



(86) **Date de dépôt PCT/PCT Filing Date:** 2008/11/27
(87) **Date publication PCT/PCT Publication Date:** 2009/06/04
(45) **Date de délivrance/Issue Date:** 2017/01/17
(85) **Entrée phase nationale/National Entry:** 2010/05/25
(86) **N° demande PCT/PCT Application No.:** AU 2008/001751
(87) **N° publication PCT/PCT Publication No.:** 2009/067743
(30) **Priorité/Priority:** 2007/11/27 (AU2007906490)

(51) **Cl.Int./Int.Cl. C12Q 1/68** (2006.01),
C12N 15/11 (2006.01), **C12P 19/34** (2006.01)
(72) **Inventeur/Inventor:**
MILLAR, DOUGLAS SPENCER, AU
(73) **Propriétaire/Owner:**
HUMAN GENETIC SIGNATURES PTY LTD, AU
(74) **Agent:** GOWLING WLG (CANADA) LLP

(54) **Titre :** ENZYMES POUR L'AMPLIFICATION ET LA COPIE D'ACIDES NUCLEIQUES MODIFIES PAR DU BISULFITE
(54) **Title:** ENZYMES FOR AMPLIFICATION AND COPYING BISULPHITE MODIFIED NUCLEIC ACIDS

(57) **Abrégé/Abstract:**

The invention relates to the use of enzymes for copying or amplifying bisulphite modified or treated nucleic acids, wherein the enzymes are more effective in copying or amplifying the nucleic acid compared with native Taq polymerase under substantially the same conditions.



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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
4 June 2009 (04.06.2009)

PCT

(10) International Publication Number
WO 2009/067743 A1

(51) International Patent Classification:

C12Q 1/68 (2006.01) *C12P 19/34* (2006.01)
C12N 15/11 (2006.01)

(21) International Application Number:

PCT/AU2008/001751

(22) International Filing Date:

27 November 2008 (27.11.2008)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

2007906490 27 November 2007 (27.11.2007) AU

(71) Applicant (for all designated States except US): **HUMAN GENETIC SIGNATURES PTY LTD** [AU/AU]; Riverside Corporate Park, Level 4, 11 Julius Avenue, North Ryde, NSW 2113 (AU).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **MILLAR, Douglas, Spencer** [AU/AU]; 8 Rowland Avenue, Revesby, NSW 2212 (AU).

(74) Agent: **ALLENS ARTHUR ROBINSON PATENT & TRADE MARKS ATTORNEYS**; Cnr Hunter and Phillip Streets, Sydney, NSW 2000 (AU).

(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, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, 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, 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, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, 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:

- with international search report
- with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

(54) Title: ENZYMES FOR AMPLIFICATION AND COPYING BISULPHITE MODIFIED NUCLEIC ACIDS

(57) Abstract: The invention relates to the use of enzymes for copying or amplifying bisulphite modified or treated nucleic acids, wherein the enzymes are more effective in copying or amplifying the nucleic acid compared with native Taq polymerase under substantially the same conditions.

WO 2009/067743 A1

ENZYMES FOR AMPLIFICATION AND COPYING BISULPHITE MODIFIED NUCLEIC ACIDS

Technical Field

5 The present invention relates to use of enzymes for copying or amplifying nucleic acids, particularly nucleic acids treated with bisulphite.

Background Art

10 As a result of advances in automated sequencing technology, much work has been carried out on determining coding regions of DNA resulting in the full sequencing of many animal genomes including the human genome. It has been realised for many years that the majority of genomic DNA, however, is non-coding and this material was once considered as "junk" DNA. Analysis of the non-coding regions of DNA is now being considered as important in the study of gene expression and function. Methylation states
15 or patterns in nucleic acid, particularly genomic DNA, is thought to have a functional or regulatory role in gene expression and control in animals.

 It has been demonstrated that, in single stranded DNA, sodium bisulphite preferentially deaminates cytosine to uracil, compared to a very slow rate of deamination of 5-methylcytosine to thymine (Shapiro, R., DiFate, V., and Welcher, M, (1974) J. Am.
20 Chem. Soc. 96: 906-912). This observation served as the basis for the development of the bisulphite genomic sequencing protocol of Frommer *et al* 1992 (Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL and Paul CL. PNAS 89: 1827-1831 (1992)). In summary, this method as presently practiced involves the following general steps: alkaline denaturation of DNA; deamination using sodium
25 bisulphite; desulphonation by desalting followed by sodium hydroxide treatment; neutralization and desalting.

 One of the major disadvantages of the bisulphite modification procedure and the established variation thereof is that it has been shown that the procedure results in the degradation of between 84-96% of the original input DNA (Grunau et al. Nucleic Acids
30 Research 29 (13) e65, (2001)). The high loss associated with the procedure means that practically it is very difficult to successfully analyse small numbers of cells for their genomic methylation status, or successfully analyse ancient archival specimens in which the DNA is already in a partially degraded state. In addition, due to inherent nucleic acid

degradation of the current methods, it is not possible to sequence and assemble the complete genome of an organism to determine its genome-wide methylation profile in the same manner as has been successfully applied by the public Human Genome Project (International Human Genome Sequencing Consortium, 2001, *Nature*, 409, 860-921) or the private CELERA sequencing project (J Craig Venter et al., 2001, *Science*, 291, 1304-1351) owing to the huge number of "gaps" in the sequence.

A further disadvantage with the bisulphite method as presently practiced is that, in general, only small fragments of DNA can be amplified. Experience shows that generally less than about 500 base pairs (bp) can be successfully treated and amplified. The present technique is not applicable to new molecular biological methods such as Long Distance polymerase chain reaction (PCR) which has made it possible to amplify large regions of untreated genomic DNA, generally up to about 50 kb. At present, it is not even possible to analyse the methylation status of intact genes, as a large number of genes in mammalian genomes exceed 50 kb in length.

Thermostable polymerases in widespread use are unable to bypass the abasic sites generated during the bisulphite conversion and generally this causes stalling of the amplification reaction (Sikorsky, J.A., Primerano, D.A., Fenger, T.W. and Denvir, J. (2004) *Biochem. Biophys. Res. Commun.* 323, 823-230). In addition, these polymerases are also unable to successfully and efficiently amplify DNA which contains bulky adducts such as sulphonate groups. This necessitates desulphonation of the bisulphite converted nucleic acid at high temperatures in an alkaline medium prior to PCR amplification and results in the majority of the nucleic acid damage and loss seen during this procedure (Munson, K., Clark, J., Lamparska-Kupsik, K. and Smith, S.S. (2007) *Nucl. Acids. Res.* 35(9), 2893-2903). Furthermore, the generation of effectively a T-rich 3 base genome (as non-methylated C's are converted to U's and then into Ts during PCR amplification, giving rise to a genome comprised predominantly of bases A, T, G) results in significant difficulties for currently available polymerases and causes frequent slippage during extension. An additional problem encountered during PCR amplification of bisulphite converted DNA is that the single stranded template contains uracil which some polymerases, such as the archaeobacterial DNA polymerases like *Pfu Pwo* and *Vent*, are unable to process (Lasken, R.S., Schuster, D.M. and Rashtchian, A. (1996) *J. Biol. Chem.* 271 (30), 17692-17696). Currently, therefore, in order to investigate the methylation status of even relatively small genes (<4 kb), PCR reactions have had to be staggered across the gene region of interest (D.S Millar, K.K Ow, C.L. Paul, P.J. Russell,

P.L. Molloy, S.J. Clarke, 1999, *Oncogene*, 18(6):1313-24; Millar DS, Paul CL, Molloy PL, Clarke SJ. (2000). *J Biol Chem*; 275(32):24893-9).

In some instances it is desirable to bisulphite modify RNA prior to reverse transcription into cDNA and subsequent PCR. However, RNA is even more sensitive to degradation at the high temperatures and pH required for desulphonation and this results in a further reduction in sensitivity. There is a need for enzymes that are capable of efficiently processing bisulphite modified, treated or converted nucleic acids.

Disclosure of Invention

10 The present inventor has surprisingly found that it is possible to identify or modify enzymes which are more effective in their ability to efficiently process bisulphite modified or treated nucleic acid.

The present invention generally relates to use of new or modified enzymes which are more effective in their ability to copy or amplify bisulphite modified or treated nucleic acid than currently available enzymes such as the eubacterial enzyme *Taq* polymerase, Superscript™ III reverse transcriptase, Klenow exo- polymerase, Bst polymerase or Bca polymerase under substantially the same reaction conditions.

In a first aspect, the invention provides use of an enzyme for copying or amplifying bisulphite modified or treated nucleic acids, wherein the enzyme is more effective in copying or amplifying the nucleic acid compared with native *Taq* polymerase under substantially the same conditions.

Preferably, the enzyme is capable of copying or amplifying nucleic acid having abasic sites, bulky adducts such as sulphonate groups, substantially only having A, T, G and U bases, or substantially only having A, T and G bases.

25 The enzymes can be a thermophilic or mesophilic polymerase, a reverse transcriptase, an endonuclease or modified or chimeric forms thereof.

Preferably, the enzyme is selected from enzymes described in EP 1801213, 5D4 as herein defined, HIV-RT, and modified forms thereof.

Enzymes suitable for use in the present invention maybe obtained using the modification methods disclosed in WO 99/02671, WO 00/40712, WO 02/22869, WO 03/044187, WO 05/045015 and EP 1801213 (Medical Research Council).

In a second aspect, the present invention provides a method for copying or amplifying bisulphite treated nucleic acid comprising:

bisulphite treating a nucleic acid; and

5 copying or amplifying the bisulphite treated nucleic acid using an enzyme which is more effective in copying or amplifying bisulphite treated nucleic acid compared with native Taq polymerase under substantially the same conditions..

Preferably the bisulphite treatment uses sodium bisulphite or sodium metabisulphite.

Preferably the bisulphite treatment is substantially without a desulphonation step.

10 Preferably the method further comprising denaturing the nucleic acid prior to bisulphite treatment.

Preferably the denaturing step is carried out by providing an alkali environment or heating the nucleic acid.

15 Preferably any methylated nucleotides in the sample remain unchanged while unmethylated nucleotides are converted to uracil by the bisulphite treatment.

Preferably the method further comprises desalting or isolating the treated nucleic acid sample.

20 Preferably the enzyme is capable of copying or amplifying nucleic acid having abasic sites, bulky adducts including sulphonate groups, having substantially only A, T, G and U bases, or substantially only having A, T and G bases.

Preferably the enzyme is selected from the group consisting of thermophilic polymerase, mesophilic polymerase, reverse transcriptase, endonuclease, and modified and chimeric forms thereof.

