APPLICATION NOTE

Same-Day Sample to Enriched Sequencing-Ready Library using LoopCap™

A PROOF-OF-CONCEPT PERFORMANCE EVALUATION

Summary

Targeted next-generation sequencing (NGS) has been widely adopted in research and clinical settings due to the cost-effective and time-efficient benefits of only sequencing specific regions of interest when compared to genome-wide analyses. Despite this benefit compared to whole-genome and whole-exome sequencing, the protocols to prepare samples for targeted sequencing typically require multiple days, with hybridization times of ~16 hours being standard for most commercial providers.[1] Fast hybridization has gained prominence as a method to streamline targeted sequencing workflows, allowing for quicker turnaround times and efficient utilization of laboratory resources. LoopCap technology offers an alternative to complex targeted sequencing protocols, with less than 75 minutes of hands-on time required to take a sample from extracted DNA to a sequencing-ready library. To assess the impact of reduced hybridization times on the performance of the LoopCap target capture technology, we conducted a proof-of-concept evaluation using an internally curated hereditary oncology panel. This application note presents the outcomes of our initial proof-of-concept study, shedding light on the performance characteristics of a LoopCap-based targeted panel under varying hybridization durations.

molecular loop

Methodology and Experimental Design

For our evaluation, libraries were constructed from 150 ng of high molecular weight human genomic DNA (gDNA) (Sigma-Aldrich 11691112001) with the LoopCap DNA Target Capture Kit. We utilized an internally designed 160.5 kb hereditary oncology panel, and tested hybridization times of 16, 4, 2, and 1 hr, which represents total workflow times from sample to sequencing-ready library of 19, 7, 5, and 4 hrs, respectively. Each condition was replicated eight times for robust statistical analysis. Bead cleanup of the final library pool was performed using a ratio of 0.67X, and the PCR cycle number was gradually increased from 18 to 21 cycles with decreasing hybridization time. Sequencing was conducted on the MiniSeq platform using 2 x 151 paired-end reads, followed by downsampling to 1M reads per sample. Sequencing data were analyzed according to Molecular Loop's *Data Analysis User Guide v.3* (available upon request).

Results and Discussion

We observed mean deduplicated coverage between 350X and 370X for all samples. Coverage completeness (defined as the percentage of target covered at ≥25X, ≥50X, and ≥100X) and the percentage of target at zero coverage was assessed for all time points.

The percent minimum coverage over target regions was consistent across all hybridization times and coverage depths, with >94% of bases covered at ≥100X for hybridization times as short as 4 hours (Figure 1). In the 16-hour condition, 0% of target bases were left uncovered and <0.1% of bases were left uncovered in the three other timepoints (0.03%, 0.06%, and 0.05% in the 4 hr, 2 hr, and 1 hr, respectively), all well within the 99.9% coverage needed for a majority of somatic and inherited disease applications. As this panel is being used as a proof-of-concept, and was not optimized for shorter hybridization times, we believe we can improve our coverage even further with denser tiling of LoopCap probes.

