Comparison of library cleanup reagents for the AgriSeq HTS Library Kit

Introduction

Applied Biosystems™ AgriSeq™ targeted genotyping by sequencing (GBS) is designed to help accelerate plant and animal breeding programs. Targeted GBS utilizes large sets of primers (tens to thousands) in highly multiplexed PCR to generate amplicons that provide sequences to make genotype calls on a predetermined set of alleles. We have demonstrated the efficacy of this procedure on DNA from a large sampling of both animal and plant species. In addition, through a sequence-tagging method known as barcoding, this methodology can be applied to analyze large numbers of samples in the same sequencing run. This workflow requires a cleanup step, which removes small DNA fragments as well as other products that can interfere with the output of the final sequencing analysis. For ease of use and to enable the sample to be run on common liquid-handling platforms, the method of choice for this cleanup step is based on solid-phase reversible immobilization (SPRI) technology. In this article, we compare the performance of two SPRI-based cleanup reagents used with the Applied Biosystems™ AgriSeq™ HTS Library Kit.

Materials and methods

Overview of workflow

The library preparation workflow is described in detail in the AgriSeq HTS Library Kit user guide (Pub. No. MAN0015971, available at **thermofisher.com**). In brief, the workflow begins with aliquots of DNA from each sample (optimally, 10 ng) being subjected separately to massively parallel PCR amplification with target-specific AgriSeq primer panels. Following amplification, the libraries are treated with an enzyme to prepare amplicons for ligation to adapters containing unique barcodes attached to universal priming sites. Libraries are then pooled into smaller sets to allow for simplified cleanup that is performed in a 96-well plate using significantly less reagents than the standard library preparation protocols. The pooled libraries then go through two rounds of bead-based cleanup in order to remove the non–nucleic acid components of the amplification reaction as well as small nonspecific PCR products and other DNA fragments, which can interfere with the sequencing process. The pooled libraries are each normalized to a concentration of 100 pM using a bead-based normalization reagent. Equal volumes of these intermediate library pools are then combined to form a master pool, which is used for template preparation on the Ion Chef™ Instrument and subsequently sequenced using the Ion 540™ or Ion 550™ Chip on the Ion GeneStudio™ S5 System. Since they all contain unique barcodes, data for each sample can be parsed at the analysis stage.

Testing alternative SPRI systems

To evaluate the performance of library purification reagents used with the AgriSeq HTS Library Kit, we tested three GBS panels (Table 1). Triplicate library preparations were performed with each panel. All evaluations were conducted with the 96-well plate protocol, with the exception that sample pooling volumes were increased so that the final pools could be split to create two identical sets. One set was processed with the AMPure™ XP reagent from Beckman Coulter and the other processed with the CleanNGS™ reagent from Bulldog Bio. Library cleanup steps were performed according to the AgriSeq HTS

Library Kit user guide regardless of the reagent used. Libraries from each protocol were normalized and then pooled for template preparation on the Ion Chef Instrument and sequencing on the Ion GeneStudio S5 System with the Ion 540 Chip. Analysis was performed using the Coverage Analysis and Torrent Variant Caller (TVC) plugins with the default settings for germline analysis. Statistical analysis was conducted using JMP™ software from SAS. Workflow performance was assessed by comparing call rate, run uniformity, and call concordance between replicates.

Table 1. Summary of GBS panels tested.

Panel number	Number of markers	Number of samples	Number of replicates/ sample
1	154	3	32
2	1,500	96	1
3	2,800	3	32

Results

Call rate

Call rate, the percent of markers generating a genotype call for a specific sample, was calculated for all samples tested. Mean call rates across sample technical replicates as well as a grand mean across all samples were determined to evaluate workflow performance. No significant difference in call rate performance was observed between the two SPRI methods using the 154-marker panel (Table 2). Although the difference observed between run 1 and run 2 or 3 was significant relative to the variation in each run, in practical terms it was less than what we routinely see in run-to-run variation. In addition, no significant difference between the two SPRI methods was observed when using the 1,500-marker panel or the 2,800-marker panel. For the 2,800-marker panel, the statistical difference observed between run 2 and run 1 or 3 was not practically significant. The mean call rate was >97.5% for all panels tested.

Run uniformity

Run uniformity, the percentage of target bases covered by at least 0.2x of the average base read depth, measures how evenly target amplicons are covered with reads. Uniformity less than 90% can cause marker drop-off and poor call rates. No significant difference in run uniformity was observed between the two cleanup reagents (Figure 1).

