

Claims

1. Method for determining the sequence of a genomic region of interest comprising a target nucleotide sequence comprising, fragmenting a crosslinked DNA, ligating the fragmented cross linked DNA, reversing the crosslinking and determining at least part of the sequences of ligated DNA fragments which comprise a target nucleotide
5 sequence, and using the determined sequences to build a sequence of the genomic region of interest.
2. Method for determining the sequence of a genomic region of interest comprising a target nucleotide sequence, comprising the steps of:
10
 - a) providing a sample of cross-linked DNA;
 - b) fragmenting the crosslinked DNA;
 - c) ligating the fragmented crosslinked DNA;
 - d) reversing the crosslinking;
 - e) optionally fragmenting the DNA of step d), preferably with a restriction
15 enzyme;
 - f) optionally, ligating the fragmented DNA of step d) or e) to at least one Adaptor;
 - g) optionally and preferably, amplifying the ligated DNA fragments of step d) or e) comprising the target nucleotide sequence using at least one primer which
20 hybridises to the target nucleotide sequence, or amplifying the ligated DNA fragments of step f) using at least one primer which hybridises to the target nucleotide sequence and at least one primer which hybridises to the at least one adaptor;
 - h) determining at least part of the sequence of the (amplified) ligated DNA
25 fragments of step d), e), f) or g) comprising the target nucleotide sequence preferably using high throughput sequencing;
 - i) building a contig of the genomic region of interest from the determined sequences.
- 30 3. Method for determining the sequence of a genomic region of interest comprising a target nucleotide sequence, comprising the steps of:
 - a) providing a sample of cross-linked DNA;
 - b) fragmenting the crosslinked DNA;
 - c) ligating the fragmented crosslinked DNA;
 - 35 d) reversing the crosslinking;

- e) optionally fragmenting the DNA of step d), preferably with a restriction enzyme
f) circularising the DNA of step d) or e);
g) optionally and preferably, amplifying the circularised DNA comprising the target nucleotide sequence using preferably at least one primer which hybridises to the target nucleotide sequence;
h) determining at least part of the sequence of the (amplified) ligated DNA fragments comprising the target nucleotide using high throughput sequencing;
i) building a contig of the genomic region of interest from the determined sequences .

10

4. Method for determining the sequence of a genomic region of interest according to claims 2 or 3, wherein the genomic region of interest comprises one or more target nucleotide sequences in addition, and wherein in the amplification step (g) a primer is provided that hybridises with the target nucleotide sequence and one or more primers are provided for the corresponding one or more additional target nucleotides, wherein the ligated DNA fragments are amplified, or circularized DNA is amplified, using the primers.

15

5. Method according to claim 2-4, wherein the fragmenting step b) comprises sonication, followed by enzymatic DNA end repair.

20

6. Method according to claim 2-4, wherein the fragmenting step b) comprises fragmenting with a restriction enzyme.

25

7. Method according to claim 5 or 6, wherein the ligation step c) is performed in the presence of an adaptor, ligating adaptor sequences in between fragments.

30

8. Method according to claim 6 or 7, wherein in step b) a plurality of subsamples is processed, and for each subsample restriction enzymes with different recognition sites are used.

35

9. Method according to claim 8, wherein the fragmenting step e) comprises a restriction enzyme with a recognition sequence which is longer than the recognition sequence of the restriction enzyme of step b).

35

10. Method according to any of the previous claims, wherein the sequences of multiple genomic regions of interests are determined.

11. Method according to any of the previous claims, wherein an identifier is included in at least one of the oligonucleotide primers of step g).
- 5 12. Method according to any of the previous claims, wherein prior to or after the amplification step g) a size selection step is performed.
13. Method according to claim 12, wherein the size selection step is performed using gel extraction chromatography, gel electrophoresis or density gradient centrifugation.
- 10 14. Method according to claim 12-13, wherein DNA is selected of a size between 20-20,000 base pairs, preferably 50-10,000 base pairs, most preferably between 100-3,000 base pairs.
- 15 15. Method according to claim 1-14, wherein in case the ploidy in a cell of a genomic region of interest is greater than 1, for each ploidy a contig is built in step i).
16. Method according to any of the preceding claims, wherein the step i) of building a contig comprises the steps of:
- 20 1) identifying the fragments of step b);
- 2) assigning the fragments to a genomic region;
- 3) building a contig for the genomic region.
17. Method according to claim 16, wherein the step 2) of assigning the fragments to a genomic region comprises identifying the different ligation products of step f) and coupling of the different ligation products to the identified fragments.
- 25 18. Method for identifying the presence or absence of a genetic mutation, comprising the steps a)-i) of any of the claims 1-17, wherein contigs are built for a plurality of samples, comprising the further steps of:
- 30 j) aligning the contigs of a plurality of samples;
- k) identifying the presence or absence of a genetic mutation in the genomic regions of interest from the plurality of samples.
- 35 19. Method for identifying the presence or absence of a genetic mutation, comprising the steps a)-i) of any of the claims 1-17, comprising the further steps of:
- j) aligning the contig to a reference sequence.

k) identifying the presence or absence of a genetic mutation in the genomic region of interest.

