

sequence that is not complementary to the universal 3' adaptor sequence and that comprises an R1 sequencing primer site, and (ii) a sequence complementary to the annealing site for the 5' adaptor; c) extending the 5' adaptors across the exogenous UID sequence and the first segment, thereby generating the complement of the exogenous UID sequence and complement of the first segment, and d) covalently linking the 3' end of the complement of the first segment to the 5' ends of the Watson and Crick strands of the double-stranded DNA fragments, thereby generating a plurality of adaptor-ligated double-stranded DNA fragments.

In one aspect, the disclosure provides a method for assessing nucleic acids, comprising: a) attaching partially double-stranded 3' adaptors (3'PDSAs) to 3' ends of both Watson and Crick strands of a population of double-stranded DNA fragments in an analyte DNA sample, wherein a first strand of each of the 3'PDSAs comprises, in the 5'-3' direction, (i) a first segment, (ii) a single stranded unique identifier (ssUID) sequence, (iii) an annealing site for a universal 5' adaptor, and (iv) a universal 3' adaptor sequence, and wherein the second strand of the 3'PDSAs comprises, in the 5' to 3' direction, (i) a segment complementary to the first segment, and (ii) a 3' blocking group, b) annealing universal 5' adaptors to said annealing site, wherein the universal 5' adaptors comprise, in the 5' to 3' direction, (i) a universal 5' adaptor sequence that is not complementary to the universal 3' adaptor sequence, and (ii) a sequence complementary to the annealing site for the 5' adaptor; c) extending the universal 5' adaptors across the ssUID sequence and said first segment, thereby generating the complement of said ssUID sequence and complement of said first segment; d)

covalently linking the 3' end of said complement of said first segment to the 5' ends of the Watson and Crick strands of the double-stranded DNA fragments, thereby generating a plurality of adapter-ligated double-stranded DNA fragments labelled with a double-stranded unique identifier (dsUID) sequence on each end; e) amplifying the dsUID-labelled double-stranded DNA fragments, thereby generating a duplex sequencing library sample; f) amplifying a first aliquot from the duplex sequencing library sample with a first Watson target-selective primer pair, thereby generating dsUID-labelled double-stranded target Watson amplification products (dsUID-dsTWAPs) and not generating dsUID-labelled double-stranded target Crick amplification products (dsUID-dsTCAPs), g) performing hemi-nested amplification of the dsUID-dsTWAPs, thereby creating a first sample comprising a plurality of dsUID-labelled double-stranded target Watson library members (dsUID-dsTWLMs); h) amplifying a second aliquot from the duplex sequencing library sample with a first Crick target-selective primer pair, thereby generating dsUID-dsTCAPs and not generating dsUID-dsTWAPs; i) performing hemi-nested amplification of the dsUID-dsTCAPs, thereby creating a second sample comprising a plurality of dsUID-labelled double-stranded target Crick family members (dsUID-dsTCLMs); and j) sequencing the dsUID-dsTWLMs and the dsUID-dsTCLMs, thereby generating a plurality of UID-labelled Watson strand reads and a plurality of UID-labelled Crick strand reads.

In some examples, provided herein are methods comprising: a) attaching partially double-stranded 3' adaptors (3'PDSAs) to 3' ends of both Watson and Crick strands of a population of double-stranded DNA fragments in an analyte DNA sample, wherein a first strand of the 3'PDSAs comprises, in the 5'-3' direction, (i) a first segment, (ii) an exogenous UID sequence, (iii) an annealing site for a 5' adaptor, and (iv) a universal 3' adaptor sequence comprising an R2 sequencing primer site, and wherein the second strand of the 3'PDSAs comprises, in the 5' to 3' direction, (i) a segment complementary to the first segment, and optionally (ii) a 3' blocking group, b) annealing 5' adaptors to the annealing site, wherein the 5' adaptors comprise, in the 5' to 3' direction, (i) a universal 5' adaptor sequence that is not complementary to the universal 3' adaptor sequence and that comprises an R1 sequencing primer site, and (ii) a sequence complementary to the annealing site for the 5' adaptor; c) extending the 5' adaptors across the exogenous UID sequence, thereby generating the complement of the exogenous UID sequence, and d) covalently linking the 3' end of the complement of the exogenous UID sequence to the 5' ends of the segment complementary to the first segment on each of the Watson and Crick strands of the double-stranded DNA fragments, thereby generating a plurality of adaptor-ligated double-stranded DNA fragments.

