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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 March 2006 (30.03.2006)

PCT

(10) International Publication Number
WO 2006/034264 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01) *C07H 21/04* (2006.01)

(21) International Application Number:
PCT/US2005/033639

(22) International Filing Date:
21 September 2005 (21.09.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/611,779 21 September 2004 (21.09.2004) US

(71) Applicant *(for all designated States except US)*: **APPLERA CORPORATION** [US/US]; 850 Lincoln Centre Drive, Foster City, CA 94404-1128 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **ZON, Gerald** [US/US]; 199 Fir Street, San Carlos, CA 94070 (US).

(74) Agent: **PATANE, Michael, A.**; Cozen O'Connor, 1900 Market Street, Philadelphia, PA 19103 (US).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2006/034264 A2

(54) Title: METHODS OF USING SULFUR NUCLEOPHILES AS IMPROVED ALTERNATIVES TO SODIUM BISULFITE FOR METHYLATED DNA ANALYSIS

(57) Abstract: The invention provides for the use of sulfur nucleophiles in analyzing methylated DNA and novel sulfur nucleophiles suitable for such us.

METHODS OF USING SULFUR NUCLEOPHILES AS IMPROVED ALTERNATIVES TO SODIUM BISULFITE FOR METHYLATED DNA ANALYSIS

FIELD OF INVENTION

[0001] The invention relates generally to sulfur nucleophiles and methods of using them for analysis of methylated DNA.

BACKGROUND OF THE INVENTION

[0002] Assessment of the methylation of DNA is useful in many research, diagnostic, medical, forensic, and industrial fields. Particularly, methylation of cytosine in genomic DNA has been correlated with lack of gene expression, and in some instances can be indicative of early and frequent alterations found in some cancers. Thus, the ability to assess the methylation status of DNA is significant.

[0003] Key to this assessment is converting cytosine to uracil. One basic method for such conversion, employing sodium bisulfite ($NaHSO_3$), is well known. Over the years, the method has been improved in attempts to overcome disadvantages that include tedious procedures, lengthy reaction times, and DNA degradation. The most commonly used protocol is taught by J. Herman, *Proc. Natl. Acad. Sci.* **93**, 9821-26 (1996), incorporated herein by reference in its entirety. This method involves denaturation, reaction with sodium bisulfite in the presence of hydroquinone, and subsequent completion of the modification by treatment with NaOH. Despite the attempts to improve the protocol, it is still required to pre-denature the genomic DNA (gDNA) to single stranded DNA (ssDNA), prepare fresh solutions of sodium bisulfite, typically about 3M, and include an antioxidant (e.g.,

hydroquinone). The protocol also involves long reaction times and tedious clean-up procedures.

[0004] In addition, the currently employed sodium bisulfite protocols are plagued by reports of incomplete conversion, irreproducible results, and other problems. In some cases, the reaction can result in significant DNA degradation (reportedly as high as 96%), making it difficult to obtain enough sample for further analysis. See. S.J. Clark et al. *Nucleic Acid Research* 2001, **29** no. 13, e65.

[0005] Other methods exist to assess methylation status. Many of these methods use labeling technology. For example, radio-labeled samples can be compared to internal standards by GC-MS (P. F. Crain and J. A. McCloskey. *Anal. Biochem.* (1983) 132, 124-131). Fluorescent or chemiluminescent moieties may be used to assess methylation status through optical detection means. These usually require sophisticated and expensive HPLC or CE equipment operated by experts (M. Wirtz et al. *Electrophoresis* (2004) 25, 839-845; D. Stach et al. *Nucleic Acids res.* (2003) 31, E2.). One current approach, useful in analyzing CpG islands, is restriction landmark genome scanning (RLGS), which is based on digestion of DNA with methylation-sensitive restriction enzymes, radiolabeling and then 2D-gel separation (D. J. Smiraglia et al. *Genomics* (1999) 58, 254-262.). RLGS is therefore limited to only those CpG islands which contain sites compatible with available restriction enzymes.

[0006] Given the importance of assessment of DNA methylation, it can be seen that there is a need for improved methodologies for conversion of cytosine to uracil and for assessing the methylation status of DNA.

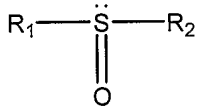
SUMMARY OF THE INVENTION

[0007] In some embodiments, a method for converting cytosine to uracil in a nucleic acid comprises the steps of:

providing a nucleic acid comprising at least one cytosine nucleobase; and

reacting said nucleic acid with a nucleophilic organo-sulfur compound.

