Application note

Simplified targeted genotyping by sequencing using the Amp-Seq Reagent System

Introduction

Targeted sequencing methods are being adopted at an increasing pace for genotyping at many agricultural biotechnology companies due to their cost-effectiveness and ease of use for the targeting and sequencing of intermediate numbers (up to thousands) of single nucleotide polymorphisms (SNPs) or insertions/deletions (indels). As new sequencing platforms become available, driving down the costs associated with sequencing, greater demand for targeted sequencing methods is anticipated.

LGC Biosearch Technologies™ developed a method to simplify and streamline targeted genotyping workflows to ensure that it is easy to adopt and scale up to industrial throughput. Here we describe the development of Amp-Seq, a new technology for targeted genotyping by sequencing (GBS), and discuss how this new technology can be used to lower cost per sample, is easily automatable for high-throughput genotyping of crop samples and produces high-quality, uniform data.

BIOSEARCH™ TECHNOLOGIES

GENOMIC ANALYSIS BY LGC

Application note

Simplified targeted genotyping by sequencing using the Amp-Seq Reagent System

Methods

A simple, cost-effective protocol for multiplexed genotyping of samples must include a workflow that:

a. is easily automatable for high-throughput processing;
b. minimises hands on time;
c. utilises reagents that can be produced in bulk at a low cost, as only a single enzyme is used;
d. reduces the amount of plasticware needed;
e. allows for pooling of indexed samples prior to cleaning and sequencing;
f. requires low volumes of reagents that can be pre-combined into a complete master mix and then stored frozen either as a ready-to-use master mix or as pre-dispensed 384-well plates;
g. utilises indexing primer pairs that are supplied as dried dual indexed primers in 384-well plates that can be stored at room temperature;
h. uses reagents that are stable at room temperature both prior to and after thermal cycling, allowing for high-throughput automation.

The workflow for Amp-Seq, our simplified targeted GBS process, is detailed in figure 1.

