Rapid, Ultra-Multiplexed Amplicon-Based Targeted Sequencing Approach for High-Throughput Genotyping and Phylogenetics

Wissam Hamou, Yang Lily Liu, Lucie S. Lee, Jeffery Liu, Lifeng Lin, Guoying Liu, Zhitong Liu*

Paragon Genomics, Inc. Hayward, CA 94545 *Corresponding author

Abstract

Next-generation sequencing (NGS) technologies have accelerated research efforts in plant and animal genomics. While whole genome sequencing is commonly used in agriculture, a targeted sequencing approach is often more suitable, time- and cost-effective in many applications[1]. Targeted genotyping by sequencing (GBS) can deliver reliable, high marker call rates for specified SNPs at high throughput and lowered cost. Amplicon-based targeted NGS library preparation methods are preferred by many as they are more convenient and efficient to perform. However, conventional amplicon-based methods suffer from drawbacks such as high background, poor coverage uniformity, and limited multiplexing of amplicons.

Paragon Genomics' AgriType™ targeted genotyping solution is powered by the patented CleanPlex® technology. AgriType is a high-performing, amplicon-based target enrichment and library preparation solution for NGS. Its unique background cleaning chemistry enables tens of thousands of amplicons to be pooled in a single reaction to enrich for a large number of targets required by many high-throughput genotyping studies. AgriType features a streamlined workflow with Illumina or Ion Torrent compatible libraries via a single tube in 3 steps and under 3 hours. We present results from experiments demonstrating the robust performance of the AgriType solution. AgriType target-enriched NGS libraries have low PCR background, no apparent GC bias, and can achieve >95% coverage uniformity using as little as 1 ng of DNA. The technology has a limit of detection below 1% and high variant detection accuracy, as well as high mapping and on-target rates over a wide range of target numbers and input DNA amount.

Simple and Fast Work Flow

The Paragon Genomics CleanPlex workflow is designed for straightforward and rapid preparation of NGS libraries for the Illumina platform. Compared to other methods which can take up to 4 days, this workflow yields sequencing- ready libraries in under 3 hours with just 45 minutes of hands on time. The process is easily automatable on standard liquid handling platforms via standard NGS protocols.

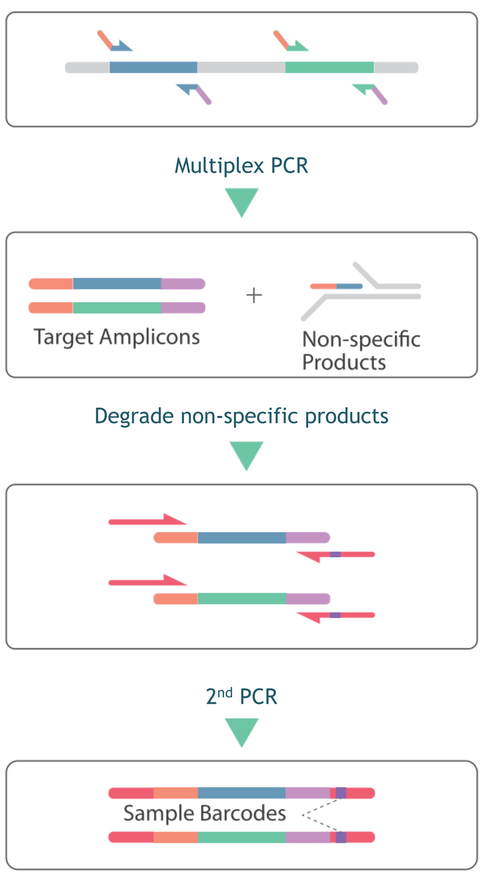

Figure 1: The Paragon Genomics Library Preparation workflow begins with amplification of target sequences through multiplex PCR using oligonucleotides that contain a partial adaptor sequence. Primer-dimers, heteroduplexes, and other side products are then degraded and removed from the reaction solution using the AgriType Digestion Reagent. Finally, Illumina or Ion Torrent adaptor sequences and sample barcodes are added to the library through a second round of PCR, which anneals to the partial adaptor sequence added to each amplicon during the multiplex PCR step.

