A streamlined and modular workflow solution enhances coverage uniformity and achieves high library complexity from genomic DNA/total RNA using both short and long read sequencers

Bahram Arezi[1], Adam Janssen[1], Nikki Liu[2], Sangwon Lee[2], Sergey Shiryaev[1], Keith Chen[1], Bilan Hsue[1], Brandyn Clark[1], Jeff Fox[1], Nedda Saremi[2], Inlora Jingga[2], Kyeong-soo Jeong[2], Khine Win[2], Gilbert Amparo[2], Neelima Mehendale[2], Jayati Ghosh[2], Manuel Gomez[2], Doug Roberts[2], Cindy Hon[2] , Karen Chapman[1]

[1]Agilent Technologies, La Jolla, CA [2]Agilent Technologies, Santa Clara, CA

AACR 2025

POSTER # 23
ABSTRACT# 4086

PR#7001-4055

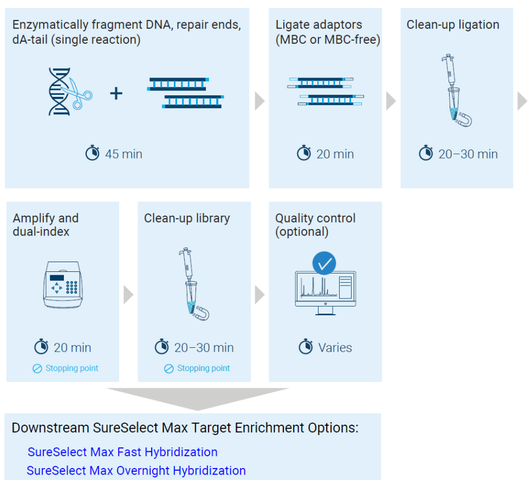
Agilent
Trusted Answers

Introduction

The application of NGS-based assays can be limited by the availability and quality of genomic DNA from certain sample sources. Practical factors such as turnaround time, workflow simplicity, target enrichment capability, and overall performance are crucial in determining whether NGS is the preferred method for germline or somatic mutation detection. In addition, due to certain short read sequencer limitations (e.g., mapping rate, repetitive regions, etc.), enabling long read capture and sequencing allows for better detection of structural variations, complex rearrangements, and variants in highly repetitive or polymorphic regions. To address these challenges, we have developed a flexible and automation-compatible workflow solution that supports both WGS (PCR-free and PCR-based, depending on the application) and target enrichment sequencing. Our modular workflow accommodates a wide range of DNA and RNA inputs (10–500 ng), standard and in-line molecular-barcoded adaptors, enzymatic and mechanical shearing methods, fast or overnight probe hybridization, and up to 384 unique dual sample indices to eliminate index hopping. Our enzymatic library preparation method combines fragmentation, end repair, and dA-tailing in a single step, allowing the use of samples with low concentrations. We demonstrate high base coverage uniformity and minimal bias across all GC content (Fold 80 of 1.12-1.13) using three bacteria of varying GC content (30%-67%) and human gDNA. Finally, our novel fast hybridization buffer formulation provides maximum complexity while maintaining high uniformity and base coverage, specifically addressing the issue of complexity loss when multiple libraries are pooled during bait hybridization. As we demonstrate, with small changes in the protocol, efficient long read capture (3-5kb) and sequencing can be successfully accomplished.

Experimental

A



Experimental

B

Figure 1. SureSelect Max workflow

(A) SSEL Max One-step Enzymatic Library prep workflow with inputs from 10-200ng using gDNA. To allow maximum flexibility, the kit has been designed in modular format to process mechanically-sheared DNA, total RNA, and mRNA. (B) SSEL Max Fast Hyb module (new formulation that maintains higher complexity) is compatible with individual or pre-cap pooling up to 8 libraries for large baits (e.g., exome V8) or 16 libraries for smaller baits. The workflow is optimized to allow skipping of pre-cap QC, as well as extending the 1-day workflow to 2 days, if desired. WGS protocols are also available.

