

ABSTRACT

The efficiency of Next Generation Sequencing (NGS) depends on the quality of libraries used. Existing NGS sample preparation kits are not optimized for A-tailing, resulting in libraries that are not efficiently ligated to T-tailed adapters, are contaminated with chimeras, or require lengthy protocols to finish. In addition, because enzymes such as Taq DNA polymerase add non-templated purines (both A and G residues) to the 3' end of templates, these kits produce both A- and G-tailed templates, which also reduces the ligation efficiency to T-tailed adapters. We have developed an optimized system that repairs template ends and maximizes A-tailing and subsequent ligation to T-tailed adapters. Optimization was accomplished with an assay that allows one to quantify A- and G-tailing of target DNA by ligation to T- and C-tailed fluorescent adapters. Our 90 bp target DNA consists of a mixture of random 3' ends that must be end repaired and phosphorylated before self ligation is possible. A-tailing prevents self ligation and chimera formation. The results demonstrate an essentially chimera free target with over 80% efficiency of T-tailed adapters ligated in a single step compared to significant chimera formation and low T-tailed adapter ligation (7% to 48%) with other systems. We confirmed the efficiency of our method by constructing and sequencing Ion Torrent random shear genomic libraries. Sequence results show that Lucigen reagents produce a more complex, less biased library with fewer chimeras than libraries produced with competitor's kits.

METHODS AND RESULTS

Figure 1. Diagram of Generalized Sample Preparation for NGS Libraries. Sample preparation for the majority of NGS platforms follows a similar protocol. Genomic DNA is sheared to an appropriate size, end repaired, and in most cases A-tailed. Adapters are ligated and the final product is purified or size selected. If needed, the library will be subjected to limited amplification by PCR.