In some examples, provided herein are methods comprising: a) attaching partially double-stranded 3' adaptors (3'PDSAs) to 3' ends of both Watson and Crick strands of a population of double-stranded DNA fragments in an analyte DNA sample, wherein a first strand of the 3'PDSAs comprises, in the 5'-3' direction, (i) a first segment, (ii) an exogenous UID sequence, (iii) an annealing site for a 5' adaptor, and (iv) a universal 3' adaptor sequence comprising an R2 sequencing primer site, and wherein the second strand of the 3'PDSAs comprises, in the 5' to 3' direction, (i) a segment complementary to the first segment, and optionally (ii) a 3' blocking group, b) annealing 5' adaptors to the annealing site, wherein the 5' adaptors comprise, in the 5' to 3' direction, (i) a universal 5' adaptor

sequence that is not complementary to the universal 3' adaptor sequence and that comprises an R1 sequencing primer site, and (ii) a sequence complementary to the annealing site for the 5' adaptor; c) extending the 5' adaptors across the exogenous UID sequence and the first segment of the 3'PDSA, thereby generating the complement of the exogenous UID sequence and the complement of the first segment of the 3'PDSA, and d) covalently linking the 3' end of the complement of the first segment of the 3'PDSA to the 5' ends of the Watson and Crick strands of the double-stranded DNA fragments, thereby generating a plurality of adaptor-ligated double-stranded DNA fragments.

In some examples, provided here are systems and kits and compositions comprising: a) a population of partially double-stranded 3' adaptors (3'PDSAs) configured to be ligated to the 3' ends of both Watson and Crick strands of a population of double-stranded DNA fragments, wherein a first strand of the 3'PDSAs comprises, in the 5'-3' direction, (i) a first segment, (ii) an exogenous UID sequence, (iii) an annealing site for a 5' adaptor, and (iv) a universal 3' adaptor sequence comprising an R2 sequencing primer site, and wherein the second strand of the 3'PDSAs comprises, in the 5' to 3' direction, (i) a segment complementary to the first segment, and (ii) a 3' blocking group; and b) a population of 5' adaptors configured to anneal to the annealing site, wherein the 5' adaptors comprise, in the 5' to 3' direction, (i) a universal 5' adaptor sequence that is not complementary to the universal 3' adaptor sequence and that comprises an R1 sequencing primer site, and (ii) a sequence complementary to the annealing site for the 3' adaptor.

In another aspect, the disclosure provides a system, comprising: a) a population of partially double-stranded 3' adaptors (3'PDSAs) configured to be ligated to the 3' ends of both Watson and Crick strands of a population of double-stranded DNA fragments, wherein a first strand of each of the 3'PDSAs comprises, in the 5'-3' direction, (i) a first segment, (ii) a single-stranded UID sequence, (iii) an annealing site for a universal 5' adapter, and (iv) a universal 3' adapter sequence, and wherein the second strand of the 3'PDSAs comprises, in the 5' to 3' direction, (i) a segment complementary to the first segment, and (ii) a 3' blocking group; b) a population of universal 5' adapters configured to anneal to said annealing site, wherein the universal 5' adapters comprise, in the 5' to 3' direction, (i) a universal 5' adapter sequence that is not complementary to the universal 3' adapter sequence, and (ii) a sequence complementary to the annealing site for the universal 3' adapter; c) a first set of Watson target-selective primer pairs, comprising (i) one or more first Watson target-selective primers comprising a sequence complementary to the universal 3' adapter sequence, and (ii) one or more second Watson target-selective primers, each of the one or more second Watson target-selective primers comprising a target-selective sequence, and d) a first set of Crick target-selective primer pairs, comprising (i) one or more Crick target-selective primers comprising a sequence complementary to the complement of the universal 5' adapter sequence, and (ii) one or more second Crick target-selective primers, each of the one or more second Crick target selective primers comprising the same target-selective sequence as the second Watson target-selective primer sequence.