Results and Discussion

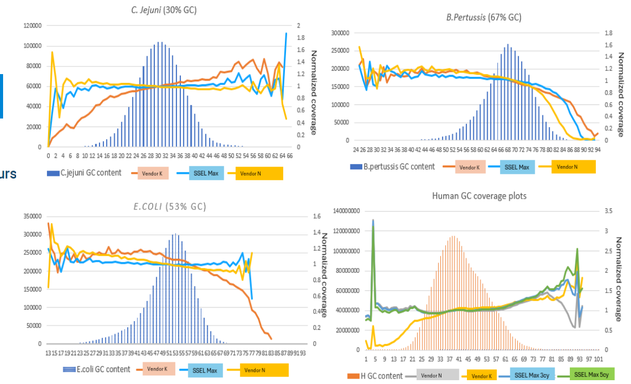
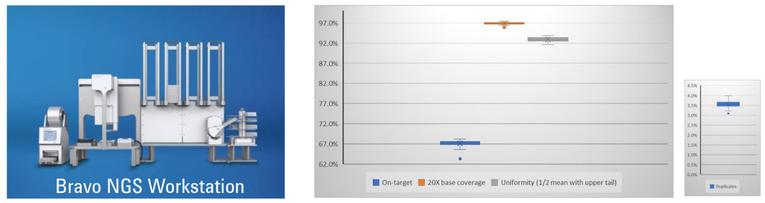
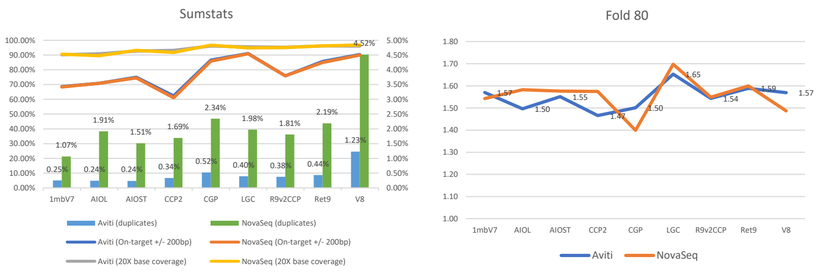


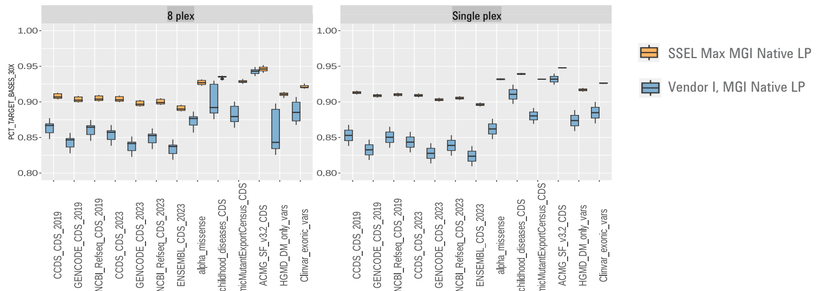
Figure 2. Uniform GC Coverage using SSEL Max WGS with minimal PCR cycles on 3 different microbial as well as human gDNA with varying GC content.

Agilent's SSEL Max WGS libraries were prepared in conjunction with library preps using vendors K and N. SSEL Max shows higher coverage in high AT regions with *C. jejuni (30%GC)*/human genome and high GC regions in *E.coli (53%GC)* in comparison to vendor K. SSEL Max outperformed vendor N at high GC regions for *B pertussis (67%GC)*. Libraries were sequenced on Novaseq 6000 and reads were normalized to 310 million (2X150bp).

