applied biosystems

AgriSeq™ ULTRA: A Fast, Simple, High-Throughput Workflow for Targeted Genotyping-By-Sequencing of Aquaculture and Plant Samples

Jennifer Lavoie, Krishna Reddy Gujjula, Angela Burrell, Tia Dunbar, Kellsie Kice, Gail Ulloa, Tiffany Wang
Thermo Fisher Scientific, 2130 Woodward Street, Austin, TX, USA, 78744

ABSTRACT

Targeted genotyping-by-sequencing (GBS) is a robust agrigenomics tool that is used to identify SNPs for breeding selection or trait identification. High-throughput targeted GBS applications rely on rapid, high-accuracy sequencing results of thousands of markers per day with minimal labor and laboratory inputs.

The AgriSeq™ ULTRA targeted GBS kit is the next generation of AgriSeq™ library prep products. The updated protocol's partial-combinatorial dual barcodes allow for processing of up to 3,072 uniquely barcoded sample libraries, while reducing the amount of barcode material needed to only 4 plates and 8 tubes. Up to 3,072 libraries can then be sequenced on a single Ion 550™ chip, with two chips sequenced per day, allowing processing of a total of up to 6,144 samples daily per Ion GeneStudio™ S5 System. Updates to the Torrent Suite analysis software and the addition of the AgriSeqVariantCallerLite™ plugin allow full data processing in approximately 24 hours or less, allowing for reduced wait time between sequencing and acquiring results.

AgriSeq™ ULTRA is compatible with plant leaf punches, and aquaculture sample types, such as pleopods and fin punches. Reliable genotype calls are generated with both purified nucleic acids and direct lysates due to our novel AgriSeq™ Anti-Inhibitor Solution, which negates effects of PCR inhibitors carried over from crude extraction and allows highly inhibited samples to be used for targeted GBS library preparation. Here we illustrate how AgriSeq™ ULTRA successfully generates data from plant leaf punch, and salmon fin punch lysates extracted using the AgriSeq™ Genomic DNA Extraction Kit.

INTRODUCTION

With targeted GBS becoming a go-to method for SNP genotyping, there is an increasing need for higher throughput workflows with faster analysis times. The ability to multiplex thousands of samples in a single sequencing run is critical for cost-effective marker-assisted breeding.

To allow for a more flexible workflow, the AgriSeq™ ULTRA Library Kit (*A40002302*) is optimized to work both with direct lysate sample and extracted, purified DNA. Using the AgriSeq™ Genomic DNA Extraction Kit (*A66428*), direct lysis can be performed at ~10 min per plate and multiple plates can processed at a time using automation and multiple thermal cyclers. The total time from sample to normalized library pool is ~6 hours for 3072 samples using direct lysis and automation.

In this study, we explore the workflow for the new AgriSeq Ultra™ Library Kit and detail some results generated using the new kit for plant leaf punches and aquaculture samples.

MATERIALS AND METHODS

Kit and Contents

AgriSeq™ ULTRA Library Kit (**Figure 1**) contains enough reagent for library preparation of 9600 samples, with overage for automated workflows (**Table 1**)

Figure 1. AgriSeq™ ULTRA Library Kit

Table 1. Kit Components

Component	Quantity per kit	Storage
Library reagents[1]		
AgriSeq™ Genomic Platinum™ SuperFi™ U Multiplex Master Mix	54 mL	–25°C to –15°C
AgriSeq™ ULTRA Enhancer	40 mL	
AgriSeq™ Anti-Inhibitor Solution	30 mL	
Normalization Master Mix	13.75 mL	
AgriSeq™ Nuclease-Free Water	30 mL	15°C to 30°C[2]

[1] Shipped at –20°C. Store as directed.
[2] Can be stored at –25°C to –15°C or 2°C to 8°C for convenience.

Barcodes

The AgriSeq™ ULTRA workflow utilizes a partial combinatorial barcoding method (**Figure 2**). This method allows for higher multiplexing with less barcode material. The AgriSeq™ ULTRA Dual Barcode Adapters Kit (*A40001408*) provides material for 9600 reactions and up to 3072 combinations per sequencing run. The kit contains 4 plates of START barcode adapters that identify the well on the library plate and 8 tubes of END barcode adaptors that identify the library prep plate from which the sample came (**Figure 3**).

