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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 October 2017 (12.10.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/176834 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2017/026061

(22) International Filing Date:
5 April 2017 (05.04.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/318,879 6 April 2016 (06.04.2016) US

(71) Applicant: **UNIVERSITY OF FLORIDA RESEARCH FOUNDATION, INCORPORATED** [US/US]; 223 Grinter Hall, Gainesville, FL 32611 (US).

(72) Inventor: **OKUNIEFF, Paul, Gerson**; 4010 S.W. 93rd Drive, Gainesville, FL 32608 (US).

(74) Agents: **EISENSCHENK, Frank, C.** et al.; Saliwanchik, Lloyd & Eisenschenk, P.O. Box 142950, Gainesville, FL 32614-2950 (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, DJ, 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, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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.

(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, ST, 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).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

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

— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: MEASUREMENT OF GENOMIC AGE FOR PREDICTING THE RISK OF CANCER

(57) Abstract: The subject invention pertains to materials and methods of determining accumulated mutations and the rate of mutations in a target genomic sequence which is a part of a short interspersed element (SINE). The assay utilizes a combination of a target sequence clamp and digital PCR (dPCR). The target sequence clamp prevents PCR amplification of the wild- type target sequence and permits PCR amplification of only the mutated target sequence. The dPCR provides the number of mutated target sequences per genome, which can be used to determine the rate of mutations in the target sequence, the accumulated mutations in the genome and/or the rate of mutations in the genome. The accumulated mutations and the rate of mutations in the target sequence and/or the genome can be used to determine the genomic age and/or the risk of cancer of a subject. A kit for performing the assay is also provided.

WO 2017/176834 A2

DESCRIPTION

MEASUREMENT OF GENOMIC AGE FOR PREDICTING THE RISK OF CANCER

CROSS-REFERENCE TO RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application Serial No. 62/318,879, filed April 6, 2016, the disclosure of which is hereby incorporated by reference in its entirety, including all figures, tables and amino acid or nucleic acid sequences.

The Sequence Listing for this application is labeled "Seq-List.txt" which was created on April 4, 2017 and is 124 KB. The entire contents of the sequence listing are incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

Due to a large number silent mutations and the absence of tools to detect them, the measurement of genome-wide chromosomal DNA abnormalities is not routine. The current next-generation sequencing (NGS) methods make inherent sequencing errors. The sequencing errors can be partially alleviated by increasing the number of runs and improving the purity of the sample; however, even deep sequencing methods suffer from false detection rates. Detection of genome-wide variance is rarely better than 1%. This is far from the 10^6 sensitivity expected to be required for detecting the rate of silent mutations accumulating with each cell division or after an exposure to a mutagen, for example, a low-dose particle exposure. Therefore, the current NGS methods are insufficient to estimate genome-wide accumulated mutations and/or the rate of mutations, for example, point mutations and insertion/deletion (indel) mutations.

The premalignant genome-wide accumulated mutations and/or the rate of mutations usually cause few or no phenotypic effects but stochastically (randomly but extremely rarely) lead to driver mutations which cause cancer. The driver mutations accelerate the oncogenic mutational process and lead to evolutionary selection of more drivers in cancer-causing genes. The genome-wide accumulated mutations and/or the rate of mutations are dominated by point mutations (~95%) and short indel mutations (~1-3%) and include a small component of translocations when high-linear energy transfer (LET) radiation is considered (~1%). Driver mutations are far too complex to measure and interpret and are too infrequent to detect in small tissue volumes.

BRIEF SUMMARY OF THE INVENTION

The invention relates to materials and methods of determining accumulated mutations and the rate of mutations in a target genomic sequence (also referred to herein as "a target sequence"), particularly a target sequence which is a part of a short interspersed element (SINE), a long interspersed element (LINE), any highly repeated sequence in a cell's genome, and/or the mitochondrial genome. Thus, the target sequence is present in a large number of copies per genome as SINEs, LINEs, or other highly repeated sequences within the genome of the cell, or the mitochondrial genome which is highly repeated in each cell in each mitochondrion. Because the entire mitochondrial genome is highly repeated due to the number of genome copies in each mitochondria and the number of mitochondria in each cell, the mitochondrial genome, in addition to a SINE and/or LINE, can serve as a target sequence in one embodiment of the invention. The assay to determine the number of accumulated mutations in a target sequence utilizes a combination of a target sequence clamp with digital PCR (dPCR). The target sequence clamp binds only to the wild-type target sequence, prevents PCR amplification of only the amplicons that have the wild-type target sequence and permits PCR amplification of only the amplicons that have the mutated target sequence. The dPCR, where the sample is separated into a large number of partitions, detects the presence of the DNA fragments containing the mutant target sequences in the large number of analyzed genomic DNA fragments. The PCR amplification in a partition of the sample indicates the presence of the mutant target sequence in that partition. The accumulated mutations in the target sequence can be calculated based on the number of fragments of the genomic DNA that arise from one genome and the number of fragments of the genomic DNA per genome that contain the mutated target sequence.

The accumulated mutations in a target sequence can be used to determine the rate of mutations in the target sequence, the accumulated mutations in the genome (genome-wide mutations) and the rate of mutations in the genome (genome-wide rate of mutations).

Accumulated mutations and the rate of mutations in the genome are directly proportional to genomic age and the risk of cancer. Accordingly, the invention also provides a method of calculating the genomic age and/or the risk of cancer in a subject. In one embodiment, the risk of cancer in a tissue or organ is determined.

Furthermore, the invention provides a kit to carry out the methods of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication, with color drawing(s), will be provided by the Office upon request and payment of the necessary fee.

Figure 1. Natural Aging: baseline accumulation of mutations in a healthy astronaut with no risk factors. Nat+Direct Rad: incremental mutations due to the 3-year voyage. Nat+Dir+RMD+RRI: impact on the rates of accumulation of mutations due to radiation-induced mitochondrial damage (RMD) and radiation-related inflammation (RRI). Cancer distributions are based on reported levels of non-synonymous mutations in these types of tumors.

Figure 2. Genomic age versus chronological age.

Figure 3. Computing genomic age and rate of aging.

Figure 4. Age is the dominant risk factor for cancer, outpacing smoking and PSA. Individuals under 20 years old have a 130-fold reduced risk of cancer compared to individuals over 75 years old.

Figure 5. The passenger mutation (also known as silent or non-synonymous) rate increases with age, allowing for random cells to become cancerous. Different tissues can tolerate different numbers of mutations before cancers become common.

Figure 6. QClamp™ utilizes a sequence-specific and wild-type template xenonucleotide "clamp" (XNA) that suppresses PCR amplification of only the wild-type template DNA and allows selective PCR amplification of only the mutant template DNA. This allows the detection of mutant DNA in the presence of a large excess of wild-type templates from any sample, including FFPE tissues and whole blood.

Figure 7. dPCR results for QClamp™. Initial results demonstrating detection of 1 mutant target in 100,000 copies of wild-type targets is achievable with standard dPCR. Clamp2 was used for this study.

Forward Primer B1 F001 5' CTTTAATCCCAGCACTCGGG-3' (SEQ ID NO: 1).

Reverse Primer B1 R001 5'-CTCTGTAGCCCTGGTGTCCTGG-3' (SEQ ID NO: 2).

Clamp1: GGCTGGCCTCGAACTC (SEQ ID NO: 4); Tm = 79.4°C, 68.8% GC.

Clamp2: GTGTCCTGGAACTCACTCTG (SEQ ID NO: 5); Tm = 78.9°C, 55% GC.

Figure 8. B1 SINE Sequence Choice. The two XNA Clamps and the forward and reverse PCR primers were designed based on the highly conserved sequence shown as double underlined text (SEQ ID NO: 482).

Figure 9. Example calculation of the fraction of B1 SINE sites with conserved CE3, LE2, or [LE2 orCE3] using an XNA multiplex. CE3 and LE2 are sequential (serially placed) in the PCR reading frame. DNA was collected from the livers of male NIH Swiss mice. PCR for B1 SINE was performed using primers alone, and the cycles required for detection in a 50-ng sample were set at 0.00. XNA CE3 was highly conserved among B1 loci at 79%, LE2 was also highly conserved at 86%, and [CE3 or LE2] were conserved on 97%. Alternatively to serial XNA placement, overlaying "allelic" XNA on the same locus can further improve sensitivity to detect mutations at that locus. Likewise, parallel clamps with independent PCR primers can be used to evaluate different chromosomes or DNA repair mechanisms.

Figure 10. Fractions of total B1 SINE SITES with homology to the CE3 and LE2 regions. Potential to measure differential repair in areas with high and low baseline allelic conformity using serial XNA from the previous example. For example, B1 sequences that are highly conforming can provide a template on fidelity of repair. Serial XNA sequences, therefore, can be used to separately evaluate the 68% of the mouse genome that is very highly conforming.

Figure 11. Impact of 1 Gy Total Body Irradiation on Liver B1 Mutation Levels Measured at 2 hr. DNA was collected from the livers of male NIH Swiss mice. PCR for B1 SINE was performed using primers alone, and the cycles required for detection in a 50-ng sample were set at 0.00. Mutations per base assumes 20 bases at risk for mutation for LE2, 16 bases for CE3, and a mouse genome size of 2.9 E9 bases. At 2 hours, DNA repair is known to only partially complete. Liver epithelium is known to repair damage with only a minority of cells undergoing apoptosis in the first 24 hr. Loci convergent with one or another majority loci repaired better than the loci that were divergent from the majority loci as evidenced by the reduced mutations in the CE3&LE2 group following radiation. This is presumed to be due to the availability of the majority template. This phenomenon allows for study of specific DNA repair mechanisms, in a personalized way.

Figure 12. Impact of 9 Gy Total Body Irradiation on Spleen B1 Mutation Levels Measured at 2 hr. DNA was collected from the spleens of male NIH Swiss mice. PCR for B1 SINE was performed using primers alone, and the cycles required for detection in a 50-ng sample were set at 0.00. Mutations per base assumes 20 bases at risk for mutation for LE2, 16 bases for CE3, and a mouse genome size of 2.9 E9 bases. It is known that at 2 hr, little DNA repair occurs in lymphocytes; instead, they responded to 9 Gy at ≈4-6 hr with simultaneous apoptosis. Thus, the highly damaged lymphocytes were still present at 2 hr.

Figure 13. Different organs from the same animal exhibited different frequencies and rates of mutation repair. DNA was collected from various organs of male NIH Swiss mice at 2 hr or 6 days after 1 or 9 Gy irradiation. Persistent mutation levels were intermediate in frequency and increased with dose for the liver. The mutations maintained for the 6-day endpoint for this slowly proliferative and low-apoptotic epithelial tissue. The spleen cleared mutations at day 6, presumably through lymphocyte apoptosis. The brain was resistant to mutations at any time point, and the small bowel, a rapidly proliferating tissue, had increasing mutation frequency with time as expected from silent mutations.

Figures 14A and 14B. Brain tissue of male NIH Swiss mice were resistant to mutations measured both early after exposure (2h – Fig. 14A) and at later times (6d – Fig. 14B). This was true for low doses (1 Gy) and high doses (9 Gy).

Figures 15A-15C. Brain tissue of male and female mice of radiation sensitive (BALB/c) mice (Fig. 15B) and radiation-resistant (C57BL/6) mice (Fig. 15A) were resistant to mutations in the brain measured at 24 hours. As seen with the serial XNA, the more resistant C57BL/6 strain appeared to use the majority allele to "repair" some mutated sequences (Fig. 15C). The resistance to mutations was seen in both sexes. Thus we can detect innate ability to repair DNA mutations.

Figure 16. PCR primers were designed to flank highly conserved regions on the human LINE1. In each case, human genomic DNA was used (50 ng). Many produced fairly homogenous products and extremely high copy numbers, some detected in as little as 5-7 cycles. Primer set 6059 produced both homogeneous and plentiful product. Primer 279 also produced a homogeneous product. Both are excellent options for LINE1 regions for measuring genotoxicity. Examples for 6059 will be shown and primers for the detection of LINE1 are provided in Figure 20.

Figure 17. LINE1 abundance in 3 human gDNA sources (Ken, HFL1, HEK293) using primer set 6059. All had high abundance in a 50-ng gDNA sample, and all produced a reliable melting curve, indicating a homogeneous product.

Figure 18. Shown are the 6059 LINE1 target sequence (SEQ ID NO: 483), the sections used to design PCR proprimerbes (in double underlining) and XNAs (single underlining) and the sequencing data from all 3 human DNA sources. The single underlined segment was 100% conserved, as determined by using deep gene sequencing and can be used to design XNA clamps. Deep gene sequencing is sensitive to about 1% per base. Several XNAs have can and have been developed from this lengthy and highly conserved sequence.

Figure 19. Mitochondrial clamps. There are many suitable clamps, as the mitochondrial genome is well conserved and inherited almost exclusively from the mother. An example of a potential XNA from mouse Cyt A is shown. The XNA (SEQ ID NO: 484) and primer sequences (SEQ ID NOs: 485 and 486) are free from similar genomic sequences.

Figure 20. PCR primers used for the amplification of LINE1.

Figures 21-23. Alignments of LINE1 sequences for the identification of clamp and primers identified in Figures 16 and 20. Continuous stars indicate conserved region suitable for clamp and primer design used in the experiments shown in Figure 16 (sequence 6059; Figure 21, sequence 279, Figure 22, and sequence 139, Figure 23).

Figure 24. Line1 sequence for designing probes (double underlined) and clamps (single underlined sequence) (SEQ ID NO: 501). PCR primers are shown in bold.

BRIEF DESCRIPTION OF THE SEQUENCES

SEQ ID NO: 1: Forward Primer B1.

SEQ ID NO: 2: Reverse Primer B1 R001.

SEQ ID NO: 3: Clamp1.

SEQ ID NO: 4: Clamp2.

SEQ ID NO: 5: Clamp3.

SEQ ID NO: 6: Sequence of mouse B1.

SEQ ID NOs: 7-10: Example of clamp sequences for Alu SINEs.

SEQ ID NOs: 11-98: Examples of target amplicons in Alu SINEs.

SEQ ID NOs: 99-274: Forward and reverse primers for clamp sequence of SEQ ID NOs: 7-10 and Alu SINE sequences of 11-98.

SEQ ID NOs: 275-316: Sequence of SINEs indicated in Table 1.

SEQ ID NOs: 317-416: Examples of frequent sequences in human genomes that can be used as clamp sequences.

SEQ ID NOs: 417-478: Examples of Alu sequences in humans.

SEQ ID NO: 479: An example of a forward primer for B1 having the sequence of SEQ ID NO: 281.

SEQ ID NO: 480: An example of a reverse primer for B1 having the sequence of SEQ ID NO: 281.

SEQ ID NO: 481: An example of a clamp for B1 having the sequence of SEQ ID NO: 281.

DETAILED DISCLOSURE OF THE INVENTION

As used herein, the singular forms "a", "an" and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. Furthermore, to the extent that the terms "including", "includes", "having", "has", "with", or variants thereof are used in either the detailed description and/or the claims, such terms are intended to be inclusive in a manner similar to the term "comprising".

The term "about" or "approximately" means within an acceptable error range for the particular value as determined by one of ordinary skill in the art, which will depend in part on how the value is measured or determined, *i.e.*, the limitations of the measurement system. For example, "about" can mean within 1 or more than 1 standard deviation, per the practice in the art. Alternatively, "about" can mean a range of up to 0 to 20%, 0 to 10%, 0 to 5%, or up to 1% of a given value. Alternatively, particularly with respect to biological systems or processes, the term can mean within an order of magnitude, preferably within 5-fold, and more preferably within 2-fold, of a value. Where particular values are described in the application and claims, unless otherwise stated, the term "about" means within an acceptable error range for the particular value. In the context of compositions containing amounts of ingredients where the terms "about" or "approximately" are used, these compositions contain the stated amount of the ingredient with a variation (error range) of 0 to 10% around the value (X ± 10%).

In the present disclosure, ranges are stated in shorthand to avoid having to set out at length and describe each and every value within the range. Any appropriate value within the range can be selected, where appropriate, as the upper value, lower value, or the terminus of the range. For example, a range of 0.1-1.0 represents the terminal values of 0.1 and 1.0, as well as the intermediate values of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and all intermediate ranges encompassed within 0.1-1.0, such as 0.2-0.5, 0.2-0.8, 0.7-1.0, *etc*. Values having at least two significant digits within a range are envisioned, for example, a range of 5-10 indicates all the values between 5.0 and 10.0 as well as between 5.00 and 10.00, including the terminal values.

When ranges are used, such as for length of a SINE or a LINE or target sequence within a genome, primer or target sequence clamp, combinations and subcombinations of ranges (*e.g.*, subranges within the disclosed ranges) and specific embodiments therein are intended to be explicitly included.

As used herein, the term "cancer" refers to the presence of cells possessing abnormal growth characteristics, such as uncontrolled proliferation, immortality, metastatic potential, rapid growth and proliferation rate, perturbed oncogenic signaling, and certain characteristic morphological features. This includes but is not limited to the growth of: (1) benign or malignant cells (*e.g.*, tumor cells) that correlate with overexpression of a serine/threonine kinase; or (2) benign or malignant cells (*e.g.*, tumor cells) that correlate with abnormally high levels of serine/threonine kinase activity or lipid kinase activity. Non-limiting serine/threonine kinases implicated in cancer include but are not limited to PI-3K mTOR and AKT. Exemplary lipid kinases include but are not limited to PI3 kinases such as PBKα, PBKβ, PBKδ, and PBKγ.

"Subject" refers to an animal, such as a mammal, for example a human. The methods described herein can be useful in both humans and non-human animals. In some embodiments, the subject is a mammal, and in some embodiments, the subject is human. The invention can be used in a subject selected from non-limiting examples of a human, non-human primate, rat, mouse, pig, dog or cat. Additional embodiments of the animals in which the invention can be practiced are well-known to a person of ordinary skill in the art and such embodiments are within the purview of the invention.

Where the term "and/or" is used within the application, it is intended that the elements recited within the phrase where the term "and/or" is used can be assess individually or in any combination of the recited elements. For example, the phrase "target SINE, target LINE and/or the genome" is meant to convey that each element of the phrase can be assessed individually or in any possible combination (e.g., target SINE alone, target LINE alone, genome alone, target SINE and target LINE in combination, target SINE and genome in combination, target LINE and genome in combination, or target SINE, target line, and genome in combination).

The invention relates to materials and methods of determining accumulated mutations and the rate of mutations in a target genomic sequence (also referred to herein as "a target sequence"), particularly a target sequence which is a part of a short interspersed element (SINE), a long interspersed element (LINE), any highly repeated sequence in a cell's genome, and/or the mitochondrial genome. Thus, the target sequence is present in a large number of copies per genome as SINEs, LINEs, or other highly repeated sequences within the genome of the cell. Thus, the invention provides an assay to determine the accumulated mutations and/or the rate of mutations in a target sequence within a short interspersed

element (SINE), long interspersed elements (LINEs), mitochondrial genome, and/or the genome (as used herein, the target sequence is any highly repeated sequence within the genome of a cell). The accumulated mutations and/or the rate of mutations in a target SINE, target LINE and/or the genome integrate various causes of DNA damage. As used herein, the term "genome" refers to highly repeated sequences within the genome of a cell, such as mitochondrial genomes. "Highly repeated sequences" are a nucleotide sequence that is repeated hundreds to thousands of times within the genome of a cell. As such, the invention provides materials and methods for quantitative estimation of mutations and rate of mutations. The invention also provides an assay to measure point mutations and indels (which together comprise > 95% of all mutations) in a target sequence within a cell a target SINE and/or target LINE for example, Alu in humans or B1 in mice or a target LINE, such as those provided in Tables 6-7 (which provides GenBank Accession numbers for partial LINE1 sequences and full length LINE1 sequences for humans and mice, each of which is hereby incorporated by reference in their entirety). The accumulated mutations and/or the rate of mutations in a target sequence within a target SINE or genome of the cell can be extrapolated to measure point mutations and indels in the genome.

The current approaches to estimating cancer risk require screening many animals for long periods of time. These methods generally feature genetically defined animals with driver mutations that cause cancers and incidence rates that are not representative of spontaneously occurring human cancers. The invention provides cancer risk assessments that can be performed on a subject-by-subject basis and on an organ-by-organ basis, thus allowing for subject-specific and organ-specific estimates of cancer risk. The methods of the invention can also be used to determine the effects on cancer risk of genetic and non-genetic factors, for example, race, family lineage, and environmental factors such as food, lifestyle choices, smoking, *etc*.

The accumulated mutations and/or the rate of mutations in a target sequence within the target SINE, target LINE, and/or the genome is directly proportional to the individual's chronological age. Specifically, an individual with a higher chronological age has more accumulated mutations and/or a higher rate of mutations in a target sequence within a target SINE, target LINE, and/or the genome compared to an individual with a relatively lower chronological age. An embodiment of the invention provides an assay to determine accumulated mutations and/or the rate of mutations in a target sequence in a target SINE, target LINE and/or the genome. The genome-wide accumulated mutations and the rate of

mutations can be used to estimate genomic age of an individual for correlation with the individual's chronological age.

For the purpose of the invention, the phrase "chronological age" is the age of a subject based on the subject's date of birth. Accordingly, compared to a chronologically younger subject, a chronologically older subject has a date of birth earlier in time.

Age-related DNA damage is random, is not confined to coding regions and increases with age. Therefore, the number of mutations in certain tumors is directly proportional to the age of the patient (Fig. 4). Ninety-five percent (95%) of these premalignant mutations (passengers) are point mutations (*e.g.*, C>G), and many of the remaining mutations are short indels (*e.g.*, CTT>CT or CT>CTT). The inter-individual rate of mutations, which indicates different genomic age between individuals, is determined by a host of hereditary, acquired, and environmental factors, including radiation exposure.

While almost all mutations are "passenger" mutations, *i.e.*, silent mutations, occasionally a "driver" mutation occurs. Driver mutations confer a competitive advantage upon the reproduction of the affected cell. Thus, while it would be difficult to detect cells with a driver mutation without complete organs to study and without unrealistic databases and sequencing resources, passenger mutation frequencies can be used to mathematically determine driver mutation frequency and downstream cancer incidence. The process from driver mutation to tumor is typically estimated to be 1 to 15 years, and tumors typically have only a few driver mutations (≤ 10). Notably, 10 times fewer driver mutations are affected by chromosome changes than by point mutations, and high-LET radiation induces both types of mutations.

Since the rate of mutations, and consequently, the chances of the occurrence of driver mutations increase with increasing chronological age, chronological age typically correlates with cancer risk in the general population, *i.e.*, higher chronological age of a subject typically, but not necessarily, indicates higher risk of cancer in the subject. For example, cancer incidence per 100,000 is 17 for ages less than 20, which increases by a factor of 10 to 157 for ages 20-49, increases another 5-fold for ages 50 to 64, and further increases another 3-fold for ages over 75 (> 2,200/100,000), for a total increase of more than 130-fold (Fig. 3). These fold-changes indicate that increasing age impacts death and incidence of cancer more than smoking.

For the purpose of the invention, the "genomic age" indicates the accumulated mutations and/or the rate of mutations in the genome of a subject (Fig. 2) as they relate to the

average accumulated mutations and/or the average rate of mutations in the genome of a subject of a particular chronological age. For example, if the accumulated mutations and/or the rate of mutations in the genome of a subject 30 years of age is equal to the average accumulated mutations and/or the average rate of mutations in the genome of a subject 40 years of age, then the genomic age of the subject 30 years of age is 40 years. Therefore, a genome having more accumulated mutations and/or a higher rate of mutations is an "older genome" compared to a "younger genome" having relatively fewer accumulated mutations and/or a lower rate of mutations.

The comparison of a subject's chronological age and genomic age can be expressed as genomic age – the chronological age (Δ_{age}). Therefore, a positive Δ_{age} indicates that a subject is aging at a higher rate than average, whereas a negative Δ_{age} indicates that a subject is aging at a lower rate than average

A standard scale for the genomic age for a particular species can be determined based on the average accumulated mutations and/or the average rate of mutations in different groups of individuals of varying ages. As such, the standard scale for the genomic age for the species indicates the average accumulated mutations and/or the average rate of mutations in a target sequence in a target SINE, target LINE, and/or the genomes of individuals belonging to the species at increasing chronological ages.

Accordingly, the invention provides methods for measuring accumulated mutations and/or the rate of mutations in a target sequence in a target SINE, target LINE and/or the genome to determine the genomic age of a subject and, consequently, to determine the cancer risk of the subject. Genomic age and chronological age of a subject can be compared to the known chronological age of the subject and the standard scale of genomic age to identify the subject's risk for cancer and offer enhanced cancer screening if the subject has a higher risk of cancer. Low-risk groups, on the other hand, can be spared from unnecessary screening tests. The accumulated mutations and/or mutation rates can also be used to evaluate the impact of environment (*e.g.*, insecticides), lifestyle changes (*e.g.*, weight loss or smoking cessation), and therapies (*e.g.*, X-rays, medications) on genotoxic load, mutation rate, and, consequently, cancer risk.

Accordingly, an embodiment of the invention provides an assay to determine the number of accumulated mutations in a target sequence within a target SINE and/or target LINE and/or genome of a subject. For the purpose of this invention, this assay is called the clamp/dPCR combination assay.

The clamp/dPCR combination assay comprises the steps of:

a) obtaining a genomic DNA sample from the subject and fragmenting the genomic DNA sample, or obtaining a fragmented genomic DNA sample from the subject,

b) mixing a predetermined number of fragments of the genomic DNA that arise from a predetermined number of genomes with a reagent mixture to produce a reaction mixture, the reagent mixture comprising:

i) a pair of polymerase chain reaction primers that amplify a target amplicon comprising the target sequence within the target SINE,

ii) a target sequence clamp which binds only to the wild-type target sequence within the SINE, wherein the target sequence clamp prevents the PCR amplification of only those target amplicons that have the target wild-type sequence within the SINE and permits the PCR amplification of only those target amplicons that have the target mutated sequence within the SINE, and

iii) a DNA polymerase enzyme and the reactants for a digital PCR (dPCR),

c) subjecting the reaction mixture to the dPCR,

d) identifying the number of fragments of the genomic DNA comprising the target amplicon having the target mutated sequence within the SINE based on the number of positive PCR amplifications in the dPCR,

e) calculating the number of accumulated mutations per genome in the target sequence within the target SINE and/or target LINE and/or genome based on the number of fragments of the genomic DNA that arise from one genome and the number of fragments of the genomic DNA per genome that comprise the target amplicons having the target mutated sequence within the target SINE and/or target LINE and/or genome wherein the presence of the target mutated sequence within the target SINE and/or target LINE and/or genome is indicated by the positive PCR amplification in the dPCR.

LINEs are transposons that are about 5-6 kb long, contain an internal polymerase II promoter and encode two open reading frames (ORFs). Upon translation, a LINE RNA assembles and moves to the nucleus, where an endonuclease activity makes a single-stranded nick and the reverse transcriptase uses the nicked DNA to prime reverse transcription from the 3′ end of the LINE RNA. Reverse transcription frequently fails to proceed to the 5′ end, resulting in many truncated, nonfunctional insertions. Most LINE-derived repeats are short, with an average size of 900 bp for all LINE1 copies, and a median size of 1,070 bp for copies of the currently active LINE1 element. Three distantly related LINE families are found in the

human genome: LINE1, LINE2 and LINE3, with LINE1 being the only remaining active LINE. Exemplary target LINE1 sequences are provided in Tables 6-7, which provide both partial and full length LINE1 sequences for humans and mice, identified by GenBank accession number. Other LINE1 sequences, including those of other animal species, are known in the art and can be easily identified in various databases, such as GenBank.

A SINE is a highly repetitive sequence that retrotransposes into a eukaryotic genome through intermediates transcribed by RNA polymerase III (pol III). In many species, SINEs are ubiquitously dispersed throughout the genome and can constitute a significant mass fraction of total genome, for example, typically about 10% or even above 10% in some cases. SINEs cause mutations both by their retrotransposition within genes and by unequal recombination.

SINEs are relatively short (< 700 bp) nonautonomous retroposons transcribed by pol III from an internal promoter and reverse transcribed by the reverse transcriptase of long interspersed elements. Eukaryotic genomes typically contain hundreds of thousands, and sometimes even more, of SINE copies (see Table 1, column: copy number). A SINE typically consists of a head, body and tail. The 5'-terminal head originates from one of the cellular RNAs synthesized by pol III: tRNA, 7SL RNA or 5S rRNA; the body can contain a central domain which may be shared by distant SINE families; and the 3'-terminal tail is a sequence of variable length consisting of simple and often degenerate repeats. Various aspects of SINE structure, biology and evolution have been reviewed in Vassetzky *et al.* (2012). Vassetzky *et al.* also provide a database of SINE families from animals, flowering plants and green algae (see sines.eimb.ru). Non-limiting examples of SINEs from certain animals from the database provided by Vassetzky *et al.* are provided in Table 1. Several variants of the SINEs provided in Table 1 are also described by Vassetzky *et al.* (see sines.eimb.ru). Additional examples of SINEs suitable according to the methods of the claimed invention are well-known to a person of ordinary skill in the art and such embodiments are within the purview of the invention.

Table 1. Non-limiting examples of SINEs from certain animals.