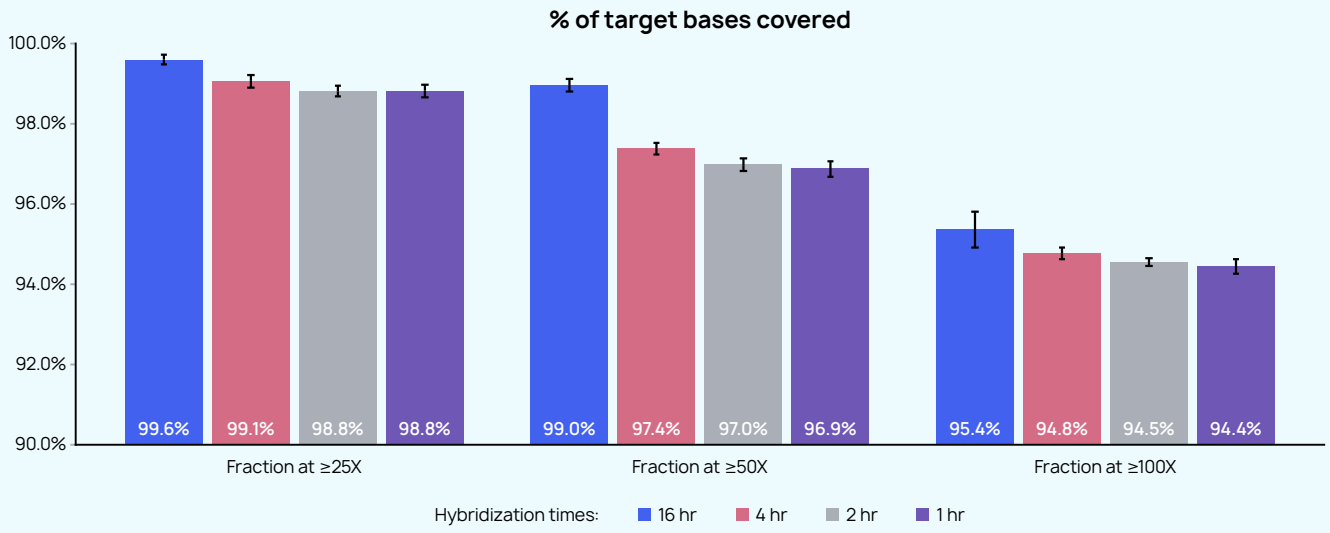

Figure 1. Shortened hybridization time shows minimal impacts on coverage completeness. Shortening the hybridization time does not appear to substantially affect the percentage of target bases covered at ≥25X, ≥50X, or ≥100X.

To assess the overall efficiency of sequencing, we evaluated on-target read rates, coverage uniformity (fold-80 base penalty), and PCR duplicates (Figure 2).

- **On-target read rates** (the fraction of reads that map to target regions) were similar between the hybridization times tested, with on-target read rates of 86 – 88% observed for all four time points.
- **Coverage uniformity** (how evenly target regions are represented in the sequencing data) was evaluated with the fold-80 base penalty, defined as the fold change of additional sequencing necessary to raise 80% of target bases to the observed mean coverage. All four hybridization times showed low (<1.7) fold-80 base penalty scores, with no material difference observed between the four timepoints.
- **PCR duplicate rates** (the percentage of reads that were removed by UMI-based deduplication) were evaluated as a measure of library complexity. Unlike coverage uniformity and completeness, we observed a mild increase in the PCR duplicate rate as hybridization time decreased, with the 1-hour protocol resulting in a 4.8% duplicate rate. While this trend is indicative of reduced library complexity with decreased hybridization times, the values observed for the short hybridization times were all still well within the ~2 – 10% duplicate rate range typically observed in germline targeted sequencing,[2] and could be further optimized by increasing input DNA amount and minimizing the number of PCR cycles.

