

of exposure to environmental mutagens, to monitor endogenous DNA repair, and to study the accumulation of somatic mutations in aging individuals. Additionally, more sensitive methods to detect rare variants can enhance prenatal diagnosis, enabling the characterization of fetal cells present in maternal blood.

## SUMMARY

**[0005]** In view of the foregoing, there is a need for improved methods of detecting rare sequence variants. The compositions and methods of the present disclosure address this need, and provide additional advantages as well. In particular, the various aspects of the disclosure provide for highly sensitive detection of rare or low frequency nucleic acid sequence variants (sometimes referred to as mutations). This includes identification and elucidation of low frequency nucleic acid variations (including substitutions, insertions and deletions) in samples that may contain low amounts of variant sequences in a background of normal sequences, as well as the identification of low frequency variations in a background of sequencing errors.

**[0005a]** In one aspect, the disclosure provides a method for processing or analyzing a biological sample of a subject, comprising: (a) obtaining said biological sample of said subject; (b) using a ribonucleic acid (RNA) molecule or derivative thereof derived from said biological sample to generate a circular RNA molecule; (c) subjecting said circular RNA molecule or derivative thereof to reverse transcription to generate a complementary deoxyribonucleic acid (cDNA) concatemer; (d) creating an amplicon of a target sequence of said cDNA molecule using linear amplification to amplify said cDNA concatemer or derivative thereof using a primer that comprises (i) a sequence that binds to said target sequence or a sequence that binds to a complement of a junction sequence between a 5' end and a 3' end of said RNA molecule in said circular RNA molecule comprising said target sequence; and (ii) an adaptor sequence; and (e) subjecting said amplicon of said cDNA concatemer or derivative thereof to sequencing to identify a sequence of said RNA molecule.

**[0006]** In one aspect, the disclosure provides a method of identifying a sequence variant, such as in a nucleic acid sample. In some embodiments, each polynucleotide of the plurality has a 5' end and a 3' end, and the method comprises: (a) circularizing individual polynucleotides of said plurality to form a plurality of circular polynucleotides, each of which having a junction between the 5' end and 3' end; (b) amplifying the circular polynucleotides of (a); (c) sequencing the amplified polynucleotides to produce a plurality of sequencing reads; (d) identifying sequence differences between sequencing reads and a reference sequence; and (e) calling a sequence difference that occurs in at least two circular polynucleotides having different junctions as the sequence variant. In some embodiments, the method comprises identifying sequence differences between sequencing reads and a reference sequence, and calling a sequence difference that occurs in at least two circular polynucleotides having different junctions as the sequence variant, wherein: (a) the sequencing reads correspond to amplification products of the at least two circular polynucleotides; and (b) each of the at least two

circular polynucleotides comprises a different junction formed by ligating a 5' end and 3' end of the respective polynucleotides.

**[0007]** The plurality of polynucleotides can be single- or double-stranded. In some embodiments, the polynucleotides are single-stranded. In some embodiments, circularizing is effected by subjecting the plurality of polynucleotides to a ligation reaction. In some embodiments, an individual circular polynucleotide has a junction that is unique among the circularized polynucleotides. In some embodiments, the sequence variant is a single nucleotide polymorphism (SNP). In some embodiments, the reference sequence is a consensus sequence formed by aligning the sequence reads with one another. In some embodiments, the reference sequence is a known reference sequence, such as a reference genome or portion thereof. In some embodiments, circularizing comprises the step of joining an adapter polynucleotide to the 5' end, the 3' end, or both the 5' end and the 3' end of a