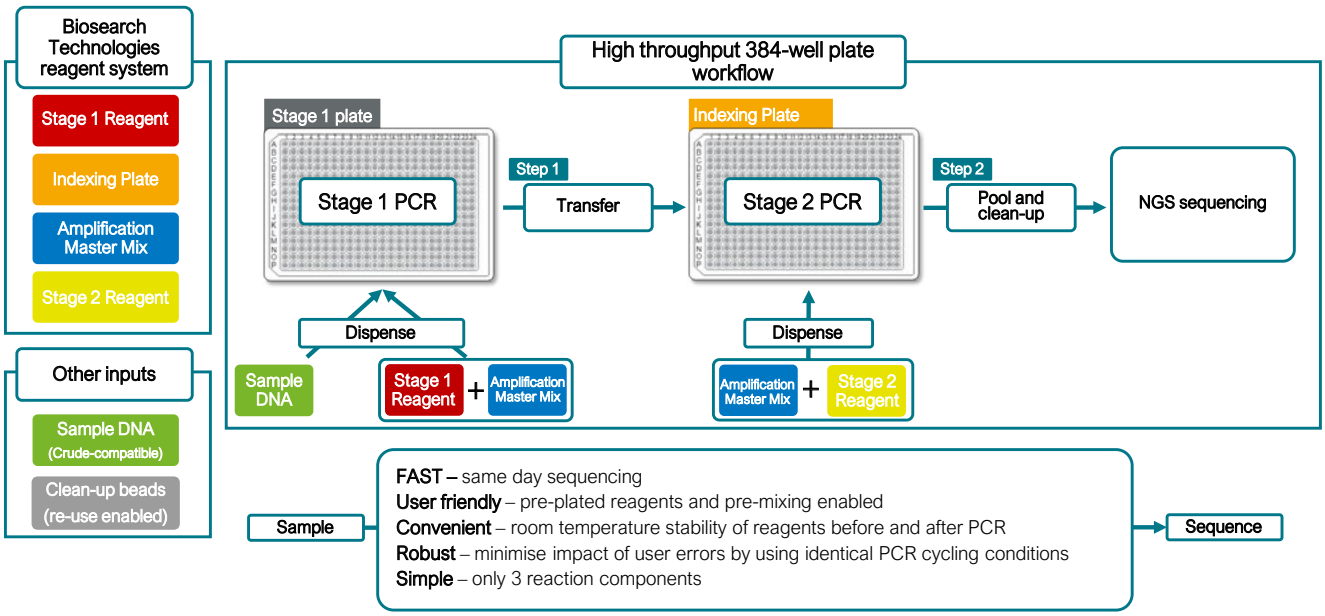


Figure 1. Schematic of the Amp-Seq Reagent System workflow. Although the workflow can be modified for specific needs, the process shown above will be described here. The process begins with the dispensing of a master mix that includes Stage 1 Reagent and Amplification Master Mix combined. This mix is aliquoted into a 384-well plate (stage 1 plate) that can either be stored at -20 °C or used within 4 hours of being dispensed when kept at room temperature. Sample DNA, either clean or crude, is transferred to the stage 1 plate and the plate is vortexed and placed into a thermal cycler for approximately 50 minutes. While the stage 1 plate is in the cycler, a second plate (Indexing Plate) is prepared. The Indexing Plate contains dried indexing primers at the required concentration. A second master mix is prepared by combining the Amplification Master Mix and Stage 2 Reagent and is dispensed into the Indexing Plate, which contains the dry primers. The prepared Indexing Plate is stable at room temperature for at least 4 hours. When the stage 1 plate has completed the thermal cycling program, an aliquot of each of the 384 completed reactions is transferred to the Indexing Plate. Note that completed stage 1 plate reactions are stable at room temperature for at least 20 hours after cycling. When Indexing Plate dispensing is complete, it is vortexed, spun down and transferred to a thermal cycler for approximately 50 minutes. When the cycling program is complete, the 384 reactions are pooled, bead-cleaned, quantified and sequenced.

Application note

Simplified targeted genotyping by sequencing using the Amp-Seq Reagent System

Amp-Seq workflow simplicity

The protocol and reagents described in figure 1 have many properties that make them suited for a simple, low cost, high-throughput simple assay. The combined reagents are robust and stable at room temperature for several hours, which allows for a flexible workflow. Stability at room temperature permits high-throughput automation of the assay where plates are stacked before being transferred into the workflow.

Certain reagents, described in figure 1, can be combined into a master mix that remains stable at room temperature, but is also stable when frozen in bulk or dispensed into 384-well plates for short term storage at -20 °C allowing for convenient handling. It is not recommended to store frozen pre-plated reagents for more than a few weeks due to the risk of sublimation of the small reaction volumes. Please note that Biosearch Technologies do not supply pre-mixed frozen reagents; this is an option for the customer to consider if appropriate for their workflow, noting the caveats detailed above.

The Amp-Seq library preparation protocol consists of two amplification stages, with a simple transfer of material between Stage 1 and Stage 2, and may be completed in less than 2 hours. Post-processing in preparation for sequencing, including library pooling and clean-up, size estimation, quantitation and dilution may take an additional 1.5 to 2 hours.

After completion of the thermal cycling program (which is standardised across stages 1 and 2 of the workflow for to minimise impact of user errors), the process can be paused for up to 20 hours at room temperature, ensuring maximum flexibility for laboratory staff and automation. Total assay time is also reduced because of overlapping workflows for stage 1 and stage 2 plates. While the stage 1 plate is in the thermal cycler, the stage 2 plate (Indexing Plate) reagents can be assembled and plated, and are stable for at least 4 hours at room temperature before being placed in the thermal cycler. After stage 2 cycling is completed, the plates can be left in the cycler at room temperature for at least 20 hours.

This method utilises dual indexing primers that are dried in ready-to-use 384-well plates, providing even more cost and time savings. These plates can be stored at room temperature, saving space and energy that would otherwise be required for cold storage. Reaction volumes of 5 µL, for both stage 1 and stage 2 plates, also saves of the cost of reagents. There is, however, flexibility in reaction volumes, especially if a Hydrocycler^{2}™ is used for high-throughput processing; reaction well evaporation is minimal and thus reduced volumes (and associated evaporation) do not interfere with data quality.

