Overcoming sample quality and quantity limitations for NGS analysis of clinically relevant samples

Martin Ranik, Giulia Corbet, Philip Benson, Josh Haimes, Andrew Gaines, Thomas Harrison, Skyler Mishkin, Kailee Reed, Kristin Scott, Zane Jaafar, Kristina Giorda, Brian Kudlow
Watchmaker Genomics – Boulder, CO USA

Introduction

The use of NGS in clinical applications has skyrocketed in recent years, facilitating early cancer detection and monitoring of disease progression. Two common sample types in clinical applications, cell-free DNA (cfDNA) and formalin-fixed, paraffin-embedded (FFPE), present various challenges for use in NGS, including limited quantity and quality of nucleic acids. Challenging samples necessitate robust and high-fidelity library preparation methods to minimize artifacts and maximize sensitivity. Watchmaker Genomics pioneers innovative library preparation kits for these challenging samples, using in-house enzyme engineering and high-resolution NGS assays to enhance scalability and sensitivity, revolutionizing clinical genomics.

Table 1. Challenges in working with FFPE and cfDNA samples and their impacts

Challenges	Impact
Poor extraction yields	Limited mass available for Library Prep which results in reduced library complexity, coverage, and sensitivity
Chemical modifications and cross-links in FFPE limit nucleic acid available for library prep	Lower conversion efficiency resulting in lower complexity, coverage, and sensitivity
Increased sample failure rates	Inefficient and reduced economy
Sequencing artifacts	Increased false positives

Enhancing cfDNA Coverage & Sequencing Efficiency

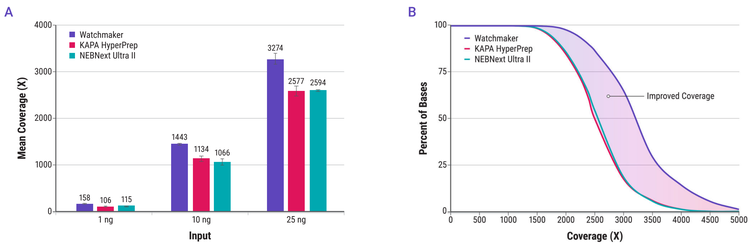

Figure 1. Watchmaker Kit produces higher coverage and efficient sequencing. (A) Mean coverage across a 37 kb panel was higher with the Watchmaker kit for all input masses tested. Read pair subsampling for each input was: 1 ng – 500,000; 10 ng – 2,500,000; 25 ng – 25,000,000. **(B)** Schematic demonstrating the percent of bases with the indicated coverage across library prep methods.

Table 2. The Watchmaker kit produces high-quality libraries for efficient sequencing and target enrichment

Input	Subsampled Read Pairs	Library Prep	Passing Filter Reads Aligned (%)	Improper Pairs (%)	Duplication Rate (%)	Bases On + Near Bait (%)
1 ng	500,000	Watchmaker	99.2%	1.7%	88.5%	70.1%
		KAPA	99.2%	2.1%	90.6%	62.3%
		NEB	99.2%	2.1%	91.3%	68.2%
10 ng	2,500,000	Watchmaker	99.5%	1.1%	81.9%	79.2%
		KAPA	99.3%	1.7%	85.2%	78.2%
		NEB	99.3%	1.5%	87.1%	80.0%
25 ng	25,000,000	Watchmaker	98.4%	4.2%	96.2%	80.7%
		KAPA	98.4%	5.3%	96.9%	79.5%
		NEB	98.2%	4.8%	96.9%	79.4%

Boosting Confidence in cfDNA Variant Calling

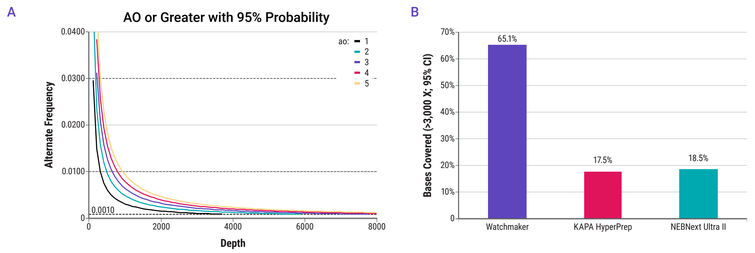

Figure 2. Higher coverage gives higher confidence for variant calls. (A) Theoretical model demonstrating the depth of coverage required for 95% confidence in detecting a given number of alternate observations (AOs) for given alternate frequencies. **(B)** Percentage of bases covered at 3000X or higher in libraries made from 25 ng cfDNA subsampled to 25M read-pairs.