25 In a preferred form, the enzyme is selected from chimeric enzymes, 5D4 as herein defined, HIV-RT, and modified forms thereof.

In a preferred form, the enzyme is 5D4 or modified forms thereof.

In a preferred form, the enzyme is HIV-RT or modified forms thereof.

30 Preferably the method further comprises processing or analysing the treated nucleic acid to determine nucleotide sequence, methylation status, identify of source of nucleic acid, or detect a microorganism.

Preferably the amplification of the bisulphite treated nucleic acid is carried out by polymerase chain reaction (PCR), Reverse-Transcriptase PCR, qPCR, isothermal amplification, or signal amplification.

Preferably the treated nucleic acid comprises bisulphite modified DNA, bisulphite modified RNA, or a combination of bisulphite modified DNA and bisulphite modified RNA.

In a third aspect, the present invention provides a method for copying or amplifying bisulphite treated nucleic acid comprising:

denaturing a nucleic acid sample;

treating the nucleic acid sample with a bisulphite reagent and incubating the sample to form a treated nucleic acid sample where any methylated nucleotides in the sample remain unchanged while unmethylated nucleotides are converted to another form;

desalting or isolating the treated nucleic acid sample; and

using an enzyme which are more effective in their ability to efficiently process bisulphite modified or treated nucleic acid to amplify, reverse transcribe, digest, or otherwise enzymatically process the bisulphite modified nucleic acid.

The method may further comprise:

processing or analysing the treated nucleic acid to determine nucleotide sequence or methylation status.

Amplification of bisulphite treated nucleic acid can be carried out by any suitable means such as polymerase chain reaction (PCR), Reverse-Transcriptase PCR, qPCR, isothermal amplification, or signal amplification.

Preferably, the denaturing step is carried out by providing an alkali environment or heating the nucleic acid sample.

Preferably, the treated nucleic acid comprises bisulphite modified DNA, bisulphite modified RNA, or a combination of bisulphite modified DNA and bisulphite modified RNA.

The sample can be obtained from tissue, organ, cell, microorganism, biological sample, or environmental sample. Preferably, the tissue or organ is selected from brain, colon, urogenital, lung, renal, hematopoietic, breast, thymus, testis, ovary, uterus, or mixtures thereof. Preferably, the microorganism is selected from bacteria, virus, fungi, protozoan, viroid, or mixtures thereof. Preferably, the biological sample is selected from

the group consisting of blood, urine, faeces, semen, cerebrospinal fluid, lavage or mixtures thereof.

Unlike prior art methods, the method uses enzymes, including, but not limited to, thermophilic and mesophilic polymerases, reverse transcriptases and endonucleases that will specifically have an enhanced ability to process bisulphite modified nucleic acids compared to currently available enzymes. Primarily, the enzymes will have enhanced function in their ability to bypass abasic sites, bulky adducts on the DNA such as sulphonate groups and to work more effectively on a 3-base genome compared to standard enzymes such as *Taq* polymerase, Superscript III reverse transcriptase, Klenow exo- polymerase, Bst polymerase or Bca polymerase, but other modifications may also be generated. The generation of new enzymes which are able to process nucleic acids with sulphonate groups attached will abrogate the need to de-sulphonate the nucleic acid following bisulphite modification and will hence eliminate the damage and loss of nucleic acid seen in current methods, as well as saving time. Furthermore, the use of polymerases that are more capable of bypassing abasic sites will enable more efficient PCR reactions and the generation of longer PCR products than is currently possible using the standard enzymes that are currently available.

Preferably, the bisulphite is sodium bisulphite, a reagent which, in the presence of water, modifies cytosine into uracil.

Sodium bisulphite (NaHSO_3) reacts readily with the 5,6-double bond of cytosine to form a sulfonated cytosine reaction intermediate which is susceptible to deamination, and in the presence of water gives rise to uracil sulphonate. If necessary, the sulphonate group can be removed under mild alkaline conditions, resulting in the formation of uracil. Thus, potentially all unprotected cytosines will be converted to uracils. Any methylated cytosines, however, cannot be converted by bisulphite due to protection by methylation.

Bisulphite treatment of DNA results in modified nucleic acid having a reduced total number of cytosines compared with the corresponding untreated nucleic acid where any unmethylated cytosines (C) will be converted to uracil (U). Thus, a bisulphite treated nucleic acid having no methylation will substantially contains bases adenine (A), guanine (G), thymine (T) and uracil (U) but still have substantially the same total number of bases as the corresponding untreated nucleic acid. If this treated nucleic acid undergoes amplification, for example by PCR, a simplified form of the nucleic acid is formed with substantially only the bases adenine (A), guanine (G) and thymine (T).

For double stranded DNA which contains cytosines, the bisulphite treatment results in two nucleic acids (one for each complementary strand), each containing the

bases adenine, guanine, thymine and uracil. The two nucleic acids are produced from the two single strands of the double stranded DNA. The two nucleic acids can have no cytosines but still have the same total number of bases and sequence length as the original untreated DNA molecule. Importantly, the two treated nucleic acids are not complimentary to each other and form a top and a bottom strand template for amplification. One or more of the strands can be used as the target for amplification or further processing. During amplification of the treated nucleic acids, uracils in the top (or bottom strand) are replaced by thymines in the corresponding amplified product of the nucleic acid. As amplification continues, the top (and / or bottom strand if amplified) will be diluted out as each new complimentary strand will have only bases adenine, guanine, thymine.

In a fourth aspect, the present invention provides use of an enzyme selected from the group consisting of 5D4 or modified forms thereof, or HIV-RT reverse transcriptase or modified forms thereof in a copying or amplifying bisulphite treated nucleic acid.

Preferably, the bisulphite treatment is substantially without a desulphonation step.

Various embodiments of the present invention relate to use of enzyme Pol A DNA polymerase variant 5D4 for copying or amplifying bisulphite modified or treated nucleic acids.

Various embodiments of the present invention relate to A method for copying or amplifying bisulphite treated nucleic acid comprising bisulphite treating a nucleic acid; and copying or amplifying the bisulphite treated nucleic acid using Pol A DNA polymerase variant 5D4.

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is solely for the purpose of providing a context for the present invention. It is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present invention as it existed in Australia prior to development of the present invention.

In order that the present invention may be more clearly understood, preferred embodiments will be described with reference to the following drawings and examples.

7a

Brief Description of the Drawings

Figure 1 is a diagram showing expected results of PCR amplification of (Fig A) desulphonated and (Fig B) non-desulphonated bisulphite treated DNA with *Taq* polymerase (lanes 1, 4 and 7), *Dpo4* polymerase (lanes 2, 5 and 8) and a hybrid polymerase (lanes 3, 6 and 9). 10 pg of bisulphite treated DNA was amplified using

primers that yield an amplicon of 200 bp (lanes 1-3), 500 bp (lanes 4-6) or 2000 bp (lanes 7-9).

Figure 2 is a diagram showing expected results of reverse transcription efficiency. ³²P-labelled cDNA was synthesized in 20 µl reactions at the temperatures indicated for 60 minutes, by Superscript III (Fig 2A and Fig 2C) and a chimeric reverse transcriptase (B and D), from a mixture of 0.25 µg each of 1.3 kb, 2.5 kb, 4 kb, 6.5 kb and 10 kb total RNA with 200 units of each RT and 0.5 µg random hexamers in Superscript III buffer. The RNA was bisulphite modified and either desulphonated (Fig 2A and Fig 2B) or non-desulphonated (C and D). The cDNA products were electrophoresed on a 1.4% alkaline gel and exposed to X-ray film.

Figure 3 shows results of amplification of bisulphite treated RNA using HIV-RT reverse transcriptase. Lane 1. 0 minutes desulphonation @ 76°C; Lane 2. 1 minutes desulphonation @ 76°C; Lane 3. 2 minutes desulphonation @ 76°C; Lane 4. 5 minutes desulphonation @ 76°C; Lane 5. 10 minutes desulphonation @ 76°C; Lane 6. 15 minutes desulphonation @ 76°C; Lane 7. 15 minutes desulphonation @ 76°C (control reverse transcriptase); Lane 8. RT negative control; Lane 9. PCR negative control.

Figure 4 shows comparison of PCR amplification efficiency between Taq and 5D4 enzymes.

Figure 5 shows the amino acid sequence of 5D4.

20 Mode(s) for Carrying Out the Invention

Enzymes

The recently developed technique of compartmentalised self-replication (CSR) or *in vitro* compartmentalisation (IVC) is one technique which can be manipulated to generate new enzymes with specific characteristics (Tawfik, D.S and Griffiths, A.D. (1998) Nature Biotech. 16, 652-656; Ghadessy, F.J., Ong, J.L. and Holliger, P. (2001) PNAS. 98(8), 4552-4557; d'Abbadie, M., Hofreiter, M., Vaisman, A., Loakes, D., Gasparutto, D., Cadet, J., Woodgate, R., Paabo, S. and Holliger, P. (2007) Nature Biotech. 25(8), 939-943). In addition, new enzymes such as the DinB family of polymerases (Boudsocq, F., Iwai, S., Hanaoka, F. and Woodgate, R. (2001) Nucl. Acids. Res. 29, (22), 4607-4616) which are found in all three kingdoms of life, have been reported that have new properties, which to date have only been exploited for the amplification of ancient DNA, and may be useful for copying or amplifying bisulphite treated nucleic acids.

Examples of enzymes suitable for use in the present invention may be obtained using the modification methods disclosed in WO 99/02671, WO 00/40712, WO 02/22869, WO 03/044187, WO 05/045015 and EP 1801213 (Medical Research Council).

5 A number of modified enzymes are disclosed in EP 1801213 which are potential candidates for use in the present invention or be further modified to develop or enhance activity on bisulphate treated nucleic acid. Examples include enzymes designated 2F3, 1A10, 1A9, 2F12, 1C2, 2G6, 1A8, 2F11, 2H4, 2H9, 1B12, 2H2, 1C8, 2H10X, 3A10, 3B5, 3B6, 3B8, 3B10, 3C12, 3D1, 4D1 and 5D4. The enzyme 5D4 has been found by the
10 present inventor to be particularly suitable for the present invention.

Various mutagenesis and/or recombination techniques can be used to generate mutant/hybrid enzymes. Error-prone PCR mutagenesis (Zaccolo, M. and Gherardi, E. (1999) J. Mol. Biol. 285, 775-783), staggered extension process (StEP; Zhao, H., Giver, L., Shao, Z. Affholter, J.A. and Arnold, F.H. (1998) Nature biotech. 16, 258-261) and
15 molecular breeding (d'Abbadie, M., Hofreiter, M., Vaisman, A., Loakes, D., Gasparutto, D., Cadet, J., Woodgate, R., Paabo, S. and Holliger, P. (2007) Nature Biotech. 25(8), 939-943) are examples of suitable techniques, but other techniques may also be used.