Indexed samples can be pooled prior to cleaning and sequencing which helps to further reduce turnaround times and can facilitate same-day sequencing.

Application note

Simplified targeted genotyping by sequencing using the Amp-Seq Reagent System

DNA samples

A variety of DNA samples have been used to develop and optimise the Amp-Seq Reagent System. Across the plant species tested, both purified (clean) DNA and crudely-extracted DNA from leaves and seeds have been tested (table 1). However, the quality of data can depend on the protocol for extracting the crude lysates. Due to the variability of crude extract sample quality and yield from different tissues and species, and the inability to accurately quantitate DNA in extracts, the use of crude extracts with Amp-Seq may require significant optimisation of the extraction process and sample input amounts to achieve consistent results. A manual HotSHOT[1]-like protocol for crude extraction of DNA from both leaf punches and seeds was used, whilst purified sample DNA was obtained using automated protocols for paramagnetic bead-based chemistry (the oKtopure™ from Biosearch Technologies and the KingFisher Flex from ThermoFisher) or spin column-based kits (Qiagen, Zymo Research). The ability to use undiluted crude input material further simplifies the workflow by reducing upfront labour and logistics costs and reduces total turnaround times.

Species	Tissue type	DNA purification type	SNP panel sizes
Maize	Seed	Clean	600; 950; 1,150; 5,000
Maize	Seed	Crude	600; 1,150; 5,000
Maize	Leaf	Crude	1,150; 5,000
Soy	Seed	Clean	950; 1,000; 1,150; 4,600
Soy	Seed	Crude	1,000; 4,600
Soy	Leaf	Crude	1,000; 4,600
Wheat	Seed	Clean	950; 2,750
Strawberry	Leaf	Clean	400
Strawberry	Leaf	Crude	400
Sorghum	Leaf	Clean	1,100
Sorghum	Leaf	Crude	1,100
Canola	Seed	Clean	1,100
Canola	Seed	Crude	1,100
Canola	Leaf	Crude	1,100

Table 1. Sources of genomic DNA used for development and optimisation of the Amp-Seq protocol. Clean DNA indicates that a purification column (Qiagen or Zymo Research) or automated system (KingFisher Flex or oKtopure) was used. Crude DNA indicates that a rapid protocol such as HotSHOT was used. Crude DNA extract was not diluted for the assay unless very small volumes were required. Plant DNA was obtained from leaf punches or seeds. The number of specific primer pairs for each sample type is indicated in the SNP panel sizes column and ranges from 400 to 5,000 SNP locations. All data met specifications of >90% correct SNP call and >80% SNP depth uniformity of data.

Application note

Simplified targeted genotyping by sequencing using the Amp-Seq Reagent System

Quality control

Assay-to-assay consistency is established with rigorous quality control. The assay reagents are optimised so that a high percentage of calls from the SNP panel design (over 90%) and that a high SNP depth uniformity* for sample-to-sample data is achieved (over 80%). All reagents are subjected to functional assays to establish high standards. Dry indexing primer plates are assayed for both function and for cross contamination. Reagents were optimised during development to be robust and suitable for generating high quality, uniform data across a range of sample input types. Newly designed primer pair pools are quality checked for optimal DNA input and optimal primer concentration to complete the robustness of the Amp-Seq assay.

* SNP depth uniformity for Amp-Seq is calculated as the percentage of SNPs with a depth above 20% of the mean depth for that sample.

Assay stability

Table 2 illustrates the experimental setup for testing the stability of Amp-Seq components and workflow intermediates at room temperature. Using soy DNA, stage 1 plate reactions were prepared and incubated at room temperature for 4 hours. The plates were then placed in a programmed thermal cycler and run to completion. Completed stage 1 plate reactions were incubated at room temperature for 20 hours before the protocol was continued. Subsequently, the Indexing Plate was prepared by combining Amplification Master Mix and Stage 2 Reagent, indexing primers and 2 µL of each completed stage 1 plate reaction. Prepared Indexing Plate reactions were incubated at room temperature for 4 hours. After this incubation, Indexing Plate reactions were then placed in a thermal cycler run to completion. A set of Indexing Plate reactions were pooled and bead-cleaned immediately (Set A) whilst a second set (Set B) were incubated at room temperature for 20 hours before being processed in the same way as the first set. A control set of reactions were run that had neither 20-hour room temperature incubation (Set C).

