Chimeric DNA byproducts in strand displacement amplification using the T7 replisome

Dillon B. Nye, Nathan A. Tanner

Nucleic Acid Replication, New England Biolabs, Inc.

Abstract

Long read sequencing technologies motivate the development of DNA amplification methods capable of producing specific amplicons larger than those achievable by PCR. Here, we investigate isothermal strand-displacement amplification reactions using the T7 replisome, a macromolecular complex of a helicase, a single-stranded DNA binding protein, and a DNA polymerase. **Sequence analysis revealed chimeric DNA molecules corresponding to long inverted repeats, which in some cases accounted for the bulk of amplified DNA.** These chimeric reads originate from template switching at short, inverted repeat sequences within the amplicon. Nanopore sequencing revealed that template switching occurs in part through an intermolecular mechanism via a cruciform intermediate and highlighted a role for polymerase exonuclease activity in promoting template switching. These findings demonstrate challenges to targeted isothermal DNA amplification and guide development of novel amplification methods for long read sequencing.

T7 Replisome

The T7 replisome is a dynamic macromolecular assembly of four proteins

- T7 gp2.5, a single-stranded DNA binding protein
- T7 gp4, a hexameric helicase and primase
- **T7 gp5, a DNA polymerase**
- *E. coli* thioredoxin which binds T7 gp5 with nM affinity

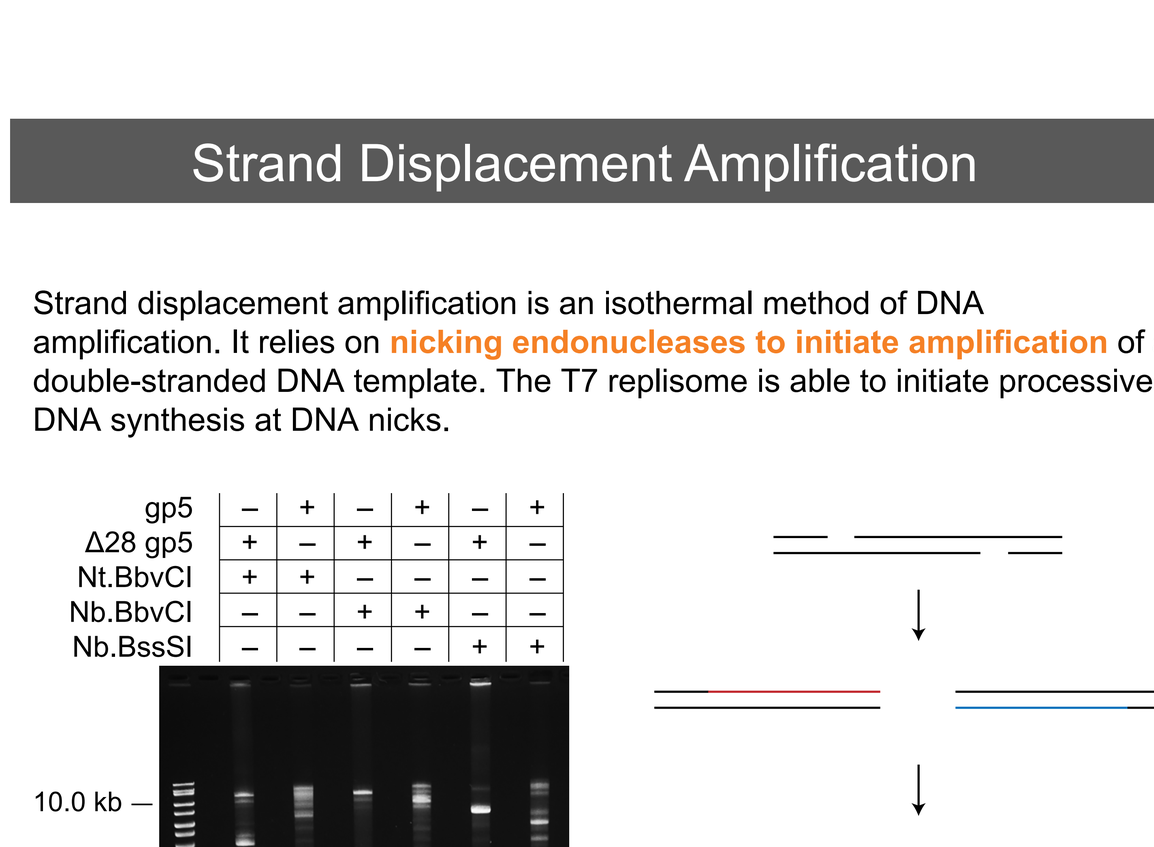

Figure 1. (Left) A schematic of coupled leading- and lagging-strand synthesis by the T7 replisome. (Right) A structural model of the gp4 / gp5-Trx complex from cryo-EM data (Gao, et al. 2019).

Strand Displacement Amplification

Strand displacement amplification is an isothermal method of DNA amplification. It relies on **nicking endonucleases to initiate amplification** of a double-stranded DNA template. The T7 replisome is able to initiate processive DNA synthesis at DNA nicks.

Figure 2. (Left) Agarose gel showing discrete amplicons in the reaction of a nicking endonuclease, λ bacteriophage genomic DNA, and either the **WT or exonuclease-deficient T7 replisome**. (Right) A scheme for linear strand-displacement amplification.