Results and Discussion







A

Analysis tool	Metric	AmpPure XP bead size-selection	SureSelect Max bead size-selection
	Bamcounts	1885881	2662125
	Normalization	1800000	1800000
NanoPlot (trimmed)	median_read_length	4506	3965
	mean_read_length	4367.8	3981.9
	n50	5291	4580
	mean_qual	16.4	16.6
	Reads >Q10	1643815 (96.0%)	2337363 (96.4%)
Picard tools	TOTAL_READS	1633886	1640104
	PF_READS_ALIGNED	1620968	1628146
	PCT_PF_READS_ALIGNED	99.21%	99.27%
	MEAN_READ_LENGTH	4367.5	3981.6
	PCT_CHIMERAS	2.47%	2.13%
	PCT_ADAPTER	0.03%	0.03%
Picard hsmetrics	PCT_SELECTED_BASES	86.37%	86.79%
	MEAN_BAIT_COVERAGE	303.1	296.6
	FOLD_80_BASE_PENALTY	2.39	2.44
	PCT_TARGET_BASES_1X	99.15%	99.30%
	PCT_TARGET_BASES_20X	96.14%	96.53%
	PCT_TARGET_BASES_50X	90.85%	91.18%
	AT_DROPOUT	0.29	0.35
	GC_DROPOUT	9.35	8.88

B

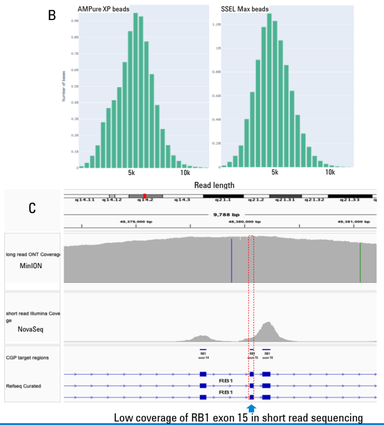


Figure 3. SSEL Max LP automated on Bravo NGS workstation Generates Highly Reproducible Sequencing Metrics Across 96 Wells

SSEL Max Enzymatic LP and Fast Hyb buffer were used to enrich with 250 Kb custom cancer panel. Data is normalized to 120X raw seq depth. Average of on-target, base coverage, uniformity, duplicates, and SD are plotted.

Figure 4. SSEL Max Sequencing Metrics on Novaseq 6000 and Element Aviti Cloudbreak Across Multiple Panels

SSEL Max Enzymatic LP and Fast hyb buffer were used to enrich using 9 panels of varying GC content and sizes. Libraries were sequenced on both Illumina NovaSeq 6000 and Element Aviti Cloudbreak (no conversion). Reads were normalized to 40 million for V8 and 100X raw depth for all other panels.

Figure 5. SSEL Max shows highest 30X Base Coverage (HS Metrics) on MGISEQ 2000 (aka, DNBSEQ-G400) Against Vendor I Mapping to Various Reference Databases

SSEL Max Library Prep and Fast Hyb buffer with exome V8 enrichment shows the highest percentage of targeted bases at 30X in 13 commonly used databases with the lowest variability across replicates when compared to vendor I Fast hyb and enrichment using its relevant exome.

Figure 6. Compatibility of SSEL Max library prep and capture with Long Read sequencing.

SSEL Max enzymatic LP and Fast hyb enrichment were used with CGP panel. Covaris G-tubes were used for fragmentation followed by size selection using AMPure or SSEL Max beads. Final libraries were converted using ONT native barcoding kit and run on MinION sequencer.

(A) Sequencing stats (B) Weighted histogram of read length distribution (C) Examples of sequencing coverage for challenging regions using short read vs. long read sequencing (e.g., RB1 exon 15 coverage)

Conclusions

SSEL Max generates higher library complexity, deeper and more uniform coverage in both pre-cap and post cap pooling. Libraries are compatible with sequencing on Illumina, Element, and MGI sequencers.

SSEL Max provides a more streamlined one-step enzymatic LP in higher fragmentation volume (50ul), optional pre-cap QC step, and fast or overnight hyb options.

SSEL Max is fully modularized and automation compatible, supporting both DNA and RNA, as well as low cycle (for FFPE and ct DNA) and PCR-free WGS.

With small changes in the protocol, efficient long read library prep and capture, suitable for sequencing on Oxford Nanopore and PacBio sequencers can be accomplished.