

Josh Franken, Max Boeck, Cole Rohrbaugh, Ariele Hanek, Thomas Harrison, Amy Emerman, and Heather Maasdorp – Watchmaker Genomics, Boulder, CO

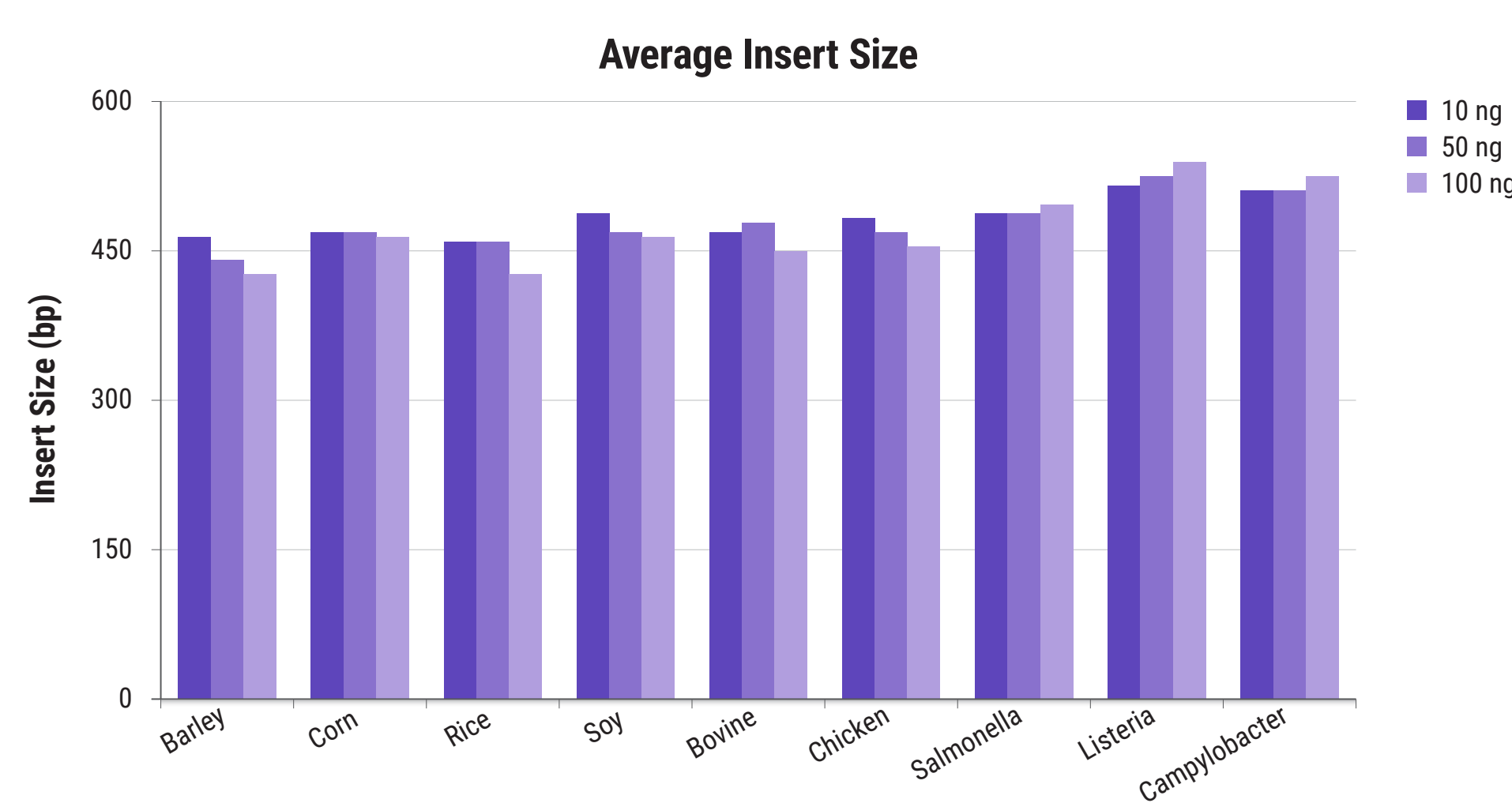
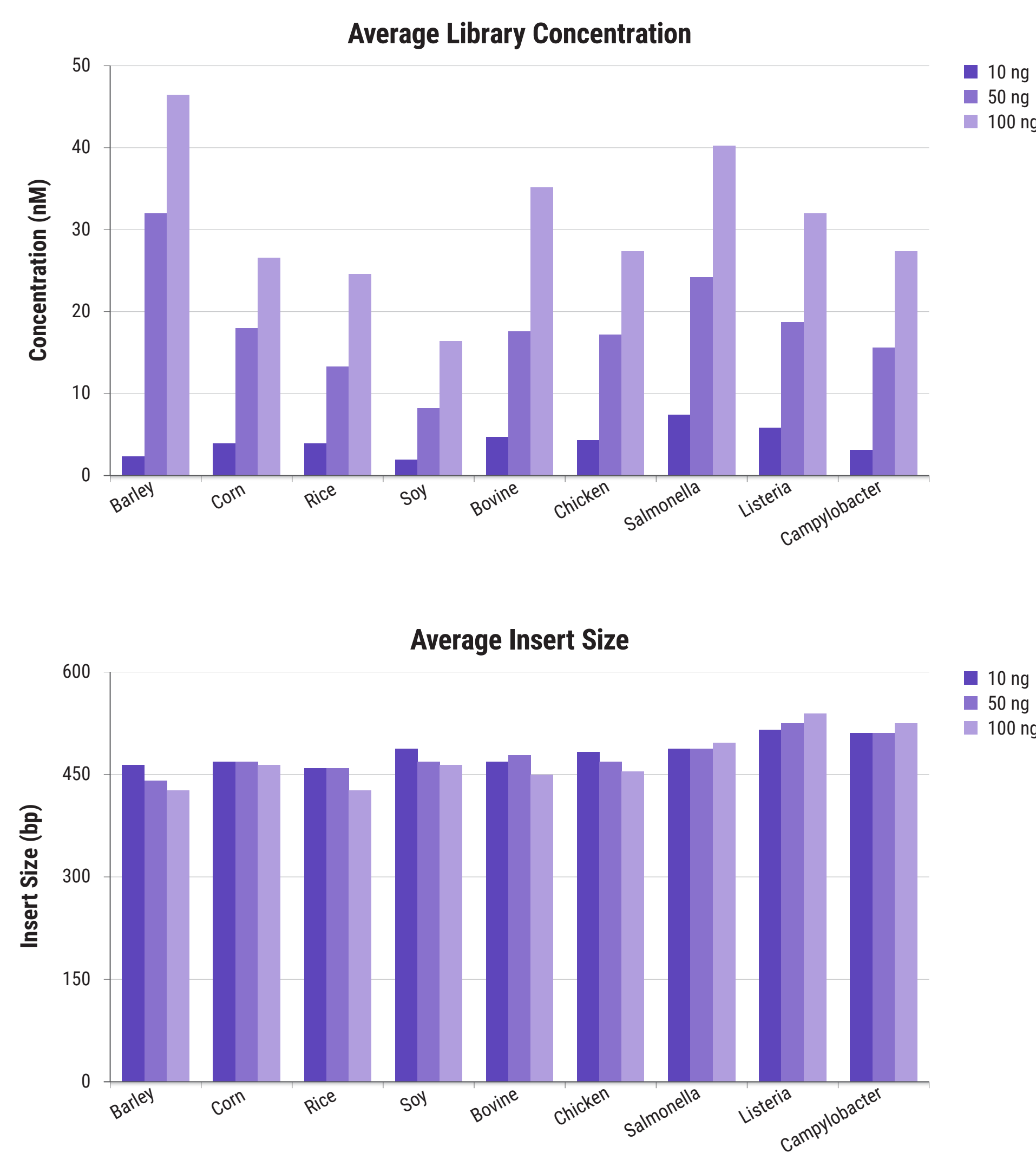
Introduction