[0008] In some embodiments, a nucleophilic organo-sulfur compound Formula I:



I

wherein R_1 and R_2 are each independently selected from the group consisting of hydroxyl, alkyl, aryl, amino, alkoxy, and aryloxy, each of which may be optionally substituted; and

or R_1 and R_2 can be concatenated to form a 4-8 membered ring optionally having 1 or 2 additional hetero ring atoms selected from N, S, and O, wherein said ring can be optionally substituted with one or more substituents;

or a salt thereof

is reacted with a nucleic acid comprising at least one cytosine nucleobase, prior to assessment of methylation status.

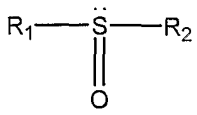
[0009] In some embodiments, the methods herein are carried out with a salt of formula I where one or both of R_1 and R_2 forms an ionic bond (or salt pair) with a cation selected from lithium, sodium, magnesium and ammonium. In such embodiments, one or both R_1 and R_2 may comprise(s) an anionic group capable of forming such ionic bond or salt pair.

[0010] In some embodiments, a method for assessing the methylation status of cytosine comprises the steps of:

providing a sample nucleic acid comprising at least one cytosine nucleobase of unknown methylation status; and

reacting said nucleic acid with a nucleophilic organo-sulfur compound comprising a radio-labeled substituent.

[0011] In some embodiments, the nucleophilic organo-sulfur compound is a compound of formula I:



I

wherein R_1 and R_2 are each independently selected from the group consisting of hydroxyl, alkyl, aryl, amino, alkoxy, and aryloxy, and a radiolabel substituent, wherein each of said alkyl, aryl, amino, alkoxy, and aryloxy can be optionally substituted;

wherein at least one of R_1 and R_2 comprises a radio-labeled substituent;

or a salt thereof.

[0012] In some embodiments, such methods further provide the steps of:

providing a control nucleic acid comprising at least one cytosine nucleobase of known non-methylated status;

reacting said nucleic acid with the same said nucleophilic organo-sulfur compound; and

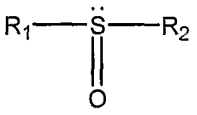
comparing the level of radioactivity of the sample and control to determine the relative content of methylated cytosine in the sample based on the rates of reaction of the methylated cytosine and unmethylated cytosine.

[0013] In some embodiments, a method for assessing the methylation status of cytosine comprises the steps of:

providing a sample nucleic acid comprising at least one cytosine nucleobase of unknown methylation status; and

reacting said nucleic acid with a nucleophilic organo-sulfur compound comprising a fluorescent or chemiluminescent moiety.

[0014] In some embodiments, the nucleophilic organo-sulfur compound comprising a fluorescent or chemiluminescent moiety is a compound of formula I:



I

wherein R_1 and R_2 are each independently selected from the group consisting of hydroxyl, alkyl, aryl, amino, alkoxy, and aryloxy, and a radiolabel substituent, wherein each of said alkyl, aryl, amino, alkoxy, and aryloxy can be optionally substituted;

wherein at least one of R_1 and R_2 comprises a fluorescent or chemiluminescent moiety;

or a salt thereof.

DETAILED DESCRIPTION

[0015] It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention. In this application, the use of the singular includes the plural unless specifically stated otherwise. In this application, the use of "or" means "and/or" unless stated otherwise. The use of the term "comprising," as well as other forms, such as "comprises" and "comprise," will be considered inclusive, in that the term "comprising" leaves open the possibility of including additional elements. Furthermore, the use of the terms "including" or "having", as well as other forms, such as "includes", "has", "included", and "have" is not intended to be limiting. Also, terms such as "element" or "component" encompass both elements and components comprising one unit and elements and components that comprise more than one subunit unless specifically stated otherwise.

[0016] The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described.