Figure 2. Partial combinatorial barcode schematic

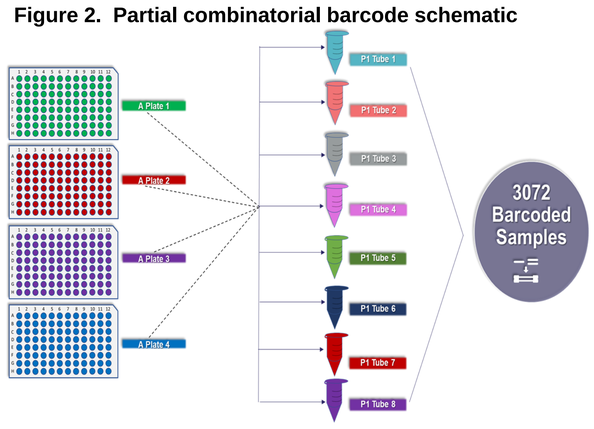
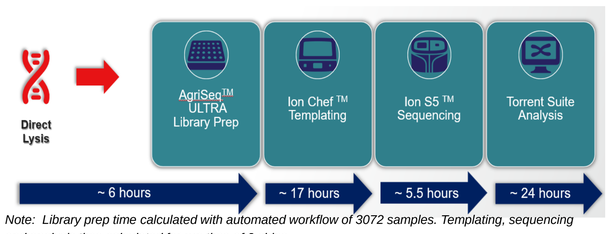


Figure 3. AgriSeq™ ULTRA Dual Barcode Adapters

AgriSeq™ Ultra Workflow

For this set of experiments, direct lysate sample was used. Direct lysates were not quantified prior to library preparation and 3 uL of the lysates were used in library preparation. The full workflow for 3072 samples requires ~6 hours, which includes the time the plates were on automated instrumentation and the thermal cycling times. The full workflow for 3072 samples from punch to genotype can be completed in ~2 days **(Figure 4).**

Figure 4. The AgriSeq™ ULTRA workflow



Note: Library prep time calculated with automated workflow of 3072 samples. Templating, sequencing and analysis time calculated for run time of 2 chips.

Direct Lysis

For all samples, direct lysis was performed using the AgriSeq™ Genomic DNA Extraction Kit using two ~3 mm punches or one ~5-6 mm punch and the conditions shown below (**Table 2**).

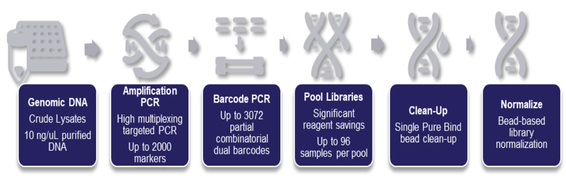
Table 2. Direct lysis conditions

	Aquaculture	Plants
High Temperature Cell Lysis	5-10 minutes	5-15 minutes
Reduced Temperature Rest Incubation	3 minutes	3 minutes

Library Preparation

Using panels custom designed for the AgriSeq™ ULTRA workflow, libraries were prepared using the AgriSeq™ ULTRA Library Kit along with the AgriSeq™ ULTRA Dual Barcode Adapters Kit as indicated in the user guide. Master mix was prepared using the AgriSeq™ Anti-Inhibitor Solution since direct lysates were being used as the input sample. All library preparation was performed in 384-well plates.

Figure 5. AgriSeq™ ULTRA library prep



Libraries were pooled 96-samples per pool prior to library clean-up and normalization. Library clean up was performed using our room-temperature stable AgriSeq™ Genomic Pure Bind Beads (*A40001417*). Libraries were normalized to ~200 pM using the AgriSeq™ Normalization Reagents (*A34140*). Both protocols were completed as directed in the AgriSeq™ ULTRA user guide.

Custom AgriSeq™ ULTRA Panels

The experiments were performed using primer pools, designed specifically for the ULTRA workflow, testing both a representative plant and aquaculture sample type (**Table 3**)**.** Forward and Reverse pools are stored in separate tubes to reduce instances of primer dimerization. Because these experiments were meant to validate the workflow rather than the panels, no markers were redesigned after the initial design for panel, and thus, were not

Table 3. Experiment amplification panels

Panel	Markers
Pepper	131
Salmon	993

NGS Sequencing and Data Analysis

Prior to sequencing, pools were diluted 1:1 with TE buffer to allow for sequencing at a concentration of ~ 100 pM. Libraries were pooled and sequenced to accomplish a minimum of ~100x target coverage.

