Highly scalable and accurate library preparation for high sensitivity applications

Martin Ranik[1], Zane Jaafar[1], Josh Haimes[1], Thomas D. Harrison[1], Lindsay Peterkin[1], Kristina Giorda[1], Kristin Scott[2], Brian Kudlow[1]
1 Watchmaker Genomics – Boulder, CO USA | 2 Colorado State University, Department of Cell & Molecular Biology – Fort Collins, CO USA

Introduction

High-throughput sequencing has brought a paradigm shift in large-scale precision medicine and clinical research initiatives. To meet this demand, library preparation methods need to have a robust and automation-friendly workflow.

Table 1. Conventional trade-offs between Covaris sonication and enzymatic fragmentation library preparation methods

	Sonication	Enzymatic fragmentation
Pros	High sequence accuracy Protocol independent of input amount Performance independent of contaminants	Scalable Workflow simplicity Minimal template damage
Cons	Costly equipment investment Costly consumables Time consuming Difficult to scale Template damage	Introduces sequence artifacts May need optimization for input amount May be sensitive to contaminants

Simplified and Easily Automated Workflow

Figure 1. Simplified and automation friendly workflow. The Watchmaker DNA Library Prep Kit with Fragmentation streamlines library construction by using combined enzymatic steps for Frag/AT (fragmentation and A-tailing) and a Ligation Master Mix that were designed with automation in mind.

Dependable and Tunable Fragmentation

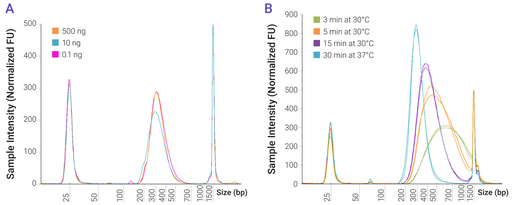

Figure 2. Consistent yet flexible fragmentation. (A) Libraries were constructed in duplicate using human gDNA fragmented for 20 minutes at 30°C, as indicated. **(B)** Libraries were constructed from 50 ng of human gDNA using a fragmentation reaction time course at 30°C or 37°C, as indicated. Final library distributions were assessed using a D1000 assay by TapeStation (Agilent).

Methods

PCR-free WGS. Sonication control libraries were constructed from 475 ng of Covaris sheared gDNA (NA12878, Coriell Institute) using the KAPA HyperPrep Kit. Enzymatic fragmentation libraries were constructed from 300 ng of high molecular weight gDNA (NA12878, Coriell Institute) using either the Watchmaker DNA Library Prep Kit with Fragmentation, KAPA HyperPlus, NEBNext Ultra II FS DNA Library Preparation Kit, QIAseq FX DNA Library Kit, xGen DNA Library Prep EZ Kit (formerly Swift FS Turbo DNA), per manufacturer's recommendations. Watchmaker DNA Library Prep Kits with Fragmentation utilized a 3 minute at 30°C fragmentation followed by a 0.5X post-ligation SPRI for longer fragments.Libraries for all conditions were prepared in duplicate.

Ultra-low input. NCI-60 cell lines were used for library construction using a titration from 100 ng (PCR-free), 1 ng, 0.1 ng, and 0.01 ng. Ultra-low input 0.01 ng samples utilized tandem post-PCR SPRI. All samples were deduplicated for downstream analysis. Read counts were binned into 15 kb windows and normalized by the total number reads in the window's respective sample.

PCR-Free Whole Genome Sequencing

Table 2. Mean and median library fragment lengths from sequencing

Method	Kit	Mean insert size (bp)	Median insert size (bp)
Sonication	KAPA	384	366
Enzymatic fragmentation	Watchmaker	368	325
	KAPA	315	264
	NEB	369	307
	QIAGEN	326	286
	IDT/Swift	332	293

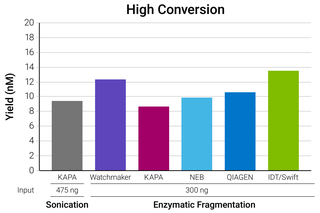

Figure 3. Efficient library construction. All kits yielded adequate library concentrations for sequencing whereas insert sizes varied significantly (Table 2).

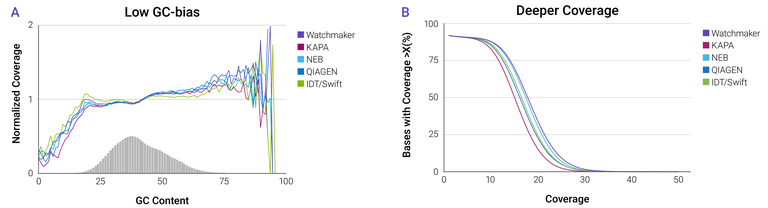

Figure 4. Unbiased GC-coverage and deeper per-base sequencing coverage improves sequencing economy. (A) GC-bias and **(B)** cumulative coverage plots for PCR-free libraries were prepared using the manufacturer's recommendations, as indicated. Watchmaker DNA Library Prep Kits with Fragmentation produced similar GC uniformity to other enzymatic methods, but higher coverage using the same input mass and number of starting reads. Higher coverage using the same number of sequencing reads should enhance variant calling sensitivity and reduce sequencing costs.