[0017] The term "alkyl" refers to straight and branch chain hydrocarbon groups, such as, but not limited to, methyl, ethyl, propyl, butyl, pentyl, hexyl, heptyl, octyl, nonyl, decyl, undecyl, dodecyl and the like. The term also includes branched chain isomers of straight chain alkyl groups, including but not limited to, the following which are provided by way of example: $-CH(CH_3)_2$, $-CH(CH_3)(CH_2CH_3)$, $-CH(CH_2CH_3)_2$, $-C(CH_3)_3$, $-C(CH_2CH_3)_3$, $-CH_2CH(CH_3)_2$, $-CH_2CH(CH_3)(CH_2CH_3)$, $-CH_2CH(CH_2CH_3)_2$, $-CH_2C(CH_3)_3$, $-CH_2C(CH_2CH_3)_3$, $-CH(CH_3)CH(CH_3)(CH_2CH_3)$, $-CH_2CH_2CH(CH_3)_2$, -

$CH_2CH_2CH(CH_3)(CH_2CH_3)$, $-CH_2CH_2CH(CH_2CH_3)_2$, $-CH_2CH_2C(CH_3)_3$, $-CH_2CH_2C(CH_2CH_3)_3$, $-CH(CH_3)CH_2CH(CH_3)_2$, $-CH(CH_3)CH(CH_3)CH(CH_3)_2$, $-CH(CH_2CH_3)CH(CH_3)CH(CH_3)(CH_2CH_3)$, and others. The term also includes cyclic alkyl groups such as cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, cycloheptyl, and cyclooctyl and such rings substituted with straight and branched chain alkyl groups as defined above. Thus alkyl groups include primary alkyl groups, secondary alkyl groups, and tertiary alkyl groups. Preferred alkyl groups include straight and branched chain alkyl groups and cyclic alkyl groups having 1 to 12 carbon atoms.

[0018] The term "alkoxy" refers to a group of formula –O-alkyl, where alkyl is as defined above. Examples include but are not limited to –OMe, -O Et, and the like.

[0019] The term "aryl" is intended to denote a radical derived from a compound that contains at least one aromatic ring. Thus, aryl groups include, but are not limited to, groups such as phenyl and biphenyl, and groups containing condensed rings such as naphthalene and anthracene. A preferred unsubstituted aryl group is phenyl.

[0020] The term "amino" refers to a nitrogen having two substituents. The substituents are independently selected and include, but are not limited to, hydrogen, hydroxyl, alkyl, aryl, etc. and may be optionally substituted. Most preferred are hydrogen, methyl, ethyl, propyl, isopropyl, 2-hydroxyethyl, and 2-methoxyethyl.

[0021] The term "aryloxy" refers to a group of formula –O-aryl, where aryl is as defined above. One non-limiting example of an aryloxy group is a phenoxy group; i.e., a group of formula –OPh where Ph is phenyl.

[0022] The term "bisulfite ion," as used herein, has its accustomed meaning of HSO_3^-. Typically, bisulfite is used as an aqueous solution of a bisulfite salt, for example

magnesium bisulfite, which has the formula $Mg(HSO_3)_2$, and sodium bisulfite, which has the formula $NaHSO_3$.

[0023] The phrase "optionally substituted" refers to groups in which one or more hydrogen atoms have been replaced by a non-hydrogen substituent group. Such groups include, but are not limited to, halogen atoms such as F, Cl, Br, and I; hydroxyl groups, alkyl groups, alkenyl groups, alkynyl groups, aryl groups, alkoxy groups, aryloxy groups, ester groups; thiol groups, alkyl and aryl sulfide groups, sulfone groups, sulfonyl groups, sulfoxide groups, amines, amides, alkylamines, dialkylamines, arylamines, alkylarylamines, diarylamines, N-oxides, imides, enamines, trialkylsilyl groups, dialkylarylsilyl groups, alkyldiarylsilyl groups, and triarylsilyl groups.

[0024] The term "PCR" is intended to denote polymerase chain reaction, as is well known in the art. The term "MSP" denotes methylation specific PCR, such as described by J. Herman, *Proc. Natl. Acad. Sci.* **93**, 9821-26 (1996), incorporated herein by reference in its entirety.

[0025] The term "nucleic acid sample" is intended to denote a sample (e.g., a composition, mixture, suspension or solution) that contains at least one nucleic acid.

[0026] As used herein, the term "nucleic acid" includes nucleobase-containing polymeric compounds, including naturally occurring and non-naturally occurring forms thereof, for example and without limitation, genomic DNA, cDNA, hnRNA, mRNA, rRNA, tRNA, fragmented nucleic acids, nucleic acids obtained from subcellular organelles such as mitochondria or chloroplasts, and nucleic acids obtained from microorganisms, or DNA or RNA viruses that may be present on or in a biological sample.

[0027] As used herein, the term "gDNA" refers to genomic DNA.

[0028] "Fluorescent moiety," as used herein, means a moiety that fluoresces (i.e. emits light of a certain wavelength) when exposed to radiation. Examples of such moieties include but are not limited to 6-carboxyfluorescein or 6-carboxytetramethylrhodamine.

