

**METHODS FOR GENERATING, AND SEQUENCING FROM, ASYMMETRIC
ADAPTORS ON THE ENDS OF POLYNUCLEOTIDE TEMPLATES COMPRISING
HAIRPIN LOOPS**

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claim priority under 35 U.S.C. §119(e) to U.S. Provisional Application No. 62/926,360, filed October 25, 2019, the disclosures of which are incorporated herein by reference for all purposes.

FIELD OF THE INVENTION

[0002] The disclosure relates generally to methods for generating asymmetric adaptors on the ends of polynucleotides comprising hairpin loops, and sequencing therefrom.

INCORPORATION BY REFERENCE OF SEQUENCE LISTING

[0003] Accompanying this filing is a Sequence Listing entitled, "Sequence-Listing_ST25.txt" created on October 19, 2020 and having 784 bytes of data, machine formatted on IBM-PC, MS-Windows operating system. The sequence listing is hereby incorporated by reference in its entirety for all purposes.

BACKGROUND

[0004] Sequencing platforms that employ closed-ended double stranded templates ideally require two different adaptor sequences at the ends of sequencing templates. Such end asymmetry enables a single sequencing primer and polymerase to bind to only one end of the close-ended templates and generate only one sequence read per template molecule. This is not currently possible with standard library preparations, where the same sequences are appended to both ends of a template, thus introducing Poisson artefacts to sequencing, *i.e.*, some templates have no primer/polymerase bound, some have one and some have to (one at each end).

SUMMARY

[0005] The disclosure describes methods for generating asymmetric adaptors on the ends of polynucleotide templates comprising hairpin loops. By enabling asymmetric adaptors, Poisson artefacts seen with standard methodologies can be avoided, thereby increasing the conversion yield of sequencing libraries. Further, the methods disclosed herein simplify library preparation workflow and allow for

some steps of library preparation to be conducted on the sequencing platform.

[0006] In one example, the disclosure describes a method of generating an asymmetric closed-ended double stranded nucleic acid template from a double stranded nucleic acid template having free 5' and 3' ends, comprising: (A) attaching a first nucleic acid-based hairpin or dumbbell adaptor to the 3' ends of a double stranded nucleic acid template comprising free 5' and 3' ends; (B) extending from each 3' end of the nucleic acid hairpin a sequence complementary to the double stranded nucleic acid template using a processive polymerase to generate two long hairpin duplex templates, wherein one end of the duplex template comprises a closed hairpin (the "hairpin end") and the other end of the duplex comprises a free 3'-strand end and a free 5'-strand end (the "free end"); (C) closing the free end of each long hairpin duplex template to form an asymmetric closed-ended double stranded nucleic acid template by: (i) ligating a second nucleic acid-based hairpin or dumbbell adaptor to the free end of the duplex template; or (ii) using a TelN Protelomerase to close the free end of the duplex template, wherein the free end of the duplex template is designed to include a TelN recognition sequence. In a further example presented herein, the double stranded nucleic acid template is a double stranded DNA template. In another example presented herein, the 5' and 3' ends of the double stranded nucleic acid template are dephosphorylated and end repaired. In yet another example presented herein, the double stranded nucleic acid template has blunt 5' and 3' ends. In another example presented herein, the double stranded nucleic acid template has A-tailed 3' ends. In a further example presented herein, the first nucleic acid-based hairpin or dumbbell adaptor is ligated to the 3' ends of the double stranded nucleic acid template using a ligase. In yet a further example presented herein, the ligase is a T4 DNA Ligase or a T3 DNA ligase. In another example presented herein, the first nucleic acid-based hairpin or dumbbell adaptor comprises blunt ends or T-tailed ends. In yet another example presented herein, the