SINE name	Organism	Copy number	SEQ ID NO:
AFC	African cichlids	2×10^3–2×10^4	275
AFC-2	African cichlids	2.4×10^3	276

AFC-3	African cichlids	8×10^2	277
AfroLA	elephants and mammoths (Proboscidea)	7×10^5 (mammoth)	278
AfroSINE	afrotherians	8×10^5	279
Alu	primates	1.1×10^6 (human) 1.5×10^5 (galago)	280
B1	Mus musculus	8×10^3- 6.5×10^5	281
B1-dID	Gliridae, Sciuridae & Aplodontidae	1.1×10^5 (dormouse) 4.5×10^4 (squirrel) 2.8×10^4 (beaver)	282
B4	different rodent families	4×10^5 (mouse) 3.6×10^5 (rat)	283
Bovc-tA2	Bovidae (cattle, goats & sheep)	2×10^4	284
Bov-tA	Bovidae (cattle, goats & sheep)	2×10^5	285
C	Lagomorpha	5.4×10^5	286
CAN	Carnivora	2×10^5	287
CHRS-1	cetaceans, hippopotamuses, ruminants & suiforms	10^4-10^5	288
CHRS-2	cetaceans, hippopotamuses, ruminants & suiforms	10^4-10^5	289
EC-1	horse (*Equus caballus*)	7.7×10^3	290
ERE-A	horses (*Equus* spp.)	9×10^4	291
ERE-B1	horse (*Equus caballus*)	3.8×10^5	292
ERE-B2	horse (*Equus caballus*)	3.8×10^5	293
Fc-1	Carnivora	4×10^4- 6×10^4 (cat, dog, panda)	294

ID	rodents	2.5×10^4- 5×10^5	295
LF	coelacanths & tetrapods (Sarcopterygii)	10^5 (*Latimeria menadoensis*) 3×10^2 (mammals)	296
Mac1	kangaroos (Macropodoidea)	4.5×10^4 (wallaby)	297
MamSINE1	mammals	5×10^2 (platypus)	298
Mar1	marsupials (Metatheria)	5×10^5 (opossum)	299
Mar3	marsupials (Metatheria)	5.3×10^4 (wallaby)	300
Mare3	mammals	1×10^2 (human) 1.4×10^3 (opossum)	301
MEN	squirrels (*Menetes* & *Callosciurus*)	1×10^5	302
Mon-1	monotremes	2×10^6 (platypus)	303
Pca-1	hyrax (*Procavia capensis*)	9.3×10^3	304
Ped-1	springhare (*Pedetes capensis*)	6×10^4	305
Ped-2	springhare (*Pedetes capensis*)	2.5×10^5	306
PinS-I	spruces (*Picea* spp.)	2.7×10^2	307
PRE-1	pigs and peccaries	10^6 (pig)	308
Rhin-1	bat families Rhinolophidae & Hipposideridae		309
RSINE-1	mouse & rat	3×10^4 (mouse) 1.6×10^4 (rat)	310
SS-1	pig (*Sus scrofa*)	1.9×10^5	311
STRIDM	thirteen-lined ground squirrel (Spermophilus tridecemlineatus)	5.4×10^3	312
TAL	moles (Talpidae)	1×10^5	313

Ther-1	mammals, birds & reptiles	4×10^5 (human) 1×10^5 (mouse, rat)	314
Ther-2	marsupials & placentals	8×10^4 (human)	315
WallSI4	wallaby (*Macropus eugenii*)	1×10^5	316

A SINE and/or LINE and/or target sequence within the genome used in an embodiment of the invention is spread throughout the genome and covers at least a certain percentage of the entire genome, and thus forms a representative sample of the entire genome. For example, a preferred SINE and/or LINE and/or target sequence within the genome for use in the assay of the invention typically covers about 4-15%, about 5-14%, about 6-13%, about 7-12%, about 8-11%, about 9-10% or about 10% of the genome.

The SINE and/or LINE and/or target sequence within the genome used for a particular assay depends on the species to which the subject belongs and the prevalence of the SINE and/or LINE and/or target sequence within the genome in the species' genome. A preferred SINE and/or LINE and/or target sequence within the genome for use in the assay of the invention typically comprises about 50-500, about 100-400, about 100-250, about 200-300, about 250-350 or about 300 bp.

A person of ordinary skill in the art can identify a suitable SINE for use in a particular species. Also, depending on the sequence of the SINE and/or LINE and/or target sequence within the genome, a person of ordinary skill in the art can design a target sequence clamp and a primer pair to conduct the assay of the invention. Such embodiments are within the purview of the invention.

The target sequence and, accordingly, the clamp can be designed based on the sequence of the SINE and/or LINE and/or target sequence within the genome. A person of ordinary skill in the art can appreciate that variation exists even within the large number of copies of a particular SINE and/or LINE and/or target sequence within the genome throughout the genome. However, within the larger sequence of a SINE and/or LINE and/or target sequence within the genome, certain portions of the SINE and/or LINE and/or target sequence within the genome do not show much variability among the copies of the SINE and/or LINE and/or target sequence within the genome throughout the genome. For example, in humans, Alu is about 300 bp long; however, portions of Alu that are about 5-50 bp, about

10-40 bp, about 20-30 bp, about 25 bp or about 10-15 bp are highly conserved among the large number of copies of Alu throughout the genome. In certain embodiments, the clamp sequences are designed based on sequences that are conserved across different human races or animal breeds/strains.

In one embodiment, the clamp sequence is selected from the sequences provided in Table 2 below. These sequences are the most common sequences that occur in human Chromosome 1.

Table 2. Examples of frequent sequences in *Homo sapiens* genome, Chromosome 1.

SEQ ID No.	Frequency	Percentage	SEQ ID NO:
1	120976	0.0536	317
2	120236	0.0532	318
3	30120	0.0133	319
4	30056	0.0133	320
5	29754	0.0132	321
6	29720	0.0132	322
7	28146	0.0125	323
8	27970	0.0124	324
9	21179	0.0094	325
10	20808	0.0092	326
11	20167	0.0089	327
12	20161	0.0089	328
13	19956	0.0088	329
14	19794	0.0088	330
15	19704	0.0087	331
16	19690	0.0087	332
17	19685	0.0087	333
18	19590	0.0087	334
19	19389	0.0086	335
20	19336	0.0086	336
21	19243	0.0085	337
22	19076	0.0084	338
23	19061	0.0084	339
24	19035	0.0084	340
25	19002	0.0084	341
26	18958	0.0084	342
27	18926	0.0084	343
28	18851	0.0083	344
29	18837	0.0083	345

30	18734	0.0083	346
31	18624	0.0082	347
32	18563	0.0082	348
33	18526	0.0082	349
34	18498	0.0082	350
35	18397	0.0081	351
36	18390	0.0081	352
37	18385	0.0081	353
38	18335	0.0081	354
39	18312	0.0081	355
40	18231	0.0081	356
41	18150	0.008	357
42	18046	0.008	358
43	17676	0.0078	359
44	17616	0.0078	360
45	17498	0.0077	361
46	17486	0.0077	362
47	16988	0.0075	363
48	16945	0.0075	364
49	16913	0.0075	365
50	16856	0.0075	366
51	16797	0.0074	367
52	16758	0.0074	368
53	16518	0.0073	369
54	16483	0.0073	370
55	16452	0.0073	371
56	16409	0.0073	372
57	16387	0.0073	373
58	16380	0.0073	374
59	16378	0.0073	375
60	16342	0.0072	376
61	16278	0.0072	377
62	16245	0.0072	378
63	16011	0.0071	379
64	15994	0.0071	380
65	15978	0.0071	381
66	15974	0.0071	382
67	15942	0.0071	383
68	15941	0.0071	384
69	15920	0.007	385
70	15882	0.007	386

71	15854	0.007	387
72	15838	0.007	388
73	15820	0.007	389
74	15798	0.007	390
75	15797	0.007	391
76	15770	0.007	392
77	15725	0.007	393
78	15687	0.0069	394
79	15658	0.0069	395
80	15650	0.0069	396
81	15490	0.0069	397
82	15444	0.0068	398
83	14881	0.0066	399
84	14799	0.0066	400
85	14407	0.0064	401
86	14283	0.0063	402
87	14227	0.0063	403
88	14176	0.0063	404
89	13398	0.0059	405
90	13334	0.0059	406
91	13320	0.0059	407
92	13309	0.0059	408
93	13303	0.0059	409
94	13283	0.0059	410
95	13273	0.0059	411
96	13135	0.0058	412
97	12905	0.0057	413
98	12792	0.0057	414
99	12748	0.0056	415
100	12682	0.0056	416

Therefore, in one embodiment, the target sequence, and, accordingly, the clamp, is designed based on the highly conserved portion of a particular SINE.

Examples of Alu SINES, suitable clamp sequences for particular Alu SINES and corresponding primer pairs are given in Table 3.

Table 3. Examples of combinations of clamp sequences, target amplicon SINEs and primer pairs.

Clamp Sequence (SEQ ID NO:)	Target amplicon in Alu SINE (SEQ ID NO:)	Primer 1 (SEQ ID NO:)	Primer 2 (SEQ ID NO:)
7	11	99	100
7	12	101	102
7	13	103	104
7	14	105	106
7	15	107	108
7	16	109	110
7	17	111	112
7	18	113	114
7	19	115	116
7	20	117	118
7	21	119	120
7	22	121	122
7	23	123	124
7	24	125	126
7	25	127	128
7	26	129	130
7	27	131	132
7	28	133	134
7	29	135	136
7	30	137	138
7	31	139	140
7	32	141	142
8	33	143	144
8	34	145	146
8	35	147	148
8	36	149	150

8	37	151	152
8	38	153	154
8	39	155	156
8	40	157	158
8	41	159	160
8	42	161	162
8	43	163	164
8	44	165	166
8	45	167	168
8	46	169	170
8	47	171	172
8	48	173	174
8	49	175	176
8	50	177	178
8	51	179	180
8	52	181	182
8	53	183	184
9	54	185	186
9	55	187	188
9	56	189	190
9	57	191	192
9	58	193	194
9	59	195	196
9	60	197	198
9	61	199	200
9	62	201	202
9	63	203	204
9	64	205	206
9	65	207	208
9	66	209	210
9	67	211	212

9	68	213	214
9	69	215	216
9	70	217	218
9	71	219	220
9	72	221	222
9	73	223	224
9	74	225	226
9	75	227	228
9	76	229	230
9	77	231	232
9	78	233	234
9	79	235	236
9	80	237	238
9	81	239	240
9	82	241	242
9	83	243	244
9	84	245	246
9	85	247	248
9	86	249	250
9	87	251	252
9	88	253	254
9	89	255	256
9	90	257	258
9	91	259	260
9	92	261	262
9	93	263	264
10	94	265	266
10	95	267	268
10	96	269	270
10	97	271	272
10	98	273	274

Table 4 below shows the frequencies of each of the clamp sequences having SEQ ID NOs: 7 to 10 on each of the human chromosomes.

Chromosome number	Clamp sequence (SEQ ID NO: 7)	Clamp sequence (SEQ ID NO: 8)	Clamp sequence (SEQ ID NO: 9)	Clamp sequence (SEQ ID NO: 10)
Chr1	27970	16945	19076	12682
Chr2	22806	13905	15603	10538
Chr3	17925	11051	12513	8269
Chr4	14495	9131	10063	6718
Chr5	15599	9538	10604	7378
Chr6	Not checked	Not checked	Not checked	Not checked
Chr7	18755	11155	12792	8068
Chr8	12928	7890	8812	5785
Chr9	13452	8224	9096	6158
Chr10	14992	8969	10127	6803
Chr11	13129	7956	8913	5957
Chr12	15808	9465	10684	7182
Chr13	7752	4848	5400	3640
Chr14	10086	5998	6654	4435
Chr15	10659	6466	7212	4794
Chr16	13523	7870	9055	5669
Chr17	16867	9989	11071	7424
Chr18	6556	4000	4565	2949
Chr19	15661	9239	10598	6459
Chr20	7933	4623	5301	3482
Chr21	3450	2155	2417	1590
Chr22	6869	4121	4509	3062
ChrX	13474	8043	9512	5828
ChrY	1667	1214	1311	817
Total	302356	182795	205888	135687

In certain embodiments, a number of clamps encompassing a highly conserved region are designed by "walking across" the highly conserved region. Each of these clamps can be tested to identify the clamp which exhibits maximum experimental ease, accuracy and reproducibility. For the purposes of this invention, the term "walking across" indicates the process of designing a plurality of clamps that bind to an area of interest, wherein each of the plurality of clamps has a specific length, for example, 10-15 bp, each of the plurality of clamps begins at a particular nucleotide of the area of interest and the plurality of clamps as a

whole cover the entire are of interest. In general, the clamps are about 15-50 bp, about 15-40 bp, about 15-30 bp, about 16-28 bp, about 17-26 bp, about 18-24 bp, about 19-22 bp, or about 20 bp.

For example, if a SINE is about 300 bp containing a highly conserved region of 50 bp, a plurality of clamps, each of about 10-15 bp, is designed by walking across the highly conserved region of 50 bp. For example, a plurality of clamps of 10 bp are designed based on a highly conserved region of 50 bp by designing clamps that have the sequence of 1-10 bases, 2-11 bases, 3-12 bases, . . . , and 41-50 bases of the highly conserved region. As such, 41 different clamps can be designed and tested to identify the preferred clamp for use in an assay according to the invention.

A person of ordinary skill in the art can design a plurality of clamps of a particular length within an area of interest, and such embodiments are within the purview of the invention.

For example, the sequence for Alu is highly conserved, particularly among the first 50 bp, and other sites in the Alu sequence. A sequence alignment between certain Alu sequences is provided in Figure 1 of Batzer *et al.* (2002), Nature Reviews Genetics, 3(5):370-379, which is herein incorporated by reference in its entirety. Based on Figure 1 of Batzer *et al.*, a person of ordinary skill in the art can appreciate that a clamp can be designed to cover the maximum genomic region based on the sequences that correspond to the sequences that are conserved among various versions of Alu sequences. On the other hand, a clamp for an Alu sequence of interest can be designed based on the variation present in the specific Alu sequence of interest.

Additional examples of Alu sequences are provided as SEQ ID NOs: 417 to 478. Conserved domains in these sequences can be determined by a person of ordinary skill in the art to design appropriate clamp sequences and primer sequences.

In a further embodiment, multiple versions of clamps can be used in an assay where the different versions of the clamps are directed to variants of Alu sequences. Therefore, as a whole, the multiple versions of clamps can be used to identify mutations in a region of interest beyond the variability naturally observed in the region of interest.

In preferred embodiments, the clamps are designed based on a region of about 16 to 20 highly conserved bp. The clamp sites also require the presence of suitable 5' and 3' primers for the PCR component.

Examples of 14 bp clamps derived from the areas of greatest frequency in Alu, which occur in both mice and humans, are provided in SEQ ID NOs: 7-10. The clamp sequence CCTGTAATCCCAGC (SEQ ID NO: 7) has about 16,000 repeats on chromosome 12; the clamp sequences CTAAAAATACAAAA (SEQ ID NO: 8) and TGCACTCCAGCCTG (SEQ ID NO: 9) each have approximately 10,000 repeats on chromosome 12; and the clamp sequence TCTCAAAAAAAAAA (SEQ ID NO: 10) has approximately 7,000 repeats on chromosome 12. All of these are clamps are suitable for appropriate 3' and 5' PCR primers.

In one embodiment, the subject is a mouse and the SINE is B1 having the sequence of SEQ ID NO: 6. An example of the primer pair used in mice with this B1 (SEQ ID NO: 6) as the SINE comprises the primers having the sequences of SEQ ID NOs: 1 and 2 and the target sequence clamp having a sequence selected from SEQ ID NO: 3, 4 and 5.

In another embodiment, the subject is a mouse and the SINE is B1 having the sequence of SEQ ID NO: 281. An example of the primer pair used in mice with this B1 (SEQ ID NO: 281) as the SINE comprises the primers having the sequences of SEQ ID NOs: 479 and 480 and the target sequence clamp having a sequence selected from SEQ ID NO: 3, 4 and 481.

In a further embodiment of the invention, the subject is human and the SINE is Alu.

The step of obtaining a genomic DNA sample from the subject and fragmenting the DNA sample can be performed based on methods well known in the art. Methods and parameters used in fragmenting genomic DNA depend on the size of the genome and the desired average and median size of the fragments. The desired size of the fragments depends on the size of the target SINE, target LINE and/or target sequence within the genome *i.e.*, most of the fragments must allow binding of the primers and the target sequence clamps for the assay to be successful. For example, for a target SINE, target LINE and/or target sequence within the genome of about 100 to 500 bp, the substantial number of DNA fragments is about 800-1500 bp each. For example, in a typical genomic DNA fragment sample, each fragment from at least about 80-99%, about 82-98%, about 84-96%, about 86-94%, about 88-92%, or about 90% of all of the genomic DNA fragments have the desired size. A person of ordinary skill in the art can determine the optimal size of the genomic fragments for a particular assay. Also, the techniques of producing genomic fragments of a desired size are well-known in the art and such embodiments are within the purview of the invention.

A person of ordinary skill in the art would appreciate that since the fragmenting of the genomic DNA is random, a fragment can contain more than one SINE and/or one or more LINE and/or target sequence within the genome. A fragment can also contain only a part of the SINE or part of a LINE and/or target sequence from within the genome. These issues can be addressed by using a large sample and/or running multiple repeats of the assay so that the possible errors are diluted and a more accurate estimation of the mutations is obtained.

Mixing a predetermined number of fragments of the genomic DNA that arise from a predetermined number of genomes with a reagent mixture to produce a reaction mixture is intended to provide the values used in the calculations of the accumulated mutations and/or rate of mutations in the target sequence within the target SINE and/or target LINE and/or the genome. For example, a person of ordinary skill in the art can calculate the accumulated mutations and/or the rage of mutations in the target SINE and/or target LINE and/or target sequence within the genome based on the number of fragments arising from one genome, the number of fragments subjected to dPCR, the number of mutated target sequences as indicated by the positive PCR amplification results and the size and frequency of the target SINE and/or target LINE and/or target sequence within the genome in the genome. Also, based on the size and frequency of the target SINE and/or target LINE and/or target sequence within the genome and the accumulated mutations and/or the rate of mutations in the target SINE and/or target LINE and/or target sequence within the genome, a person of ordinary skill in the art can calculate the accumulated mutations and/or the rate of mutations in the genome.

In addition to the primer pair, the target sequence clamp and the DNA polymerase, the reagent mixture contains reagents for the dPCR. The reagent mixture comprises deoxyribonucleotides (dNTPs), metal ions (for example, $Mg^{2+}$and Mn^{2+}), and a buffer. Additional reagents which may be used in a dPCR reaction are well-known to a person of ordinary skill in the art and such embodiments are within the purview of the invention.

The pair of polymerase chain reaction primers that amplify a target amplicon comprises the target sequence within the target SINE and/or target LINE and/or target sequence within the genome. The primers are designed so that an amplicon is not produced when the target sequence clamp is bound to the target sequence within the target SINE and/or target LINE and/or target sequence within the genome.

Based on a particular target SINE and/or target LINE and/or target sequence within the genome , for example, a target SINE selected from Table 1, a person of ordinary skill in the art can design appropriate primers and the target sequence clamp. For a particular target

SINE and/or target LINE and/or target sequence within the genome, person of ordinary skill in the art can test multiple primer pairs and/or target sequence clamps to identify the optimal combination of primers and target sequence clamps and such embodiments are within the purview of the invention.

A target sequence within a target SINE and/or target LINE and/or target sequence within the genome is the sequence to which the target sequence clamp binds. Particularly, the target sequence clamp is complementary to the target sequence within the target SINE and/or target LINE and/or target sequence within the genome.

A wild-type target sequence does not contain any mutations. A mutated target sequence contains one or more point mutations and/or indel mutations. Accordingly, a wild-type target amplicon contains the wild-type target sequence and a mutated target amplicon contains a mutated target sequence.

In one embodiment, the target sequence clamp is designed based on the sequence of the SINE, LINE and/or genomic sequence and is about 15-50 bp, about 15-40 bp, about 15-30 bp, about 16-28 bp, about 17-26 bp, about 18-24 bp, about 19-22 bp, or about 20 bp.

In an embodiment of the invention, the target sequence clamp is designed so that the melting temperature of the target sequence clamp with the target sequence is higher than the temperatures used in the PCR cycle. The higher melting temperature of the clamp ensures that the clamp is bound to the clamp target sequence during the PCR cycles when the clamp is perfectly matched with the target sequence. A mutation in a target sequence reduces the melting temperature of the target sequence clamp with the mutated target sequence and the target sequence clamp is not bound to the mutated target sequence at the temperatures of the PCR cycles, particularly the annealing steps and the amplification steps of the PCR cycles. Therefore, the target sequence clamp prevents PCR amplification of the target amplicon when the amplicon contains the wild-type target sequence and the clamp permits PCR amplification of the target amplicon when the amplicon contains a mutated target sequence.

In another embodiment of the invention, the target sequence clamp comprises xenonucleotide (XNA). A variety of XNA are known in the art. The target sequence XNA clamp also suppresses PCR amplification of the amplicons containing wild-type clamp target sequences and allows selective PCR amplification of only the amplicons containing mutated target clamp sequences. XNA, for example, can contain an amino acid linkages rather than a phosphate between bases, which causes it to bind tightly with the wild-type clamp target sequence and reduces hydration and heat instability. Therefore, a target XNA sequence

clamp does not melt off the wild-type clamp target sequence at the usual PCR temperatures when the match is perfect.

In a further embodiment, a target XNA clamp of about 13-20 bp is used. In other embodiments, the XNA clamp is about 15-50 bp, about 15-40 bp, about 15-30 bp, about 16-28 bp, about 17-26 bp, about 18-24 bp, about 19-22 bp, or about 20 bp. When a 13-20-bp XNA is used, a single-point mutation in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome lowers the melting point between the two binding sequences by 15-20°C. Indel mutations lower the melting point between a target XNA sequence clamp and the wild-type clamp target sequence by more than 15-20°C. Because the mutated target sequence does not bind the target XNA sequence clamp, only the amplicons containing the mutated target sequences are amplified during PCR.

Accordingly, when the target sequence clamp binds to the wild-type target sequence within the target SINE and/or target LINE and/or target sequence within the genome the target sequence clamp prevents the PCR amplification of the target amplicons that have the target wild-type sequence within the SINE. In contrast, when the target sequence contains a mutation, the target sequence clamp cannot bind to the target sequence, which allows the PCR amplification of the target amplicons that have the target mutated sequence within the SINE, LINE and/or target genomic sequence.

dPCR, as used in the claimed invention, refers to a PCR where the PCR reaction is carried out as a single reaction within a sample; however, the sample is separated into a large number of partitions and the reaction is carried out in each partition individually and separately from the other partitions. dPCR involves identification of the amplification of the target amplicons in each of the large number of partitions. dPCR enables precise and highly sensitive quantification of nucleic acids. An overview of dPCR is provided by Baker (2012), the contents of which are incorporated herein in their entirety.

In one embodiment of the invention, the dPCR used in the assay is droplet digital PCR (ddPCR). In ddPCR, a PCR sample is partitioned into a large number of droplets, for example, 20,000 droplets, using water-oil emulsion droplet technology. After amplification, droplets containing the target sequence are detected by fluorescence and scored as positive, and droplets without fluorescence are scored as negative. Poisson statistical analysis of the numbers of positive and negative droplets yields absolute quantitation of the target sequences. An overview of ddPCR is provided by Hundson *et al.* (2011), the contents of which are incorporated herein in their entirety.

When the dPCR results are obtained, the number of accumulated mutations per genome in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome is calculated based on the number of fragments of the genomic DNA that arise from one genome and the number of fragments per genome that comprise the target amplicons having the target mutated sequence within the target SINE and/or target LINE and/or target sequence within the genome. As discussed above, the presence of the target mutated sequence within the target SINE and/or target LINE and/or target sequence within the genome is indicated by the positive PCR amplification in the dPCR.

In standard dPCR or a ddPCR mix, the assay of the invention enables the detection of 1-2 mutant DNA fragments in a pool of 100,000 wild-type amplicons (Fig. 6). The use of the target sequence clamp in combination with dPCR provides an extraordinary mutation screening method. For example, in a 45-min cycle of a ddPCR (BioRad QX200 AutoDG ddPCR, Hercules, CA) over 10^6 DNA fragments can be analyzed in each of the 8 channels for the presence of mutations. Since about 10% of those DNA fragments likely contain Alu, ~10^5 Alu are analyzed at one fragment per well or drop. Using 100 or 1,000 fragments per drop instead of ~1-2 per drop improves the screening of Alu fragments by 2 or 3 orders of magnitude (up to 10^{-8} Alu/channel).

As such, an embodiment of the invention provides a method of estimating the accumulated mutations in a target sequence within a target SINE and/or target LINE . The number of accumulated mutations in a target sequence within a target SINE and/or target LINE and/or target sequence within the genome can be used to determine the rate of mutations in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome. For example, the accumulated mutations in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome can be determined at two time points and the rate of mutations can be calculated based on the difference in the number of accumulated mutations in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome and the duration between the two time points.

In another embodiment, a first sample is obtained from the subject at Time 1 and a second sample is obtained from the subject at Time 2. The accumulated mutations in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome are estimated in the first and the second samples according to the clamp/dPCR combination assay and the rate of mutations in the target sequence within the target SINE

and/or target LINE and/or target sequence within the genome can be calculated based on the difference in the number of accumulated mutations in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome and the duration between Time 1 and Time 2.

In a specific embodiment, a sample is obtained from the subject at birth. This sample provides accumulated mutations in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome at birth, which can be considered as baseline mutations or the state of no mutations. The accumulated mutations estimated in a sample obtained from the subject at a later time can be compared to the baseline mutations or the state of no mutations.

Accordingly, an embodiment of the invention provides a method for calculating the rate of mutations in a target sequence within a target SINE and/or target LINE and/or target sequence within the genome in a subject. The method comprises the steps of:

a) according to the clamp/dPCR combination assay, determining the number of accumulated mutations in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome in a first sample obtained from the subject at a first time point,

b) according to the clamp/dPCR combination assay, determining the number of accumulated mutations in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome in a second sample obtained from the subject at a second time point,

c) calculating the rate of mutations in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome in the subject based on the difference between the number of accumulated mutations in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome in the subject at the first time point and the second time point and the duration between the first time point and the second time point.

The number of accumulated mutations and/or the rate of mutations in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome can be used to estimate the accumulated mutations and/or the rate of mutations in the genome of a subject. For example, the number of accumulated mutations in the genome of a subject can be calculated based on the frequency of occurrence of a target SINE and/or target LINE and/or target sequence within the genome throughout the genome and the number of

accumulated mutations in the target sequence within the target SINE and/or target LINE and/or target sequence within the genome.

Similarly, the rate of mutations in a target sequence within a target SINE and/or target LINE and/or target sequence within the genome can be used to estimate the rate of mutations in the genome of the subject. For example, the rate of accumulated mutations in the genome of the subject can be calculated based on the frequency of occurrence of the target SINE and/or target LINE and/or in the genome and the rate of mutations in the target sequence within the target SINE.

As such, an embodiment of the invention provides a method for determining accumulated mutations and/or the rate of the mutations in a target sequence within a target SINE and/or target LINE and/or the genome of the subject.

Accumulated mutations and/or the rate of mutations typically increase with age. For example, the number of accumulated mutations and/or the rate of mutations in a chronologically older subject are typically higher than the corresponding values in a chronologically younger subject. Also, different individuals age at a different rates, *i.e.*, the accumulated mutations and/or the rate of mutations in two individuals of the same age can be different. For example, individuals exposed to higher levels of mutagens like carcinogens, mutagenic chemicals, radiation, stress, *etc.* typically have more accumulated mutations and/or a higher rate of mutations compared to individuals not exposed to such mutagens or exposed to relatively lower levels of mutagens.

A standard scale for genomic age for a particular species can be determined based on average accumulated mutations and/or the average rate of mutations in different groups of individuals of varying ages that are living under the conditions of exposure to only natural mutagens and/or the conditions of minimal exposure to man-made mutagens.

For the purpose of the invention, the phrase "the conditions of exposure to only natural mutagens" indicates exposure to only unavoidable natural mutagens, for example, cosmic radiation, ultraviolet rays from the sun, mutagens that may be naturally (*i.e.*, without interference from humans) present in soil, air, water, and food or other environmental factors. Additional examples of unavoidable natural mutagens can be readily envisioned by a person of ordinary skill in the art.

For example, an individual living in the conditions of exposure to only natural mutagen is living in conditions that are free from:

a) exposure to man-made mutagens, such as synthetic carcinogens, synthetic pollutants, radiation from man-made sources, *etc.*, and

b) avoidable/unnecessary exposure to natural mutagens, for example, smoking, using tobacco and other avoidable/unnecessary exposure to natural carcinogens.

Similarly, for the purpose of the invention, the phrase "the conditions of minimal exposure to man-made mutagens" indicates minimal exposure to unavoidable natural mutagens (discussed above) and minimal exposure to man-made mutagens, such as synthetic carcinogens, synthetic pollutants and radiation from man-made sources. The conditions of minimal exposure to man-made mutagens are also free from avoidable/unnecessary exposure to natural mutagens, for example, smoking, using tobacco and other avoidable/unnecessary exposure to natural carcinogens.

