

CLAIMS

1. A library of multimeric barcoding reagents comprising at least 10 multimeric barcoding reagents for labelling a target nucleic acid for sequencing, in which each multimeric barcoding reagent comprises:

(a) first and second barcode molecules comprised within a nucleic acid molecule, in which each of the barcode molecules comprises a nucleic acid sequence comprising a barcode region; and

(b) first and second barcoded oligonucleotides, in which the first barcoded oligonucleotide comprises, optionally in the 5' to 3' direction, a barcode region complementary and annealed to the barcode region of the first barcode molecule and a target region capable of annealing or ligating to a first sub-sequence of the target nucleic acid, and in which the second barcoded oligonucleotide comprises, optionally in the 5' to 3' direction, a barcode region complementary and annealed to the barcode region of the second barcode molecule and a target region capable of annealing or ligating to a second sub-sequence of the target nucleic acid;

and in which the barcode regions of the first and second barcode molecules of each multimeric barcoding reagent are different to the barcode regions of at least 9 other multimeric barcoding reagents in the library.

2. A library of multimeric barcoding reagents according to Claim 1, in which the barcode regions uniquely identify each of the barcode molecules.

3. A library of multimeric barcoding reagents according to Claim 1 or 2, in which:

(a) each of the barcode molecules comprises a nucleic acid sequence comprising in the 5' to 3' direction an adapter region and a barcode region;

(b) each first barcoded oligonucleotide comprises in the 5' to 3' direction a barcode region annealed to the barcode region of the first barcode molecule, an adapter region annealed to the adapter region of the first barcode molecule and a target region capable of annealing to a first sub-sequence of the target nucleic acid; and

(c) each second barcoded oligonucleotide comprises in the 5' to 3' direction a barcode region annealed to the barcode region of the second barcode molecule, an adapter region

annealed to the adapter region of the second barcode molecule and a target region capable of annealing to a second sub-sequence of the target nucleic acid.

4. A library of multimeric barcoding reagents according to Claim 1, 2 or 3, in which each multimeric barcoding reagent comprises:

5 (a) at least 5, at least 10, at least 20, at least 25, at least 50, at least 75 or at least 100 barcode molecules linked together, in which each barcode molecule is as defined in Claim 1, 2 or 3; and

(b) a barcoded oligonucleotide annealed to each barcode molecule, in which each barcoded oligonucleotide is as defined in Claim 1, 2 or 3.

10 5. A library of multimeric barcoding reagents according to any preceding claim, in which the library comprises at least 20, at least 25, at least 50, at least 75, at least 100, at least 250, at least 500, at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 or at least 10^8 multimeric barcoding reagents as defined in any preceding claim.

6. A kit for labelling a target nucleic acid for sequencing, in which the kit comprises:

15 (a) a library of multimeric barcoding reagents comprising at least 10 multimeric barcoding reagents, in which each multimeric barcoding reagent comprises:

20 (i) first and second barcode molecules comprised within a nucleic acid molecule, in which each of the barcode molecules comprises a nucleic acid sequence comprising, optionally in the 5' to 3' direction, an adapter region and a barcode region, and

25 (ii) first and second barcode oligonucleotides, in which the first barcode oligonucleotide comprises a barcode region complementary and annealed to the barcode region of the first barcode molecule, and in which the second barcode oligonucleotide comprises a barcode region complementary and annealed to the barcode region of the second barcode molecule;

and in which the barcode regions of the first and second barcode molecules of each multimeric barcoding reagent are different to the barcode regions of at least 9 other multimeric barcoding reagents in the library, and

30 (b) first and second adapter oligonucleotides for each of the multimeric barcoding reagents, in which the first adapter oligonucleotide comprises, optionally in the 5' to 3'

direction, an adaptor region capable of annealing to the adapter region of the first barcode molecule and a target region capable of annealing or ligating to a first sub-sequence of the target nucleic acid, and in which the second adapter oligonucleotide comprises, optionally in the 5' to 3' direction, an adaptor region capable of annealing to the adapter region of the second barcode molecule and a target region capable of annealing or ligating to a second sub-sequence of the target nucleic acid.

7. A kit according to Claim 6, in which the barcode regions uniquely identify each of the barcode molecules and/or in which the adapter regions of the adapter oligonucleotides comprise an identical constant region.

8. A kit according to Claim 6 or 7, in which each multimeric barcoding reagent comprises at least 5, at least 10, at least 20, at least 25, at least 50, at least 75 or at least 100 barcode molecules linked together, in which each barcode molecule is as defined in Claim 6 or 7; and in which the kit comprises an adaptor oligonucleotide capable of annealing to each barcode molecule, in which each adaptor oligonucleotide is as defined in Claim 6 or 7.