first nucleic acid-based hairpin adaptor is a Y-shaped adaptor. In a certain example presented herein, dimers formed from two first nucleic acid-based hairpin or dumbbell adaptors being bound to each other are removed by using size-selection or size-exclusion techniques. In another example presented herein, the processive polymerase is Phi29 polymerase. In yet another example presented herein, the second nucleic acid-based hairpin or dumbbell adaptor comprises blunt ends or T-tailed ends. In a further example presented herein, prior to step (C) the long hairpin duplex templates are digested with a restriction enzyme that creates a 5' overhang. In yet a further example presented herein, prior to step (C) the long hairpin duplex templates are digested with a restriction enzyme that creates a 3' overhang. In another example presented herein, the second nucleic acid-based hairpin or dumbbell adaptor comprises an overhang sequence that is complementary to the overhang sequence of the digested long hairpin duplex templates. In another example presented herein, the second nucleic acid-based hairpin or dumbbell adaptor are ligated to the free end of the duplex template using a polynucleotide kinase and a ligase. In yet another example presented herein, dimers formed from two second nucleic acid-based hairpin or dumbbell adaptors being bound to each other are removed by using size-selection or size-exclusion techniques. In a further example presented herein, the TelN Protelomerase is from phage N15, and the TelN Protelomerase cuts the long hairpin duplex templates at the TelN recognition sequence and leaves covalently closed ends at the site of cleavage. In yet a further example presented herein, a method disclosed herein further comprises the step of: (C') generating nanoball complexes comprising polycistronic amplified asymmetric closed-ended double stranded nucleic acid templates using rolling circle replication. In another example presented herein, a method disclosed herein further comprises the step of: (D) sequencing the asymmetric closed-ended double stranded nucleic acid template or nanoball complexes using a sequencing primer and a polymerase. In yet

another example presented herein, step (B) step (C)(ii) and step (D) can be combined together as a one-pot reaction. In a further example presented herein, step (B), step (C)(ii) and step (D) are carried out in a well of an automated sequencing platform.

[0007] In one aspect, the disclosure describes a method of generating an asymmetric double stranded nucleic acid template from tagged DNA comprising complementary hairpin loops, comprising: (I) generating tagged DNA comprising complementary hairpin loops at the 5' ends of each strand, wherein the hairpin loops comprise a base paired transposase recognition sequence, and wherein there is a gap of single stranded sequence between the 5' ends and the 3' ends of the tagged DNA; (II) filling in the gaps between the 5' ends and the 3' ends of the tagged DNA using a gap-fill-ligation reaction to form closed ended tagged DNA; (III) generating a nick in the top strand at each hairpin region of the closed ended tagged DNA; (IV) extending from each nick a sequence complementary to the double stranded nucleic acid template using a processive polymerase to generate two long hairpin duplex templates, wherein one end of the duplex template comprises a closed hairpin (the "hairpin end") and the other end of the duplex template comprises a 3'-strand end and a 5'-strand end (the "free end"); (V) closing the free end of each long hairpin duplex template to form an asymmetric closed-ended double stranded nucleic acid template by: (a) ligating a nucleic acid-based hairpin or dumbbell adaptor to the free end of the duplex template; or (b) using a TelN Protelomerase to close the free end of the duplex template, wherein the free end of the duplex template is designed to include a TelN recognition sequence. In a further example presented herein, the transposase recognition sequence is a 19-bp Mosaic End sequence. In yet a further example presented herein, the gap is 9 base pairs in length. In another example presented herein, the gap-fill-ligation reaction comprises Klenow fragment. In yet another example presented herein, the gap-fill-ligation reaction comprises T4 DNA polymerase and a thermostable DNA ligase. In a certain example presented herein, the nick is generated by using a site-specific endonuclease.

