RENEW

White Paper

Direct Targeted Methylation Sequencing Feature Characterization

Summary

This study characterizes Renew Biotechnologies' (Renew) Direct Targeted Methylation Sequencing (dTMS) assay for integrated detection of genetic variants and DNA methylation within defined target regions. Using Agilent SureSelect hybrid capture technology, proprietary Renew library preparation methods, and native long-read nanopore sequencing, we enriched targeted loci and evaluated performance in established reference DNA. The assay delivered robust on-target efficiency and sufficient coverage to accurately resolve SNPs, STRs, haplotypes, and regional methylation in a single workflow, supporting scalable, targeted multi-omic applications.

Introduction

Native long-read sequencing technologies, including Oxford Nanopore Technologies (ONT) and Pacific Biosciences (PacBio), have transformed genomics by resolving repetitive and structurally complex regions of the human genome that short-read sequencing technologies have historically missed (1-3). For many applications, however, whole-genome sequencing remains unnecessary and inefficient.

As long-read platforms have matured, multiple targeted strategies have emerged to concentrate sequencing depth to regions of interest while, when possible, preserving long-range genomic context and native epigenetic information. These include long-range PCR amplicon sequencing, CRISPR/Cas9-mediated enrichment approaches, and ONT adaptive sampling (4). While powerful, each presents tradeoffs.

Long-range PCR is rapid and requires lower input, but erases native methylation and other epigenetic signatures during amplification. CRISPR-Cas9-based enrichment preserves methylation but is typically limited in multiplexing capacity and requires substantial DNA input. Adaptive sampling offers programmable, real-time enrichment without additional wet-lab steps but relies on high-speed basecalling and capable computational hardware to maintain accuracy. Because targeting decisions are made within the first few hundred bases, adaptive sampling performance is sensitive to fragment length, reference design, and target proportion, and may be affected by pore blockages and setup complexity, often limiting coverage (~20-30x) and multiplexing capacity (5).

Renew Biotechnologies (Renew), a certified Oxford Nanopore Technologies (ONT) sequencing service provider, offers comprehensive cDNA sequencing to analyze full-length cDNA transcripts. Leveraging ONT's long-read sequencing technology, this service enables the detection of complete gene isoforms, splice variants, and fusion genes. Designed

Hybridization-based enrichment addresses many of these limitations. However, because they were optimized for short-read platforms, conventional hybrid capture workflows produce short reads and do not preserve native epigenetic modifications due to PCR amplification of the enriched fragments. Our recently developed Direct Targeted Methylation Sequencing (dTMS) extends hybrid capture into the long-read domain by integrating modified Agilent SureSelect hybrid capture with native long-read nanopore sequencing. This enables high-depth, multiplexed enrichment across discontinuous genomic regions while retaining native methylation information and long-range haplotypic context.

This white paper describes the characterization of dTMS, outlining enrichment efficiency, variant detection accuracy, repeat resolution, phasing performance, and methylation concordance. By combining the scalability of hybridization-based targeting with the structural and epigenetic advantages of native long reads, dTMS is designed to support high-confidence, multi-omic profiling for research and translational assay development.

Methods

Sample types and reference material

To evaluate assay performance, we used the HG002 Genome in a Bottle (GIAB) consortium (NIST) cell line, which has a publicly available and well-characterized human genome generated using nanopore sequencing (6) as a benchmark for variant and methylation accuracy. In addition, the fully CpG-methylated control cell line (HCT116; Takara Bio) was included to assess methylation detection performance and signal fidelity.

SureSelect panel designs

Target enrichment was performed using Agilent SureSelect hybridization-based capture technology.
Two panels were selected to assess performance across diverse genomic features, including single-nucleotide variants, repeat regions, and differentially methylated regions:

1. A developing short tandem repeat (STR) panel (Agilent x ONT) that infers STR sequence from flanking regions.
2. Agilent's SureSelect MethylSeq DMR probe set for targeted methylation analysis.

Library prep and sequencing

gDNA was extracted and processed following Renew's dTMS workflow, which incorporates a proprietary modified SureSelect hybridization protocol to enrich regions of interest. Captured libraries were sequenced on the ONT PromethION platform.

