

CLAIMS

1. A method for introducing substitution mutations into at least one target nucleic acid molecule, in which the at least one target nucleic acid molecule is at least one target DNA molecule, comprising:

5 (a) providing at least one sample comprising the at least one target DNA molecule; and

(b) amplifying the at least one target DNA molecule using a low bias DNA polymerase, in which the low bias DNA polymerase is a high fidelity DNA polymerase, and in which the step of amplifying the at least one target DNA molecule is carried out
10 in the presence of a nucleotide analog and comprises at least 2 rounds of replicating the at least one target DNA molecule, in which in a first round of replication the DNA polymerase incorporates the nucleotide analog in place of a nucleotide and in a second round of replication, the nucleotide analog pairs with a natural nucleotide to introduce a substitution mutation in the complementary strand.

15 2. Use of a low bias DNA polymerase, in which the low bias DNA polymerase is a high fidelity DNA polymerase in a method for introducing mutations into at least one target nucleic acid, in which the at least one target nucleic acid molecule is at least one target DNA molecule, in which the method comprises:

20 (a) providing at least one sample comprising the at least one target DNA molecule; and

(b) amplifying the at least one target DNA molecule using the DNA polymerase; in which the step of amplifying the at least one target DNA molecule is carried out in the presence of a nucleotide analog and comprises at least 2 rounds of replicating the at least one target DNA molecule, in which in a first round of replication the DNA
25 polymerase incorporates the nucleotide analog in place of a nucleotide and in a second round of replication, the nucleotide analog pairs with a natural nucleotide to introduce a substitution mutation in the complementary strand.

3. A method or use according to Claim 1 or 2, in which the low bias DNA polymerase has low template amplification bias.

4. A method or use according to Claim 1, 2 or 3, in which the DNA polymerase mutates adenine, thymine, guanine, and cytosine nucleotides in the at least one target DNA molecule at a rate ratio of 0.5-1.5:0.5-1.5:0.5-1.5:0.5-1.5, 0.6-1.4:0.6-1.4:0.6-1.4:0.6-1.4, 0.7-1.3:0.7-1.3:0.7-1.3:0.7-1.3, 0.8-1.2:0.8-1.2:0.8-1.2:0.8-1.2, or around 1:1:1:1 respectively.
5. A method or use according to any preceding claim, in which the DNA polymerase mutates between 1% and 15%, between 2% and 10%, or around 8% of the nucleotides in the at least one target DNA molecule and/or in which the DNA polymerase mutates between 0% and 3%, or between 0% and 2% of the nucleotides in the at least one target DNA molecule per round of replication.
6. A method or use according to any preceding claim, in which the DNA polymerase mutates adenine, thymine, guanine, and/or cytosine in the at least one target DNA molecule using a nucleotide analog and/or in which the DNA polymerase replaces guanine, cytosine, adenine, and/or thymine with a nucleotide analog, and/or in which the DNA polymerase introduces guanine or adenine nucleotides using a nucleotide analog at a rate ratio of 0.5-1.5:0.5-1.5, 0.6-1.4:0.6-1.4, 0.7-1.3:0.7-1.3, 0.8-1.2:0.8-1.2, or around 1:1 respectively, and/or in which the DNA polymerase introduces guanine or adenine nucleotides using a nucleotide analog at a rate ratio of 0.7-1.3:0.7-1.3 respectively.
7. A method or use according to any preceding claim, in which the nucleotide analog is dPTP, optionally in which the DNA polymerase introduces guanine to adenine substitution mutations, cytosine to thymine substitution mutations, adenine to guanine substitution mutations, and thymine to cytosine substitution mutations, optionally in which the DNA polymerase introduces guanine to adenine substitution mutations, cytosine to thymine substitution mutations, adenine to guanine substitution mutations, and thymine to cytosine substitution mutations at a rate ratio of 0.5-1.5:0.5-1.5:0.5-1.5:0.5-1.5, 0.6-1.4:0.6-1.4:0.6-1.4:0.6-1.4, 0.7-1.3:0.7-1.3:0.7-1.3:0.7-1.3, 0.8-1.2:0.8-1.2:0.8-1.2:0.8-1.2, or around 1:1:1:1 respectively, optionally in which the DNA polymerase introduces guanine to adenine substitution mutations, cytosine to thymine substitution mutations, adenine to guanine substitution mutations, and thymine to cytosine substitution mutations at a rate ratio of 0.7-1.3:0.7-1.3:0.7-1.3:0.7-1.3 respectively.