In further examples, the systems, kits, and compositions further comprise: c) the population of double-stranded DNA fragments from a biological sample, and/or c) reagents for degrading the second strand of the 3'PDSAs to generate single-stranded 3' adaptors (3'SSAs); and/or c) a first primer complementary to the universal 3' adaptor sequence, and second primer complementary to the complement of the universal 5' adaptor sequence; and/or c) a sequencing system; and/or c) a Watson anchor primer complementary to the universal 3' adapter sequence, and d) a Crick anchor primer complementary to the complement of the universal 5' adapter sequence; and/or c) a first set of Watson target-selective primer pairs, comprising (i) one or more first Watson target-selective primers comprising a sequence complementary to a portion of the universal 3' adapter sequence, and (ii) one or more second Watson target selective primers, each of the one or more second Watson target-selective primers comprising a target-selective sequence; and/or c) a first set of Crick target-selective primer pairs, comprising (i) one or more Crick target-selective

b. performing an initial amplification to amplify the adaptor-ligated, double-stranded DNA fragments to produce amplicons; c. determining sequence reads of one or more amplicons of the one or more of the adaptor-ligated, double-stranded DNA fragments; d. assigning the sequence reads into UID families, wherein each member of a UID family comprises the same exogenous UID sequence; e. identifying a nucleotide sequence as accurately representing an analyte DNA fragment when a threshold percentage of members of a UID family contain the sequence; and f. identifying a mutation in the analyte DNA fragment when the sequence identified as accurately representing an analyte DNA fragment differs from a reference sequence that lacks the mutation. In some examples, the error rate associated with the identification of one or more mutations in analyte DNA fragments according to a method of any one of the preceding claims is no more than 1×10^{-2} , no more than 1×10^{-3} , no more than 1×10^{-4} , no more than 1×10^{-5} , no more than 1×10^{-6} , no more than 5×10^{-6} , or no more than 1×10^{-7} .

Also provided herein are computer readable media comprising computer executable instructions for analyzing sequence read data from a nucleic acid sample, wherein the data is generated by a method of any one of the preceding claims. In one aspect, the disclosure provides a computer readable medium comprising computer executable instructions for analyzing sequence read data from a nucleic acid sample, wherein the data is generated by a method as disclosed herein, and wherein the computer executable instructions comprise instructions for:

a. assigning sequence reads into UID families, wherein each member of a UID family comprises the same UID sequence; b. assigning sequence reads of each UID family into Watson and Crick subfamilies based on spatial relationship of the UID sequence to the R1 and R2 read sequence; c. identifying a first Watson strand read as accurately representing a Watson strand of a dsUID-labelled DNA fragment when a threshold percentage of members of the Watson subfamily contain the same nucleic acid sequence; d. identifying a first Crick strand read as accurately representing a Crick strand of an a dsUID-labelled DNA fragment when a threshold percentage of members of the Crick subfamily contain the same nucleic acid sequence; e. identifying a mutation in the first Watson strand read as accurately representing the Watson strand when the first Watson strand read differs from a reference sequence that lacks the mutation; f. identifying a mutation in the first Crick strand read as accurately representing the Crick strand when the first Crick strand read differs from a reference sequence that lacks the mutation; and g. identifying a mutation in the dsUID-labelled double-stranded DNA fragment when the is present in a target Watson strand read and a corresponding mutation is present in the target Crick strand read.

In some examples, the computer readable medium comprises executable instructions for

a. assigning sequence reads into UID families, wherein each member of a UID family comprises the same exogenous UID sequence;
b. assigning sequence reads of each UID family into Watson and Crick subfamilies based on spatial relationship of the exogenous UID sequence to the R1 and R2 read sequence;
c. identifying a nucleotide sequence as accurately representing a Watson strand of an analyte DNA fragment when a threshold percentage of members of the Watson subfamily contain the sequence;
d. identifying a nucleotide sequence as accurately representing a Crick strand of an analyte DNA fragment when a threshold percentage of members of the Crick subfamily contain the sequence;

WHAT IS CLAIMED IS:

1. A method for assessing nucleic acids, comprising:
 - a) attaching partially double-stranded 3' adapters (3'PDSAs) to 3' ends of both Watson and Crick strands of a population of double-stranded DNA fragments in an analyte DNA sample,

wherein a first strand of each of the 3'PDSAs comprises, in the 5'-3' direction, (i) a first segment, (ii) a single stranded unique identifier (ssUID) sequence, (iii) an annealing site for a universal 5' adapter, and (iv) a universal 3' adapter sequence, and