Post-sequencing processing and analysis

Raw sequencing reads were aligned to the GRCh38 human reference genome using minimap2. and general alignment metrics, including read length distribution and coverage, were assessed to evaluate sequencing performance. Target-specific quality control analyses were performed to assess enrichment efficiency and probe behavior. In cases where probe-derived contamination peaks exceeded 10% of pulldown reads, a minimum read-length filter of ≥250 bp was applied to remove short probe-derived fragments and improve downstream variant-calling accuracy.

Variants, including single-nucleotide variants (SNV), copy number variants (CNV), and larger structural variants (SV), were called using ONT's wf-human-variation workflow (7), incorporating Clair3 for diploid variant detection. Variant calls were benchmarked against the HG002 Genome in a Bottle truth sets to assess precision and recall. Modified base detection for methylation was performed using Modkit within the same workflow.

STRs were analyzed using StragIR (8) supplemented with curated PacBio STR reference alignments (9) and Agilent-ONT STR target lists to expand locus coverage beyond the standard workflow. Methylation accuracy was evaluated by comparing Modkit-derived methylation calls to the HG002 Nanopore methylation reference dataset, enabling assessment of both single-CpG and regional methylation concordance.

Results

Sequencing Statistics and Quality Characterization

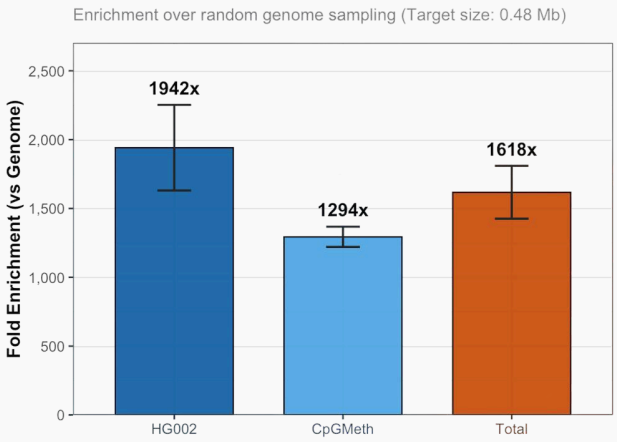
Across the targeted features, comprising 0.47 Mb (~0.5 Mb) of the human genome (Table 1), dTMS achieved stable sequencing across runs and exhibited 20-30% on-target efficiency, corresponding to >1,200x fold target enrichment (Figure 1A, 1B). N50 values of approximately 1.0-1.1 kb were maintained, preserving roughly 50kb of phased sequence per sample for haplotype analysis within target regions.

As anticipated, coverage depth varied by target feature, reflecting panel design and genomic composition (Figure 1C). The methylation panel achieved the highest median coverage (~200-250x), consistent with its larger genomic range (~320 kb compared to ~45 kb). STR loci demonstrated moderate coverage (~100-200x), supporting accurate genotyping and repeat sizing. Phasing and haplotype regions showed coverage comparable to SNV detection (~100-150x), confirming sufficient depth for phase-based variant calling.

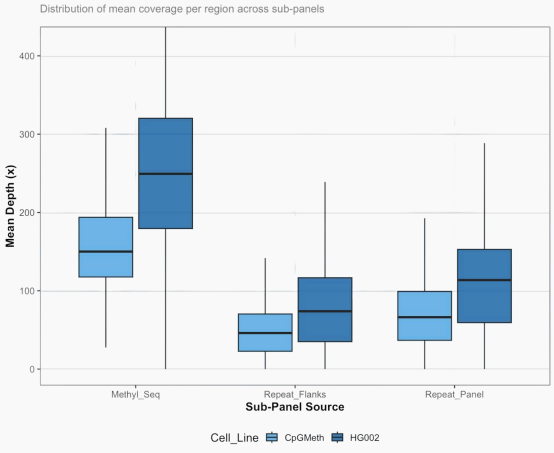
Table 1: dTMS probe set sizes

Region	Size
STR Repeat	2,363 bp
STR Flanking Regions	44,852 bp
MethylSeq Targets	320,480 bp
Total	467, 695bp (~468kb)

A Targeted Fold Enrichment



B Sequencing Depth by Panel Source (Q20 Reads)



