Single-Step Library Preparation Workflow Enables High-Throughput Low-Pass Whole-Genome Sequencing Without Compromising Accuracy

Sabina Gude[1], Samuel Widmayer[2], Mary Barter[2], Yanyan Liu[1], Michelle Rahardja[1], Dan Gatti[2], Qingchang Meng[2], Rebecca Feeley[1], Jenna Couture[1], Stella Huang[1], Maura Costello[1]
[1]seqWell, Inc. Beverly, MA USA *[2]The Jackson Laboratory, Bar Harbor, ME*

Introduction

Accurate and scalable genotyping is foundational to population genetics and large-cohort genomic studies. While microarray-based platforms have historically been used for genotyping, genotyping-by-sequencing (GBS) approaches, particularly low-pass whole-genome sequencing (lpWGS) combined with statistical imputation, have emerged as cost-effective alternatives that capture genome-wide variation with high accuracy. By sequencing samples at low coverage and leveraging dense reference panels, lpWGS enables robust genotype inference across millions of variants. Despite its advantages, the widespread adoption of lpWGS has been limited by the complexity and time demands of traditional library preparation workflows. Multi-step protocols are labor-intensive, prone to variability, and challenging to scale to hundreds or thousands of samples, creating bottlenecks in high-throughput genotyping efforts. AgriPrep™ Library Prep Kit was developed to address these challenges through a streamlined, single-step library preparation workflow that integrates tagmentation and amplification into a single thermal cycling reaction. This approach minimizes hands-on time while automatically normalizing libraries across a broad range of DNA input concentrations. Here, we evaluate the performance of AgriPrep Library Prep for low-pass whole-genome sequencing by benchmarking sequencing metrics and imputation-based genotyping accuracy against a multi-step tagmentation library preparation workflow. Our results demonstrate that AgriPrep enables rapid, high-throughput lpWGS without compromising genotype quality.

