

PRIMERS AND ASSAYS FOR LINKING REGIONS USING POLYMERASES

BACKGROUND

[0001] Humans have two copies of each chromosome, one from each of the parents. Understanding of the combination of variants on the same parental chromosome, i.e., the haplotype, can provide valuable clinical implications. In particular, the haplotype information is useful for noninvasive prenatal testing of monogenic diseases and deciphering the genomic inheritance of the fetus (Hui et al. Clin Chem. 2017;63:513-524; Lo et al. Sci Transl Med. 2010;2:61ra91). However, current techniques for determining haplotypes of a particular individual can be costly, have low accuracy, and provide low throughput.

BRIEF SUMMARY

[0002] PCR assays using high throughput polymerases can be used to link distant regions (e.g., separated by 100 bp to 100 kbp) of a same large DNA molecule into a smaller DNA molecule, e.g., so a haplotype of the two regions can more easily be measured. The high throughput polymerases can add an extra nucleotide to only one end of each strand of a double stranded DNA (e.g., an A at the 3' end). This may preclude or cause a low yield for generating the smaller, linked DNA molecules (also referred to as extended amplicons).

[0003] To address these problems, particular forward and reverse primers can be used. For example, a reverse primer R1 can have a first portion complementary to an ending sequence of region A and can have a second portion having an overlapping sequence. A forward primer F2 can have a first portion complementary to a starting sequence of region B, where the forward primer includes a complementary overlapping sequence (e.g., the same first portion or a second portion) that is complementary to the overlapping sequence. The first portion of F2 may be the entire primer.

[0003a] In one aspect, provided herein is a method for connecting separated regions of DNA molecules in a biological sample of a subject, the method comprising: partitioning a plurality of DNA molecules from the biological sample into a plurality of compartments, wherein a first compartment of the plurality of compartments includes: a first DNA molecule having a region A

and a region B separated by an intermediate region, a polymerase having a bias for adding a particular nucleotide at an overhang position, a reverse primer R1 having a first portion complementary to an ending sequence of the region A, the reverse primer R1 including a second portion having an overlapping sequence, and a forward primer F2 having a first portion complementary to a starting sequence of the region B, the forward primer F2 including a complementary overlapping sequence that is complementary to the overlapping sequence; amplifying, using the polymerase and the reverse primer R1, the region A to obtain a first set of amplicons having the particular nucleotide at a first 3' end and having a complementary nucleotide between sequence strings corresponding to the first portion and the second portion of the reverse primer R1, the complementary nucleotide being complementary to the particular nucleotide; amplifying, using the polymerase and the forward primer F2, the region B to obtain a second set of amplicons having the particular nucleotide at a second 3' end and having the complementary nucleotide on a 5' end of the overlapping sequence; and extending, using the polymerase, the first set of amplicons to include the region B and the second set of amplicons to include the region A, thereby obtaining extended amplicons that include the region A and the region B and exclude the intermediate region.

[0003b] In another aspect, provided herein is a method for connecting separated regions of DNA molecules in a biological sample of a subject, the method comprising: partitioning a plurality of DNA molecules from the biological sample into a plurality of compartments, wherein a first compartment of the plurality of compartments includes: a first DNA molecule having a region A and a region B separated by an intermediate region, a polymerase, a reverse primer R1 having a first portion complementary to an ending sequence of the region A, the reverse primer R1 including a second portion having an overlapping sequence, a forward primer F2 having a first portion complementary to a starting sequence of the region B, the forward primer F2 including a complementary overlapping sequence that is complementary to the overlapping sequence, a forward primer F1 having a first portion complementary to a starting sequence of the region A and having a second portion, a reverse primer R2 having a first portion complementary to an ending sequence of the region B and having a second portion, a forward primer Fz having a first portion identical to the second portion of the forward primer F1 and having a second portion, and a reverse primer Rz having a first portion identical to the second portion of the reverse primer R2 and having a second portion; amplifying, using the polymerase and the reverse primer R1, the

