Genotyping of Bovine Tissue Samples Using High-Throughput Multiplexed NGS Library Preparation Workflow With Self-Normalizing Adapters

Owen Smith, Kristin Butcher, Ryan O'Donnell, Cibelle Nassif, Mathieu Chaluleau, Tiffany Truong, Sean Tighe, Tong Liu, Danny Antaki, Michael Bocek, Derek Murphy, Elian Lee, Ramsey Zeitoun, Esteban Toro, Siyuan Chen
Twist Bioscience, Research and Development, South San Francisco, California, USA

Abstract

High-throughput next-generation sequencing (NGS) workflows are required for population-level genomic research that aid SNP detection and discovery for genetic verification, breeding, and microbiome analysis. The cost and effort required in processing individual samples remain a barrier to efficient NGS library preparation. We present innovative and automation-friendly technologies that reduce cost and time: Normalization by Ligation™ (NBL) and library preparation multiplexing with inline barcodes. We perform experiments using our high-throughput workflow on real-life bovine samples and demonstrate concordance of SNP detection with ground truth datasets.

Enzymatic fragmentation (EF) of samples is the foundation of our ultra-high-throughput method. The EF module allows researchers to adjust the insert size based on sequencing format and to use samples of varying quality and GC content without significant base composition bias. Normalization using NBL produces uniform coverage for samples of 30 ng to 300 ng DNA input, eliminating the need for initial input quantification. NBL integrates inline barcodes to permit multiplexing post-ligation. As a result, the reaction footprint and reagent usage of downstream steps decrease 12-fold, increasing throughput and generating libraries in less time when compared with other standard workflows. By controlling conversion from each sample, NBL enables 96-plex in the target enrichment step versus the 8-plex commonly used in other protocols. When paired with automated methods, sample capacity is greatly increased with minimal hands-on time needed.

To demonstrate the utility of the workflow, we extracted genomic DNA (gDNA) from bovine ear punches using the BioEcho ECHOlution Tissue DNA Micro kit for high-throughput NGS library preparation and 96-plex target enrichment to detect variant alleles at ~74k SNP sites without upfront normalization. We observed great performance in coverage uniformity across samples with varying input masses and high SNP concordance when compared with a ground truth reference genotype built from WGS data.These results show that NBL and inline barcodes allow for a high-throughput streamlined workflow that saves on time and resources.

High-Throughput Workflow

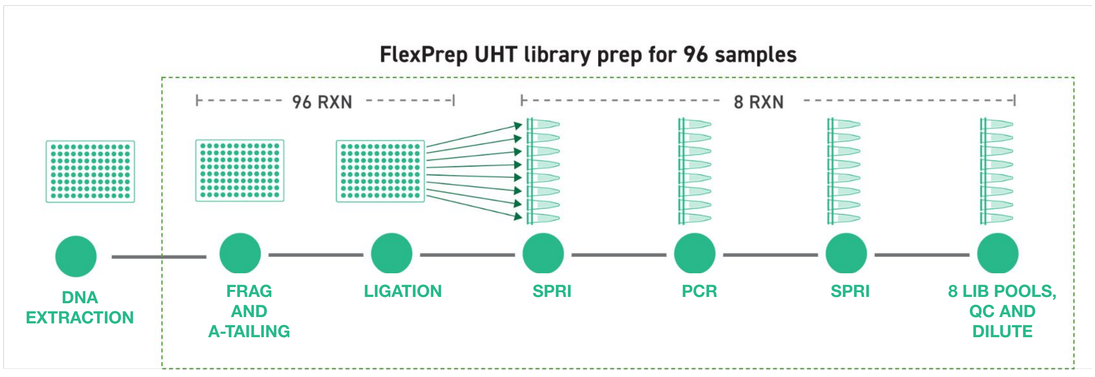

Figure 1. FlexPrep Ultra High-Throughput (UHT) Library Preparation Workflow for 96 Samples. Workflow begins with a fragmentation and A-tailing module in 96-sample format. Next, ligation proceeds for each reaction individually using self-normalizing and inline barcoded adaptors. Inline barcodes allow 96 samples to be pooled into eight library pools.

