applied biosystems

Development of AgriSeq™ targeted GBS panels for breeding and parentage applications in dogs

Angela Burrell[1], Haktan Suren[1], Michelle Swimley[1], Chris Willis[1], Prasad Siddavatam[1], Krishna Gujjula[1], and Rick Conrad[1]
[1]Thermo Fisher Scientific, 2130 Woodward Street, Austin, TX, USA, 78744.

ABSTRACT

Utilization of genetic information for breeding and parentage verification purposes is a common tool of veterinary management. The need for highly consistent detection of informative genetic markers is critical for parentage verification and genetic trait detection. Targeted GBS methods, like the AgriSeq workflow, have the advantage over non-targeted GBS approaches (e.g. RADSeq) that are highly susceptible to allele drop-outs and missing data. Using the AgriSeq workflow, we can target hundreds of markers simultaneously in a highly reproducible manner across diverse sample sets.

We developed two targeted sequencing panels, one for canine parentage/ID verification and one for canine genetic defect/trait identification. The AgriSeq Canine SNP Parentage & ID Panel targets 381 markers utilized for parentage verification (Figure 1). The AgriSeq Canine Traits & Disorders Panel targets 154 markers involved in inherited disease and commonly tested genetic traits (Figure 1).

Utilizing the AgriSeq™ HTS Library Kit, a high-throughput targeted amplification and re-sequencing workflow, each panel's performance was tested on ≥180 diverse DNA samples. Libraries were sequenced on the Ion S5™ using an Ion 540™ chip with genotyping calling generated using the Torrent Variant Caller (TVC) plugin. The mean genotype call rate of markers across the samples was >95% for both panels. Concordance across replicate library preparations and independent sequencing runs was >99% for both panels. Each panel's results were compared with results from a DNA array, qPCR, and/or CE sequencing for orthogonal confirmation of genotype accuracy and the genotype calls were >99% concordant with the AgriSeq workflows.

The data demonstrates the utility of the AgriSeq targeted GBS approach for canine SNP genotyping applications.

Figure 1. Canine Panels Summary

Panel	SNPs	MNPs	Insertions	Deletions
Parentage & ID	379	0	0	2
Traits & Disorders	97	6	13	38

Figure 1. The Canine Parentage and ID panel targets 381 parentage markers, include 379 SNPs and 2 deletions. The Canine Traits and Disorders panel targets 154 markers including SNPs, MNPs, insertions, and deletions.

MATERIALS AND METHODS

Both Canine AgriSeq panels were validated by three methods: Orthogonal Testing, Robustness Testing, and Field Testing in order to verify performance of the panels met all requirements.

Orthogonal testing was completed by running up to six samples on the Axiom Canine HD Array and/or by CE sequencing. Samples were also run using the AgriSeq workflow and genotype concordance was calculated between orthogonal technologies.

Figure 2. AgriSeq Library Prep workflow

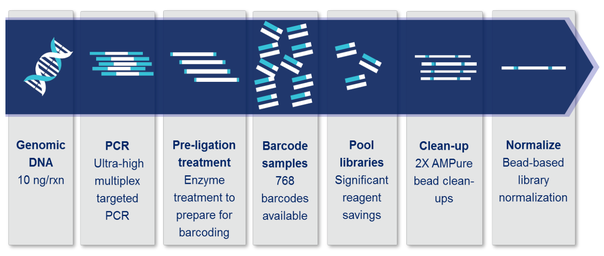


Figure 2. Using the AgriSeq HTS Library Kit, 10ng/rxn of canine DNA was amplified using each of the Canine AgriSeq GBS panels separately. Each sample was then treated with a Pre-ligation Enzyme to remove residual primer dimers allowing for more efficient sequencing. Samples were ligated with unique barcoded adapters allowing them to be pooled for subsequent clean-up and sequencing while retaining traceability to the original sample during analysis for significant cost savings. Libraries were cleaned-up by a two-round AMPure purification. A final bead-based normalization step helps ensure each library is at a consistent final concentration suitable for direct input into template prep on the Ion Chef™ instrument. All libraries were pooled 1:1 for templating in a single reaction.

AgriSeq libraries were sequenced on the Ion S5™ sequencing system using an Ion 540™ or Ion 550™ chip. Data was analyzed using the Torrent Variant Caller (TVC) plugin as part of the Torrent Suite™ software package to determine the genotype call for each marker and sample (Figure 3).