An example of an individual living under the conditions of exposure to only natural mutagens and/or the conditions of minimal exposure to man-made mutagens is an individual living in the countryside. Because of the industrialized lifestyle of almost everyone in the world, it is very difficult and almost impossible to find individuals living under the conditions of exposure to only natural mutagens. Therefore, the standard scale for the genomic age for a particular species can be determined based on the average accumulated mutations and/or the average rate of mutations in different groups of individuals of varying ages that are living under the conditions of minimal exposure to man-made mutagens.

Accordingly, a standard scale for the genomic age of humans can be produced by determining the average accumulated mutations and/or the average rate of mutations in humans of varying ages that live in the conditions of minimal exposure to man-made mutagens, for example, people living in the countryside. Such a scale of genomic age can be used to determine the genomic age of an individual based on the individual's accumulated mutations and/or rate of mutations in the genome.

The exposure to avoidable/unnecessary natural mutagens and/or the exposure to man-made mutagens typically increase the accumulated mutations and/or the rate of mutations in the genome of a subject. For example, a person living in the countryside typically has fewer accumulated mutations and/or a lower rate of mutations compared to a person living in a city, particularly a polluted city. Therefore, a person of a particular chronological age living in the countryside typically has a lower genomic age compared to the genomic age of a person of the same chronological age living in a city.

Accordingly, the clamp/dPCR combination assay for determining the accumulated mutations and/or the rate of the mutations in a target sequence in a target SINE, target LINE and/or the genome of a subject can be used to determine the genomic age of the subject. The method comprises the steps of:

a) preparing a standard genomic age scale for individuals belonging to the species of the subject and living under the conditions of exposure to only natural mutagens or the conditions of minimal exposure to man-made mutagens, or obtaining a pre-determined standard genomic age scale for the species of the subject,

b) determining the accumulated mutations and/or the rate of mutations in the subject according to the clamp/dPCR combination assay, and

c) estimating the genomic age of the subject based on the comparison of the accumulated mutations and/or the rate of mutations in the subject with the standard scale for the genomic age of the subject.

Similar to increasing age, a subject having cancer or having a higher risk of developing cancer exhibits an increase in the accumulated mutations and/or the rate of mutations in a target sequence in a target SINE, target LINE and/or the genome. For example, a person of a particular chronological age having more accumulated mutations and/or a higher rate of mutations in a target sequence in a target SINE, target LINE and/or the genome is at a higher risk of developing cancer compared to a person of the same chronological age who has relatively fewer accumulated mutations and/or a lower rate of mutations in a target sequence in a target SINE, target LINE and/or the genome.

Also, chronologically older individuals are at a higher risk of developing cancer because the accumulated mutations and/or the rate of mutations in a target sequence in a target SINE, target LINE and/or the genome are typically higher in chronologically older individuals compared to the corresponding values in chronologically younger individuals.

Accordingly, an embodiment of the invention provides a method of identifying a higher risk of cancer development in a subject based on the accumulated mutations and/or the rate of mutations in a target sequence in a target SINE, target LINE and/or the genome of the individual and a standard scale of cancer risk in the species to which the subject belongs.

The standard scale of cancer risk indicates the risk of cancer in a subject based on the accumulated mutations and/or the rate of mutations in a target sequence in a target SINE, target LINE and/or the genome of the subject. For example, more accumulated mutations and/or a higher rate of mutations in a target sequence in a target SINE, target LINE and/or the

genome in a subject indicates a higher risk of cancer development in the subject compared to an individual of the same chronological age as the subject and having relatively fewer accumulated mutations and/or a lower rate of mutations in a target sequence in a target SINE, target LINE and/or the genome.

A standard scale of the cancer risk for a species, for example, humans, can be produced by determining the average accumulated mutations and/or the average rate of mutations in a target sequence in a target SINE, target LINE and/or the genome in individuals of varying ages that are free from cancer and/or are known to have a low risk of cancer development. The standard scale of cancer risk can be used to determine the risk of cancer development in a subject based on the subject's accumulated mutations and/or rate of mutations in a target sequence in a target SINE, target LINE and/or the genome and the subject's chronological age. As such, the standard scale of cancer risk in the species indicates, at increasing chronological age, the average accumulated mutations and/or the average rate of mutations in the target sequence in the target SINE, target LINE and/or the genomes of individuals of varying ages that belong to the species and are free from cancer and/or are known to have a low risk of cancer development.

Accordingly, an embodiment of the invention provides a method for determining the risk of cancer development of a subject based on accumulated mutations and/or rate of mutations in a target sequence in a target SINE, target LINE and/or the genome of the subject. The method comprises the steps of:

a) preparing a standard scale for cancer risk by determining the average accumulated mutations and/or the average rate of mutations in a target sequence in a target SINE, target LINE and/or the genome in individuals of varying ages that are free from cancer and/or are known to have a low risk of cancer development, or obtaining a pre-determined standard scale of cancer risk,

b) determining the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome in the subject, and

c) estimating the risk of cancer development of the subject based on the comparison of the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome of the subject with the standard scale of cancer risk.

The step of estimating the risk of cancer development of the subject based on the comparison of the accumulated mutations and/or the rate of mutations in the target sequence

in the target SINE, target LINE and/or the genome of the subject with the standard scale can be:

a) identifying the subject as having a higher risk of cancer development if the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome of the subject are higher than the corresponding values in the standard scale of cancer risk, or

b) identifying the subject as having a lower risk or no risk of cancer development if the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome of the subject are lower than or equal to the corresponding values in the standard scale of cancer risk.

Every individual always carries some risk of cancer development. For example, spontaneous oncogenic mutations occur even among individuals living with minimal exposure to mutagens. Therefore, for the purpose of this invention, a higher risk of cancer development of a subject refers to a higher risk of cancer development compared to the risk of cancer development in the population of the same chronological age as the subject that is free from cancer and/or is known to have a low risk of cancer development. Similarly, a lower risk of cancer development of a subject refers to a lower risk of cancer development compared to the risk of cancer development in the average population of the same chronological age as the subject that is free from cancer and/or is known to have a low risk of cancer development.

If a subject is identified as having a higher risk of cancer development, enhanced screening for cancer can be administered to the subject for early detection and treatment of cancer. As is well-established in the art, early detection and treatment of cancer typically results in cancer-free survival. Therefore, administering enhanced screening to a subject based on the subject's identification as having a higher risk of cancer development ensures that the cancer, if developed, is identified during the early stages, thereby increasing the chances of cancer-free survival of the subject.

Enhanced screening for cancer indicates that the cancer screening is administered more frequently than recommended for a healthy individual. For example, if recommended screening frequency for cancer for a healthy individual is once a year, an individual identified as having a higher risk of cancer development can be screened every six months. Recommended cancer screening schedules for various cancers and the modifications which can be done to the recommended schedules to produce an enhanced screening schedule are

well-known to a person of ordinary skill in the art and such embodiments are within the purview of the invention.

Also, if a subject is identified as having a higher risk of cancer development, lifestyle changes can be recommended to the subject to reduce the risk of cancer development. Non-limiting examples of lifestyle changes which can reduce the risk of cancer development include cessation of smoking, reducing the exposure to a known carcinogen, or changing a profession or job which poses increased exposure to a particular carcinogen. Additional examples of lifestyle changes which can reduce the risk of cancer development are well-known to a person of ordinary skill in the art and such embodiments are within the purview of the invention.

If a subject is identified as having a low risk or no risk of cancer development, enhanced screening for cancer is withheld from the subject and, optionally, routine screening is administered. Withholding enhanced screening for cancer from a subject based on the subject's identification as having a lower risk or no risk of cancer development ensures that the subject does not receive any unnecessary cancer screening. Avoiding unnecessary cancer screening may be significant because sometimes the cancer screening itself uses mutagens, for example, x-rays for the identification of breast cancer.

Personal lifestyle and environmental exposures affect a subject's risk for cancer. These factors are extrinsic to a subject's inherited genetics. For example, smoking causes as much as a 10-fold increased rate of accumulation of pulmonary epithelial mutations. Therefore, an embodiment of the invention provides a tissue; or organ-specific prediction of the risk of cancer development. Most solid tumors in adults have 33 to 66 genes with subtle somatic mutations expected to alter their proteins (Fig. 7); the exceptions include lung cancers (smoking; ~150) and melanomas (sun exposure; ~135). Thus, tissues naturally exposed to powerful carcinogens appear to require more mutations, suggesting they are relatively resistant to cancer. The level of passengers/silent mutations can be established for cancer types in specific tissues or organs and this level can be used to predict the risk of organ-specific cancer development.

Accordingly, a further embodiment of the invention provides a method of identifying a higher risk of cancer development in a tissue or organ of a subject based on the accumulated mutations and/or the rate of mutations in a target sequence in a target SINE, target LINE and/or the genome of the cells of the tissue or organ.

A standard scale of the cancer risk for a tissue or organ in a species, for example, breast cancer in humans, can be produced by determining the average accumulated mutations and/or the average rate of mutations in a target sequence in a target SINE, target LINE and/or the genome in the cells of the tissue or organ from individuals of varying ages that are free from cancer and/or are known to have a low risk of cancer development. The standard scale of cancer risk for a tissue or organ can be used to determine the risk of cancer development in the tissue or organ of a subject based on the subject's accumulated mutations and/or rate of mutations in a target sequence in a target SINE, target LINE and/or the genome in the cells of the tissue or organ and the subject's chronological age. As such, a standard scale of cancer risk for a tissue or organ in a species indicates, at increasing chronological ages, the average accumulated mutations and/or the average rate of mutations in a target sequence in a target SINE, target LINE and/or the genomes of the cells in the tissues or organs of individuals of varying ages that belong to the species and are free from cancer and/or are known to have a low risk of cancer development in the tissue or organ.

Non-limiting examples of the tissue or organ which can be used in the methods of the invention include placenta, brain, eyes, pineal gland, pituitary gland, thyroid gland, parathyroid glands, thorax, heart, lung, esophagus, thymus gland, pleura, adrenal glands, appendix, gall bladder, urinary bladder, large intestine, small intestine, kidneys, liver, pancreas, spleen, stoma, ovaries, uterus, testis, skin, blood or buffy coat sample of blood. Additional examples of organs and tissues are well-known to a person of ordinary skill in the art and such embodiments are within the purview of the invention.

In certain embodiments, the methods of current invention are practiced to determine the risk of cancer, wherein the cancer is selected from acute lymphoblastic lukemia, acute myeloid leukemia, adrenocortical carcinoma, AIDS-related cancers, AIDS-related lymphoma, anal cancer, appendix cancer, astrocytoma, cerebellar cstrocytoma, basal cell carcinoma, bile duct cancer, extrahepatic bladder cancer, bladder cancer, bone cancer, osteosarcoma and malignant fibrous histiocytoma, brain stem glioma, brain tumor, central nervous system embryonal tumors, cerebral astrocytoma/malignant glioma, ependymoblastoma, medulloblastoma, medulloepithelioma, pineal parenchymal tumors of intermediate differentiation, supratentorial primitive neuroectodermal tumors and pineoblastoma, visual pathway and hypothalamic glioma, brain and spinal cord tumors, breast cancer, bronchial tumors, Burkitt lymphoma, carcinoid tumor, gastrointestinal cancer, carcinoma of the head and neck, central nervous system embryonal tumors, central nervous

system lymphoma, cervical cancer, chordoma, chronic lymphocytic leukemia, chronic myelogenous leukemia, chronic myeloproliferative disorders, colorectal cancer, cutaneous T-cell lymphoma, embryonal tumors, endometrial cancer, ependymoblastoma, ependymoma, esophageal cancer, Ewing family of tumors, extracranial germ cell tumor, extrahepatic bile duct cancer, eye cancer, intraocular melanoma, retinoblastoma, gallbladder cancer, gastric (stomach) cancer, gastrointestinal carcinoid tumor, gastrointestinal stromal tumor (GIST), extracranial germ cell tumor, germ cell tumor, extragonadal germ cell tumor, ovarian cancer, gestational trophoblastic tumor, glioma, brain stem glioma, hairy cell leukemia, head and neck cancer, hepatocellular (liver) cancer, Hodgkin's lymphoma, hypopharyngeal cancer, hypothalamic and visual pathway glioma, intraocular melanoma, islet cell tumors (endocrine pancreas), Kaposi sarcoma, kidney (renal cell) cancer, kidney cancer, laryngeal cancer, chronic lymphocytic leukemia, chronic leukemia, myelogenous leukemia, lip and oral cavity cancer, lung cancer, non-small cell lung cancer, small cell lymphoma, AIDS-related lymphoma, cutaneous T-cell lymphoma, non-Hodgkin's lymphoma, macroglobulinemia, Waldenström macroglobulinemia, malignant fibrous histiocytoma of bone and osteosarcoma, medulloblastoma, medulloepithelioma, melanoma, intraocular Merkel cell carcinoma, mesothelioma, metastatic squamous neck cancer with occult primary, mouth cancer, multiple endocrine neoplasia syndrome, multiple myeloma/plasma cell neoplasm, mycosis fungoides, myelodysplastic syndromes, myelodysplastic/myeloproliferative diseases, myelogenous leukemia, multiple myeloproliferative disorders, nasal cavity and paranasal sinus cancer, nasopharyngeal cancer, neuroblastoma, non-small cell lung cancer, oral cancer, oral cavity cancer, lip and oropharyngeal cancer, osteosarcoma and malignant fibrous histiocytoma of bone, ovarian epithelial cancer, ovarian germ cell tumor, ovarian low malignant potential tumor, pancreatic cancer, papillomatosis, paranasal sinus and nasal cavity cancer, parathyroid cancer, penile cancer, pharyngeal cancer, pheochromocytoma, pineal parenchymal tumors of intermediate differentiation, pineoblastoma and supratentorial primitive neuroectodermal tumors, pituitary tumor, plasma cell neoplasm/multiple myeloma, pleuropulmonary blastoma, primary central nervous system lymphoma, prostate cancer, rectal cancer, renal pelvis and ureter raner, transitional cell cancer, respiratory tract carcinoma involving the NUT gene on chromosome 15, retinoblastoma, rhabdomyosarcoma, salivary gland cancer, sarcoma, soft tissue sarcoma, uterine Sézary syndrome, skin cancer (nonmelanoma), skin carcinoma, small cell lung cancer, small intestine cancer, squamous cell carcinoma, squamous neck cancer with occult primary cancer, supratentorial primitive neuroectodermal tumors, T-cell

lymphoma, testicular cancer, throat cancer, thymoma and thymic carcinoma, thyroid cancer, transitional cell cancer of the renal pelvis and ureter, gestational trophoblastic tumor, carcinoma of unknown primary site, urethral cancer, uterine cancer, endometrial uterine sarcoma, vaginal cancer, visual pathway and hypothalamic glioma, vulvar cancer, and Wilms' tumor.

Accordingly, an embodiment of the invention provides a method for determining the risk of cancer development of a tissue or organ of a subject based on accumulated mutations and/or rate of mutations in a target sequence in a target SINE, target LINE and/or the genome in the cells of the tissue or organ of the subject. The method comprises the steps of:

a) preparing a standard scale for the cancer risk for the tissue or organ by determining the average accumulated mutations and/or the average rate of mutations in a target sequence in a target SINE, target LINE and/or the genome in the cells of the tissue or organ in individuals of varying ages that are free from cancer and/or are known to have a low risk of cancer development, or obtaining a pre-determined standard scale of cancer risk,

b) according to the clamp/dDNA combination assay of the invention, determining the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome in the cells of the tissue or organ of the subject, and

c) estimating the risk of cancer development of the tissue or organ of the subject based on the comparison of the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome in the cells of the tissue or organ of the subject with the standard scale of cancer risk for the tissue or organ.

A person of ordinary skill in the art would appreciate that the steps described above for estimating the risk of cancer development in the subject and the steps taken after a subject is identified as having a higher or lower risk of cancer development are relevant to cancer development of a tissue or organ and are within the purview of the invention as it relates to the estimation of a risk of cancer development of a tissue or organ.

Lifestyle changes are prescribed by medical professionals as means of improving overall health and well-being of humans, including decreasing the chances of cancer development or slowing down the rate of aging. An embodiment of the invention provides a method of determining whether a particular lifestyle change or a combination of lifestyle changes effectively reduced the risk of cancer development in a subject or effectively slowed down the rate of aging.

For example, a person's risk of cancer development, such as a cancer of a particular tissue or organ, can be determined before and after a lifestyle change is initiated. Similarly, a person's rate of aging, for example, rate of increase in the genomic age, can be determined before and after a lifestyle change is initiated.

Non-limiting examples of lifestyle changes that are recommended for reducing the risk of cancer and/or slowing down the rate of aging include weight loss, cessation of smoking, limiting the exposure to a known carcinogen, change of a profession or job to avoid exposure to a particular carcinogen, dietary changes, *etc*. Additional examples of lifestyle changes that can be prescribed to reduce the risk of cancer development and/or slow down the rate of aging are well-known to a person of ordinary skill in the art and such embodiments are within the purview of the invention.

If the risk of cancer development and/or the rate of aging are reduced, indicated, for example, by a reduced rate of mutations in a target sequence in a target SINE, target LINE and/or the genome of the subject, the lifestyle change is considered to be successful in achieving the intended goal.

On the other hand, if the risk of cancer development and/or the rate of aging are not reduced or are not reduced to the desired extent, indicated, for example, by no reduction or lack of reduction to the preferred extent in the rate of mutations in a target sequence in a target SINE, target LINE and/or the genome of the subject, the lifestyle change is considered to be unsuccessful. In such cases, different and/or additional lifestyle changes can be recommended to the subject for achieving the desired result.

Exposure to mutagens, changes in environment or changes in lifestyle can affect the overall health and well-being of humans, including altering the chances of cancer development and/or changing the rate of aging. An embodiment of the invention provides a method of determining whether a lifestyle change or a combination of lifestyle changes, exposure to mutagens, or changes in environment altered the risk of cancer development of a subject and/or changed the rate of aging. For example, a person's risk of cancer development, such as a cancer of a particular tissue or organ, and/or the rate of aging as indicated by the genomic age can be determined before and after the lifestyle change was initiated. Non-limiting examples of lifestyle changes which can alter a subject's risk of cancer development and/or rate of aging include weight gain, smoking, exposure to a known carcinogen, change of a profession or job causing increased exposure to a particular carcinogen, dietary changes, *etc*. Additional examples of lifestyle changes that can alter the

risk of cancer development and/or change the rate of aging are well-known to a person of ordinary skill in the art and such embodiments are within the purview of the invention.

Accordingly, an embodiment of the invention provides a method of identifying the effect of a lifestyle change on the risk of cancer development and/or the rate of aging of a subject. The method comprises:

a) according to the clamp/dPCR combination assay, determining the rate of mutations in a target sequence within a target SINE or target LINE and/or in the genome of the subject immediately before the lifestyle change is initiated,

b) according to the clamp/dPCR combination assay, determining the rate of mutations in the target sequence within the target SINE or target LINE and/or in the genome of the subject after the lifestyle change is initiated, and

c) comparing the rates of mutations in the target sequence within the target SINE or target LINE and/or the genome of the subject before and after the lifestyle change is initiated to determine the effect of the lifestyle change on the risk of cancer development and/or the rate of aging of the subject.

If the risk of cancer development increases and/or the rate of aging increases, indicated, for example, by a higher rate of mutations in a target sequence in a target SINE, target LINE and/or the genome of the subject, the lifestyle change is considered to increase the risk of cancer development and/or increase the rate of aging. The subject can then be recommended to either eliminate the lifestyle change which increased the risk of cancer and/or increased the rate of aging. Alternately, another lifestyle change intended to counter the earlier lifestyle change can be recommended.

On the other hand, if the risk of cancer development and/or the rate of aging does not increase or does not increase to an alarming extent, indicated, for example, by no increase or lack of increase to an alarming extent in the rate of mutations in a target sequence in a target SINE, target LINE and/or the genome of the subject, the lifestyle change is considered to be harmless. In such cases, no unnecessary changes in lifestyle are recommended.

Similarly, an embodiment of the invention provides a method of identifying the effect of an exposure to a mutagen on the risk of cancer development and/or the rate of aging of a subject. The method comprises:

a) according to the clamp/dPCR combination assay, determining the rate of mutations in a target sequence within a target SINE or target LINE and/or in the genome of the subject immediately before the exposure to the mutagen,

b) according to the clamp/dPCR combination assay, determining the rate of mutations in the target sequence within the target SINE or target LINE and/or in the genome of the subject after the exposure to the mutagen, and

c) comparing the rates of mutations in the target sequence within the target SINE or target LINE and/or the genome of the subject before and after the exposure to the mutagen to determine the effect of the exposure to the mutagen on the risk of cancer development and/or the rate of aging of the subject.

A further embodiment of the invention provides a kit comprising reagents to carry out the clamp/dPCR assay of the invention. In one embodiment, the kit comprises primers and/or probes specific for a SINE of interest in a species of interest. The kit can also comprise chemicals for treating the tissue or the genomic DNA sample obtained from the subject, for example, deproteination, degradation of non-DNA nucleotides, removal of other impurities, *etc*. The kit can further contain reagents and/or instrumentation for fractionating the genomic DNA into fragments of a desired size. Additionally, the kit can contain reagents and/or instrumentation for conducting the dPCR reaction. A manual containing instructions to carry out various methods of the invention can also be included in the kit.

Table 5. LINE1 sequences for humans (Homo sapiens)

1. Homo sapiens SLCO1B3 gene for solute carrier organic anion transporter family, member 1B3, partial cds, exon 7 is excluded due to the insertion of LINE1

7,685 bp linear DNA

AB896715.1 GI:1108831814

2. Homo sapiens partial LINE1 retrotransposon, clone HS3_6A

573 bp linear DNA

LT593639.1 GI:1127900357

3. Homo sapiens partial LINE1 retrotransposon, clone HS3_5H

574 bp linear DNA

LT593638.1 GI:1127900354

4. Homo sapiens partial LINE1 retrotransposon, clone HS3_5G

570 bp linear DNA

LT593637.1 GI:1127900351

5. Homo sapiens partial LINE1 retrotransposon, clone HS3_5F

566 bp linear DNA

LT593636.1 GI:1127900349

6. Homo sapiens partial LINE1 retrotransposon, clone HS3_5E
573 bp linear DNA
LT593635.1 GI:1127900342
7. Homo sapiens partial LINE1 retrotransposon, clone HS3_5D
570 bp linear DNA
LT593634.1 GI:1127900339
8. Homo sapiens partial LINE1 retrotransposon, clone HS3_5C
565 bp linear DNA
LT593633.1 GI:1127900335
9. Homo sapiens partial LINE1 retrotransposon, clone HS3_5B
573 bp linear DNA
LT593632.1 GI:1127900331
10. Homo sapiens partial LINE1 retrotransposon, clone HS3_5A
570 bp linear DNA
LT593631.1 GI:1127900327
11. Homo sapiens partial LINE1 retrotransposon, clone HS3_4G
567 bp linear DNA
LT593630.1 GI:1127900323
12. Homo sapiens partial LINE1 retrotransposon, clone HS3_4D
573 bp linear DNA
LT593629.1 GI:1127900320
13. Homo sapiens partial LINE1 retrotransposon, clone HS3_4C
576 bp linear DNA
LT593628.1 GI:1127900317
14. Homo sapiens partial LINE1 retrotransposon, clone HS3_4B
571 bp linear DNA
LT593627.1 GI:1127900313
15. Homo sapiens partial LINE1 retrotransposon, clone HS3_4A
573 bp linear DNA
LT593626.1 GI:1127900309
16. Homo sapiens partial LINE1 retrotransposon, clone HS3_3G
573 bp linear DNA
LT593625.1 GI:1127900304

17. Homo sapiens partial LINE1 retrotransposon, clone HS3_3F
573 bp linear DNA
LT593624.1 GI:1127900300
18. Homo sapiens partial LINE1 retrotransposon, clone HS3_3E
573 bp linear DNA
LT593623.1 GI:1127900296
19. Homo sapiens partial LINE1 retrotransposon, clone HS3_3D
573 bp linear DNA
LT593622.1 GI:1127900292
20. Homo sapiens partial LINE1 retrotransposon, clone HS3_3B
573 bp linear DNA
LT593621.1 GI:1127900288
21. Homo sapiens partial LINE1 retrotransposon, clone HS3_3A
572 bp linear DNA
LT593620.1 GI:1127900283
22. Homo sapiens partial LINE1 retrotransposon, clone HS3_2H
565 bp linear DNA
LT593619.1 GI:1127900278
23. Homo sapiens partial LINE1 retrotransposon, clone HS3_2G
579 bp linear DNA
LT593618.1 GI:1127900272
24. Homo sapiens partial LINE1 retrotransposon, clone HS3_2F
572 bp linear DNA
LT593617.1 GI:1127900268
25. Homo sapiens partial LINE1 retrotransposon, clone HS3_2E
573 bp linear DNA
LT593616.1 GI:1127900264
26. Homo sapiens partial LINE1 retrotransposon, clone HS3_2D
573 bp linear DNA
LT593615.1 GI:1127900259
27. Homo sapiens partial LINE1 retrotransposon, clone HS3_2C
575 bp linear DNA
LT593614.1 GI:1127900255

28. Homo sapiens partial LINE1 retrotransposon, clone HS3_2B
571 bp linear DNA
LT593613.1 GI:1127900250
29. Homo sapiens partial LINE1 retrotransposon, clone HS3_2A
573 bp linear DNA
LT593612.1 GI:1127900245
30. Homo sapiens partial LINE1 retrotransposon, clone HS3_1F
573 bp linear DNA
LT593611.1 GI:1127900241
31. Homo sapiens partial LINE1 retrotransposon, clone HS3_1E
572 bp linear DNA
LT593610.1 GI:1127900237
32. Homo sapiens partial LINE1 retrotransposon, clone HS3_1D
569 bp linear DNA
LT593609.1 GI:1127900234
33. Homo sapiens partial LINE1 retrotransposon, clone HS3_1B
573 bp linear DNA
LT593608.1 GI:1127900229
34. Homo sapiens partial LINE1 retrotransposon, clone HS3_1A
573 bp linear DNA
LT593607.1 GI:1127900225
35. Homo sapiens partial LINE1 retrotransposon, clone HS2_6H
572 bp linear DNA
LT593606.1 GI:1127900221
36. Homo sapiens partial LINE1 retrotransposon, clone HS2_6G
574 bp linear DNA
LT593605.1 GI:1127900217
37. Homo sapiens partial LINE1 retrotransposon, clone HS2_6F
573 bp linear DNA
LT593604.1 GI:1127900214
38. Homo sapiens partial LINE1 retrotransposon, clone HS2_6D
571 bp linear DNA
LT593603.1 GI:1127900210

39. Homo sapiens partial LINE1 retrotransposon, clone HS2_6A
564 bp linear DNA
LT593602.1 GI:1127900209
40. Homo sapiens partial LINE1 retrotransposon, clone HS2_5H
569 bp linear DNA
LT593601.1 GI:1127900206
41. Homo sapiens partial LINE1 retrotransposon, clone HS2_5G
567 bp linear DNA
LT593600.1 GI:1127900202
42. Homo sapiens partial LINE1 retrotransposon, clone HS2_5F
568 bp linear DNA
LT593599.1 GI:1127900199
43. Homo sapiens partial LINE1 retrotransposon, clone HS2_5D
572 bp linear DNA
LT593598.1 GI:1127900195
44. Homo sapiens partial LINE1 retrotransposon, clone HS2_5C
573 bp linear DNA
LT593597.1 GI:1127900191
45. Homo sapiens partial LINE1 retrotransposon, clone HS2_5B
573 bp linear DNA
LT593596.1 GI:1127900187
46. Homo sapiens partial LINE1 retrotransposon, clone HS2_5A
569 bp linear DNA
LT593595.1 GI:1127900183
47. Homo sapiens partial LINE1 retrotransposon, clone HS2_4G
573 bp linear DNA
LT593594.1 GI:1127900178
48. Homo sapiens partial LINE1 retrotransposon, clone HS2_4F
575 bp linear DNA
LT593593.1 GI:1127900173
49. Homo sapiens partial LINE1 retrotransposon, clone HS2_4E
573 bp linear DNA
LT593592.1 GI:1127900167

50. Homo sapiens partial LINE1 retrotransposon, clone HS2_4D
573 bp linear DNA
LT593591.1 GI:1127900162
51. Homo sapiens partial LINE1 retrotransposon, clone HS2_4C
568 bp linear DNA
LT593590.1 GI:1127900156
52. Homo sapiens partial LINE1 retrotransposon, clone HS2_4B
572 bp linear DNA
LT593589.1 GI:1127900153
53. Homo sapiens partial LINE1 retrotransposon, clone HS2_4A
572 bp linear DNA
LT593588.1 GI:1127900148
54. Homo sapiens partial LINE1 retrotransposon, clone HS2_3F
567 bp linear DNA
LT593587.1 GI:1127900146
55. Homo sapiens partial LINE1 retrotransposon, clone HS2_3D
573 bp linear DNA
LT593586.1 GI:1127900142
56. Homo sapiens partial LINE1 retrotransposon, clone HS2_3C
573 bp linear DNA
LT593585.1 GI:1127900139
57. Homo sapiens partial LINE1 retrotransposon, clone HS2_3B
561 bp linear DNA
LT593584.1 GI:1127900135
58. Homo sapiens partial LINE1 retrotransposon, clone HS2_3A
573 bp linear DNA
LT593583.1 GI:1127900132
59. Homo sapiens partial LINE1 retrotransposon, clone HS2_2G
573 bp linear DNA
LT593582.1 GI:1127900127
60. Homo sapiens partial LINE1 retrotransposon, clone HS2_2F
573 bp linear DNA
LT593581.1 GI:1127900123