8. A method or use according to any preceding claim, in which , in the absence of nucleotide analogs, the high fidelity DNA polymerase introduces less than 0.01%, less than 0.0015%, less than 0.001%, between 0% and 0.0015%, or between 0% and 0.001% mutations per round of replication.

5 9. A method or use according to any preceding claim, in which the method comprises a further step of amplifying the at least one target DNA molecule comprising nucleotide analogs in the absence of nucleotide analogs, optionally in which the step of amplifying the at least one target DNA molecule comprising nucleotide analogs in the absence of nucleotide analogs is carried out using the DNA polymerase.

10 10. A method or use according to any preceding claim, in which the method provides a mutated at least one target DNA molecule and the method further comprises a further step of amplifying the mutated at least one target DNA molecule using the DNA polymerase.

15 11. A method or use according to any preceding claim, in which the DNA polymerase comprises a proof-reading domain and/or a processivity enhancing domain, and/or in which the DNA polymerase comprises a fragment of at least 400, at least 500, at least 600, at least 700, or at least 750 contiguous amino acids of:

- (a) a sequence of SEQ ID NO. 2;
- (b) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 2;
- (c) a sequence of SEQ ID NO. 4;
- (d) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 4;
- (e) a sequence of SEQ ID NO. 6;
- (f) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 6;
- (g) a sequence of SEQ ID NO. 7; or
- (h) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 7;

25 optionally in which the DNA polymerase comprises:

- (a) a sequence of SEQ ID NO. 2;
- (b) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 2;

- (c) a sequence of SEQ ID NO. 4;
- (d) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 4;
- (e) a sequence of SEQ ID NO. 6;
- (f) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 6;
- (g) a sequence of SEQ ID NO. 7; or
- (h) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 7;

optionally in which

- (i) the DNA polymerase comprises a sequence at least 98% identical to SEQ ID NO. 2;
- (ii) the DNA polymerase comprises a sequence at least 98% identical to SEQ ID NO. 4;
- (iii) the DNA polymerase comprises a sequence at least 98% identical to SEQ ID NO. 6; or
- (iv) the DNA polymerase comprises a sequence at least 98% identical to SEQ ID NO. 7;

and/or

in which the low bias DNA polymerase is a thermococcal polymerase, or derivative thereof, optionally in which the thermococcal polymerase is derived from a thermococcal strain selected from the group consisting of *T.kodakarensis*, *T.siculi*, *T.celer* and *T.sp* KS-1.

12. A method or use according to any preceding claim, further comprising:

(A) introducing barcodes and/or sample tags into the at least one target DNA molecule, optionally in which:

- (a) a group of sample tags is used and target DNA molecules from different samples are labelled with different sample tags from the group, optionally in which each sample tag differs from substantially all other sample tags in the group by at least 1 low probability mutation difference or at least 3 high probability

mutation differences, optionally in which each sample tag differs from substantially all other sample tags in the group by at least 3 low probability mutation differences and/or each sample tag differs from substantially all other sample tags in the group by between 3 and 25, or between 3 and 10 low probability mutation differences and/or the low probability mutation is a transversion mutation or an indel mutation and/or in which the high probability mutation is a transition mutation;

or

(b) each sample tag differs from substantially all other sample tags in the group by at least 5 high probability mutation differences, optionally in which each sample tag differs from substantially all other sample tags in the group by between 5 and 25, or between 5 and 10 high probability mutation differences, optionally in which the high probability mutation is a transition mutation;

and/or

(B) introducing adapters into each of the at least one target DNA molecules, optionally in which introducing adapters into each of the at least one target DNA molecules comprises introducing a first adapter at the 3' end of the at least one target DNA molecule and a second adapter at the 5' end of the at least one target DNA molecule, in which the first adapter and the second adapter can anneal to one another, optionally in which the at least one target DNA molecule is amplified using primers that are identical to one another and complementary to a portion of the first adapter and/or in which the first adapter is complementary to a DNA molecule that is at least 80%, at least 90%, at least 95%, at least 99%, or 100% identical to the second adapter and/or in which the primers comprise a second primer binding site, and the method comprises amplifying the at least one target DNA molecule using the primers, removing the primers and further amplifying the at least one target DNA molecule using a second set of primers that anneal to the second primer binding site;

and/or

(C) introducing barcodes, sample tags and adapters into each of the target DNA molecules, and/or in which the barcodes, sample tags and/or adapters are introduced by tagmentation or by shearing and ligation.