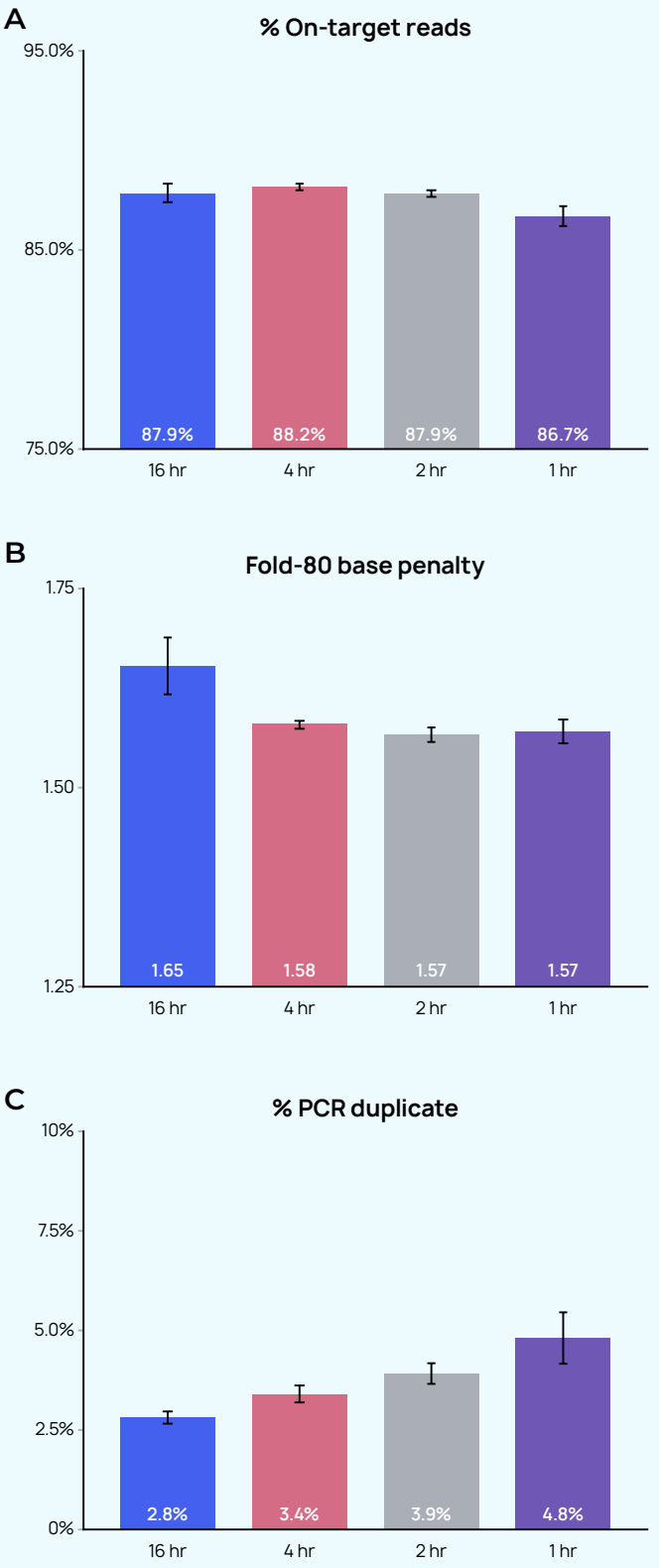

Figure 2. Hybridization time appears to have minimal impact on the efficiency of sequencing (% on-target, % zero coverage bases, and fold-80 base penalty) but does seem to affect library complexity (% PCR duplicate reads). **(A)** The percent of on-target reads is the percent of mapped, non-duplicate reads overlapping the target region by at least 1 base. **(B)** Coverage uniformity expressed as fold-80 base penalty. **(C)** Percent of duplicate reads removed.

Conclusion

Target enrichment NGS workflows will continue to play an important role in the proliferation of NGS into more laboratory settings. Allowing flexibility in hybridization time provides laboratories with another tool to maximize the efficiencies of these traditionally lengthy protocols.

This proof-of-concept study shows that there is significant promise in achieving a true single-day, single-shift target enrichment workflow using LoopCap. The post-hybridization portion of the LoopCap workflow is only 3 hours, enabling the preparation of sequencing-ready libraries directly from extracted samples in as little as 4 hours with a 1-hour hybridization. This is in stark contrast to the multi-day library preparation and target enrichment protocols needed to prepare a sample for targeted sequencing using a Hybridization Capture workflow.

We observed substantially equivalent performance with all four hybridization times tested (16 hr, 4 hr, 2 hr, and 1 hr), with similar coverage completeness and uniformity, and a slight increase in zero-coverage bases and PCR duplicates as hybridization times decreased. For zero-coverage bases, we believe we can utilize the unique tiling density feature of our LoopCap chemistry to include panel modifications for denser tiling of our LoopCap probes in uncovered regions in future development efforts. For the duplicate rate—despite a trend of increasing rates with decreasing hybridization time—the duplicate rate was below 5.0% across all hybridization times, and we believe that with additional protocol modifications we can further improve the performance of this panel.

When combined with our proven ability to design targeted germline panels ranging from 60 kb – 700 kb without affecting performance—as demonstrated in our previous application note *Custom Enrichment for Germline Targeted Sequencing with LoopCap™ DNA Target Capture Technology*—we believe that this protocol can be applied to any germline application and continue to show equivalent performance to the panel tested in this study.

We are excited to share this initial proof-of-concept data that supports the use of LoopCap technology to substantially decrease the total target enrichment workflow time—from ~19 hours to ~4 hours by performing a 1-hour hybridization with minimal impact to panel performance—which offers a compelling turnaround time advantage over existing technologies for many NGS application areas.

References

1. Pel, J. et al. Rapid and highly-specific generation of targeted DNA sequencing libraries enabled by linking capture probes with universal primers. *PLoS One* 13.12 (2018): e0208283.
2. Bansal, V. A computational method for estimating the PCR duplication rate in DNA and RNA-seq experiments. *BMC Bioinformatics* 18.3 (2017): 113 – 123.

molecularloop

Molecular Loop Biosciences, Inc.
300 Tradecenter Drive, Suite 5400, Woburn, MA 01801

sales@molecularloop.com
www.molecularloop.com

P101 MLAN101_v2-1023