213690	F	rs35474509	
AAGAAATACGGTGAAGGAGTCCTTCCCAAAGTTGTCTAGGTCCTTCCGCG	6	25760581	R		
TGCCAGTCTAGATCCGTCCTGGTGCCTTACTGTGCATACAGTTCTACTCG	6	30403445	R		
CGACAAAATCTGACCAGGCTAGATCTGAATCTGTTAATTTCTGAGACGGT	6	147570214	F		
CGGCGGCTCAACGTCCACGGAGCCCCAGGAATACCCACCCGCTGCCCAGA	6	11152863	F		
CGAGCTGCTCCCAGTCAGGTGAGGAATAAAATGCTATGATGGCGTGAAAA	6	10618775	F		

FIGURE 9B-3

SourceSeq	Chromosome_36	Coordinate_36	Strand	Probe_SNPs	Probe_SNPs_10
CTCGAGTGCAAGGTGTGCTTTGAGAAGTTTGGCCACCGGCAGCAGCGGCG	6	18230698	F	rs28940575	
CGTATGTGTTTGAGATAGCAGTTGTTTACTATCACTTGAAAATTCTGAAT	7	130659283	R		rs76007347
CGGCCCTCACTACACGAGGCCTGGGCGCCTGCACGCCCCCGTGCTTCAGC	7	156125869	F		rs74424274
CGCTCTGTTACCATTACCTGGCTCGCCGGCAGAAGAAAGAACGCGGAGAC	7	130069699	R		
TGTTTAACAGCCTCAGAAATTATCTTGTCTCCGCGTTCTTTCTTCTGCCG	7	130069673	R		
CGCGCTCCGCCCTTTGCCTGCAGAGCGCTGGGGGTTTAAAGTCCTGAACC	7	152222631	F		
GCCCCCCGGCTAAGTCATGTTTAACAGCCTCAGAAATTATCTTGTCTCCG	7	130069656	R		
GTGGCGCTTGGCAGCAGGTGTGACAGACCTCCTCCGGGGCGCCTGATCCG	7	130068864	R		
TCTTCTGCCGGCGAGCCAGGTAATGGTAACAGAGCGAAACTCCCCAGTCG	7	130069713	R		
CAGCCCCATTTAAGGTTTTTGATACACTGAGGATCATTCAGAAAACTTCG	8	82806567	R		
CAAAGGGCTGATTTCTACAGTCGCTAGGACCTGCAGCGGCGCTGCTCCCG	8	57523167	R		
CGGTTGTGAGGTGCTCACGTGTTTTGGAGATAGCAAAAGTCTCAAATAAT	10	98052677	F		
CGACCACACCAGGCACCCAGGAGCAAGTGCTTTGAAATGCGGCTTTCTCC	10	49343540	F	rs11101335	
TTTCCACACGTGGCTCCTCGCTGCTTTGTTGCCTGGGCCAGGGAAGGGCG	11	133303121	F		
ACACACCCTCGGGCGCCTTGGACGGGGTGCGCTGGGGAGCCAGAAGTTCG	11	12653334	R		

FIGURE 9C-3

SourceSeq	Chromosome_36	Coordinate_36	Strand	Probe_SNPs	Probe_SNPs_10
TGGTCACGTGTCCTGCCGCTGCACTGAGGGCTTCCGGCTGGCAGCAGACG	11	65840273	F		
CGGCAATCTGGCACTCATCTGTGTCCACACAGCGGTGCGGATCATCCTCC	11	65840117	F		
ATCTGCTGGCACACACCGGCAATCTGGCACTCATCTGTGTCCACACAGCG	11	65840149	R		
GGAGGCGGCGGCGCTGGTGGGGACTGACCCGGCAGTCCGAGAATCCACCG	11	8982343	R		
CACGAAGCTTTGGGGAGCACTCTAGCCCCTGCTACTCACCCATGCAAGCG	11	10280478	R		
GCTCTACCTCAAGGAGCTCAGGGCCATCGTGCTGAACCAACAGAGGCTCG	12	11544753	R		
CGTTAACCTCTGCTAGTGATGACCAAACCTGGTAAAGATTGTAAAGTGGG	12	79992363	F		
TTACCCATTCTCGCTCGTAAATCCAGTTCAATTGTGCTAACCCAGAGTCG	12	52734532	F		
AACAAGTCTATCTTGGAACAGGCCCAGTTGCGATGCCCATACAGTTCTCG	12	30740381	F		
CAGTAGTCCAGGTGAGAGGTAATAATGGCCTGCCTAACCCAGGGAAGCAG	13	39564907	F		
ATCTAACTCAACCCCTTTAGATATTCTTCCAGGTGGAATTATTGGATTCG	14	60177980	R		
CCATGCTCGGCCTTCTGGAAGATGCCCACAGACACTGGCAATAATGGACG	15	77363115	R		
CGCCACGTACCCGCAGCAGAACCGCTCGCTGTCGTCGCAGAGCTACAGCC	15	29563187	R		
ACTGCACCCAGACAATTTAGAACTTTAAAATACAATATAGTCCAGCATCG	15	70312066	F		