## CLAIMS

### WHAT IS CLAIMED IS:

1. A method for processing or analyzing a biological sample of a subject, comprising:
  - (a) obtaining said biological sample of said subject;
  - (b) using a ribonucleic acid (RNA) molecule or derivative thereof derived from said biological sample to generate a circular RNA molecule;
  - (c) subjecting said circular RNA molecule or derivative thereof to reverse transcription to generate a complementary deoxyribonucleic acid (cDNA) concatemer;
  - (d) creating an amplicon of a target sequence of said cDNA molecule using linear amplification to amplify said cDNA concatemer or derivative thereof using a primer that comprises (i) a sequence that binds to said target sequence or a sequence that binds to a complement of a junction sequence between a 5' end and a 3' end of said RNA molecule in said circular RNA molecule comprising said target sequence; and (ii) an adaptor sequence; and
  - (e) subjecting said amplicon of said cDNA concatemer or derivative thereof to sequencing to identify a sequence of said RNA molecule.
2. The method of claim 1, wherein said biological sample is a cell-free biological sample.
3. The method of claim 1, further comprising, subsequent to (c), capturing said cDNA concatemer or derivative thereof using a capture probe.
4. The method of claim 3, wherein said capture probe comprises one or more copies of a target sequence of said cDNA concatemer.
5. The method of claim 1, wherein said adaptor sequence comprises a molecular barcode sequence.
6. The method of claim 5, wherein said primer further comprises a first sequencing adaptor sequence.
7. The method of claim 5, further comprising subjecting said amplicon to a primer extension reaction using a first primer comprising a first sequencing adaptor sequence and a second primer comprising a second sequencing adaptor sequence.

8. The method of claim 1, wherein said reverse transcription of (c) is performed with a primer comprising: (i) a sequence that binds to a target sequence of said circular RNA molecule; and (ii) a molecular barcode sequence.
9. The method of claim 1, wherein said reverse transcription of (c) is performed with a primer comprising (i) a random sequence that binds to said circular RNA molecule; and (ii) a molecular barcode sequence.
10. The method of claim 1, wherein said reverse transcription of (c) is performed with a primer comprising (i) a sequence that binds to a junction sequence between a 5' end and a 3' end of said RNA molecule in said circular RNA molecule; and (ii) a molecular barcode sequence.
11. The method of claim 1, wherein said RNA molecule is a micro-RNA (miRNA).
12. The method of claim 1, wherein said sequence comprises a junction sequence in which a 5' end and a 3' end of said RNA molecule are joined.
13. The method of claim 1, wherein said sequence identifies said 5' end or said 3' end of said RNA molecule.
14. The method of claim 1, wherein (b) comprises ligating ends of said RNA molecule or derivative thereof to one another.
15. The method of claim 1, wherein said sequencing comprises (i) bringing said amplicon of said cDNA concatemer or derivative thereof in contact with a plurality of nucleotides in presence of a polymerase to incorporate one or more nucleotides of said plurality of nucleotides into a growing strand complementary to a strand of said amplicon of said cDNA concatemer or derivative thereof, and (ii) detecting one or more signals indicative of incorporation of said one or more nucleotides into said growing strand.
16. The method of claim 1, wherein said RNA molecule is from a tumor.
17. The method of claim 1, further comprising processing said sequence against a reference.
18. The method of claim 1, further comprising adding a molecular barcode sequence to (i) said RNA molecule by ligation prior to circularization; (ii) said cDNA concatemer during reverse

transcription; or (iii) a reverse complement strand of said cDNA concatemer during synthesis of a strand complementary to said cDNA concatemer.

19. The method of claim 18, further comprising processing said sequence to identify a consensus sequence when said sequence occurs in (i) a majority of repeats or at least two repeats of said concatemer; and/or (ii) at least two circular RNA molecules or derivatives thereof having different molecular barcode sequences.
20. The method of claim 18, further comprising quantifying said RNA molecule by counting sequencing reads from circular RNA molecules having different molecular barcodes.
21. The method of claim 1, further comprising processing said sequence to identify a consensus sequence when said sequence occurs in (i) at least two copies of said RNA molecule in said cDNA concatemer and/or (ii) a majority of copies of said RNA molecule in said cDNA concatemer.
22. The method of claim 1, further comprising comparing a consensus sequence of said RNA molecule to a reference RNA sequence to identify a sequence difference.
23. The method of claim 1, further comprising processing said sequence to identify a disease in said subject.
24. The method of claim 23, wherein said disease is cancer.
25. The method of claim 24, wherein said sequence is computer processed against a reference to identify one or more genetic variants, and wherein said one or more genetic variants identify said cancer.
26. The method of claim 1, further comprising processing said sequence to monitor a progression of a disease in said subject in response to treatment.
27. The method of claim 1, wherein said biological sample comprises a bodily fluid.
28. The method of claim 27, wherein said bodily fluid is urine, saliva, blood, serum, or plasma.