	Experimental step	Steps used for each sample set		
		Set A	Set B	Control Set C
Stage 1	Stage 1 plate set-up	✓	✓	✓
	Room temperature incubation (4 hr)	✓	✓	n/a
	PCR	✓	✓	✓
	Room temperature incubation (20 hr)	✓	✓	n/a
Stage 2	Index plate set-up	✓	✓	✓
	Room temperature incubation (4 hr)	✓	✓	n/a
	PCR	✓	✓	✓
	Room temperature incubation (20 hr)	n/a	✓	n/a
Post-PCR	Pool and clean up	✓	✓	✓
	Sequencing	✓	✓	✓

Table 2. Experimental setup for Amp-Seq stability at room temperature. This assay was run with a soy (Williams 82) panel (the soy 1,000 panel). The stage 1 plate was prepared using a 10 µL assay volumes and, following PCR, 2 µL of product was transferred from completed stage 1 plate reactions to the Indexing Plate reactions. All incubations were performed for the times shown at room temperature. Completed Indexing Plate reactions were either processed immediately (Set A), or further incubated at room temperature for 20 hours prior to processing (Set B). Set C constituted the control reactions that had neither 20-hour room temperature incubation.

Application note

Simplified targeted genotyping by sequencing using the Amp-Seq Reagent System

Software pipelines

Software pipelines were developed to design specific primer pools and for data analysis. Primer pool software was rigorously tested to ensure that minimal primer interactions would occur, and that background signal is reduced to a level that does not significantly impact on-target mapping of amplicon reads. This level of stringency ensures that sequencing runs are efficient and very little sequence is wasted. A data analysis pipeline was built to detect variants ensuring high variant calling accuracy (>98%) and high genotype call concordance of diverse sample data (>95%) when compared to other genotyping technologies.

Protocol automation

A Beckman Biomek i7 liquid handler workstation was used to run purified soy genomic DNA (Williams 82) in the Amp-Seq assay. All reactions were set up in a 384-well plate using the liquid handler and then placed in a 384-well thermal cycler. Dried Indexing Plates were used in the protocol automation. After completion of the cycler program, the plate was again processed with the liquid handler and the final reaction was placed in a 384-well cycler. Upon completion of the assay, reactions in all 384 wells were pooled and the final library was bead-cleaned by hand and sequenced.

Results and discussion

Assay stability

High-throughput automation of a large set of sample DNA requires that, at some stages in the protocol, 384-well plates of reactions must sit at room temperature for several hours.

Figure 2 demonstrates that the reagents for the Amp-Seq assay are very stable at room temperature. Both the percentage of mapped reads (sequencing reads that mapped to the soy genome) and percentage of SNPs detected (percentage of the soy 1,000 panel that was called) are over 90%. The SNP depth uniformity for all conditions is over 80%, indicating an even sequence depth across the targets. The stability of the Amp-Seq reagents allows for flexibility in running this assay at high-throughput due to the convenient handling of reagents. This simplification of the workflow contributes to the reduced turnaround times offered by Amp-Seq.

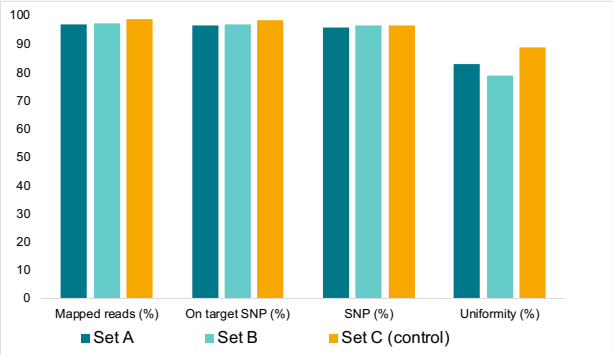


Figure 2. Amp-Seq assay stability at room temperature. Sequencing data quality characteristics for soy samples processed through the workflow at different rates. Set A; stage 1 plate reactions were incubated at room temperature for 20 hours prior to proceeding to stage 2. Completed Indexing Plate (stage 2) reactions were bead-cleaned and sequenced immediately upon completion. Set B: stage 1 plate reactions were incubated at room temperature for 20 hours prior to proceeding to stage 2. Completed Indexing Plate (stage 2) reactions were further incubated at room temperature for 20 hours prior to being bead-cleaned and sequencing. Set C; control, all reactions were processed immediately with no room temperature hold for during either stage 1 or stage 2.