61. Homo sapiens partial LINE1 retrotransposon, clone HS2_2E
573 bp linear DNA
LT593580.1 GI:1127900119
62. Homo sapiens partial LINE1 retrotransposon, clone HS2_2D
574 bp linear DNA
LT593579.1 GI:1127900115
63. Homo sapiens partial LINE1 retrotransposon, clone HS2_2C
573 bp linear DNA
LT593578.1 GI:1127900110
64. Homo sapiens partial LINE1 retrotransposon, clone HS2_2B
569 bp linear DNA
LT593577.1 GI:1127900108
65. Homo sapiens partial LINE1 retrotransposon, clone HS2_1H
573 bp linear DNA
LT593576.1 GI:1127900104
66. Homo sapiens partial LINE1 retrotransposon, clone HS2_1G
572 bp linear DNA
LT593575.1 GI:1127900101
67. Homo sapiens partial LINE1 retrotransposon, clone HS2_1F
574 bp linear DNA
LT593574.1 GI:1127900097
68. Homo sapiens partial LINE1 retrotransposon, clone HS2_1E
570 bp linear DNA
LT593573.1 GI:1127900095
69. Homo sapiens partial LINE1 retrotransposon, clone HS2_1D
569 bp linear DNA
LT593572.1 GI:1127900090
70. Homo sapiens partial LINE1 retrotransposon, clone HS2_1C
573 bp linear DNA
LT593571.1 GI:1127900087
71. Homo sapiens partial LINE1 retrotransposon, clone HS2_1B
573 bp linear DNA
LT593570.1 GI:1127900083

72. Homo sapiens partial LINE1 retrotransposon, clone HS1_5G
571 bp linear DNA
LT593569.1 GI:1127900079
73. Homo sapiens partial LINE1 retrotransposon, clone HS1_5D
573 bp linear DNA
LT593568.1 GI:1127900075
74. Homo sapiens partial LINE1 retrotransposon, clone HS1_5B
573 bp linear DNA
LT593567.1 GI:1127900072
75. Homo sapiens partial LINE1 retrotransposon, clone HS1_5A
573 bp linear DNA
LT593566.1 GI:1127900069
76. Homo sapiens partial LINE1 retrotransposon, clone HS1_4H
572 bp linear DNA
LT593565.1 GI:1127900066
77. Homo sapiens partial LINE1 retrotransposon, clone HS1_4G
573 bp linear DNA
LT593564.1 GI:1127900061
78. Homo sapiens partial LINE1 retrotransposon, clone HS1_4F
572 bp linear DNA
LT593563.1 GI:1127900057
79. Homo sapiens partial LINE1 retrotransposon, clone HS1_4E
599 bp linear DNA
LT593562.1 GI:1127900054
80. Homo sapiens partial LINE1 retrotransposon, clone HS1_4D
573 bp linear DNA
LT593561.1 GI:1127900051
81. Homo sapiens partial LINE1 retrotransposon, clone HS1_4C
568 bp linear DNA
LT593560.1 GI:1127900048
82. Homo sapiens partial LINE1 retrotransposon, clone HS1_4B
567 bp linear DNA
LT593559.1 GI:1127900044

83. Homo sapiens partial LINE1 retrotransposon, clone HS1_4A
573 bp linear DNA
LT593558.1 GI:1127900041
84. Homo sapiens partial LINE1 retrotransposon, clone HS1_3H
573 bp linear DNA
LT593557.1 GI:1127900037
85. Homo sapiens partial LINE1 retrotransposon, clone HS1_3G
569 bp linear DNA
LT593556.1 GI:1127900033
86. Homo sapiens partial LINE1 retrotransposon, clone HS1_3F
573 bp linear DNA
LT593555.1 GI:1127900029
87. Homo sapiens partial LINE1 retrotransposon, clone HS1_3D
573 bp linear DNA
LT593554.1 GI:1127900028
88. Homo sapiens partial LINE1 retrotransposon, clone HS1_2H
573 bp linear DNA
LT593553.1 GI:1127900024
89. Homo sapiens partial LINE1 retrotransposon, clone HS1_2G
569 bp linear DNA
LT593552.1 GI:1127900021
90. Homo sapiens partial LINE1 retrotransposon, clone HS1_2F
573 bp linear DNA
LT593551.1 GI:1127900017
91. Homo sapiens partial LINE1 retrotransposon, clone HS1_2D
573 bp linear DNA
LT593550.1 GI:1127900013
92. Homo sapiens partial LINE1 retrotransposon, clone HS1_2C
569 bp linear DNA
LT593549.1 GI:1127900009
93. Homo sapiens partial LINE1 retrotransposon, clone HS1_2B
570 bp linear DNA
LT593548.1 GI:1127900005

94. Homo sapiens partial LINE1 retrotransposon, clone HS1_2A
573 bp linear DNA
LT593547.1 GI:1127900002
95. Homo sapiens partial LINE1 retrotransposon, clone HS1_1H
573 bp linear DNA
LT593546.1 GI:1127899997
96. Homo sapiens partial LINE1 retrotransposon, clone HS1_1F
573 bp linear DNA
LT593545.1 GI:1127899992
97. Homo sapiens partial LINE1 retrotransposon, clone HS1_1E
573 bp linear DNA
LT593544.1 GI:1127899988
98. Homo sapiens partial LINE1 retrotransposon, clone HS1_1D
573 bp linear DNA
LT593543.1 GI:1127899984
99. Homo sapiens partial LINE1 retrotransposon, clone HS1_1C
573 bp linear DNA
LT593542.1 GI:1127899980
100. Homo sapiens partial LINE1 retrotransposon, clone HS1_1B
573 bp linear DNA
LT593541.1 GI:1127899975
101. Homo sapiens partial LINE1 retrotransposon, clone HS1_1A
573 bp linear DNA
LT593540.1 GI:1127899972
102. Homo sapiens LINE1 type transposase domain containing 1 (L1TD1), transcript variant 2, mRNA
3,894 bp linear mRNA
NM_019079.4 GI:258679512
103. Homo sapiens LINE1 type transposase domain containing 1 (L1TD1), transcript variant 1, mRNA
4,005 bp linear mRNA
NM_001164835.1 GI:258679510

104. Homo sapiens transcribed RNA, LINE1 antisense promoter (L1ASP)-derived chimeric transcript, clone: LCT26-SP6

136 bp linear transcribed-RNA

AB755817.1 GI:457866444

105. Homo sapiens transcribed RNA, LINE1 antisense promoter (L1ASP)-derived chimeric transcript, clone: LCT25-SP6

195 bp linear transcribed-RNA

AB755816.1 GI:457866443

106. Homo sapiens transcribed RNA, LINE1 antisense promoter (L1ASP)-derived chimeric transcript, clone: LCT24-SP6

312 bp linear transcribed-RNA

AB755815.1 GI:457866442

107. Homo sapiens transcribed RNA, LINE1 antisense promoter (L1ASP)-derived chimeric transcript, clone: LCT22-SP6

338 bp linear transcribed-RNA

AB755814.1 GI:457866441

108. Homo sapiens LINE1 element inserted in B-globin gene intron 2

6,045 bp linear DNA

AF149422.1 GI:5052949

109. Human LINE1 (L1.4) repetitive element DNA sequence

6,053 bp linear DNA

L19092.1 GI:307102

110. Human LINE1 (L1.4) repetitive element DNA sequence

351 bp linear DNA

L19091.1 GI:307101

111. Human LINE1 (L1.4) repetitive element DNA sequence

279 bp linear DNA

L19090.1 GI:307100

112. Human LINE1 (L1.4) repetitive element DNA sequence

389 bp linear DNA

L19089.1 GI:307099

113. Human LINE1 (L1.3) repetitive element DNA sequence

6,059 bp linear DNA

L19088.1 GI:307098

114. Human LINE1 (L1.3) repetitive element DNA sequence

139 bp linear DNA

L19087.1 GI:307097

115. Human LINE1 (L1.3) repetitive element DNA sequence

541 bp linear DNA

L19086.1 GI:307096

116. Homo sapiens 5' flanking region of LINE1 element within the apo(a)-plasminogen intergenic region

543 bp linear DNA

AJ316226.2 GI:40644101

117. Homo sapiens fragile X mental retardation syndrome protein (FMR1) gene, alternative splice products, complete cds; and pseudogene, complete sequence

185,775 bp linear DNA

L29074.1 GI:1668818

118. Homo sapiens cosmid HGAB from chromosome 13, complete sequence

36,188 bp linear DNA

AC002982.1 GI:2443900

119. Homo sapiens germline beta T-cell receptor locus

684,973 bp linear DNA

L36092.2 GI:38492353

120. Human cosmid insert containing polymorphic marker DXS455

38,409 bp linear DNA

L31948.1 GI:473756

121. Homo sapiens chromosome 4 Morf4 protein (MORF4) gene, complete cds

4,370 bp linear DNA

AF100614.2 GI:6960302

122. Homo sapiens chromosome 15 clone RP11-947O24 map 15q21.1, complete sequence

188,439 bp linear DNA

AC025919.8 GI:19071579

123. Homo sapiens serine/threonine kinase 32B (STK32B), RefSeqGene on chromosome 4

456,679 bp linear DNA

NG_051593.1 GI:1067605104

124. Homo sapiens serine/threonine kinase 32B (STK32B), transcript variant 3, mRNA

3,420 bp linear mRNA

NM_001345969.1 GI:1066566378

125. Homo sapiens serine/threonine kinase 32B (STK32B), transcript variant 2, mRNA

3,639 bp linear mRNA

NM_001306082.1 GI:807045872

126. Homo sapiens serine/threonine kinase 32B (STK32B), transcript variant 1, mRNA

3,529 bp linear mRNA

NM_018401.2 GI:807045869

127. Homo sapiens long intergenic non-protein coding RNA 1587 (LINC01587), transcript variant 2, long non-coding RNA

888 bp linear ncRNA, lncRNA

NR_126518.1 GI:723802116

128. Homo sapiens long intergenic non-protein coding RNA 1587 (LINC01587), transcript variant 3, long non-coding RNA

797 bp linear ncRNA, lncRNA

NR_126519.1 GI:723802114

129. Homo sapiens long intergenic non-protein coding RNA 1587 (LINC01587), transcript variant 1, long non-coding RNA

909 bp linear ncRNA, lncRNA

NR_126517.1 GI:723802112

130. Homo sapiens insertion sequence TMF1, complete sequence

5,839 bp linear DNA

KJ027511.1 GI:572609680

131. Homo sapiens DNA, replication enhancing element (REE1)

10,199 bp linear DNA

D50561.1 GI:1167504

132. Homo sapiens transposon L1PMA2 5' LINE1-like repetitive element

795 bp linear DNA

AJ426051.1 GI:22035757

133. Homo sapiens retrotransposon L1 insertion in X-linked retinitis pigmentosa locus, complete sequence

6,019 bp linear DNA

AF148856.1 GI:5070620

134. Human chromosome X cosmid, clones 196B12, 9H11 and 43H9, repeat units and sequence tagged sites

106,000 bp linear DNA

U40455.1 GI:1079752

135. Homo sapiens pre-integration site of LINE1-mediated deletion LH28 genomic sequence

876 bp linear DNA

DQ017971.1 GI:67848469

136. Homo sapiens pre-integration site of LINE1-mediated deletion LH26 genomic sequence

448 bp linear DNA

DQ017970.1 GI:67848468

137. Homo sapiens pre-integration site of LINE1-mediated deletion LH24 genomic sequence

700 bp linear DNA

DQ017969.1 GI:67848467

138. Homo sapiens pre-integration site of LINE1-mediated deletion LH19 genomic sequence

504 bp linear DNA

DQ017968.1 GI:67848466

139. Homo sapiens pre-integration site of LINE1-mediated deletion LH15 genomic sequence

199 bp linear DNA

DQ017967.1 GI:67848465

140. Homo sapiens tyrosinase related protein 1 (TYRP1) gene, complete cds
24,667 bp linear DNA
AF001295.1 GI:2735661
141. Homo sapiens son-pseudogene
7,290 bp linear DNA
X71604.1 GI:296950
142. Homo sapiens partial gene for novel KRAB protein domain, exons 1-2
1,488 bp linear DNA
AJ245586.1 GI:5730193
143. H.sapiens sequence involved in X;Y translocation
2,563 bp linear DNA
X70412.1 GI:515818
144. H.sapiens gene for immunoglobulin kappa light chain variable region A6
2,834 bp linear DNA
X71886.1 GI:434965
145. Homo sapiens clone CC36281C1C4C7 NADP-dependent retinol dehydrogenase/reductase (DHRS4) gene, exons 6.2 through 8.1
3,087 bp linear DNA
DQ149231.1 GI:73918039
146. Homo sapiens clone CA16281C4 NADP-dependent retinol dehydrogenase/reductase (DHRS4) gene, exons 6.2 through 8.1
3,087 bp linear DNA
DQ149227.1 GI:73918035
147. Homo sapiens clone CA16282C8 NADP-dependent retinol dehydrogenase/reductase (DHRS4) gene, exons 6.2 through 8.2
3,083 bp linear DNA
DQ149225.1 GI:73918033
148. Homo sapiens Yp11.2/Xp22.3 junction fragment in 46,X,t(X;Y) patient lhda
2,874 bp linear DNA
AJ309278.1 GI:14270371
149. Homo sapiens STS BAC clone L0974, SP6 end, sequence tagged site
703 bp linear DNA
AJ250574.1 GI:6688215

150. Homo sapiens clone LS535G M3 hypoxanthine phosphoribosyltransferase (hprt) 50 kb deletion mutant mRNA, partial cds, containing human LINE element
471 bp linear mRNA
U31732.1 GI:940938

151. H.sapiens sequence involved in X;Y translocation
2,469 bp linear DNA
X70413.1 GI:515034

Table 6. LINE1 sequences for mice (Mus musculus)

1. Mus musculus clone ma15 L1 retrotransposon LINE1 repeat region
204 bp linear DNA
U80266.1 GI:1737245

2. Mus musculus clone ma14 L1 retrotransposon LINE1 repeat region
207 bp linear DNA
U80265.1 GI:1737244

3. Mus musculus clone ma13 L1 retrotransposon LINE1 repeat region
187 bp linear DNA
U80264.1 GI:1737243

4. Mus musculus clone ma12 L1 retrotransposon LINE1 repeat region
201 bp linear DNA
U80263.1 GI:1737242

5. Mus musculus clone ma11 L1 retrotransposon LINE1 repeat region
197 bp linear DNA
U80262.1 GI:1737241

6. Mus musculus clone ma10 L1 retrotransposon LINE1 repeat region
201 bp linear DNA
U80261.1 GI:1737240

7. Mus musculus clone ma9 L1 retrotransposon LINE1 repeat region
206 bp linear DNA
U80260.1 GI:1737239

8. Mus musculus clone ma8 L1 retrotransposon LINE1 repeat region
205 bp linear DNA

U80259.1 GI:1737238

9. Mus musculus clone ma7 L1 retrotransposon LINE1 repeat region

205 bp linear DNA

U80258.1 GI:1737237

10. Mus musculus clone ma6 L1 retrotransposon LINE1 repeat region

206 bp linear DNA

U80257.1 GI:1737236

11. Mus musculus clone ma5 L1 retrotransposon LINE1 repeat region

206 bp linear DNA

U80256.1 GI:1737235

12. Mus musculus clone ma4 L1 retrotransposon LINE1 repeat region

205 bp linear DNA

U80255.1 GI:1737234

13. Mus musculus clone ma3 L1 retrotransposon LINE1 repeat region

199 bp linear DNA

U80254.1 GI:1737233

14. Mus musculus clone ma2 L1 retrotransposon LINE1 repeat region

204 bp linear DNA

U80253.1 GI:1737232

15. Mus musculus clone ma1 L1 retrotransposon LINE1 repeat region

194 bp linear DNA

U80252.1 GI:1737231

16. Mus musculus LINE1 repeat DNA

324 bp linear DNA

X58284.1 GI:52937

17. Mus musculus proteasome activator PA28 beta subunit (PSME2b) gene, complete cds

2,261 bp linear DNA

AF115502.1 GI:5031227

18. Mus musculus protease (prosome, macropain) activator subunit 2B (Psme2b), mRNA

1,066 bp linear mRNA

NM_001281472.1 GI:527317389

19. Mus musculus hypothetical protein RDA63 gene, complete cds

19,216 bp linear DNA

AF442737.1 GI:17298528

20. Mus musculus strain 129X1/SvJ tyrosinase locus control region

13,806 bp linear DNA

AF364302.1 GI:22506679

21. Mus musculus piwi-like RNA-mediated gene silencing 2 (Piwil2), mRNA

4,913 bp linear mRNA

NM_021308.1 GI:10946609

22. Mus musculus tudor domain containing 1 (Tdrd1), transcript variant 3, mRNA

5,187 bp linear mRNA

NM_001002241.2 GI:268607548

23. Mus musculus tudor domain containing 1 (Tdrd1), transcript variant 2, mRNA

4,887 bp linear mRNA

NM_001002240.2 GI:268607546

24. Mus musculus tudor domain containing 1 (Tdrd1), transcript variant 1, mRNA

4,914 bp linear mRNA

NM_001002238.2 GI:268607545

25. Mus musculus tudor domain containing 1 (Tdrd1), transcript variant 4, mRNA

5,047 bp linear mRNA

NM_031387.3 GI:268607543

26. Mus musculus tudor domain containing 9 (Tdrd9), mRNA

4,809 bp linear mRNA

NM_029056.1 GI:198278550

27. Mus musculus piwi-like RNA-mediated gene silencing 4 (Piwil4), mRNA

2,637 bp linear mRNA

NM_177905.3 GI:52138555

28. Mus musculus Apexp1, Cbx3p1 psudogenes for Apex nuclease, chromobox homolog 3

14,773 bp linear DNA

AB084238.2 GI:51699492

29. Mus musculus DNA, clone:pINS_hur170, insertion mutant

2,747 bp linear DNA

AB104439.1 GI:28569959

30. Mus musculus CC chemokine LEC pseudogene exons

5,728 bp linear DNA

AB018250.1 GI:4033628

31. Mus musculus ALDR gene including 5'UTR and promoter, partial

3,812 bp linear DNA

AJ009992.2 GI:7209181

32. Mus musculus beige gene, LINE 1 repetitive element

2,068 bp linear DNA

U78038.1 GI:2209022

33. M.musculus SPRR3 gene

2,474 bp linear DNA

Y09227.1 GI:3157400

All patents, patent applications, provisional applications, and publications referred to or cited herein are incorporated by reference in their entirety, including all figures and tables, to the extent they are not inconsistent with the explicit teachings of this specification.

Following are examples which illustrate procedures for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

EXAMPLE 1 – CLAMP/DPCR of ALU, A HUMAN SINE AND B1, A MOUSE SINE

Protein-coding regions of the human genome occupy only ~1.5% of the DNA, accounting for approximately 21,000 genes on the 23 chromosomes. A large component of the remaining DNA is composed of SINEs. Alu elements are the most abundant SINE in the human genome. Similarly, B1 elements are the most abundant SINE in the mouse genome. Alu elements are short with approximately 300-350 base pairs and contain a restriction enzyme site. With approximately 500,000 to 1,500,000 copies, B1 elements and Alu make up about 11% of the mouse and human genomes, respectively.

An embodiment of the invention provides assaying point mutations in 11% of the genome formed by the Alu elements. The rate of mutations in the genome-wide Alu elements can be used to obtain an accurate estimation of mutations in the genome.

Accordingly, in one embodiment, the invention provides a clamp/dPCR assay for Alu, a human SINE, to serially quantify the total Alu mutations in an organ in a subject, for example, examining 10^9 Alu loci in an hour. In another embodiment, the invention provides a method of determining the accumulated mutations in Alu. In an even further embodiment, the invention provides a method of using the accumulated mutations in conjunction with chronological age and Surveillance, Epidemiology and End Results (SEER) cancer statistics to quantitatively predict cancer risk (Fig. 1).

The clamp/dPCR assay of the invention can detect 1-2 mutant DNA fragments in a pool of 100,000 wild-type fragments (Fig. 6), and quantitatively differentiate 1-2 mutations from 5-10 mutations in a pool of 100,000 wild-type fragments. An appropriately chosen target SINE, for example, a SINE sequence that is about 10% prevalent in the genomic DNA and is about 300-400 bp, assures that in a mixture of genomic DNA fragments more than 10% will contain the target SINE. Accordingly, the prevalence and short lengths of Alu and B1 assure that in a mix of genomic DNA fragments from a human and a mouse, respectively, more than 10% will contain the sequence.

In a 45-min cycle of a dPCR, for example, ddPCR (BioRad QX200 AutoDG ddPCR, Hercules, CA), over 10^6 DNA fragments can be analyzed in each of the 8 channels for the presence of mutations. The clamp/dPCR assay of the invention allows using 100 or 1,000 fragments per drop instead of ~1-2 fragments per drop because the target sequence clamp prevents amplification in most of the target SINE sequences that are likely to be wild-type. Since about 10% of those DNA fragments likely contain the target SINE, ~10^5 target SINEs are analyzed at one fragment per well or drop. Therefore, the clamp/dPCR combination assay of the invention improves the screening of Alu fragments by 2 or 3 orders of magnitude, *i.e.*, up to 10^8 Alu/channel. Assuming 10^{-6} mutations per cell division and/or per week of age, the assay has the capacity to estimate the rate of mutations in genome-wide Alu sequences.

Target sequence clamps for mouse B1 and human Alu with base sizes of 16-20 are provided. In a test run, excess idealized wild-type B1 with a trace of mutant B1 was tested using a 16-base Clamp2 (SEQ ID NO: 4) at the dilutions of 1:1000 and 1:10,000. Additional clamps can be designed, and alleles for any clamp can be prepared and multiplexed to eliminate common variations in SINE sequences.

According to an embodiment, any tissue can be tested with the clamp/dPCR combination assay. For example, in humans, a skin or mucosal scrape or 1-2 μl of blood is

sufficient. Since reagents used in routine PCR are used, the assay provides high quality and reproducibility. Therefore, an assay can be repeated economically in the same subject or organ.

EXAMPLE 2 – ESTIMATION OF GENOME-WIDE MUTATIONS AFTER PARTICLE IRRADIATION IN VARIOUS TISSUES

Like natural age-related genome-wide accumulated mutations and/or rate of mutations, radiation-related genome-wide accumulated mutations and/or rate of mutations occur at different rates in individuals and directly lead to different risks and rates for cancer. Radiation-induced mutations should be random with respect to their distribution across the genome. Although some location-specific effects likely arise due to repair mechanisms, testing for mutations near specific genes is likely to be futile. On the other hand, overall mutation load, *e.g.*, the genome-wide accumulated mutations and/or the rate of mutations and the incidence of driver mutations have a linear relationship.

A dose- and LET-dependent increase in genome-wide mutations is expected. Basal and serial studies can be performed following ^{1}H, ^{1}n, ^{28}Si, or ^{56}Fe irradiation (dose $\leq$ 0.5 Gy). Both sexes and individual organs can be studied. Several assays can be designed based on Table 7.

Table 7: Assay designs

Hypothesis	Assay design
Age, radiation dose, and LET affect the genome-wide mutations.	Clamp/dPCR combination assay for B1 sequence in mice and Alu sequence in humans after particle radiation with, for example, ^{1}n, ^{1}H, ^{28}Si, and ^{56}Fe.
Genome-wide mutations in spontaneous tumors can be similar or higher than the genome-wide mutations in normal tissues.	Clamp/dPCR combination assay for B1 in mice with spontaneous tumors. Tumor and normal tissue can be examined from the same subject and same organ.

Point and indel mutations increase with aging and radiation of different qualities and doses affect the rate of mutations. Point and indel mutation comprises over 90-95% of all mutations. The test provided in this example is simple and inexpensive and requires only 1-2

ng of DNA. As such, it can be performed on a number of mouse strains, on each individual animal and in different organs. Strains representing a range of cancer predilections (including sex-related cancers, such as breast and ovary) can also be studied. A preliminary evaluation of the various potential clamp sites can be made to determine the most robust set of target sequence clamps for use in an animal of interest, for example, a mouse strain.

Various tissues, including blood, muscle, brain, heart, lung, skin, breast, large bowel, liver, and spleen, can be tested for the effect of radiation. Because cancer rates increase in progeny after high LET, testes and ovarian tissues can be tested to evaluate germline genome-wide mutation levels. These organs are chosen because these organs are all known to have cancer predispositions following radiation exposure or, like skin, might sustain the highest radiation exposure.

Non-limiting examples of specific organs which can be tested according to this example of the invention include skin, lung, breast, and WBC. Skin receives a higher exposure than most organs and leukemia is common after irradiation. Also, WBC genome-wide mutations are needed for human comparison and lung and breast tumors are relatively common in mice and humans. These tissues are particularly preferred to study the effects of radiation.

The test can be carried on tissues obtained at 0 hours, 24 hours, 1 month, 6 months, 1 year, 2 to 5 years or longer after exposure to radiation. These analyses can be done on individual mice for animal-specific organ comparisons. Thus, intra- and inter-strain comparisons to compute the difference between the genomic age and the chronological age (Δ_{age}) can be calculated.

EXAMPLE 3 – MEASUREMENTS OF GENOME-WIDE MUTATIONS IN SPONTANEOUS TUMORS AND NORMAL TISSUE

The number of the genome-wide mutations in a spontaneous tumor is similar to or larger than the number of these mutations in the normal tissue. A spontaneous tumor refers to a tumor which arises in a subject that is not exposed to known carcinogens or tumor-promoting factors, *e.g.*, ionizing radiation, mutagens, oncogenic viruses, *etc.*

Tumors and the source tissue can be examined from the same subject. NIH Swiss white mice with a female to male ratio of 1:2 can be used. Having fewer females is also logical as breast and ovarian cancers are common in this strain, leading to good representation of females in the final tumor population. This strain has a ~10-20%

cumulative lifetime risk of malignancy, with lung > ovary > breast > leukemia > sarcoma > gastrointestinal (GI) cancers. The GI cancers include an even mix of stomach, colon and liver.

Genetically defined animals with cancer predilection can also be used. For example, the *Kras*^{LA1} model of lung cancer, the *Apc* heterozygote knockout model for GI cancers, as well as other cancer-predisposed models featuring *Trp53*$^{-/-}$ can be used. However, as most of these mice develop cancer without radiation and simply demonstrate shorter latency or more aggressive pathology when irradiated, they are good models for radiation-induced cancer progression but less effective for emulating spontaneous carcinogenesis. Therefore, the strain chosen for studying passenger DNA damage (PDD) in spontaneous tumors should be a typically healthy strain.

Spontaneous oncogenesis studies can be long-term, with latency to cancer of 300-800 days in mice, and can involve large and laborious animal cohorts as the lifetime risks are only 10-20% in non-irradiated and 15-30% in irradiated animals. Certain algorithms can be used which do not require validation by correlation with the rate of cancers in animals, thereby allowing the use of smaller cohorts. Also, algorithms can be used which require the accumulation of DNA damage after irradiation to be allometrically scalable between species so that genomic age and Δ_{age} can be calculated. Accordingly, human epidemiological statistics can be applied to predict driver frequency and human cancer risk.

EXAMPLE 4 – Measurement of PDD in normal human aging

PDD mutations increase with normal human aging processes and can progress in different subjects at different rates. Subjects in the various age groups/ranges can be studied. PDD mutations are expected to correlate with increasing age.

Framework of PDD and Cancer Risk:

Subject-specific rates of genomic aging as markers for cancer risk can be calculated. Genomic aging can be quantified through serial measurements of point and indel mutations. A subject's genomic aging status can be tracked against age-related cancer incidence trajectories at the population level for risk estimation. The rate of PDD accumulation is a primary quantity of interest. Germline variations can be implicitly adjusted to confound the risk estimates. Specifically, at birth, there exist some number of germline abnormalities (point and indels) in each cell in the body. Notably, the overwhelming majority of germline abnormalities behave as passengers and do not confer heightened cancer risk. A subject's

neonatal blood can be used as the subject's mutation-free state, which can be used to fully disentangle PDD accumulation over the subject's lifetime from benign germline variations. However, sequential measurements of PDD over a period can be used to estimate rates of future damage accumulation that can be compared against population averages to determine relative risk for a subject.

Established age-specific cancer incidence data which show that most cancers follow a power law in chronological age can be used as a frame of reference. In particular, incidence curves are commonly observed as $I(t) \propto t^k$, where t and I(t) respectively denote chronological age and age-specific incidence. The constant k is estimated from population data and is specific to cancer type, sex, race, and other epidemiologic factors. Appropriate functional forms for population subgroups can be determined using the SEER data.