Error-prone PCR mutagenesis is a random process of mutating sequences during a PCR reaction in which the triphosphate derivatives of a pyrimidine (6-(2-deoxy-β-D-
20 ribofuranosyl)-3,4-dihydro-8H-pyrimido-[4,5-c][1,2]oxazin-7-one; dP) and a purine (8-oxo-2'-deoxyguanosine; 8-oxodG) nucleoside analogue are employed in DNA synthesis reactions *in vitro* and act as mutagenic agents. The primers used are typically biotinylated (or carry a different label) to enable the mutant PCR products to be purified from the wild type DNA, so that the mutated sequences can be cloned and screened for
25 desirable characteristics and then subjected to further rounds of mutagenesis and screening. The procedure yields higher frequencies of amino acid residue substitutions, at least one order of magnitude greater, than those previously achieved.

The staggered extension process (StEP) combines *in vitro* mutagenesis and recombination of template sequences in a PCR format. The template sequence(s) are
30 primed with common flanking primers and undergo PCR cycling with extremely short annealing and extension steps. The short extension times result in incomplete primer extension during each round. In subsequent cycles, the growing fragments anneal to different templates based on sequence complementarity and extend further, effectively leading to recombination between different genes. The process is continued until full-
35 length sequences are made and fine tuning the extension times gives some control over the length of the gene segments that are swapped.

Molecular breeding has been demonstrated to be superior for the selection of damage-tolerant polymerases than random mutagenesis procedures alone. This procedure utilizes StEP to recombine segments of various polymerases from different organisms. This yields a variety of chimeric polymerases which can then be tested for the desired attributes. This protocol would also be suitable for the generation of other chimeric enzymes such as reverse transcriptases, endonucleases etc.

The enzymes of interest for use with bisulphite modified nucleic acids include, but are not limited to, thermophilic and mesophilic DNA polymerases (for example, *Taq*, *Pfu*, *Tth*, *Tfl*, *Pfx*, *Pfx50*[™], *Tko*, *Bst*, *Bca*, Vent_R[®], Deep Vent[™], Phusion[™], ABV, UITima, DyNAzyme EXT[™], Terminator, polk, pol IV, Dbh, Dpo4 and Dpo4-like enzymes, DNA I, Klenow fragment of DNA I polymerase, Klenow exo-, Phi 29, T4 and T7 DNA polymerases), reverse transcriptases (for example, AMV RT, M-MuLV RT, ThermoX RT[™], ThermoScript RT[™], Superscript III), and endonucleases (for example, Endonuclease III, IV, V, VIII, T7 Endonuclease I, hOGG1, UDG, Fpg, USER) and mutants or chimeras thereof.

New enzymes suitable for the present invention that are superior in their ability to process bisulphite modified DNA or RNA can be made using one or more of the techniques set out above, or a combination of these techniques or other suitable techniques of mutagenesis and/or recombination or by the isolation of new types of polymerases from diverse bacterial species such as the DinB family of polymerases (Boudsocq, F., Iwai, S., Hanaoka, F. and Woodgate, R. (2001) Nucl. Acids. Res. 29, (22), 4607-4616) that have been shown to have unique properties that may be useful when using bisulphite treated nucleic acid samples. In particular, but not exclusively, thermophilic and mesophilic polymerases, reverse transcriptases and endonucleases are targeted. Two or more enzymes from the same class or different classes of enzymes will undergo recombination to yield large numbers of mutant/hybrid enzymes. Hybrid enzymes will be screened and those exhibiting high processivity and high tolerance to DNA damage will be selected and examined further. The assay of choice will depend on the enzyme type being examined. For example, polymerase activity can be assayed for by PCR and comparing the ability of the "new" enzyme to amplify wild-type and a bisulphite modified template with that of the parent polymerases. In this case, clones that exhibit a superior ability to amplify bisulphite modified DNA will be selected for further rounds of recombination and reselection until a polymerase or polymerases exhibiting the ability to efficiently amplify bisulphite converted DNA to a level that is better than the parental polymerases whilst displaying satisfactory levels of processivity and

thermostability is generated. The bisulphite modified nucleic acid may or may not be desulphonated.

An example where chimeras of two different enzymes may be beneficial for the amplification of bisulphite modified DNA is the recombination of the thermostable
5 polymerase *Taq*, which has high processivity but low tolerance to DNA damage, and
Dpo4-like enzymes which have low processivity but are capable of replicating past
several types of DNA lesions, including abasic sites and bulky adducts (McDonald, J.P.,
Hall, A., Gasparutto, D., Cadet, J., Ballantyne, J. and Woodgate, R. (2006) 34(4), 1102-
1111). Any combination of polymerases, reverse transcriptases, endonucleases, DNA
10 modifying enzymes, and other enzymes, may be recombined together using these
techniques and their use in processing bisulphite converted nucleic acids is covered
herein by this invention.

Applications

15 The present invention provides methods for the utilisation of enzymes that have
enhanced abilities to process bisulphite modified nucleic acids. The methods are
advantageous in that they can identify a new array of enzymes that will be capable, for
example, but not limited to, of improved PCR amplification efficiency and increased
amplicon length of bisulphite modified DNA, and of efficient reverse transcription and/or
20 one-step reverse-transcription PCR of bisulphite modified RNA. The methods of the
present invention provide a simplified procedure by which bisulphite modified nucleic
acids can be manipulated using common molecular biological techniques, without the
need to de-sulphonate the samples, which will greatly improve yield and maintain higher
molecular weight DNA relative to previously known methods. These enzymes, coupled
25 with bisulphite modification techniques, will enable the successful analysis of the
methylation state of the genomes in small numbers of cells, of archival nucleic acids and
of large stretches of nucleic acids or whole genes, all of which have been hitherto
impossible.

The terms "more efficient activity" or "improved activity" as used herein mean that
30 the isolated enzyme according to the present invention can process bisulphite modified
nucleic acid more effectively with less or no substantial errors as compared with standard
enzymes presently in use, such as *Taq* polymerase, Superscript III reverse transcriptase
Klenow exo- polymerase, Bst polymerase or Bca polymerase, under substantially the
same reaction conditions.

The bisulphite modified sample may include DNA or RNA or a combination of both DNA and RNA.

The sample can be prepared from tissue, cells or can be any biological sample such as blood, urine, faeces, semen, cerebrospinal fluid, lavage, cells or tissue from sources such as brain, colon, urogenital, lung, renal, hematopoietic, breast, thymus, testis, ovary, uterus, tissues from embryonic or extra-embryonic lineages, environmental samples, plants, microorganisms including bacteria, intracellular parasites, virus, fungi, protozoan, viroid and the like. The best described mammalian cell types suitable for treatment by the present invention are summarized in B. Alberts et al., 1989, The Molecular Biology of the Cell, 2nd Edition, Garland Publishing Inc New York and London, pp 995-997.

The analysis of 5-methyl cytosine residues in DNA from samples of human, animal, plant, bacterial, fungal and viral origin is meant to cover all life cycle stages, in all cells, tissues and organs from fertilization until 48 hours post mortem, as well as samples that may be derived from histological sources, such as microscope slides, samples embedded in blocks, or samples extracted from synthetic or natural surfaces or from liquids.

The analyses include the naturally occurring variation between cells, tissues and organs of healthy individuals, (health as defined by the WHO), as well as cells, tissues and organs from diseased individuals. Diseased in this sense includes all human diseases, afflictions, ailments and deviant conditions described or referred to in Harrison's Principles of Internal Medicine, 12th Edition, edited by Jean D Wilson et al., McGraw Hill Inc, and subsequent later editions; as well as all diseases, afflictions ailments and deviant conditions described in OMIM (Online Mendelian Inheritance in Man), but with emphases on the leading causes of death, namely, malignant neoplasms, (cancer), ischaemic heart disease, cerebrovascular disease, chronic obstructive pulmonary disease, pneumonia and influenza, diseases of arteries, (including atherosclerosis and aortic aneurysm), diabetes mellitus, and central nervous system diseases, together with socially debilitating conditions such as anxiety, stress related neuropsychiatric conditions and obesity, and all conditions arising from abnormal chromosome number or chromosome rearrangements, (aneuploidy involving autosomes as well as sex chromosomes, duplications, deficiencies, translocations and insertions), as well as similar abnormalities of the mitochondrial genomes.

The normal or diseased individuals may be from (i) populations of diverse ethnicity and evolutionary lineages; (ii) strains and geographical isolates; (iii) sub

species; (iv) twins or higher order multiplets of the same or different sex; (v) individuals arising from normal methods of conjugation, artificial insemination, cloning by embryonic stem cell methods, or by nuclear transfer, (from somatic or germ line nuclei), or from the input or modification of mitochondrial or other cellular organelles; (vi) individuals deriving
5 from transgenic knock-out, knock-in or knock-down methods, (either *in vivo*, *ex vivo*, or by any method in which gene activity is transiently or permanently altered, e.g., by RNAi, ribozyme, transposon activation, drug or small molecule methodologies, Peptide Nucleic Acid (PNA), Intercalating Nucleic Acid (INA), Altritol Nucleic Acid (ANA), Hexitol Nucleic Acid (HNA), Locked Nucleic Acid (LNA), Cyclohexanyl Nucleic Acid (CNA), and the like,
10 or nucleic acid based conjugates, including but not restricted to Trojan peptides, or individuals at any stages of pregnancy, normal or ectopic.