Application note

Simplified targeted genotyping by sequencing using the Amp-Seq Reagent System

Genome complexity and SNP panel size
The robustness of the Amp-Seq assay was assessed with genomic DNA from several crops, including canola, maize, sorghum, soy, strawberry and wheat, using SNP panels of different sizes (table 1). Figure 3 illustrates data from maize panels of 1,152, 1,920 and 5,000 SNPs that clearly demonstrates the flexibility and robustness of the assay as all sizes of panels produced values at or above the specifications of 90% for percent SNP called and 80% for SNP depth uniformity. Data for each of these panels represents performance of an initial design, with no optimisation.

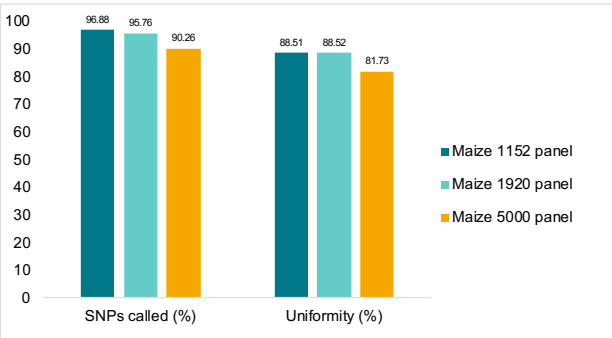
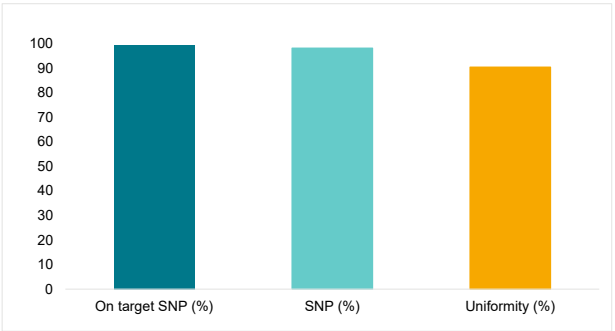


Figure 3. Amp-Seq assay robustness across a range of panel sizes. Maize B73 DNA was assayed with panels of 1,152, 1,920 and 5,000 primer pairs. Sequence data was downsampled to an average of 130.3, 131.1 and 115.7 reads respectively. All panels passed our specification of 90% for percentage of SNPs called and 80% for SNP depth uniformity of amplicon reads.

Figure 4 provides an example of data from an assay with crude extracts from maize leaf tissue. Although the quality of data will depend on the exact method of extraction, good results were obtained using a HotSHOT-like extraction involving NaOH and heat. This compatibility with undiluted crude input material allows for increased throughput and turnaround times by reducing upfront labour and logistics costs.