region A to obtain a first set of amplicons; amplifying, using the polymerase and the forward primer F2, the region B to obtain a second set of amplicons; extending, using the polymerase, the first set of amplicons to include the region B and the second set of amplicons to include the region A, thereby obtaining first extended amplicons that include the region A and the region B and exclude the intermediate region, wherein each first extended amplicon includes the second portion of the forward primer F1 and the second portion of the reverse primer R2; and amplifying, using the polymerase, the forward primer Fz, and the reverse primer Rz, the first extended amplicons to obtain second extended amplicons, wherein each second extended amplicon includes the second portion of the forward primer Fz and the second portion of the reverse primer Rz.

[0003c] In another aspect, provided herein is a method for connecting separated regions of DNA molecules in a biological sample of a subject, the method comprising: partitioning a plurality of DNA molecules from the biological sample into a plurality of compartments, wherein a first compartment of the plurality of compartments includes: a first DNA molecule having a region A and a region B separated by a first intermediate region, a polymerase having a bias for adding a particular nucleotide at an overhang position, a reverse primer R1 having a first portion complementary to an ending sequence of the region A, the reverse primer R1 including a second portion having a first overlapping sequence, a forward primer F2 having a first portion complementary to a starting sequence of the region B, the forward primer F2 including a first complementary overlapping sequence that is complementary to the first overlapping sequence, a reverse primer R2 having a first portion complementary to an ending sequence of the region B, the reverse primer R2 including a second portion having a second overlapping sequence, and a forward primer F3 having a first portion complementary to a starting sequence of a region C of the first DNA molecule, the region C separated from the region B by a second intermediate region, the forward primer F3 including a second complementary overlapping sequence that is complementary to the second overlapping sequence; amplifying, using the polymerase and the reverse primer R1, the region A to obtain a first set of amplicons; amplifying the region B to obtain a second set of amplicons using the polymerase, the forward primer F2, and the reverse primer R2; amplifying, using the polymerase and the forward primer F3, region C to obtain a third set of amplicons; extending, using the polymerase: the first set of amplicons to include the region B and the region C, the third set of amplicons to include the region A and the region B,

thereby obtaining extended amplicons that include the region A, the region B, and the region C, and exclude the first intermediate region and the second intermediate region.

[0003d] In another aspect, provided herein is a method for connecting separated regions of DNA molecules in a biological sample of a subject, the method comprising: partitioning a plurality of DNA molecules from the biological sample into a plurality of compartments, wherein a first compartment of the plurality of compartments includes: a first DNA molecule having a region A and a region B separated by a first intermediate region, a polymerase, a reverse primer R1 having a first portion complementary to an ending sequence of the region A, the reverse primer R1 including a second portion having an overlapping sequence, a forward primer F2 having a first portion complementary to a starting sequence of the region B, the forward primer F2 including a second portion having a complementary overlapping sequence that is complementary to the overlapping sequence, and a forward primer F4 having a first portion complementary to a starting sequence of a region D of the first DNA molecule, the region D separated from the region A by a second intermediate region, the forward primer F4 including a second portion having the complementary overlapping sequence; amplifying, using the polymerase and the reverse primer R1, the region A to obtain a first set of amplicons; amplifying, using the polymerase and the forward primer F2, the region B to obtain a second set of amplicons; amplifying, using the polymerase and the forward primer F4, the region D to obtain a third set of amplicons; extending, using the polymerase, the first set of amplicons to include the region B and the second set of amplicons to include the region A, thereby obtaining first extended amplicons that include the region A and the region B and exclude the first intermediate region; and extending, using the polymerase, the first set of amplicons to include the region D and the third set of amplicons to include the region A, thereby obtaining second extended amplicons that include the region A and the region D and exclude the second intermediate region.