FIGURE 9D-3

SourceSeq	Chromosome_36	Coordinate_36	Strand	Probe_SNPs	Probe_SNPs_10
GCGGGGGTGAATTGTGAAGAACCATCGCGGGGTCCTTCCTGCTGAGGCCG	15	30797828	R		
TCGAGATCAAGCTCCGATGAGGACCCAGGGCCCCTGCCCTCTGGGGAGCG	16	29982693	R		rs36029842
CGGGCCTTGGTGTGCCCACCTGTGAAACGGGAGATTCCTATGGTCCTTCT	16	83986536	R		
ATGGCTAAAAGGTCTGTCTGAATCCAGGACGCTGGCTTTAGCCTTCCTCG	17	59752425	F	rs71377739	
CGGCCTGCAGAGTAAGCTTATTACCCACAACTGTGCCCGCTTTGTGCTTC	17	37430941	F	rs1128719	
GGCTGCGCGGGGAGGCTGGTCCCGGGCTGGGCAGGCGGGCTGGCCTCGCG	17	25586811	R		
CGTGCCGCCTCCGCATCTGCTCTGGAAACCACAGAGCCCGGGGCGGCTTT	17	53018224	R		
GATTTCTCCTTGAACAATCCCCGCAAAGATAGCAGCCAAAAAAGGATGCG	18	64540400	F		
GAGAACACCAGGCTCCACATGAAGGCGCGCAGCAGCTTCAGCGACAGGCG	19	15203982	R		
GCTCAGCTCCATTGGAATGCTCCGGGCGCTGTCCAAGGTGCTGGAATGCG	20	44091632	R		
TCTCACTCTGTGCCTCTTAGTTATCTTGCATGCTCTGGTCTTTGCATACG	22	19699010	R		
CGCATCCTCAGGGCTGGGGGTCTGCTGTGGTCTCAAGGAGAAAGACCTGC	22	34978919	R		

FIGURE 9E-3

Random_Loci	Methyl27_Loci	UCSC_RefGene_Name	UCSC_RefGene_Accession
NA	NA	TRIM45;TRIM45	NM_025188;NM_001145635
NA	TRUE	EDARADD;EDARADD;EDARADD	NM_080738;NM_145861;NM_145861
NA	NA		
NA	NA	F5	NM_000130
NA	NA	ATP13A2;ATP13A2;ATP13A2	NM_022089;NM_001141974;NM_001141973
NA	NA	RPS6KA1	NM_002953
NA	NA	RPA2	NM_002946
NA	NA	ARID5A	NM_212481
NA	NA	FHL2;FHL2;FHL2;FHL2	NM_001039492;NM_001450;NM_201557;NM_201555
NA	NA	DNER	NM_139072
NA	NA	LOC375196;LOC100271715	NR_028386;NM_001145451
NA	NA	FHL2;FHL2;FHL2;FHL2	NM_001039492;NM_001450;NM_201557;NM_201555
NA	NA	FAM123C;FAM123C;FAM123C;FAM123C;FAM123C;FAM123C	NM_152698;NM_001105194;NM_001105195;NM_001105194;NM_001105193;NM_001105195
NA	NA	FHL2;FHL2;FHL2;FHL2	NM_001039492;NM_001450;NM_201557;NM_201555
NA	NA		

FIGURE 9A-4

Random_Loci	Methyl27_Loci	UCSC_RefGene_Name	UCSC_RefGene_Accession
NA	NA	SST	NM_001048
NA	NA	SFMBT1;SFMBT1;SFMBT1	NM_001005159;NM_016329;NM_001005158
NA	TRUE	ABHD14B;ABHD14B;ABHD14B;ABHD14A;ABHD14B;ABHD14B	NM_001146314;NR_027476;NM_032750;NM_015407;NM_032750;NM_001146314
NA	NA	TRIM59	NM_173084
NA	NA	SST	NM_001048
NA	NA		
NA	NA	HPSE;HPSE;HPSE	NM_001098540;NM_001166498;NM_006665
NA	NA		
NA	NA		
NA	NA	PSORS1C1;PSORS1C2	NM_014068;NM_014069
NA	TRUE	SCGN;SCGN	NM_006998;NM_006998
NA	NA	HCG18;TRIM39;TRIM39;HCG18;TRIM39	NR_024053;NM_172016;NM_021253;NR_024052;NM_172016
NA	NA	STXBP5;STXBP5	NM_001127715;NM_139244
NA	NA	ELOVL2	NM_017770
NA	NA		

FIGURE 9B-4

Random_Loci	Methyl27_Loci	UCSC_RefGene_Name	UCSC_RefGene_Accession
NA	TRUE	NHLRC1	NM_198586
NA	NA	MKLN1	NM_001145354
NA	NA	C7orf13;RNF32	NR_026865;NM_030936
NA	NA	KLF14	NM_138693
NA	NA	KLF14	NM_138693
NA	NA		
NA	NA	KLF14	NM_138693
NA	NA	KLF14	NM_138693
NA	NA	KLF14	NM_138693
NA	NA	CHMP4C	NM_152284
NA	NA	PENK	NM_001135690
NA	NA	DNTT;DNTT	NM_004088;NM_001017520
NA	NA	ARHGAP22	NM_021226
NA	NA	IGSF9B	NM_014987
NA	NA	TEAD1	NM_021961

FIGURE 9C-4

Random_Loci	Methyl27_Loci	UCSC_RefGene_Name	UCSC_RefGene_Accession
NA	NA	CD248	NM_020404
NA	NA	CD248	NM_020404
NA	NA	CD248	NM_020404
NA	NA	NRIP3	NM_020645
NA	NA		
NA	NA		
NA	NA		
NA	NA	HOXC4;HOXC4	NM_153633;NM_014620
NA	TRUE	IPO8	NM_006390
NA	NA		
NA	NA		
NA	NA	ANKRD34C	NM_001146341
NA	NA	OTUD7A	NM_130901
NA	NA	PKM2;PKM2;PKM2	NM_182470;NM_182471;NM_002654

FIGURE 9D-4

Random_Loci	Methyl27_Loci	UCSC_RefGene_Name	UCSC_RefGene_Accession
NA	TRUE	GREM1	NM_013372
NA	NA	ALDOA;ALDOA;ALDOA	NM_001127617;NM_184043;NM_000034
NA	NA		
NA	NA		
NA	NA	NKIRAS2;NKIRAS2;NKIRAS2;NKIRAS2;NKIRAS2	NM_001144927;NM_001144928;NM_017595;NM_001144929;NM_001001349
NA	NA	SLC6A4;SLC6A4	NM_001045;NM_001045
NA	NA	MSI2;MSI2	NM_138962;NM_170721
NA	NA	CCDC102B	NM_001093729
NA	NA	EPHX3;EPHX3	NM_024794;NM_001142886
NA	NA	SLC12A5;SLC12A5	NM_020708;NM_001134771
NA	TRUE	MGC16703;P2RX6;P2RX6	NR_003608;NM_005446;NM_001159554
NA	NA	APOL1;APOL1;APOL1;APOL1	NM_001136541;NM_003661;NM_001136540;NM_145343

