

In additional examples, a hybrid tagged DNA library comprises hybrid tagged DNA sequences for use in targeted genetic analysis, wherein each hybrid tagged DNA sequence comprises: i) fragmented end-repaired DNA; ii) a random nucleotide tag sequence; iii) a sample code sequence; iv) a PCR primer sequence; and v) a multifunctional capture probe module tail sequence.

In further examples, a multifunctional adaptor module comprises: i) a first region comprising a random nucleotide tag sequence; ii) a second region comprising a sample code sequence; and iii) a third region comprising a PCR primer sequence.

In particular examples, a multifunctional capture probe module comprises: i) a first region capable of hybridizing to a partner oligonucleotide; ii) a second region capable of hybridizing to a specific target region; and iii) a third region comprising a tail sequence. In some examples, the first region of the capture probe module is bound to a partner oligonucleotide. In some examples, the partner oligonucleotide is chemically modified.

In one example, compositions comprise a tagged DNA library, a multifunctional adaptor module and a multifunctional capture probe module.

In particular examples, a composition comprises a hybrid tagged genomic library according to the methods of the present disclosure.

In certain examples, a composition comprises a reaction mixture for performing the methods contemplated herein.

In one aspect, the present disclosure refers to a plurality of compositions for generating a captured library comprising hybrid nucleic acid molecules, the plurality of compositions comprising: (a) a first composition comprising a multifunctional adaptor module comprising: (i) a first region comprising a random nucleotide tag sequence, (ii) a second region comprising a sample code sequence, and (iii) a third region comprising a PCR primer sequence; characterized in that the multifunctional adaptor module is capable of ligating to each end of a DNA fragment, and (b) a second composition comprising a multifunctional capture probe module comprising: (i) a first region capable of hybridizing to a partner oligonucleotide, and (ii) a second region capable of hybridizing to a specific target region in the DNA fragment.

In particular examples, a reaction mixture capable of generating a tagged DNA library comprises: a) fragmented DNA and b) DNA end-repair enzymes to generate fragmented end-repaired DNA.

In certain examples, a reaction mixture further comprises a multifunctional adaptor module.

In additional examples, a reaction mixture further comprises a multifunctional capture probe module.

In some examples, a reaction mixture further comprises an enzyme with 3'-5' exonuclease activity and PCR amplification activity.

In one example, the reaction mixture comprises a FLAP endonuclease, a DNA polymerase, and DNA ligase.

In one aspect, the present disclosure refers to a complex comprising: (a) a tagged DNA molecule comprising a DNA fragment, wherein each end of the DNA fragment is ligated to a multifunctional adaptor module comprising: (i) a first region comprising a random nucleotide tag sequence; (ii) a second region comprising a sample code sequence; and (iii) a third region comprising a PCR primer sequence; and (b) a hybrid nucleic acid molecule hybridized to the tagged DNA molecule, wherein the hybrid nucleic acid molecule comprises, from 5' to 3': (i) a tail sequence comprising a PCR primer binding site; (ii) a capture probe to which a portion of the DNA fragment is hybridized; and (iii) the complement of a region of the DNA molecule that is located 3' to the capture probe, wherein the region comprises a multifunctional adaptor module.

In one aspect, the present disclosure refers to a complex comprising: (a) a multifunctional capture probe module comprising, from 3' to 5': (i) a first region capable of hybridizing to a partner oligonucleotide; (ii) a second region capable of hybridizing to a specific target region in a DNA fragment; and (iii) a third region comprising a tail sequence; and (b) a hybrid nucleic acid molecule comprising (i) the DNA fragment and (ii) at its 3' end, the complement of the third region of the multifunctional capture probe module, wherein the DNA fragment is hybridized to the multifunctional capture probe module, and wherein the 5' end of the DNA fragment is ligated to a multifunctional adaptor module comprising: (i) a first region comprising a random nucleotide tag sequence; (ii) a second region comprising a sample code sequence; and (iii) a third region comprising a PCR primer sequence.

In one aspect, the present disclosure refers to a complex comprising: (a) a tagged DNA molecule comprising a DNA fragment, wherein each end of the DNA fragment is ligated to a multifunctional adaptor module; wherein the multifunctional adaptor module comprises: (i) a first region comprising a random nucleotide tag sequence; (ii) a second region comprising a sample code sequence; and (iii) a third region comprising a PCR primer sequence; and (b) a multifunctional capture probe module comprising: (i) a first region capable of hybridizing to a partner oligonucleotide; and (ii) a second region capable of hybridizing to a specific target region in the DNA fragment; wherein the tagged DNA molecule is hybridized to the multifunctional capture probe module.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

Figure 1: Construction of an amplifiable, sample coded, tagged genomic DNA library. Purified genomic DNA was isolated from a source such as whole blood or a buccal cheek swab. The DNA was fragmented (e.g. by mechanical, enzymatic or chemical means) and the ends of the DNA were repaired, in this example, to blunt ends. The repaired DNA was ligated to a multifunctional adaptor module that contained a universal amplification sequence, a random nucleotide tag sequence and a sample code sequence. A specific example of a typical adaptor duplex molecule is shown by way of example.