The analyses also include 5-methyl cytosine and cytosine residues in DNA or RNA from prokaryotic or eukaryotic organisms and viruses (or combinations thereof), that are associated with human diseases in extracellular or intracellular modes, for the
15 purposes of diagnostics and disease state monitoring or determining, and therapeutically altering, in both normally varying and diseased systems, the changed parameters and underlying mechanisms of:

- (i) genetic diseases;
- (ii) non-genetic or epigenetic diseases caused by environmentally induced factors,
20 be they of biological or non-biological origin, (environmental in this sense being taken to also include the environment within the organism itself, during all stages of pregnancy, or under conditions of fertility and infertility treatments);
- (iii) predisposition to genetic or non genetic diseases, including effects brought about by the "prior" class of factors, by exposure to pressure changes and weightlessness, or
25 by radiation effects;
- (iv) 5-methyl cytosine changes in the processes of aging in all cell types, tissues, organ systems and biological networks, including age related depression, pain, neuropsychiatric and neurodegenerative conditions and pre- and post-menopausal conditions, (including reduced fertility; in both sexes);
- 30 (v) 5-methyl cytosine changes in cancer, (including changes in cells with abnormal karyotypes arising from DNA amplification, deletion, rearrangement, translocation and insertion events), and their variations or alterations in different cell cycle phenomena (including cell cycle effects on diurnal rhythms, photoperiod, sleep, memory, and "jet lag";

- (vi) 5-methyl cytosine changes in metabolic networks defined in the broadest sense, from the zygote through embryogenesis, foetal development, birth, adolescence, adulthood and old age (including metabolic effects brought about by hypoxia, anoxia, radiation of any type, (be it ionizing or non ionizing, or arising from chemotherapeutic treatments, high altitude exposure radiation from nearby natural sources, such as rocks or from "fallout" from military or government sponsored activities), stress, or by imbalances between the mitochondrial, nuclear or organellar genomes;
- (vii) 5-methyl cytosine alterations due to responses at the molecular, cellular, tissue, organ and whole organism levels to proteins, polypeptides, peptides, and DNA, RNA, PNA, INA, ANA, HNA, LNA, CNA, and the like, or peptide aptamers (including any with post translational additions, post translational cleavage products, post translational modifications (such as inteins and exeins, ubiquination and degradation products); proteins, polypeptides and peptides containing rare natural amino acids, as well as single rare amino acids such as D-serine involved in learning, brain growth and cell death; drugs, biopharmaceuticals, chemical entities (where the definitions of Chemical Entities and Biopharmaceuticals is that of G. Ashton, 2001, Nature Biotechnology 19, 307-3111)), metabolites, new salts, prodrugs, esters of existing compounds, vaccines, antigens, polyketides, non-ribosomal peptides, vitamins, and molecules from any natural source (such as the plant derived cyclopamine);
- (viii) 5-methyl cytosine or cytosine alterations due to responses at the molecular, cellular, tissue, organ and whole organism levels to RNA and DNA viruses be they single or double stranded, from external sources, or internally activated such as in endogenous transposons or retrotransposons, (SINES and LINES);
- (ix) 5-methyl cytosine alterations due to responses at the molecular, cellular, tissue, organ and whole organism levels to reverse transcribed copies of RNA transcripts be they of genic or non genic origins, (or intron containing or not);
- (x) 5-methyl cytosine alterations due to responses at the molecular, cellular, tissue, organ and whole organism levels to: (a) DNA, RNA, PNA, INA, ANA, HNA, LNA, CNA, and the like (or DNA, RNA, PNA, INA, ANA, HNA, LNA, CNA, aptamers of any in all combinations); including DNA, RNA, PNA, INA, ANA, HNA, LNA, CNA, and the like molecules circulating in all fluids including blood and cerebrospinal fluid as well as maternal fluids before, during and after pregnancy (b) combinations of conjugated biomolecules that are chimeras of peptides and nucleic acids; or chimeras of natural molecules such as cholesterol moieties, hormones and nucleic acids; and

(xi) 5-methyl cytosine alterations due to responses of stem cells, (either *in vivo*, *ex vivo* or in association with novel environments or natural and synthetic substrates (or combinations thereof), from human and animal origin to any of the perturbations described in (i) to (x) above.

5

Modified nucleic acid

Any suitable method for obtaining nucleic acid material can be used. Examples include, but are not limited to, commercially available DNA, RNA kits or reagents, workstation, standard cell lysis buffers containing protease reagents and organic
10 extraction procedures, which are well known to those of skill in the art.

The method can be carried out in a reaction vessel. The reaction vessel can be any suitable vessel such as tube, plate, capillary tube, well, centrifuge tube, microfuge tube, slide, coverslip, bead, membrane or any suitable surface. The method is generally carried out in one reaction vessel in order to reduce the likelihood of degradation or loss
15 of the nucleic acid sample.

Generally, the alkali environment is provided to the sample by adding an alkali such as NaOH. If the nucleic acid material is RNA then heat is used instead of alkali to produce single stranded material without secondary structure. The alkali environment is provided to denature double stranded nucleic acid molecules into a state where the
20 molecules are readily reactive with the bisulphite reagent. It will be appreciated, however, that any other denaturation method such as heat treatment or other suitable alkali or denaturing agent could be added or used such as KOH and any other alkali.

Bisulphite treatment

Generally, the bisulphite reagent is sodium metabisulphite. The bisulphite reagent is used to cause sulphonation of cytosine bases to cytosine sulphonate followed by hydrolytic deamination of the cytosine sulphonate to uracil sulphonate. It will be appreciated, however, that any other suitable bisulphite reagent could be used such as sulphite or acetate ions (see Shapiro, R., DiFate, V., and Welcher, M, (1974) J. Am.
30 Chem. Soc. 96: 906-912).

The incubation with the sulphonating reagent can be carried out at pH below 7 and at a temperature which favors the formation of the uracil sulphonate group. A pH below 7 is optimal for carrying out the sulphonation reaction, which converts the cytosine

bases to cytosine sulphonate and subsequently to uracil sulphonate. However, the methods can be performed with the sulphonation reaction above pH 7, if desired.

The sulphonation reaction can be carried out in the presence of an additive capable of enhancing the bisulphite reaction. Examples of suitable additives include, but not limited to, quinol, urea, DTT and methoxyamine. Of these reagents, quinol is a reducing agent. Urea and methoxyamine are agents added to improve the efficiency of the bisulphite reaction. In addition, DTT can be used in the reaction to prevent the degradation of RNA by endogenous RNases. It will be appreciated that other additives or agents can be provided to assist in the bisulphite reaction. The sulphonation reaction results in methylated cytosines in the nucleic acid sample remaining unchanged while unmethylated cytosines are converted to uracils.

Reaction conditions found to work well are as follows. The DNA, or other nucleic acids, to be treated is made up to a volume of 20 μ l and denatured by incubating with 2.2 μ l freshly prepared 3 M sodium hydroxide (BDH AnalaR #10252.4X) solution for 15 minutes at 37°C. The concentration of sodium hydroxide and incubation times can be adjusted as necessary to ensure complete denaturation of the template nucleic acid. 220 μ l of a freshly prepared solution of 3 M sodium metabisulphite (BDH AnalaR #10356.4D) pH 5.0 (the pH is adjusted by the addition of 10M sodium hydroxide (BDH AnalaR #10252.4X) along with 12 μ l of a 100 mM quinol solution (BDH AnalaR #103122E) is added. The concentration of quinol added can be anything in the range of about 10 to 500 mM as determined experimentally. The solution is then vortexed and overlaid with 208 μ l of mineral oil (Sigma molecular biology grade M-5904). The sample is then incubated at a suitable temperature and for sufficient time, to allow time for full bisulphite conversion, for example at 80°C for 45 minutes. It is understood by those skilled in the art that the volumes, concentrations and incubation time and temperature described above can be varied so long as the reaction conditions are suitable for sulphonation of the nucleic acids.

The converted nucleic acids are then desalted either by use of a desalting column, such as Zymo-Spin I columns according to the manufacturer's instructions, or by precipitation. For precipitation, samples are diluted so that the salts inhibitory to subsequent reactions are not co-precipitated with the sulphonated nucleic acids. The salt concentration is diluted to less than about 1 M. Generally, the dilution step is carried out using water or buffer to reduce the salt concentration to below about 0.5M. For example, the salt concentration is generally diluted to less than about 1 mM to about 1 M, in particular, less than about 0.5 M, less than about 0.4 M, less than about 0.3 M, less

than about 0.2 M, less than about 0.1 M, less than about 50 mM, less than about 20 mM, less than about 10 mM, or even less than about 1 mM, if desired. One skilled in the art can readily determine a suitable dilution that diminishes salt precipitation with the nucleic acids so that subsequent steps can be performed with minimal further clean up or
5 manipulation of the nucleic acid sample. The dilution is generally carried out in water but can be carried out in any suitable buffer, for example Tris/EDTA or other biological buffers so long as the buffer does not precipitate significantly or cause the salt to precipitate significantly with the nucleic acids so as to inhibit subsequent reactions. Generally, precipitation is carried out using a precipitating agent such as an alcohol. An
10 exemplary alcohol for precipitation of nucleic acids can be selected from isopropanol, ethanol or any other suitable alcohol.

Using enzymes that are capable of processing DNA containing bulky adducts such as sulphonate groups should eliminate the need to de-sulphonate the bisulphite converted nucleic acids. As the desulphonation step is the point at which most nucleic
15 acid damage and loss is seen during the bisulphite conversion method, being able to omit or modify this step, will greatly improve yield and maintain higher molecular weight DNA relative to previously known methods. However, there may be instances where desulphonation of the bisulphite modified nucleic acid is still required, and this can be carried out using standard methods or modified methods as set out below.

20 The desulphonation step can be carried out by adjusting the pH of the precipitated treated nucleic acid up to about 12.5. Exposure to alkaline environments tends to promote strand breaks in apurinic sites in the DNA induced by the previous exposure to an acidic pH. Therefore, the alkaline pH treatment is minimized if strand breaks are to be avoided. This step can be carried out efficiently at around pH 10.5-11.5
25 with a suitable buffer or alkali reagent. Examples of suitable buffers or alkali reagents include buffers having a pH 7.0 -12.5. It will be appreciated by persons skilled in the art that suitable buffers or alkali reagents can be selected from the vast range of known buffers and alkali reagents available.

Temperature ranges for the desulphonation step are room temperature to about
30 96°C and times can vary from 2 minutes to 96 hours or longer depending on the conditions used. One skilled in the art can readily determine a suitable time and temperature for carrying out the desulphonation reaction. Temperatures below room temperature can also be used so long as the incubation time is increased to allow sufficient desulphonation. Thus, the incubation step can be carried out at about 10°C,
35 about 20°C, about 22°C, about 25°C, about 30°C, about 35°C, about 37°C, about 40°C,

about 45°C, about 50°C, about 55°C, about 60°C, about 65°C, about 70°C, about 75°C, about 80°C, about 85°C, about 90°C, about 95°C, and about 96°C. A particularly useful temperature for carrying out the desulphonation reaction is in the temperature range of about 75 to 95°C.

5 The present invention provides methods for characterization of methylated nucleic acids. The methods allow efficient sulphonation (and desulphonation steps if desired) to be carried out on the nucleic acid sample. However, it is understood that neither of the sulphonation or desulphonation steps need be carried out to completion, only sufficiently to subsequently characterize methylation of the nucleic acid, as
10 disclosed herein. One skilled in the art can readily determine whether these steps should be carried out to near completion or whether incomplete reactions are sufficient for a desired analysis. For example, when a small number of cells or a small amount of nucleic acid sample are used, it is generally desired that a more complete reaction be performed. When larger quantities of nucleic acid sample are being characterized, a less
15 complete reaction can be carried out while still providing sufficient reaction products for subsequent analysis of the methylation state of the nucleic acid sample.