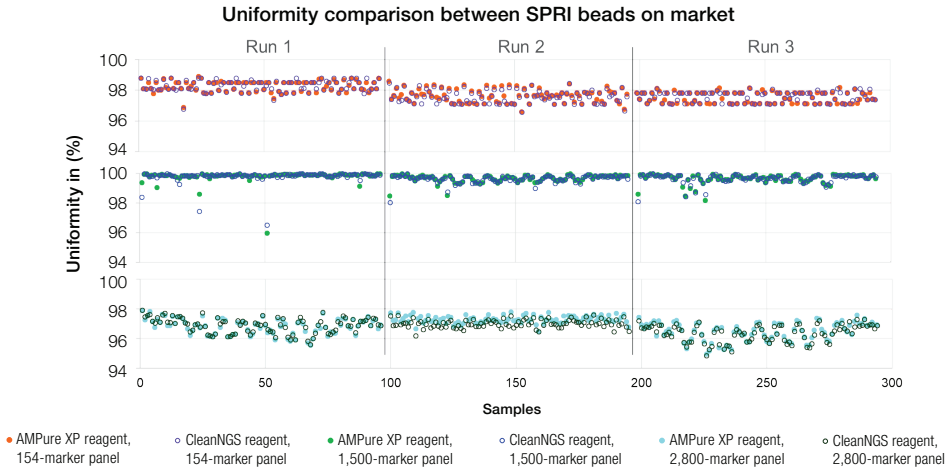

Figure 1. Comparison of run uniformities between libraries prepared with AMPure XP and CleanNGS reagents. Mean read uniformity was high (~96%) across all panels and all runs.

Table 2. Comparison of call rates between libraries prepared with AMPure XP and CleanNGS reagents.

SPRI method	Run 1 (%)	Run 2 (%)	Run 3 (%)
154-marker panel			
AMPure XP reagent	99.33 ±0.09	98.62 ±0.20	98.7 ±0.17
CleanNGS reagent	99.29 ±0.17	98.70 ±0.06	98.7 ±0.16
1,500-marker panel			
AMPure XP reagent	99.79 ±0.53	99.87 ±0.25	99.84 ±0.33
CleanNGS reagent	99.72 ±0.64	99.89 ±0.23	99.87 ±0.26
2,800-marker panel			
AMPure XP reagent	98.04 ±0.62	98.35 ±0.50	97.90 ±1.67
CleanNGS reagent	98.08 ±0.72	98.35 ±0.42	97.81 ±1.69

applied
biosystems

Concordance

Concordance of genotype calls between cleanup reagents was analyzed and showed robust, reproducible results (Table 3). Concordance of genotype calls was also analyzed between all three runs for each method and showed consistent, reproducible results within replicate runs.

Table 3. Concordance of genotype calls.

Panel size	Comparison between runs	Average concordance	Std. dev.
154	CleanNGS reagent vs. AMPure XP reagent	99.99%	0.01%
	CleanNGS reagent	99.95%	0.01%
	AMPure XP reagent	99.99%	0.01%
1,500	CleanNGS reagent vs. AMPure XP reagent	99.75%	0.06%
	CleanNGS reagent	99.72%	0.03%
	AMPure XP reagent	99.71%	0.02%
2,800	CleanNGS reagent vs. AMPure XP reagent	99.84%	0.02%
	CleanNGS reagent	99.79%	0.02%
	AMPure XP reagent	99.82%	0.01%

Conclusions

SPRI technology in NGS applications allows workflows to scale to higher sample numbers due to its ease of use and potential for automation. The AgriSeq HTS Library Kit, which includes the use of SPRI technology, is a reliable and economical solution for genotyping applications in production agriculture. Using the workflow for the AgriSeq HTS Library Kit, CleanNGS and AMPure XP reagents, which are both based on SPRI technology, were compared across amplicon panels of varying sizes (154, 1,500, and 2,800 amplicons). Equivalent performance was observed between cleanup kits with mean call rates >97.5% for all panels tested. Additionally, both cleanup reagents demonstrated robust run-to-run reproducibility and high mean call concordance between reagents. In conclusion, the CleanNGS and AMPure XP reagents can be used interchangeably in the library cleanup step of the AgriSeq HTS Library Kit.

Find out more at **thermofisher.com/agriseq**

ThermoFisher SCIENTIFIC