[0003e] In another aspect, provided herein is a kit for connecting separated regions of DNA molecules in a biological sample of a subject, the kit comprising: a polymerase having a bias for adding a particular nucleotide at an overhang position; a reverse primer R1 having a first portion complementary to an ending sequence of a region A of a first DNA molecule, the reverse primer R1 including a second portion having an overlapping sequence, wherein the reverse primer R1 includes the particular nucleotide between the first portion and the second portion; and a forward

primer F2 having a first portion complementary to a starting sequence of a region B of the first DNA molecule, the region B separated from the region A by an intermediate region, the forward primer F2 including a complementary overlapping sequence that is complementary to the overlapping sequence.

[0003f] In another aspect, provided herein is a kit for connecting separated regions of DNA molecules in a biological sample of a subject, the kit comprising: a polymerase; a reverse primer R1 having a first portion complementary to an ending sequence of a region A of a first DNA molecule, the reverse primer R1 including a second portion having an overlapping sequence; a forward primer F2 having a first portion complementary to a starting sequence of a region B of the first DNA molecule, the region B separated from the region A by an intermediate region, the forward primer F2 including a complementary overlapping sequence that is complementary to the overlapping sequence; a forward primer F1 having a first portion complementary to a starting sequence of the region A and having a second portion, a reverse primer R2 having a first portion complementary to an ending sequence of the region B and having a second portion, a forward primer Fz having a first portion identical to the second portion of the forward primer F1 and having a second portion, and a reverse primer Rz having a first portion identical to the second portion of the reverse primer R2 and having a second portion.

[0003g] In another aspect, provided herein is a kit for connecting separated regions of DNA molecules in a biological sample of a subject, the kit comprising: a polymerase; a reverse primer R1 having a first portion complementary to an ending sequence of a region A of a first DNA molecule, the reverse primer R1 including a second portion having an overlapping sequence; a forward primer F2 having a first portion complementary to a starting sequence of a region B of the first DNA molecule, the region B separated from the region A by an intermediate region, the forward primer F2 including a second portion having a complementary overlapping sequence that is complementary to the overlapping sequence; and a forward primer F4 having a first portion complementary to a starting sequence of a region D of the first DNA molecule, the region D separated from the region A by a second intermediate region, the forward primer F4 including a second portion having the complementary overlapping sequence.

[0004] A better understanding of the nature and advantages of embodiments of the present disclosure may be gained with reference to the following detailed description and the accompanying drawings.

WHAT IS CLAIMED IS:

- 1 1. A method for connecting separated regions of DNA molecules in a
2 biological sample of a subject, the method comprising:
3 partitioning a plurality of DNA molecules from the biological sample into a
4 plurality of compartments, wherein a first compartment of the plurality of compartments
5 includes:
6 a first DNA molecule having a region A and a region B separated by an
7 intermediate region,
8 a polymerase having a bias for adding a particular nucleotide at an overhang
9 position,
10 a reverse primer R1 having a first portion complementary to an ending sequence
11 of the region A, the reverse primer R1 including a second portion having an overlapping
12 sequence, and
13 a forward primer F2 having a first portion complementary to a starting sequence
14 of the region B, the forward primer F2 including a complementary overlapping sequence that is
15 complementary to the overlapping sequence;
16 amplifying, using the polymerase and the reverse primer R1, the region A to
17 obtain a first set of amplicons having the particular nucleotide at a first 3' end and having a
18 complementary nucleotide between sequence strings corresponding to the first portion and the
19 second portion of the reverse primer R1, the complementary nucleotide being complementary to
20 the particular nucleotide;
21 amplifying, using the polymerase and the forward primer F2, the region B to
22 obtain a second set of amplicons having the particular nucleotide at a second 3' end and having
23 the complementary nucleotide on a 5' end of the overlapping sequence; and
24 extending, using the polymerase, the first set of amplicons to include the region B
25 and the second set of amplicons to include the region A, thereby obtaining extended amplicons
26 that include the region A and the region B and exclude the intermediate region.
- 1 2. The method of claim 1, wherein the complementary nucleotide in the first
2 set of amplicons is in the first DNA molecule.