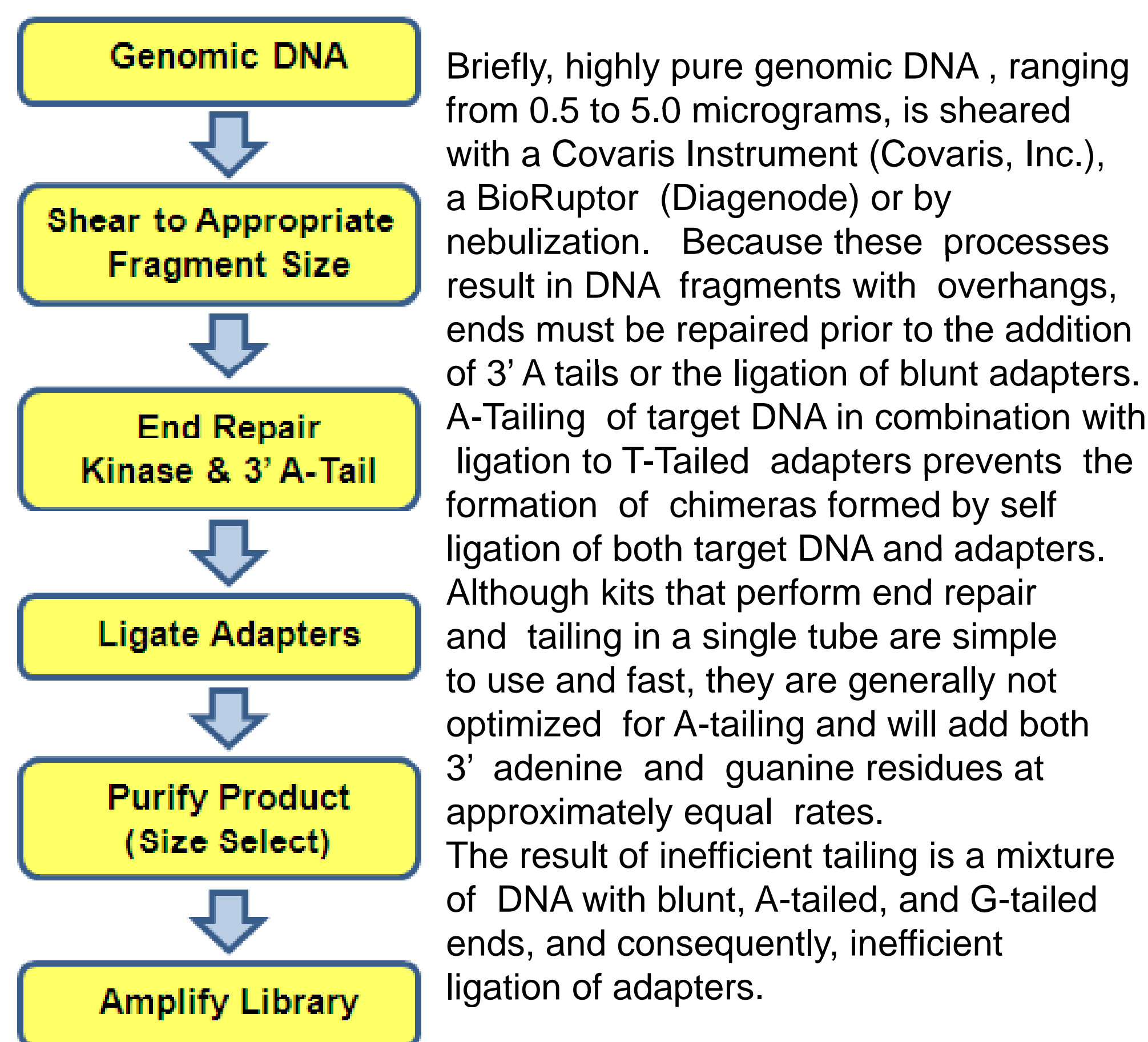


Figure 3. Blunt versus Tailed Ligation.

