HIGH THROUGHPUT GENOTYPING-BY-SEQUENCING IN LIVESTOCK - ION AMPLISEQ™

Kristian Ridley[1], Jeremy Walker[2], Matthew Hickenbotham[3]
[1]Thermo Fisher Scientific, 33 Marsiling Industrial Estate Rd, Singapore
[2]GeneSeek, 48th St, Lincoln, NE, USA
[3]Thermo Fisher Scientific, 81 Wyman St, Waltham, MA, USA

SUMMARY

Genotyping-By-Sequencing (GBS) is rapidly gaining popularity for high throughput applications in livestock genetics and agriculture biotechnology. Ion Ampliseq™ is a highly multiplexed, PCR-based resequencing technology that enables the targeting of hundreds to thousands of markers across hundreds of samples in a single sequencing run using the Ion Torrent™ Sequencing. We recently demonstrated the power of Ion Ampliseq™ GBS for high throughput cattle genotyping by designing a panel targeting over 4800 markers in a single pool. Furthermore, we have developed a simple, low cost, high throughput and rapid protocol ideal for commercial testing environments. The data shows excellent reproducibility, accuracy and >95% concordance with existing microarray data even at high (384) sample multiplex. GBS may also reveal novel variants within the targeted region that can enrich existing mapping data. Ion Ampliseq™ offers the flexibility to design customized GBS panels for any species with a reference genome, and with little to no optimization required.

INTRODUCTION

Ion AmpliSeq™ Designer is a simple assay design tool which can be used to create primer pools targeting any region of interest within a reference sequence. Amplicons ranging between 125 and 375 bp can be designed to target individual genetic loci such as single nucleotide polymorphisms (SNPs), or tiled across a total targeted region of more than 30,000 bp. For GBS applications, the reference genomes for mouse, cow, pig, sheep, dog, Chinese hamster, corn, rice, soybean, and tomato are preloaded into the Ion AmpliSeq™ Designer. Additionally, Ion AmpliSeq™ Designer allows the creation of Ion AmpliSeq™ panels for private reference genomes or known target regions in a secure cloud computing environment.

With the simplicity and speed of PCR, Ion AmpliSeq™ technology allows automated preparation of sequencing libraries. In combination with the Ion Proton™ Sequencer, rapid molecular marker screening by GBS can be performed for hundreds of samples and thousands of targets in a single run. This study describes a collaborative investigation into the use of Ion Ampliseq™ technology for developing a bovine SNP panel for a high throughput commercial testing application.

METHODS

Ion Ampliseq™ panel design

Using the BosTau6 assembly with Y chromosome sequences from BosTau7 plus 2 custom contigs, a total of 4,874 SNP loci were submitted for to the Ion Ampliseq™ Designer tool, of which 4,818 satisfied the design tool's *in silico* requirements (>99%).

Ion Torrent™ Sequencing

Automated library preparation, including normalization using the Ion Library Equalizer™ Kit was performed using the Tecan Freedom EVO® 150 platform using Ion Ampliseq™ 2.0 reagents for 384

bovine gDNA samples using this custom primer pool. 192 or 384 libraries were then pooled prior to automated template preparation and chip loading using the Ion PI™ IC 200 Kit with the Ion Chef™ platform. Finally, samples were sequenced in duplicate using Ion PI™ Chip Kit v2 BC on the Ion Proton™ Sequencer.

Data analysis

Sequencing reads were aligned to the reference sequence using Torrent Suite™ Software under standard parameters and variants identified using the Torrent Variant Caller plug-in under default germline variant call parameters.

RESULTS

Call rates of 88% and 80% were observed at 192 and 384 sample multiplex respectively. The average genotype concordance to previous microarray data was >95% ($r>0.95$) for 4,469 positions in common between the two assays (excluding no-calls by both assays), with robust genotyping, regardless of sample multiplexing, of 192 or 384 libraries per sequencing run (Figure 1.). A set of 96 'beef diversity' samples, representing a wide range of cattle breeds was included to demonstrate the ability of GBS to identify polymorphisms within the amplicon target sequences surrounding the targeted SNPs. The numbers of novel SNPs identified by alignment of the sequencing reads to the bosTau6 genome assembly for all 384 samples is presented in Figure 2.

DISCUSSION

The study describes a novel genotyping-by-sequencing approach using Ion Ampliseq™ technology. This is the first account of multiplexed PCR targeted resequencing for a panel of SNPs of similar number to those currently used in low density bovine arrays (3000-7000 markers) for applications such as genomic selection. The high level of concordance and sample multiplexing demonstrated here indicate that Ion Ampliseq™ may represent an ideal solution for developing both smaller and larger bovine SNP panels for a range of livestock applications including parentage, inherited disease and key production trait testing. Ion Ampliseq™ panels can easily be modified by addition or subtraction of primers 'on-the-fly' without the need for extensive re-optimization, and this represents a major advantage over existing array-based or mass spec-based genotyping platforms. Furthermore, genotyping-by-sequencing generates additional SNP data from the amplicon fragments surrounding the targeted SNP, providing researchers with additional variant data that may aid mapping studies.