13. A method or use according to any preceding claim, in which the at least one target DNA molecule is greater than 1 kbp, greater than 1.5 kbp, greater than 2 kbp, greater than 4 kbp, greater than 5 kbp, greater than 7 kbp, or greater than 8 kbp.

14. A method for determining a sequence of at least one target DNA molecule, comprising steps of:

(a) performing a method according to any of Claims 1 or 3 to 13 to provide at least one mutated target DNA molecule;

(b) optionally fragmenting and/or amplifying the at least one mutated target DNA molecule to provide at least one fragmented and/or amplified mutated target DNA molecule;

(c) sequencing regions of the at least one mutated target DNA molecule or the at least one fragmented and/or amplified mutated target DNA molecule to provide mutated sequence reads; and

(d) assembling a sequence for at least a portion of the at least one target DNA molecule using the mutated sequence reads.

15. A method for engineering a protein comprising steps of:

(a) performing the method according to any of Claims 1 or 3 to 13 to provide at least one mutated target DNA molecule;

(b) inserting the at least one mutated target DNA molecule into a vector; and

(c) expressing a protein encoded by the at least one mutated target DNA molecule optionally in which:

(i) the method comprises steps of:

(a) amplifying the at least one target DNA molecule comprising the nucleotide analog in the absence of the nucleotide analog to provide at least one mutated target DNA molecule;

(b) inserting the at least one mutated target DNA molecule into a vector; and

5 (c) expressing a protein encoded by the at least one mutated target DNA molecule; and/or

(ii) the method further comprises a step of testing the activity or assessing the structure of the protein encoded by the at least one mutated target DNA molecule; and/or

10 (iii) the vector is a plasmid, a virus, a cosmid, or an artificial chromosome; and/or

(iv) the step of expressing a protein encoded by the at least one mutated target DNA molecule is achieved by transforming bacterial cells, transfecting eukaryotic cells or transducing eukaryotic cells with the vector.

16. A method or use according to any of Claims 1 to 15, in which:

15 (I) the step of amplifying the at least one target DNA molecule using a low bias high fidelity DNA polymerase is carried out using dNTPs at unequal concentrations; and/or

(II) (A)

20 (i) the method comprises a further step of amplifying the at least one target DNA molecule comprising nucleotide analogs in the absence of nucleotide analogs and the further step of amplifying the at least one target DNA molecule comprising nucleotide analogs in the absence of nucleotide analogs is carried out using dNTPs at unequal concentrations; or

25 (ii) the method provides a mutated at least one target DNA molecule, the method comprises a further step of amplifying the mutated at least one target DNA molecule using the DNA polymerase, optionally in the absence of nucleotide analogs and the further step of amplifying the mutated at least one target DNA molecule using the DNA polymerase is carried out using dNTPs at unequal concentrations; and/or

(B) dNTPs at unequal concentrations are used to alter the profile of mutations that are introduced, optionally in which dNTPs at unequal concentrations are used to reduce bias in the profile of mutations that are introduced and/or

(C) the dNTPs at unequal concentrations comprise dATP, dCTP, dTTP and dGTP and one or two of dATP, dCTP, dTTP or dGTP are at a lower concentration compared to other dNTPs; and/or

(D) using dNTPs at unequal concentrations further comprises a step of identifying a dNTP whose level should be increased or decreased in order to reduce bias in the profile of mutations that are introduced; and/or

(E) the dNTPs at unequal concentrations comprise dTTP at a lower concentration than other dNTPs, optionally in which the dNTPs at unequal concentrations comprise dTTP at a concentration less than 75%, less than 70%, less than 60%, less than 55%, between 25% and 75%, between 25% and 70, between 25% and 60%, or around 50% of the concentration of dATP, dCTP or dGTP, optionally in which the dNTPs at unequal concentrations comprise dTTP at a concentration less than 75%, less than 70%, less than 60%, less than 55%, between 25% and 75%, between 25% and 70, between 25% and 60%, or around 50% of the concentration of dCTP, optionally in which the dNTPs at unequal concentrations comprise dTTP at a concentration less than 60% of the concentration of dCTP, optionally the dNTPs at unequal concentrations comprise dTTP at a concentration between 25% and 60% of the concentration of dCTP; and/or