In a further example presented herein, the processive polymerase is Phi29 polymerase. In yet a further example presented herein, the nucleic acid-based hairpin or dumbbell adaptor comprises blunt ends or T-tailed ends. In another example presented herein, prior to step (V) the long hairpin duplex templates are digested with a restriction enzyme that creates a 5' overhang. In yet another example presented herein, prior to step (V) the long hairpin duplex templates are digested with a restriction enzyme that creates a 3' overhang. In a further example presented herein, the nucleic acid-based hairpin or dumbbell adaptor comprises an overhang sequence that is complementary to the overhang sequence of the digested long hairpin duplex templates. In yet a further example presented herein, the nucleic acid-based hairpin or dumbbell adaptor are ligated to the free end of the duplex template using a polynucleotide kinase and a ligase. In another example presented herein, dimers formed from two nucleic acid-based hairpin or dumbbell adaptors being bound to each other are removed by using size-selection or size-exclusion techniques. In yet another example presented herein, the TelN Protelomerase is from phage N15, and the TelN Protelomerase cuts the long hairpin duplex templates at the TelN recognition sequence and leaves covalently closed ends at the site of cleavage. In a further example presented herein, the method further comprises the step of: (V') generating nanoball complexes comprising polycistronic amplified asymmetric closed-ended double stranded nucleic acid templates using rolling circle replication. In yet a further example presented herein, the method further comprises the step of: (VI) sequencing the asymmetric closed-ended double stranded nucleic acid template or nanoball complexes using a sequencing primer and a polymerase. In another example presented herein, step (V)(ii) and step (VI) can be combined together into a single step. In yet another example presented herein, step (V)(ii) and step (VI) are carried out in a well of an automated sequencing platform.

[0008] In another aspect, the disclosure also describes a method of sequencing tagmented DNA comprising complementary hairpin loops, comprising: (I) generating tagmented DNA comprising complementary hairpin loops at the 5' ends of each strand, wherein the hairpin loops comprise a base paired transposase recognition sequence, and wherein there is a gap of single stranded sequence between the 5' ends and the 3' ends of the tagmented DNA; (II) forming two extended strands of tagmented DNA comprising a transposase recognition sequence and complementary transposase recognition sequence at 5' and 3' ends of each extended strand of tagmented DNA by using a polymerase; (III) separating the extended strands of tagmented DNA and rehybridizing the transposase recognition sequences and complementary transposase recognition sequences at the ends of each of the extended strands of tagmented DNA to form complementary hairpin loops; and (IV) sequencing the extended strand of tagmented DNA comprising complementary hairpin loops by using a sequencing polymerase. In another example presented herein, step (IV) is carried out in a well of an automated sequencing platform. In yet another example presented herein, prior to sequencing of step (IV): (III') generating nanoball complexes comprising polycistronic amplified extended strands of tagmented DNA.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] **Figure 1** provides a schematic illustrating a method for generating template end asymmetry by creating a hairpin end and a free end by ligating the 5' of a hairpin to the 3' ends of a double stranded polynucleotide template and extending the 3' end of the hairpin using a processive polymerase. An additional hairpin or dumbbell can be ligated to free end of the asymmetric template to provide a closed ended template comprising asymmetrical regions. Small adaptor dimers can be removed using size-selection or size exclusion technologies.

[0010] **Figure 2** provides a schematic illustrating a method for generating template end asymmetry by creating a hairpin end and a free end by ligating the 5' of a Y adaptor to the 3' ends of a double stranded polynucleotide template and extending the 3' end of the Y adaptor using a processive polymerase.

WHAT IS CLAIMED:

1. A method of generating an asymmetric double stranded nucleic acid template from tagmented DNA comprising complementary hairpin loops, comprising:

(I) generating tagmented DNA comprising complementary hairpin loops at the 5' ends of each strand, wherein the hairpin loops comprise a base paired transposase recognition sequence, and wherein there is a gap of single stranded sequence between the 5' ends and the 3' ends of the tagmented DNA;

(II) filling in the gaps between the 5' ends and the 3' ends of the tagmented DNA using a gap-fill-ligation reaction to form closed ended tagmented DNA;

(III) generating a nick in the top strand at each hairpin region of the closed ended tagmented DNA;

(IV) extending from each nick a sequence complementary to the double stranded nucleic acid template using a processive polymerase to generate two long hairpin duplex templates, wherein one end of the duplex template comprises a closed hairpin (the "hairpin end") and the other end of the duplex template comprises a 3'-strand end and a 5'-strand end (the "free end"); and

(V) closing the free end of each long hairpin duplex template to form an asymmetric closed-ended double stranded nucleic acid template by:

(a) ligating a nucleic acid-based hairpin or dumbbell adaptor to the free end of the duplex template; or

(b) using a TelN Protelomerase to close the free end of the duplex template, wherein the free end of the duplex template is designed to include a TelN recognition sequence.