Figure 3. Complete AgriSeq Sequencing Workflow

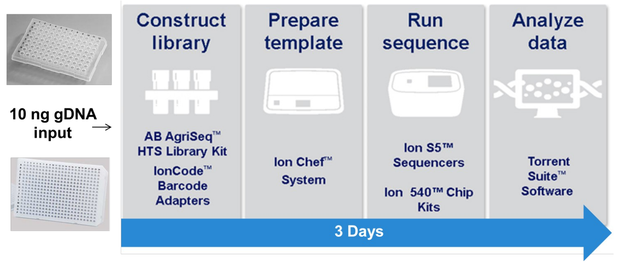


Figure 3. Following library prep, libraries were pooled into a single tube and run overnight on the Ion Chef instrument for template prep. The following day, libraries were sequenced on the Ion S5 XL instrument and data was analyzed using the Torrent Suite Software v5.10. Genotypes for all markers were obtained from the Torrent Variant Caller plugin.

To test robustness of our sequencing results we processed 12 high-quality commercial DNA samples in replicates of n=64 for a total of 768 barcoded samples with both canine AgriSeq panels using our standard workflow (Figures 2 and 3). Each library was sequenced twice on an Ion 540 and Ion 550 chip (Figure 3). Replicate genotype concordance, the percent of genotype calls across all replicate samples that are identical, was also determined.

The performance of the AgriSeq Canine Parentage and AgriSeq Canine Trait/Disorders panels were also tested using ≥180 diverse oral swab canine DNA samples. The libraries were prepared using the AgriSeq HTS Library Kit (Figure 2) and sequenced on the Ion S5 XL instrument using a 540 chip (Figure 3). Call rate and read uniformity were calculated for all libraries.

RESULTS

Figure 4. Canine Parentage Orthogonal Testing Results Summary

DNA	# Concordant Markers	# Discordant Markers	% Concordance
Testis DNA Lot 041718	381	0	100%
Uterus DNA Lot 041718	365	0	100%
Female DNA Lot 041718	364	2	99.4%
Male DNA Lot 041718	360	0	100%
Male DNA Lot 2	361	0	100%
Female DNA Lot 2	362	1	99.5%

Figure 4. Orthogonal concordance was determined for the Canine Parentage panel by testing 359 of the 381 Canine Parentage panel markers with the Axiom Canine HD Array. The remaining 39 makers not present on the array were tested by CE sequencing. Up to 6 DNA samples were used for testing each method and results were compared to GBS sequencing results using the AgriSeq workflow. Of the 2193 genotypes obtained, 2191 were concordant with the AgriSeq workflow resulting in >99.9% concordance.

Figure 5. Canine Trait/Disorders Orthogonal Testing Results Summary

Orthogonal Method	# Concordant Markers to GBS	# Discordant Markers to GBS	# No Calls	Concordance
CE Sequencing Only	107	0	5	100%
Axiom Array Only	39	0	0	
Both CE and Array	8	0	0	

Figure 5. Orthogonal concordance was determined for the Canine Trait/Disorders panel by testing 39 of the 154 Canine Trait/Defect panel markers with the Axiom Canine HD Array. The remaining 107 makers not present on the array were tested by CE sequencing. 5 markers were tested by both orthogonal technologies. Up to 6 DNA samples were used for testing each method and results were compared to GBS sequencing results using the AgriSeq workflow. There were 5 markers that were unable to be genotyped by CE testing due to poor sequencing quality. Of the 149 markers that were able to be genotyped by an orthogonal method, concordance to the AgriSeq workflow was 100%.

Figure 6. Robustness Testing Mean Sample Call Rate Results

Panel	Sequencing Kit	Mean Call Rate	stdev
Canine Parentage/ID	540 Kit	99.9%	0.4%
	550 Kit	99.7%	0.4%
Canine Trait/Defect	540 Kit	99.6%	0.4%
	550 Kit	99.4%	0.5%

Figures 6. Twelve canine DNA samples were tested in replicates (n=64) for a total of 768 barcoded libraries with the Canine Parentage and Canine Trait/Disorders panels. Libraries were sequenced twice on the Ion 540 and Ion 550 chips to look at genotype call robustness and consistency. Mean call rate for both panels and all sequencing kits was >99% with minimum variation between samples demonstrating the robustness of the genotype calls.