C Coverage Suitability by Analysis Type

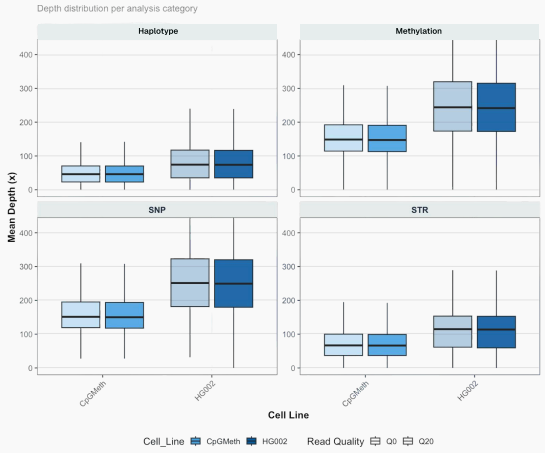


Figure 1. Performance across targeted panels. When normalized to panel size relative to the human genome, total enrichment ranged from approximately 1,200x to 1,900x (A). Analysis of individual panels by cell type revealed substantial variability in mean depth, with some targets achieving greater than 300x coverage (B). Coverage suitability by analysis type (C).

Analyses By Target Types

Single-nucleotide variants

SNVs were evaluated using a long-read HG002 reference dataset (ONT) to benchmark performance. After applying a minimum depth filter of >20x, SNV detection achieved F1 scores exceeding 99% (Figure 2), demonstrating high-confidence variant calling within adequately covered regions.

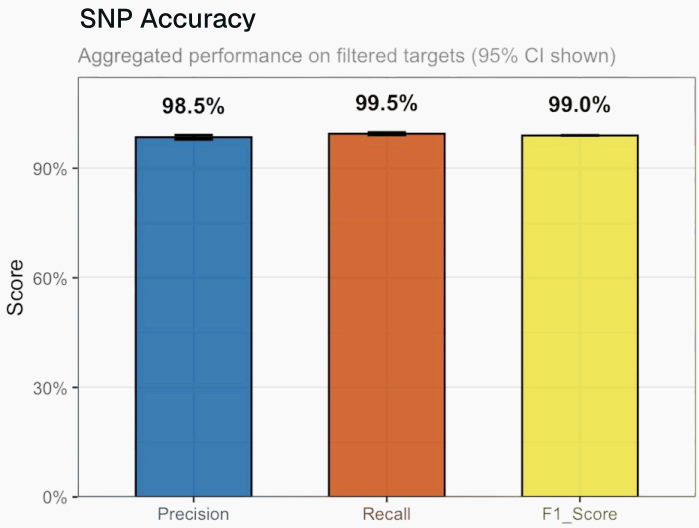
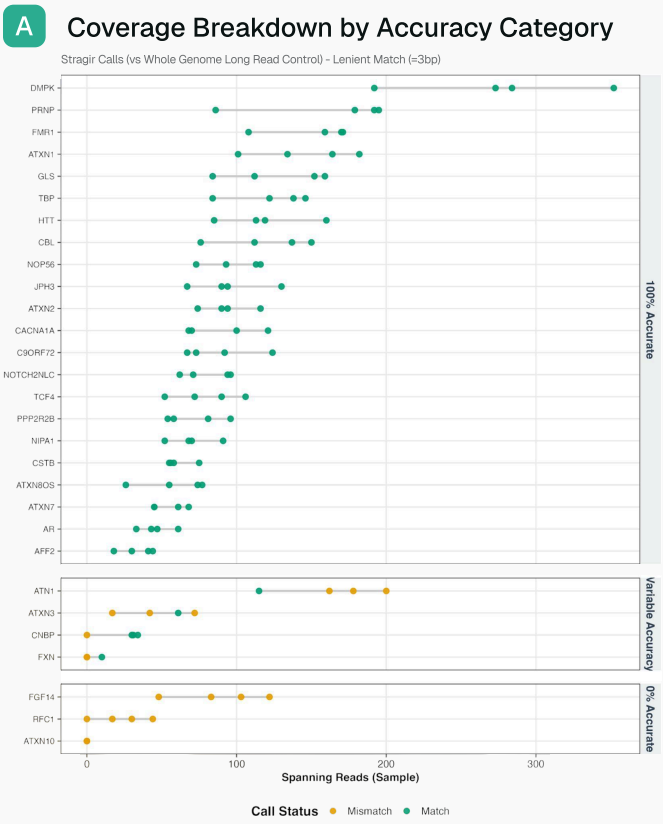


Figure 1: SP comparison statistics post filtering from ONT HG002 reference dataset. MethylSeq sites reflect > 98% accuracy after filtering for low depth or poor mapping quality.