Twist FlexPrep UHT Library Preparation Kit provides streamlined workflow that simplifies library preparation and reduces reagent usage.

Genomic DNA Extraction

- Collect tissue punches from bovine ears using Allflex TSU.
- Extract gDNA with BioEcho ECHOlution Tissue DNA 96 kit.

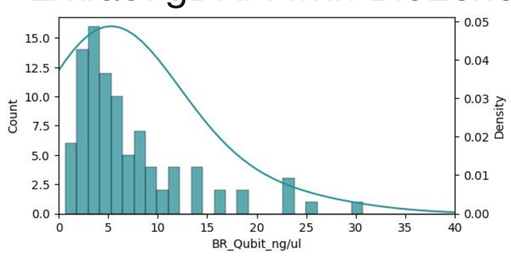

Figure 2. DNA Concentration After BioEcho Extraction. Overlay of histogram and density plots of extracted gDNA concentrations from bovine tissue samples.

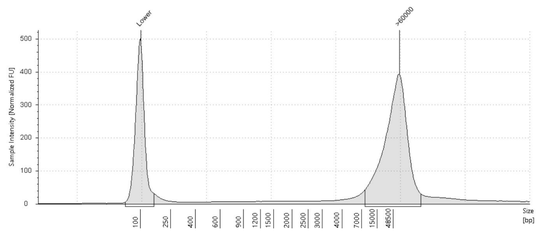

Figure 3. Genomic DNA Tapestation of One Sample From BioEcho Extraction. Representative trace of gDNA size distribution after BioEcho extraction. Clear high molecular weight peak is observed.

BioEcho extraction kit suitable for gDNA purification that can be used in Twist FlexPrep UHT Library Preparation.

Inline Barcode Adapters & Normalization

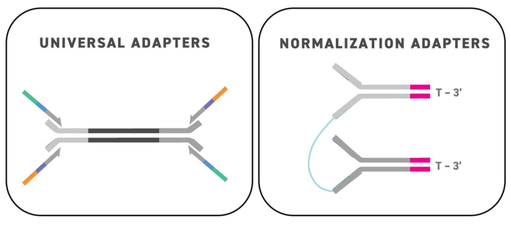

Figure 4. Structure of Normalization Adapters in Comparison to Universal Adapters. Twist FlexPrep Normalization Adapters feature a linked adapters and introduce an inline barcode (pink) to allow for multiplexed NGS library pooling for cleanup and PCR. Twist FlexPrep Normalization Adapters are compatible with Twist UDI primer system.

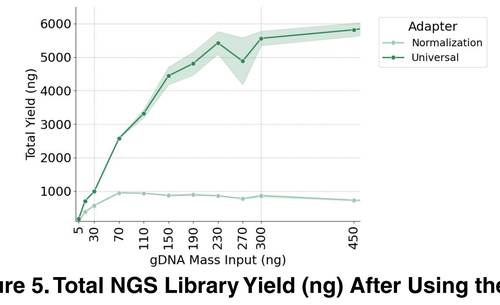

Figure 5. Total NGS Library Yield (ng) After Using the Twist FlexPrep UHT Library Preparation Kit with Various Input gDNA Masses. Individual gDNA library pools were prepared with the Twist FlexPrep UHT Library Preparation Kit and ligated with Twist Universal Adapters or Twist FlexPrep Normalization Adapters. After six cycles of PCR, libraries were quantified and total yield was calculated.

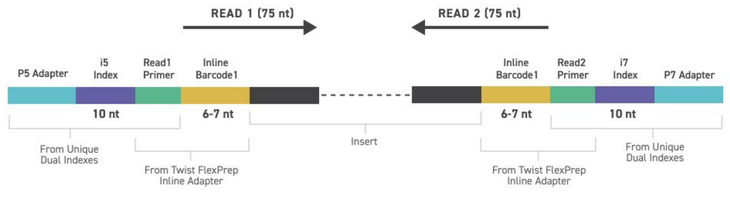

Figure 6. Twist FlexPrep UHT Library Structure. After ligation with Twist FlexPrep Normalization Adapters and indexing with UDI primers by PCR, library molecules will be flanked by hybridization adapters (P5/P7). Molecules will contain standard i5 and i7 indices used for pool-level demultiplexing, and inline with read 1 and read 2 will contain 6-7 bp inline barcodes for sample-level demultiplexing within each pool.