[0029] "Chemiluminescent moiety" means a moiety that allows chemiluminescent activity (i.e. generation of light by chemical reaction) to be detected by optical means. Examples of such moieties include but are not limited to acridinium esters and derivatives thereof.

[0030] "Nucleophilic organo-sulfur compound" as used herein refers to those compounds having a lone pair of electrons at sulfur. Preferred nucleophilic organo-sulfur compounds are substituted derivatives of sulfinic acid. Most preferred are those of formula I, discussed below.

[0031] There are a wide variety of compounds which can formally be viewed as derivatives of the HO-S(:)(O)-OH moiety that preserve the nucleophilic lone-pair of electrons (:) at sulfur. While not wishing to be bound by a particular theory, it is believed that this nucleophilic lone-pair of electrons at sulfur modulates the specificity and rate of the reversible adduct formation with cytosine which in turn influences the subsequent irreversible hydrolysis to generate uracil. Consequently, certain derivatives of HO-S(:)(O)-OH may have desirable features with regard to cytosine-to-uracil conversion prior to analyses of methylated DNA.

[0032] The nucleophilicity of the sulfur compounds has been indicated as the basis of attack of sulfur at carbon in an aromatic ring (A. Ulman and E. Urankar, J. Org. Chem. (1989) 54, 4691-4692), at an unsaturated (acetylenic) carbon (T. Kataoka et al. Phosphorus, Sulfur and Silicon and the Related Elements (1998) 136/138, 497-500), and at the carbon-carbon double bond in acrylonitrile (I. V. Bodrikov et al. Z. Org. Khim. (1985) 21, 1017-

1022). Each supports the present invention that the nucleophilicity of the sulfur compounds provides the basis for reaction with cytosine to yield uracil. None, however, teaches the conversion of cytosine to uracil.

General Preparation Of Nucleophilic Organo-Sulfur Compounds

[0033] Mono-substituted organo-sulfur nucleophiles are made by replacing one -OH moiety attached to sulfur, S, with alkyl, aryl, amino, alkoxy, or aryloxy groups, which may in turn be substituted with various other groups. The remaining -OH group may be used to form a salt, preferably lithium or magnesium, more preferably sodium, and therefore be ionic.

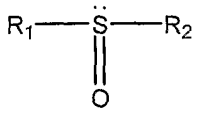
[0034] Bis-substituted, non-ionic compounds may also be formed where both -OH groups are replaced, independently, with alkyl, aryl, amino, alkoxy, or aryloxy groups, which in turn may be substituted with various other groups.

[0035] A variety of mono-substituted and bis-substituted derivatives of HO-S(:)(O)-OH, including sodium, lithium, and magnesium salts thereof, are known in the art. Form example, derivatives of HO-S(:)(O)-OH are found in FR 2,288,086, hereby incorporated by reference. FR 2,288,086 also discloses sulfinic acids where one -OH is replaced with an alkyl group and sulfinic esters where one OH is replaced by an alkyl group and the other by an alkoxy group are disclosed.

[0036] Dialkyl sulfates where both -OH moieties of the bisulfite are replaced with alkoxy groups are disclosed in M. Mikolyczyk and coworkers in *Tetrahedron* (1988) 44 (16) 5243, which is hereby incorporated in its entirety by reference.

Exemplary Sulfur Nucleophiles

[0037] Some embodiments of the methods of the present invention employ sulfur nucleophiles according to formula I:



I

wherein R_1 is selected from the group consisting of hydroxyl, alkyl (R), aryl (Ar), amino (NR_3R_4), alkoxy (OR_5), and aryloxy (ArO), each of which may be optionally substituted, and each of which optionally may be labeled with one of a radio-marker, a fluorescent moiety, and a chemiluminescent moiety;

wherein R_2 is selected from the group consisting of hydroxyl, alkyl (R), aryl (Ar), amino (NR_3R_4), alkoxy (OR_5), and aryloxy (ArO), each of which may be optionally substituted and each of which optionally may be labeled with one of a radio-marker, a fluorescent moiety, and a chemiluminescent moiety;

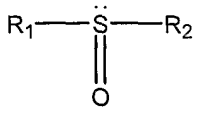
or, wherein R_1 and R_2 are concatenated to form a 4-8 membered ring optionally having 1-2 additional hetero ring atoms selected from N, S, and O, and optionally substituted with one or more substituents;

R_3 and R_4 are each independently selected from the group consisting of alkyl, substituted alkyl, aryl, and substituted aryl;

R_5 is an alkyl or substituted alkyl;

or, a salt thereof, such as a lithium, sodium, ammonium or magnesium salt wherein one of R_1 and R_2 forms an ionic bond with a halide ion.