Short Tandem Repeats (STR)

STR performance was also evaluated using HG002 benchmark references. Targeted loci achieved up to ~200 spanning reads per region, enabling direct repeat sizing from long-read data (Figure 3A). Applying a >20x spanning-read depth filter, consistent with the SNV analysis, increased overall STR accuracy to 89% (Figure 3), indicating that >20x repeat-spanning coverage supports reliable repeat inference in targeted long-read datasets.



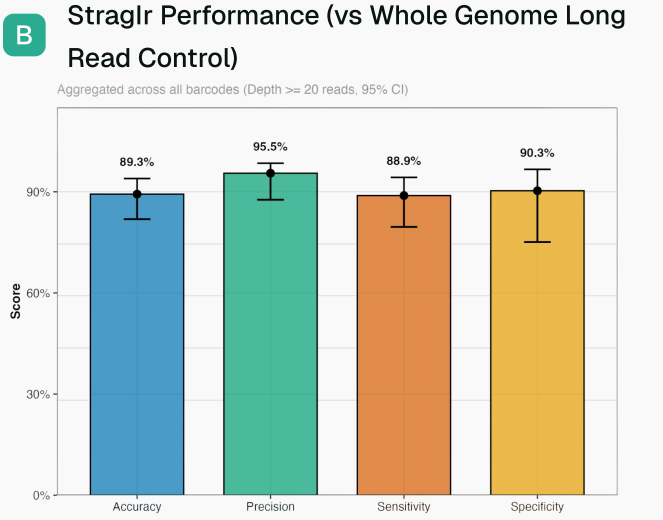
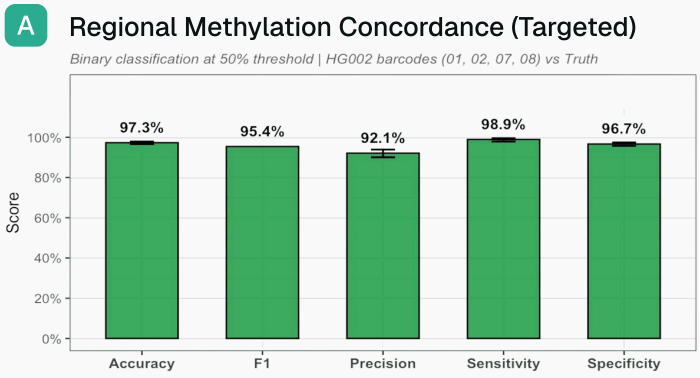


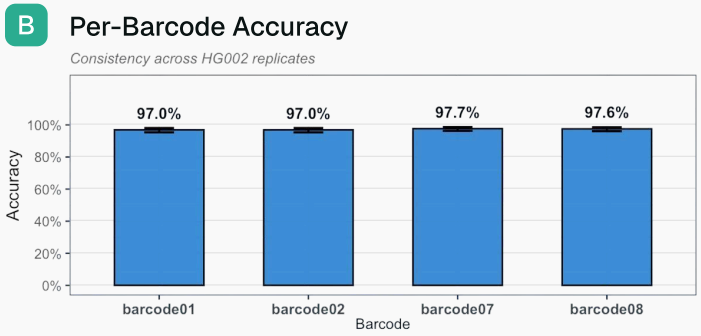
Figure 2. STR accuracy improves with increased spanning-read depth. dTMS STR data was compared to the HG002 PacBio reference dataset. Analyzed via Strag R, regions with >20 spanning reads demonstrated higher sizing accuracy (A). Filtering out loci with fewer than 20 spanning reads yielded accuracy >89% (B).

Methylation (5mC)

Methylation performance was evaluated at both regional and single-CpG resolution in HG002 samples relative to the whole-genome reference datasets. While single-CpG analysis showed modest reductions in sensitivity in mixed-methylation contexts, regional methylation profiling achieved an average accuracy of approximately 97% (Figure 4A) and was highly reproducible across replicates (Figure 4B). Strong concordance with the nanopore reference (R2 ≥ 0.99, Figure 4C) demonstrates reliable detection of broader epigenetic shifts across targeted regions.