9. A kit according to Claim 6, 7 or 8, in which the kit comprises a library of at least 20, at least 25, at least 50, at least 75, at least 100, at least 250, at least 500, at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 , at least 10^8 or at least 10^9 multimeric barcoding reagents, in which each multimeric barcoding reagent is as defined in Claim 6, 7 or 8, and adaptor oligonucleotides for each of the multimeric barcoding reagents, in which each adaptor oligonucleotide is as defined in Claim 6, 7 or 8.

10. A kit for labelling a target nucleic acid for sequencing, in which the kit comprises:

(a) a library of multimeric barcode molecules comprising at least 10 multimeric barcode molecules, each multimeric barcode molecule comprising first and second barcode molecules comprised within a nucleic acid molecule, in which each of the barcode molecules comprises a nucleic acid sequence comprising, optionally in the 5' to 3' direction, an adaptor region, a barcode region, and a priming region, and in which the barcode regions of the first and second barcode molecules of each multimeric barcode molecule are different to the barcode regions of at least 9 other multimeric barcode molecules in the library;

(b) first and second extension primers for each of the multimeric barcode molecules, in which the first extension primer comprises a sequence capable of annealing to the

priming region of the first barcode molecule, and in which the second extension primer comprises a sequence capable of annealing to the priming region of the second barcode molecule; and

(c) first and second adapter oligonucleotides for each of the multimeric barcode molecules, in which the first adapter oligonucleotide comprises, optionally in the 5' to 3' direction, an adaptor region capable of annealing to the adapter region of the first barcode molecule and a target region capable of annealing or ligating to a first sub-sequence of the target nucleic acid, and in which the second adapter oligonucleotide comprises, optionally in the 5' to 3' direction, an adaptor region capable of annealing to the adapter region of the second barcode molecule and a target region capable of annealing or ligating to a second sub-sequence of the target nucleic acid.

11. A kit for labelling a target nucleic acid for sequencing, in which the kit comprises:

(a) a library of multimeric barcode molecules comprising at least 10 multimeric barcode molecules, each multimeric barcode molecule comprising first and second barcode molecules comprised within a nucleic acid molecule, in which each of the barcode molecules comprises a nucleic acid sequence comprising, optionally in the 5' to 3' direction, an adapter region and a barcode region, and in which the barcode regions of the first and second barcode molecules of each multimeric barcode molecule are different to the barcode regions of at least 9 other multimeric barcode molecules in the library; and

(b) first and second adapter oligonucleotides for each of the multimeric barcode molecules, in which the first adapter oligonucleotide comprises, optionally in the 5' to 3' direction, an adaptor region capable of annealing to the adapter region of the first barcode molecule and a target region capable of annealing or ligating to a first sub-sequence of the target nucleic acid, and in which the second adapter oligonucleotide comprises, optionally in the 5' to 3' direction, an adaptor region capable of annealing to the adapter region of the second barcode molecule and a target region capable of annealing or ligating to a second sub-sequence of the target nucleic acid.

12. A kit according to Claim 10 or 11, in which the barcode regions uniquely identify each of the barcode molecules, and/or in which the adapter regions of the adapter oligonucleotides comprise an identical constant region.

13. A kit according to Claim 10, 11 or 12, in which each multimeric barcode molecule comprises at least 5, at least 10, at least 20, at least 25, at least 50, at least 75 or at least 100 barcode molecules linked together, in which each barcode molecule is as defined in Claim 10, 11 or 12; and in which the kit comprises an adaptor oligonucleotide capable of annealing to
5 each barcode molecule, in which each adaptor oligonucleotide is as defined in Claim 10, 11 or 12.

14. A kit according to any of Claims 10 to 13, in which the kit comprises a library of at least 20, at least 25, at least 50, at least 75, at least 100, at least 250, at least 500, at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 , at least 10^8 or at least 10^9 multimeric barcode
10 molecules, in which each multimeric barcode molecule is as defined in any of Claims 10 to 13, and adaptor oligonucleotides for each of the multimeric barcode molecules, in which each adaptor oligonucleotide is as defined in any of Claims 10 to 13.