[0038] Some representative mono-substituted sulfur nucleophiles of formula I where R_2 is –OH, and salts thereof, are listed in table 1:



Compound	Structure (R_2-S(O)-R_1)
sulfurous acid, monomethyl ester (monomethyl sulfite)	HO-S(O)-OCH_3
sulfurous acid, monomethyl ester, lithium salt (lithium methyl sulfite)	Li • O-S(O)-OCH_3
sulfurous acid, monomethyl ester, sodium salt (methyl sodium sulfite)	Na • O-S(O)-OCH_3
sulfurous acid, monomethyl ester, magnesium salt	½Mg • O-S(O)-OCH_3
sulfurous acid, monoethyl ester (monoethyl sulfite)	HO-S(O)-OCH_2CH_3
sulfurous acid monoethyl ester, lithium salt	Li • O-S(O)-OCH_2CH_3
sulfurous acid monoethyl ester, sodium salt (ethyl sodium sulfite; sodium ethyl sulfite)	Na • O-S(O)-OCH_2CH_3
sulfurous acid monoethyl ester, magnesium salt	½ Mg • O-S(O)-OCH_2CH_3
sulfurous acid, monopropyl ester	HO-S(O)-$OCH_2CH_2CH_3$
sulfurous acid, monopropyl ester, lithium salt	Li • O-S(O)-O $CH_2CH_2CH_3$
sulfurous acid, monopropyl ester, sodium	Na • O-S(O)-O $CH_2CH_2CH_3$

Compound	Structure (R_2-S(O)-R_1)
salt	
sulfurous acid, monopropyl ester, magnesium salt	½ Mg • O-S(O)-$OCH_2CH_2CH_3$
sulfurous acid, monophenyl ester (phenyl hydrogen sulfite)	HO-S(O)-OPh
sulfurous acid, monophenyl ester, sodium salt (sodium phenyl sulfite)	Na • O-S(O)-OPh
methanesulfinic acid (methylsulfinic acid)	HO-S(O)-CH_3
sodium methanesulfinate	Na • O-S(O)-CH_3
ethanesulfinic acid (ethylsulfinic acid)	HO-S(O)-CH_2CH_3
sodium benzenesulfinate	Na • O-S(O)-Ph
methylamidosulfurous acid	HO-S(O)-$NHCH_3$
sodium dialkylamidosulfinate	Na • O-S(O)-NRR' R = R' = Me or Et or $(CH_2)_5$

[0039] This list of compounds is exemplary only and is not intended to be limiting in any manner.

[0040] A representative synthesis of the monomethyl, monoethyl, and monoisopropyl ester sodium salts, AND THE synthesis of the corresponding dimethyl, diethyl, and di-isopropyl esters, is described by A. B. Foster et al. J. of the Chemical Soc. (1956) 2589-2592, incorporated herein by reference in its entirety. Synthesis of sodium dialkylamidosulfinate compounds (Na • O-S(O)-NRR') wherein R = R' = Me or Et or $(CH_2)_5$ was reported by A.

Blaschette and H. Safari, Z. *fuer Naturforsch.*(1970) 25, 319-320, also incorporated herein by reference in its entirety.

[0041] Some representative bis-substituted sulfur nucleophiles of formula I are listed in table 2:

R1	**R2**	**COMPOUND**
OMe	OMe	Sulfurous acid dimethyl ester
OEt	OEt	Sulfurous acid diethyl ester
$N(Me)_2$	OMe	Dimethyl-sulfinamic acid methyl ester
$N(Me)_2$	$N(Me)_2$	
Me	Me	Methanesulfinylmethane
CMe_3	NEt_2	2-Methyl-propane-2-sulfinic acid diethylamide
$-O-(CH_2)_2-O-$ (forming a ring with S)		[1,3,2]Dioxathiolane 2-oxide
$-N(Et)-(CH_2)_2N(Et)-$ (forming a ring with S)		2,5-Diethyl-[1,2,5]thiadiazolidine 1-oxide

[0042] Of course these are non-limiting examples and are presented by way of illustration only. The table is not intended to limit the scope of the invention.

