BovineSNP50 Genotyping BeadChip

Featuring 54,609 evenly spaced SNP probes that span the bovine genome.

Highlights

- **Unrivaled Call Rates and Accuracy**
  >99% average call rates and 99.9% reproducibility
- **Comprehensive and Uniform Coverage**
  Evenly distributed polymorphic SNPs with a mean gap of 49.4 kb
- **Simple Workflow**
  PCR- and ligation-free protocol
- **High-Throughput Format**
  Up to 24 samples can be interrogated in parallel

Introduction

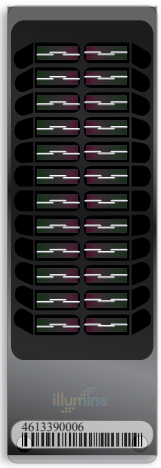
Illumina, in collaboration with the USDA ARS, University of Missouri, and the University of Alberta, has developed the second generation of a proven high-density, genome-wide genotyping array for the interrogation of genetic variation in cattle. The BovineSNP50 v2 BeadChip (Figure 1) features 54,609 informative SNP probes that uniformly span the entire bovine genome, empowering applications such as genome-wide–enabled selection, identification of quantitative trait loci, evaluation of genetic merit of individuals, and comparative genetic studies.

More than half of the probes on this BeadChip are designed to target novel SNPs discovered using Illumina's Genome Analyzer. Additional high-value content is derived from publicly available sources such as the bovine reference genome, Btau[1], and the Bovine HapMap Consortium data set[2]. The BovineSNP50 BeadChip covers common SNPs validated in economically important beef and dairy cattle breed types and presents an average minor allele frequency (MAF) of 0.25 across all loci. Importantly, this BeadChip offers uniform coverage with an average probe spacing of 49.4 kb to provide more than sufficient SNP density for robust genome-association studies in cattle.

The BovineSNP50 BeadChip is a multi-sample genotyping panel powered by Illumina's Infinium® HD Assay. This revolutionary assay provides the industry's highest call rates, allows for flexible content deployment, and enables the detection and measurement of copy number variation. In addition, the assay's PCR-free, single-tube sample preparation[3,4] significantly reduces labor and potential sample handling errors. Illumina's multi-sample format further reduces experimental variability and overall project cost by allowing researchers to interrogate up to 24 samples in parallel.

The combination of Illumina's proprietary assay technologies, unconstrained locus selection, and high-throughput format offers a comprehensive solution for whole-genome studies in beef and dairy cattle.

Figure 1: BovineSNP50 BeadChip



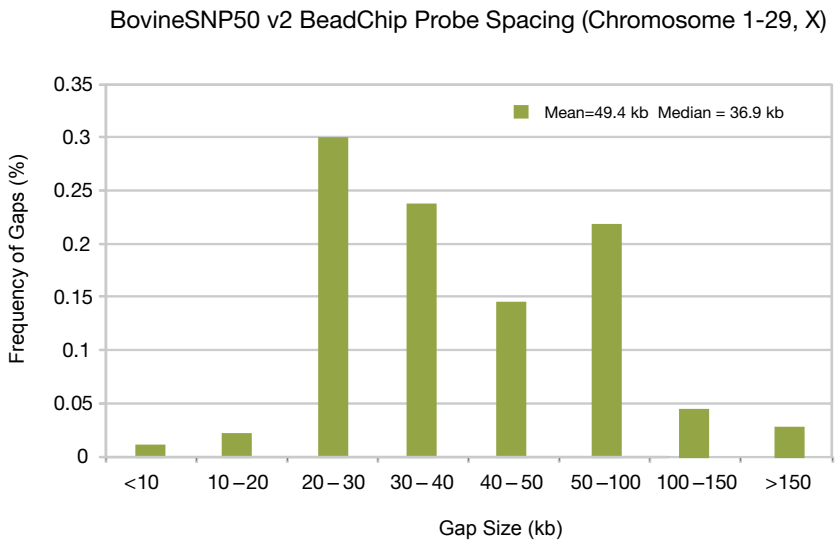
The BovineSNP50 BeadChip features more than 54,000 evenly spaced SNPs across the entire bovine genome

BovineSNP50 v2 BeadChip Content

Illumina scientists and collaborators set out to develop an informative and high-density SNP genotyping microarray that could be used to investigate genetic variation in any cattle breed. More than 12,000 probes were designed to target validated common SNPs (MAF ≥ 0.05) described by the Bovine HapMap Consortium. BeadChip developers also mined publicly available resources for common SNPs including the bovine reference genome[1] and whole-genome shotgun reads assembled by researchers at Baylor College of Medicine[5]. Additional content sources include parentage markers identified by researchers from the U.S. Meat Animal Research Center, Clay Center[6], and SNPs identified by researchers at the USDA ARS through the comparison of Holstein bacterial artificial chromosomes (BAC) sequence data[7] to the bovine genome assembly (Table 1).