The occurrence of PDD is proportional to the occurrence of driver mutations and, hence, proportional to cancer incidence. Therefore, $\frac{dPDD}{dt} = ct^{k-1}$, where c is a constant (Fig. 7). Conveniently, the parameter of interest, c, may be estimated from sequential measurements for each subject, since $\frac{dPDD}{dt}$ can be calculated as a simple difference, and k and t have been derived from the population analysis. This analysis does not require an assumption-laden back-calculation to cellular abnormalities at birth. In short, subjects with large estimated values of c will be identified as having a high genomic age and an elevated future risk of cancer, since their high mutation accumulation rate implies a history of harmful exposure and an enhanced likelihood of driver mutations.

Genomic aging can predict adverse responses to sustained low-dose irradiation. Flexible regression strategies (*e.g.*, spline fitting) will be used to optimally link molecular markers of radiation sensitivity to PDD. Data from the non-irradiated animals can be paired with analogous data from human subjects to permit allometric scaling of marker quantities from mouse to human. The rate of PDD accumulation at age *t* and also PDD itself at age *t* through the system of equations specified separately for chronological age can be estimated. PDD is expected to represent a more comprehensive measure of accumulated DNA damage.

It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and the scope of the appended claims. In addition, any elements or

limitations of any invention or embodiment thereof disclosed herein can be combined with any and/or all other elements or limitations (individually or in any combination) or any other invention or embodiment thereof disclosed herein, and all such combinations are contemplated within the scope of the invention without limitation thereto.

REFERENCES

1. Bozic I, Antal T, Ohtsuki H, Carter H, Kim D, Chen S, Karchin R, Kinzler KW, Vogelstein B, Nowak MA. Accumulation of driver and passenger mutations during tumor progression. *Proc Natl Acad Sci U S A.* 2010;107(43):18545-50.

2. Tomasetti C, Marchionni L, Nowak MA, Parmigiani G, Vogelstein B. Only three driver gene mutations are required for the development of lung and colorectal cancers. *Proc Natl Acad Sci U S A.* 2015;112(1):118-23.

3. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., Kinzler KW. Cancer genome landscapes. *Science.* 2013;339(6127):1546-58.

4. Surveillance, Epidemiology, and End Results Program. Citations for SEER Databases and SEER*State Software. 2015. Available from: http://seer.cancer.gov/data/citation.html.

5. Teschendorff AE, West J, Beck S. Age-associated epigenetic drift: implications, and a case of epigenetic thrift? *Hum Mol Genet.* 2013;22(R1):R7-R15.

6. Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H, Whittaker P, McCann OT, Finer S, Valdes AM, Leslie RD, Deloukas P, Spector TD. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res.* 2010;20(4):434-9.

7. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. *Nat Rev Genet.* 2009;10(10):691-703.

8. Chen JC, Alvarez MJ, Talos F, Dhruv H, Rieckhof GE, Iyer A, Diefes KL, Aldape K, Berens M, Shen MM, Califano A. Identification of causal genetic drivers of human disease through systems-level analysis of regulatory networks. *Cell.* 2014;159(2):402-14.

9. Vandin F, Upfal E, Raphael BJ. Finding driver pathways in cancer: models and algorithms. *Algorithms Mol Biol.* 2012;7(1):23.

10. Rodriguez-Brenes IA, Komarova NL, Wodarz D. Evolutionary dynamics of feedback escape and the development of stem-cell-driven cancers. *Proc Natl Acad Sci U S A.* 2011;108(47):18983-8.

11. Rithidech KN, Honikel L, Whorton EB. mFISH analysis of chromosomal damage in bone marrow cells collected from CBA/CaJ mice following whole body exposure to heavy ions (56Fe ions). *Radiat Environ Biophys.* 2007;46(2):137-45.

12. Tucker JD, Marples B, Ramsey MJ, Lutze-Mann LH. Persistence of chromosome aberrations in mice acutely exposed to 56Fe+26 ions. *Radiat Res.* 2004;161(6):648-55.

13. Little MP, Kwon D, Doi K, Simon SL, Preston DL, Doody MM, Lee T, Miller JS, Kampa DM, Bhatti P, Tucker JD, Linet MS, Sigurdson AJ. Association of chromosome translocation rate with low dose occupational radiation exposures in U.S. radiologic technologists. *Radiat Res.* 2014;182(1):1-17.

14. Vorobtsova IE, Tucker JD, Timofeeva NM, Bogomazova AN, Semenov AV, Pleshanov PG. [Effect of age and radiation exposure on the frequency of translocations and dicentrics detected by FISH method in human lymphocytes]. *Radiat Biol.* 2000;40(2):142-8.

15. Lindholm C, Makelainen I, Paile W, Koivistoinen A, Salomaa S. Domestic radon exposure and the frequency of stable or unstable chromosomal aberrations in lymphocytes. *Int J Radiat Biol.* 1999;75(8):921-8.

16. Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, Mangino M, Zhai G, Zhang F, Valdes A, Shin SY, Dempster EL, Murray RM, Grundberg E, Hedman AK, Nica A, Small KS, Dermitzakis ET, McCarthy MI, Mill J, Spector TD, Deloukas P. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet.* 2012;8(4):e1002629.

17. Bollati V, Schwartz J, Wright R, Litonjua A, Tarantini L, Suh H, Sparrow D, Vokonas P, Baccarelli A. Decline in genomic DNA methylation through aging in a cohort of elderly subjects. *Mech Ageing Dev.* 2009;130(4):234-9.

18. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL, Nelson HH, Karagas MR, Padbury JF, Bueno R, Sugarbaker DJ, Yeh RF, Wiencke JK, Kelsey KT. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. *PLoS Genet.* 2009;5(8):e1000602.

19. Fraga MF, Esteller M. Epigenetics and aging: the targets and the marks. *Trends Genet.* 2007;23(8):413-8.

20. Rodriguez-Rodero S, Fernandez-Morera JL, Fernandez AF, Menendez-Torre E, Fraga MF. Epigenetic regulation of aging. *Discov Med.* 2010;10(52):225-33.

21. Mathers JC, Strathdee G, Relton CL. Induction of epigenetic alterations by dietary and other environmental factors. *Adv Genet.* 2010;71:3-39.

22. Bell JT, Spector TD. DNA methylation studies using twins: what are they telling us? *Genome Biol.* 2012;13(10):172.

23. Kilpinen H, Dermitzakis ET. Genetic and epigenetic contribution to complex traits. *Hum Mol Genet.* 2012;21(R1):R24-8.

24. Leonhardt H, Page AW, Weier HU, Bestor TH. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell.* 1992;71(5):865-73.

25. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, Klotzle B, Bibikova M, Fan JB, Gao Y, Deconde R, Chen M, Rajapakse I, Friend S, Ideker T, Zhang K. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell.* 2013;49(2):359-67.

26. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol.* 2013;14(10):R115.

27. Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, Gibson J, Henders AK, Redmond P, Cox SR, Pattie A, Corley J, Murphy L, Martin NG, Montgomery GW, Feinberg AP, Fallin MD, Multhaup ML, Jaffe AE, Joehanes R, Schwartz J, Just AC, Lunetta KL, Murabito JM, Starr JM, Horvath S, Baccarelli AA, Levy D, Visscher PM, Wray NR, Deary IJ. DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biol.* 2015;16:25.

28. West J, Beck S, Wang X, Teschendorff AE. An integrative network algorithm identifies age-associated differential methylation interactome hotspots targeting stem-cell differentiation pathways. *Sci Rep.* 2013;3:1630.

29. Maegawa S, Hinkal G, Kim HS, Shen L, Zhang L, Zhang J, Zhang N, Liang S, Donehower LA, Issa JP. Widespread and tissue specific age-related DNA methylation changes in mice. *Genome Res.* 2010;20(3):332-40.

30. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H, Campan M, Noushmehr H, Bell CG, Maxwell AP, Savage DA, Mueller-Holzner E, Marth C, Kocjan G, Gayther SA, Jones A, Beck S, Wagner W, Laird PW, Jacobs IJ, Widschwendter M. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res.* 2010;20(4):440-6.

31. Horvath S, Zhang Y, Langfelder P, Kahn RS, Boks MP, van Eijk K, van den Berg LH, Ophoff RA. Aging effects on DNA methylation modules in human brain and blood tissue. *Genome Biol.* 2012;13(10):R97.

32. Aubert G, Lansdorp PM. Telomeres and aging. *Physiol Rev.* 2008;88(2):557-79.

33. Meissner C, Ritz-Timme S. Molecular pathology and age estimation. *Forensic Sci Int.* 2010;203(1-3):34-43.

34. Zubakov D, Liu F, van Zelm MC, Vermeulen J, Oostra BA, van Duijn CM, Driessen GJ, van Dongen JJ, Kayser M, Langerak AW. Estimating human age from T-cell DNA rearrangements. *Curr Biol.* 2010;20(22):R970-1.

35. Lin J, Epel E, Blackburn E. Telomeres and lifestyle factors: roles in cellular aging. *Mutat Res.* 2012;730(1-2):85-9.

36. Pardo CE, Nabilsi NH, Darst RP, Kladde MP. Integrated DNA methylation and chromatin structural analysis at single-molecule resolution. *Methods Mol Biol.* 2015;1288:123-41.

37. Nabilsi NH, Deleyrolle LP, Darst RP, Riva A, Reynolds BA, Kladde MP. Multiplex mapping of chromatin accessibility and DNA methylation within targeted single molecules identifies epigenetic heterogeneity in neural stem cells and glioblastoma. *Genome Res.* 2014;24(2):329-39.

38. Chen Y, Hyrien O, Williams J, Okunieff P, Smudzin T, Rubin P. Interleukin (IL)-1A and IL-6: applications to the predictive diagnostic testing of radiation pneumonitis. *Int J Radiat Oncol Biol Phys.* 2005;62(1):260-6.

39. Rubin P, Johnston CJ, Williams JP, McDonald S, Finkelstein JN. A perpetual cascade of cytokines postirradiation leads to pulmonary fibrosis. *Int J Radiat Oncol Biol Phys.* 1995;33(1):99-109.

40. Xiao Z, Yang S, Su Y, Wang W, Zhang H, Zhang M, Zhang K, Tian Y, Cao Y, Yin L, Zhang L, Okunieff P. Alteration of the inflammatory molecule network after irradiation of soft tissue. *Adv Exp Med Biol.* 2013;765:335-41.

41. Zhang M, Yin L, Zhang K, Sun W, Yang S, Zhang B, Salzman P, Wang W, Liu C, Vidyasagar S, Zhang L, Ju S, Okunieff P. Response patterns of cytokines/chemokines in two murine strains after irradiation. *Cytokine.* 2012;58(2):169-77.

42. Liu W, Ding I, Chen K, Olschowka J, Xu J, Hu D, Morrow GR, Okunieff P. Interleukin 1beta (IL1B) signaling is a critical component of radiation-induced skin fibrosis. *Radiat Res.* 2006;165(2):181-91.

43. Zhang SB, Okunieff P. Mitochondrial genetic abnormalities after radiation exposure. *Adv Exp Med Biol.* 2014;812:1-7.

44. Zhang SB, Maguire D, Zhang M, Tian Y, Yang S, Zhang A, Casey-Sawicki K, Han D, Ma J, Yin L, Guo Y, Wang X, Chen C, Litvinchuk A, Zhang Z, Swarts S, Vidyasagar S, Zhang L, Okunieff P. Mitochondrial DNA and functional investigations into the radiosensitivity of four mouse strains. *Int J Cell Biol.* 2014;2014:850460.

45. Zhang SB, Maguire D, Zhang M, Zhang Z, Zhang A, Yin L, Zhang L, Huang L, Vidyasagar S, Swarts S, Okunieff P. The murine common deletion: mitochondrial DNA 3,860-bp deletion after irradiation. *Radiat Res.* 2013;180(4):407-13.

46. Zhang SB, Maguire D, Zhang M, Zhang A, Zhang L, Swarts S, Okunieff P. Maternal bias in mouse radiosensitivity: the role of the mitochondrial PTP. *Adv Exp Med Biol.* 2013;789:251-6.

47. Zhang SB, Zhang M, Cao Y, Yang S, Zhang AP, Yin L, Tian Y, Zhang K, Guo Y, Wang X, Chen C, Ma J, Han D, Litvinchuk A, Swarts S, Vidyasagar S, Maguire D, Zhang L, Okunieff P. Delayed effects of radiation on mitochondrial DNA in radiation-sensitive organs. *Adv Exp Med Biol.* 2012;737:139-45.

48. Scanlon SE, Glazer PM. Multifaceted control of DNA repair pathways by the hypoxic tumor microenvironment. *DNA Repair (Amst).* 2015;32:180-9.

49. Luoto KR, Kumareswaran R, Bristow RG. Tumor hypoxia as a driving force in genetic instability. *Genome Integr.* 2013;4(1):5.

50. Maybury BD. Mitochondrial DNA damage is uncommon in cancer but can promote aggressive behaviour. *Anticancer Res.* 2013;33(9):3543-52.

51. Kim JS, Qian T, Lemasters JJ. Mitochondrial permeability transition in the switch from necrotic to apoptotic cell death in ischemic rat hepatocytes. *Gastroenterology.* 2003;124(2):494-503.

52. Kim JS, He L, Lemasters JJ. Mitochondrial permeability transition: a common pathway to necrosis and apoptosis. *Biochem Biophys Res Commun.* 2003;304(3):463-70.

53. Kim JS, Jin Y, Lemasters JJ. Reactive oxygen species, but not Ca2+ overloading, trigger pH- and mitochondrial permeability transition-dependent death of adult rat myocytes after ischemia-reperfusion. *Am J Physiol Heart Circ Physiol.* 2006;290(5):H2024-34.

54. Kim JS, Ohshima S, Pediaditakis P, Lemasters JJ. Nitric oxide protects rat hepatocytes against reperfusion injury mediated by the mitochondrial permeability transition. *Hepatology.* 2004;39(6):1533-43.

55. Kon K, Kim JS, Jaeschke H, Lemasters JJ. Mitochondrial permeability transition in acetaminophen-induced necrosis and apoptosis of cultured mouse hepatocytes. *Hepatology.* 2004;40(5):1170-9.

56. Kim JS, He L, Qian T, Lemasters JJ. Role of the mitochondrial permeability transition in apoptotic and necrotic death after ischemia/reperfusion injury to hepatocytes. *Curr Mol Med.* 2003;3(6):527-35.

57. Kim JS, Nitta T, Mohuczy D, O'Malley KA, Moldawer LL, Dunn WA, Jr., Behrns KE. Impaired autophagy: A mechanism of mitochondrial dysfunction in anoxic rat hepatocytes. *Hepatology.* 2008;47(5):1725-36.

58. Wang JH, Ahn IS, Fischer TD, Byeon JI, Dunn WA, Jr., Behrns KE, Leeuwenburgh C, Kim JS. Autophagy suppresses age-dependent ischemia and reperfusion injury in livers of mice. *Gastroenterology.* 2011;141(6):2188-99 e6.

59. Kim JS, Wang JH, Biel TG, Kim DS, Flores-Toro JA, Vijayvargiya R, Zendejas I, Behrns KE. Carbamazepine suppresses calpain-mediated autophagy impairment after ischemia/reperfusion in mouse livers. *Toxicol Appl Pharmacol.* 2013;273(3):600-10.

60. Biel TG, Lee S, Flores-Toro JA, Dean JW, Go KL, Lee MH, Law BK, Law ME, Dunn WA, Jr., Zendejas I, Behrns KE, Kim JS. Sirtuin 1 suppresses mitochondrial dysfunction of ischemic mouse livers in a mitofusin 2-dependent manner. *Cell Death Differ.* 2015).

61. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43(5):491-8.

62. Kilgore JA, Hoose SA, Gustafson TL, Porter W, Kladde MP. Single-molecule and population probing of chromatin structure using DNA methyltransferases. *Methods.* 2007;41(3):320-32.

63. Kelly TK, Liu Y, Lay FD, Liang G, Berman BP, Jones PA. Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res.* 2012;22(12):2497-506.

64. Pondugula S, Kladde MP. Single-molecule analysis of chromatin: changing the view of genomes one molecule at a time. *J Cell Biochem.* 2008;105(2):330-7.

65. Xu M, Kladde MP, Van Etten JL, Simpson RT. Cloning, characterization and expression of the gene coding for a cytosine-5-DNA methyltransferase recognizing GpC. *Nucl Acid Res.* 1998;26(17):3961-6.

66. Darst RP, Haecker I, Pardo CE, Renne R, Kladde MP. Epigenetic diversity of Kaposi's sarcoma-associated herpesvirus. *Nucl Acid Res.* 2013;41(5):2993-3009.

67. Blaber E, Marcal H, Burns BP. Bioastronautics: the influence of microgravity on astronaut health. *Astrobiology.* 2010;10(5):463-73.

68. Li K, Huang T, Zheng J, Wu K, Li D. Effect of marine-derived n-3 polyunsaturated fatty acids on C-reactive protein, interleukin 6 and tumor necrosis factor alpha: a meta-analysis. *PLoS One.* 2014;9(2):e88103.

69. National Council on Radiation Protection and Measurements. Radiation Protection Guidance for Activities in Low-Earth Orbit. 2000. Bethesda, MD: NCRP Report No. 132.

70. International Commission on Radiation Protection. Relative Biological Effectiveness, Radiation Weighting and Quality Factor. *Ann ICRP.* 2003;33(4).

71. Adewoye AB, Lindsay SJ, Dubrova YE, Hurles ME. The genome-wide effects of ionizing radiation on mutation induction in the mammalian germline. *Nat Commun.* 2015;6:6684.

72. Mouse Tumor Biology Database. Tumor Frequency Grid (Inbred Strain Family x Organ). 2015. Available from: http://tumor.informatics.jax.org/mtbwi/dynamicGrid.do?organKey=|60|&currentOrganKey=|60|&strainFamilyKey=&currentStrainFamilyKey=&expandColapse=false&organGrpChk=18&organGrpChk=15&organGrpChk=193&organGrpChk=199&organGrpChk=363&organGrpChk=17&organGrpChk=26&organGrpChk=245&organGrpChk=14&organGrpChk=9&organGrpChk=159&organGrpChk=31&organGrpChk=56&organGrpChk=28&organGrpChk=251&organGrpChk=58&organGrpChk=29&organGrpChk=60&organGrpChk=117&organGrpChk=66&organGrpChk=92&organGrpChk=119&organGrpChk=5&organGrpChk=348&organGrpChk=45&organGrpChk=130&organGrpChk=3&organGrpChk=25&organGrpChk=342&organGrpChk=129&organGrpChk=32&organGrpChk=157&organGrpChk=6&organGrpChk=194&organGrpChk=140&organGrpChk=11&organGrpChk=39&organGrpChk=112&organGrpChk=40&organGrpChk=120&organGrpChk=380&organGrpChk=254&organGrpChk=118&organGrpChk=4&organGrpChk=71&organGrpChk=116&organGrpChk=12&organGrpChk=95&organGrpChk=2&organGrpChk=153&organGrpChk=61&organGrpChk=530&organGrpChk=132&organGrpChk=242&organGrpChk=108&organGrpChk=269&organGrpChk=36&organGrpChk=156&organGrpChk=177&organGrpChk=94&strainFamilyChk=2&strainFamilyChk=70&strainFamilyChk=9&strainFamilyChk=12&strainFamilyChk=4&strainFamilyChk=32&strainFamilyChk=23&strainFamilyChk=43&strainFamilyChk=44&strainFamilyChk=58&strainFamilyChk=112&strainFamilyChk=86&strainFamilyChk=84&strainFamilyChk=87&strainFamilyChk=31&strainFamilyChk=30&strainFamilyChk=29&strainFamilyChk=26&strainFamilyChk=27&strainFamilyChk=28&strainFamilyChk=73&strainFamilyChk=126&strainFamilyChk=116&strainFamilyChk=54&strainFamilyChk=82&strainFamilyChk=8

3&strainFamilyChk=100&strainFamilyChk=101&strainFamilyChk=20&strainFamilyChk=36&strainFamilyChk=94&strainFamilyChk=65&strainFamilyChk=145&showAllStrains=showAllStrains.

73. Link A, Lenz M, Legner D, Bohm M, Nickenig G. Telmisartan inhibits beta2-integrin MAC-1 expression in human T-lymphocytes. *J Hypertens.* 2006;24(9):1891-8.

74. Moertl D, Hammer A, Steiner S, Hutuleac R, Vonbank K, Berger R. Dose-dependent effects of omega-3-polyunsaturated fatty acids on systolic left ventricular function, endothelial function, and markers of inflammation in chronic heart failure of nonischemic origin: a double-blind, placebo-controlled, 3-arm study. *Am Heart J.* 2011;161(5):915 e1-9.

75. Nodari S, Triggiani M, Campia U, Manerba A, Milesi G, Cesana BM, Gheorghiade M, Dei Cas L. Effects of n-3 polyunsaturated fatty acids on left ventricular function and functional capacity in patients with dilated cardiomyopathy. *J Am Coll Cardiol.* 2011;57(7):870-9.

76. Trevelyan J, Brull DJ, Needham EW, Montgomery HE, Morris A, Mattu RK. Effect of enalapril and losartan on cytokines in patients with stable angina pectoris awaiting coronary artery bypass grafting and their interaction with polymorphisms in the interleukin-6 gene. *Am J Cardiol.* 2004;94(5):564-9.

77. Pavlatou MG, Mastorakos G, Margeli A, Kouskouni E, Tentolouris N, Katsilambros N, Chrousos GP, Papassotiriou I. Angiotensin blockade in diabetic patients decreases insulin resistance-associated low-grade inflammation. *Eur J Clin Invest.* 2011;41(6):652-8.

78. Medhora M, Gao F, Wu Q, Molthen RC, Jacobs ER, Moulder JE, Fish BL. Model development and use of ACE inhibitors for preclinical mitigation of radiation-induced injury to multiple organs. *Radiat Res.* 2014;182(5):545-55.

79. Moulder JE, Cohen EP, Fish BL. Mitigation of experimental radiation nephropathy by renin-equivalent doses of angiotensin converting enzyme inhibitors. *Int J Radiat Biol.* 2014;90(9):762-8.

80. Moulder JE. 2013 Dade W. Moeller lecture: medical countermeasures against radiological terrorism. *Health Phys.* 2014;107(2):164-71.

81. Hesson LB, Patil V, Sloane MA, Nunez AC, Liu J, Pimanda JE, Ward RL. Reassembly of nucleosomes at the MLH1 promoter initiates resilencing following decitabine exposure. *PLoS Genet.* 2013;9(7):e1003636.

82. Marcus CS. Time to Reject the Linear-No Threshold Hypothesis and Accept Thresholds and Hormesis: A Petition to the U.S. Nuclear Regulatory Commission. *Clin Nucl Med.* 2015;40(7):617-9.

83. Lehrer S, Rosenzweig KE. Lung cancer hormesis in high impact states where nuclear testing occurred. *Clin Lung Cancer.* 2015;16(2):152-5.

84. Morgan WF, Bair WJ. Issues in low dose radiation biology: the controversy continues. A perspective. *Radiat Res.* 2013;179(5):501-10.

85. Varley KE, Mitra RD. Bisulfite Patch PCR enables multiplexed sequencing of promoter methylation across cancer samples. *Genome Res.* 2010;20(9):1279-87.

86. Weidner CI, Lin Q, Koch CM, Eisele L, Beier F, Ziegler P, Bauerschlag DO, Jockel KH, Erbel R, Muhleisen TW, Zenke M, Brummendorf TH, Wagner W. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biol.* 2014;15(2):R24.

87. Naugler WE, Karin M. The wolf in sheep's clothing: the role of interleukin-6 in immunity, inflammation and cancer. *Trends Mol Med.* 2008;14(3):109-19.

88. Barcellos-Hoff MH. How do tissues respond to damage at the cellular level? The role of cytokines in irradiated tissues. *Radiat Res.* 1998;150(5 Suppl):S109-20.

89. Manning CM, Johnston CJ, Reed CK, Lawrence BP, Williams JP, Finkelstein JN. Lung irradiation increases mortality after influenza A virus challenge occurring late after exposure. *Int J Radiat Oncol Biol Phys.* 2013;86(1):128-35.

90. Johnston CJ, Hernady E, Reed C, Thurston SW, Finkelstein JN, Williams JP. Early alterations in cytokine expression in adult compared to developing lung in mice after radiation exposure. *Radiat Res.* 2010;173(4):522-35.

91. Gridley DS, Dutta-Roy R, Andres ML, Nelson GA, Pecaut MJ. Acute effects of iron-particle radiation on immunity. Part II: Leukocyte activation, cytokines and adhesion. *Radiat Res.* 2006;165(1):78-87.

92. Pecaut MJ, Dutta-Roy R, Smith AL, Jones TA, Nelson GA, Gridley DS. Acute effects of iron-particle radiation on immunity. Part I: Population distributions. *Radiat Res.* 2006;165(1):68-77.

93. Dong C, He M, Tu W, Konishi T, Liu W, Xie Y, Dang B, Li W, Uchihori Y, Hei TK, Shao C. The differential role of human macrophage in triggering secondary bystander effects after either gamma-ray or carbon beam irradiation. *Cancer Lett.* 2015;363(1):92-100.

94. Chatterjee A, Borak TH. Physical and biological studies with protons and HZE particles in a NASA supported research center in radiation health. *Phys Med.* 2001;17 Suppl 1:59-66.

95. Chen Y, Williams J, Ding I, Hernady E, Liu W, Smudzin T, Finkelstein JN, Rubin P, Okunieff P. Radiation pneumonitis and early circulatory cytokine markers. *Semin Radiat Oncol.* 2002;12(1 Suppl 1):26-33.

96. Evans ES, Kocak Z, Zhou SM, Kahn DA, Huang H, Hollis DR, Light KL, Anscher MS, Marks LB. Does transforming growth factor-beta1 predict for radiation-induced pneumonitis in patients treated for lung cancer? *Cytokine.* 2006;35(3-4):186-92.

97. Zhang H, Maguire DJ, Zhang M, Zhang L, Okunieff P. Elevated mitochondrial DNA copy number and POL-gamma expression but decreased expression of TFAM in murine intestine following therapeutic dose irradiation. *Adv Exp Med Biol.* 2011;701:201-6.

98. Zhang H, Maguire D, Swarts S, Sun W, Yang S, Wang W, Liu C, Zhang M, Zhang D, Zhang L, Zhang K, Keng P, Okunieff P. Replication of murine mitochondrial DNA following irradiation. *Adv Exp Med Biol.* 2009;645:43-8.

99. Kim JS, Lemasters JJ. Opioid receptor-independent protection of ischemic rat hepatocytes by morphine. *Biochem Biophys Res Commun.* 2006;351(4):958-64.

100. Trovati S, Ballarini F, Battistoni G, Cerutti F, Fasso A, Ferrari A, Gadioli E, Garzelli MV, Mairani A, Ottolenghi A, Paretzke HG, Parini V, Pelliccioni M, Pinsky L, Sala PR, Scannicchio D, Zankl M. Human exposure to space radiation: role of primary and secondary particles. *Radiat Prot Dosim.* 2006;122(1-4):362-6.

101. Badhwar GD, O'Neill PM. Response of silicon-based linear energy transfer spectrometers: implication for radiation risk assessment in space flights. *Nucl Instrum Methods Phys Res A.* 2001;466(3):464-74.

102. Bahadori AA, Sato T, Slaba TC, Shavers MR, Semones EJ, Van Baalen M, Bolch WE. A comparative study of space radiation organ doses and associated cancer risks using PHITS and HZETRN. *Phys Med Biol.* 2013;58(20):7183-207.

103. Armstrong TW, Colborn BL. Predictions of secondary neutrons and their importance to radiation effects inside the international space station. *Radiat Meas.* 2001;33:229-34.

104. Koshiishi H, Matsumoto H, Chishiki A, Goka T, Omodaka T. Evaluation of the neutron environment inside the International Space Station based on the bonner ball neutron detector experiment. *Radiat Meas.* 2007;42:1510-20.

105. Mitchell JB, Bedford JS, Bailey SM. Dose-rate effects in mammalian cells in culture III. Comparison of cell killing and cell proliferation during continuous irradiation for six different cell lines. *Radiat Res.* 1979;79(3):537-51.

106. Yi PN, Stanley WS, Lee W. Relationship between mitotic delay and the minimum dose rate of X irradiation required to stop cell proliferation. *Radiat Res.* 1993;133(2):163-9.

107. Williams JP, Brown SL, Georges GE, Hauer-Jensen M, Hill RP, Huser AK, Kirsch DG, Macvittie TJ, Mason KA, Medhora MM, Moulder JE, Okunieff P, Otterson MF, Robbins ME, Smathers JB, McBride WH. Animal models for medical countermeasures to radiation exposure. *Radiat Res.* 2010;173(4):557-78.

108. Williams JP, Jackson IL, Shah JR, Czarniecki CW, Maidment BW, DiCarlo AL. Animal models and medical countermeasures development for radiation-induced lung damage: report from an NIAID Workshop. *Radiation research.* 2012;177(5):e0025-39.

109. Doll R. The Age Distribution of Cancer: Implications for Models of Carcinogenesis. *J R Stat Soc Series A.* 1971;134(2):133-66.