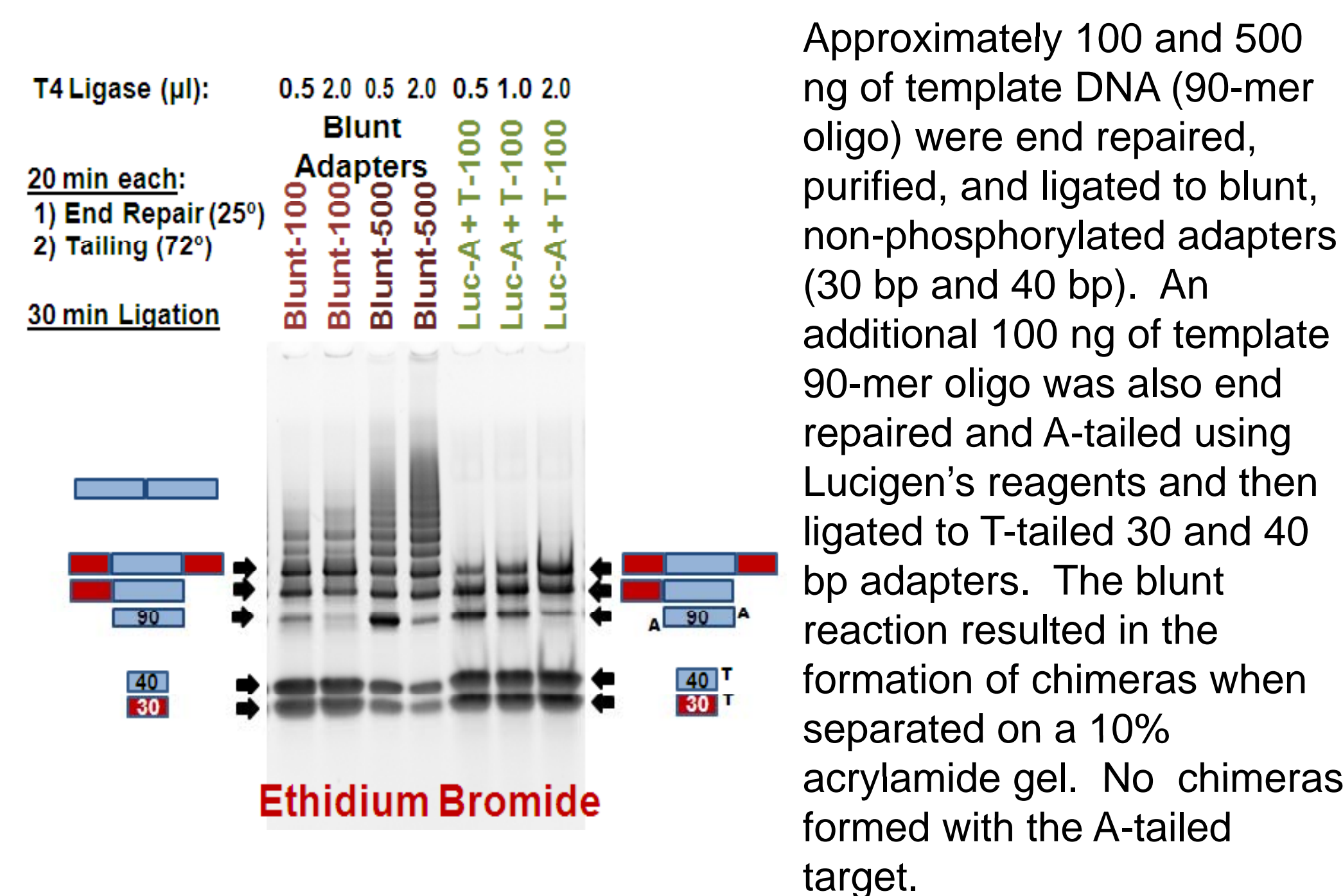


Figure 4. T4 Ligase Titration.

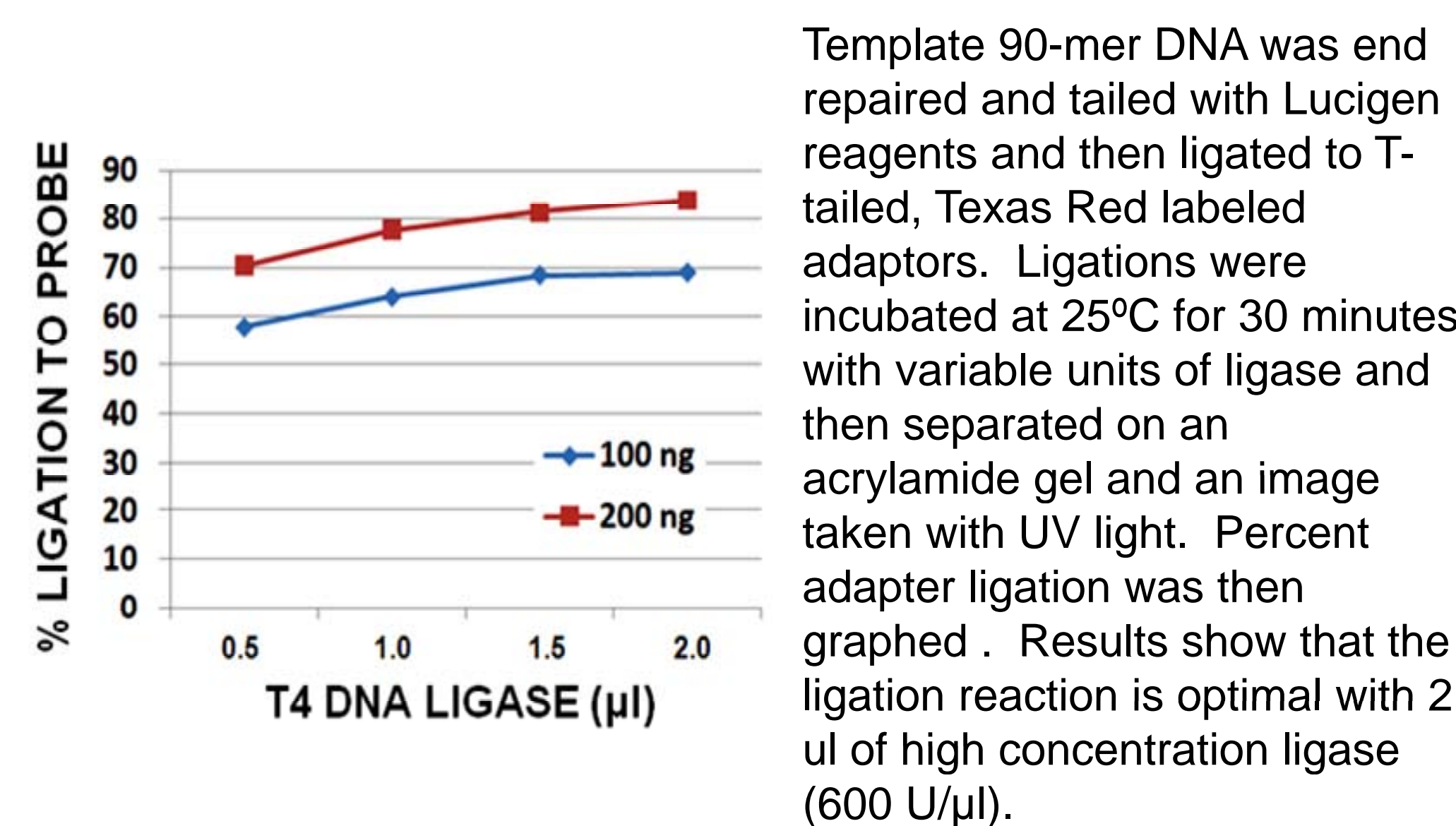


Figure 5. Oligo Assay: Comparison of Lucigen Reagents with competitor N and R reagents.

