

WHAT IS CLAIMED IS:

1. A computer-implemented method of aligning RNA comprising:
receiving onto a data storage unit a plurality of primer sequences and a plurality of transcript sequences from a reference genome, the transcript sequences being transcribable from the reference genome based on a gene model;
generating, using a microprocessor, a plurality of target sequences to be amplified from a combination of the plurality of primer sequences and the plurality of transcript sequences;
generating, using a microprocessor, a modified reference genome based on the plurality of target sequences;
aligning, using a microprocessor, sequence reads generated from a test sample comprising RNA amplicon molecules to the modified reference genome; and
generating an alignment profile for the test sample based on the aligning.
2. The method of claim 1, further comprising assigning primer sequences individual loci corresponding to loci of respective transcript sequences.
3. The method of claim 2, further comprising removing one or more of the generated target sequences based on the one or more of the generated target sequences spanning more than one on-target sequence.
4. The method of claim 2, wherein the plurality of primer sequences comprises a plurality of primer pairs, and a first primer pair comprises a first primer and a second primer for a first locus, and a second primer pair comprises the first primer and a second primer for a second locus.
5. The method of claim 1, wherein the gene model comprises identification of splice junctions, fusion junctions, or both, in the modified reference genome.
6. The method of claim 5, further comprising translating sequence reads aligned to targets derived from splice and fusion junctions.
7. The method of claim 1, wherein the plurality of target sequences comprise on-target sequences and off-target sequences.
8. The method of claim 7, further comprising reducing a number of off-target sequences by excluding one or more primer sequence from the plurality of primer sequences.
9. The method of claim 1, further comprising computationally comparing gene expression of two or more samples, wherein aligned reads generated from a first sample of RNA are compared to aligned reads generated from a second sample of RNA, wherein the alignment is performed using the plurality of target sequences.

10. The method of claim 1, wherein the alignment profile includes at least one of placement, a quality score, and sequence integrity for the sequence reads of the test sample.

11. The method of claim 1, further comprising:

translating the sequence reads from the test sample to a whole reference genome using the mapped target sequences and the modified reference genome.

12. The method of claim 1, wherein generating an alignment profile further comprises aligning a sequence read comprising an unaligned fusion junction to non-contiguous sequences of the reference genome, wherein the unaligned fusion junction was not identified in the gene model.

13. The method of claim 5, wherein the alignment profile comprises a fusion junction and the fusion junction was identified in the gene model.

14. A computer-implemented method of aligning RNA comprising:

receiving onto a data storage unit a plurality of primer sequences and a plurality of transcript sequences from a reference genome, the transcript sequences being transcribable from the reference genome using a gene model comprising identification of splice junctions, fusion junctions, or both, in the reference genome;

assigning primer sequences individual loci corresponding to loci of respective transcript sequences;

generating, using a microprocessor, a plurality of target sequences to be amplified from a combination of the plurality of transcript sequences and the plurality of primer sequences;

generating, using a microprocessor, a modified reference genome based on the plurality of target sequences;

aligning, using a microprocessor, sequence reads generated from a test sample comprising RNA amplicon molecules to the modified reference genome;

generating an alignment profile wherein the alignment profile includes at least one of placement, a quality score, and sequence integrity for the sequence reads of the test sample; and

translating the sequence reads from the test sample to a whole reference genome using the mapped target sequences and the modified reference genome.

15. A computer system of aligning RNA comprising:

one or more microprocessors,

one or more memories storing a plurality of primer sequences and a plurality of transcript sequences from a reference genome, and a gene model, the transcript sequences being transcribable from the reference genome based on the gene model;

the one or more memories storing instructions that, when executed by the one or more microprocessors, cause the computer system to:

generate a plurality of target sequences to be amplified from a combination of the plurality of primer sequences and the plurality of transcript sequences;

generate a modified reference genome based on the plurality of target sequences;

align sequence reads generated from a test sample comprising RNA amplicon molecules to the modified reference genome; and

generate an alignment profile for the test sample based on the aligning.

16. The computer system of claim 15, wherein the instructions cause the computer system to assign primer sequences individual loci corresponding to loci of respective transcript sequences.

17. The computer system of claim 16, wherein the instructions cause the computer system to remove one or more of the generated target sequences based on the one or more of the generated target sequences spanning more than one on-target sequence.

18. The computer system of claim 16, wherein the plurality of primer sequences comprises a plurality of primer pairs, and a first primer pair comprises a first primer and a second primer for a first locus, and a second primer pair comprises the first primer and a second primer for a second locus

19. The computer system of claim 15, wherein the gene model comprises identification of splice junctions, fusion junctions, or both, in the modified reference genome.

20. The computer system of claim 15, wherein the plurality of target sequences comprise on-target sequences and off-target sequences.

21. The computer system of claim 20, wherein the instructions cause the computer system to reduce a number of off-target sequences by excluding one or more primer sequence from the plurality of primer sequences.

22. The computer system of claim 21, wherein the instructions cause the computer system to compare gene expression of two or more samples, whereby aligned reads generated from a first sample of RNA are compared to aligned reads generated from a second sample of RNA.

23. The computer system of claim 15, wherein generating an alignment profile further comprises aligning a sequence read comprising an unaligned fusion junction to non-

contiguous sequences of the reference genome, wherein the unaligned fusion junction was not identified in the gene model.

24. The computer system of claim 19, wherein the alignment profile comprises a fusion junction and the fusion junction was identified in the gene model.