[0043] As discussed above, the substitutents of R_1 and R_2 may include various markers. These markers may be radio-labels, fluorescent moieties, or chemiluminescent moieties.

[0044] Radio labels are atoms or compounds that contain an atom that undergoes a process resulting in the emission of a photon, electron or other nuclear constituent, thus

allowing their detection. Suitable radio-labels include, but are not limited to, ^{3}H and ^{14}C. These markers may be incorporated into any of the various substituents of R_1 and R_2.

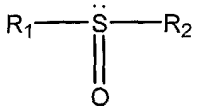
[0045] The present invention is amenable to the use of a wide variety of fluorescent and chemiluninescent moieties, as are known in the art. Non-limiting examples of suitable fluorescent moieties include 6-carboxyfluorescein or 6-carboxytetramethylrhodamine. Suitable chemilumiscent moieties include, but are not limited to acridinium esters and derivatives thereof.

[0046] Those of ordinary skill in the art will readily recognize appropriate methods of making mono- and bis-substituted sulfur nucleophiles as well as salts and labeled versions thereof..

Methods

[0047] According to some embodiments of the methods of the invention, a nucleic acid sample, containing a nucleic acid comprising at least one cytosine nucleobase is reacted with a nucleophilic organo-sulfur compound to facilitate conversion of cytosine to uracil for further assessment according to known techniques to determine methylation status. Such reactions may be performed by suitable adaptation of standard techniques for converting cytosine to uracil by using organo-sulfur compounds of the present invention in place of (or in addition to) bisulfite. For example, genomic DNA (1 microgram or less) is denatured for 15 to 30 minutes at 45°C with NaOH (2M to 3M), followed by incubation with 0.1M hydroquinone and 3.6M sodium bisulfite (pH 5.0) at 55°C for 12 hours or overnight. The DNA is then purified from the reaction mixture using standard miniprep columns, for example. For desulfuration, the purified DNA sample is resuspended in aqueous 0.25M NaOH (60 microliters) is incubated at 40°C for 5-10 minutes. The desulfurated DNA can then be ethanol-precipitated and washed, followed by resuspension in water.

[0048] In some embodiments of the invention, a method for converting cytosine to uracil includes the step of reacting a nucleic acid comprising at least one cytosine nucleobase with a nucleophilic organo-sulfur compound, or a salt thereof, according to Formula I:



I

wherein R_1 and R_2 are as described above.

[0049] Reaction of the nucleophilic organo-sulfur compound with the cytosine containing nucleic acid results in specific conversion of cytosine, but not 5-methyl cytosine, to uracil. Upon conversion, known techniques, such as PCR, MSP, and other techniques, may be used to assess the methylation status of the sample.

[0050] In some embodiments, the nucleophilic organo-sulfur compound is a mono-substituted compound where R_2 is –OH and R_1 is as described above. Preferably, R_1 is selected from methyl or ethyl.

[0051] Again, upon completion of the reaction, known techniques may be used to assess the methylation status of the sample by analyzing conversion data.

[0052] Salts of the mono-substituted nucleophilic organo-sulfur compounds may also be employed. Particularly, the –OH group of R_2 is hydrolyzed to form an ionic bond with a cation. Preferred cations are lithium, sodium, ammonium or magnesium, and more particularly sodium, with the proviso that the compound is not a bisulfite compound.

[0053] In some embodiments, the nucleophilic compound is a bis-substituted compound where each of R_1 and R_2 is other than hydroxyl. R_1 is preferably methyl or ethyl and R_2 is preferably methyl or ethyl. As with the other embodiments herein, upon conversion, methylation status is assessed by known techniques.

[0054] Other embodiments employ the use of labels and detection of differing levels of labeling to determine methylation status. Labeling schemes in conjunction with existing single-molecule DNA-scanning procedures, or AFM (atomic force microscopy) technology, and other technologies, provides a powerful tool for discovery and analysis of, for example, methylated promoters of genes without the limitations associated with currently used RLGS methodology.