(F) the dNTPs at unequal concentrations comprises dATP at a lower concentration compared to other dNTPs, optionally in which the dNTPs at unequal concentrations comprises dATP at a concentration less than 75%, less than 70%, less than 60%, less than 55%, between 25% and 75%, between 25% and 70%, between 25% and 60%, or around 50% of the concentration of dTTP, dCTP or dGTP, optionally in which the dNTPs at unequal concentrations of dNTPs comprises dATP at a concentration less than 75%, less than 70%, less than 60%, less than 55%, between 25% and 75%, between 25% and 70%, between 25% and

60%, or around 50% of the concentration of dGTP, optionally in which the dNTPs at unequal concentrations comprises dATP at a concentration less than 60% of the concentration of dGTP, optionally in which the dNTPs at unequal concentrations comprises dATP at a concentration between 25% and 60% of the concentration of dGTP;

and/or

(III) the method is a method for introducing mutations in a low bias mutation profile.

17. A group of sample tags, in which each sample tag differs from substantially all other sample tags in the group by at least one low probability mutation difference or at least three high probability mutation differences.

18. A method for designing a group of sample tags suitable for use in a method for introducing mutations into at least one target nucleic acid molecule comprising:

(a) analysing the method for introducing mutations into at least one target nucleic acid molecule and determining the average number of low probability mutations that take place during the method for introducing mutations into at least one target nucleic acid molecule; and

(b) determining sequences for a group of sample tags in which each sample tag differs from substantially all sample tags in the group by more low probability differences than the average number of low probability mutations that take place during the method for introducing mutations into at least one target nucleic acid molecule.

19. A method for preferentially amplifying target nucleic acid molecules that are larger than 1 kbp in length comprising:

(a) providing at least one sample comprising target nucleic acid molecules;

(b) introducing a first adapter at the 3' end of target nucleic acid molecules and a second adapter at the 5' end of target nucleic acid molecules; and

(c) amplifying the target nucleic acid molecules using primers that are complementary to a portion of the first adapter,

in which the first adapter and the second adapter can anneal to one another.

20. A method for introducing mutations into at least one target nucleic acid molecule comprising:

(a) providing at least one sample comprising at least one target nucleic acid molecule;
and

5 (b) introducing mutations into the at least one target nucleic acid molecule
by amplifying the at least one target nucleic acid molecule using a DNA polymerase to
provide a mutated at least one target nucleic acid molecule,
in which step (b) is carried out using dNTPs at unequal concentrations.

CLAIMS

1. A method for introducing substitution mutations into at least one target nucleic acid molecule, in which the at least one target nucleic acid molecule is at least one target DNA molecule, comprising:

5 (a) providing at least one sample comprising the at least one target DNA molecule;
and

(b) amplifying the at least one target DNA molecule using a low bias DNA polymerase, in which the low bias DNA polymerase is a high fidelity DNA polymerase, and in which the step of amplifying the at least one target DNA molecule is carried out
10 in the presence of a nucleotide analog and comprises at least 2 rounds of replicating the at least one target DNA molecule, in which in a first round of replication the DNA polymerase incorporates the nucleotide analog in place of a nucleotide and in a second round of replication, the nucleotide analog pairs with a natural nucleotide to introduce a substitution mutation in the complementary strand.

15 2. Use of a low bias DNA polymerase, in which the low bias DNA polymerase is a high fidelity DNA polymerase in a method for introducing mutations into at least one target nucleic acid, in which the at least one target nucleic acid molecule is at least one target DNA molecule, in which the method comprises:

20 (a) providing at least one sample comprising the at least one target DNA molecule;
and

(b) amplifying the at least one target DNA molecule using the DNA polymerase;
in which the step of amplifying the at least one target DNA molecule is carried out in the presence of a nucleotide analog and comprises at least 2 rounds of replicating the at least one target DNA molecule, in which in a first round of replication the DNA
25 polymerase incorporates the nucleotide analog in place of a nucleotide and in a second round of replication, the nucleotide analog pairs with a natural nucleotide to introduce a substitution mutation in the complementary strand.