FIGURE 9E-4

UCSC_RefGene_Group	UCSC_CpG_Islands_Name	Relation_to_UCSC_CpG_Island
TSS1500;TSS1500	chr1:117664180-117665148	Island
TSS1500;5'UTR;1stExon	chr1:236558459-236559336	N_Shore
TSS1500		
TSS1500;TSS1500;TSS1500	chr1:17337829-17338590	S_Shore
TSS1500	chr1:26856191-26856684	N_Shore
TSS1500	chr1:28240584-28241535	S_Shore
TSS1500	chr2:97202313-97203787	N_Shore
TSS200;TSS200;5'UTR;TSS200	chr2:106014878-106015884	Island
Body		
TSS200;Body	chr2:39186777-39187968	Island
TSS200;TSS200;5'UTR;TSS200	chr2:106014878-106015884	Island
5'UTR;5'UTR;1stExon;1stExon;5'UTR;5'UTR	chr2:131513363-131514183	Island
TSS200;TSS200;5'UTR;TSS200	chr2:106014878-106015884	Island

FIGURE 9A-5

UCSC_RefGene_Group	UCSC_CpG_Islands_Name	Relation_to_UCSC_CpG_Island
Body	chr3:187387914-187388176	N_Shore
TSS1500;TSS1500;TSS1500	chr3:53078956-53081101	Island
1stExon;Body;5'UTR;TSS1500;1stExon;5'UTR	chr3:52008943-52009339	N_Shore
TSS1500	chr3:160167184-160168200	Island
Body	chr3:187387914-187388176	N_Shore
	chr4:147558231-147558583	Island
Body;Body;Body	chr4:84255726-84256399	N_Shore
	chr5:139017133-139017668	N_Shelf
Body;3'UTR		
1stExon;5'UTR	chr6:25652380-25652709	Island
TSS1500;5'UTR;5'UTR;TSS1500;1stExon	chr6:30294169-30295071	S_Shore
Body;Body	chr6:147525374-147525848	S_Shelf
TSS1500	chr6:11043913-11045206	Island

FIGURE 9B-5

UCSC_RefGene_Group	UCSC_CpG_Islands_Name	Relation_to_UCSC_CpG_Island
1stExon	chr6:18122250-18122994	Island
Body	chr7:131012460-131013190	N_Shelf
Body;TSS1500	chr7:156432433-156433670	Island
TSS1500	chr7:130417912-130419378	Island
TSS1500	chr7:130417912-130419378	Island
	chr7:152591458-152591706	Island
TSS1500	chr7:130417912-130419378	Island
1stExon	chr7:130417912-130419378	Island
TSS1500	chr7:130417912-130419378	Island
TSS1500	chr8:82644603-82644849	N_Shore
TSS1500	chr8:57360585-57360815	Island
TSS1500;TSS1500		
Body	chr10:49674243-49674776	N_Shore
Body	chr11:133800684-133800931	N_Shelf
5'UTR	chr11:12695414-12696981	Island

FIGURE 9C-5

UCSC_RefGene_Group	UCSC_CpG_Islands_Name	Relation_to_UCSC_CpG_Island
1stExon	chr11:66083572-66083782	Island
1stExon	chr11:66083572-66083782	N_Shore
1stExon	chr11:66083572-66083782	Island
TSS200	chr11:9025095-9026315	Island
	chr11:10324353-10324828	N_Shore
	chr12:11653232-11653775	Island
	chr12:81471569-81472119	N_Shelf
Body;Body	chr12:54447744-54448091	S_Shore
TSS1500	chr12:30848264-30849016	S_Shore
	chr14:61108954-61109786	N_Shore
5'UTR	chr15:79576059-79576270	Island
Body	chr15:31775540-31776988	Island
TSS1500;TSS1500;TSS1500	chr15:72522131-72524238	S_Shore

FIGURE 9D-5

UCSC_RefGene_Group	UCSC_CpG_Islands_Name	Relation_to_UCSC_CpG_Island
5'UTR	chr15:33009530-33011696	Island
TSS1500;TSS1500;5'UTR	chr16:30076310-30077872	N_Shore
3'UTR;3'UTR;3'UTR;3'UTR;3'UTR		
1stExon;5'UTR	chr17:28562387-28563186	Island
Body;Body		
5'UTR		
1stExon;Body	chr19:15342626-15343181	Island
Body;Body	chr20:44657463-44659243	Island
TSS1500;TSS1500;TSS1500	chr22:21368197-21368771	S_Shore
TSS200;TSS200;TSS200;TSS200		

FIGURE 9E-5

Phantom	DMR	Enhancer	HMM_Island	Regulatory_Feature_Name
		NA	1:117465578-117466781	1:117663907-117665512
		NA		
	CDMR	TRUE	1:206063625-206063801	
		NA		1:169555452-169556050
		TRUE		1:17336920-17338827
		NA		1:26855689-26857507
		NA	1:28113187-28114165	1:28240552-28241702
		NA		2:97202196-97202767
		NA	2:105381311-105382817	2:106014507-106016259
		TRUE		2:230451331-230452578
		TRUE	2:39040222-39041697	2:39187021-39187940
		NA	2:105381311-105382817	2:106014507-106016259
	DMR	NA	2:131229834-131230653	2:131513688-131513993
		NA	2:105381311-105382817	2:106014507-106016259
		NA		