1 3. The method of claim 1, wherein the reverse primer R1 includes the
2 particular nucleotide between the first portion and the second portion.

1 4. The method of claim 1, wherein the complementary nucleotide is in the
2 region A of the first set of amplicons.

3 5. The method of claim 1, wherein the compartments are droplets.

1 6. The method of claim 1, further comprising:
2 using a forward primer F1 and a reverse primer R2 to amplify the extended
3 amplicons.

1 7. The method of claim 1, further comprising:
2 detecting a first allele in the region A and a second allele in the region B, thereby
3 detecting that the first allele and the second allele are on a same haplotype.

1 8. The method of claim 7, wherein detecting the first allele in the region A
2 comprises:
3 measuring an amount of the first allele in the region A, and
4 comparing the amount to a cutoff value.

1 9. The method of claim 1, wherein the biological sample is diluted such that
2 on average only one DNA molecule is in a compartment.

1 10. The method of claim 1, further comprising:
2 repeating the method for other compartments to determine the haplotype of other
3 DNA molecules in the other compartments.

1 11. The method of claim 1, wherein:
2 the first compartment further comprises:
3 a forward primer F1 having a first portion complementary to a starting sequence
4 of the region A and having a second portion,
5 a reverse primer R2 having a first portion complementary to an ending sequence
6 of the region B and having a second portion,
7 a forward primer Fz having a first portion identical to the second portion of the
8 forward primer F1 and having a second portion, and
9 a reverse primer Rz having a first portion identical to the second portion of the
10 reverse primer R2 and having a second portion,
11 the extended amplicons are first extended amplicons, and

12 each first extended amplicon includes the second portion of the forward primer F1
13 and the second portion of the reverse primer R2,
14 the method further comprising:
15 amplifying the first extended amplicons to obtain second extended amplicons
16 using the polymerase, the forward primer Fz, and the reverse primer Rz, wherein each second
17 extended amplicon includes the second portion of the forward primer Fz and the second portion
18 of the reverse primer Rz.

1 12. The method of claim 11, wherein:
2 the forward primer Fz comprises a complementary nucleotide, the complementary
3 nucleotide being complementary to the particular nucleotide, and
4 the reverse primer Rz comprises the complementary nucleotide.

1 13. The method of claim 1, wherein:
2 the overlapping sequence is a first overlapping sequence,
3 the complementary overlapping sequence is a first complementary overlapping
4 sequence,

5 the intermediate region is a first intermediate region,
6 the extended amplicons are first extended amplicons; and
7 the first compartment further includes:

8 a reverse primer R2 having a first portion complementary to an ending sequence
9 of the region B, the reverse primer R2 including a second portion having a second overlapping
10 sequence, and a forward primer F3 having a first portion complementary to a starting sequence
11 of a region C of the first DNA molecule, the region C separated from the region B by a second
12 intermediate region, the forward primer F3 including a second complementary overlapping
13 sequence that is complementary to the second overlapping sequence, and

14 amplifying the region B to obtain the second set of amplicons comprises using the
15 reverse primer R2,

16 the method further comprising:

17 amplifying, using the polymerase and the forward primer F3, region C to obtain a
18 third set of amplicons having the particular nucleotide at a third 3' end and having the
19 complementary nucleotide on a 5' end of the second overlapping sequence,

20 extending, using the polymerase:

21 the first extended amplicons to include the region C, and
22 the third set of amplicons to include the region A and the region B,
23 thereby obtaining second extended amplicons that include the region A, the region
24 B, and the region C and exclude the first intermediate region and the second intermediate region.

1 14. The method of claim 13, wherein the forward primer F3 includes a second
2 portion having the second complementary overlapping sequence.

1 15. The method of claim 13, wherein the first portion of the forward primer F3
2 includes second complementary overlapping sequence.

1 16. The method of claim 13, wherein:
2 extending the first extended amplicons to include the region C comprises
3 hybridizing to the second overlapping sequence toward a fourth 3' end.