wherein the second strand of the 3'PDSAs comprises, in the 5' to 3' direction, (i) a segment complementary to the first segment, and (ii) a 3' blocking group,
 - b) annealing universal 5' adapters to said annealing site, wherein the universal 5' adapters comprise, in the 5' to 3' direction, (i) a universal 5' adapter sequence that is not complementary to the universal 3' adapter sequence, and (ii) a sequence complementary to the annealing site for the 5' adapter;
 - c) extending the universal 5' adapters across the ssUID sequence and said first segment, thereby generating the complement of said ssUID sequence and complement of said first segment;
 - d) covalently linking the 3' end of said complement of said first segment to the 5' ends of the Watson and Crick strands of the double-stranded DNA fragments, thereby generating a plurality of adapter-ligated double-stranded DNA fragments labelled with a double-stranded unique identifier (dsUID) sequence on each end;
 - e) amplifying the dsUID-labelled double-stranded DNA fragments, thereby generating a duplex sequencing library sample;
 - f) amplifying a first aliquot from the duplex sequencing library sample with a first Watson target-selective primer pair, thereby generating dsUID-labelled double-stranded target Watson amplification products (dsUID-dsTWAPs) and not generating dsUID-labelled double-stranded target Crick amplification products (dsUID-dsTCAPs),
 - g) performing hemi-nested amplification of the dsUID-dsTWAPs, thereby creating a first sample comprising a plurality of dsUID-labelled double-stranded target Watson library members (dsUID-dsTWLMs);

h) amplifying a second aliquot from the duplex sequencing library sample with a first Crick target-selective primer pair, thereby generating dsUID-dsTCAPs and not generating dsUID-dsTWAPs;

i) performing hemi-nested amplification of the dsUID-dsTCAPs, thereby creating a second sample comprising a plurality of dsUID-labelled double-stranded target Crick family members (dsUID-dsTCLMs); and

j) sequencing the dsUID-dsTWLMs and the dsUID-dsTCLMs, thereby generating a plurality of UID-labelled Watson strand reads and a plurality of UID-labelled Crick strand reads.

2. The method of claim 1, wherein:

A) the first Watson target-selective primer pair comprises: (i) a first Watson target-selective primer comprising a sequence complementary to the universal 3' adapter sequence, and (ii) a second Watson target-selective primer comprising a first target-selective sequence; and

B) the first Crick target-selective primer pair comprises: (i) a first Crick target-selective primer comprising a sequence complementary to the complement of the universal 5' adapter sequence, and (ii) a second Crick target-selective primer comprising a first Crick target-selective sequence that comprises the same target-selective sequence as the first Watson target-selective sequence.

3. The method of claim 2, wherein:

A) said hemi-nested amplification of the dsUID-dsTWAPs is performed using second set of Watson target-selective primers comprising: (i) a third Watson target-selective primer comprising: a sequence complementary to the universal 3' adapter sequence, and a Watson sample index barcode, and (ii) a fourth Watson target-selective primer comprising, in the 5' to 3' direction: an R1 sequencing primer binding site, and a second Watson target-selective sequence that is nested with respect to the first Watson target-selective sequence; and

B) said hemi-nested amplification of the dsUID-dsTCAPs is performed using a second set of Crick target-selective primers comprising:

(i) a third Crick target-selective primer comprising a sequence complementary to the complement of the universal 5' adapter sequence, and a Crick sample barcode index, and

(ii) a fourth Crick target-selective primer comprising, in the 5' to 3' direction: an R2 sequencing primer site, and a second Crick target-selective sequence that is nested with respect to said first Crick target-selective sequence and comprises the same target-selective sequence as said second Watson target-selective sequence.

4. The method of any one of claims 1-3, further comprising: removing the 3' blocking group from the second strand of the 3'PDSAs, wherein the 3' blocking group comprises one or more deoxyuridines, and wherein said removing comprises contacting the analyte DNA sample with uracil-DNA glycosylase (UDG) to degrade the 3'blocking group.

5. The method of claim any one of claims 1-4, further comprising assigning sequencing reads into UID families, wherein each member of a UID family comprises the same UID sequence.

6. The method of claim 5, wherein UID-labelled Watson strand reads comprise an R1 read sequence, a target Watson strand read, and a Watson sample index barcode read, and wherein UID-labelled Crick strand reads comprise an R2 read sequence, a target Crick strand read, and a Crick sample index barcode read.

7. The method of claim 6, further comprising assigning sequencing reads of each UID family into a Watson subfamily and Crick subfamily based on: i) the spatial relationship of the UID sequence to the R1 and R2 read sequence, and/or ii) the presence of said Watson sample index barcode or said Crick sample index barcode.

8. The method of claim 7, further comprising: identifying a first target Watson strand read as accurately representing a Watson strand of a dsUID-labelled double-stranded DNA fragment when at least 50% of the members of the Watson subfamily contain the same nucleic acid sequence, and identifying a first target Crick strand read as accurately representing a Crick strand of a dsUID-labelled double-stranded DNA fragment when at least 50% of the members of the Crick subfamily contain the same nucleic acid sequence.