Representative IGV visualizations further illustrate integrated variant and methylation resolution across repetitive and promoter regions (Figure 5). Together, these findings confirm that regional methylation patterns are robustly preserved within the dTMS workflow and identifiable in conjunction with genomic variants.





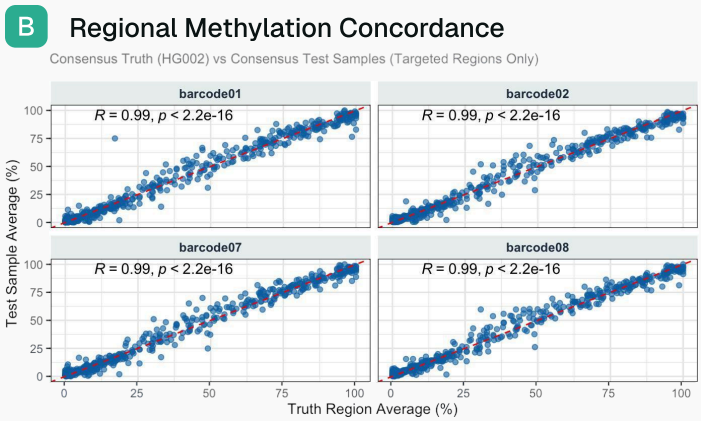


Figure 4. Regional CpG accuracy. Regional methylation accuracy was assessed by comparing HG002 whole-genome nanopore reference data to targeted samples. Overall regional accuracy reached 97%, with higher precision at the regional level (92%) than at single-CpG resolution (A,B). Concordance of regional methylation values was high (R2 ≥ 0.99) (C).