Agrigenomics is transforming crop and livestock improvement by using genomic data to quickly identify traits linked to productivity, resilience, and disease resistance. As agriculture faces rising food demand, climate pressure, and the need for more sustainable practices, Next Generation Sequencing (NGS) has the opportunity to help identify genes and aid in breeding programs, helping researchers and producers develop healthier, more resilient plants and animals.

In order to facilitate NGS of a range of agriculturally important species, this study assessed the suitability of a single, standardized NGS library prep protocol for a range of animal, plant, and bacterial species. To enable high-throughput processing, the method was demonstrated using an automated liquid handling platform.

Library Preparation and Characterization

Using the Watchmaker Library Prep Kit with Fragmentation automated on a Hamilton NGS STAR MOA, we generated over 90 libraries from eight genomically diverse organisms in a single run. All samples and input masses generated sequenceable (>4 nM) library concentrations. As expected, when using a fixed number of PCR cycles, higher input masses yielded proportionally higher concentration libraries (Figure 1) while maintaining consistent library sizes (Figure 2).



DNA Library Prep with Fragmentation is tunable for 200 – 700 bp inserts by adjusting conditions, with optimal size depending on sequencing goals. For WGS, inserts must average >300 bp and avoid short fragments that bias sequencing toward shorter reads. This is achieved through lighter fragmentation and a stricter SPRI size selection, improving sequencing efficiency by reducing wasted bases. Consistent fragmentation across a 10-fold input range delivered average fragment sizes >450 bp (Figure 2), supporting reliable WGS performance.

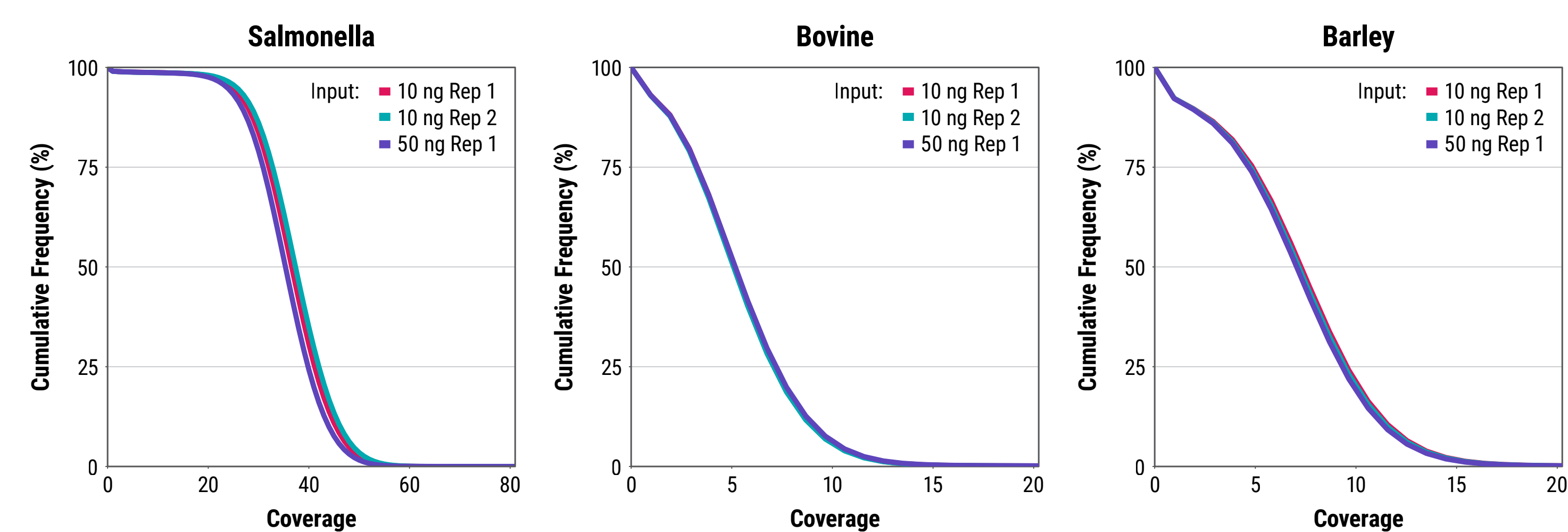


Figure 3. Cumulative coverage plots demonstrate high consistency within genome replicates. Cumulative coverage was plotted for each replicate of salmonella, bovine, and barley genomes. Coverage is shown for two replicates for 10 ng inputs (pink and teal) and one replicate of 50 ng (purple) for each genome.

Sequencing Results

Libraries were pooled to target a 5X mean coverage for eukaryotic genomes and 20X mean coverage for prokaryotic genomes, based on estimated genome sizes and library concentrations (Table 1). Sequencing confirmed average insert sizes ≥ 300 bp with strong reproducibility and >99% alignment for most samples. Rice DNA was the exception with less than 70% aligned reads. This was later determined to be due to contamination by *Pseudomonas fuscovaginae* (~30% of reads). This illustrates a common agrigenomics challenge of pathogen co-isolation and its impact on sequencing depth requirements. After filtering, >99% of rice reads aligned correctly.