Background Cleaning

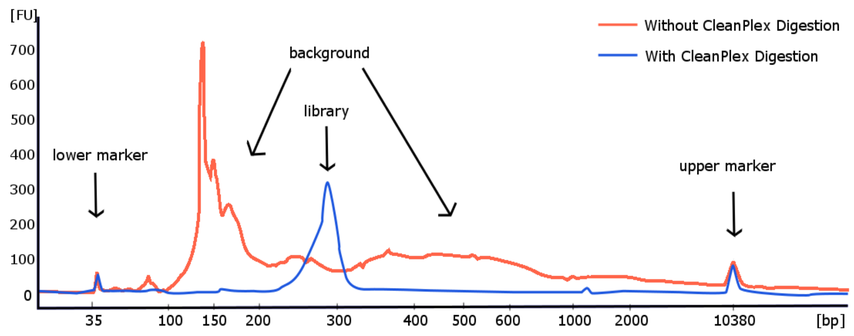

Figure 2: Comparison of Agilent BioAnalyzer traces of samples treated (blue plot) and not treated (red plot) with AgriType Digestion Reagent. The proprietary CleanPlex Digestion Reagent is essential for removing undesired side products formed during the multiplex PCR amplification of target sequences. The sample not treated with the AgriType Digestion Reagent contains a large adaptor-primer peak as well as side products throughout the measurable size spectrum, effectively burying the target library. Whereas, the sample that was treated with the AgriType Digestion Reagent shows a library peak at the expected size range, with little side product.

Panel Performance Metrics

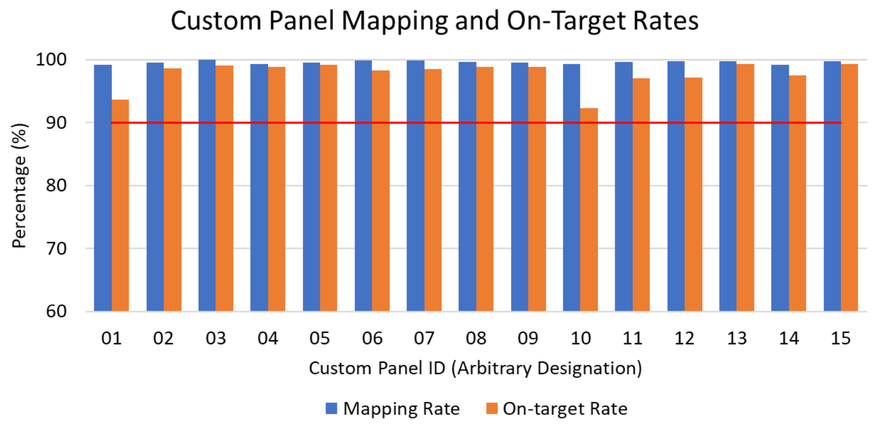

Figure 3: Examples of mapping and on-target rate for custom panels designed by Paragon Genomics showing the consistency in performance in various custom panels. Every panel shown has mapping and on-target rates above 90%.

Uniformity and GC Bias

AgriType mPCR Reagent is optimized to minimize high GC bias and therefore consistently yields high uniformity for panels of various sizes and also for samples with a wide range of DNA input. For certain applications, input as low as 0.1 ng DNA can still give quality sequencing data. NGS library preparation is a fundamental and critical part of sequencing studies, as any information lost or bias introduced during the process will not be restored and the sequencing data is only as good as the quality of the NGS libraries prepared[2]

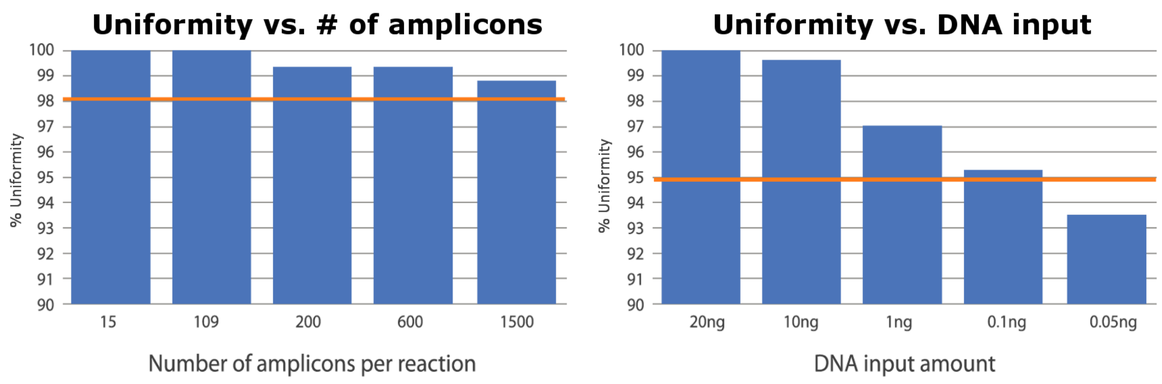

Figure 4: Libraries prepared using Paragon Genomics' AgriType solution produce > 95% uniformity at 0.2X mean reads. **(Left)** Uniformity of >98% above 0.2X mean reads is maintained in Paragon Genomics' designed libraries of 15 to 1500 amplicons. **(Right)** With a 600 amplicon panel, AgriType solution also produces >95% uniformity at 0.2X means reads using as low as 0.1ng of input genomic DNA.