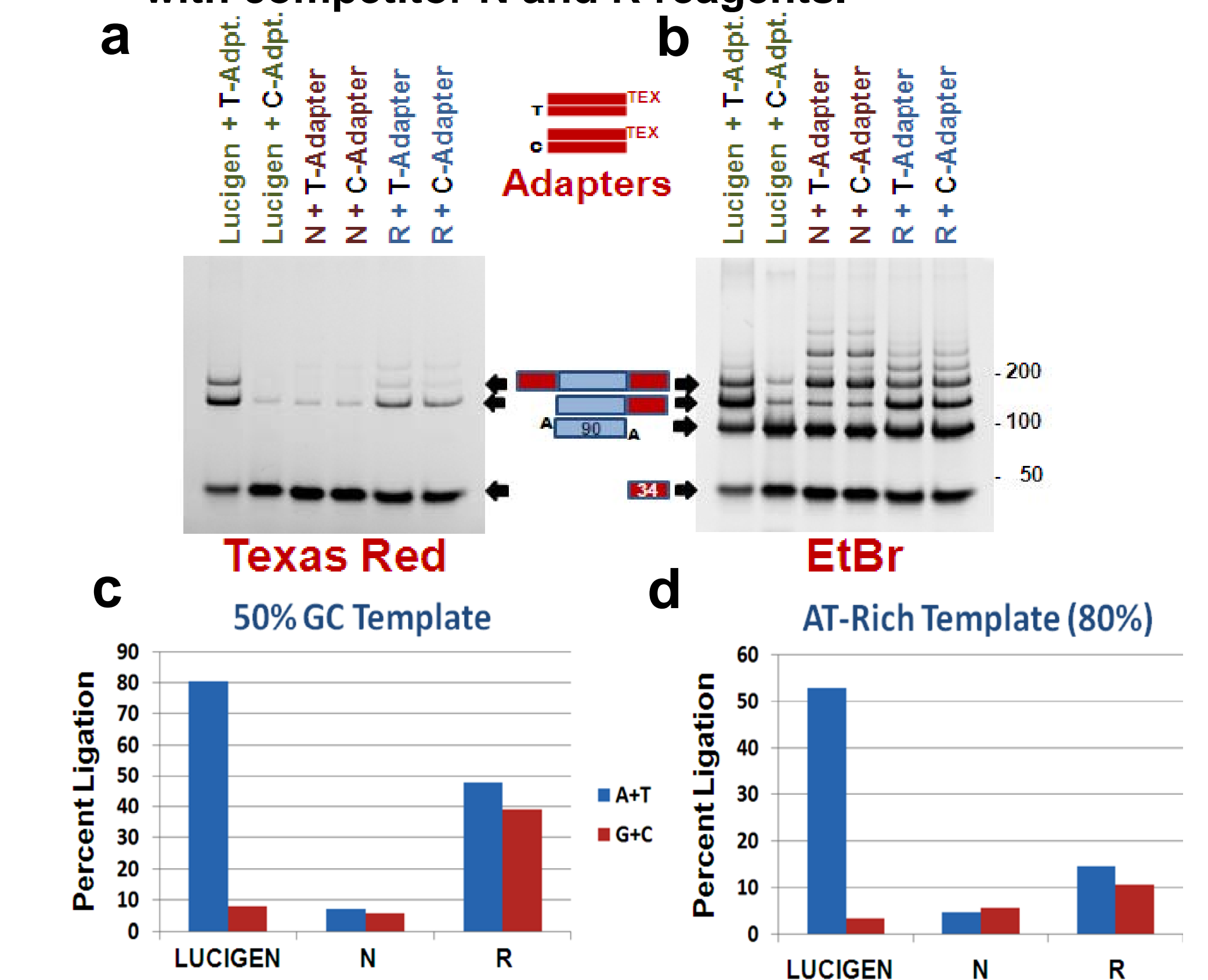
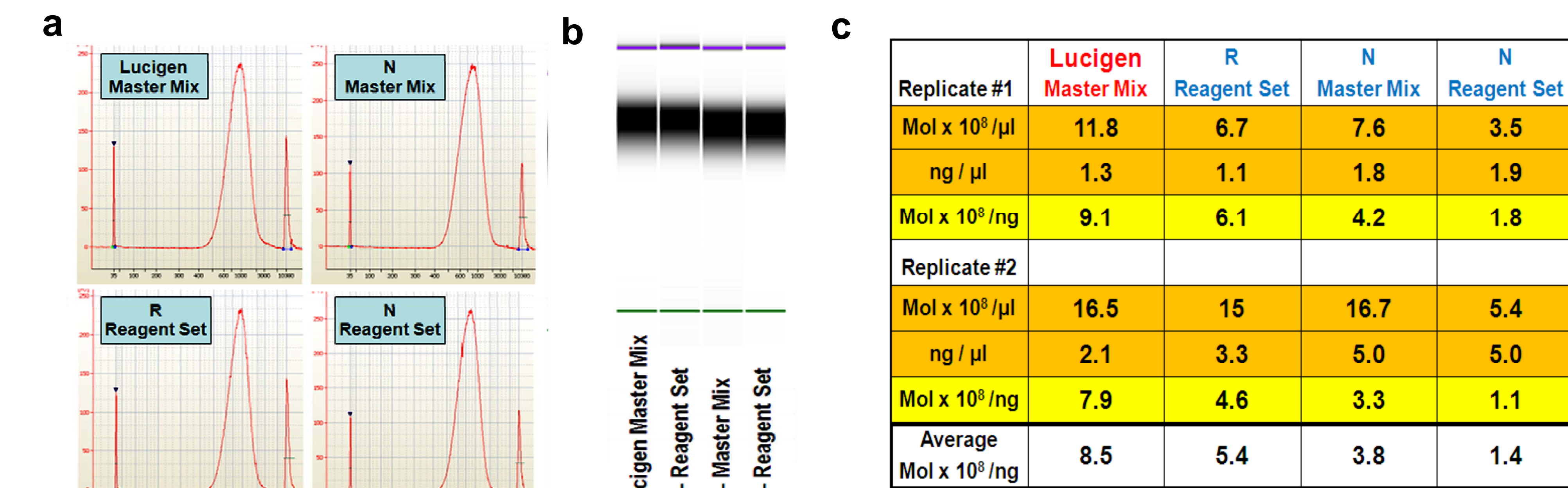