FIGURE 9A-6

Phantom	DMR	Enhancer	HMM_Island	Regulatory_Feature_Name
	RDMR	TRUE	3:188870246-188870359	
high-CpG:53054999-53055997		NA	3:53053856-53056190	
		TRUE	3:51982751-51985402	3:52007848-52009460
		NA	3:161649892-161650878	3:160166409-160168278
	RDMR	TRUE	3:188870501-188870889	
		NA	4:147777501-147778016	4:147557996-147558356
		NA		4:84255006-84256489
		TRUE		5:139013288-139013559
	RDMR	NA	5:140400003-140400154	
		NA		6:31104495-31106407
	DMR	NA	6:25760360-25760750	6:25652510-25652746
		NA		
		NA		6:147528521-147528685
		NA	6:11151611-11153237	6:11044102-11044892
		TRUE		6:10510346-10511316

FIGURE 9B-6

Phantom	DMR	Enhancer	HMM_Island	Regulatory_Feature_Name
		TRUE	6:18230230-18231229	6:18122473-18123542
		NA		7:131008672-131009115
	DMR	NA	7:156125195-156126707	7:156432754-156434135
		NA	7:130068467-130069793	7:130418325-130419878
		NA	7:130068467-130069793	7:130418325-130419878
		NA	7:152222028-152222744	7:152590901-152592150
		NA	7:130068467-130069793	7:130418325-130419878
	DMR	NA	7:130068467-130069793	
		NA	7:130068467-130069793	7:130418325-130419878
		NA		
		NA	8:57522950-57523369	8:57360377-57362115
		NA		
		NA		
		TRUE	11:133303046-133303162	11:133797084-133799070
		NA	11:12651991-12653557	11:12695339-12696865

FIGURE 9C-6

Phantom	DMR	Enhancer	HMM_Island	Regulatory_Feature_Name
		NA	11:65840060-65841164	
	RDMR	NA	11:65840060-65841164	
		NA	11:65840060-65841164	
	DMR	NA	11:8981699-8983012	
		NA		
		NA	12:11544500-11545229	12:11653353-11654101
		NA		
		NA	12:52734084-52734533	12:54447856-54448358
		NA	12:30739425-30740382	
		NA		
	RDMR	TRUE	14:60177929-60179820	
	RDMR	NA	15:77363046-77363443	
		NA	15:29562601-29564280	
		NA		

FIGURE 9D-6

Phantom	DMR	Enhancer	HMM_Island	Regulatory_Feature_Name
		NA	15:30796823-30799072	
		NA		
		TRUE		
		NA	17:59752099-59752445	17:62397669-62399390
		NA		17:40176771-40177869
	DMR	NA	17:25586344-25587312	17:28562266-28563419
		NA	17:53018213-53018286	
	DMR	NA		18:66388995-66389733
	DMR	NA	19:15203635-15204238	19:15341951-15343455
		NA	20:44090882-44092713	20:44657985-44658436
		TRUE		22:21368080-21369261
		NA		22:36648819-36649508

FIGURE 9E-6

Regulatory_Feature_Group	DHS	RANGE_START	RANGE_END	RANGE_GB	SPOT_ID
Promoter_Associated	NA	117665053	117665176	NC_000001.10	
	NA	236557682	236557805	NC_000001.10	
	TRUE	207997020	207997143	NC_000001.10	
Unclassified_Cell_type_specific	NA	169556022	169556145	NC_000001.10	
Promoter_Associated	NA	17338766	17338889	NC_000001.10	
Promoter_Associated	NA	26855765	26855888	NC_000001.10	
NonGene_Associated	NA	28241577	28241700	NC_000001.10	
Promoter_Associated	NA	97202260	97202383	NC_000002.11	
Unclassified_Cell_type_specific	TRUE	106015739	106015862	NC_000002.11	
Unclassified_Cell_type_specific	NA	230452311	230452434	NC_000002.11	
Unclassified	TRUE	39187539	39187662	NC_000002.11	
Unclassified_Cell_type_specific	TRUE	106015767	106015890	NC_000002.11	
Unclassified_Cell_type_specific	TRUE	131513927	131514050	NC_000002.11	
Unclassified_Cell_type_specific	TRUE	106015771	106015894	NC_000002.11	
	NA	30561970	30562093	NC_000002.11	

FIGURE 9A-7

Regulatory_Feature_Group	DHS	RANGE_START	RANGE_END	RANGE_GB	SPOT_ID
	NA	187387650	187387773	NC_000003.11	
	NA	53080440	53080563	NC_000003.11	
Promoter_Associated	NA	52008487	52008610	NC_000003.11	
Promoter_Associated	NA	160167977	160168100	NC_000003.11	
	NA	187387866	187387989	NC_000003.11	
Unclassified_Cell_type_specific	TRUE	147558239	147558362	NC_000004.11	
Unclassified_Cell_type_specific	NA	84255411	84255534	NC_000004.11	
Unclassified	NA	139013317	139013440	NC_000005.9	
	NA	140419819	140419942	NC_000005.9	
Unclassified	NA	31105711	31105834	NC_000006.11	
Unclassified_Cell_type_specific	TRUE	25652602	25652725	NC_000006.11	
	NA	30295466	30295589	NC_000006.11	
Unclassified_Cell_type_specific	NA	147528521	147528644	NC_000006.11	
Unclassified_Cell_type_specific	TRUE	11044877	11045000	NC_000006.11	
Unclassified	NA	10510789	10510912	NC_000006.11	