20. Method for identifying the presence or absence of a genetic mutation, comprising the steps a)-h) of any of claims 1-14, comprising the further steps of:

i) aligning the determined sequences of the (amplified) ligated DNA fragments to a reference sequence.

j) identifying the presence or absence of a genetic mutation in the determined sequences.

10

21. Method for identifying the presence or absence of a genetic mutation, comprising the steps a)-h) of any of claims 1-14, wherein of a plurality of samples sequences of (amplified) ligated DNA fragments are determined comprising the further steps of:

15

i) aligning the determined sequences of the (amplified) ligated DNA fragments of a plurality of samples.

j) identifying the presence or absence of a genetic mutation in the determined sequences.

20

22. Method for determining the ratio of fragments carrying a genetic mutation from a cell population suspected of being heterologous comprising the steps a)-h) of any of claims 1-14, comprising the further steps of:

i) identifying the fragments of step b);

j) identifying the presence or absence of a genetic mutation in the fragments;

k) determine the number of fragments carrying the genetic mutation;

25

l) determine the number of fragments not carrying the genetic mutation;

m) calculating the ratio of fragments carrying the genetic mutation.

30

23. Method for determining the ratio of ligation products carrying a fragment with a genetic mutation from a cell population suspected of being heterologous comprising the steps a)-h) of any of claims 1-14, comprising the further steps of:

i) identifying the fragments of step b);

j) identifying the presence or absence of a genetic mutation in the fragments;

k) identifying the ligated DNA fragments carrying the fragments with or without the genetic mutation;

35

l) determine the number of ligated DNA fragments carrying the fragments with the genetic mutation;

- m) determine the number of ligation products carrying the fragments without the genetic mutation;
- n) calculating the ratio of ligation products carrying the genetic mutation.
- 5 24. Method according to claims 22 - 23, wherein in step j) the presence or absence of a genetic mutation is identified by aligning to a reference sequence and/or by comparing fragment sequences of a plurality of samples.
- 10 25. Method according to any of claims 18-24, wherein a genetic mutation is a SNP, a deletion, an insertion, an inversion and/or a translocation.
- 15 26. Method according to claim 25, wherein a deletion and/or insertion is identified by comparing the number of fragments and/or ligation products from a sample carrying the deletion and/or insertion with a reference sample.
- 20 27. Method according to claim 25, wherein a deletion, insertion, inversion and/or translocation is identified based on the presence of chromosomal breakpoints in analyzed fragments.
28. Method according to any of the previous claims, wherein the presence or absence of methylated nucleotides is determined in DNA fragments, ligated DNA fragments, and/or genomic regions of interest.

Claims

1. Method for building a contig of a genomic region of interest comprising a target nucleotide sequence comprising, fragmenting a crosslinked DNA, ligating the fragmented cross linked DNA, reversing the crosslinking and determining at least part of the sequences of ligated DNA fragments which comprise a target nucleotide sequence, and using the determined sequences to build a contig of the genomic region of interest.
2. Method for building a contig of a genomic region of interest comprising a target nucleotide sequence, comprising the steps of:
 - a) providing a sample of cross-linked DNA;
 - b) fragmenting the crosslinked DNA;
 - c) ligating the fragmented crosslinked DNA;
 - d) reversing the crosslinking;
 - e) optionally fragmenting the DNA of step d), preferably with a restriction enzyme;
 - f) optionally, ligating the fragmented DNA of step d) or e) to at least one Adaptor;
 - g) optionally and preferably, amplifying the ligated DNA fragments of step d) or e) comprising the target nucleotide sequence using at least one primer which hybridises to the target nucleotide sequence, or amplifying the ligated DNA fragments of step f) using at least one primer which hybridises to the target nucleotide sequence and at least one primer which hybridises to the at least one adaptor;
 - h) determining at least part of the sequence of the (amplified) ligated DNA fragments of step d), e), f) or g) comprising the target nucleotide sequence preferably using high throughput sequencing;
 - i) building a contig of the genomic region of interest from the determined sequences.
3. Method for building a contig of a genomic region of interest comprising a target nucleotide sequence, comprising the steps of:
 - a) providing a sample of cross-linked DNA;
 - b) fragmenting the crosslinked DNA;
 - c) ligating the fragmented crosslinked DNA;
 - d) reversing the crosslinking;

