HIGH QUALITY AND HIGH YIELD OF ENRICHED DNA-SEQ LIBRARIES POSSIBLE USING AUTOMATED LIBRARY PREPARATION AND CAPTURE METHODS

Emma White[1], *Joshua Tjokrosurjo[1], Marsha McMakin[1], Rachel Kasinskas[1], Nicole Madamba[2], and Dawn Obermoeller[2]
[1]Roche Diagnostics, [2]Perkin Elmer, Inc.
*Author is no longer an employee of Roche

OVERVIEW

- Target Enrichment of prepared NGS libraries maximizes the value of sequencing and computational resources by targeting and capturing regions of interest prior to sequencing.
- Automation of target enrichment workflows increases throughput, reduces handling errors, and increases overall workflow efficiency.
- The KAPA HyperCap Workflow v3 combines KAPA HyperPrep or KAPA HyperPlus library preparation kits with the KAPA Target Enrichment Kit, and is compatible with pre-designed or custom probe panels.
- This study demonstrates the successful automation of the KAPA HyperCap Workflow v3 on the Sciclone G3 NGSx and Sciclone NGSx iQ platforms from Perkin Elmer, Inc.
- In this study, the liquid hander platforms processed low- and high-throughput sample preparations using replicates of intact human gDNA and KAPA HyperExome, a human exome probe panel.
- Pre- and post-sequencing metrics demonstrate that both the Sciclone G3 NGSx and NGSx iQ platforms generate consistent, high-quality enriched libraries in both low-throughput and high-throughput conditions using the HyperCap Workflow v3 and KAPA HyperExome.

WORKFLOW

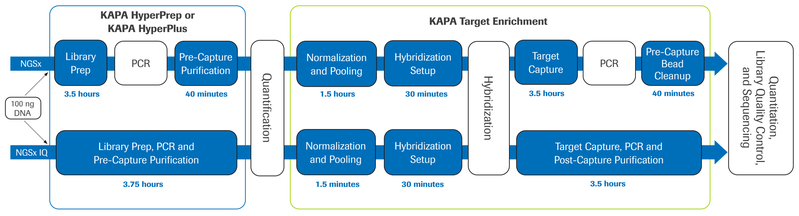

Figure 1: KAPA HyperCap Workflow v3 on Sciclone G3 NGSx and NGSx iQ Workstations. This diagram shows the KAPA HyperPrep/HyperPlus Kits (encircled in blue) and KAPA Target Enrichment Kit (encircled in green) implemented into sample prep workflows on the Sciclone G3 NGSx (top) and NGSx iQ (bottom) workstations following the KAPA HyperCap Workflow v3 for preparing enriched sequencing libraries. Blue blocks indicate on-deck processes, and the white blocks identify actions occurring off-deck. The provided run times have been estimated for processing 96 samples on each platform.

INTRODUCTION

Next-Generation Sequencing can be optimized by narrowing the breadth of a sample through selective capture of targeted sequences. This approach exclusively processes specific regions of interest during sample preparations, thus reducing demands for sequencing and computational capacity per sample.

The **KAPA HyperCap Workflow v3** combines two validated KAPA Kits into one standardized target enrichment workflow:

1. Either the **KAPA HyperPrep Kit** (for mechanically fragmented input DNA) or the **KAPA HyperPlus Kit** (with up-front enzymatic DNA fragmentation) for ligation-based library preparation
2. **KAPA Target Enrichment Kit**: an in-solution probe hybridization kit compatible with the KAPA HyperExome human exome probe panel and other KAPA Target Enrichment standard or custom probe panels

In this study, the KAPA HyperCap Workflow v3 was implemented on two Perkin Elmer, Inc. automated liquid handler platforms: the Sciclone G3 NGSx and NGSx iQ workstation. Workflows on both instruments yielded high-quality sequencing libraries and exemplary target capture specificity and uniformity. In addition, automation of the KAPA HyperCap Workflow under both low- and high-throughput conditions yielded high-quality libraries and sequencing results.

METHODS

LIBRARY PREPARATION AND TARGET ENRICHMENT:

- Enriched libraries were prepared following the standard KAPA HyperCap Workflow v3.
- Libraries were prepared from 100 ng of intact NA12878 human gDNA using the KAPA HyperPlus Kit, including the enzymatic fragmentation step.
- Libraries were normalized and combined into pools of 8 libraries for 8-plex target capture.
- Pre-capture libraries and post-capture pools were quantified using the QubitTM 2, and fragment distribution was analyzed using the LabChip® GXII Touch™ HT. A subset of post-capture pools were further quantified with the KAPA Library Quantification Kit prior to sequencing.

SEQUENCING AND ANALYSIS:

- Selected pools were sequenced on an Illumina® NextSeq® 500 with high-output flowcells, generating 75 bp pair-end reads.
- Sequencing data for each sample was down-sampled to 30 M paired-end reads for target enrichment analysis.

Table 1: Parameters for low-throughput and high-throughput experiments performed on the Sciclone G3 NGSx and NGSx iQ platforms.