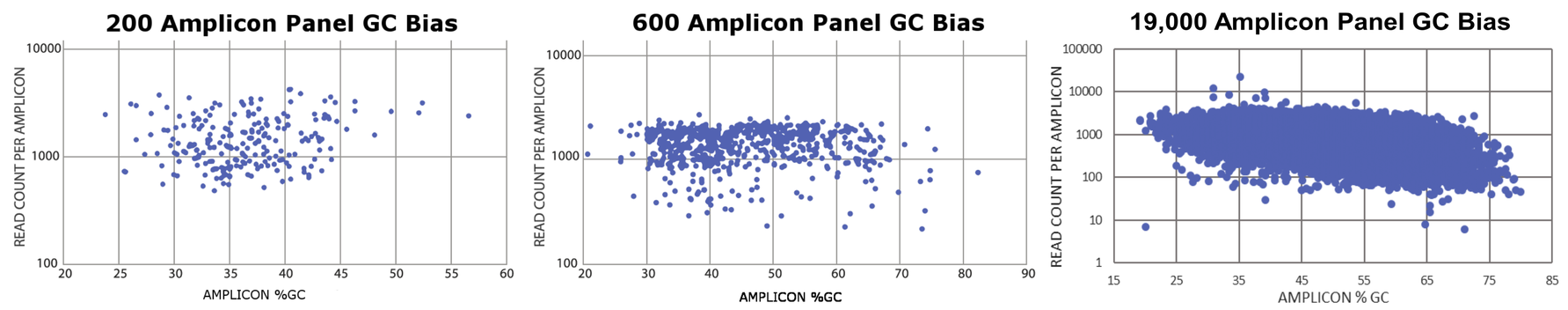

Figure 5: Coverage Depth vs. Amplicon %GC Content of libraries made with a 200 amplicon panel, a 600 amplicon panel, and a 19,000 amplicons panel. All libraries demonstrate even coverage of amplicons with ~ 25-45% GC content, ~20-75% GC content. And ~20-80% GC content respectively, with little to no GC bias.

Applications

Organisms	Targets	Amplicons	Purpose
Mouse	71	72	Gene editing (CRISPR)
Bacteria	107	685	Species identification
Bacteria	342	348	Species identification
Wolf	475	425	SNP identification
Tomato	436	436	Genotyping
Fruit Flies	879	879	Phylogenomics analysis[3]
Rhino	1017	1017	Genotyping
Cotton	2088	2088	Genotyping
Human	19766	19766	SNP identification

Figure 6: The Paragon Genomics panel design, technology, and workflow have been successfully applied to human, animal, insect, and microbial genomes. Table shows examples of custom libraries designed by Paragon Genomics spanning a variety of organisms and a wide range in number of covered amplicons.

References

1) Target-enrichment strategies for next-generation sequencing. Mamanova et al. Nat. Methods. (2010) 7(2):111-8

2) Library construction for next-generation sequencing: Overviews and challenges. Head et al. Biotechniques. (2014) 56(2): 61-77

3) HiMAP: Robust phylogenomics from highly multiplexed amplicon se quencing. Dupuis JR et al. (2018) Mol Ecol Resour. 18(5):1000-1019

Conclusions

Paragon Genomics' NGS library preparation chemistry system reduces the workflow of creating high quality and high amplicon count libraries for agrigenomics applications.

- AgriType Digestion Reagent significantly reduces background products prior to sequencing, allowing for more samples per sequencing run, reducing costs.
- AgriType panels produce high mapping and on-target rates over a wide range of target numbers, even at low sample input.
- Proprietary multiplex PCR reagents and expert primer design result in >95% uniformity and little to no GC bias.
- AgriType solution demonstrates a limit of detection below 1%, allowing for detecting of rare SNPS
- There are no specialized equipment requirements other than a thermocycler.
- The workflow is a simple one-tube protocol, minimizing sample loss, reducing consumable waste, and streamlining automation.
- DNA sample to sequence ready libraries in 3 hours with minimum hands on time.
- AgriType panels are compatible for sequencing on both Illumina and Ion Torrent platforms for additional flexibility.
- Custom panels can be designed to target DNA from any organism in less than 2 weeks.