A



B

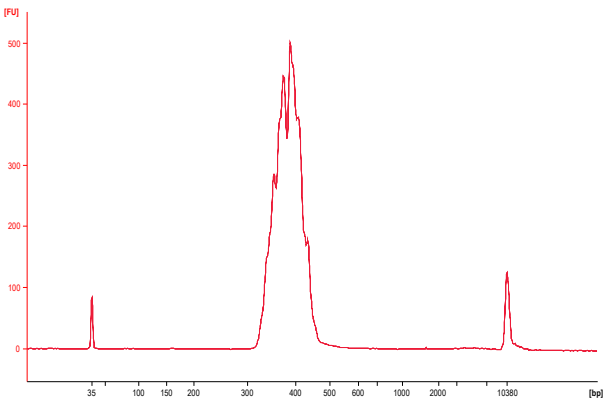


Figure 4A and B. Amp-Seq results for crudely extracted DNA. Crude maize leaf punch extract was directly used in the Amp-Seq protocol with a Maize 1,152 SNP panel. Figure A shows sequence analysis specifications for: i) percentage of reads that mapped to the correct sequence region (On target SNP %); ii) Percent of the 1,152 SNPs that were detected (SNP %) and; iii) SNP depth uniformity (%) of mapped reads for the 1,152 SNPs at an average depth of 134 reads. Figure B shows a Bioanalyzer scan of the pooled and cleaned Maize 1,152 library prior to sequencing. This scan illustrates that the majority of DNA signal is within an area (300 to 500 bp) that includes all the specific amplicons, and also shows that there was very little off target product.

Application note

Simplified targeted genotyping by sequencing using the Amp-Seq Reagent System

Automation of the Amp-Seq workflow

Figure 5 displays the suitability of the Amp-Seq Reagent System for automation. Consistently high performance was shown across 384 replicates of soy DNA, with a SNP call rate of over 98% and a SNP depth uniformity of amplicon coverage of over 90% on average. The flexibility to automate the workflow without impacting data quality further contributes to the reductions in turnaround times offered by Amp-Seq.

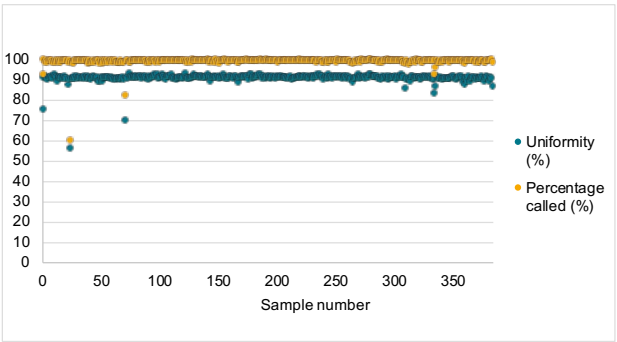


Figure 5. Automation of the Amp-Seq assay. Performance of the soy 1,000 panel illustrating consistently high performance of the Amp-Seq assay when the protocol was automated using a Biomek liquid handler and dried Indexing Plates were used.

Summary

Here we have presented our method for simplifying and streamlining targeted genotyping workflows. The method presented for Amp-Seq comprises a simple and user-friendly 2-stage protocol in 384-well format, using pre-mixed and pre-plated reagents. User error is minimised by using identical PCR cycling conditions for both stages and involving a very limited number of reaction components.

The data illustrated covers the key qualities of the Amp-Seq assay protocol.

These qualities include:

- Performance - The data presented illustrate the high-quality data that can be obtained using Amp-Seq, with very low off target sequence values, a high percentage of SNPs called, and SNP depth uniformity of read coverage for all amplicons;
- Robustness of reagents to handle a wide variety of DNA sample types;
- Stability of reagents to permit room temperature storage during high-throughput automation;
- Cost-effectiveness of the assay through:
    - Inexpensive reagents that include a single enzyme;
    - An assay that can be run with very small volumes;
    - Simplicity of the protocol with minimal hands-on time;
    - The use of automation for high-throughput sample processing;
    - Drying of indexing primers into plates for room temperature storage;
    - Master mixes that can be pre-plated for short term storage at -20 °C
- A robust assay that can tolerate crudely extracted DNA**.

** Please note that use of crude extract with Amp-Seq may require significant optimisation due to variability in quality and yield.

The Amp-Seq library preparation protocol may be completed in less than 2 hours, with an additional 1.5 to 2 hours required for post-processing in preparation for sequencing.

The simplicity of the workflow, combined with the robustness of data, make the Amp-Seq Reagent System an excellent choice for increasing genotyping by sequencing throughput in your laboratory to generate more data for your crop species.

Application note

Simplified targeted genotyping by sequencing using the Amp-Seq Reagent System

References

1. Truett, G.E., Heeger, P., Mynatt, R.L., Truett, A.A., Walker, J.A. and Warman, M.L., 2000. Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *Biotechniques*, 29(1), pp.52-54.

@LGCBiosearch | biosearchtech.com

BIOSEARCH™ TECHNOLOGIES

GENOMIC ANALYSIS BY LGC