 The present invention provides methods for the use of enzymes that have enhanced abilities to efficiently process bisulphite modified nucleic acids. These enzymes can be used in the analysis of the methylation state of a genome as a measure
20 of the state of a cell, tissue or organism, as disclosed herein or used to determine the sequence of a given region of nucleic acid. The outcomes of the invention provide several advantages over enzymes currently used for processing methylated nucleic acids. The methods are advantageous in that they will provide a new array of enzymes that will be capable, for example, but not exclusively, of improved PCR amplification
25 efficiency and increased amplicon length of bisulphite modified DNA, and of efficient reverse transcription and/or one-step reverse-transcription PCR of bisulphite modified RNA. The enzymes will have a greater ability to process the abasic sites, sulphonate groups and 3 base genome generated during the bisulphite modification, as well as other nucleic acid damage sites, and also to bypass other bulky adducts, such as, but not
30 limited to, intercalating nucleic acids.

 The methods of the present invention may also be used to detect or diagnose microorganisms in accordance with WO 2006/058393 'Methods for simplifying microbial nucleic acids by chemical modification of cytosine' and WO 2006/066353 'Detection of Human Papilloma virus' (Human Genetic Signatures Pty Ltd, Australia).

The methods of the present invention provide a simplified procedure by which bisulphite modified nucleic acids can be manipulated using common molecular biological techniques, without the need to de-sulphonate the samples, or to de-sulphonate using more gentle methods. As the desulphonation step is the point at which most nucleic acid damage and loss is seen, being able to omit or modify this step, will greatly improve yield and maintain higher molecular weight DNA relative to previously known methods. These enzymes will enable the successful analysis of the methylation state of small numbers of cells, of archival nucleic acids and of large stretches of nucleic acids or whole genes, or be able to measure the presence or absence of small numbers of infectious agents whether DNA or RNA based in clinical samples all of which have been hitherto impossible.

Accordingly, the methods of the present invention provide the additional advantage of allowing even smaller amounts of starting material to be used and efficiently characterized with respect to methylation. The method for determining the methylation state of a sample can be carried out in parallel with a test sample and a control sample so that the methylation state of the sample can be compared and determined relative to a reference sample. For example, the samples can be compared to determine whether there is an increase or decrease of methylation in general or at particular sites. Such a determination can be used to diagnose and/or determine the prognosis of a disease, as discussed herein. The method can further include reporting of the methylation state of a sample, for example, in a diagnostic application.

It is understood that the components for use of the invention can be provided in the form of a kit for efficient processing of bisulphite modified nucleic acids.

Embodiments for utilising engineered enzymes specific for enhanced processing of bisulphite modified nucleic acids are described in non-limiting detail below.

EXAMPLES

Methods and reagents

Chemicals can be obtained as follows: Ethanol from Aldrich (St. Louis MO; 200 proof E702-3); Isopropanol from Sigma (St. Louis MO; 99%+ Sigma I-9516); Mineral oil from Sigma (M-5904); Quinol from BDH (AnalaR #103122E); Sodium acetate solution 3M from Sigma (S-7899); Sodium chloride from Sigma (ACS reagent S9888); and Sodium hydroxide from BDH (AnalaR #10252.4X); Sodium metabisulphite from BDH (AnalaR #10356); Diethyl ether from Sigma (St. Louis MO; 309958); Hexane from Sigma

(St. Louis MO; 650420); Luria broth from Oxoid (Liverpool; CM0996B); Magnesium chloride from Sigma (St. Louis MO; 63069); Mineral oil from Sigma (M-5904); Potassium chloride from Sigma (St. Louis MO; 60142); Span 80 From Fluka (Buchs CH; 85548); Tetracycline hydrochloride from Sigma (St. Louis MO; T8032); Triton X™-100 from
 5 Sigma (St. Louis MO; 93426); Trizma™ hydrochloride from Sigma (St. Louis MO; T5941); Tween™ 80 from Sigma (St. Louis MO; P8074).

Enzymes/Reagents can be obtained as follows: dNTPs from Promega (Madison WI; C1145); Glycogen from Roche (Indianapolis IN; #10 901 393 001); tRNA from Roche (Indianapolis IN; #10 109 495 001); RNase, DNase-free from Roche (Castle Hill
 10 NSW; 11 119 915 001); *Sa*I from New England Biolabs (Beverly MA; #R0138L, 20 units/μl); *Xba*I from New England Biolabs (Beverly MA; #R0145L, 20 units/μl); and DNA markers from Sigma (Direct load PCR low ladder 100-1000 bp, Sigma D-3687 and 100-10 Kb, Sigma D-7058); *Eco*R1 from Roche (Indianapolis IN; #87930626, 10 units/μl); *Hind*III from Biolabs (Beverly MA; #R01045, 10 units/μl); PCR master mix from Promega
 15 (Madison WI; #M7505); and DNA markers from Sigma (Direct load PCR low ladder 100-1000 bp, Sigma D-3687 and 100-10 Kb, Sigma D-7058).

Solutions are as follows: (1) 10 mM Tris/0.1M EDTA, pH 7.0-12.5; (2) 3 M NaOH (6g in 50 ml water; BDH AnalaR #10252.4X); (3) 3M Metabisulphite (7.6 g in 20 ml water with 416 μl 10 N NaOH (BDH AnalaR #10356.4D); (4) 100 mM Quinol (0.55 g in 50 ml
 20 water; BDH AnalaR #103122E); (5) 50X TAE gel electrophoresis buffer (242 g Trizma™ base, 57.1 ml glacial acetic acid, 37.2g EDTA and water to 1l); (6) 5X Agarose gel loading buffer (1 ml 1% Bromophenol blue (Sigma B6131), 1 ml Xylene Cyanol (Sigma X-4126), 3.2 ml Glycerol (Sigma G6279), 8 μl 0.5M EDTA pH 8.0, 200 μl 50X TAE buffer and water to 10 ml); and (7) 1 x *Taq* buffer (50mM KCl, 10mM Tris-HCl, pH 9.0, 0.1%
 25 Triton X™-100, 1.5 mM MgCl₂).

Bisulphite treatment of nucleic acid

An exemplary protocol for the bisulphite treatment of nucleic acids is set out below and can be used to generate template nucleic acids for amplification or copying by
 30 the enzymes according to the present invention. This protocol successfully results in retaining substantially all DNA treated. It will be appreciated that the volumes or amounts of sample or reagents can be varied.

To 2 μg of nucleic acid in a volume of 20 μl, 2.2 μl of 3 M NaOH (6 g in 50 ml water, freshly made) is added. This step denatures the double stranded nucleic acid

molecules into a single stranded form, since the bisulphite reagent preferably reacts with single stranded molecules. The mixture is incubated at 37°C for 15 minutes. Incubation at temperatures above room temperature can be used to improve the efficiency of denaturation.

5 After the incubation, 220 µl 3M Sodium Metabisulphite (3.35 g in 4.68 ml water with 320µl 10 N NaOH; BDH AnalaR #10356.4D; freshly made) and 12 µl of 100 mM Quinol (0.55 g in 50 ml water, BDH AnalaR #103122E; freshly made) are added in succession. Quinol is a reducing agent and helps to reduce oxidation of the reagents. Other reducing agents can also be used, for example, dithiothreitol (DTT),
10 mercaptoethanol, quinone (hydroquinone), or other suitable reducing agents. Likewise, additives which enhance the reaction, such as methoxyamine or urea, may also be incorporated. The sample is overlaid with 200 µl of mineral oil which prevents evaporation and oxidation of the reagents, but is not essential. The sample is then incubated for 45 minutes at 80°C. Other temperatures from 25°C to 90°C may also be
15 used with incubation lengths varying from 5 minutes to 8 hours, or longer.

 After the treatment with sodium metabisulphite, the oil is removed, and 2 µl glycogen (20 mg/ml; Roche # 10 901 393 001) or tRNA (Roche #10 109 495 001) are added if the nucleic acid concentration is low. These additives are optional and can be used to improve the yield of nucleic acid obtained by co-precipitating with the target
20 nucleic acid especially when the nucleic acid is present at low concentrations. Typically, glycogen is used in the precipitation of DNA whereas tRNA is used as a coprecipitant with RNA, although other coprecipitants may also be used.

 Bisulphite modified nucleic acids are then desalted by use of a desalting spin column such as Zymo-spin columns (Zymo # C1003) according to the manufacturer's
25 instructions. Alternatively, the samples can be isopropanol precipitated as follows: 800 µl of water is added to the sample, mixed and then 1 ml isopropanol is added. The water or buffer reduces the concentration of the bisulphite salt in the reaction vessel to a level at which the salt will not precipitate along with the target nucleic acid of interest. The sample is mixed again and left at 4°C for 60 minutes, although other temperatures and
30 lengths of incubation can be used as long as it effectively results in precipitation of the nucleic acid. The sample is centrifuged at 15,000xg for 10-15 minutes at 4°C and the pellet washed with 70% EtOH. This washing treatment removes any residual salts that precipitated with the nucleic acids.

The pellet is allowed to dry and then resuspended in a suitable volume of buffer or water, depending on the downstream application. If desulphonation is desired, resuspension in TE buffer (10 mM Tris, 0.1 mM EDTA) pH 10.5 and incubation at 95°C for 20 minutes has been found to be particularly effective for desulphonation of DNA samples. Buffers at pH 7.0-12.5 can also be used and the sample may be incubated at 37°C to 95°C for 1 min to 96 hr, as needed to facilitate desulphonation of the nucleic acid to a level that is acceptable by the user.

The method described above can be preceded by digestion with one or more restriction enzymes. Two independent restriction enzyme digests are set up of the same sample of DNA as described below. The enzymes selected for digestion are dependent upon the sequence to be amplified. For example, digest 2 µg genomic DNA with EcoRI in a 20 µl volume for 1 hr at 37°C. This step is used to digest the genomic DNA into smaller fragments which are more amenable to bisulphite conversion than genomic DNA. Sonication or physical forces can also be used to shear the DNA into smaller sized fragments. The intensity of sonication and the length of sonication is selected based on the desired size of DNA fragments. A separate digestion reaction is carried out, for example, by digesting 2 µg genomic DNA with *Hind*III as described above. These or other suitable restriction enzymes can be selected for pretreatment digestion. The digested DNA is treated with metabisulfite as described above.

20

Generation and utilization of enzymes

An example of how to generate a thermostable DNA polymerase for use in amplification of bisulphite converted DNA is given to illustrate the technique and is performed according to the method of d'Abbadie *et al* (d'Abbadie, M., Hofreiter, M., Vaisman, A., Loakes, D., Gasparutto, D., Cadet, J., Woodgate, R., Paabo, S. and Holliger, P. (2007) *Nature Biotech.* 25(8), 939-943) with adaptations specific for generating enzymes that are superior in their ability to process bisulphite modified nucleic acid as described below.