Figure 2: Genomic capture probe design. (A) Structure of a prototypical 114 nt probe. Region 1 is a 34 nt sequence that shares perfect complementarity to a highly modified, 34 nt partner oligonucleotide. The partner oligonucleotide was modified on its 5' end with a biotin – TEG chemical entity that enabled capture on streptavidin-coated magnetic beads. B stands for the “Bio-TEG” modification. Region 2 is the 60 nt probe region designed to interact with genomic DNA targets. Region 3 is a 20 nt tail that introduced PCR amplification sequences into the captured genomic fragments. (B) An example of the highly modified partner strand that is complementary to region 1 of each individual probe.

Figure 3: Sequence “spreading” in conventional, hybridization-based capture experiments. (A) Flanking fragments can “hitch-hike” into capture libraries by way of legitimate hybridization triplexes involving probe-fragment-flanking fragment interactions. (B) The net result of cross-fragment hybridization is sequence “spreading”, meaning sequences that are within proximity to the target region (dashed lines) but fall outside the boundaries of the desired target.

Figure 4: Enzymatic processing of fragment: probe hybridized complexes. (A) The purified complex of fragment (light gray) and probe (black) prior to processing. B – biotin affinity modification. (B) DNA polymerase (e.g. T4 DNA polymerase) encoded 3' → 5' exonuclease activity removes the 3' segment of the captured fragment. (C) Upon encountering the probe: fragment duplex region, the polymerase copies the probe tail segment onto the hybridized genomic fragment. (D) The final modified

CLAIMS

1. A plurality of compositions for generating a captured library comprising hybrid nucleic acid molecules, the plurality of compositions comprising:

(a) a first composition comprising a multifunctional adaptor module comprising:

(i) a first region comprising a random nucleotide tag sequence,

(ii) a second region comprising a sample code sequence, and

(iii) a third region comprising a PCR primer sequence;

characterized in that the multifunctional adaptor module is capable of ligating to each end of a DNA fragment, and

(b) a second composition comprising a multifunctional capture probe module comprising:

(i) a first region capable of hybridizing to a partner oligonucleotide, and

(ii) a second region capable of hybridizing to a specific target region in the DNA fragment.

2. The plurality of compositions of claim 1, wherein the first composition comprises a plurality of multifunctional adaptor modules.

3. The plurality of compositions of claim 2, wherein each multifunctional adaptor module of the plurality of multifunctional adaptor modules comprises a random nucleotide tag sequence, wherein the multifunctional adaptor module and a nucleic acid sequence of the DNA fragment ligated thereto is capable of being sequenced to obtain one or more sequencing reads, and wherein a quantity of random nucleotide tag sequences determined in the sequencing reads indicates the copy number of the nucleic acid sequence.

4. The plurality of compositions of claim 2 or claim 3, wherein the PCR primer sequence of each multifunctional adaptor module of the plurality of multifunctional adaptor modules is the same.

5. The plurality of compositions of any one of claims 1-4, wherein the sample code sequence is capable of identifying a DNA sample from which the captured library is generated.

6. The plurality of compositions of any one of claims 1-5, wherein the multifunctional adaptor module comprises a ligation strand and a complementary strand, wherein the complementary strand is hybridized to the 3' end of the ligation strand and forms a partial duplex therewith.

7. The plurality of compositions of claim 6, wherein the partial duplex of the multifunctional adaptor module is blunt ended.

8. The plurality of compositions of claim 6, wherein the partial duplex of the multifunctional adaptor module comprises an overhang at the 3' end of the ligation strand, wherein the overhang is complementary to a 3'-overhang of a DNA fragment to be ligated thereto.

9. The plurality of compositions of any one of claims 1-8, wherein the partner oligonucleotide comprises a specific member of a binding pair that is configured to facilitate isolation and/or purification of a complex comprising the multifunctional capture probe module hybridized to the DNA fragment, wherein the specific member of a binding pair is a biotin or a hapten.

10. The plurality of compositions of claim 9, wherein the hapten is a dinitrophenol molecule, or a digoxigenin molecule.

11. The plurality of compositions of any one of claims 1-10, wherein the first region of the multifunctional capture probe module comprises a tail sequence at the 5' end.

12. The plurality of compositions of any one of claims 1-10, wherein the multifunctional capture probe module comprises a third region comprising a tail sequence at the 5' end.