[0055] For example, either R_1 or R_2, or both, can include ^3H or ^{14}C labels for measurement of total 5-methyl cytosine vs. non-methylated cytosine content. Other radio-labels may be used as well. In this type of assay the DNA sample of interest (S) and a control sample (C) are separately reacted with the labeled reagent under identical reaction conditions. Following removal of excess labeled reagent, the difference in radioactivity of these two samples (R_S and R_C, respectively) provides a relative measure of 5-Methyl Cytosine content based on the expected differential reactivity of 5-methyl cytosine compared to Cytosine, namely, Cytosine reacts much more rapidly than 5-Methyl Cytosine. For example, if R_S = 1000 counts/nucleotide-equivalent and R_C = 2000 counts/nucleotide-equivalent, then sample of interest, S, is 50% methylated. Synthetic internal standards comprised of fully-methylated and non-methylated oligonucleotide sequences may be used as controls to normalize the raw data by correcting for low-levels of non-specific or incomplete reaction, respectively.

[0056] The labeling technique can be extended beyond radio-labels to marking with fluorescent moieties or moieties that allow chemiluminescence to be detected. Current optical methods, not employing differential labeling require use of sophisticated and expensive HPLC or CE equipment, and experienced operators. A significant advantage of differential labeling of methylated DNA using such reagents is that it provides a means of optically detecting sites of methylation such as CpG islands in promoter regions of genes.

DNA reacted with fluorescent-labeled sulfur nucleophiles may be used with existing single-molecule DNA-scanning methods (S. Zhou et al. Genome res. (2003) 13, 2142-2151) to enable a method for genome-wide analysis of methylated promoters that does not require the use of radio labels and, moreover, is not limited to promoter regions having methylation-specific sites.

[0057] In this new approach, the elongated single-molecules of DNA are first imaged using YOYO-1 dye, as described (S. Zhou et al. Genome res. (2003) 13, 2142-2151), followed by removal of this dye and reaction with fluorescently labeled sulfur nucleophiles such that the DNA of interest is labeled in one color and the control DNA is labeled in a second color. The latter images are electronically subtracted such that 5-methyl cytosine is seen as a positive signal, which is then overlayed on the whole-genome map derived from the YOYO-1 data, as described (S. Zhou et al. Genome res. (2003) 13, 2142-2151.). In this manner, the methylated promoter regions of all genes are seen and identified by comparison with the relevant genome sequence.

[0058] Improved signal-to-noise ratios can be obtained by use of sulfur nucleophiles of formula I where R_1 and/or R_2 provide moieties that allow chemiluminescent imaging.

[0059] A fundamentally different scanning approach would use an adaptation of atomic force microscopy (K. Virnik et al. J. Mol. Biol. (2003) 334, 56-63.). The basic idea is to analyze differentially reacted DNA as an AFM-difference readout.

[0060] Analysis of total methylated DNA content by the above methods is relatively simple and does not require sophisticated, costly equipment operated by experts. These features are particularly advantageous in clinical settings.

[0061] In some embodiments, a method for converting cytosine to uracil includes the step of reacting a nucleic acid comprising at least one cytosine nucleobase with a mixture

including a bisulfite ion and a nucleophilic organo-sulfur compound according to formula I above, or a salt thereof, according to Formula I. In this reaction, it is contemplated that the bisulfite ion reacts more quickly than the nucleophillic organo-sulfur compounds. The bisulfite is then displaced by the nucleophillic organo-sulfur compound. Methylation status may then be assessed according to known techniques. This approach may be used with labeled and unlabeled nucleophiles, but is particularly preferred with the labeled nucleophiles.

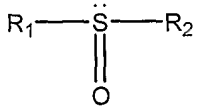
[0062] The examples described herein have been chosen to illustrate the invention, and are not intended to be limiting. Those reasonably skilled in the art will readily recognize additional embodiments that do not differ from the scope and spirit of the invention disclosed herein.

CLAIMS

What is claimed is:

1. A method for converting cytosine to uracil in a nucleic acid comprising the steps of:
providing a nucleic acid comprising at least one cytosine nucleobase; and
reacting said nucleic acid with a nucleophilic organo-sulfur compound.

2. The method of claim 1 wherein said nucleophilic organo-sulfur compound is a compound of Formula I:



I

wherein

R_1 and R_2 are each independently selected from the group consisting of hydroxyl, alkyl, aryl, amino, alkoxy, and aryloxy, each of which may be optionally substituted; and

or R_1 and R_2 can be concatenated to form a 4-8 membered ring optionally having 1 or 2 additional hetero ring atoms selected from N, S, and O, wherein said ring can be optionally substituted with one or more substituents;

or a salt thereof.

3. The method of claim 2 wherein said amino of said R_1 and said R_2 has the formula NR_3R_4, and said alkoxy of said R_1 and said R_2 has the formula OR_5; and

wherein R_3, R_4 and R_5 are each independently selected from the group consisting of hydrogen, methyl, ethyl, propyl, isopropyl, 2-hydroxyethyl, and 2-methoxyethyl.