AgriPrep Library Preparation Kit Workflow

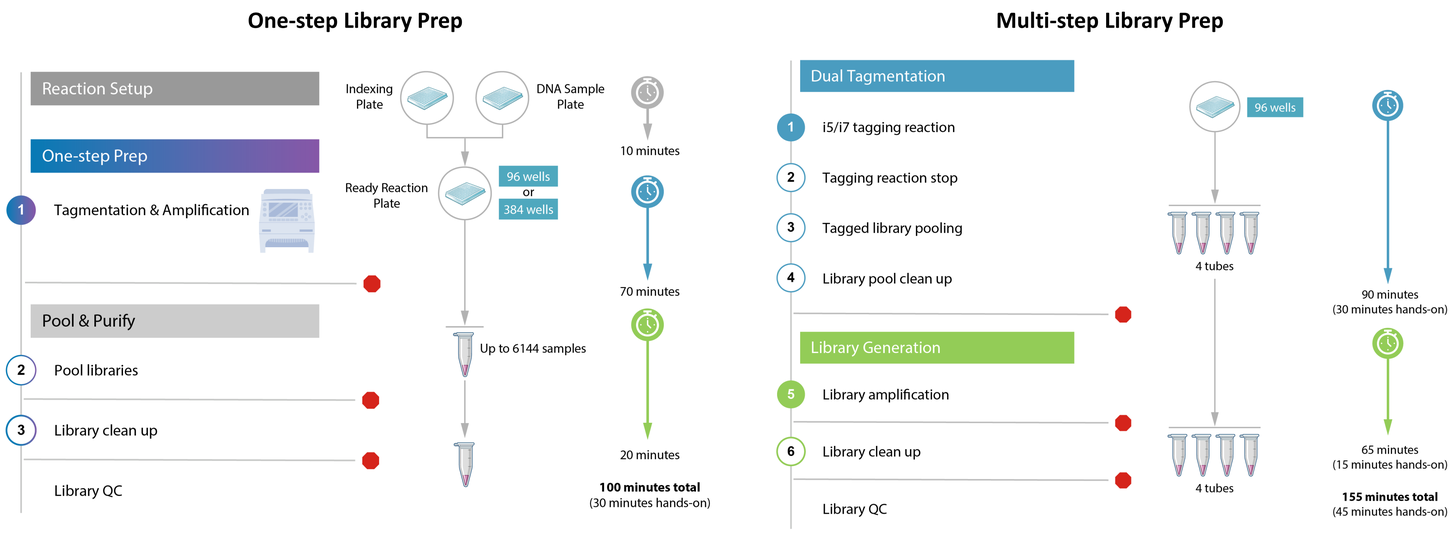

Figure 1. Streamlined AgriPrep workflow for scalable low-pass WGS library preparation. Traditional multi-step lpWGS library prep workflows *(right)* are labor-intensive, time-consuming, and prone to contamination and batch effects, limiting scalability and reproducibility in large-cohort studies. The AgriPrep workflow *(left)* simplifies library preparation by integrating tagmentation and amplification into a single thermal cycling reaction, reducing hands-on time while enabling high-throughput, consistent processing of hundreds to thousands of samples. The AgriPrep library prep kits utilize a proprietary mixture of enzymes to tag input DNA with full length indexed adapters and amplify libraries all in a single reaction. Different full-length i7 indexed adapters tag the 96 DNA samples and barcoded libraries are amplified in separate wells, making for a highly efficient, one-step multiplexed library prep workflow (right). Additional index sets achieve multiplex levels >96. Samples are pooled volumetrically, purified, and converted into libraries to complete the 100-minute workflow, which includes 30 minutes of hands-on time. Built-in auto-normalization obviates the need to normalize sample input.

High SNP Call Concordance and Genotyping Accuracy

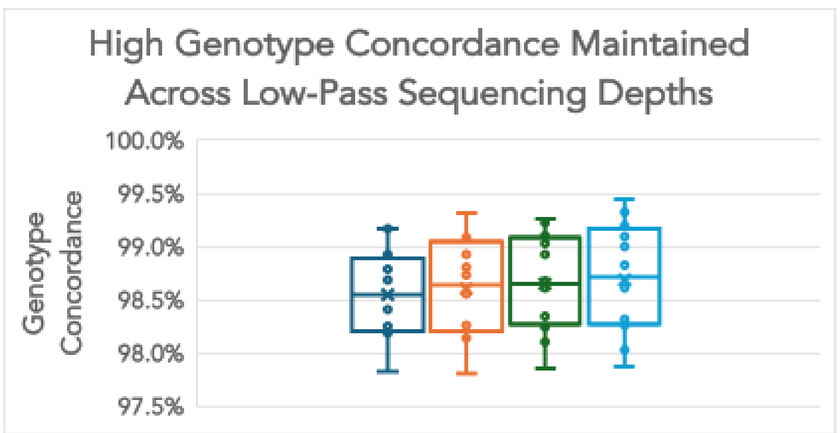

Figure 2. Genotype concordance across sequencing depths. Genotype imputation of lpWGS data generated using a conventional multi-step Tn5 tagmentation library preparation workflow produced high-density variant calls with strong concordance to Giga Mouse Universal Genotyping Array (GigaMUGA) genotypes. Imputed genotypes showed strong concordance with GigaMUGA results across sequencing depths, with approximately >98% concordance maintained at 0.1× coverage. These results establish lpWGS as a reliable alternative to microarray-based genotyping.

Figure 3. Total pass-filter read counts for 24 low-pass whole-genome sequencing libraries prepared using the AgriPrep workflow. Read yield is consistent (CV of 8.4%) across 24 libraries, demonstrating reproducible library preparation without manual normalization.