13. The plurality of compositions of claim 11 or claim 12, wherein the tail sequence comprises a PCR primer binding site.

14. A complex comprising:

(a) a tagged DNA molecule comprising a DNA fragment, wherein each end of the DNA fragment is ligated to a multifunctional adaptor module comprising:

(i) a first region comprising a random nucleotide tag sequence;

(ii) a second region comprising a sample code sequence; and

(iii) a third region comprising a PCR primer sequence; and

(b) a hybrid nucleic acid molecule hybridized to the tagged DNA molecule, wherein the hybrid nucleic acid molecule comprises, from 5' to 3':

(i) a tail sequence comprising a PCR primer binding site;

(ii) a capture probe to which a portion of the DNA fragment is hybridized; and

(iii) the complement of a region of the DNA molecule that is located 3' to the capture probe, wherein the region comprises a multifunctional adaptor module.

15. The complex of claim 14, wherein the random nucleotide tag sequence is configured to designate the DNA fragment ligated thereto.

16. The complex of claim 14 or claim 15, wherein the DNA fragment is derived or isolated from a DNA sample, wherein the sample code sequence is capable of identifying the DNA sample.

17. The complex of any one of claims 14-16, further comprising:

(c) a partner oligonucleotide hybridized to the 5' end of the hybrid nucleic acid molecule.

18. The complex of claim 17, wherein the partner oligonucleotide comprises a specific member of a binding pair that is configured to facilitate isolation and/or purification of the complex, wherein the specific member of a binding pair is a biotin or a hapten.

19. The complex of claim 18, wherein the hapten is a dinitrophenol molecule, or a digoxigenin molecule.

20. A complex comprising:

- (a) a multifunctional capture probe module comprising, from 3' to 5':
 - (i) a first region capable of hybridizing to a partner oligonucleotide;
 - (ii) a second region capable of hybridizing to a specific target region in a DNA fragment; and
 - (iii) a third region comprising a tail sequence; and
- (b) a hybrid nucleic acid molecule comprising (i) the DNA fragment and (ii) at its 3' end, the complement of the third region of the multifunctional capture probe module, wherein the DNA fragment is hybridized to the multifunctional capture probe module, and wherein the 5' end of the DNA fragment is ligated to a multifunctional adaptor module comprising:
 - (i) a first region comprising a random nucleotide tag sequence;
 - (ii) a second region comprising a sample code sequence; and
 - (iii) a third region comprising a PCR primer sequence.

21. The complex of claim 20, further comprising:

- (c) a partner oligonucleotide hybridized to the first region of the multifunctional capture probe module.

22. The complex of claim 20 or claim 21, wherein the random nucleotide tag sequence is configured to designate the DNA fragment ligated thereto.

23. The complex of any one of claims 20-22, wherein the DNA fragment is derived or isolated from a DNA sample, wherein the sample code sequence is capable of identifying the DNA sample.

24. The complex of any one of claims 20-23, wherein the partner oligonucleotide comprises a specific member of a binding pair that is configured to facilitate isolation and/or

purification of the complex, wherein the specific member of a binding pair is a biotin or a hapten.

25. The complex of claim 24, wherein the hapten is a dinitrophenol molecule, or a digoxigenin molecule.

26. A complex comprising:

(a) a tagged DNA molecule comprising a DNA fragment, wherein each end of the DNA fragment is ligated to a multifunctional adaptor module;

wherein the multifunctional adaptor module comprises:

(i) a first region comprising a random nucleotide tag sequence;

(ii) a second region comprising a sample code sequence; and

(iii) a third region comprising a PCR primer sequence; and

(b) a multifunctional capture probe module comprising:

(i) a first region capable of hybridizing to a partner oligonucleotide; and

(ii) a second region capable of hybridizing to a specific target region in the DNA fragment;

wherein the tagged DNA molecule is hybridized to the multifunctional capture probe module.

27. The complex of claim 26, wherein the random nucleotide tag sequence is configured to designate the DNA fragment ligated thereto.

28. The complex of claim 26 or claim 27, wherein the sample code sequence is capable of identifying a DNA sample from which the tagged DNA molecule is derived or isolated.

29. The complex of any one of claims 26-28, wherein the partner oligonucleotide comprises a specific member of a binding pair that is configured to facilitate isolation and/or purification of the complex, wherein the specific member of a binding pair is a biotin or a hapten.

30. The complex of claim 29, wherein the hapten is a dinitrophenol molecule, or a digoxigenin molecule.

31. The complex of any one of claims 26-30, wherein the first region of the multifunctional capture probe module comprises a tail sequence.

32. The complex of any one of claims 26-30, wherein the multifunctional capture probe module comprises a third region comprising a tail sequence.

33. The complex of claim 31 or claim 32, wherein the tail sequence comprises a PCR primer binding site.