- e) optionally fragmenting the DNA of step d), preferably with a restriction enzyme
 - f) circularising the DNA of step d) or e);
 - g) optionally and preferably, amplifying the circularised DNA comprising the target nucleotide sequence using preferably at least one primer which hybridises to the target nucleotide sequence;
 - h) determining at least part of the sequence of the (amplified) ligated DNA fragments comprising the target nucleotide using high throughput sequencing;
 - i) building a contig of the genomic region of interest from the determined sequences.
4. Method for determining the sequence of a genomic region of interest according to claims 2 or 3, wherein the genomic region of interest comprises one or more target nucleotide sequences in addition, and wherein in the amplification step (g) a primer is provided that hybridises with the target nucleotide sequence and one or more primers are provided for the corresponding one or more additional target nucleotides, wherein the ligated DNA fragments are amplified, or circularized DNA is amplified, using the primers.
5. Method according to claim 2-4, wherein the fragmenting step b) comprises sonication, followed by enzymatic DNA end repair.
6. Method according to claim 2-4, wherein the fragmenting step b) comprises fragmenting with a restriction enzyme.
7. Method according to claim 5 or 6, wherein the ligation step c) is performed in the presence of an adaptor, ligating adaptor sequences in between fragments.
8. Method according to claim 6 or 7, wherein in step b) a plurality of subsamples is processed, and for each subsample restriction enzymes with different recognition sites are used.
9. Method according to any one of claims 6-8, wherein the fragmenting step e) comprises a restriction enzyme with a recognition sequence which is longer than the recognition sequence of the restriction enzyme of step b).
10. Method according to any of the previous claims, wherein the sequences of multiple genomic regions of interests are determined.

11. Method according to any one of claims 2 to 10, wherein an identifier is included in at least one of the oligonucleotide primers of step g).
12. Method according to any of the previous claims, wherein prior to or after the amplification step g) a size selection step is performed.
13. Method according to claim 12, wherein the size selection step is performed using gel extraction chromatography, gel electrophoresis or density gradient centrifugation.
14. Method according to claim 12-13, wherein DNA is selected of a size between 20-20,000 base pairs, preferably 50-10,000 base pairs, most preferably between 100-3,000 base pairs.
15. Method according to any one of claims 1-4, wherein (amplified) ligated DNA fragments comprising the target nucleotide sequence are captured with a capture probe to separate (amplified) ligated DNA fragments comprising the target nucleotide sequence from (amplified) ligated DNA fragments not comprising the target nucleotide sequence.
16. Method according to claim 1 -15, wherein in case the ploidy in a cell of a genomic region of interest is greater than 1 , for each ploidy a contig is built in step i).
17. Method according to any of claims 2-16, wherein the step i) of building a contig comprises the steps of:
 - 1) identifying the fragments of step b);
 - 2) assigning the fragments to a genomic region;
 - 3) building a contig for the genomic region.
18. Method according to claim 17, wherein the step 2) of assigning the fragments to a 25 genomic region comprises identifying the different ligation products of step f) and coupling of the different ligation products to the identified fragments.
19. Method for identifying the presence or absence of a genetic mutation, comprising the steps a)-i) of any of the claims 2-18, wherein contigs are built for a plurality of samples, comprising the further steps of:
 - j) aligning the contigs of a plurality of samples;

- k) identifying the presence or absence of a genetic mutation in the genomic regions of interest from the plurality of samples.
20. Method for identifying the presence or absence of a genetic mutation, comprising the steps a)-i) of any of the claims 2-19, comprising the further steps of:
- j) aligning the contig to a reference sequence.
 - k) identifying the presence or absence of a genetic mutation in the genomic region of interest.
21. Method according to any of claims 19-20, wherein a genetic mutation is a SNP, a deletion, an insertion, an inversion and/or a translocation.
22. Method according to claim 21, wherein a deletion and/or insertion is identified by comparing the number of fragments and/or ligation products from a sample carrying the deletion and/or insertion with a reference sample.
23. Method according to claim 21, wherein a deletion, insertion, inversion and/or translocation is identified based on the presence of chromosomal breakpoints in analyzed fragments.
24. Method according to any of the previous claims, wherein the presence or absence of methylated nucleotides is determined in DNA fragments, ligated DNA fragments, and/or genomic regions of interest.