DNA polymerase genes *Taq* and *Dpo4* are amplified from *Thermus aquaticus* (Accession no J04639) and *Sulfolobus solfataricus* (Accession no. N002754), respectively, using gene-specific primers with flanking *Xba*I and *Sal*I restriction sites. Purified PCR products are cloned into pre-digested pASK75 and constructs are transformed into *E. coli* and expressed according to the methods of Ghadessy *et al.*, (Ghadessy, F.J *et al.*, *PNAS* (1998), 98, (8), 4552-4557) and Skerra (Skerra, A. (1994)

Gene, 151, 131-135). For purification, *Taq* and *Dpo4* clones are subcloned with an N-terminal hexahistidine tag introduced via primers, expressed as above, and lysed and purified using Ni-NTA spin columns according to the manufacturer's instructions (Qiagen, #31014).

5 Molecular breeding is used to yield chimeras of the two enzymes. Briefly, equal concentrations of the two polymerase genes are cycled 40 times (94°C, 30 sec; 55°C, 1 sec) using primers to homologous flanking regions. The short annealing/extension time results in incomplete primer extension during each round. In subsequent cycles, the growing fragments anneal to different templates based on sequence complementarity and extend further, effectively leading to recombination between different genes. The process is continued until 10 full-length sequences are made and more or less cycles may be needed. Fine tuning the extension times gives some control over the length of the gene segments that are swapped and can be altered from the times given here. The products are then gel purified using the QIAquick gel extraction kit according to the manufacturer's instructions (Qiagen, #28706), reamplified and cloned into pASK75 to generate a hybrid polymerase library.

15 Emulsification and compartmentalised self-replication reactions (CSRs) are then performed as described by d'Abbadie *et al* (d'Abbadie *et al.* (2007) *Nature Biotech.* 25(8), 939-943) using either matched primers, primers containing abasic sites, desulphonated bisulphite-converted primers or bisulphite-converted primers which have not been desulphonated, or combinations thereof. The CSRs are undertaken by cycling at 94°C for 5 20 min, followed by 20 cycles of 94°C for 1 min, 50°C for 1 min and 72°C for 8 min. The initial denaturation for 5 minutes ruptures the bacterial cell harbouring the hybrid polymerase which is then released into the aqueous compartment. This allows the self-replication of the polymerase if the enzyme is active. Non-active enzymes will fail to self-replicate, will not be selected for subsequent rounds and hence will be eliminated from the gene pool.

25 To recover the reaction mixtures, the emulsions are solvent-extracted and the purified selection products are re-amplified and re-cloned and then screened for their ability to amplify various sized products from bisulphite modified DNA generated as indicated above (either with or without desulphonation) by PCR using 10pg template DNA, 0.2 mM dNTPs, 1 µM primers in 1x *Taq* buffer and cycling for 94°C, 5 min, followed by 30 cycles of 94°C, 1 min; 30 55°C, 1 min; 68-72°C, 30 secs-2min, depending on the expected product size. Promising clones from selection rounds one and two are StEP shuffled and backcrossed with parent polymerase genes as described above. Recombination, recloning, rescreening and reselection continues until one or more

enzymes are generated that have superior ability to amplify bisulphite modified DNA compared to the parent enzymes. In particular, superiority is demonstrated by the ability to generate amplicons more efficiently, in larger quantities, of larger size and/or from non-desulphonated bisulphite-treated DNA than the parent polymerases.

5 The various mutagenesis and/or recombination techniques available can be used to generate a variety of mutant/chimeric enzymes specifically enhanced for their ability to amplify or copy bisulphite modified nucleic acids including, but not limited to, thermophilic and mesophilic polymerases, reverse transcriptases, and endonucleases. Furthermore, newly discovered enzymes such as the DinB family of polymerases (eg Dpo4 from
10 *Sulfolobus solfatarius*), may exhibit superior ability to amplify bisulphite modified DNA without the need for further mutagenesis or recombination.

Figure 1 demonstrates the expected results with new polymerases, selected for their ability to process bisulphite converted nucleic acids, in comparison to the parental polymerases. The hybrid/mutant polymerases will exhibit enhanced ability to amplify
15 desulphonated bisulphite modified DNA and be able to amplify larger amplicons compared to parent polymerases and will also be able to amplify from a non-desulphonated template whereas parent polymerases cannot.

Figure 2 illustrates the results with new reverse transcriptases, selected for their ability to process converted nucleic acids, in comparison to the parental reverse
20 transcriptases. The new transcriptases will be able to more efficiently copy desulphonated bisulphite converted RNA templates of various sizes compared to available transcriptases, and will also be able to copy from a non-desulphonated template whereas parental reverse transcriptases cannot.

25 EXPERIMENTAL

HIV-RT enzyme

HCV RNA was isolated OptiQual[®] HCV RNA high positive control (Acrometrix cat# 96-0203) using the QIAmp UltraSens (Qiagen) viral kit according to the manufacturers instructions and resuspended at a final concentration of 5,000 IU/ μ l.

30 Bisulphite conversion of RNA.

3.3g sodium bisulphite (Sigma 59000 500 g; lot number 116K0761) was dissolved in 5 ml Xceed reagent 1. The reagent was heated at 80°C until fully dissolved, allow to cool.

0.11g of Hydroquinone (Merck 8.22333.0250; lot number K36100033 702) was
35 dissolved in 10 ml nuclease-free water.

5 μ l of RNA was mixed with 220 μ l bisulphite reagent and 12 μ l quinol in a PCR tube and incubated at 70°C for 20 minutes in a PCR machine.

800 μ l nuclease-free water was added along with 2 μ l glycoblue (Ambion AM9515; lot number 0705003), the samples mixed well, then 1 ml isopropanol was added and the samples incubated at 4°C for 1 hour.

The RNA was pelleted by centrifuging at 16 000x g for 20 minutes at 4°C

The supernatant discarded and the pellet washed with 1 ml 70% ethanol with moderate vortexing. The sample was recentrifuge at 16 000x g for 7 minutes at 4°C

The supernatant was discarded and the pellet air dried for a few minutes.

10 The pellet was resuspended pellet in 70 μ l desulphonation buffer (Xceed reagent 5) and de-sulphonate at 76°C for 0-15 minutes in a PCR machine.

The RNA was cooled, then 8 μ l of a mastermix added comprising the following per reaction:

10 μ l converted RNA

15 1 μ l 10 mM dNTPS

1 μ l random H primers (300 ng/ μ l)

The sample was heated at 65°C for 5 minutes, then placed on ice for at least 1 minute. Add 7 μ l of a mastermix comprising the following per reaction:

20

HIV-RT

Control RT

2 μ l 10x HIV-RT Buffer

4 μ l 5x FS buffer

1 μ l HIV-RT (1U/ μ l)

1 μ l 0.1M DTT

1 μ l Rnase OUT

1 μ l Superscript III

4 μ l water

1 μ l Rnase OUT

25

The samples were mixed the reverse transcription carried out as follows; 25°C for 2 mins, 27°C for 2 mins, 29°C for 2 mins, 31°C for 2 mins, 33°C for 2 mins, 35°C for 2 mins, 37°C for 30 mins, 45°C for 10 mins, 50°C for 10 mins, 70°C for 5 mins, then soaked at 15°C.

30

Take 5 μ l of the cDNA for analysis by PCR, comprising the following per reaction:

36.5 μ l Promega mastermix

1.0 µl F1 primer (100 ng/µl)

1.0 µl R0 primer (100 ng/µl)

6.5 µl water

Note: primers are labelled RNA primers in orange box in "Herbert"

5 Cycling conditions are as follows:

95°C, 3 mins

95°C, 10 secs

52°C, 1 min 40x

68°C, 1 min

10 68°C, 7 mins

As can be seen from Figure 3 the HIV-RT was able to copy bisulphite treated RNA without the additional desulphonation step.

15 **5D4 enzyme**

Bisulphite conversion reactions

1 µg of template material was made to a final volume of 20µl. The template was denatured by the addition of 2 µl of 3M NaOH for 15 minutes at 37°C. The DNA was then bisulphite treated according to the MethylEasy Xceed bisulphite conversion kit (Human Genetic Signatures, Sydney, Australia). For desulphonation, the converted DNA was heated at 80°C for 20 minutes then stored at -20°C until required. Non-desulphonated bisulphite treated DNA was resuspended in water (instead of reagent #5 from the MethylEasy kit) and used immediately.

Primer extension reactions

25 Reactions were prepared in 50 µl volumes containing 2 µl of 100 µM template oligonucleotide, 2 µl of 2.5 µM Cy3 labelled oligonucleotide, 5 µl of polymerase buffer, 4 µl dNTPs (625 µM each gives a final concentration of 50 µM) and distilled water to 49 µl.

BStemp3 5'-TAGCACCCCTAGCCAGCTAGCTGGGTATAGTGAGTGGTATTA
30 (SEQ ID NO: 1)

BStempC8 5'-CCCCCCCCTGGGTATAGTGAGTGGTATTA (SEQ ID NO: 2)

Primer 5'-7C6C5CTATACCCA (SEQ ID NO: 3)

7=Cy3

6=LNA T

5 5=LNA A

(CTCACTATACCC (SEQ ID NO: 4))

Samples were heated to 85°C for 2 minutes then cooled slowly to 40°C for 1 minute. The samples were then heated to 55°C for a further 1 minute then 1 µl of a 1/10 dilution of each enzyme added and the samples incubated for the appropriate time. The reaction was stopped by the addition of 50 µl of 8M Urea, 50 mM EDTA.

The samples were then heated to 95°C for 2 minutes then snap chilled on ice. 10 µl of the sample was run on a 20% denaturing acrylamide gel at 30 W for 3 hours. Fluorescence was then recorded on a Typhoon imager.

PCR amplification

Amplification reactions were prepared in a final reaction volume of 25 µl containing 2.5 µl of X10 SuperTaq buffer, 1 µl of 10 mM dNTP's, 1 µl of 100n g/µl each primer, 0.5 µl SuperTaq (Cambio Ltd, Cambridge, United Kingdom) and 18 µl of distilled water. 1 µl of template was added.

20 Primers

F1 5'-GAGGTTTGGAAGTTTTATTATT (SEQ ID NO: 5)

R1 5'-TAACTTATCATCAAAATAAAC (SEQ ID NO: 6)

R3 5'-TATACTACCTCAAAAATATAAATA (SEQ ID NO: 7)

R5 5'- AAAAATCCTTACAAAACCTTATAAC (SEQ ID NO: 8)

25

Template plasmid Tgo was kindly provided by Vitor Pinheiro (MRC laboratory for Amplification was carried out on a MJ Research PTC-200 (PTC200) DNA Engine Thermal Cycler PCR using the following PCR program

94°C 1 min 30 seconds

30 94°C 20 seconds

94^oC 20 seconds

45^oC 30 seconds Repeated X24

68^oC 1 minute

PCR products were resolved on a 1.5% agarose gel containing ethidium bromide.