CLAIMS

1. Use of a library of multimeric barcoding reagents to
 - (i) produce two or more sequence reads from a target nucleic acid, wherein two or more sequence reads are identified as derived from the same target nucleic acid and combined to produce a synthetic long read,
 - (ii) label a formalin-fixed paraffin-embedded nucleic acid sample, wherein the multimeric barcoding reagents are introduced into the sample and used to label a set of two or more co-localised target nucleic acids for sequencing,
 - (iii) label target nucleic acids in an individual cell, wherein the multimeric barcoding reagents are introduced into a cell and used to label a set of two or more target nucleic acids in the cell for sequencing, or
 - (iv) label target nucleic acids in a sample of human plasma or serum, wherein the multimeric barcoding reagents are used to label a set of two or more target nucleic acids in the plasma or serum for sequencing;
- 15 in which the library of multimeric barcoding reagents comprises at least 10 multimeric barcoding reagents for labelling a target nucleic acid for sequencing, in which each multimeric barcoding reagent comprises:
 - (a) first and second barcode molecules comprised within a nucleic acid molecule, in which each of the barcode molecules comprises a nucleic acid sequence comprising a barcode region; and
 - (b) first and second barcoded oligonucleotides, in which the first barcoded oligonucleotide comprises, optionally in the 5' to 3' direction, a barcode region complementary and annealed to the barcode region of the first barcode molecule and a target region that anneals or ligates to a first sub-sequence of a target nucleic acid, and in which the second barcoded oligonucleotide comprises, optionally in the 5' to 3' direction, a barcode region complementary and annealed to the barcode region of the second barcode molecule and a target region that anneals or ligates to a second sub-sequence of the target nucleic acid;

and in which the barcode regions of the first and second barcode molecules of each multimeric barcoding reagent are different to the barcode regions of at least 9 other multimeric barcoding reagents in the library.

2. Use according to Claim 1, in which the barcode regions uniquely identify each of the
5 barcode molecules.
3. Use according to Claim 1 or 2, in which:
 - (a) each of the barcode molecules comprises a nucleic acid sequence comprising in the 5' to 3' direction an adapter region and a barcode region;
 - (b) each first barcoded oligonucleotide comprises in the 5' to 3' direction a barcode
10 region annealed to the barcode region of the first barcode molecule, an adapter region annealed to the adapter region of the first barcode molecule and a target region that anneals to a first sub-sequence of a target nucleic acid; and
 - (c) each second barcoded oligonucleotide comprises in the 5' to 3' direction a barcode
15 region annealed to the barcode region of the second barcode molecule, an adapter region annealed to the adapter region of the second barcode molecule and a target region that anneals to a second sub-sequence of the target nucleic acid.
4. Use according to Claim 1, 2 or 3, in which each multimeric barcoding reagent comprises:
 - (a) at least 5, at least 10, at least 20, at least 25, at least 50, at least 75 or at least 100
20 barcode molecules linked together, in which each barcode molecule is as defined in Claim 1, 2 or 3; and
 - (b) a barcoded oligonucleotide annealed to each barcode molecule, in which each barcoded oligonucleotide is as defined in Claim 1, 2 or 3.
5. Use according to any preceding claim, in which the library comprises at least 20, at least
25 25, at least 50, at least 75, at least 100, at least 250, at least 500, at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 or at least 10^8 multimeric barcoding reagents as defined in any preceding claim.
6. Use of a kit to

- (i) produce two or more sequence reads from a target nucleic acid, wherein two or more sequence reads are identified as derived from the same target nucleic acid and combined to produce a synthetic long read,
- (ii) label a formalin-fixed paraffin-embedded nucleic acid sample, wherein the multimeric barcoding reagents are introduced into the sample and used to label a set of two or more co-localised target nucleic acids for sequencing,
- (iii) label target nucleic acids in an individual cell, wherein the multimeric barcoding reagents are introduced into a cell and used to label a set of two or more target nucleic acids in the cell for sequencing, or
- (iv) label target nucleic acids in a sample of human plasma or serum, wherein the multimeric barcoding reagents are used to label a set of two or more target nucleic acids in the plasma or serum for sequencing;

in which the kit is for labelling a target nucleic acid for sequencing, in which the kit comprises:

- (a) a library of multimeric barcode molecules comprising at least 10 multimeric barcode molecules, each multimeric barcode molecule comprising first and second barcode molecules comprised within a nucleic acid molecule, in which each of the barcode molecules comprises a nucleic acid sequence comprising, optionally in the 5' to 3' direction, an adapter region and a barcode region, and in which the barcode regions of the first and second barcode molecules of each multimeric barcode molecule are different to the barcode regions of at least 9 other multimeric barcode molecules in the library, and
- (b) first and second adapter oligonucleotides for each of the multimeric barcode molecules, in which the first adapter oligonucleotide comprises, optionally in the 5' to 3' direction, an adaptor region that anneals to the adapter region of the first barcode molecule and a target region that anneals or ligates to a first sub-sequence of a target nucleic acid, and in which the second adapter oligonucleotide comprises, optionally in the 5' to 3' direction, an adaptor region that anneals to the adapter region of the second barcode molecule and a target region that anneals or ligates to a second sub-sequence of the target nucleic acid.