Figure 7. Canine Traits/Disorders Robustness Testing Replicate Genotype Concordance

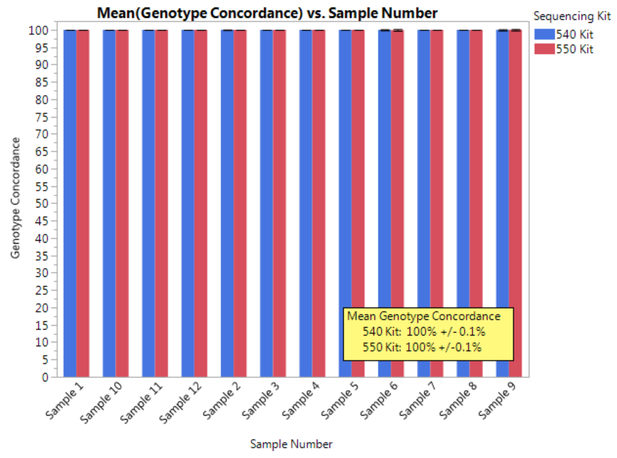
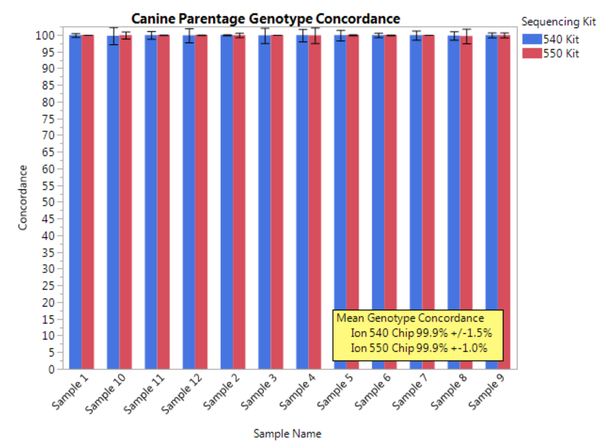


Figure 8. Canine Parentage/ID Robustness Testing Replicate Genotype Concordance



Figures 7 and 8. Replicate genotype concordance is calculated as the percent of markers that give identical genotypes for replicate samples. The top graph (Figure 7) shows the Canine Traits/Disorders panel had a mean of 100% genotype concordance and the Canine Parentage (Figure 8) had a 99.9% mean genotype concordance between replicate samples even when n=64 replicates are tested demonstrating the high robustness and consistent results obtained with each panel.