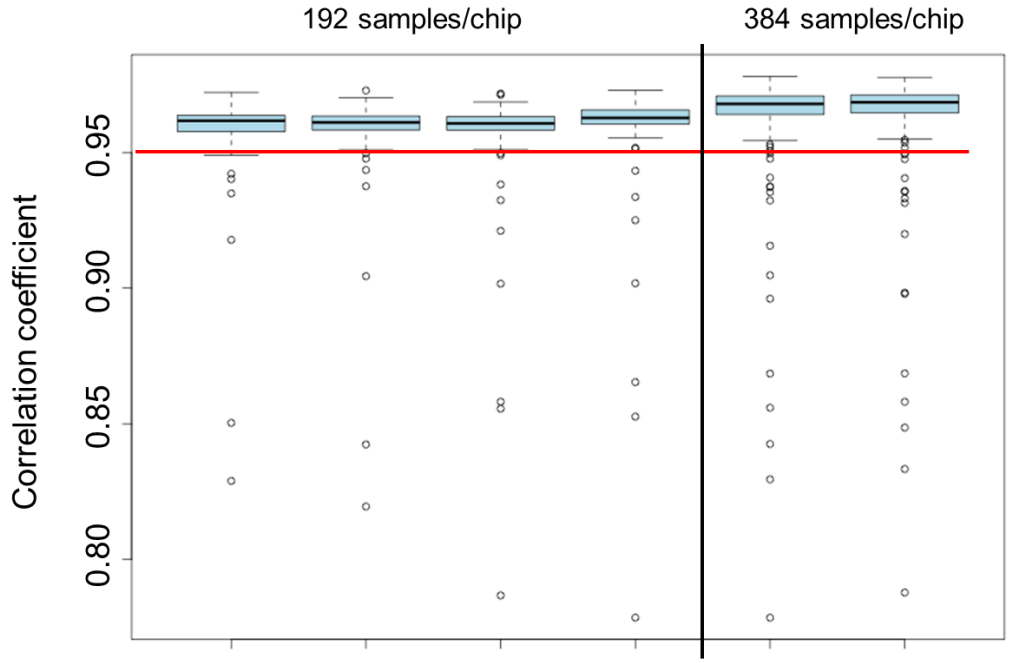

Figure 1. Genotype concordance data between an Ion AmpliSeq™ GBS panel and bead-based microarray analysis.

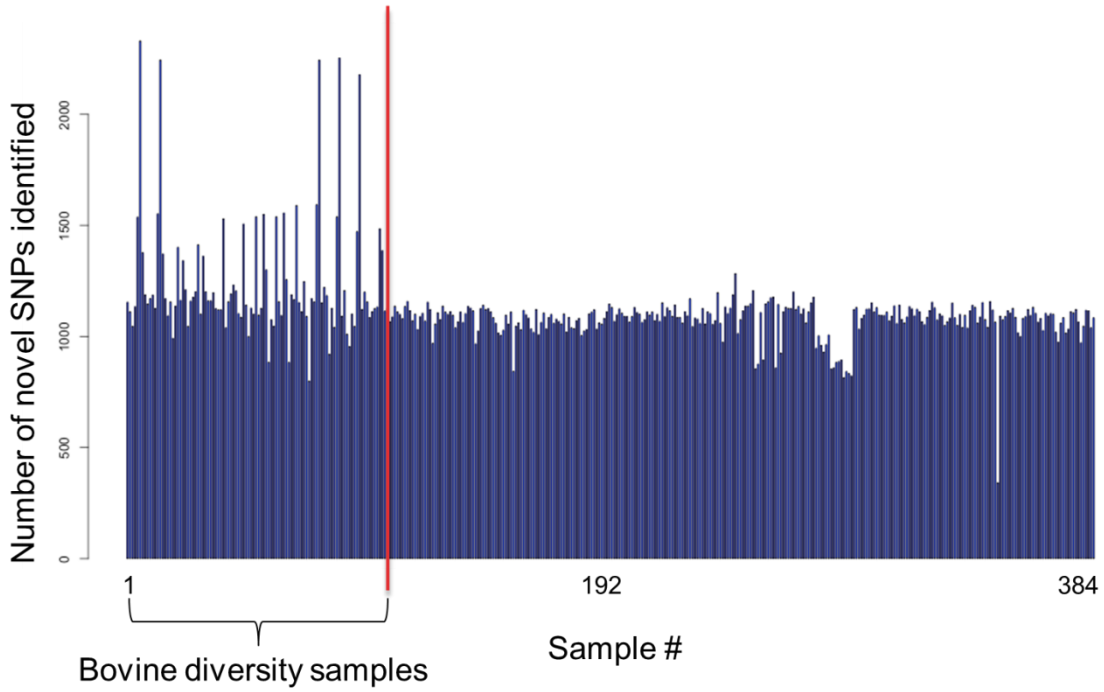

Figure 2. Number of novel SNPs identified by Ion Ampliseq™ sequencing for all 384 samples tested, including a set of 96 bovine diversity samples.

ACKNOWLEDGEMENTS

We would like to thank GeneSeek (Neogen Corporation) for providing reference sequences, target sequences and samples, and for their contribution to data analysis. The author would also like to acknowledge the staff involved in this study from Thermo Fisher Scientific (US) for design of the Ion Ampliseq™ panel and for conducting the Ion Torrent™ sequencing and analysis.