7. Use according to Claim 6, in which the kit comprises:
- (a) a library of multimeric barcode molecules comprising at least 10 multimeric barcode molecules, each multimeric barcode molecule comprising first and second barcode molecules comprised within a nucleic acid molecule, wherein each of the barcode molecules comprises a nucleic acid sequence comprising, optionally in the 5' to 3' direction, an adapter region, a barcode region, and a priming region, and wherein the barcode regions of the first and second barcode molecules of each multimeric barcode molecule are different to the barcode regions of at least 9 other multimeric barcode molecules in the library;
- (b) first and second extension primers for each of the multimeric barcode molecules, wherein the first extension primer comprises a sequence capable of annealing to the priming region of the first barcode molecule, and wherein the second extension primer comprises a sequence capable of annealing to the priming region of the second barcode molecule; and
- (c) first and second adapter oligonucleotides for each of the multimeric barcode molecules, wherein the first adapter oligonucleotide comprises, optionally in the 5' to 3' direction, an adaptor region that anneals to the adapter region of the first barcode molecule and a target region that anneals or ligates to a first sub-sequence of a target nucleic acid, and wherein the second adapter oligonucleotide comprises, optionally in the 5' to 3' direction, an adaptor region that anneals to the adapter region of the second barcode molecule and a target region that anneals or ligates to a second sub-sequence of the target nucleic acid.
8. Use according to Claim 6 or 7, in which the barcode regions uniquely identify each of the barcode molecules and/or in which the adapter regions of the adapter oligonucleotides comprise an identical constant region.
9. Use according to Claim 6, 7 or 8, in which each multimeric barcode molecule comprises at least 5, at least 10, at least 20, at least 25, at least 50, at least 75 or at least 100 barcode molecules linked together, in which each barcode molecule is as defined in Claim 6, 7 or 8; and in which the kit comprises an adaptor oligonucleotide that anneals to each barcode molecule, in which each adaptor oligonucleotide is as defined in Claim 6, 7 or 8.

10. Use according to any of Claims 6 to 9, in which the kit comprises a library of at least 20, at least 25, at least 50, at least 75, at least 100, at least 250, at least 500, at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 , at least 10^8 or at least 10^9 multimeric barcode molecules, in which each multimeric barcode molecule is as defined in Claim 6 to 9, and
- 5 adaptor oligonucleotides for each of the multimeric barcode molecules, in which each adaptor oligonucleotide is as defined in any of Claims 6 to 9.
11. Use according to Claim 6, in which the kit comprises:
- (a) a library of multimeric barcoding reagents comprising at least 10 multimeric barcoding reagents, in which each multimeric barcoding reagent comprises:
- 10 (i) first and second barcode molecules comprised within a nucleic acid molecule, in which each of the barcode molecules comprises a nucleic acid sequence comprising, optionally in the 5' to 3' direction, an adapter region and a barcode region, and
- (ii) first and second barcode oligonucleotides, in which the first barcode
- 15 oligonucleotide comprises a barcode region complementary and annealed to the barcode region of the first barcode molecule, and in which the second barcode oligonucleotide comprises a barcode region complementary and annealed to the barcode region of the second barcode molecule;
- and in which the barcode regions of the first and second barcode molecules of each
- 20 multimeric barcoding reagent are different to the barcode regions of at least 9 other multimeric barcoding reagents in the library; and
- (b) first and second adapter oligonucleotides for each of the multimeric barcoding reagents, in which the first adapter oligonucleotide comprises, optionally in the 5' to 3' direction, an adaptor region that anneals to the adapter region of the first barcode
- 25 molecule and a target region that anneals or ligates to a first sub-sequence of a target nucleic acid, and in which the second adapter oligonucleotide comprises, optionally in the 5' to 3' direction, an adaptor region that anneals to the adapter region of the second barcode molecule and a target region that anneals or ligates to a second sub-sequence of the target nucleic acid.

12. Use according to Claim 11, in which the barcode regions uniquely identify each of the barcode molecules, and/or in which the adapter regions of the adapter oligonucleotides comprise an identical constant region.

5 13. Use according to Claim 11 or 12, in which each multimeric barcoding reagent comprises at least 5, at least 10, at least 20, at least 25, at least 50, at least 75 or at least 100 barcode molecules linked together, in which each barcode molecule is as defined in Claim 11 or 12; and in which the kit comprises an adaptor oligonucleotide that anneals to each barcode molecule, in which each adaptor oligonucleotide is as defined in Claim 11 or 12.

10 14. Use according to Claim 11, 12 or 13, in which the kit comprises a library of at least 20, at least 25, at least 50, at least 75, at least 100, at least 250, at least 500, at least 10^3 , at least 10^4 , at least 10^5 , at least 10^6 , at least 10^7 , at least 10^8 or at least 10^9 multimeric barcoding reagents, in which each multimeric barcoding reagent is as defined in any of Claims 11, 12 or 13, and adaptor oligonucleotides for each of the multimeric barcoding reagents, in which each adaptor oligonucleotide is as defined in Claim 11, 12 or 13.