FIGURE 9B-7

Regulatory_Feature_Group	DHS	RANGE_START	RANGE_END	RANGE_GB	SPOT_ID
Promoter_Associated	TRUE	18122719	18122842	NC_000006.11	
Unclassified_Cell_type_specific	NA	131008743	131008866	NC_000007.13	
Unclassified	NA	156433108	156433231	NC_000007.13	
Unclassified	TRUE	130419159	130419282	NC_000007.13	
Unclassified	TRUE	130419133	130419256	NC_000007.13	
Unclassified	TRUE	152591698	152591821	NC_000007.13	
Unclassified	TRUE	130419116	130419239	NC_000007.13	
	TRUE	130418324	130418447	NC_000007.13	
Unclassified	TRUE	130419173	130419296	NC_000007.13	
	NA	82644012	82644135	NC_000008.10	
Unclassified_Cell_type_specific	TRUE	57360613	57360736	NC_000008.10	
	NA	98062687	98062810	NC_000010.10	
	NA	49673534	49673657	NC_000010.10	
Promoter_Associated	NA	133797911	133798034	NC_000011.9	
Unclassified_Cell_type_specific	NA	12696758	12696881	NC_000011.9	

FIGURE 9C-7

Regulatory_Feature_Group	DHS	RANGE_START	RANGE_END	RANGE_GB	SPOT_ID
	NA	66083697	66083820	NC_000011.9	
	NA	66083541	66083664	NC_000011.9	
	NA	66083573	66083696	NC_000011.9	
	TRUE	9025767	9025890	NC_000011.9	
	NA	10323902	10324025	NC_000011.9	
Unclassified	TRUE	11653486	11653609	NC_000012.11	
	NA	81468232	81468355	NC_000012.11	
Unclassified	TRUE	54448265	54448388	NC_000012.11	
	NA	30849114	30849237	NC_000012.11	
	NA	40666907	40667030	NC_000013.10	
	NA	61108227	61108350	NC_000014.8	
	TRUE	79576060	79576183	NC_000015.9	
	NA	31775895	31776018	NC_000015.9	
	NA	72525012	72525135	NC_000015.9	

FIGURE 9D-7

Regulatory_Feature_Group	DHS	RANGE_START	RANGE_END	RANGE_GB	SPOT_ID
	TRUE	33010536	33010659	NC_000015.9	
	NA	30075192	30075315	NC_000016.9	
	NA	85429035	85429158	NC_000016.9	
Unclassified	TRUE	62398693	62398816	NC_000017.10	
Gene_Associated	NA	40177415	40177538	NC_000017.10	
Unclassified_Cell_type_specific	TRUE	28562685	28562808	NC_000017.10	
	NA	55663225	55663348	NC_000017.10	
Unclassified_Cell_type_specific	NA	66389420	66389543	NC_000018.9	
Unclassified_Cell_type_specific	NA	15342982	15343105	NC_000019.9	
Unclassified_Cell_type_specific	TRUE	44658225	44658348	NC_000020.10	
Promoter_Associated	NA	21369010	21369133	NC_000022.10	
Unclassified	TRUE	36648973	36649096	NC_000022.10	

FIGURE 9E-7

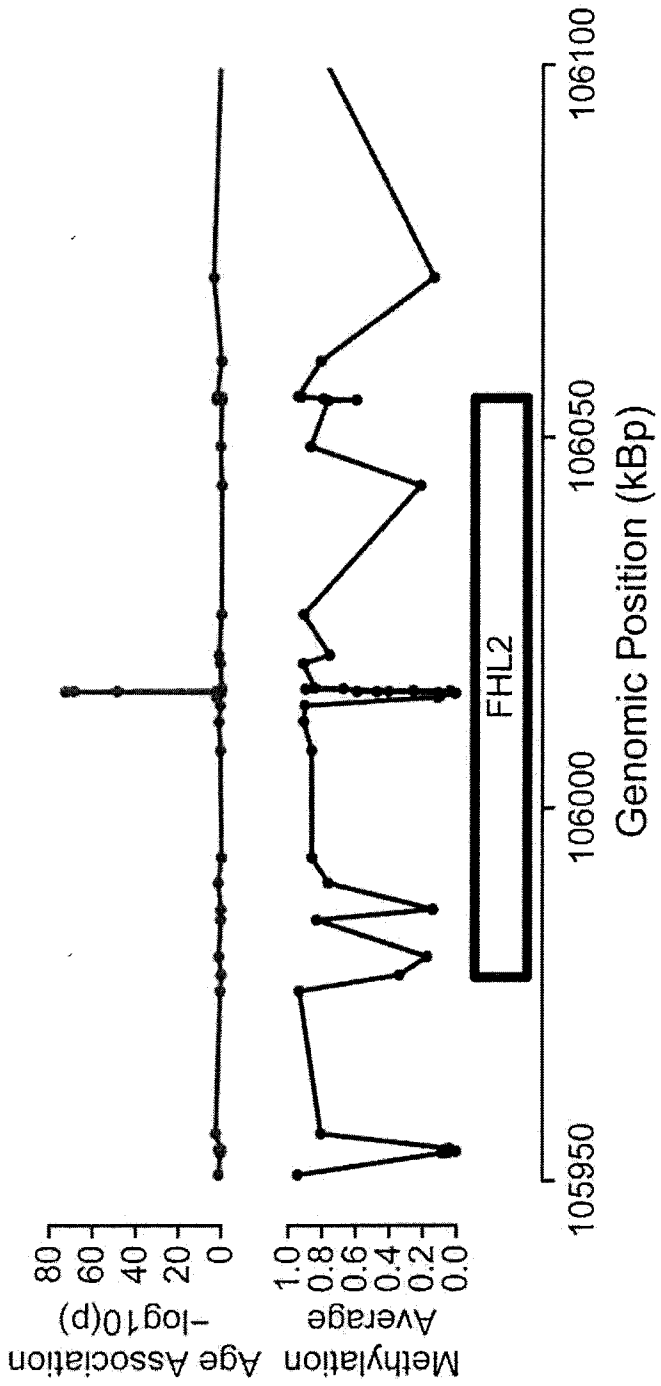

Age Association
−log10(p)
80
60
40
20
0
Methylation
Average
1.0
0.8
0.6
0.4
0.2
0.0
FHL2
105950
106000
106050
106100
Genomic Position (kBp)