1 17. The method of claim 13, wherein:
2 extending the third set of amplicons to include the region A and the region B
3 comprises hybridizing to the second complementary overlapping sequence toward the second 3'
4 end.

1 18. The method of claim 1, wherein the first portion of the forward primer F2
2 includes the complementary overlapping sequence.

1 19. The method of claim 1, wherein the forward primer F2 includes a second
2 portion having the complementary overlapping sequence.

1 20. The method of claim 19, wherein the forward primer F2 includes the
2 particular nucleotide between the first portion and the second portion.

1 21. The method of claim 19, wherein:
2 the intermediate region is a first intermediate region,
3 the extended amplicons are first extended amplicons, and
4 the first compartment further includes:
5 a forward primer F4 having a first portion complementary to a starting sequence
6 of a region D of the first DNA molecule, the region D separated from the region A by a second
7 intermediate region, the forward primer F4 including a second portion having the complementary
8 overlapping sequence,
9 the method further comprising:

10 amplifying, using the polymerase and the forward primer F4, the region D to
11 obtain a third set of amplicons having the particular nucleotide at a third 3' end and having the
12 complementary nucleotide on a 5' end of the overlapping sequence, and
13 extending, using the polymerase, the first set of amplicons to include the region D
14 and the third set of amplicons to include the region A, thereby obtaining second extended
15 amplicons that include the region A and the region D and exclude the second intermediate
16 region.

1 22. The method of claim 21, wherein:
2 the first compartment further includes:
3 a reverse primer R3 having a first portion complementary to an ending sequence
4 of a region C of the first DNA molecule, the region C separated from the region B by a third
5 intermediate region, the region C separated from the region D by a fourth intermediate region,
6 the reverse primer R3 including a second portion having the overlapping sequence,
7 the method further comprising:
8 amplifying, using the polymerase and the reverse primer R3, the region C to
9 obtain a fourth set of amplicons having the particular nucleotide at a fourth 3' end and having the
10 complementary nucleotide on a 5' end of the overlapping sequence,
11 extending, using the polymerase, the fourth set of amplicons to include the region
12 B and the second set of amplicons to include the region C, thereby obtaining third extended
13 amplicons that include the region C and the region B and exclude the third intermediate region,
14 and
15 extending, using the polymerase, the fourth set of amplicons to include the region
16 D and the third set of amplicons to include the region C, thereby obtaining fourth extended
17 amplicons that include the region C and the region D and exclude the fourth intermediate region.

1 23. The method of claim 21, further comprising:
2 detecting a first allele in the region A in the first extended amplicons,
3 detecting a second allele in the region B in the first extended amplicons,
4 detecting the first allele in the region A in the second extended amplicons, and
5 detecting a third allele in the region D in the second extended amplicons,
6 thereby detecting the first allele in the region A, the second allele in the region B,
7 and the third allele in the region D are on a same haplotype.

1 24. The method of claim 21, wherein:
2 the complementary overlapping sequence is a first complementary overlapping
3 sequence,
4 the overlapping sequence is a first overlapping sequence,
5 the first compartment further includes:
6 a reverse primer R2 having a first portion complementary to an ending sequence
7 of the region B, the reverse primer R2 including a second portion having a second overlapping
8 sequence,
9 a reverse primer R4 having a first portion complementary to an ending sequence
10 of the region D, the reverse primer R4 including a second portion having the second overlapping
11 sequence, and
12 a forward primer F3 having a first portion complementary to a starting sequence
13 of a region C of the first DNA molecule, the region C separated from the region B by a third
14 intermediate region, the forward primer F3 including a second portion having the second
15 complementary overlapping sequence,
16 amplifying the region B comprises using the reverse primer R2, and
17 amplifying the region D comprises using the reverse primer R4,
18 the method further comprising:
19 amplifying, using the polymerase and the forward primer F3, the region C to
20 obtain a fourth set of amplicons,
21 extending, using the polymerase, the second extended amplicons to include the
22 region C, thereby obtaining third extended amplicons that include the region A, the region D, and
23 the region C, and
24 extending, using the polymerase, the first extended amplicons to include the
25 region C, thereby obtaining fourth extended amplicons that include the region A, the region B,
26 and the region C.