5

Table 1 shows the results of a primer extension reaction using 6 different polymerase enzymes using a partially desulphonated BStemp3 template. As can be seen from Table 1 the only enzymes that partially copy full length sulphonated (S) bisulphite treated template are the Taq polymerase, 5D4 and 3A10 enzymes. Enzymes 10 14, 5D4, 3A10, E10 and Tgo were all generated by Compartmentalised Self Replication (CSR) and defined in EP 18012113 (Medical Research Council). 5D4 (see Figure 5) was thus chosen for further characterisation.

15 Table 1. Primer extension analysis of various mutant polymerases using desulphonated and sulphonated bisulphite treated template.

	Taq	14	5D4	3A10	E10	Tgo-L	Tgo-P
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Desulph	Yes	Yes	Yes	Yes	No	No	No
Sulph	Yes	No	Yes	Yes	No	No	No

Tgo- L = crude lysate

Tgo-P = purified enzyme.

20 Table 2 shows a time course primer extension experiment using Control (C), Desulphonated (D) and Sulphonated (S) bisulphite treated BStemp3 template. The results show that even after as little as 1 minute the modified 5D4 enzyme is able to produce full-length product from the sulphonated template. After 15 minutes the 5D4 has copied the sulphonated template to completion while the wild type Taq polymerase still shows multiple stops indicating that the enzyme does not process sulphonated templates 25 efficiently. Interestingly, 5D4 also seems to copy desulphonated material more efficiently than the Taq polymerase.

Table 2. Time course primer extension using Taq polymerase and 5D4 with desulphonated and sulphonated bisulphite treated templates.

	Full length product generated					
	Taq Polymerase			5D4		
	1 min	5min	15 min	1 min	5 min	15 min
Control	+++	+++	+++	+++	+++	+++
Desulpho	+	++	+++	++	+++	+++
Sulphon	-	+	++	+	++	+++

- No full length product generated.

+ Low levels of full length product

5 ++ Considerable full length product.

+++ All product full length product.

Table 3 shows the results of a 5 minute primer extension reaction carried out on a C8 template. Again the 5D4 enzyme can be seen to be superior to the Taq polymerase with regard to its ability to copy both desulphonated and sulphonated targets containing a run of C's. Indeed the 5D4 completely copies the template while the wild type Taq polymerase again produces multiple stops.

Table 3. Comparison of 5D4 and Taq polymerase using a poly C8 bisulphite treated desulphonated and sulphonated templates.

	Taq Polymerase			5D4		
	Control	Desulph	Sulphon	Control	Desulph	Sulphon
C#1	-	+	+	-	-	-
C#2	-	+	+	-	-	-
C#3	-	++	+	-	+	+/-
C#4	-	++	++	-	+	+/-
C#5	-	++	++	-	+	+/-
C#6	-	+	++	-	+	+/-

	Taq Polymerase			5D4		
	Control	Desulph	Sulphon	Control	Desulph	Sulphon
C#7	-	++	+++	-	+	+/-
C#8	-	+	++	-	+/-	+/-

- Indicates no blocks observed on the gel.

+/- Indicates very faint stops observed.

+ Multiple blockage products.

++ Multiple heavy blockages.

5 +++ Intense stoppage products.

Table 4 shows both the fidelity of the 5D4 and again the superior ability to bypass sulphonated DNA lesions compared to Taq polymerase which is blocked by the sulphonated C templates.

10 NB after bisulphite treatment, C residues are converted to U, which are then copied by the polymerase as a T thus the polymerase should only add A into the growing chain if the enzyme exhibits 100% fidelity.

15 Table 4. Fidelity and processivity of 5D4 and Taq polymerase using a poly C8 bisulphite treated desulphonated and sulphonated templates.

	Taq Polymerase			5D4		
	Control	Desulph	Sulphon	Control	Desulph	Sulphon
G	Yes	No	No	Yes	No	No
A	No	Yes	Blocked	No	Yes	Yes
T	No	No	No	No	No	No
C	No	No	No	No	No	No

A comparison between Taq polymerase and 5D4 in a PCR reaction was carried out and the results are shown in Figure 4. Conditions were as follows:

Template: Tgo plasmid bisulphite treated

Primers: Set 1 amplify 250 bp fragment
 Set 2 amplify 600 bp fragment
 Set 3 amplify 1050 bp fragment

94°C - 20 seconds

5 45°C - 30 seconds; 25 cycles

68°C - 1 minute

Figure 4 shows the improved amplification achieved by 5D4 in a standard PCR reaction using bisulphite treated DNA. The enzyme is also able to copy a 250 bp
10 sulphonated template as can be seen from the arrow.

The scope of the claims should not be limited by the preferred embodiments set forth in the examples, but should be given the broadest interpretation consistent with the description as a whole.

Claims:

1. Use of enzyme Pol A DNA polymerase variant 5D4 for copying or amplifying bisulphite modified or treated nucleic acids, wherein the enzyme is more effective in copying or amplifying the nucleic acid compared with native Taq polymerase under similar conditions.
2. The use according to claim 1 wherein the enzyme is capable of copying or amplifying nucleic acid having abasic sites, bulky adducts including sulphonate groups, having only A, T, G and U bases, or only having A, T and G bases.
3. A method for copying or amplifying bisulphite treated nucleic acid comprising: bisulphite treating a nucleic acid; and copying or amplifying the bisulphite treated nucleic acid using Pol A DNA polymerase variant 5D4.
4. The method according to claim 3 wherein the bisulphite treatment uses sodium bisulphite or sodium metabisulphite.
5. The method according to claim 3 or 4 wherein the bisulphite treatment is without a desulphonation step.
6. The method according to any one of claims 3 to 5 further comprising denaturing the nucleic acid prior to bisulphite treatment.
7. The method according to claim 6 wherein the denaturing step is carried out by providing an alkali environment or heating the nucleic acid.
8. The method according to any one of claims 3 to 7 wherein any methylated nucleotides in the sample remain unchanged while unmethylated nucleotides are converted to uracil by the bisulphite treatment.
9. The method according to any one of claims 3 to 8 further comprising desalting or isolating the treated nucleic acid sample.

10. The method according to any one of claims 3 to 9 further comprising processing or analysing the treated nucleic acid to determine nucleotide sequence, methylation status, identify of source of nucleic acid, or detect a microorganism.
11. The method according to any one of claims 3 to 10 wherein the amplification of the bisulphite treated nucleic acid is carried out by polymerase chain reaction (PCR), Reverse-Transcriptase PCR, qPCR, isothermal amplification, or signal amplification.
12. The method according to any one of claims 3 to 11 wherein the treated nucleic acid comprises bisulphite modified DNA, bisulphite modified RNA, or a combination of bisulphite modified DNA and bisulphite modified RNA.

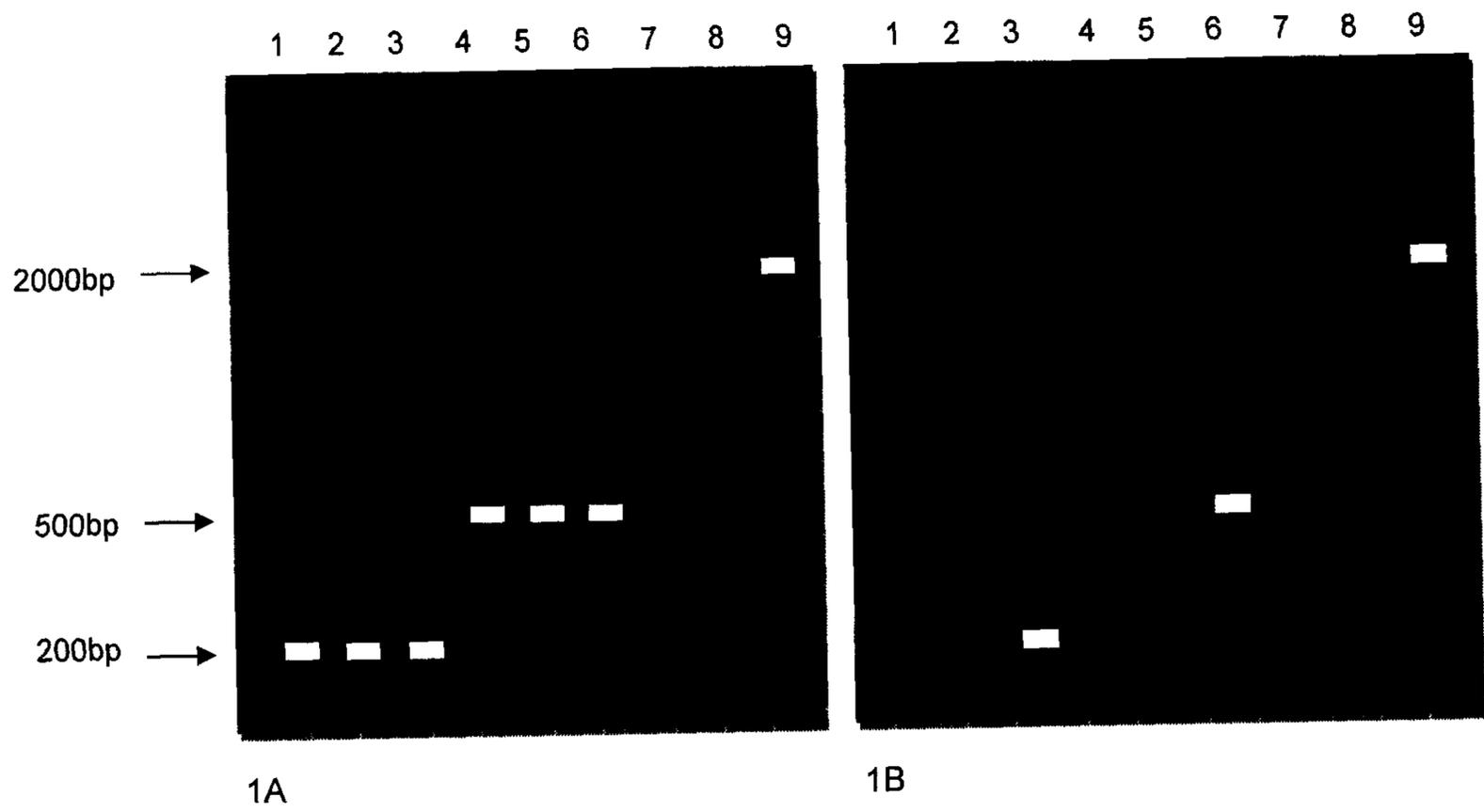
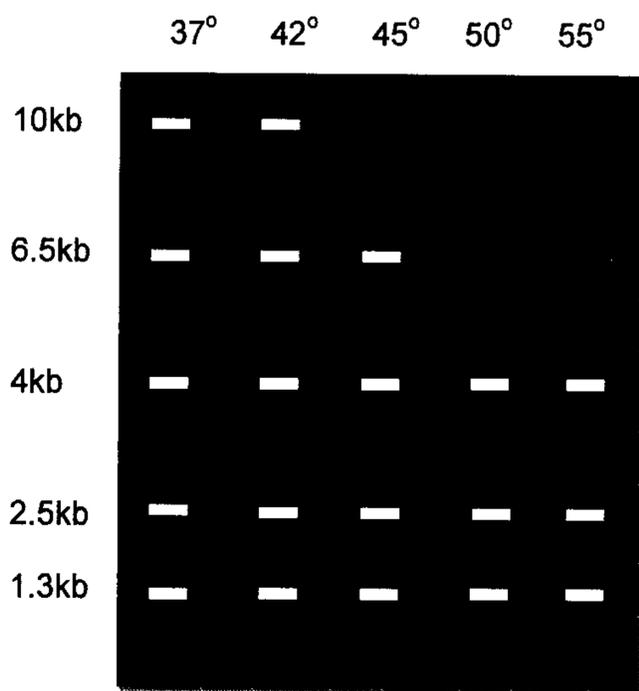
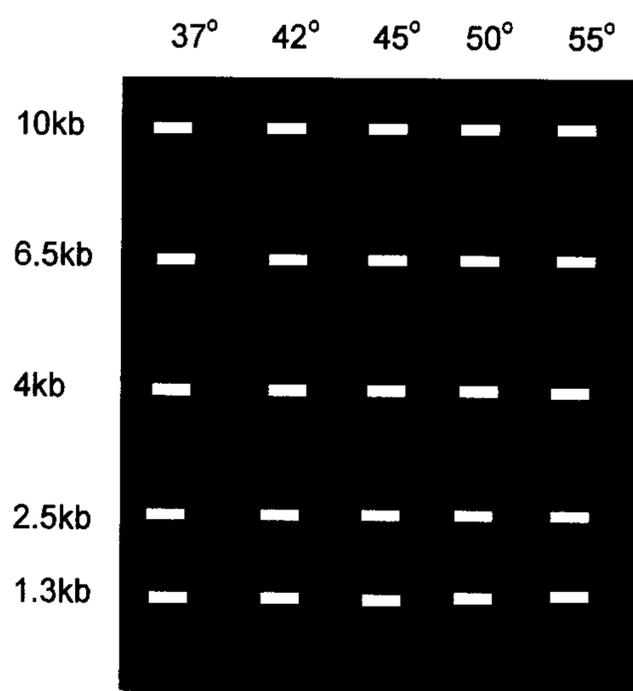