FIGURE S1

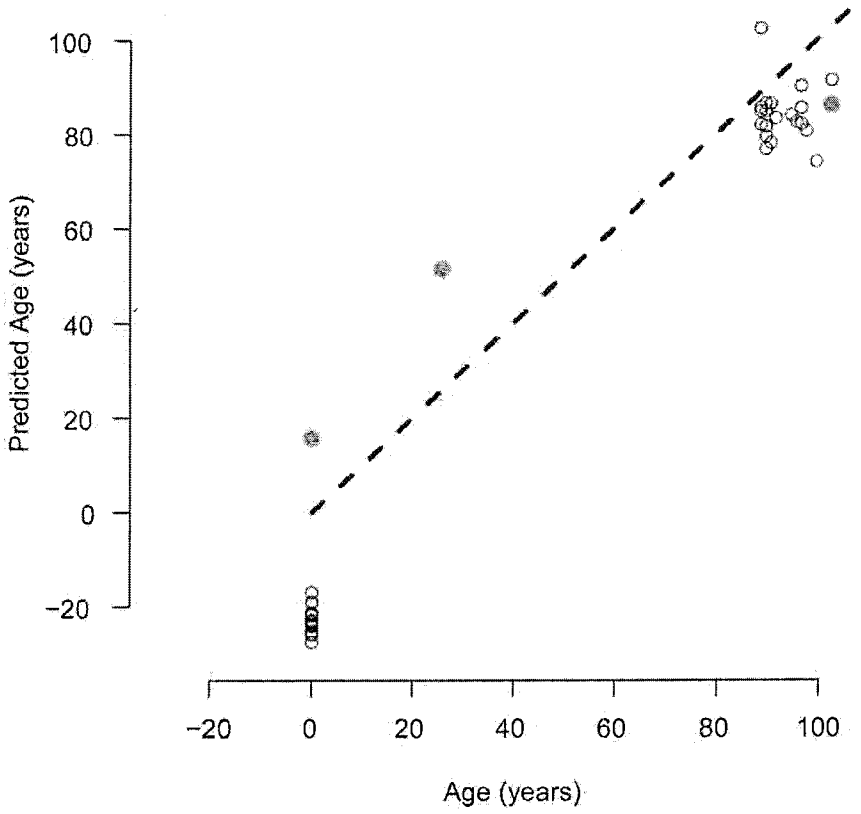

Predicted Age (years)
100
80
60
40
20
0
-20
-20
0
20
40
60
80
100
Age (years)

FIGURE S2

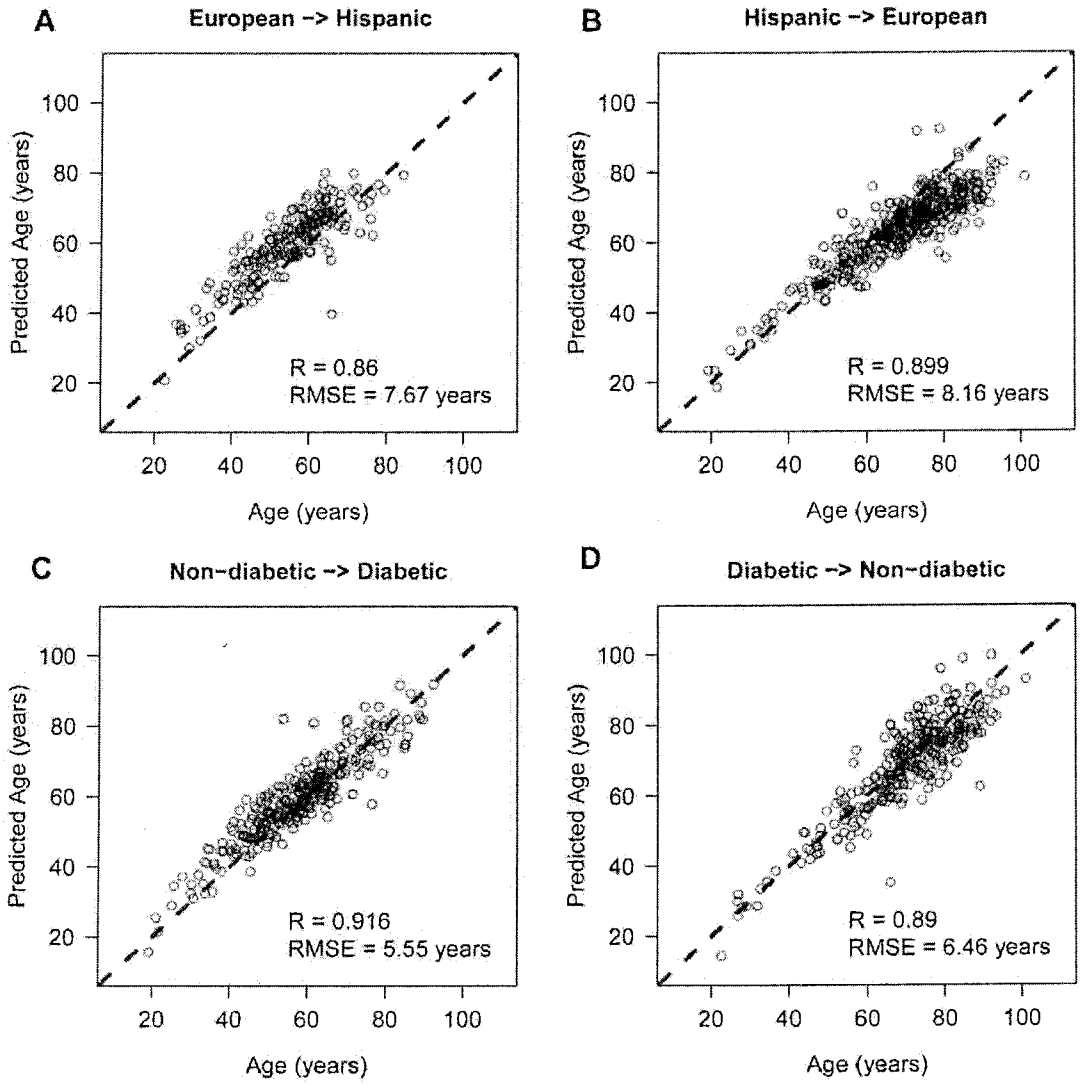

A
European -> Hispanic
Predicted Age (years)
Age (years)
R = 0.86
RMSE = 7.67 years
B
Hispanic -> European
Predicted Age (years)
Age (years)
R = 0.899
RMSE = 8.16 years
C
Non-diabetic -> Diabetic
Predicted Age (years)
Age (years)
R = 0.916
RMSE = 5.55 years
D
Diabetic -> Non-diabetic
Predicted Age (years)
Age (years)
R = 0.89
RMSE = 6.46 years