2. The method of claim 1, wherein the transposase recognition sequence is a 19-bp Mosaic End sequence.

3. The method of claim 2, wherein the gap is 9 base pairs in length.

4. The method of any of claims 1 to 3, wherein the gap-fill-ligation reaction comprises a Klenow fragment.
5. The method of any of claims 1 to 3, wherein the gap-fill-ligation reaction comprises T4 DNA polymerase and a thermostable DNA ligase.
6. The method of any of claims 1 to 5, wherein the nick is generated by using a site-specific endonuclease.
7. The method of any of claims 1 to 6, wherein the processive polymerase is Phi29 polymerase.
8. The method of any of claims 1 to 7, wherein the nucleic acid-based hairpin or dumbbell adaptor comprises blunt ends or T-tailed ends.
9. The method of any one of claims 1 to 7, wherein prior to step (V) the long hairpin duplex templates are digested with a restriction enzyme that creates a 5' overhang.
10. The method of any one of claims 1 to 7, wherein prior to step (V) the long hairpin duplex templates are digested with a restriction enzyme that creates a 3' overhang.
11. The method of claims 9 or 10, wherein the nucleic acid-based hairpin or dumbbell adaptor comprises an overhang sequence that is complementary to the overhang sequence of the digested long hairpin duplex templates.
12. The method of any one of claims 1 to 11, wherein the nucleic acid-based hairpin or dumbbell adaptor are ligated to the free end of the duplex template using a polynucleotide kinase and a ligase.
13. The method of any one of the preceding claims, wherein dimers formed from two nucleic acid-based hairpin or dumbbell adaptors being

bound to each other are removed by using size-selection or size-exclusion techniques.

14. The method of any one of claims 1 to 7, wherein the TelN Protelomerase is from phage N15, and the TelN Protelomerase cuts the long hairpin duplex templates at the TelN recognition sequence and leaves covalently closed ends at the site of cleavage.

15. The method of any one of claims 1 to 14, wherein the method further comprises the step of:

(V') generating nanoball complexes comprising polycistronic amplified asymmetric closed-ended double stranded nucleic acid templates using rolling circle replication.

16. The method of claim 15, wherein the method further comprises the step of:

(VI) sequencing the asymmetric closed-ended double stranded nucleic acid template or nanoball complexes using a sequencing primer and a polymerase.

17. The method of claim 16, wherein step (V) (ii) and step (VI) can be combined together into a single step.

18. The method of claim 17, wherein step (V) (ii) and step (VI) are carried out in a well of an automated sequencing platform.

19. A method of sequencing tagmented DNA comprising complementary hairpin loops, comprising:

(I) generating tagmented DNA comprising complementary hairpin loops at the 5' ends of each strand, wherein the hairpin loops comprise a base paired transposase recognition sequence, and wherein there is a gap of single stranded sequence between the 5' ends and the 3' ends of the tagmented DNA;

(II) forming two extended strands of tagmented DNA comprising a transposase recognition sequence and complementary transposase

recognition sequence at 5' and 3' ends of each extended strand of tagmented DNA by using a polymerase;

(III) separating the extended strands of tagmented DNA and rehybridizing the transposase recognition sequences and complementary transposase recognition sequences at the ends of each of the extended strands of tagmented DNA to form to complementary hairpin loops; and

(IV) sequencing the extended strand of tagmented DNA comprising complementary hairpin loops by using a sequencing polymerase.

20. The method of claim 19, wherein step (IV) is carried out in a well of an automated sequencing platform.

21. The method of claim 19, wherein prior to sequencing of step (IV):

(III') generating nanoball complexes comprising polycistronic amplified extended strands of tagmented DNA.