Parameters	SciClone G3 NGSx (low throughput)	SciClone G3 NGSx (high throughput)	SciClone G3 NGSx iQ (low throughput)
NA12878 human gDNA	16 Replicate Libraries	96 Replicate Libraries	16 Replicate Libraries
8-plex Target Capture	2 Pools	12 Pools	2 Pools
Target Enrichment Probe Panel	KAPA HyperExome (human exome)		
Libraries Sequenced	8	24	8

RESULTS

PRE-SEQUENCING QUALITY METRICS ARE COMPARABLE ACROSS THROUGHPUT LEVELS AND PLATFORMS

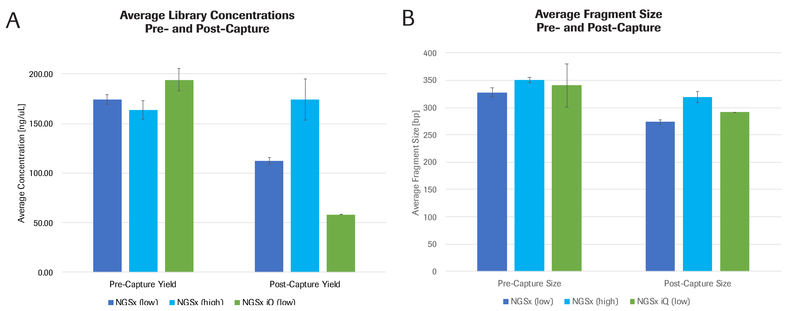

Figure 2: Comparison of average library yields and average fragment size between low-throughput and high-throughput conditions and between the SciClone G3 NGSx and SciClone NGSx iQ platforms. The average concentration of libraries **(A)** are within acceptable ranges and demonstrate minimal variability across all experimental conditions before and after target capture. Post-capture yields vary between conditions, but remain within expected ranges while also maintaining low variability within each condition. The average fragment size of pre- and post-capture libraries **(B)** are consistent and uniform across experimental conditions. Error bars represent standard deviation across averaged data.

HIGH-THROUGHPUT PROCESSING DOES NOT GENERATE PLATE EFFECTS

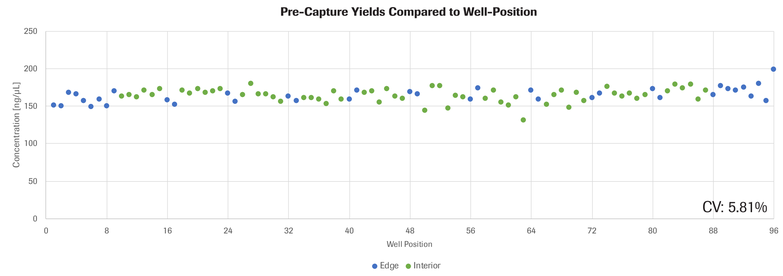

Figure 3: Well position and concentration [ng/µL] of 96 pre-capture libraries prepared on the Sciclone NGSx workstation. In a 96-well plate, well A01 is position 1 and well H12 is position 96 and all other positions are likewise identified in a sequential, columnar manner. The concentration [ng/µL] of each of the 96 replicate libraries in their corresponding well positions demonstrates even performance of library preparation regardless of positioning. Wells located on the edge of the plate (blue) are comparable to wells located in more interior positions (green). The concentrations are even across the plate, and variation is minimal, with a coefficient of variation (CV) of 5.81%.

POST-SEQUENCING QUALITY METRICS AND TARGET ENRICHMENT ANALYSIS REVEAL HIGH SPECIFICITY AND UNIFORMITY

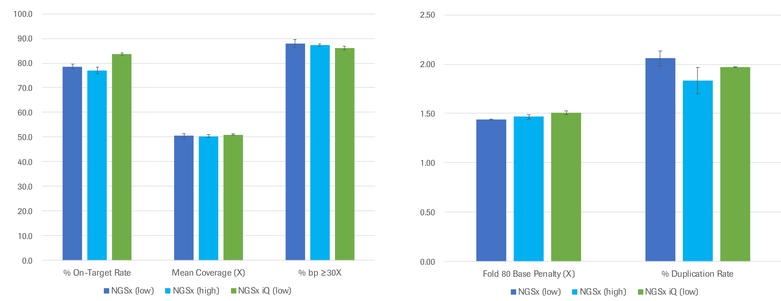

Figure 4: Accuracy, unifromity and depth of coverage for targeted regions of interest in sequencing data. Analysis of sequencing data shows high on-target rates with deep coverage, most of which is > 30X coverage **(A)**. This demonstrates that the majority of each library consisted of targeted regions of interest with significant coverage. The regions of interest were also evenly represented across the data, as shown by the Fold 80 Base Penalty **(B)**. Finally, the libraries contained a minimal percentage of duplicate reads **(B)**, indicating rich and diverse libraries unencumbered by PCR duplicates. Error bars represent standard deviation across averaged data.

CONCLUSION

Automation of the KAPA HyperCap Workflow v3 on the Sciclone G3 NGSx and Sciclone NGSx iQ workstation yielded high-quality, target-enriched sequencing libraries with the KAPA HyperExome probe panel.

Quality metrics and sequencing data demonstrate:

- The creation of sequencing libraries with acceptable yields and expected fragment distributions both pre- and post-capture
- Low duplication rates, high on-target rates, and sufficient sequencing depth with uniform coverage across the targeted regions

The automated solution of KAPA HyperCap Workflow v3 on the Sciclone G3 NGSx and Sciclone NGSx iQ workstations offers efficient preparation of consistently high-quality, target-enriched sequencing libraries in both low- and high-throughput formats.

For more information about Roche KAPA HyperCap Workflow v3, please visit: go.roche.com/HyperCap

Data on file.
For Research Use Only. Not for use in diagnostic procedures.