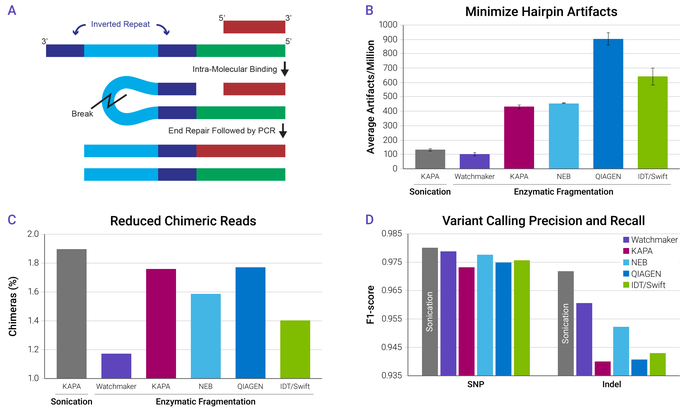

Figure 5. Increased variant calling accuracy. (A) Hairpin artifacts which may cause false positive variant calls.[1,2] Watchmaker's kit produced the lowest rate of hairpin artifacts **(B)** and **(C)** chimeric reads. **(D)** Minimal artifacts combined with deeper coverage produced higher SNP and Indel F1-scores than other enzymatic fragmentation methods, close to sonication controls which used almost 60% more input DNA.

Optimized for High Fidelity

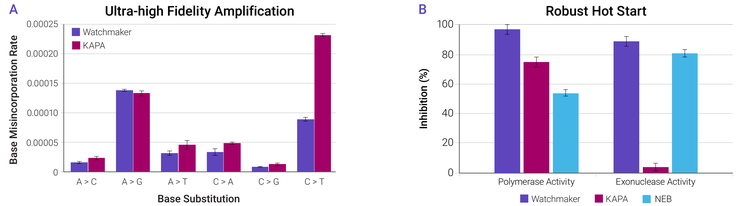

Figure 6. Excellent fidelity and optimal hot start formulation. (A) Polymerase error rates were measured using a proprietary NGS-based assay. Data not shown for NEB Q5 polymerase as it did not produce sufficient amplification for the base incorporation assay. **(B)** Polymerase and exonuclease hot start functionality was assessed for each vendor by the detection of dNTP incorporation or dNMP release, respectively, after incubation at 25°C.

Ultra-Low Input Performance

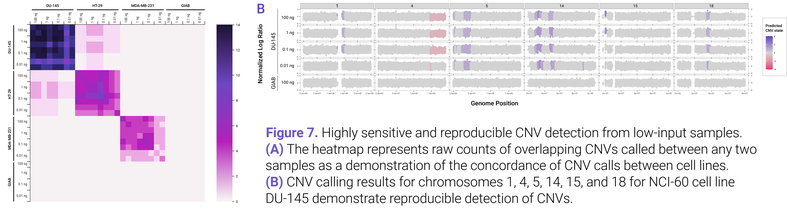

Figure 7. Highly sensitive and reproducible CNV detection from low-input samples. **(A)** The heatmap represents raw counts of overlapping CNVs called between any two samples as a demonstration of the concordance of CNV calls between cell lines. **(B)** CNV calling results for chromosomes 1, 4, 5, 14, 15, and 18 for NCI-60 cell line DU-145 demonstrate reproducible detection of CNVs.

References

1. Woyach, John C Byrd, James S Blachly, Characterization and mitigation of fragmentation enzyme-induced dual stranded artifacts, *NAR Genomics and Bioinformatics*, Volume 2, Issue 4, December 2020, lqaa070, https://doi.org/10.1093/nargab/lqaa070

2.Norio Tanaka, Akihisa Takahara, Taichi Hagio, Rika Nishiko, Junko Kanayama, Osamu Gotoh, Seiichi Mori, Sequencing artifacts derived from a library preparation method using enzymatic fragmentation, *PLoS ONE* 15(1): e0227427. https://doi.org/10.1371/journal.pone.0227427

Conclusions

Watchmaker DNA Library Prep Kit with Fragmentation offers an unparalleled combination of scalability and accuracy with:

- Robust library construction workflows to support the use of clinically relevant sample types and ultra low inputs
- High conversion, uniformity, and accuracy overcome limitations of conventional enzymatic fragmentation kits