Reducing Sequencing Artifacts in FFPE Samples

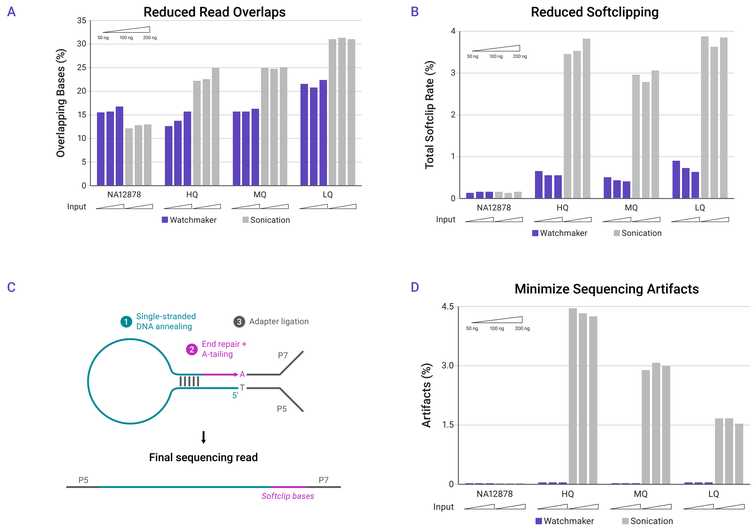

Figure 3. Watchmaker kit provides improved sequencing economy and minimal artifacts. (A) Longer insert sizes with the Watchmaker kit give fewer read overlaps, a measure of sequencing economy. **(B)** Softclipping is approximately 5-fold lower in Watchmaker FFPE libraries than sonication-based libraries. **(C)** Schematic showing a hypothesized mechanism of formation of hairpin artifacts, which hinder data interpretation (adapted from Gregory, et al.). **(D)** Watchmaker Fragmentation kit produces fewer hairpin artifacts than sonication-based library prep with 3 different FFPE samples. Highly elevated hairpin artifacts with sonication-based workflows have been observed in some, but not all, experiments and have been reported in the literature (Haile, et al.).

Maximizing FFPE Target Coverage for Sensitivity

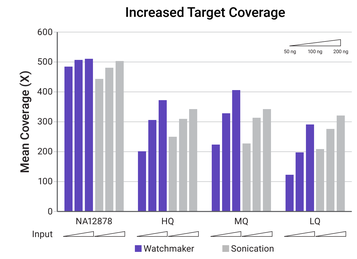

Figure 4. Fragmentation produces high-coverage libraries. Input mass for library prep was matched between workflows. Coverage with Watchmaker kit was comparable or higher than sonication at higher inputs. Sonication input mass was normalized post-shearing, which normally gives 20 – 40% sample loss, reducing coverage.

Methods

Mock cfDNA: Libraries were prepared with Horizon Discovery Multiplex I cfDNA Reference Standard Set (HD780) with 1 – 25 ng input mass. Libraries were prepared with the Watchmaker DNA Library Prep Kit, KAPA HyperPrep kit, or the NEBNext Ultra II kit according to the manufacturer's protocol with the exception of adapter concentrations which were matched across library prep kits. Libraries were target enriched with a 37 kb hybrid capture panel and sequenced on the Illumina platform with a 2 x 150 bp read length. Metrics shown are the average of 3 replicates for each input mass and kit.

FFPE: Libraries were prepared with genome in a bottle DNA (NA12878) and FFPE DNA samples using either a sonication-based library preparation method using the KAPA HyperPrep Kit or the Watchmaker DNA Library Prep Kit with Fragmentation, followed by target enrichment using a 37 kb hybrid capture panel. An input titration using 50, 100, and 200 ng was performed for each condition. Sonication library inputs were post-shearing. Libraries were sequenced on the Illumina platform with a 2 x 150 bp read length and subsampling was uniform across input masses and library preparation kits.

References

Haile S., Corbett, R.D., Bilobram, S., et al. Sources of erroneous sequences and artifact chimeric reads in next generation sequencing of genomic DNA from formalin-fixed paraffin-embedded samples. *Nucleic Acids Res.* 47(2): e12 (2018).

Gregory T., Ngankeu, A., Orwick, S., et al. Characterizationand mitigation of fragmentation enzyme-induced dualstranded artifacts. *NAR Genom. Bioinform.* 2(4): lqaa070(2020). doi: 10.1093/nargab/lqaa070.

Conclusions

Watchmaker DNA library prep kits enable clinical applications of NGS by:

- Delivering high target coverage, improved sequencing economy, and high sensitivity
- Virtually eliminating sequence artifacts associated with damaged FFPE samples.