110. Frank SA. Dynamics of Cancer. Princeton, NJ: Princeton University Press; 2007.

111. Cucinotta FA. Biophysics of Nasa Radiation Quality Factors. *Radiat Prot Dosim.* 2015).

112. Cucinotta FA. A new approach to reduce uncertainties in space radiation cancer risk predictions. *PLoS One.* 2015;10(3):e0120717.

113. Sridharan DM, Asaithamby A, Bailey SM, Costes SV, Doetsch PW, Dynan WS, Kronenberg A, Rithidech KN, Saha J, Snijders AM, Werner E, Wiese C, Cucinotta FA, Pluth JM. Understanding cancer development processes after HZE-particle exposure: roles of ROS, DNA damage repair and inflammation. *Radiat Res.* 2015;183(1):1-26.

114.Vassetzky *et al.* (2012), SINEBase: a database and tool for SINE analysis, *Nucl. Acids Res.*, 1-7.

115. Baker (2012), Digital PCR hits its stride, *Nature Methods*, 9:541-544.

116. Hundson *et al.* (2011), High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number, *Anal. Chem.*, 83:8604–8610.

CLAIMS

I claim:

1. An assay to determine the number of accumulated mutations in a target sequence within a target short interspersed element (SINE), target long interspersed element (LINE) and/or the genome of a subject, the assay comprising the steps of:

a) obtaining a genomic DNA sample from the subject and fragmenting the genomic DNA sample, or obtaining a fragmented genomic DNA sample from the subject,

b) mixing a predetermined number of fragments of the genomic DNA that arise from a predetermined number of genomes with a reagent mixture to produce a reaction mixture, wherein the reagent mixture comprises:

i) a pair of polymerase chain reaction primers that amplify a target amplicon comprising the target sequence within the target SINE, target LINE and/or target sequence within the genome,

ii) a target sequence clamp which binds only to the wild-type target sequence within the target SINE, target LINE and/or target sequence within the genome, wherein the target sequence clamp prevents the PCR amplification of only those target amplicons that have the target wild-type sequence within the target SINE, target LINE and/or target sequence within the genome and permits the PCR amplification of only those target amplicons that have the target mutated sequence within the target SINE, target LINE and/or target sequence within the genome, and

iii) a DNA polymerase enzyme and the reactants for a digital PCR (dPCR),

c) subjecting the reaction mixture to the dPCR,

d) identifying the number of fragments of the genomic DNA comprising the target amplicon having the target mutated sequence within the SINE, target LINE and/or target sequence within the genome based on the number of positive PCR amplifications in the dPCR, and

e) calculating the number of accumulated mutations per genome in the target sequence within the target SINE, target LINE and/or target sequence within the genome based on the number of fragments of the genomic DNA that arise from one genome and the number of fragments of the genomic DNA per genome that comprise the target amplicons having the target mutated sequence within the target SINE, target LINE and/or target sequence within the genome wherein the presence of the target mutated sequence within the

target SINE, target LINE and/or target sequence within the genome is indicated by PCR amplification of the target amplicon in the dPCR.

2. The assay of claim 1, wherein the target SINE, target LINE and/or target sequence within the genome is about 50-500, about 100-400, about 100-250, about 200-300, about 250-350 or about 300 bp.

3. The assay of claim 1, wherein the target SINE, target LINE and/or target sequence within the genome covers about 4%-15% of the genome of the subject.

4. The assay of claim 1, wherein each genomic DNA fragment from at least about 90% of the genomic DNA fragments is about 800-1500 bp.

5. The assay of claim 1, wherein the reactants for a digital PCR comprise deoxyribonucleotides (dNTPs), a metal ion selected from Mg^{2+} or Mn^{2+} and a buffer.

6. The assay of claim 1, wherein the target sequence clamp is about 20 bp.

7. The assay of claim 1, wherein the melting temperature of the target sequence clamp with the target sequence is higher than the temperatures used in the PCR cycle.

8. The assay of claim 1, wherein the target sequence clamp comprises xenonucleotide (XNA).

9. The assay of claim 1 wherein the dPCR is droplet dPCR (ddPCR).

10. The assay of claim 1, wherein the subject is a mammal.

11. The assay of claim 10, wherein the mammal is a human, non-human primate, rat, mouse, pig, dog or cat.

12. The assay of claim 11, wherein the mammal is a mouse and the target SINE is B1 which has the sequence of SEQ ID NO: 6.

13. The assay of claim 12, wherein the primer pair comprises the sequences of SEQ ID NOs: 1 and 2 and the target sequence clamp has the sequence of SEQ ID NOs: 3, 4 or 5.

14. The assay of claim 10, wherein the mammal is a human and the target SINE is Alu which as the sequence selected from of SEQ ID NOs: 11-98.

15. The assay of claim 14, wherein the primer pair comprises the appropriate sequences selected from SEQ ID NOs: 99-274 and the target sequence clamp has the appropriate sequence selected from SEQ ID NOs: 7-10.

16. A method for calculating the rate of mutations in a target sequence within a target SINE, target LINE and/or target sequence within the genome in a subject, the method comprising the steps of:

a) according to the assay of claim 1, determining the number of accumulated mutations in the target sequence within the target SINE, target LINE and/or target sequence within the genome in a first sample obtained from the subject at a first time point,

b) according to the assay of claim 1, determining the number of accumulated mutations in the target sequence within the target SINE, target LINE and/or target sequence within the genome in a second sample obtained from the subject at a second time point, and

c) calculating the rate of mutations in the target sequence within the target SINE, target LINE and/or target sequence within the genome of the subject based on the difference between the number of accumulated mutations in the target sequence within the target SINE, target LINE and/or target sequence within the genome in the subject at the first time point and the second time point and the duration between the first time point and the second time point.

17. A method of calculating the accumulated mutations in the genome of a subject, the method comprising the steps of:

a) according to the assay of claim 1, determining the number of accumulated mutations in the target sequence within the target SINE, target LINE and/or target sequence within the genome in a sample obtained from the subject, and

b) calculating the accumulated mutations in the genome of the subject based on the frequency of occurrence of the target SINE, target LINE and/or target sequence within the genome throughout the genome of the subject and the number of accumulated mutations in the target sequence within the target SINE, target LINE and/or target sequence within the genome.

18. A method of calculating the rate of mutations in the genome of a subject, the method comprising the steps of:

a) according to the method of claim 17, calculating the number of accumulated mutations in the genome of the subject in a first sample obtained from the subject at a first time point,

b) according to the method of claim 17, calculating the number of accumulated mutations in the genome of the subject in a second sample obtained from the subject at a second time point, and

c) calculating the rate of mutations in the genome of the subject based on the difference between the rate of mutations in the genome of the subject at the first time point and the second time point and the duration between the first time point and the second time point.

19. A method of determining a standard scale for the genomic age of a species, the method comprising the steps of:

a) according to the assay of claim 1 and/or 16-18, determining the average accumulated mutations and/or the average rate of mutations in a target sequence in a target SINE, target LINE and/or the genomes in different groups of individuals of varying ages belonging to the species that are living under conditions of exposure to only natural mutagens and/or conditions of minimal exposure to man-made mutagens, and

b) determining the standard scale for the genomic age of the species,

wherein the standard scale for the genomic age of the species indicates the average accumulated mutations and/or the average rate of mutations in the target sequence in the target SINE, target LINE and/or the genomes of individuals belonging to the species at increasing chronological ages.

20. A method of determining the genomic age of a subject, the method comprising the steps of:

a) preparing a standard scale for the genomic age of a species which indicates the average accumulated mutations and/or the average rate of mutations in the target sequence in the target SINE, target LINE and/or the genomes of individuals belonging to the species at increasing chronological ages, or obtaining the standard scale of the genomic age for a species,

a) according to the assay of claim 1 and/or 16-18, determining the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome of the subject, and

c) determining the genomic age of the subject based on the comparison of the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome of the subject with the standard scale for the genomic age of the species.

21. A method of determining a standard scale of cancer risk in a species, the method comprising the steps of:

a) according to the assay of claim 1 and/or 16-18, determining the average accumulated mutations and/or the average rate of mutations in a target sequence in a target SINE, target LINE and/or the genomes in different groups of individuals of varying ages that belong to the species and are free from cancer and/or are known to have a low risk of cancer development, and

b) determining the standard scale of cancer risk in the species,

wherein the standard scale of cancer risk in the species indicates, at increasing chronological ages, the average accumulated mutations and/or the average rate of mutations in the target sequence in the target SINE, target LINE and/or the genomes of individuals of varying ages that belong to the species and are free from cancer and/or are known to have a low risk of cancer development.

22. A method for determining the risk of cancer development in a subject, the method comprising the steps of:

a) preparing a standard scale of cancer risk in a species which indicates, at increasing chronological ages, the average accumulated mutations and/or the average rate of mutations

in a target sequence in a target SINE, target LINE and/or the genomes of individuals of varying ages that belong to the species and are free from cancer and/or are known to have a low risk of cancer development, or obtaining the standard scale of cancer risk,

b) according to the assay of claim 1 and/or 16-18, determining the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome in the subject, and

c) determining the risk of cancer development in the subject based on the comparison of the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome in the subject with the standard scale of cancer risk, wherein the step of estimating the risk of cancer development in the subject comprises:

i) identifying the subject as having a higher risk of cancer development if the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome of the subject are higher than the corresponding values in the standard scale of cancer risk, or

ii) identifying the subject as having a lower risk or no risk of cancer development if the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome of the subject are lower than or equal to the corresponding values in the standard scale of cancer risk.

23. The method of claim 22, the method further comprising administering enhanced screening for cancer to a subject identified as having a higher risk of cancer development and/or recommending lifestyle changes designed to reduce the risk of cancer development in the subject identified as having the higher risk of cancer development.

24. The method of claim 22, the method further comprising administering routine screening for cancer to a subject identified as having a lower risk of cancer development.

25. A method of determining a standard scale of cancer risk for a tissue or organ in a species, the method comprising the steps of:

a) according to the assay of claim 1 and/or 16-18, determining the average accumulated mutations and/or the average rate of mutations in a target sequence in a target SINE, target LINE and/or the genomes of the cells of the tissue or organ from different

groups of individuals of varying ages that belong to the species and are free from cancer and/or are known to have a low risk of cancer development in the tissue or organ, and

b) determining the standard scale of cancer risk for the tissue or organ in the species,

wherein the standard scale of cancer risk for the tissue or organ in the species indicates, at increasing chronological ages, the average accumulated mutations and/or the average rate of mutations in the target sequence in the target SINE, target LINE and/or the genomes of the cells of the tissue or organ in the individuals of varying ages that belong to the species and are free from cancer and/or are known to have a low risk of cancer development in the tissue or organ.

26. A method for determining the risk of cancer development for a tissue or organ in a subject, the method comprising the steps of:

a) preparing a standard scale of cancer risk for the tissue or organ in a species which indicates, at increasing chronological ages, the average accumulated mutations and/or the average rate of mutations in a target sequence in a target SINE, target LINE and/or the genomes of cells in the tissue or organ of individuals of varying ages that belong to the species and are free from cancer and/or are known to have a low risk of cancer development in the tissue or organ, or obtaining the standard scale of cancer risk for the tissue or organ,

b) according to the assay of claim 1 and/or 16-18, determining the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome of cells in the tissue or organ of the subject, and

c) estimating the risk of cancer development in the subject based on the comparison of the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome of the cells in the tissue or organ of the subject with the standard scale of cancer risk for the tissue or organ, wherein the step of estimating the risk of cancer development in the tissue or organ in the subject comprises:

i) identifying the subject as having a higher risk of cancer development in the tissue or organ if the accumulated mutations and/or the rate of mutations in the target sequence in the target SINE, target LINE and/or the genome of the cells of the tissue or organ of the subject are higher than the corresponding values in the standard scale of cancer risk for the tissue or organ, or

ii) identifying the subject as having a lower risk or no risk of cancer development if the accumulated mutations and/or the rate of mutations in the target

sequence in the target SINE, target LINE and/or the genome of the cells of the tissue or organ in the subject are lower than or equal to the corresponding values in the standard scale of cancer risk for the tissue or organ.

27. The method of claim 26, the method further comprising administering enhanced screening for cancer of the tissue or organ to a subject identified as having a higher risk of cancer development and/or recommending lifestyle changes designed to reduce the risk of cancer development of the tissue or organ to the subject identified as having the higher risk of cancer development.

28. The method of claim 26, the method further comprising administering routine screening for cancer to a subject identified as having a lower risk of cancer development of the tissue or organ.

29. A method of identifying the effect of a lifestyle change on the risk of cancer development and/or the rate of aging of a subject, the method comprising the steps of:

a) according to the assay of claim 1 and/or 16-18, determining the rate of mutations in a target sequence within a target SINE or target LINE and/or in the genome of the subject immediately before the lifestyle change was initiated,

b) according to the assay of claim 1 and/or 16-18, determining the rate of mutations in the target sequence within the target SINE or target LINE and/or in the genome of the subject after the lifestyle change was initiated,

c) comparing the rates of mutations in the target sequence within the target SINE, target LINE and/or the genome of the subject before and after the lifestyle change was initiated to determine the effect of the lifestyle change on the risk of cancer development and/or the rate of aging in the subject, wherein the step of determining the effect comprises:

i) determining an increased risk of cancer development and/or an increased rate of aging based on a higher rate of mutations in the target sequence in the target SINE, target LINE and/or the genome of the subject after the lifestyle change was initiated, or

ii) determining a decreased risk of cancer development and/or a decreased rate of aging based on a lower rate of mutations in the target sequence in the target SINE, target LINE and/or the genome of the subject after the lifestyle change was initiated.

30. A kit comprising:

a) a pair of PCR primers that amplify a target amplicon comprising a target sequence within a target SINE, target LINE and/or target sequence within the genome,

b) a target sequence clamp which binds only to the wild-type target sequence within the SINE, target LINE and/or target sequence within the genome, wherein the target sequence clamp prevents the PCR amplification of only those target amplicons that have the target wild-type sequence within the SINE, target LINE and/or target sequence within the genome and permits the PCR amplification of only those target amplicons that have the target mutated sequence within the SINE, target LINE and/or target sequence within the genome.

31. The kit of claim 30, the kit further comprising a DNA polymerase enzyme and reactants for conducting a dPCR.

32. The method or kit according to any preceding claim wherein the target sequence within the genome is a highly repeated genomic sequence or mitochondrial genomic DNA.

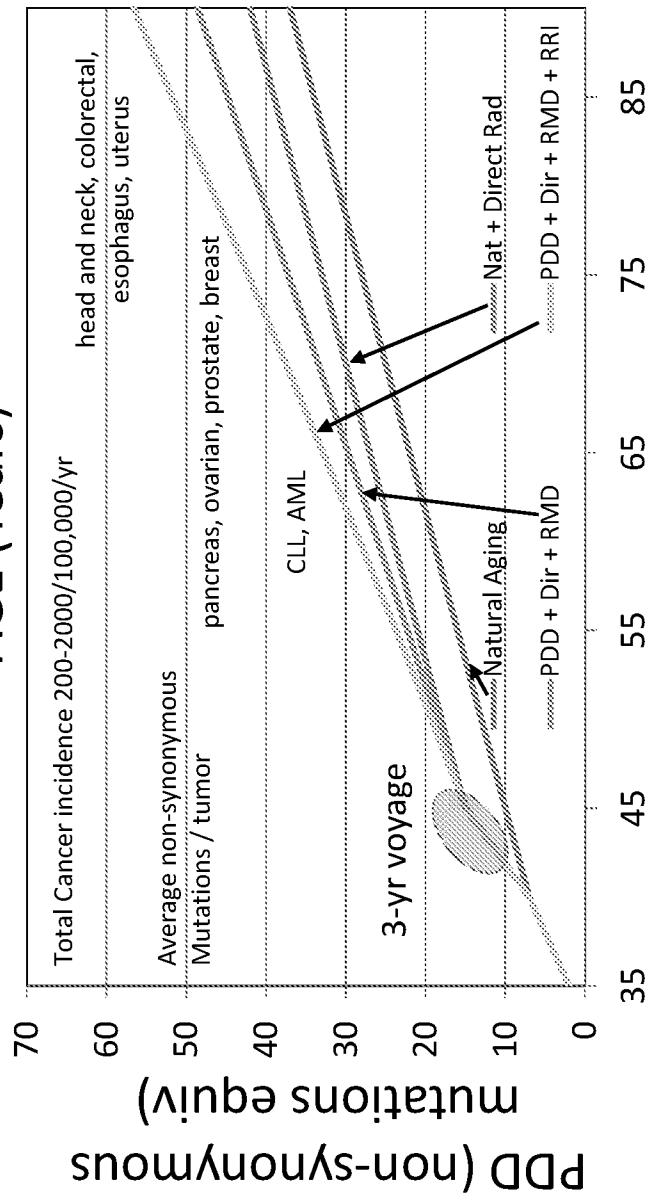

AGE (Years)
PDD (non-synonymous mutations equiv)
70
60
50
40
30
20
10
0
35
45
55
65
75
85
Total Cancer incidence 200-2000/100,000/yr
head and neck, colorectal, esophagus, uterus
Average non-synonymous Mutations / tumor
pancreas, ovarian, prostate, breast
CLL, AML
3-yr voyage
Natural Aging
PDD + Dir + RMD
Nat + Direct Rad
PDD + Dir + RMD + RRI

Figure 1

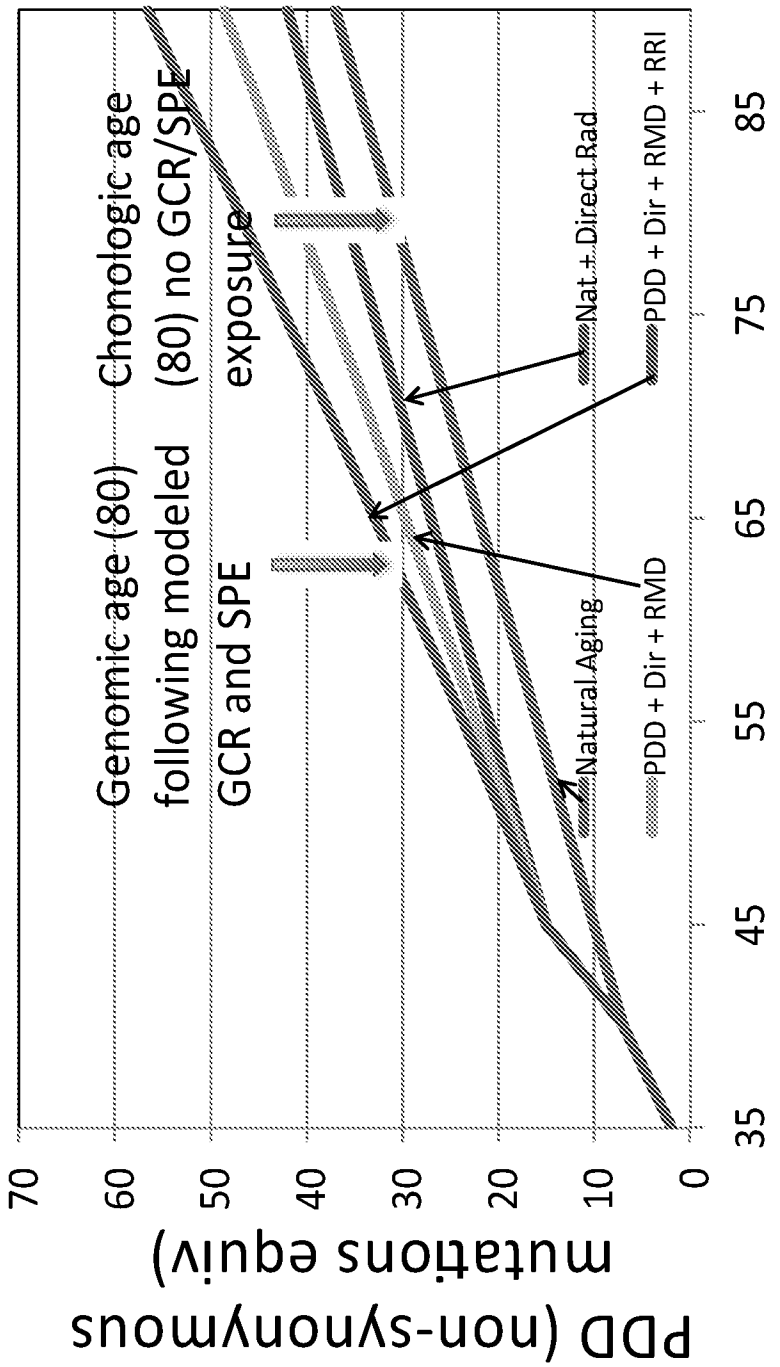

PDD (non-synonymous mutations equiv)
0
10
20
30
40
50
60
70
35
45
55
65
75
85
CHRONOLOGIC AGE (Years)
Genomic age (80) following modeled GCR and SPE
Chonologic age (80) no GCR/SPE exposure
Natural Aging
PDD + Dir + RMD
Nat + Direct Rad
PDD + Dir + RMD + RRI

Figure 2

PREMALIGNANT DNA DAMAGE AND CANCER INCIDENCE

Cancer incidence at age t $\propto$ t^k

Chronological age (t) is the most powerful predictor of cancer risk, and the relationship tends to follow a power law.

Total DNA Damage at t $\propto$ Passenger mutations at t $\propto$ PDD at t

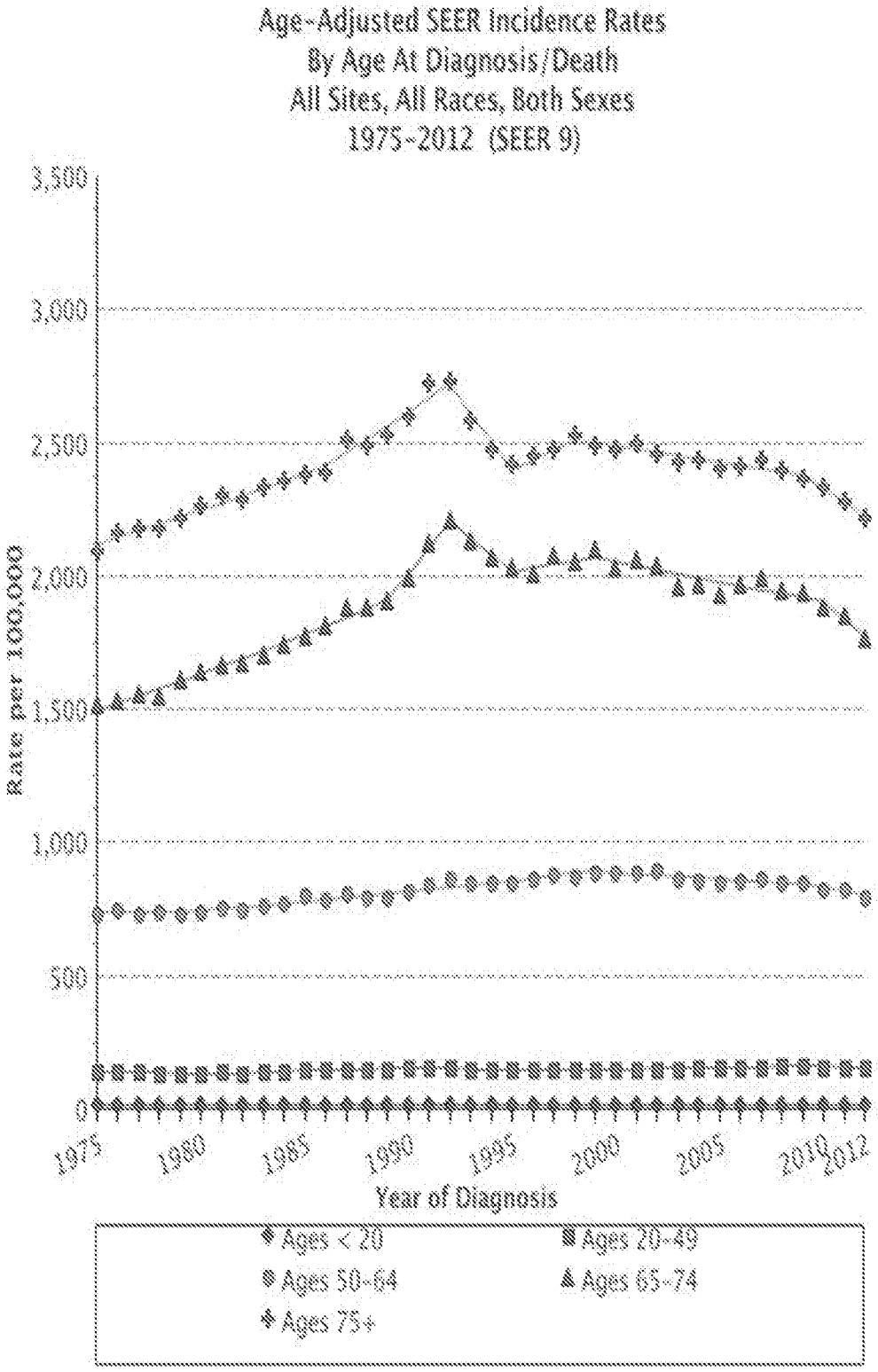
Driver mutations at t $\propto$ Cancer incidence at age t

- DNA damage accumulates in many forms from many sources during a lifetime.
- Most mutations are dark or are detected as passenger mutations.
- Certain proportions of total DNA damage are passengers and drivers.
- Driver mutations confer a reproductive advantage that can lead to oncogenesis.
- Because PDD and t^k are each proportional to cancer incidence, we see that:

$$\mathrm{PDD} \propto t^k \quad \text{and} \quad \frac{d\mathrm{PDD}}{dt} = ct^{k-1}$$

A quantitative estimate of DNA damage accumulation rate can be measured and used to predict driver frequency and current and future cancer risks.

Figure 3



Cancer sites include invasive cases only unless otherwise noted.
Rates are per 100,000 and are age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1130). Regression lines are calculated using the Joinpoint Regression Program Version 4.2.0, April 2015, National Cancer Institute.
Incidence source: SEER 9 areas (San Francisco, Connecticut, Detroit, Hawaii, Iowa, New Mexico, Seattle, Utah, and Atlanta).