- Inline barcodes on adapters allow for pooling of samples post-ligation.
- Normalization adapters convert independent of insert concentration.

Library Preparation and Target Enrichment (TE)

- 96 bovine samples processed with FlexPrep UHT Library Preparation Kit.
- All 96 samples captured together in a single TE reaction with 9 Mb panel.
- Twist developed blockers designed for the Bovine genome used in TE

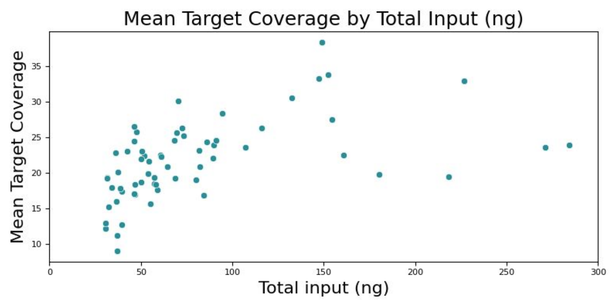

Figure 7. Mean Target Coverage by Total DNA Input. Scatter plot of mean target coverage by Total DNA input for library preparation of FlexPrep UHT samples after target enrichment. Samples processed with downsampling for 75x coverage. Only samples with input of 30-300 ng are plotted.

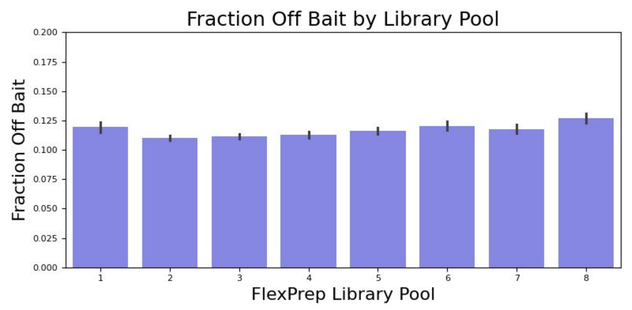

Figure 8. Fraction Off Bait by FlexPrep Library Pool. Bar plot of fraction of reads that are off bait by library pool for library preparation after target enrichment. Each FlexPrep UHT library pool consists of 12 samples. Samples processed with downsampling for 75x coverage. Blockers developed by Twist specific to the Bovine genome were used for this experiment. Only samples with input of 30-300 ng are plotted.

- Bovine tissue isolated using the Allflex TSUs and BioEcho Extraction can be successfully processed into NGS libraries with FlexPrep UHT Library Preparation Kit.
- Samples with higher DNA mass inputs do not exhibit significantly higher mean target coverage, demonstrating normalization
- Fraction off bait are low using FlexPrep UHT Library Preparation and Target Enrichment workflow with Bovine-specific blockers.

Genotyping

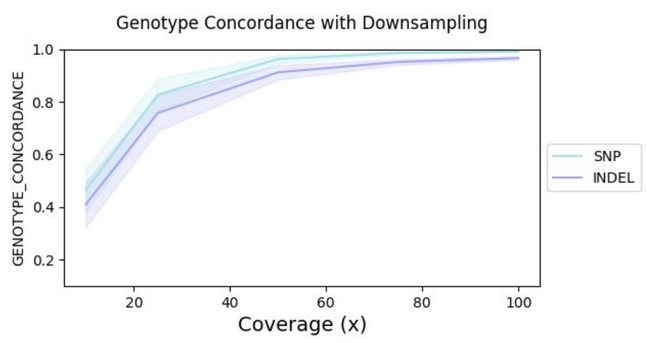

Figure 9. Genotype Concordance with Downsampling. Line plot with standard deviation of genotypes concordance between FlexPrep sample calls and ground truth PCR-free WGS calculated for both SNPs and INDELs. Concordance was calculated using the FlexPrep sample after different levels of downsampling.