4. The method of claim 2 wherein said organo-sulfur compound is a salt of formula I where one of R_1 and R_2 forms an ionic bond with a cation selected from lithium, sodium, magnesium and ammonium.

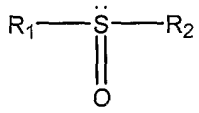
6. The method of claim 2, wherein said reacting step is carried out with a mixture comprising a bisulfite ion and said nucleophillic organo-sulfur compound.

7. A method for assessing the methylation status of cytosine comprising the steps of:

providing a sample nucleic acid comprising at least one cytosine nucleobase of unknown methylation status; and

reacting said nucleic acid with a nucleophilic organo-sulfur compound comprising a radio-labeled substituent.

8. The method of claim 7 wherein said nucleophilic organo-sulfur compound is a compound of formula I:



I

wherein

R_1 and R_2 are each independently selected from the group consisting of hydroxyl, alkyl, aryl, amino, alkoxy, and aryloxy, and a radiolabel substituent, wherein each of said alkyl, aryl, amino, alkoxy, and aryloxy can be optionally substituted;

wherein at least one of R_1 and R_2 comprises a radio-labeled substituent;

or a salt thereof.

9. The method of claim 8 further comprising the steps of:

providing a control nucleic acid comprising at least one cytosine nucleobase of known non-methylated status;

reacting said nucleic acid with the same said nucleophilic organo-sulfur compound; and

comparing the level of radioactivity of the sample and control to determine the relative content of methylated cytosine in the sample based on the rates of reaction of the methylated cytosine and unmethylated cytosine.

10. The method of claim 8 wherein said amino of said R_1 and said R_2 has the formula NR_3R_4, and said alkoxy of said R_1 and said R_2 has the formula OR_5; and

wherein R_3, R_4 and R_5 are each independently selected from the group consisting of hydrogen, methyl, ethyl, propyl, isopropyl, 2-hydroxyethyl, and 2-methoxyethyl.

11. The method of claim 8 wherein said organo-sulfur compound is a salt of formula I where one of R_1 and R_2 forms an ionic bond with a cation selected from lithium, sodium, magnesium and ammonium.

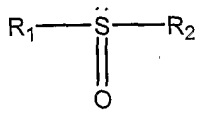
12. The method of claim 8, wherein said reacting step is carried out with a mixture comprising a bisulfite ion and said nucleophillic organo-sulfur compound.

13. A method for assessing the methylation status of cytosine comprising the steps of:

providing a sample nucleic acid comprising at least one cytosine nucleobase of unknown methylation status; and

reacting said nucleic acid with a nucleophilic organo-sulfur compound comprising a fluorescent or chemiluminescent moiety.

14. The method of claim 13 wherein said nucleophilic organo-sulfur compound is a compound of formula I:



I

wherein

R_1 and R_2 are each independently selected from the group consisting of hydroxyl, alkyl, aryl, amino, alkoxy, and aryloxy, and a radiolabel substituent, wherein each of said alkyl, aryl, amino, alkoxy, and aryloxy can be optionally substituted;

wherein at least one of R_1 and R_2 comprises a fluorescent or chemiluminescent moiety;

or a salt thereof.

15. The method of claim 14 wherein said amino of said R_1 and said R_2 has the formula NR_3R_4, and said alkoxy of said R_1 and said R_2 has the formula OR_5; and

wherein R_3, R_4 and R_5 are each independently selected from the group consisting of hydrogen, methyl, ethyl, propyl, isopropyl, 2-hydroxyethyl, and 2-methoxyethyl.

16. The method of claim 14 wherein said organo-sulfur compound is a salt of formula I where one of R_1 and R_2 forms an ionic bond with a cation selected from lithium, sodium, magnesium and ammonium.

17. The method of claim 14 wherein said radio-labeled substituents comprises one of ^{3}H and ^{14}C.

18. The method of claim 14 further comprising the steps of:

providing a control nucleic acid comprising at least one cytosine nucleobase of known non-methylated status;

reacting said control nucleic acid with the same said nucleophilic organo-sulfur compound; and

detecting and comparing the level of optical activity of the sample and control to determine the relative content of methylated cytosine.

19. The method of claim 14, wherein said reacting step is carried out with a mixture comprising a bisulfite ion and said nucleophillic organo-sulfur compound.