Figure 6. 454 NGS Library: Comparison of Lucigen's reagents with competitors N and R's protocols and reagents.



Genomic DNA was sheared and 500 ng each used in Lucigen's and competitors N and R's sample preparation processes. 25 µl End Repair and Tailing reactions were followed by the addition of FAM-labeled adapters and T4 DNA ligase. After ligation, samples were purified with SPRI beads and sizing solution and quantified with an Agilent Bioanalyzer. Results show that Lucigen's reagents and protocol produce more labeled product (labeled molecules per ng DNA) than that produced by competitors N and R. For these comparisons, all reagents used were from individual kits except for adapters, which were the same for all experiments. Fluorescence (Mol/µl) was measured with a Synergy 2 Microplate Reader using a standard from kit 'R'. Quantification of DNA was made on an Agilent 2100 Bioanalyzer with High Sensitivity chips. Data includes graphic (a), pseudo images (b) of completed libraries and quantification of library fluorescence (Mol/µl) and DNA (ng/µl) (c). Fluorescence per ng DNA (Mol/ng) was calculated by dividing values for Mol/ul by ng/ul. The average of two replicates is shown at the bottom.

Figure 2a. Schematic of Oligonucleotide Tailing Assay. Our end repair, tailing assay allows us to quantify the efficiency of tailing and differentiate between A-tailing and G-tailing, as Taq DNA polymerase will add non-templated purines (both A and G). In the absence of tailing, a ladder of chimeras will be visible indicating blunt ligation. Target DNA consists of a double strand 90-mer that must be end repaired and kinased before it is ligatable. Adapters are labeled with Texas Red and have either a T-tail or a C-tail for ligation to A-tailed target or G-tailed target, respectively. Target that is end repaired but not tailed will form chimeras when ligated. Target that is G-tailed will ligate to the C-tailed adapter. For analysis of blunt end ligations, we used a mixture of two adapters, a 30 bp FAM-labeled adapter and a 40 bp unlabeled adapter (see Fig. 3).