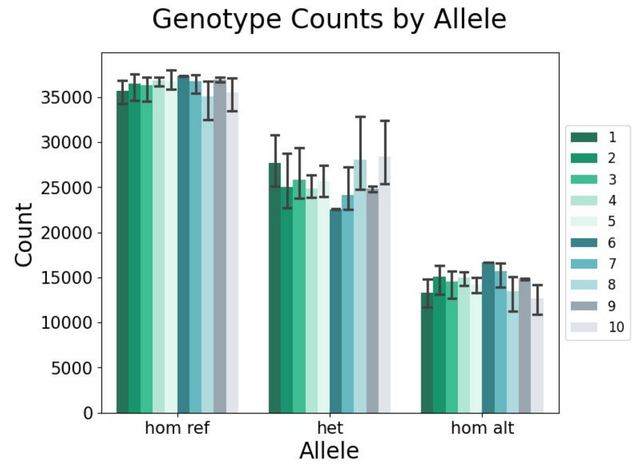

Figure 10. Genotype Counts by Allele. Bar plot of Genotype Counts for ~76k sites found in 9 Mb panel for 10 bovine samples. Each sample processed with at least six replicates. After FlexPrep UHT Library Preparation and Target Enrichment workflow, samples were processed with downsampling to 100x coverage. Allele calls: *hom ref* is homozygous reference, *het* is heterozygous, *hom alt* is homozygous alternate.

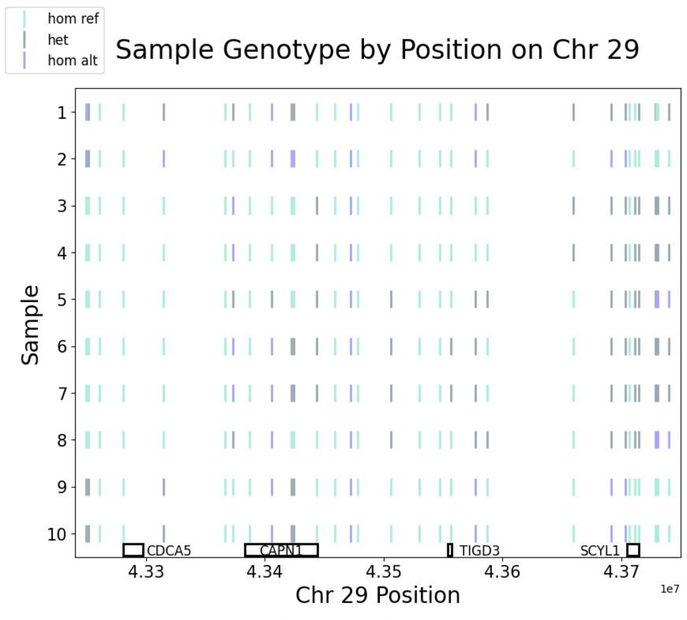

Figure 11. Genotype Calls on Section of Chromosome 29. Genome plot with genotype calls from 10 bovine samples on regions of chromosome 29 from 43.23 Mb to 43.75 Mb. Genotype calls are colored according to the legend above, demonstrating the diversity of calls observed from these individuals at this region. Allele calls: *hom ref* is homozygous reference, *het* is heterozygous, *hom alt* is homozygous alternate.

- FlexPrep UHT Library Preparation and Target Enrichment can be used to support genotyping.
- Genotype calls show high concordance with ground truth datasets.
- Genotype calls can differ by individual and demonstrate diversity observed within a population.

Conclusions

We describe the Twist FlexPrep UHT Library Preparation and Target Enrichment workflow performed on gDNA extracted from bovine ear punches. Utilizing novel normalization adapters with inline barcodes allows the early pooling of samples and self-normalization to reduce the total number of steps, reagents, and consumables needed for the user. Specifically, the multiplexing technology allows for 12 samples to be pooled after ligation and for 96 samples to be pooled into a single target enrichment reaction, which reduces the total number of workflow steps, reagents, and consumables needed. Here we demonstrate the utility of FlexPrep UHT for generating NGS libraries from bovine ear punch samples that can be easily processed for genotyping purposes.

Conflict of Interest Statement

All authors are employees and shareholders of Twist Bioscience.
Twist Bioscience and the Twist logo are trademarks of Twist Bioscience Corporation. All other trademarks are the property of their respective owners.