Figure 9. Canine Trait/Disorders Field Sample Call Rate

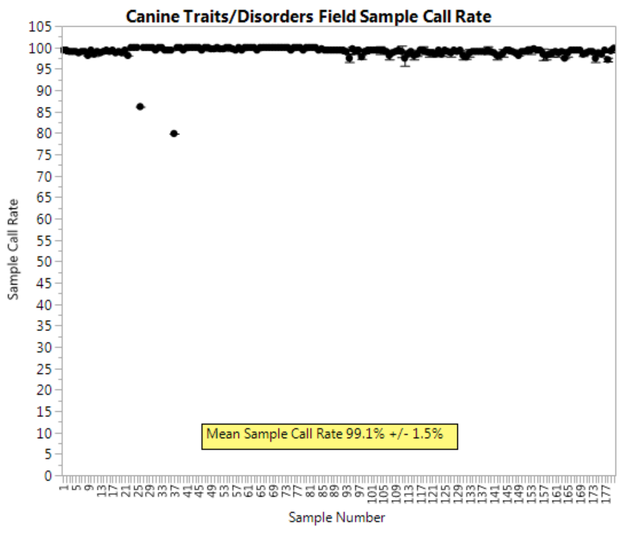
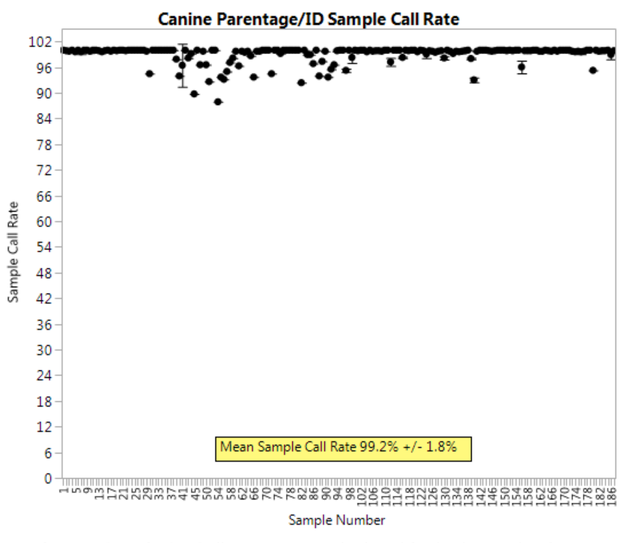
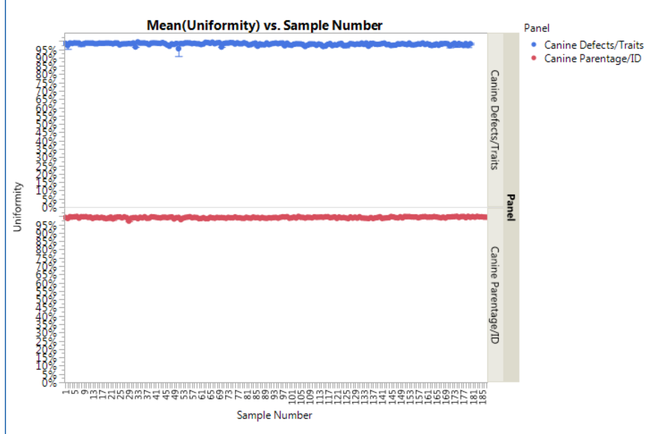


Figure 10. Canine Parentage/ID Field Sample Call Rate



Figures 9 and 10. Call rates were calculated for both panels after testing ≥180 diverse canine oral swab DNA field samples with the AgriSeq workflow. The mean call rate for the Canine Trait/Disorders panel was 99.1% and the mean call rate for the Canine Parentage panel was 99.2% demonstrating the high performance obtained from customer samples.

Figures 12. Read Uniformity



Panel	Mean Uniformity	stdev
Canine Parentage/ID	99.3%	0.4%
Canine Trait/Defect	98.2%	0.7%

Figure 12. Read uniformity is the percentage of target bases covered by at least 0.2X of the average base read depth. It is a measure of how evenly you are covering target amplicons with reads. Low uniformity (<90%) can lead to marker drop-off and poor call rates. The mean read uniformity for both panels was excellent, even when testing a set of very diverse field samples. Both the Canine Traits and Disorders panel as well as the Canine Parentage & ID panel had a mean uniformity of >98%

CONCLUSIONS

The AgriSeq Canine Traits and Disorders and AgriSeq Canine Parentage and ID panels along with the AgriSeq workflow provide a streamlined, cost-effective method for canine parentage verification and genotyping. Up to 4X 384-well plates can be processed in a single day and full sequencing results can be obtained in as little as three days. The flexibility of AgriSeq allows hundreds of samples to be pooled together into a single sequencing run targeting hundreds to thousands of markers.

Our method yields calls for the vast majority of markers (mean 99.1% for the Canine Traits/Defects panel and 99.2% for the Canine Parentage panel). Replicate genotype concordance is >99.9% and calls were highly concordant with orthogonal data (>99%). While we demonstrated the utility of AgriSeq sequencing technology for assessing parentage and genetic trait testing in dogs, our approach can be applied to other agricultural genotyping problems as well.

In conclusion, the AgriSeq library prep kit and canine GBS panels combine into a robust and efficient workflow for animal genotyping and parentage applications.

REFERENCES

1. Cariou M, Duret L, Charlat S. How and how much does RAD-seq bias genetic diversity estimates? BMC Evolutionary Biology. 2016;16:240. doi:10.1186/s12862-016-0791-0.

2. Kidd, Kenneth K. et al. Current sequencing technology makes microhaplotypes a powerful new type of genetic marker for forensics . Forensic Science International: Genetics , Volume 12 , 215 – 224.

ACKNOWLEDGEMENTS

We would like to acknowledge Maarten de Groot at VHL Genetics for supplying the canine field sample DNA used in this study.

TRADEMARKS/LICENSING

ThermoFisher SCIENTIFIC