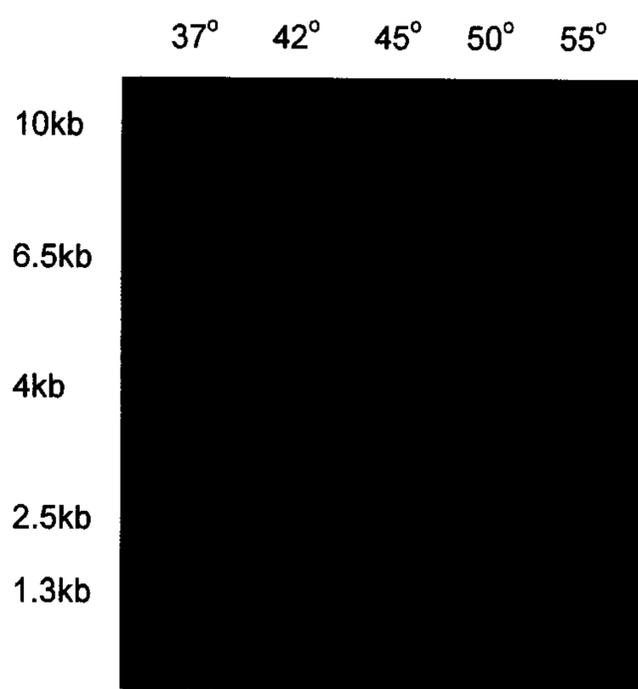
Figure 1



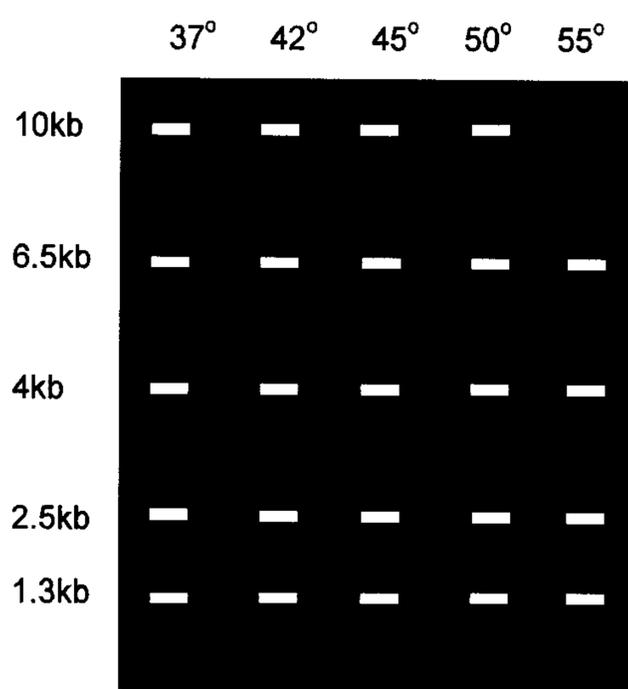
2A



2B



2C



2D

Figure 2

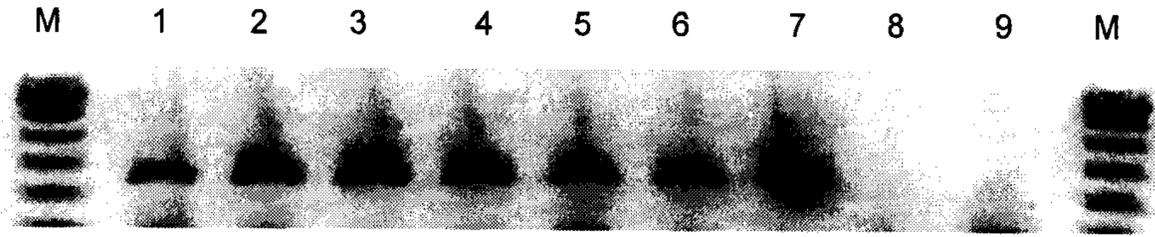


Figure 3

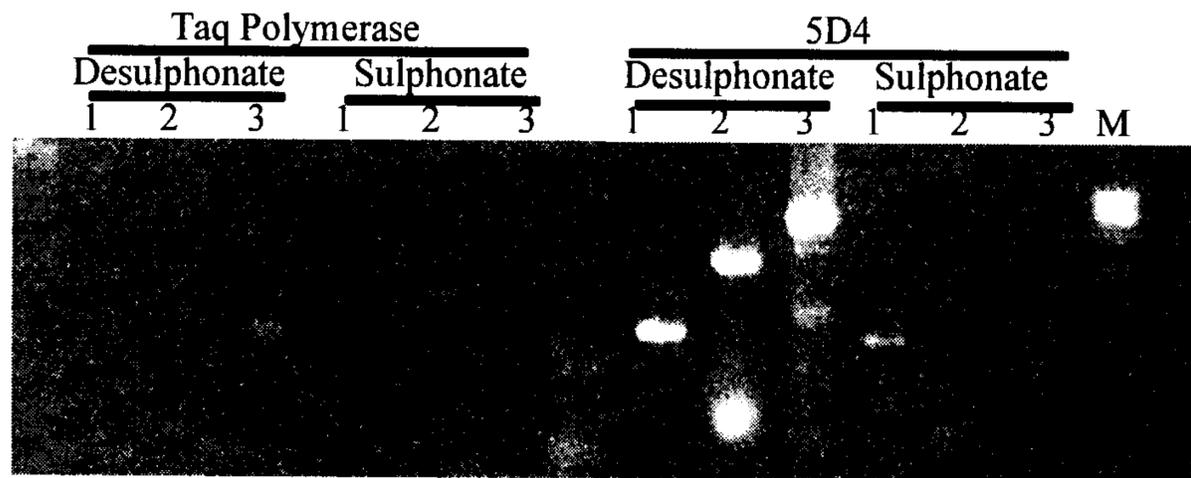


Figure 4

MAMLPLFEPKGRVLLVDGHHLAYRTFFALKGLTTSRGEVPQAVYGFASLLKALKEDGYKAI FVV
FDAQAPSRHEAHEAYKAGRAPSPEDFPRQLALIKELVDLLGFTRLEVQGYEADDVLATLAKKAE
KEGYEVRIILTADRDLYQLVSDRVAVLHPEGHLITPEWLWEKYGLRPEQWVDFRALVGDPSDNLPG
VKGIGECTALKLLKEWGSLENLLKNLDRKPAIRKILAHMDDLKLSWDLAKVRTDLPLEVDFAK
RRESDRERLRAFLELFEFGSLLHEFGLLESPKALEEAPWPPPVGAFVGFVLSRKEPMWADLLALA
AARGGRVHRAPEPYKALRDLKEARGLLAKDLSVLALREGLGLPPGDDPMLLAYLLDPSNTTPEVV
ARRYGGEWTEEAGERAAALSERLFANLWGRLEGEGRLLWLYRGVERPLSAVLAHMEATGVRLDVA
LRALSLEVAEEIARLEAEVFRLAGHPFNLSRDQLERVLFDLGLPAIGKTEKTGKRSTSAVLE
ALREAHPLVEKILQYRELTKLKSTYIDPLPDLIHPRTGRLHTRFNQTATATGRLSSSDPNLQNI
VRTPLGQRIRRAFLAEGGWLLVLDYSQMELRVLAHLSGDNLIRVVFQEGRDIHTETASWFMGVP
REAVDPLMRRAAKTINFGVLYGMSAHRLSQELAI PYEEAQAFIERYFQSFPKVRAWIEKTLEEGR
KRGYVETLFGRRRYVPDLNARVKS VREAAERMAFNMPVQGTAAADLT KLAMVKLFPRLEETGARM
LQVHDELVLEAPKERAEAVARLAKEVMGVYPLAVPLEVEVGIGEDWLSAKG

Figure 5