All sequencing was conduct on an Ion GeneStudio™ S5 System using Ion 550™ or Ion 540™ chips. Analysis was performed using the AgriSeqVariantCallerLite™ and coverageAnalysis™ plugins. Implementation of AgriSeqVariantCallerLite™, along with some modifications to the basecalling pipeline, allow for completion of the on-instrument analysis and generation of genotyping calls made available in the AgrisumToolkit™ plugin in ~24 hours or less.

RESULTS

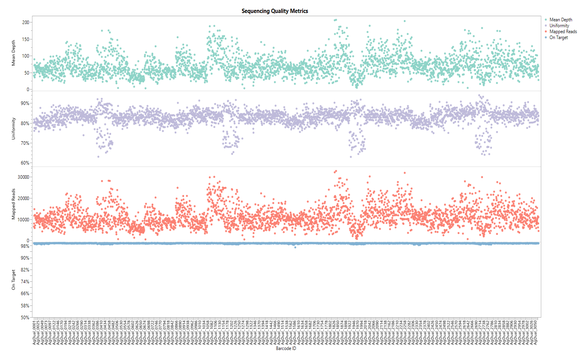
Direct Lysate Pepper Samples

In order to demonstrate the AgriSeq™ ULTRA workflow using direct lysates from plants, direct lysis was performed on 3072 leaf punches and these pepper lysates were prepped using the AgriSeq™ ULTRA Library workflow. Sequencing was then completed on an Ion 540™ chip. The leaf punches came from 768 unique individuals and thus the lysates are extraction replicates. Mean quality metrics for the run performed well with sample call rate above 90% (**Figure 6**).

Figure 6. Summary of AgrisumToolkit™ metrics for pepper.

Examining Mean Depth, Mapped Reads, Uniformity, and On Target metrics across all samples, we see excellent performance was achieved, with most samples performing well enough to make genotype calls (**Figure 7**). Genotype call performance was exceptional, with 98.7% of samples generating 80% or greater call rate and 94% of samples generating 90% or greater call rate.

Figure 7. General quality metrics for pepper.



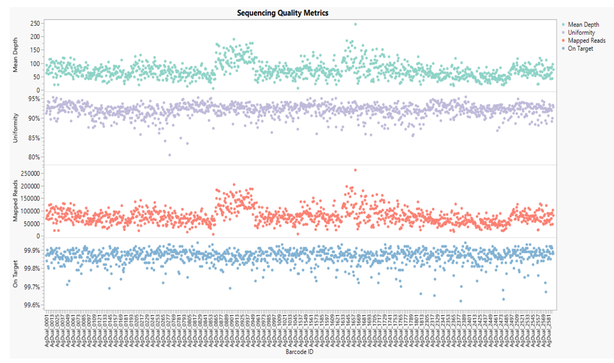
Direct Lysate Salmon Samples

In order to demonstrate the AgriSeq™ ULTRA workflow using direct lysates from aquaculture samples, direct lysis was performed on 1152 fin punches and these salmon lysates were prepped using the AgriSeq™ ULTRA Library workflow. Sequencing was then completed on an Ion 550™ chip. Mean quality metrics for the run performed well with sample call rate above 90% (**Figure 9**).

Figure 9. Summary of AgrisumToolkit™ metrics for salmon.

Examining Mean Depth, Mapped Reads, Uniformity, and On Target metrics across all samples, we see excellent performance was achieved, with most samples performing well enough to make genotype calls (**Figure 10**). Genotype call performance was exceptional, with 98.6% of samples generating 80% or greater call rate and 94.3% of samples generating 90% or greater call rate.

Figure 10. General quality metrics for salmon.



CONCLUSIONS

The AgriSeq ULTRA™ workflow is a fast, simple, high throughput workflow that allows processing of plant and aquaculture samples regardless of whether input DNA is clean extract or direct lysate. No normalization prior to library prep is required. The workflow from direct lysis through library prep can be completed in ~6 hours and from sample to genotype in ~2 days.

We have shown that the AgriSeq ULTRA™ workflow manages inhibition generally found in direct lysis samples and provides a rapid, robust genotyping solution for plant and animal customers.

TRADEMARKS/LICENSING

ThermoFisher SCIENTIFIC