Figure 4

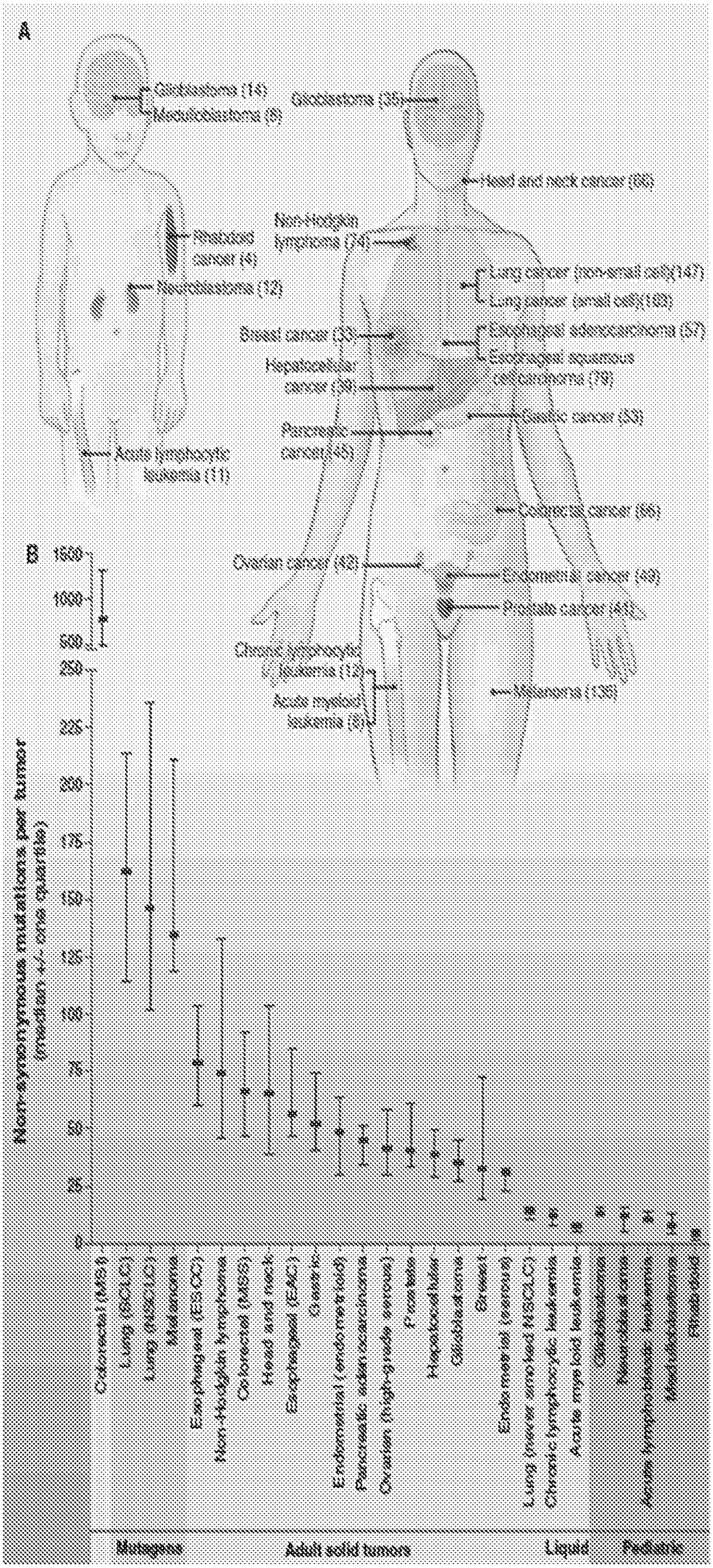

A
Glioblastoma (14)
Medulloblastoma (8)
Rhabdoid cancer (4)
Neuroblastoma (12)
Acute lymphocytic leukemia (11)
Glioblastoma (35)
Head and neck cancer (66)
Non-Hodgkin lymphoma (74)
Lung cancer (non-small cell)(147)
Lung cancer (small cell)(163)
Breast cancer (33)
Esophageal adenocarcinoma (57)
Esophageal squamous cell carcinoma (79)
Hepatocellular cancer (39)
Gastric cancer (53)
Pancreatic cancer (45)
Colorectal cancer (66)
Ovarian cancer (42)
Endometrial cancer (49)
Prostate cancer (41)
Chronic lymphocytic leukemia (12)
Acute myeloid leukemia (8)
Melanoma (135)
B
Non-synonymous mutations per tumor (median +/- one quartile)
Mutagens
Adult solid tumors
Liquid
Pediatric

Figure 5

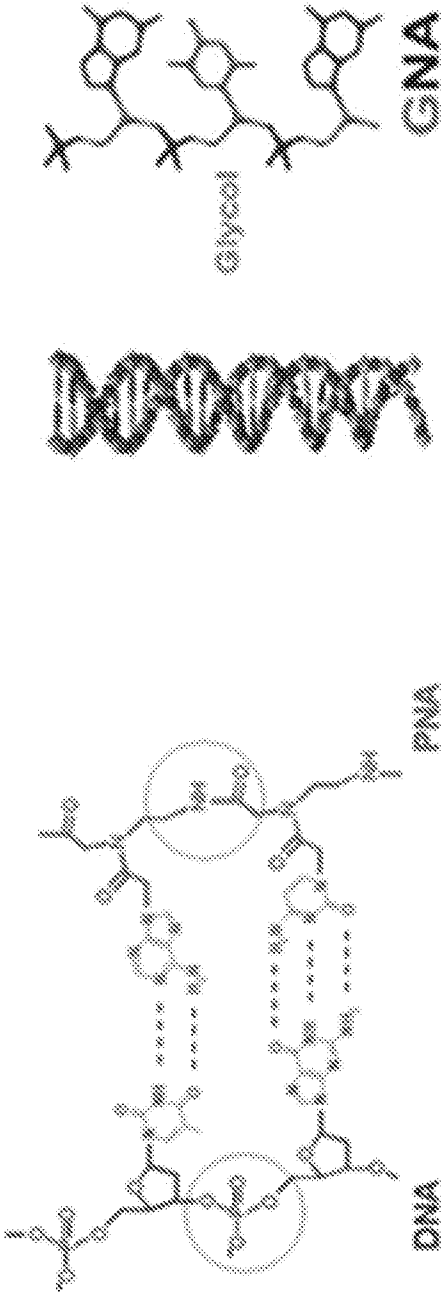

GNA
Glycol
PNA
DNA
XNA Probe
Perfectly match XNA probe for Wild-Type
Wild-Type
DNA
Real-Time PCR
Amplicon
XNA Probe
Unmatch XNA probe for Mutant-Type
Mutant-Type
DNA
Real-Time PCR
Amplicon

Figure 6

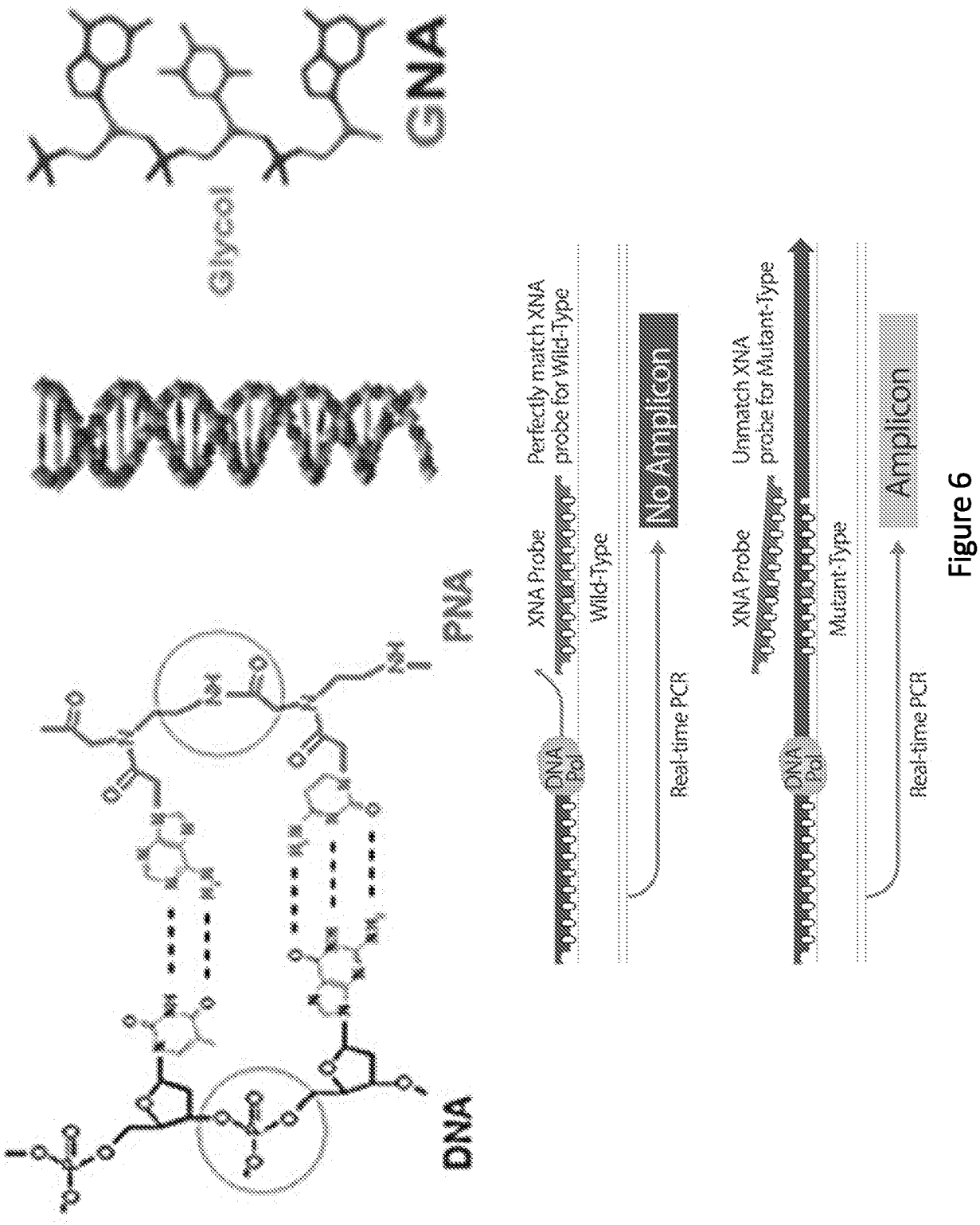

GNA
Glycol
DNA
PNA
XNA Probe
Perfectly match XNA probe for Wild-Type
Wild-Type
No Amplicon
Real-time PCR
XNA Probe
Unmatch XNA probe for Mutant-Type
Mutant-Type
Amplicon
Real-time PCR

Figure 6

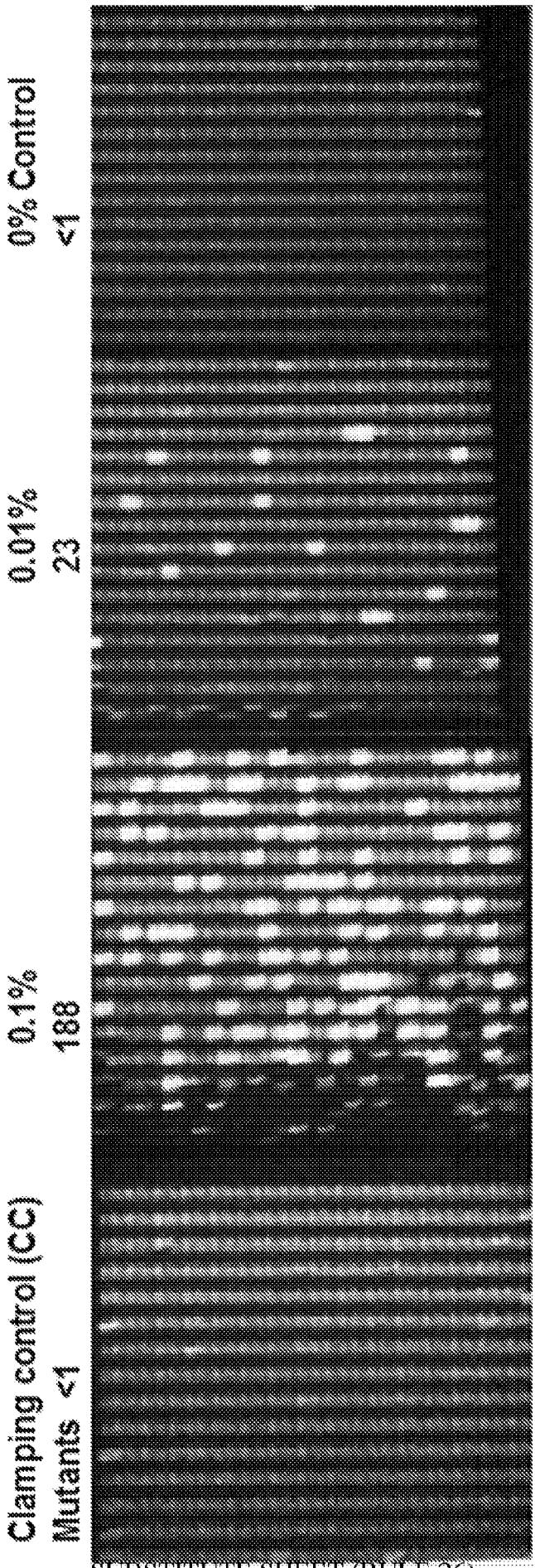


Figure 7

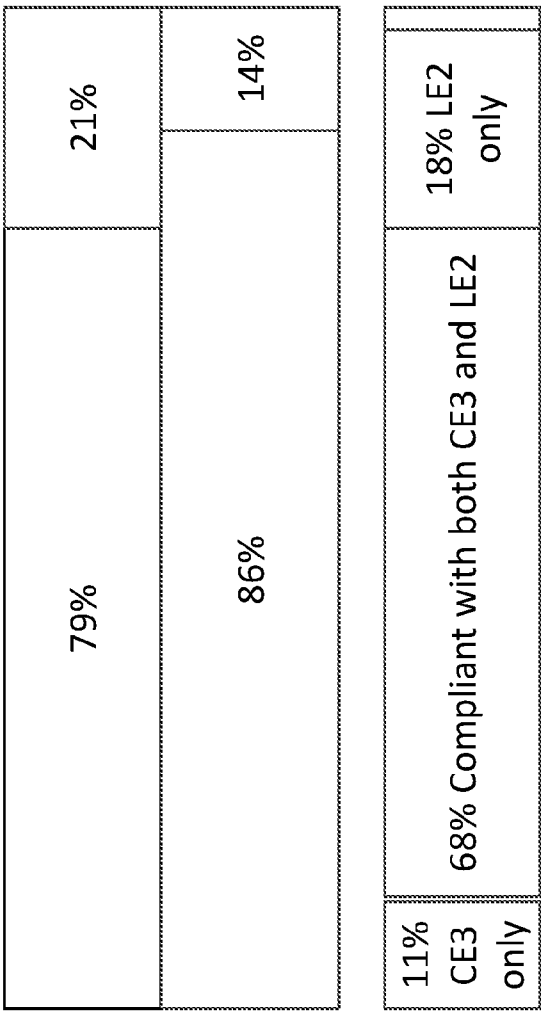
B1-gene [113 bp]

GCCTTTAATCCCAGCACTTGGGAGGCAGAGGCAGGCGGATTTCTGAGTTCGAGGCCA
GCCTGGTCTACAAAGTGAGTTCCAGGACAGCCAG**GGCTACACAGAGAAACCCTGTC**

Figure 8

XNA	Mean addl cycle number	Fraction of suppressed B1 alleles
No XNA	0.00 ± 0.04	0% ± 3%
XNA CE3	2.23 ± 0.07	79% ± 1%
XNA LE2	2.86 ± 0.31	86% ± 3%
XNA CE3 and LE2	5.26 ± 0.23	97% ± 0.4%

Figure 9



CE3: 79% compliant alleles
21% have point or larger differences

LE2: 86% compliant alleles
14% have point or larger differences

LE2 or CE3: 97% compliant alleles with one or the other
3% have point or larger differences in both CE3 and LE2

Figure 10

XNA	Delta Cycles (0 Gy)	Delta Cycles (1 Gy)	Mean Estimated mutations/base
No XNA	0.00 ± 0.10	0.00 ± 0.04	NA
CE3 XNA	2.39 ± 0.19	2.27 ± 0.05	0.08%
LE2 XNA	3.07 ± 0.10	2.91 ± 0.03	0.09%
Both CE3 & LE2	5.16 ± 0.07	4.96 ± 0.01	0.02%

Figure 11

XNA	Delta Cycles (0 Gy)	Delta Cycles (9 Gy)	Mean Estimated mutations/base
No XNA	0.00 ± 0.01	0.00 ± 0.01	NA
CE3 XNA	2.55 ± 0.01	2.01 ± 0.01	0.39%
LE2 XNA	2.79 ± 0.02	2.60 ± 0.01	0.13%
Both CE3 & LE2	4.79 ± 0.02	4.75 ± 0.02	0.01%

Figure 12

Organ	Dose	B1 mutation rate at 2 h		B1 mutation rate on Day 6	
		CE3	LE2	CE3	LE2
Liver	1 Gy	+	+	+	+
	9 Gy	+	+	+	+
Spleen	1 Gy	+	+	+	-
	9 Gy	+	+	+	+
Brain	1 Gy	+			
	9 Gy			+	
Gut	1 Gy			+	+
	9 Gy			+	

Figure 13

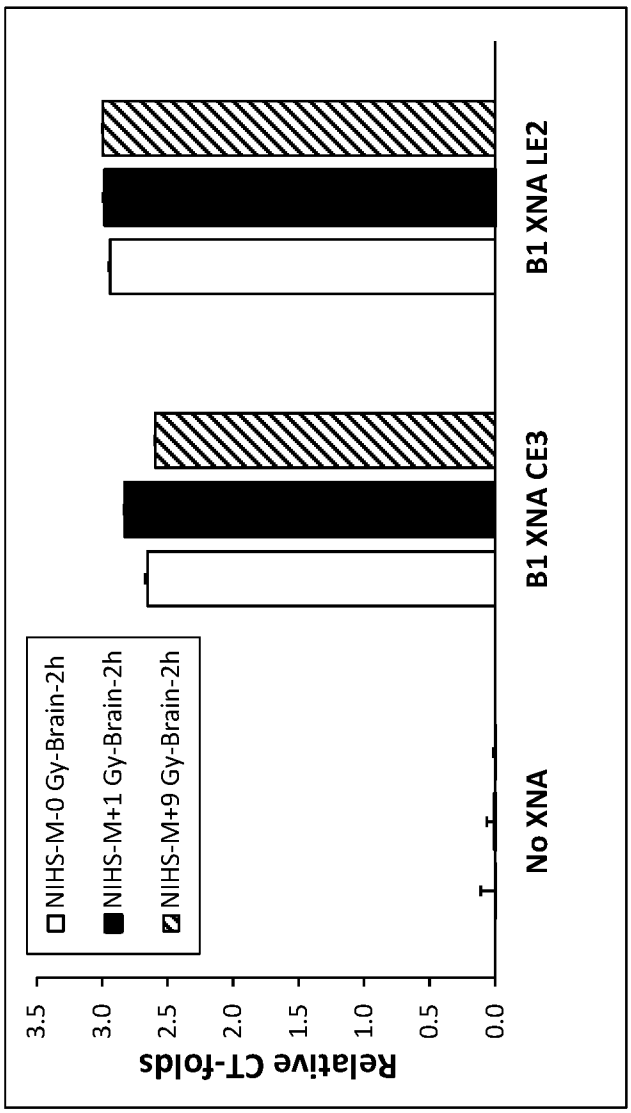


Figure 14A

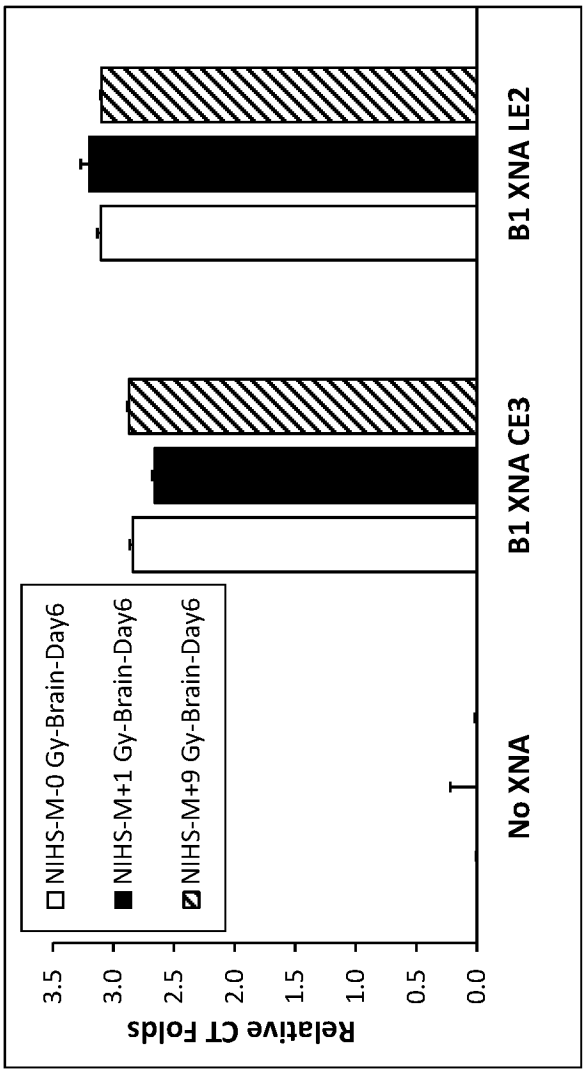


Figure 14B

Figure 15C - Change in majority B1 Alleles
Brain of the C57BL/6 mice had a reduction of non-majority B1 alleles. Balb/c had a small increase

	C57BL/6 M	C57BL/6 F
CE3	-0.05%	-0.08%
LE2	-0.01%	-0.05%
	Balb/c M	**Balb/c F**
CE3	0.02%	0.01%
LE2	0.05%	0.00%

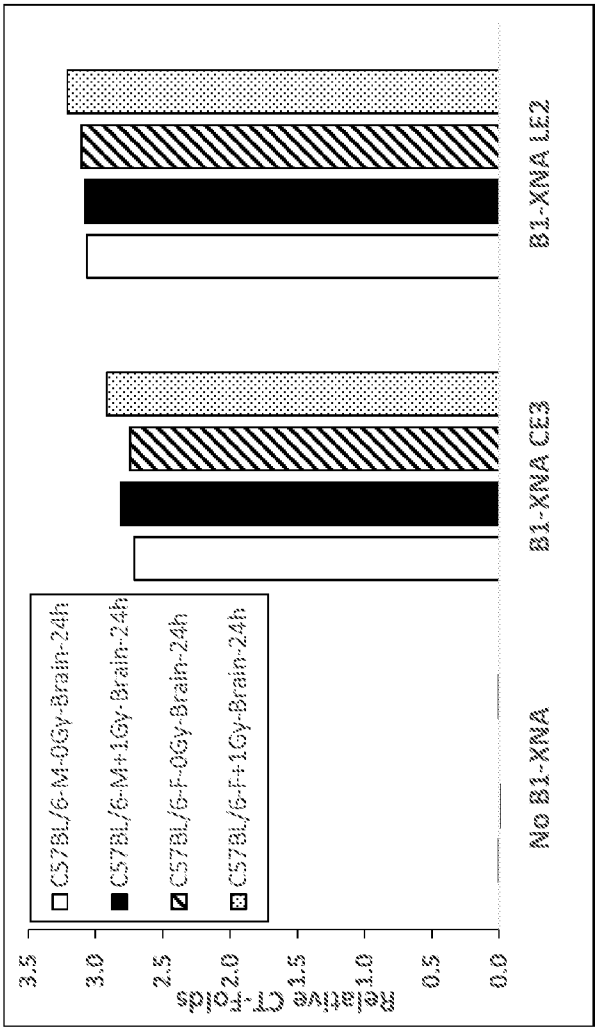


Figure 15A

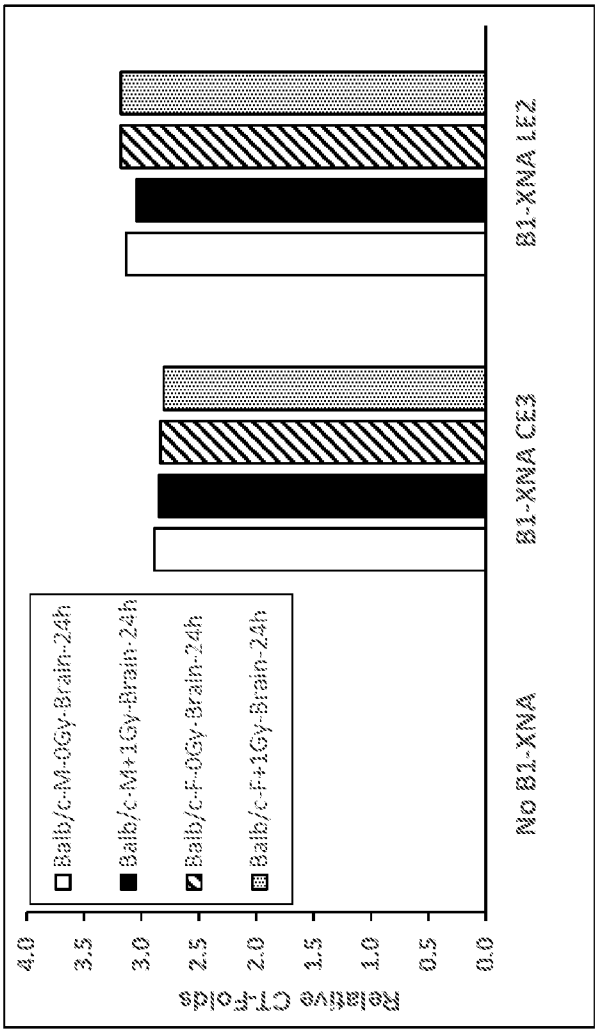


Figure 15B

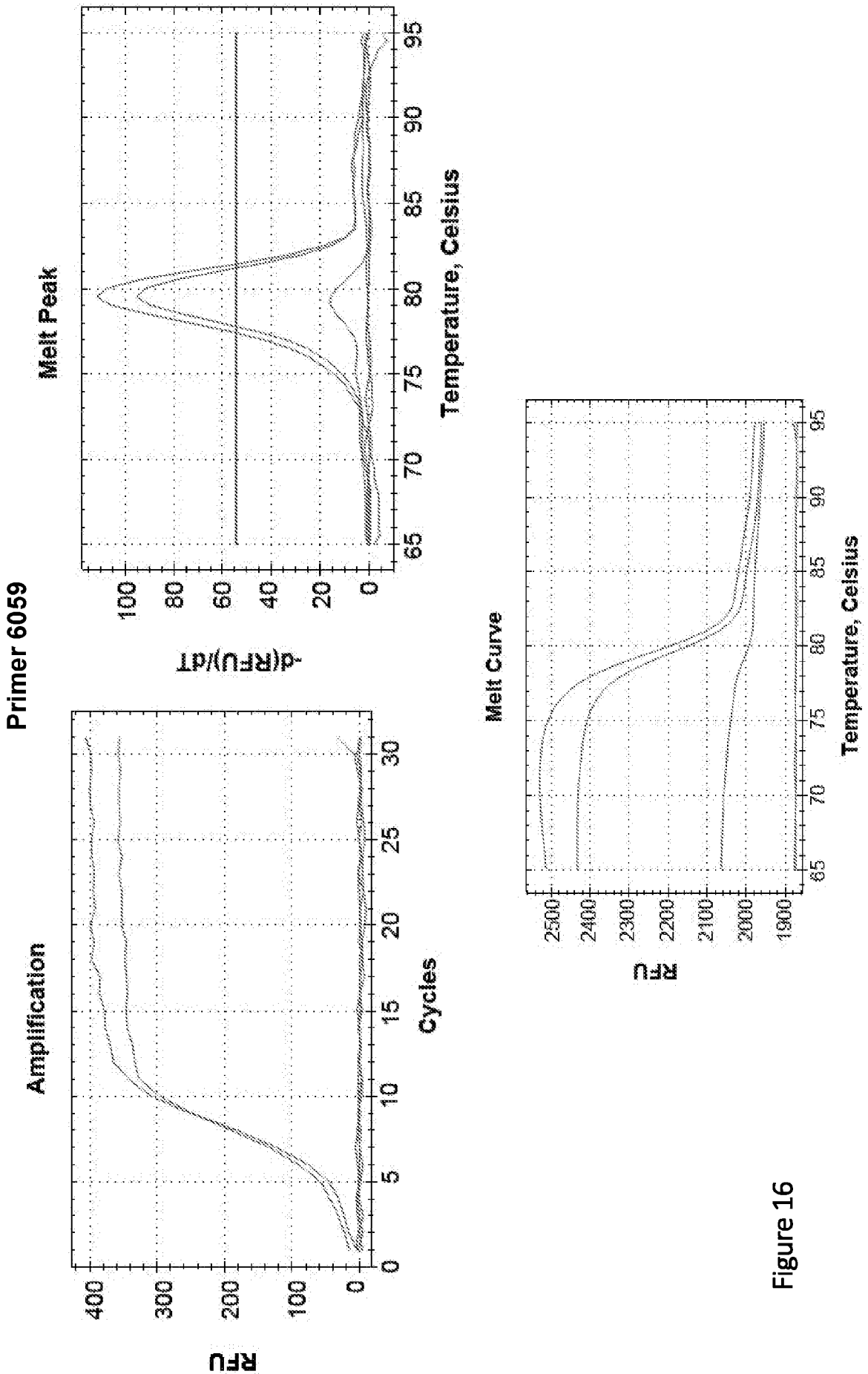

Primer 6059
Amplification
RFU
Cycles
Melt Peak
-d(RFU)/dT
Temperature, Celsius
Melt Curve
RFU
Temperature, Celsius

Figure 16

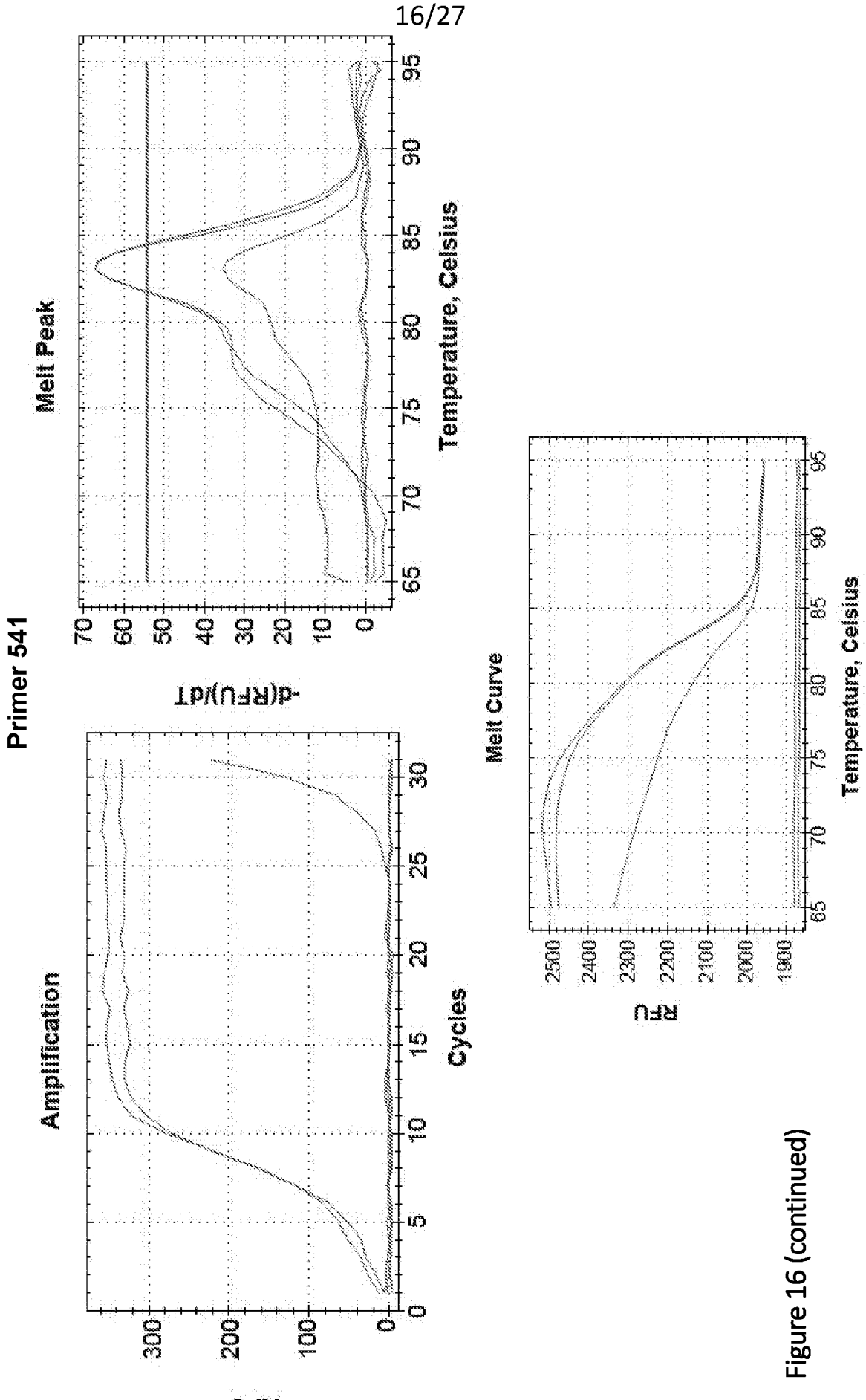


Figure 16 (continued)