9. The method of claim 8, further comprising: identifying a mutation in the first target Watson strand read as accurately representing the Watson strand when the first target

Watson strand read differs from a reference sequence that lacks the mutation, and identifying a mutation in the first target Crick strand read as accurately representing the Crick strand when the first Crick Strand read differs from a reference sequence that lacks the mutation.

10. The method of claim 9, further comprising identifying a mutation in the dsUID-labelled double-stranded DNA fragment when the mutation is present in a target Watson strand read and a corresponding mutation is present in the target Crick strand read.

11. The method of any one of claims 1-10, wherein said amplifying in step (e) employs no more than 50 cycles of PCR.

12. The method of claim 11, wherein said amplifying employs no more than 18 cycles of PCR.

13. The method of claim 11, wherein said amplifying employs no more than 11 cycles of PCR.

14. The method of claim 11, wherein said amplifying employs no more than 7 cycles of PCR.

15. The method of any one of claims 1-14, wherein the population of double-stranded DNA fragments are from a biological sample obtained from a subject, wherein the biological sample is a fluid sample or a solid sample.

16. The method of claim 15, wherein the biological sample is a fluid sample selected from whole blood, plasma, serum, sputum, urine, sweat, tears, and semen.

17. The method of claim 15, wherein the biological sample is a tumor sample.

18. The method of any one of claims 9-17, wherein the identified mutation is present in the population of double-stranded DNA fragments at a frequency of 0.1% or less.

19. The method of claim 18, wherein the identified mutation is present in the population of double-stranded DNA fragments at a frequency of 0.1% to 0.00001%.

20. A system, comprising:

a) a population of partially double-stranded 3' adaptors (3'PDSAs) configured to be ligated to the 3' ends of both Watson and Crick strands of a population of double-stranded DNA fragments,

wherein a first strand of each of the 3'PDSAs comprises, in the 5'-3' direction, (i) a first segment, (ii) a single-stranded UID sequence, (iii) an annealing site for a universal 5' adapter, and (iv) a universal 3' adapter sequence, and

wherein the second strand of the 3'PDSAs comprises, in the 5' to 3' direction, (i) a segment complementary to the first segment, and (ii) a 3' blocking group;

b) a population of universal 5' adapters configured to anneal to said annealing site, wherein the universal 5' adapters comprise, in the 5' to 3' direction, (i) a universal 5' adapter sequence that is not complementary to the universal 3' adapter sequence, and (ii) a sequence complementary to the annealing site for the universal 3' adapter;

c) a first set of Watson target-selective primer pairs, comprising (i) one or more first Watson target-selective primers comprising a sequence complementary to the universal 3' adapter sequence, and (ii) one or more second Watson target-selective primers, each of the one or more second Watson target-selective primers comprising a target-selective sequence, and

d) a first set of Crick target-selective primer pairs, comprising (i) one or more Crick target-selective primers comprising a sequence complementary to the complement of the universal 5' adapter sequence, and (ii) one or more second Crick target-selective primers, each of the one or more second Crick target selective primers comprising the same target-selective sequence as the second Watson target-selective primer sequence.

21. A computer readable medium comprising computer executable instructions for analyzing sequence read data from a nucleic acid sample, wherein the data is generated by a method of any one of claims 1-19, and wherein the computer executable instructions comprise instructions for:

a. assigning sequence reads into UID families, wherein each member of a UID family comprises the same UID sequence;

- b. assigning sequence reads of each UID family into Watson and Crick subfamilies based on spatial relationship of the UID sequence to the R1 and R2 read sequence;
- c. identifying a first Watson strand read as accurately representing a Watson strand of a dsUID-labelled DNA fragment when a threshold percentage of members of the Watson subfamily contain the same nucleic acid sequence;
- d. identifying a first Crick strand read as accurately representing a Crick strand of an a dsUID-labelled DNA fragment when a threshold percentage of members of the Crick subfamily contain the same nucleic acid sequence;
- e. identifying a mutation in the first Watson strand read as accurately representing the Watson strand when the first Watson strand read differs from a reference sequence that lacks the mutation;
- f. identifying a mutation in the first Crick strand read as accurately representing the Crick strand when the first Crick strand read differs from a reference sequence that lacks the mutation; and
- g. identifying a mutation in the dsUID-labelled double-stranded DNA fragment when the is present in a target Watson strand read and a corresponding mutation is present in the target Crick strand read.