1 25. A method for connecting separated regions of DNA molecules in a
2 biological sample of a subject, the method comprising:
3 partitioning a plurality of DNA molecules from the biological sample into a
4 plurality of compartments, wherein a first compartment of the plurality of compartments
5 includes:

6 a first DNA molecule having a region A and a region B separated by an
7 intermediate region,
8 a polymerase,
9 a reverse primer R1 having a first portion complementary to an ending sequence
10 of the region A, the reverse primer R1 including a second portion having an overlapping
11 sequence,
12 a forward primer F2 having a first portion complementary to a starting sequence
13 of the region B, the forward primer F2 including a complementary overlapping sequence that is
14 complementary to the overlapping sequence,
15 a forward primer F1 having a first portion complementary to a starting sequence
16 of the region A and having a second portion,
17 a reverse primer R2 having a first portion complementary to an ending sequence
18 of the region B and having a second portion,
19 a forward primer Fz having a first portion identical to the second portion of the
20 forward primer F1 and having a second portion, and
21 a reverse primer Rz having a first portion identical to the second portion of the
22 reverse primer R2 and having a second portion;
23 amplifying, using the polymerase and the reverse primer R1, the region A to
24 obtain a first set of amplicons;
25 amplifying, using the polymerase and the forward primer F2, the region B to
26 obtain a second set of amplicons;
27 extending, using the polymerase, the first set of amplicons to include the region B
28 and the second set of amplicons to include the region A, thereby obtaining first extended
29 amplicons that include the region A and the region B and exclude the intermediate region,
30 wherein each first extended amplicon includes the second portion of the forward primer F1 and
31 the second portion of the reverse primer R2; and
32 amplifying, using the polymerase, the forward primer Fz, and the reverse primer
33 Rz, the first extended amplicons to obtain second extended amplicons, wherein each second
34 extended amplicon includes the second portion of the forward primer Fz and the second portion
35 of the reverse primer Rz.

1 26. The method of claim 25, wherein the forward primer F2 includes a second
2 portion having the complementary overlapping sequence.

1 27. The method of claim 25, wherein the first portion of the forward primer F2
2 includes the complementary overlapping sequence.

1 28. A method for connecting separated regions of DNA molecules in a
2 biological sample of a subject, the method comprising:

3 partitioning a plurality of DNA molecules from the biological sample into a
4 plurality of compartments, wherein a first compartment of the plurality of compartments
5 includes:

6 a first DNA molecule having a region A and a region B separated by a first
7 intermediate region,

8 a polymerase having a bias for adding a particular nucleotide at an overhang
9 position,

10 a reverse primer R1 having a first portion complementary to an ending sequence
11 of the region A, the reverse primer R1 including a second portion having a first overlapping
12 sequence,

13 a forward primer F2 having a first portion complementary to a starting sequence
14 of the region B, the forward primer F2 including a first complementary overlapping sequence
15 that is complementary to the first overlapping sequence,

16 a reverse primer R2 having a first portion complementary to an ending sequence
17 of the region B, the reverse primer R2 including a second portion having a second overlapping
18 sequence, and

19 a forward primer F3 having a first portion complementary to a starting sequence
20 of a region C of the first DNA molecule, the region C separated from the region B by a second
21 intermediate region, the forward primer F3 including a second complementary overlapping
22 sequence that is complementary to the second overlapping sequence;

23 amplifying, using the polymerase and the reverse primer R1, the region A to
24 obtain a first set of amplicons;

25 amplifying the region B to obtain a second set of amplicons using the polymerase,
26 the forward primer F2, and the reverse primer R2;

27 amplifying, using the polymerase and the forward primer F3, region C to obtain a
28 third set of amplicons;

29 extending, using the polymerase:

30 the first set of amplicons to include the region B and the region C,

31 the third set of amplicons to include the region A and the region B, thereby

32 obtaining extended amplicons that include the region A, the region B, and the region C, and

33 exclude the first intermediate region and the second intermediate region.