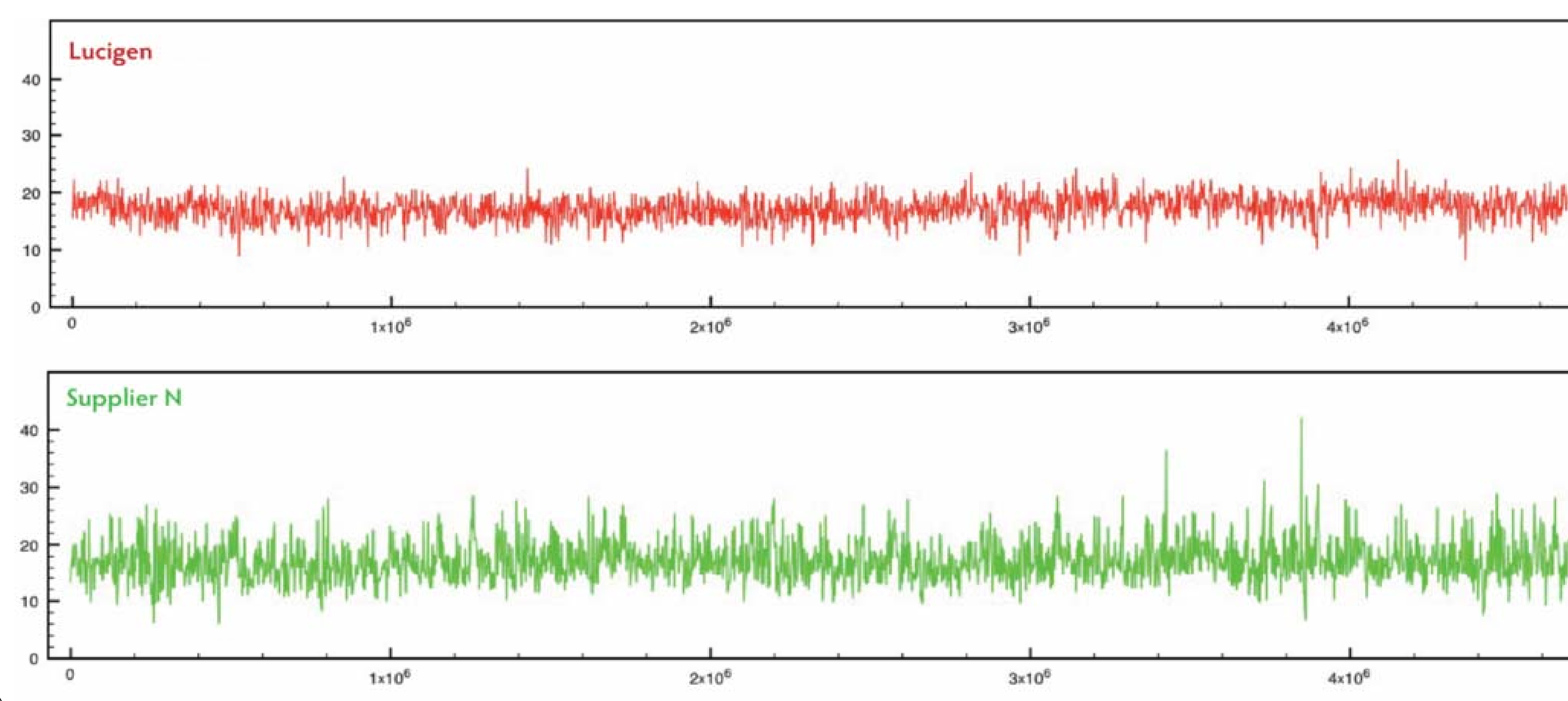
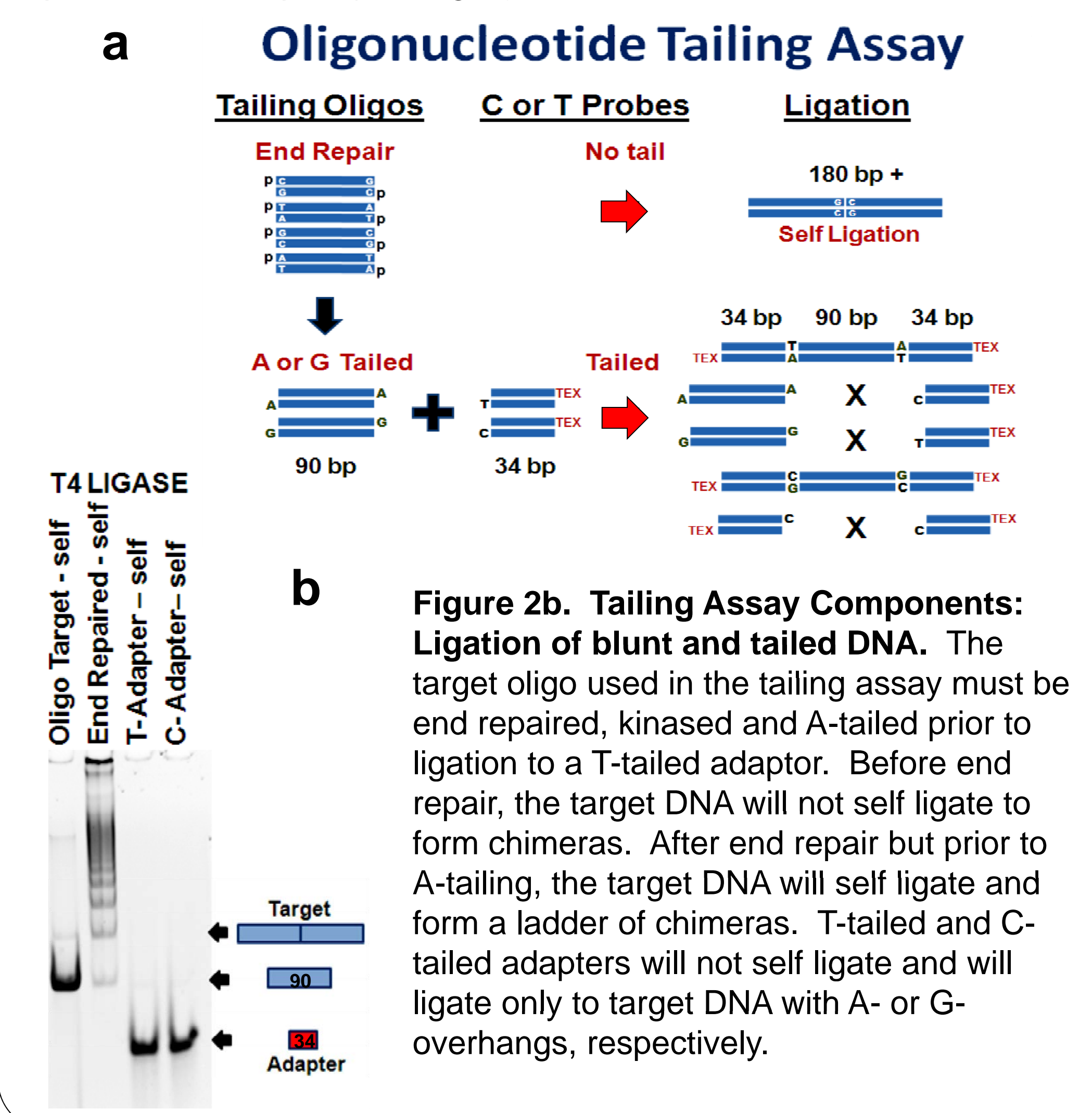


Figure 7. Ion Torrent Sequencing

Comparison of sequence bias. Genomic DH10B *E. coli* DNA was sheared and prepared for NGS library creation per manufacturers recommendations. A tighter line with less deviation from the mean represents the library with lower bias and thus more accurately matching the reference sequence.

CONCLUSION

We show here that optimization of our reagents and procedures resulted in significant improvement in single tube end repair and A-tailing of template DNA as compared to competitor reagents and procedures. Un-optimized reagents and conditions can result in:

- 1) Incomplete A-tailing
- 2) Reduced ligation of adapters
- 3) Increased G-tailing which prevents ligation to T-tailed adapters;
- 4) The formation of chimeras upon ligation.

Our optimized reagents can result in a more efficient library and improved sequence information.

ACKNOWLEDGEMENTS

This work was partly supported by an NHGRI grant to CCW.

