A highly scalable enzymatic fragmentation library preparation chemistry that greatly reduces sequencing artifacts to enable high sensitivity applications

Genomics

Zane Jaafar (zane.jaafar@watchmakergenomics.com), Watchmaker Genomics, **Josh Haimes**, Watchmaker Genomics, **Thomas Harrison**, Watchmaker Genomics, **Lindsay Peterkin**, Watchmaker Genomics, **Martin Ranik**, Watchmaker Genomics, **Kristin Scott**, Colorado State University, **Kristina Giorda**, Watchmaker Genomics, **Brian Kudlow**, Watchmaker Genomics

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. Sonication methods have been the gold standard for consistent fragmentation and uniform GC coverage, but are associated with a high upfront investment, expensive consumables, and are prone to oxidative DNA damage. Enzymatic fragmentation methods hold the potential to address this need, yet they typically suffer from reduced uniformity, are sensitive to DNA input quantity and quality variation, and introduce sequencing artifacts that lead to false variant calls which impact clinical utility.

Here, we utilized sophisticated enzyme engineering and a multidimensional Design of Experiment approach to develop a novel enzymatic library preparation method and ultra-high fidelity amplification module. The goals were to mitigate sequencing artifacts, generate a wide range of insert sizes by varying temperature or incubation time, and accommodate variable input masses. Libraries generated using our optimized workflow reduced chimeric reads and terminal hairpin artifacts 10-fold compared to other enzymatic methods, and reached comparable levels of mechanically sheared DNA controls. Fragmentation was tested across a broad sample input range from 100 pg to 500 ng. Library insert sizes were highly tunable from 150 bp to 550 bp and were consistent across the input titration. To assess the utility of ultra-low input samples, libraries were prepared using a titration from 100 ng to sub-nanogram inputs. Copy number variations (CNVs) were detected across the titration with high sensitivity and specificity using Hidden Markov Model analysis method.

Taken together, this enzymatic fragmentation and library preparation workflow avoids library preparation artifacts that convolute variant calling, is highly scalable, and suitable for ultra-low input samples.