Nanopore Sequencing

Nanopore sequencing was used to investigate the mechanism of amplification. The 5’ coverage maps demonstrate the **amplification initiates at NEase recognition sites**

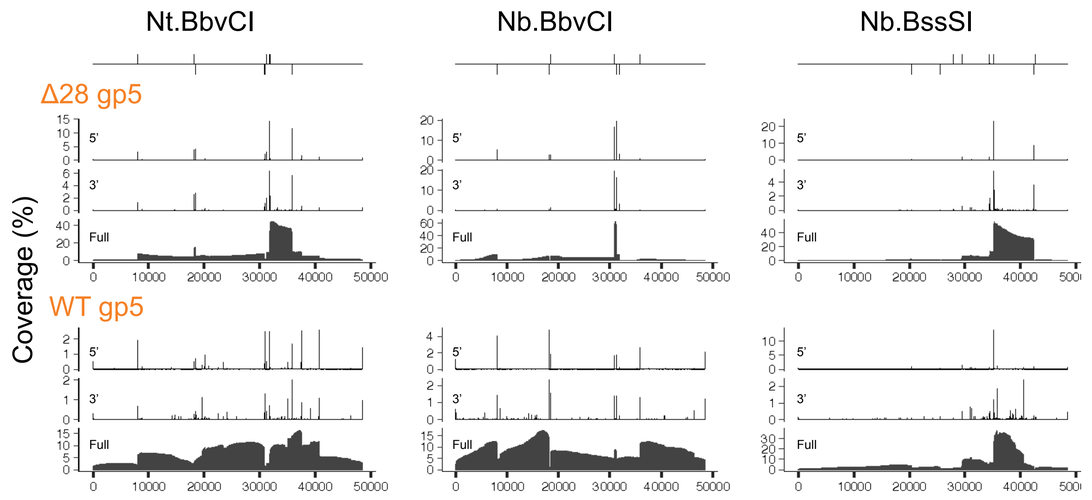

Figure 3. 5’, 3’ and full coverage maps determined by nanopore sequencing of amplicons produced by T7 SDA. At top are shown anticipated top- and bottom-strand nicks.

Plots of read length vs. alignment length demonstrate a **greater proportion of chimeric DNA products when the WT polymerase is used.**

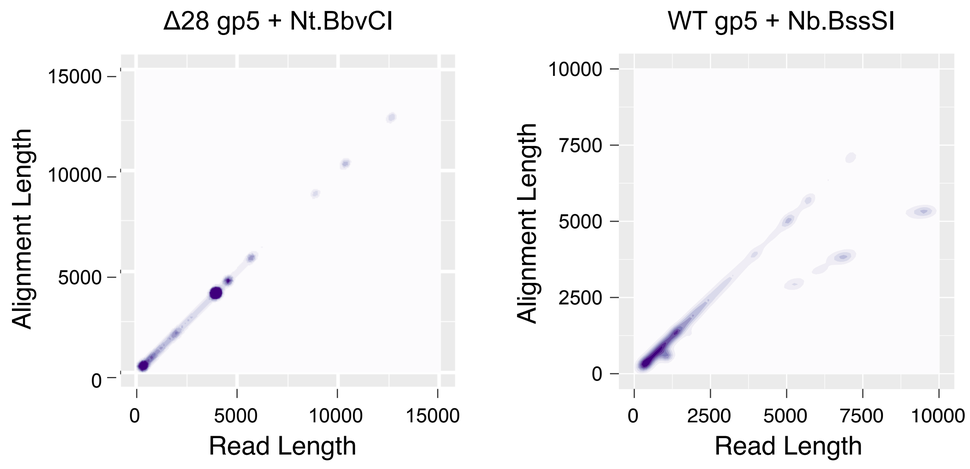

Figure 4. Plots of read length against a the length of the longest alignment made from that read to λ genomic DNA. Off-diagonal clusters of reads correspond to hairpin DNA molecules

Close inspection of 3’ coverage maps and individual reads demonstrates an **intermolecular branch-migration mechanism for hairpin production.**

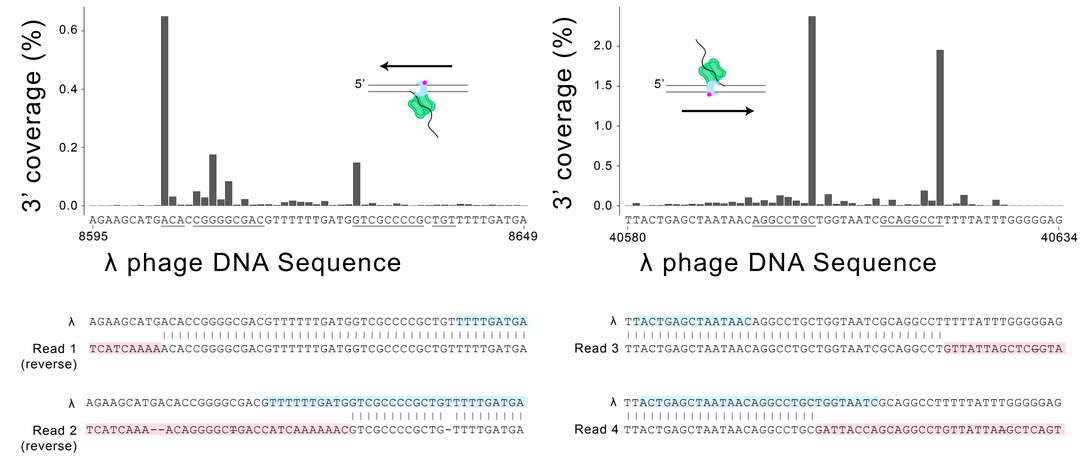

Figure 5. (Top) 3’ coverage maps highlighting inverted repeat sequences that promote template switching. (Bottom) individual reads that template-switch after the second instance of the repeat (Reads 1 and 3) or the first instance of the repeat (Reads 2 and 4).

Conclusions

The T7 replisome is a promising system for amplification of long DNA molecules. High-turnover, high-specificity nicking endnucleases will enable new amplification methods.

Isothermal amplification methods that produce long ssDNA intermediates often also produce chimeric byproducts.