FIGURE S3

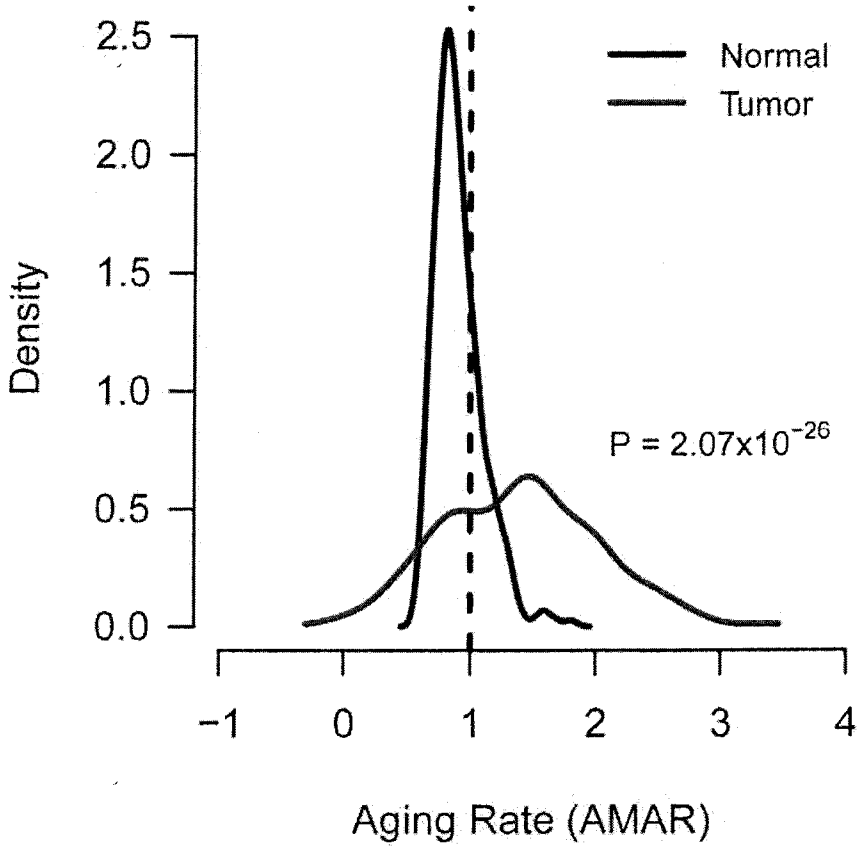

2.5
2.0
1.5
1.0
0.5
0.0
Density
Normal
Tumor
P = 2.07x10^-26
−1
0
1
2
3
4
Aging Rate (AMAR)