3. A method or use according to Claim 1 or 2, in which the low bias DNA polymerase has low template amplification bias.

4. A method or use according to Claim 1, 2 or 3, in which the DNA polymerase mutates adenine, thymine, guanine, and cytosine nucleotides in the at least one target DNA molecule at a rate ratio of 0.5-1.5:0.5-1.5:0.5-1.5:0.5-1.5, 0.6-1.4:0.6-1.4:0.6-1.4:0.6-1.4, 0.7-1.3:0.7-1.3:0.7-1.3:0.7-1.3, 0.8-1.2:0.8-1.2:0.8-1.2:0.8-1.2, or around 1:1:1:1 respectively.
5. A method or use according to any preceding claim, in which the DNA polymerase mutates between 1% and 15%, between 2% and 10%, or around 8% of the nucleotides in the at least one target DNA molecule and/or in which the DNA polymerase mutates between 0% and 3%, or between 0% and 2% of the nucleotides in the at least one target DNA molecule per round of replication.
6. A method or use according to any preceding claim, in which the DNA polymerase mutates adenine, thymine, guanine, and/or cytosine in the at least one target DNA molecule using a nucleotide analog and/or in which the DNA polymerase replaces guanine, cytosine, adenine, and/or thymine with a nucleotide analog, and/or in which the DNA polymerase introduces guanine or adenine nucleotides using a nucleotide analog at a rate ratio of 0.5-1.5:0.5-1.5, 0.6-1.4:0.6-1.4, 0.7-1.3:0.7-1.3, 0.8-1.2:0.8-1.2, or around 1:1 respectively, and/or in which the DNA polymerase introduces guanine or adenine nucleotides using a nucleotide analog at a rate ratio of 0.7-1.3:0.7-1.3 respectively.
7. A method or use according to any preceding claim, in which the nucleotide analog is dPTP, optionally in which the DNA polymerase introduces guanine to adenine substitution mutations, cytosine to thymine substitution mutations, adenine to guanine substitution mutations, and thymine to cytosine substitution mutations, optionally in which the DNA polymerase introduces guanine to adenine substitution mutations, cytosine to thymine substitution mutations, adenine to guanine substitution mutations, and thymine to cytosine substitution mutations at a rate ratio of 0.5-1.5:0.5-1.5:0.5-1.5:0.5-1.5, 0.6-1.4:0.6-1.4:0.6-1.4:0.6-1.4, 0.7-1.3:0.7-1.3:0.7-1.3:0.7-1.3, 0.8-1.2:0.8-1.2:0.8-1.2:0.8-1.2, or around 1:1:1:1 respectively, optionally in which the DNA polymerase introduces guanine to adenine substitution mutations, cytosine to thymine substitution mutations, adenine to guanine substitution mutations, and thymine to cytosine substitution mutations at a rate ratio of 0.7-1.3:0.7-1.3:0.7-1.3:0.7-1.3 respectively.

8. A method or use according to any preceding claim, in which , in the absence of nucleotide analogs, the high fidelity DNA polymerase introduces less than 0.01%, less than 0.0015%, less than 0.001%, between 0% and 0.0015%, or between 0% and 0.001% mutations per round of replication.
- 5 9. A method or use according to any preceding claim, in which the method comprises a further step of amplifying the at least one target DNA molecule comprising nucleotide analogs in the absence of nucleotide analogs, optionally in which the step of amplifying the at least one target DNA molecule comprising nucleotide analogs in the absence of nucleotide analogs is carried out using the DNA polymerase.
- 10 10. A method or use according to any preceding claim, in which the method provides a mutated at least one target DNA molecule and the method further comprises a further step of amplifying the mutated at least one target DNA molecule using the DNA polymerase.
- 15 11. A method or use according to any preceding claim, in which the DNA polymerase comprises a proof-reading domain and/or a processivity enhancing domain, and/or in which the DNA polymerase comprises a fragment of at least 400, at least 500, at least 600, at least 700, or at least 750 contiguous amino acids of:
- (a) a sequence of SEQ ID NO. 2;
 - (b) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 2;
 - (c) a sequence of SEQ ID NO. 4;
 - 20 (d) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 4;
 - (e) a sequence of SEQ ID NO. 6;
 - (f) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 6;
 - (g) a sequence of SEQ ID NO. 7; or
 - (h) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 7;
- 25 optionally in which the DNA polymerase comprises:
- (a) a sequence of SEQ ID NO. 2;
 - (b) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 2;