1 29. A method for connecting separated regions of DNA molecules in a
2 biological sample of a subject, the method comprising:

3 partitioning a plurality of DNA molecules from the biological sample into a
4 plurality of compartments, wherein a first compartment of the plurality of compartments
5 includes:

6 a first DNA molecule having a region A and a region B separated by a first
7 intermediate region,

8 a polymerase,

9 a reverse primer R1 having a first portion complementary to an ending sequence
10 of the region A, the reverse primer R1 including a second portion having an overlapping
11 sequence,

12 a forward primer F2 having a first portion complementary to a starting sequence
13 of the region B, the forward primer F2 including a second portion having a complementary
14 overlapping sequence that is complementary to the overlapping sequence, and

15 a forward primer F4 having a first portion complementary to a starting sequence
16 of a region D of the first DNA molecule, the region D separated from the region A by a second
17 intermediate region, the forward primer F4 including a second portion having the complementary
18 overlapping sequence;

19 amplifying, using the polymerase and the reverse primer R1, the region A to
20 obtain a first set of amplicons;

21 amplifying, using the polymerase and the forward primer F2, the region B to
22 obtain a second set of amplicons;

23 amplifying, using the polymerase and the forward primer F4, the region D to
24 obtain a third set of amplicons;

extending, using the polymerase, the first set of amplicons to include the region B and the second set of amplicons to include the region A, thereby obtaining first extended amplicons that include the region A and the region B and exclude the first intermediate region; and

extending, using the polymerase, the first set of amplicons to include the region D and the third set of amplicons to include the region A, thereby obtaining second extended amplicons that include the region A and the region D and exclude the second intermediate region.

30. The method of claim 29, further comprising:
detecting a first allele in the region A in the first extended amplicons,
detecting a second allele in the region B in the first extended amplicons,
detecting the first allele in the region A in the second extended amplicons, and
detecting a third allele in the region D in the second extended amplicons,
thereby detecting the first allele in the region A, the second allele in the region B, and the third allele in the region D are on a same haplotype.

31. The method of claim 29, wherein:
the complementary overlapping sequence is a first complementary overlapping sequence,
the overlapping sequence is a first overlapping sequence,
the first compartment further includes:
a reverse primer R2 having a first portion complementary to an ending sequence of the region B, the reverse primer R2 including a second portion having a second overlapping sequence,
a reverse primer R4 having a first portion complementary to an ending sequence of the region D, the reverse primer R2 including a second portion having the second overlapping sequence, and
a forward primer F3 having a first portion complementary to a starting sequence of a region C of the first DNA molecule, the region C separated from the region B by a third intermediate region, the forward primer F3 including a second portion having the second complementary overlapping sequence,
amplifying the region B comprises using the reverse primer R2, and

17 amplifying the region D comprises using the reverse primer R4,
18 the method further comprising:
19 amplifying, using the polymerase and the forward primer F3, the region C to
20 obtain a fourth set of amplicons,
21 extending, using the polymerase, the second extended amplicons to include the
22 region C, thereby obtaining third extended amplicons that include the region A, the region D, and
23 the region C, and
24 extending, using the polymerase, the first extended amplicons to include the
25 region C, thereby obtaining fourth extended amplicons that include the region A, the region B,
26 and the region C.

1 32. A kit for connecting separated regions of DNA molecules in a biological
2 sample of a subject, the kit comprising:
3 a polymerase having a bias for adding a particular nucleotide at an overhang
4 position;
5 a reverse primer R1 having a first portion complementary to an ending sequence
6 of a region A of a first DNA molecule, the reverse primer R1 including a second portion having
7 an overlapping sequence, wherein the reverse primer R1 includes the particular nucleotide
8 between the first portion and the second portion; and
9 a forward primer F2 having a first portion complementary to a starting sequence
10 of a region B of the first DNA molecule, the region B separated from the region A by an
11 intermediate region, the forward primer F2 including a complementary overlapping sequence
12 that is complementary to the overlapping sequence.