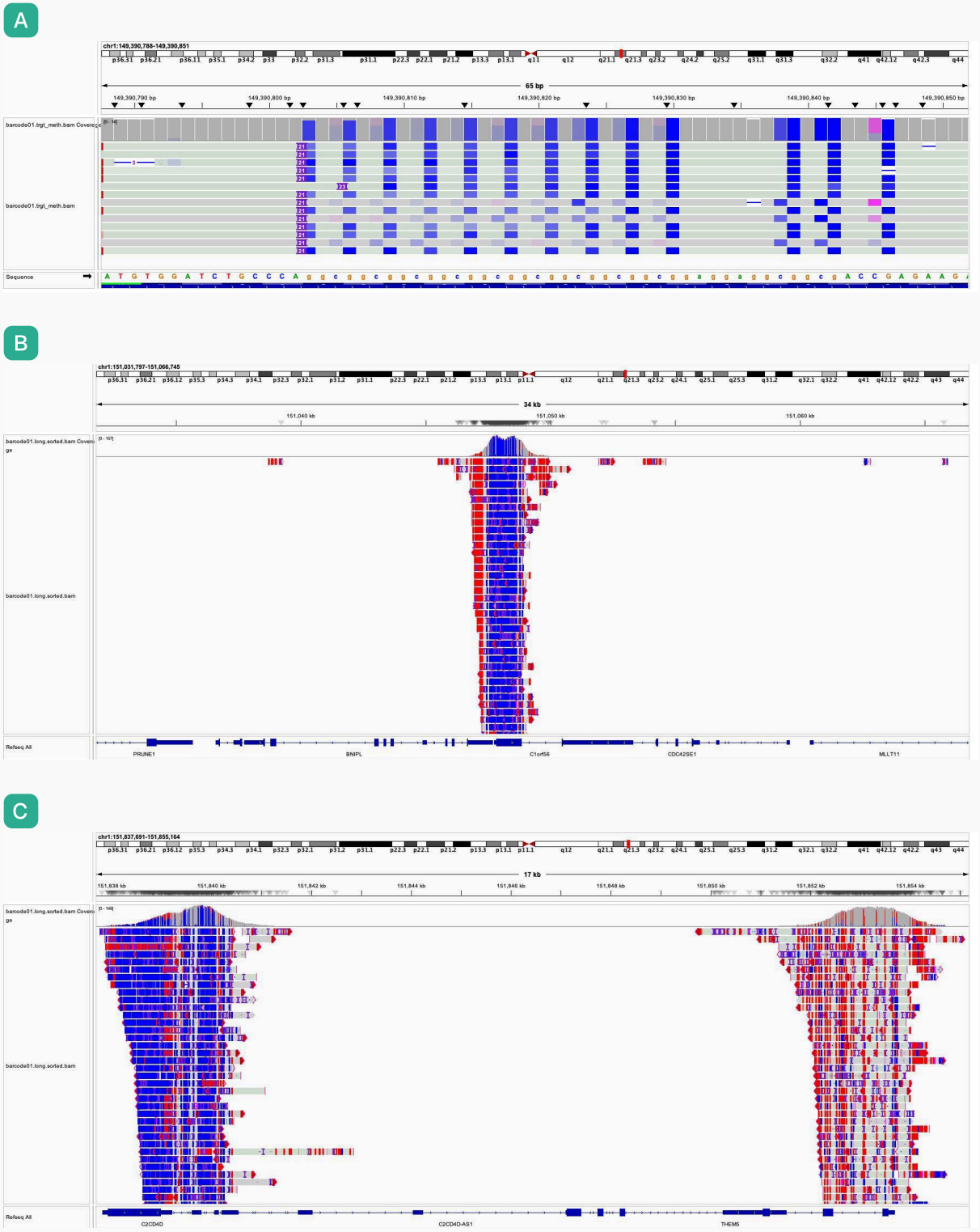


Figure 5. Representative IGV visualizations of dTMS multi-omic enrichment and methylation profiling. STR expansion at the NOTCH2NLC locus showing indels (purple) and unmethylated CpGs (blue) within the CGG repeat region (A). Promoter enrichment demonstrating an unmethylated core flanked by methylated regions (B). Two closely spaced loci with distinct methylation states, one highly methylated and one unmethylated, captured with high depth and minimal off-target enrichment (C).

Conclusions

As long-read sequencing continues to demonstrate value for applications resolving structural variation, repetitive loci, haplotypes, and native epigenetic modifications, scalable and cost-effective targeted solutions will be essential for broader research and translational adoption.

dTMS represents an emerging hybridization-based long-read enrichment platform designed to deliver high-depth variant detection while preserving native methylation context. In this characterization study, dTMS achieved >1,200x fold enrichment across target regions in GIAB HG002 and CpG Meth reference samples, supporting high-confidence detection of SNVs, STRs, haplotypes, and regional methylation. Performance was particularly strong in complex and repetitive genomic regions where short-read and PCR-based methods frequently underperform.

While comprehensive resolution of large CNVs and extended structural variants will require longer fragments than currently supported in our protocols, ongoing laboratory and bioinformatic optimization efforts are expected to expand assay capabilities.

Within the broader targeted long-read landscape, hybridization-based enrichment offers distinct advantages for scalable, multiplexed, and clinically deployable deep sequencing of discontinuous genomic regions. As such, dTMS provides a practical foundation for the development of the next generation of targeted panels, supporting oncology, rare disease, and other translational research applications requiring integrated genetic and epigenetic insight.

To learn more about Renew's NESSI-Seq technology and how it can improve your cfDNA studies

Visit renewbt.com

Contact us at info@renewbt.com

Reference List

1. Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze AV, Mikheenko A, et al. The complete sequence of a human genome. Science. 2022;376:44-53.
2. Iyer SV, Goodwin S, McCombie WR. Leveraging the power of long reads for targeted sequencing. Genome Res. 2024;34:1701-18.
3. Kanagawa T. Bias and artifacts in multitemplate polymerase chain reactions (PCR). Journal of Bioscience and Bioengineering. 2003;96:317-23.
4. Tanaka K, Okamoto A. Degradation of DNA by bisulfite treatment. Bioorganic & medicinal chemistry letters. 2007;17:1912-5.
5. Oxford Nanopore Technologies ON. EPI2ME GIAB Benchmark Workflow (v2025.01). 2025.
6. Oxford Nanopore Technologies ON. EPI2ME wf-human-variation Worktlow Documentation. n.d.
7. Chiu R, et al. (BC Cancer Genome Sciences Centre). Short Tandem Repeat Genotyping using Long Reads.
8. Pacific Biosciences P. Revio HiFi Public Dataset (HG002, 2022 Q4 Release), 2022.