- (c) a sequence of SEQ ID NO. 4;
- (d) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 4;
- (e) a sequence of SEQ ID NO. 6;
- (f) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 6;
- (g) a sequence of SEQ ID NO. 7; or
- (h) a sequence at least 95%, at least 98%, or at least 99% identical to SEQ ID NO. 7;

optionally in which

- (i) the DNA polymerase comprises a sequence at least 98% identical to SEQ ID NO. 2;
- (ii) the DNA polymerase comprises a sequence at least 98% identical to SEQ ID NO. 4;
- (iii) the DNA polymerase comprises a sequence at least 98% identical to SEQ ID NO. 6; or
- (iv) the DNA polymerase comprises a sequence at least 98% identical to SEQ ID NO. 7;

and/or

in which the low bias DNA polymerase is a thermococcal polymerase, or derivative thereof, optionally in which the thermococcal polymerase is derived from a thermococcal strain selected from the group consisting of *T.kodakarensis*, *T.siculi*, *T.celer* and *T.sp* KS-1.

12. A method or use according to any preceding claim, further comprising:

(A) introducing barcodes and/or sample tags into the at least one target DNA molecule, optionally in which:

- (a) a group of sample tags is used and target DNA molecules from different samples are labelled with different sample tags from the group, optionally in which each sample tag differs from substantially all other sample tags in the group by at least 1 low probability mutation difference or at least 3 high probability

mutation differences, optionally in which each sample tag differs from substantially all other sample tags in the group by at least 3 low probability mutation differences and/or each sample tag differs from substantially all other sample tags in the group by between 3 and 25, or between 3 and 10 low probability mutation differences and/or the low probability mutation is a transversion mutation or an indel mutation and/or in which the high probability mutation is a transition mutation;

or

(b) each sample tag differs from substantially all other sample tags in the group by at least 5 high probability mutation differences, optionally in which each sample tag differs from substantially all other sample tags in the group by between 5 and 25, or between 5 and 10 high probability mutation differences, optionally in which the high probability mutation is a transition mutation;

and/or

(B) introducing adapters into each of the at least one target DNA molecules, optionally in which introducing adapters into each of the at least one target DNA molecules comprises introducing a first adapter at the 3' end of the at least one target DNA molecule and a second adapter at the 5' end of the at least one target DNA molecule, in which the first adapter and the second adapter can anneal to one another, optionally in which the at least one target DNA molecule is amplified using primers that are identical to one another and complementary to a portion of the first adapter and/or in which the first adapter is complementary to a DNA molecule that is at least 80%, at least 90%, at least 95%, at least 99%, or 100% identical to the second adapter and/or in which the primers comprise a second primer binding site, and the method comprises amplifying the at least one target DNA molecule using the primers, removing the primers and further amplifying the at least one target DNA molecule using a second set of primers that anneal to the second primer binding site;

and/or

(C) introducing barcodes, sample tags and adapters into each of the target DNA molecules, and/or in which the barcodes, sample tags and/or adapters are introduced by tagmentation or by shearing and ligation.

13. A method or use according to any preceding claim, in which the at least one target DNA molecule is greater than 1 kbp, greater than 1.5 kbp, greater than 2 kbp, greater than 4 kbp, greater than 5 kbp, greater than 7 kbp, or greater than 8 kbp.

14. A method for determining a sequence of at least one target DNA molecule, comprising steps of:

10 (a) performing a method according to any of Claims 1 or 3 to 13 to provide at least one mutated target DNA molecule;

(b) optionally fragmenting and/or amplifying the at least one mutated target DNA molecule to provide at least one fragmented and/or amplified mutated target DNA molecule;

15 (c) sequencing regions of the at least one mutated target DNA molecule or the at least one fragmented and/or amplified mutated target DNA molecule to provide mutated sequence reads; and

(d) assembling a sequence for at least a portion of the at least one target DNA molecule using the mutated sequence reads.