FIGURE S4

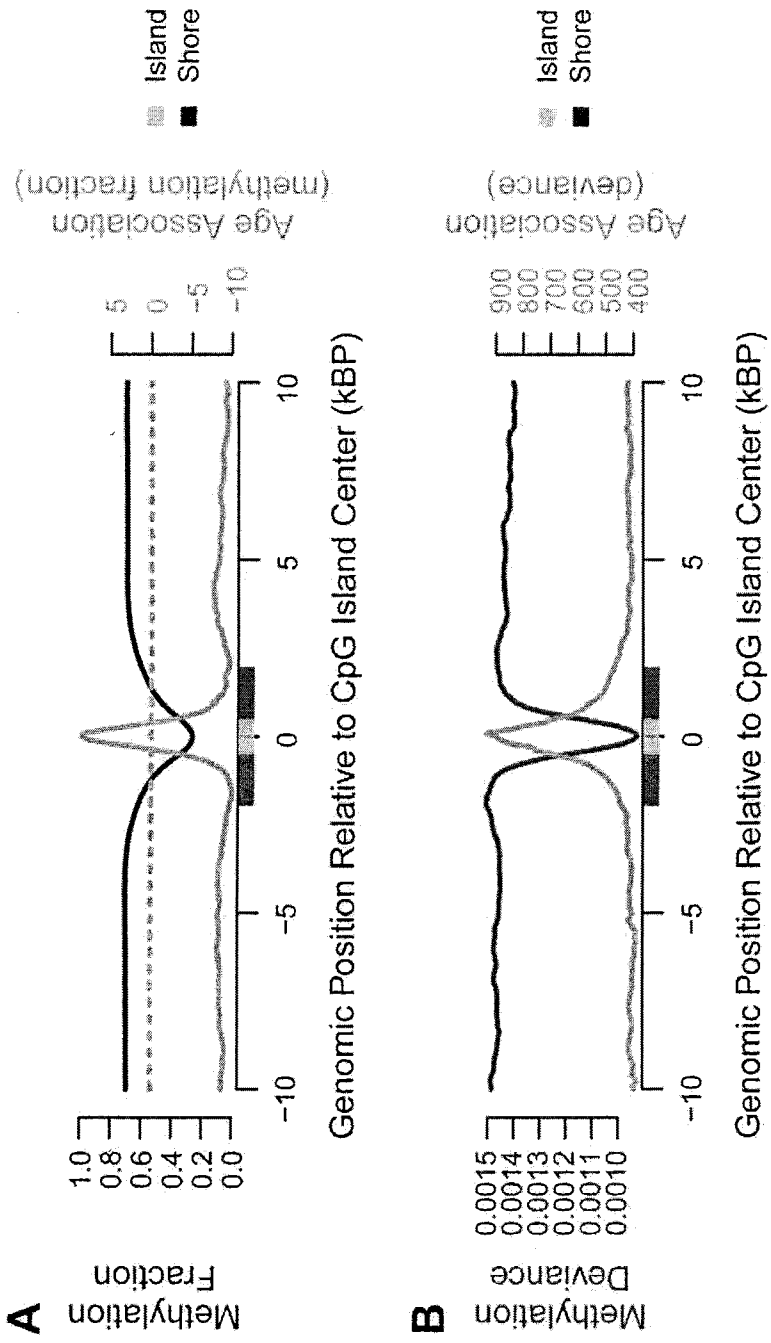

A
Methylation Fraction
1.0
0.8
0.6
0.4
0.2
0.0
−10
−5
0
5
10
Genomic Position Relative to CpG Island Center (kBP)
Age Association (methylation fraction)
5
0
−5
−10
Island
Shore
B
Methylation Deviance
0.0015
0.0014
0.0013
0.0012
0.0011
0.0010
−10
−5
0
5
10
Genomic Position Relative to CpG Island Center (kBP)
Age Association (deviance)
900
800
700
600
500
400
Island
Shore

FIGURE S5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2013/069710

A. CLASSIFICATION OF SUBJECT MATTER

C12Q l/68(2006.01)i, C12N 15/ll(2006.01)i, GOIN 33/48(2006.01)i, GOIN 33/53(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q 1/68; GOIN 33/50; C12N 15/1 1; GOIN 33/48; GOIN 33/53

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean utility models and applications for utility models
Japanese utility models and applications for utility models

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
eKOMPASS(KIPO internal) & keywords: age, methylation, epigenome, marker, methylomic aging rate

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GARAGNANI et al., `Methylation of EL0VL2 gene as a new epigenetic marker of age" Aging Cell, Vol.11, No. 6, pp. 1132-1134 (14 October 2012) See abstract and page 1113.	1-3,6-15,18,30-56 ,58,60-77
Y		4,5,16
A		17,19-22,57,59 ,78-80
Y	BOCKLANDT et al., `Epigenetic predictor of age` PLoS One, Vol.6, Issue 6, Article No. el4821, pp. 1-6 (22 June 2011) See abstract, page 3 and figure 2.	4,5,16
Y	HEYN et al., "Distinct DNA methylomes of newborns and centenarians" PNAS, Vol.109, No. 26, pp. 10522-10527 (11 June 2012) See abstract and page 10525.	4,5,16
A	US 2004-0132026 Al (OLEK) 8 July 2004 See abstract and claims 1-3.	1-22,30-80

☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
07 February 2014 (07.02.2014)	**07 February 2014 (07.02.2014)**

Name and mailing address of the ISA/KR
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701, Republic of Korea
Facsimile No. +82-42-472-7140

Authorized officer
KΓM, Seung Beom
Telephone No. +82-42-481-3371

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2013/069710

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	FLORATH et al., `Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis reveal ing over 60 novel age-associated CpG sites` Human Molecular Genetics, Advance Access, pp.1-16 (26 October 2013) See abstract, page 5 and table 2.	1-22, 30-58, 60-80
PX	HANNUM et al., "Genome-wide methylation profiles reveal quantitative views of human aging rates` Molecular Cell, Vol.49, No.2, pp.359-367 (21 November 2012) See abstract and pages 360, 361, 363.	1-3, 6-15, 18, 30-58, 60-77

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2013/069710

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 23-29
because they relate to subject matter not required to be searched by this Authority, namely:
Claims 23-29 pertain to diagnostic methods, and thus relate to a subject matter which this International Searching Authority is not required, under Article 17(2)(a)(i) of the PCT and Rule 39.1(iv) of the Regulations under the PCT, to search.

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. Ill Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required addtional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of any additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/US2013/069710

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2004-0132026 Al	08/07/2004	AT 330037 T	15/07/2006
		AU 2002-954701 A	13/02/2002
		CA 2423849 Al	27/03/2003
		DE 10038733 Al	21/02/2002
		DE 50110172 Dl	27/07/2006
		EP 1320632 A2	25/06/2003
		EP 1320632 Bl	14/06/2006
		JP 2004-504855 A	19/02/2004
		WO 02-10445 A2	07/02/2002
		wo 02-10445 A3	05/12/2002

专利名称(译)	预测年龄和鉴定诱导或抑制早衰的药剂的方法		
公开(公告)号	EP2917371A1	公开(公告)日	2015-09-16
申请号	EP2013852408	申请日	2013-11-12
[标]申请(专利权)人(译)	加利福尼亚大学董事会		
申请(专利权)人(译)	加利福尼亚大学董事会 SAGE生物网络		
当前申请(专利权)人(译)	加利福尼亚大学董事会 SAGE生物网络		
[标]发明人	ZHANG KANG HANNUM GREGORY IDEKER TREY GUINNEY JUSTIN FRIEND STEPHEN H		
发明人	ZHANG, KANG HANNUM, GREGORY IDEKER, TREY GUINNEY, JUSTIN FRIEND, STEPHEN, H.		
IPC分类号	C12Q1/68 C12N15/11 G01N33/48 G01N33/53		
CPC分类号	C12Q1/6883 C12Q1/6881 C12Q2600/118 C12Q2600/136 C12Q2600/154 C12Q2600/16 G01N33/5308 G01N33/57484 G01N2500/00		
优先权	61/724528 2012-11-09 US		
其他公开文献	EP2917371A4		
外部链接	Espacenet		

摘要(译)

本发明提供了基于受试者的表观基因组预测受试者年龄的方法。