Table 1. Average Sequencing metrics across three replicates for each genome

Species	Genome Size	Average GC Content	Percent Reads Aligned	Median Coverage
Rice	0.5 Gb	44%	67.5%*	5X
Corn	2.5 Gb	47%	99.6%	4X
Bovine	2.7 Gb	42%	99.9%	5X
Soy	1.2 Gb	43%	99.0%	5X
Chicken	1.2 Gb	50%	99.1%	7X
Barley	5.0 Gb	48%	99.7%	7X
Campylobacter	1.6 Mb	30%	99.8%	21X
Listeria	2.9 Mb	37%	99.7%	68X
Salmonella	5.0 Mb	53%	99.9%	36X

*Rice had low alignment due to contamination with a common rice pathogen, *P. fuscovaginae*.

Prokaryotic genomes exhibited reproducible coverage, with all replicates of listeria and salmonella exceeding 20X mean coverage. Campylobacter showed lower-than-expected genome coverage possibly due to a higher-than-anticipated plasmid content skewing the effective genome representation. Salmonella cumulative coverage distribution plots (Figure 3) demonstrated high reproducibility across all libraries and replicates. All eukaryotic genomes exceeded 5X coverage with the exception of corn with 4X mean coverage. We speculate this is due to either a lower quality reference genome, or possibly deviation of the reference genome from the genome that was sequenced – highlighting another challenge associated with agronomic sequencing.

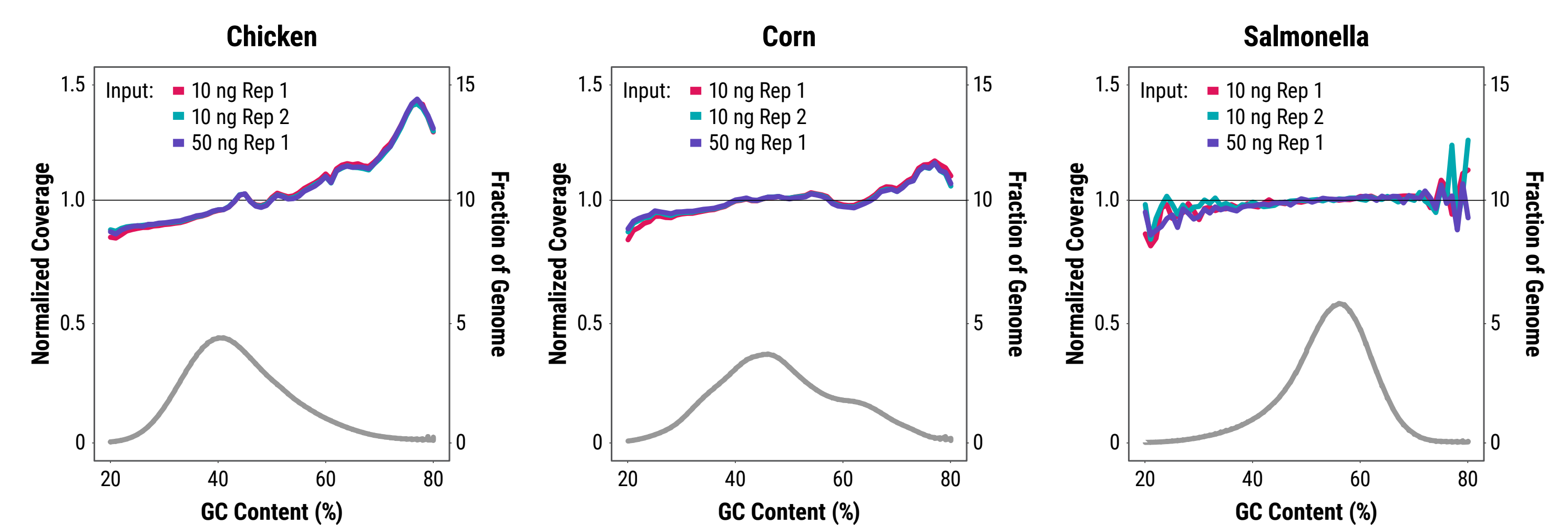


Figure 4. Consistent genome GC coverage across replicates. Normalized GC coverage for chicken, corn, and salmonella genomes. Coverage is shown for two replicates for 10 ng inputs (pink and teal) and one replicate of 50 ng (purple). GC coverage is highly consistent between replicates.

Conclusions

- A single automated workflow enabled high-throughput library prep for plant, animal, and microbial genomes
- Consistent performance across a 10X input range and diverse species eliminated the need for assay-specific optimization
- Libraries showed uniform fragment sizes, high complexity, and minimal GC bias across genomes of varying size and composition
- Compatibility with automated liquid handling solutions supports scalable, cost-effective sequencing for agrigenomic applications



To download a copy of this poster, scan code or visit watchmakergenomics.com/PAG2026.