15. A method for engineering a protein comprising steps of:

20 (a) performing the method according to any of Claims 1 or 3 to 13 to provide at least one mutated target DNA molecule;

(b) inserting the at least one mutated target DNA molecule into a vector; and

(c) expressing a protein encoded by the at least one mutated target DNA molecule optionally in which:

25 (i) the method comprises steps of:

- (a) amplifying the at least one target DNA molecule comprising the nucleotide analog in the absence of the nucleotide analog to provide at least one mutated target DNA molecule;
 - (b) inserting the at least one mutated target DNA molecule into a vector; and
 - 5 (c) expressing a protein encoded by the at least one mutated target DNA molecule; and/or
- (ii) the method further comprises a step of testing the activity or assessing the structure of the protein encoded by the at least one mutated target DNA molecule; and/or
 - 10 (iii) the vector is a plasmid, a virus, a cosmid, or an artificial chromosome; and/or
 - (iv) the step of expressing a protein encoded by the at least one mutated target DNA molecule is achieved by transforming bacterial cells, transfecting eukaryotic cells or transducing eukaryotic cells with the vector.
16. A method or use according to any of Claims 1 to 15, in which:
- 15 (I) the step of amplifying the at least one target DNA molecule using a low bias high fidelity DNA polymerase is carried out using dNTPs at unequal concentrations; and/or
 - (II) (A)
 - (i) the method comprises a further step of amplifying the at least one target DNA molecule comprising nucleotide analogs in the absence of nucleotide
 - 20 analogs and the further step of amplifying the at least one target DNA molecule comprising nucleotide analogs in the absence of nucleotide analogs is carried out using dNTPs at unequal concentrations; or
 - (ii) the method provides a mutated at least one target DNA molecule, the method comprises a further step of amplifying the mutated at least one target DNA
 - 25 molecule using the DNA polymerase, optionally in the absence of nucleotide analogs and the further step of amplifying the mutated at least one target DNA molecule using the DNA polymerase is carried out using dNTPs at unequal concentrations; and/or

(B) dNTPs at unequal concentrations are used to alter the profile of mutations that are introduced, optionally in which dNTPs at unequal concentrations are used to reduce bias in the profile of mutations that are introduced and/or

(C) the dNTPs at unequal concentrations comprise dATP, dCTP, dTTP and dGTP and one or two of dATP, dCTP, dTTP or dGTP are at a lower concentration compared to other dNTPs; and/or

(D) using dNTPs at unequal concentrations further comprises a step of identifying a dNTP whose level should be increased or decreased in order to reduce bias in the profile of mutations that are introduced; and/or

(E) the dNTPs at unequal concentrations comprise dTTP at a lower concentration than other dNTPs, optionally in which the dNTPs at unequal concentrations comprise dTTP at a concentration less than 75%, less than 70%, less than 60%, less than 55%, between 25% and 75%, between 25% and 70, between 25% and 60%, or around 50% of the concentration of dATP, dCTP or dGTP, optionally in which the dNTPs at unequal concentrations comprise dTTP at a concentration less than 75%, less than 70%, less than 60%, less than 55%, between 25% and 75%, between 25% and 70, between 25% and 60%, or around 50% of the concentration of dCTP, optionally in which the dNTPs at unequal concentrations comprise dTTP at a concentration less than 60% of the concentration of dCTP, optionally the dNTPs at unequal concentrations comprise dTTP at a concentration between 25% and 60% of the concentration of dCTP; and/or

(F) the dNTPs at unequal concentrations comprises dATP at a lower concentration compared to other dNTPs, optionally in which the dNTPs at unequal concentrations comprises dATP at a concentration less than 75%, less than 70%, less than 60%, less than 55%, between 25% and 75%, between 25% and 70%, between 25% and 60%, or around 50% of the concentration of dTTP, dCTP or dGTP, optionally in which the dNTPs at unequal concentrations of dNTPs comprises dATP at a concentration less than 75%, less than 70%, less than 60%, less than 55%, between 25% and 75%, between 25% and 70%, between 25% and

60%, or around 50% of the concentration of dGTP, optionally in which the dNTPs at unequal concentrations comprises dATP at a concentration less than 60% of the concentration of dGTP, optionally in which the dNTPs at unequal concentrations comprises dATP at a concentration between 25% and 60% of the concentration of dGTP;

5

and/or

(III) the method is a method for introducing mutations in a low bias mutation profile.