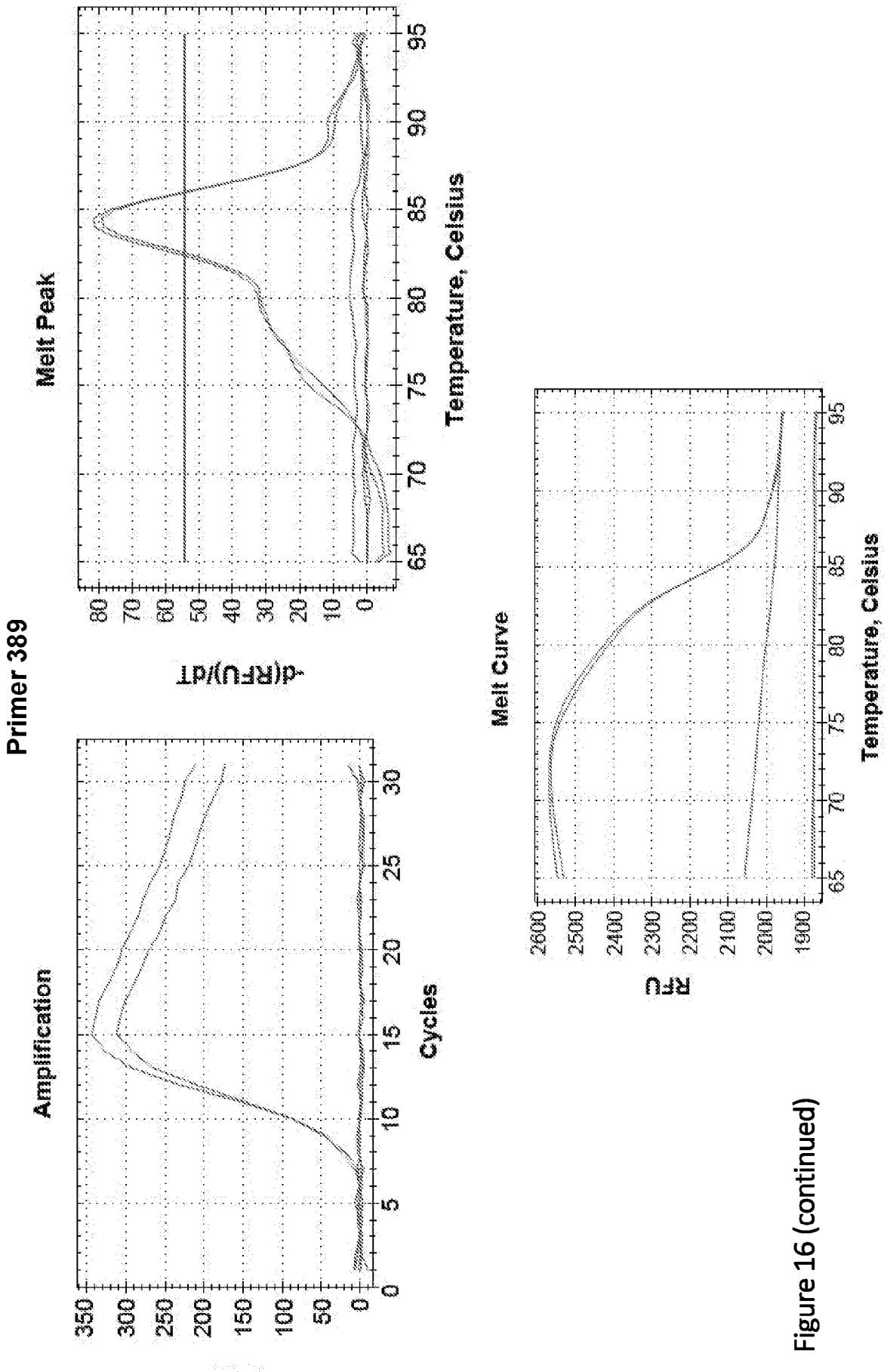


Figure 16 (continued)

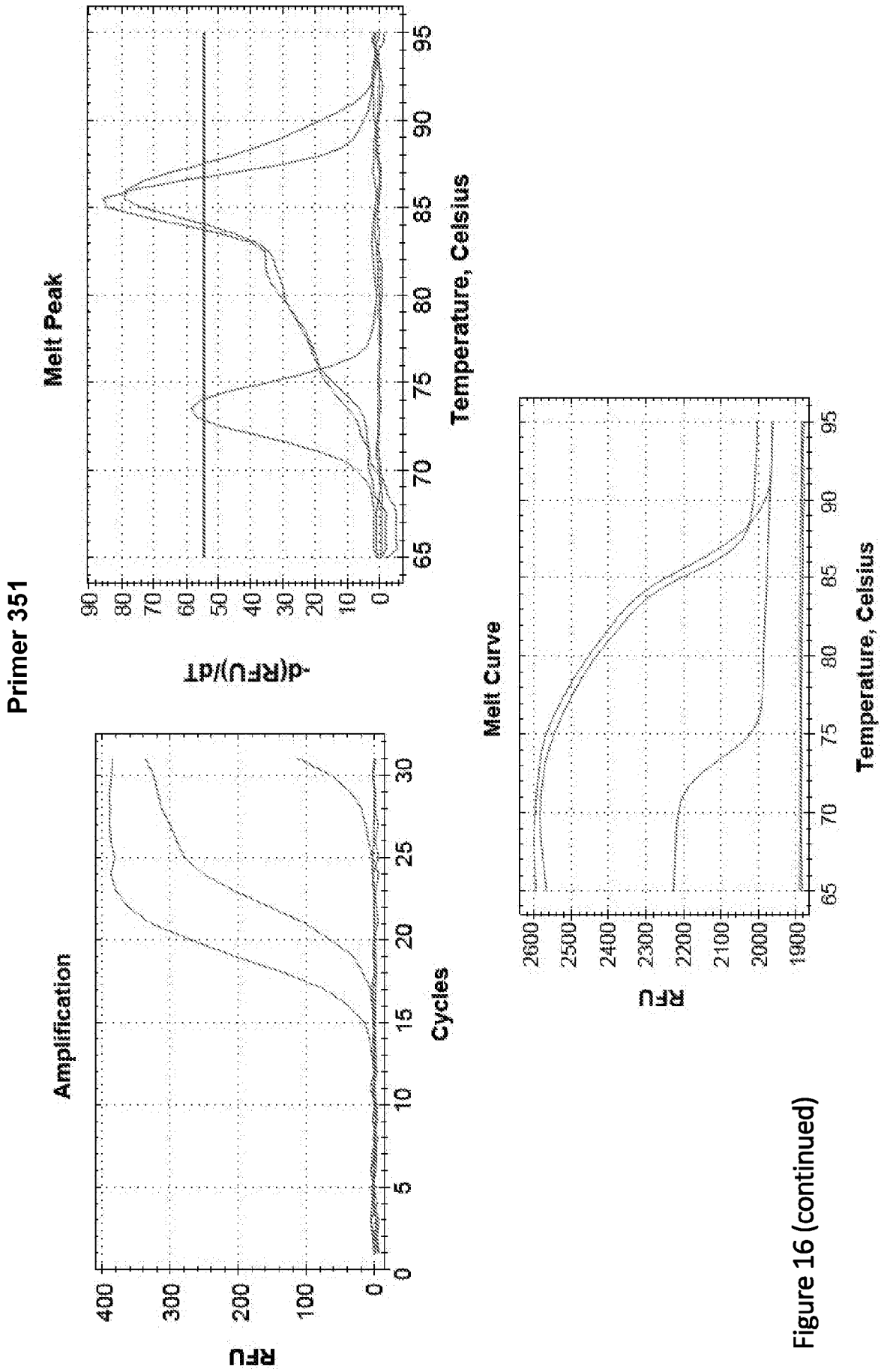


Figure 16 (continued)

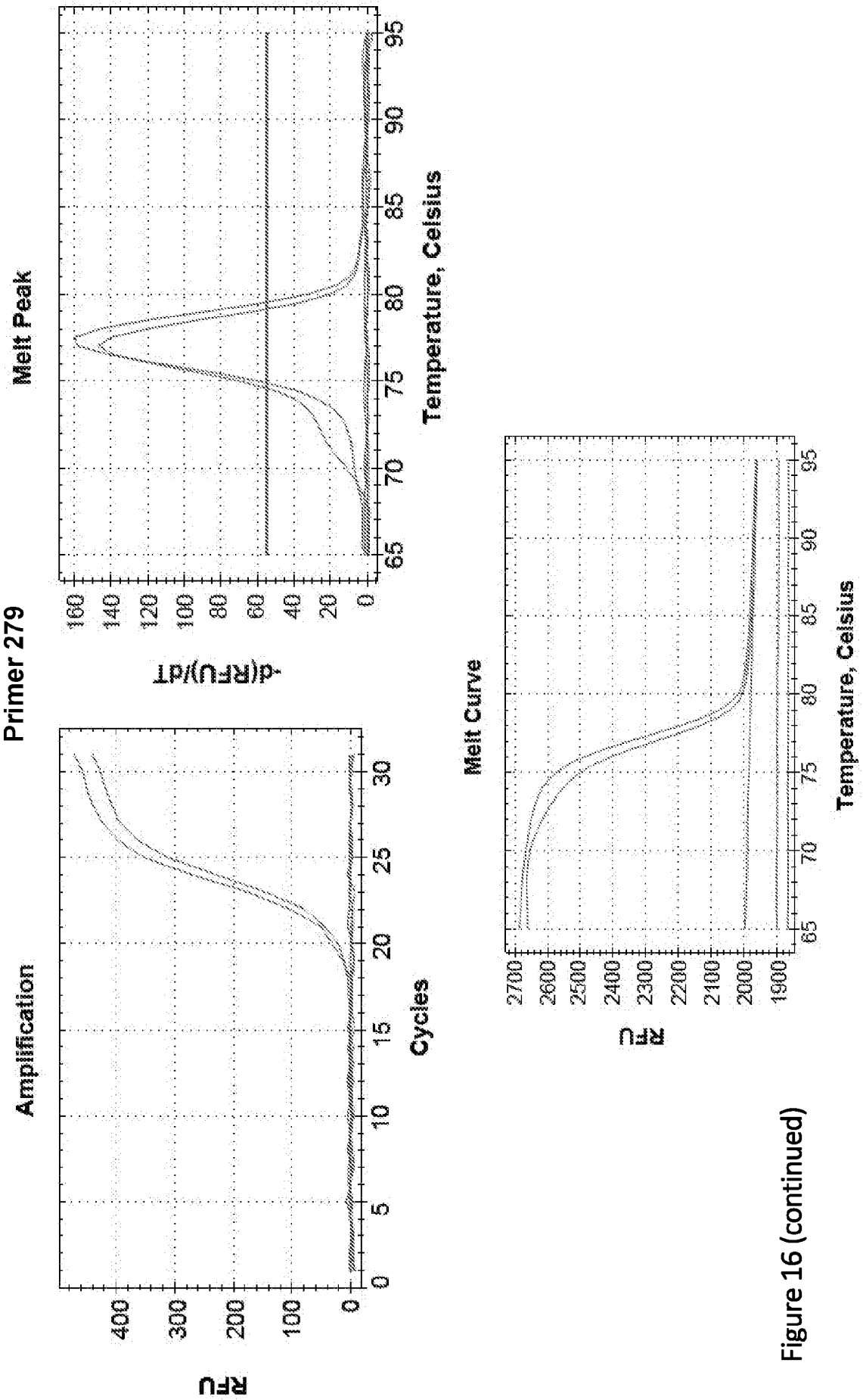


Figure 16 (continued)

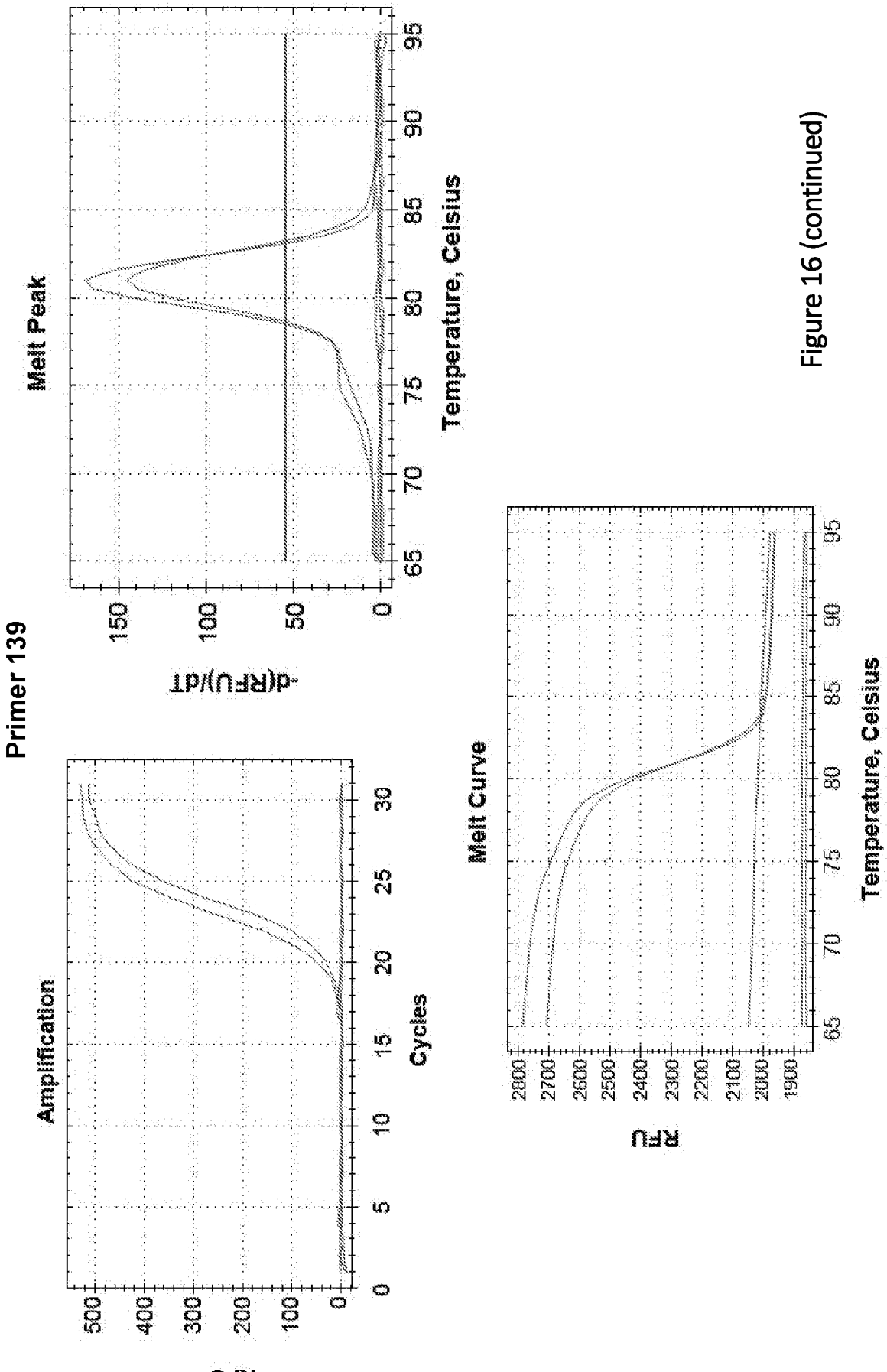


Figure 16 (continued)

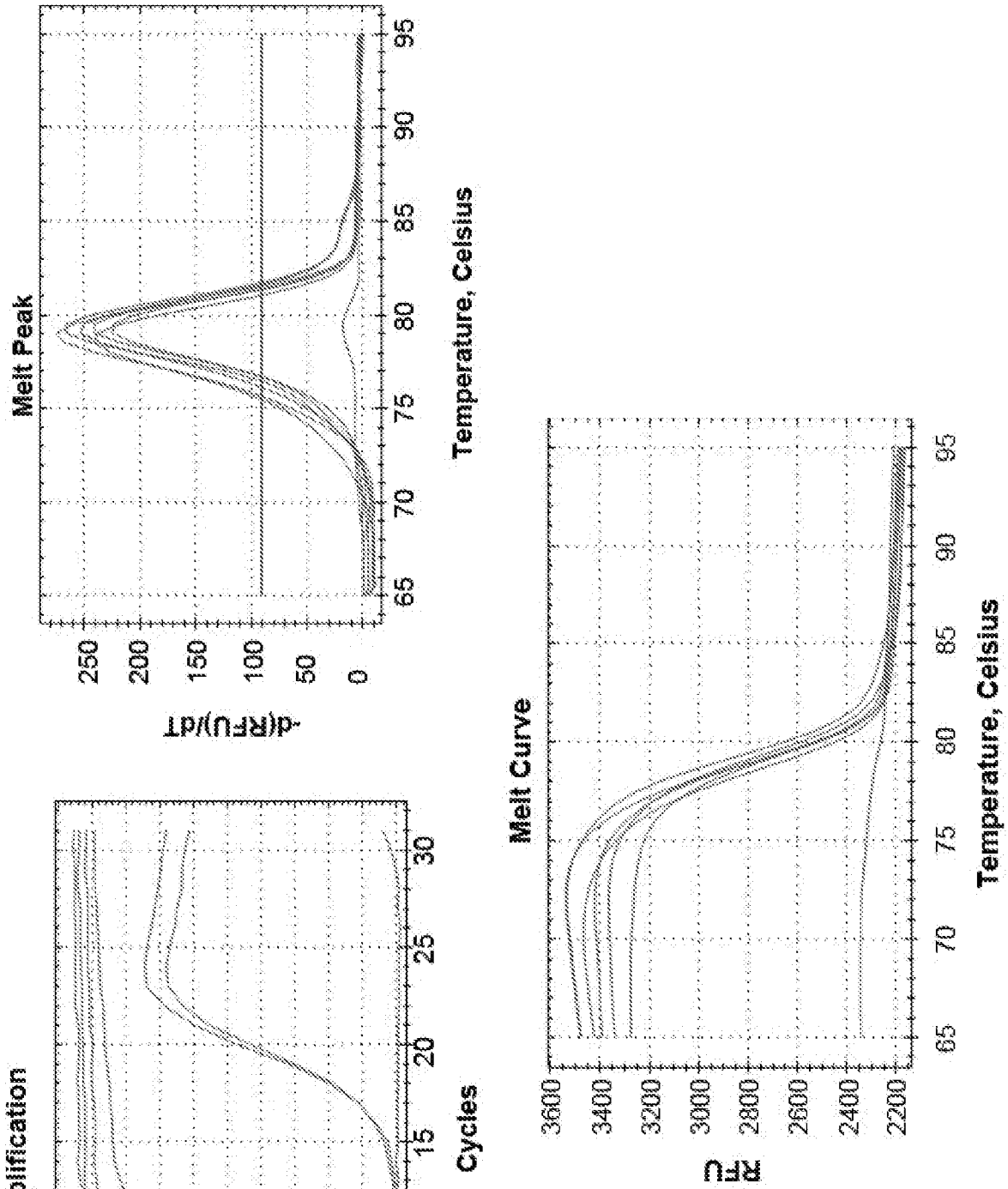

Melt Peak
-d(RFU)/dT
Temperature, Celsius
Amplification
RFU
Cycles
Melt Curve
RFU
Temperature, Celsius

Figure 17

4789 aa atggtgctgg gaaaactggc tagccatatg tagaaagctg aaactggatc ccttccttac accttataca aaaatcaatt caagatggat taaagattta aacgttaaac ctaaaaccat aaaaacccta gaagaaaacc taggcattac cattcaggac ataggcgtgg gcaaggactt catgtccaaa acaccaaaag caatggc 4997

Figure 18

5' - GCCATCCCATATATTGGAAC-3’ Tm = 77, GC = 45%

PCR Primers for this sequence include:

Forward: 5’ ctactgttcgcagtcatagcc 3’
Reverse: 5’ aatcgttcacctcctcttcctc 3’

Figure 19

Name	Sequence	PCR product size (bP)
6059L	aaatggtgctgggaaaactg (SEQ ID NO: 487)	209
6059R	gccattgcttttggtgtttt (SEQ ID NO: 488)	
6053L	aaatggtgctgggaaaactg (SEQ ID NO: 489)	209
6053R	gccattgcttttggtgtttt (SEQ ID NO: 490)	
541L	tggagtacagtggctcatgc (SEQ ID NO: 491)	151
541R	ccatcatgcctggctaattt (SEQ ID NO: 492)	
389L	gagtgcaatggtgcgatct (SEQ ID NO: 493)	152
389R	gcagcatggtgaaaccctat (SEQ ID NO: 494)	
351L	tgacgagttaatgggtgcag (SEQ ID NO: 495)	210
351R	agatccctggacacattctga (SEQ ID NO: 496)	
279L	agcaatgatacaatgtgctatctga (SEQ ID NO: 497)	150
279R	tgttaattatcttatccttgtgccatt (SEQ ID NO: 498)	
139L	acaccaccatatccctacgc (SEQ ID NO: 499)	109
139R	gggatcatggccttctaggt (SEQ ID NO: 500)	

Figure 20

```
CLUSTAL O(1.2.4) multiple sequence alignment

6059s3-6059L       NGNTCATAGTAG-AGCTGAACTGGACCCTTCCTTACACCTTATACAAAAATTAATTCAAG    59
6059s1-6059L       NGCTAATAGTAG-AGCTGAACTGGACCCTTCCTTACACCTTATACAAAAATTAATTCAAG    59
6059s2-6059L       NGCTCATAGTAGTAGCTGAACTGGACCCTTCCTTACACCTTATACAAAAATTAATTC-AG    59
                   ** *.******* ****************************************** **

6059s3-6059L       ATGGATTAAAGACTTAAATGTTAGACCTAAAACCATAAAAACCCTAGAAGAAAACCTAGG   119
6059s1-6059L       ATGGATTAAAGACTTAAACGTTAGACCTAAAACCATAAAAACCCTAGAAGAAAACCTAGG   119
6059s2-6059L       ATGGATTAAAGACTTAAATGTTAGACCTAAAACCATAAAAACCCTAGAAGAAAACCTAGG   119
                   ****************** ************************************

6059s3-6059L       CATTACCATTCAGGACATAGGCATGGGCAAGGACTTCATGTCTAAAACACCAAAAGCAAT   179
6059s1-6059L       CATTACCATTCAGGACATAGGCATGGGCAAGGACTTCATGTCTAAAACACCAAAAGCAAT   179
6059s2-6059L       CATTACCATTCAGGACATAGGCATGGGCAAGGACTTCATGTCTAAAACACCAAAAGCAAT   179
                   ************************************************************

6059s3-6059L       GGCCATCCGCGTTCTGCCCTAAGCGCTTCCTCGGGTGGCCGGCTAGCGTACAGGCAGTCC   239
6059s1-6059L       GGCA-----------------------------------------------------   183
6059s2-6059L       GGCA-----------------------------------------------------   183
                   ***.

6059s3-6059L       GACCAGCTCGAATATAGAGAGGTG       263
6059s1-6059L       -----------------------        183
6059s2-6059L       -----------------------        183
```

Figure 21

```
CLUSTAL O(1.2.4) multiple sequence alignment

279s5-279L      TATTTGAAC-TTTTTAAATATAAGTTTGAGGGGAAAATGAACTAAGTTAGTAATGTTTTG   59
279s4-279L      --AACNACCTTTTTAAAATATAAATTTGAGGGGAAAATGAACTAAGTTAGTAATGTTTTG   58
279s6-279L      ------AACT-TTTTAAATATAAGTTTGAGGGGAAAATGAACTAAGTTAGTAATGTTTTG   53
                      *.*  ***:**********************************************

279s5-279L      GGAATAAGATCATGGGTTGGCCACACCAAGCAGGTAATGGCACAAGGATAAGATAATTAA   119
279s4-279L      GGAATAAGATCATGGGTTGGCCACACCAAGCAGGTAATGGCACAAGGATAAGATAATTAA   118
279s6-279L      GGAATAAGATCATGGGTTGGCCACACCAAGCAGGTAATGGCACAAGGATAAGATAATTAA   113
                **********************************************************

279s5-279L      CAAGCAGCCGGGCCGGCGGGG--GGGGGGGGGG--CGCTCCACGCCGGGGGGGGGGCTAG-   174
279s4-279L      CAATCTTCGGTCCTGGTTTGG-TGGACTGAGGCGCCCACCCGGCCTGGGAGGGCAGTGAG   177
279s6-279L      CAACACTCGGTTGAGATTTGGGGTGGGGTTGTCCCGCCACATGGGAAGGGTAAGGA   173
                ***  .  * *   *.   **   **.   *.**   * *  . .*. ***..**   :..

279s5-279L      CGGGGCACCGGCGGGAGCTGCT-GCCGAAAAAAAGGGGGGGCCCCCCTGTGTGTTCATCGTG   233
279s4-279L      GTGGACAGCTGCAGGAGCTGGTCAACGAAAAGCAGGGGGACCACGGTTGTGTTCATCCTG   237
279s6-279L      GAAGAAAGTGGAAACAGCTGTTCAA---AAAAAAGGGGGGACACCCTGGTGTACATCCTG   230
                  .*..*   *.... ***** * ..   ***..*******..*.*     ****:**** **

279s5-279L      TGGGGGGGGCGCGAGAAGATTCTCTCTCTGTT----------------CCTCCACACAC-   275
279s4-279L      TGGGGAGGCGCCAGATTA-----CCTCGGCCCTGAA-------------CCTACACAAACC   280
279s6-279L      TTGGGGGGGCGCCACATAAT---CCCTCTGCTCCCTCCTAAAATCGAACCCCCCCTCTTGC-   286
                * ***.***** * *: *      *** *                  ** .*:*: .*

279s5-279L      ATGGAGATTCCT-----------------------   287
279s4-279L      TAGGATAATCCTTGGGGGGAAGGGTTGTAAAA   312
279s6-279L      ------------------------------   286
```

Figure 22

```
CLUSTAL O(1.2.4) multiple sequence alignment

139s8-139L      ----GANATGCATACGAAATGTTTGTGGAGCACCTACTGGGAGGCACTGGGGATACAGTA    56
139s7-139L      ATTCGCGTATGCATACCAATGTTTGTGGAGCACCTACTGGGAGGCACTGGGGATACAGTA    60
139s9-139L      -ANNNGCGACTGATACCAATGTTTGTGGAGCACCTACTGGGAGGCACTGGGGATACAGTA    59
                                *****************************************

139s8-139L      GTGACCTAGAAGGCCATGATCCCCGAGCTGTTATGTTCAGAAGGAGGAGCTCCTCGGT--    114
139s7-139L      GTGACCTAGAAGGCCATGATCCCTGCGCCCATTCCCTTCTCAGGGAGCGGAGTTCCTTGG    120
139s9-139L      GTGACCTAGAAGGCCATGATCCC-------------------------------------    82
                ***********************

139s8-139L      -GCTGACATGATTGTAGTGAGGTGGCCACCCCGCCTGGGAGGTGGGTGAGGGGTAGAGAG    173
139s7-139L      GGCTCGAAGGGTGGGGGAGGGCTCCCCCCCCCTTGTGGGGTGTGA----GGGCAACCTGT    176
139s9-139L      ------------------------------------------------------------    82

139s8-139L      GCGGCAACTGCTTAGAAA------------------------------------------    191
139s7-139L      TCAACTGCTATCGAAGAAAAGGGGGGAACAGTTGTTCCACCCTGTGGGGAGGCGCCAAAT    236
139s9-139L      ------------------------------------------------------------    82

139s8-139L      ------------------------------------    191
139s7-139L      AATCTCGCTCTGCACCCCCACAAACATAGGAAAATCT    273
139s9-139L      ------------------------------------    82
```

Figure 23

6 **agcaa tgatacaatg tgctatctga** aaaaaaactaa actttttaaa atataaattt
gaggggaaaa tgaactaagt tagtaatgtt ttgggaataa gatcatgggt tggccaccacc
aaccaggtaa **tggcacaagg ataagataat taaca** 155

Figure 24