1 33. The kit of claim 32, wherein the forward primer F2 includes a second
2 portion having the complementary overlapping sequence.

1 34. The kit of claim 32, wherein the first portion of the forward primer F2
2 includes the complementary overlapping sequence.

1 35. The kit of claim 32, further comprising:
2 a forward primer F1 having a first portion complementary to a starting sequence
3 of the region A, and
4 a reverse primer R2 having a first portion complementary to an ending sequence
5 of the region B.

1 36. The kit of claim 35, wherein:
2 the forward primer F1 includes a second portion that is not the first portion, and
3 the reverse primer R2 includes a second portion that is not the first portion;
4 the kit further comprising:
5 a forward primer Fz having a first portion identical to the second portion of the
6 forward primer F1, and
7 a reverse primer Rz having a first portion identical to the second portion of the
8 reverse primer R2.

1 37. The kit of claim 36, wherein:
2 the forward primer Fz includes a nucleotide complementary to the particular
3 nucleotide, and
4 the reverse primer Rz includes the nucleotide complementary to the particular
5 nucleotide.

1 38. The kit of claim 36, wherein:
2 the ratio of the forward primer Fz to forward primer F1 is at least 5:1, and
3 the ratio of the reverse primer Rz to reverse primer R2 is at least 5:1.

1 39. The kit of claim 32, wherein:
2 the overlapping sequence is a first overlapping sequence,
3 the complementary overlapping sequence is a first complementary overlapping
4 sequence, and
5 the intermediate region is a first intermediate region,
6 the kit further comprising:
7 a reverse primer R2 having a first portion complementary to an ending sequence
8 of the region B, the reverse primer R2 including a second portion having a second overlapping
9 sequence, and
10 a forward primer F3 having a first portion complementary to a starting sequence
11 of a region C of the first DNA molecule, the region C separate from the region B by a second
12 intermediate region, the forward primer F3 including a second complementary overlapping
13 sequence that is complementary to the second overlapping sequence.

1 40. The kit of claim 39, wherein the first overlapping sequence is not the
2 second overlapping sequence.

41. A kit for connecting separated regions of DNA molecules in a biological sample of a subject, the kit comprising:

a polymerase;

a reverse primer R1 having a first portion complementary to an ending sequence of a region A of a first DNA molecule, the reverse primer R1 including a second portion having an overlapping sequence;

a forward primer F2 having a first portion complementary to a starting sequence of a region B of the first DNA molecule, the region B separated from the region A by an intermediate region, the forward primer F2 including a complementary overlapping sequence that is complementary to the overlapping sequence;

a forward primer F1 having a first portion complementary to a starting sequence of the region A and having a second portion,

a reverse primer R2 having a first portion complementary to an ending sequence of the region B and having a second portion,

a forward primer Fz having a first portion identical to the second portion of the forward primer F1 and having a second portion, and

a reverse primer Rz having a first portion identical to the second portion of the reverse primer R2 and having a second portion.

42. A kit for connecting separated regions of DNA molecules in a biological sample of a subject, the kit comprising:

a polymerase;

a reverse primer R1 having a first portion complementary to an ending sequence of a region A of a first DNA molecule, the reverse primer R1 including a second portion having an overlapping sequence;

a forward primer F2 having a first portion complementary to a starting sequence of a region B of the first DNA molecule, the region B separated from the region A by an intermediate region, the forward primer F2 including a second portion having a complementary overlapping sequence that is complementary to the overlapping sequence; and

a forward primer F4 having a first portion complementary to a starting sequence of a region D of the first DNA molecule, the region D separated from the region A by a second

intermediate region, the forward primer F4 including a second portion having the complementary overlapping sequence.