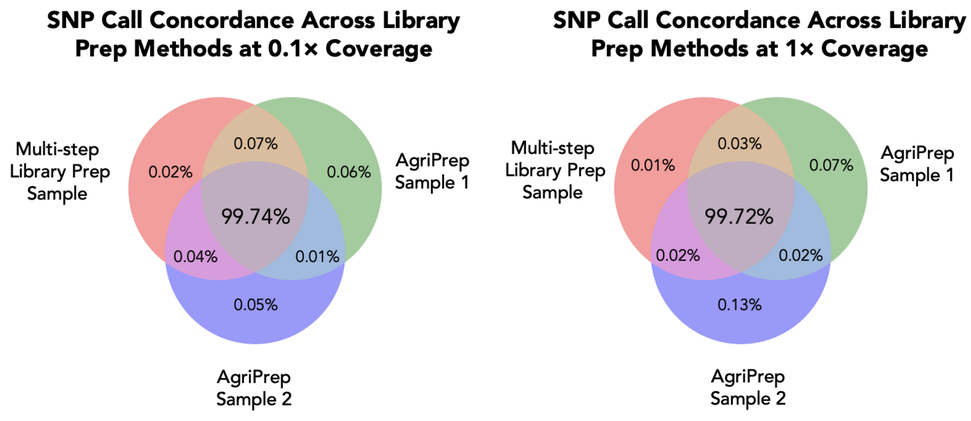

Figure 4. High-confidence SNP overlap across library preparation methods. Unweighted venn diagram show concordance of high-confidence SNP positions (INFO > 0.95) across library preparation methods at 0.1× (left) and 1× (right) coverage. SNPs were identified from merged chromosomes down-sampled to the indicated depth and compared between a multi-step library prep sample and two AgriPrep replicates. >99.7% of SNPs are shared across all three datasets at both coverage levels, indicating strong concordance and reproducibility across library prep methods, even at ultra-low coverage.

Methods

- Previously, genomic DNA from 12 Diversity Outbred (DO) mice was processed using multi-step Tn5 tagmentation library prep and sequenced on an Illumina platform to generate low-coverage whole-genome short-read data. Genotype imputation was performed, generating high-density variant calls from low-coverage sequencing data. Haplotype reconstructions derived from lpWGS showed strong agreement with Giga Mouse Universal Genotyping Array (GigaMUGA) results (Figure 2).
- Here, the AgriPrep library prep was used to processed genomic DNA from 24 DO mice samples using 20 ng input (16 individual samples and 8 technical replicates of the 16). For comparison, multi-step Tn5 tagmentation library prep was also used to processed the same 16 of samples using 45 ng input.
- The libraries prepared using the two different library prep methods were sequenced on an Illumina NovaSeq X Plus (300 cycles) using 10B reagent kit.
- Raw sequencing reads were quality filtered using FASTP and aligned to the mouse reference genome (GRCm39) using BWA-MEM. Duplicate reads were marked with Picard tools, and BAM files were down-sampled to assess performance across varying coverage depths (0.1x, 0.25x, 0.5x, and 1x). Genotype imputation was performed using QUILT, generating variant calls files (VCFs) from various coverage depths.

Table 1. Summary of sample quality assessed in the study.

Samples	DIN	Samples	DIN
Sample 1	6.3	Sample 9	7.0
Sample 2	6.8	Sample 10	7.1
Sample 3	6.7	Sample 11	6.9
Sample 4	7.4	Sample 12	7.3
Sample 5	6.9	Sample 13	6.3
Sample 6	8.6	Sample 14	6.8
Sample 7	6.9	Sample 15	7.8
Sample 8	7.9	Sample 16	7.2

Acknowledgements

We gratefully acknowledge the contribution of Samuel Widmayer, Dan Gatti, Qingchang Meng, Mary Barter, and the JAX Genome Technologies Laboratory at The Jackson Laboratory for expert assistance with the work described in this scientific poster.

Summary and Conclusions

- AgriPrep Library Prep Kit enables rapid, single-step library preparation for low-pass whole-genome sequencing with workflow time is reduced to under 100 minutes compared to multi-hour conventional protocols (Figure 1).
- Low-pass sequencing data generated using a conventional multi-step Tn5 workflow demonstrated strong concordance with microarray genotypes across sequencing depths, supporting lpWGS as a reliable alternative to array-based genotyping (Figure 2).
- AgriPrep libraries produced uniform sequencing yield across samples without manual normalization, enabling reproducible performance in high-throughput settings (Figure 3).
- Genotype concordance and high-confidence SNP recovery were equivalent between AgriPrep and conventional multi-step workflows, even at ultra-low coverage (Figure 4). these results demonstrate that AgriPrep delivers comparable genotyping accuracy and reproducibility while substantially reducing workflow complexity, supporting its use for large-scale lpWGS studies.