More than 24,000 SNP probes featured on the BovineSNP50 v2 BeadChip target novel SNP loci that were discovered by sequencing three pooled populations of economically important beef and dairy cattle. Using Illumina's Genome Analyzer, researchers at the USDA ARS discovered more than 62,000 putative SNPs by deeply sequencing approximately 2% of the *Bos taurus taurus* genome[8]. More than 23,800 SNPs derived from this novel data source were chosen based upon their spacing, expected MAF, and Infinium HD Assay performance. In addition, BeadChip developers selected loci that target the largest unmapped contigs to ensure comprehensive coverage. All 54,609 SNP probes on the BovineSNP50 BeadChip have been validated in 19 common beef and dairy breeds (Table 3). This product targets evenly distributed SNPs that are polymorphic across the breeds tested and provides an average probe spacing of 49.4 kb and a

Figure 2: BovineSNP50 v2 BeadChip Probe Spacing



The BovineSNP50 v2 BeadChip provides uniform coverage across the entire genome. Fewer than 5% of all gaps are larger than 100 kb.

median spacing of 36.9 kb (Figure 2). Current research on linkage disequilibrium (LD) in multiple breeds of cattle suggests haplotype blocks of approximately 70 kb on average[9], indicating that the resolution offered by the BovineSNP50 v2 chip is well within that of LD in cattle. More than 54,060 SNP probes on the BovineSNP50 BeadChip map to UMD 3.0, the most current bovine reference genome assembly.

Table 1: BovineSNP50 BeadChip Content Sources

Source	BovineSNP50 v1 Probes	BovineSNP50 v2 Probes
Novel SNPs derived from sequencing common cattle breeds using Illumina's Genome Analyzer	23,840	24,181
Bovine HapMap Data Set	12,298	12,342
Btau assembly SNPs	9,361	9,404
Whole-Genome Shotgun Reads*	5,808	6,038
Holstein BAC Sequence Data	1,409	1,411
Parentage**	116	120
Other†	1,169	1,113
Total	54,001	54,609

* Reads derived from six cows/breeds (Norwegian Red, Hostein, Brahman, Angus, Jersey and Limousin) compared against Btau2.0

** Highly curated parentage markers that include 30 duplicates

† Includes common SNPs validated by the Institute for Food and Agricultural Sciences Alberta and INRA, the French International Institute of Agriculture

High-Quality Data

The 54,609 SNP assays on the BovineSNP50 v2 BeadChip were subjected to rigorous functional testing to ensure strong performance using Illumina's Infinium HD Assay. Whole-genome association studies are successful, in part, due to high call rates and accurately called genotypes. Since complex traits often have relatively small gene effects, potential associations may be missed if the assayed SNP in linkage disequilibrium with the SNP of interest has a low call rate or incorrect genotype call. Illumina ensures that every BovineSNP50 BeadChip offers > 99% call rate.

Table 2 shows the outstanding results from Illumina's internal validation of the BovineSNP50 v2 BeadChip content using samples provided by the Bovine HapMap Consortium. Product developers identified and retained 4,290 loci that appear to have an adjacent or underlying deletion among the breeds sampled. Although these loci yield lower call rates when compared to the other loci on the panel, they were retained because they may provide biologically relevant information for traits of interest and for future improvements in the genome build.

The performance and content validation results clearly show the reliability and outstanding data quality the BovineSNP50 v2 BeadChip delivers. With this BeadChip, researchers can predict the genetic merit for phenotypes of interest and investigate the genetic basis of variation among a multitude of cattle breed types.

Solutions for Genotyping

As with all of Illumina's standard products, an optional Laboratory Information Management System (LIMS) and robotic automation are available to accurately and efficiently track samples throughout analysis. Researchers can genotype their own samples using the Illumina GenomeStudio DNA analysis software or Illumina's FastTrack Genotyping Service.

Table 2: BovineSNP50 v2 Performance and Specifications

Parameter	Results	Product Specification
Average Call Rate*	99.7%	> 99%
Reproducibility **	100%	> 99.9%
Mendelian Inconsistencies	0.06%	< 0.1%

* Based on 473 individuals from 19 major cattle breeds, 31 trios, and 1 replicate

** Based on 4 Holstein trios

Table 3: BovineSNP50 BeadChip Content Validation

Breed	Samples	Polymorphic Loci*	Mean MAF	Median MAF**
Angus	24	41,073	0.21	0.21
Beefmaster	23	43,114	0.22	0.22
Bos indicus Gir	21	23,567	0.11	0.02
Bos indicus Nelore	19	25,492	0.11	0.03
Brahman	22	29,444	0.13	0.07
Brown Swiss	21	35,971	0.19	0.17
Charolais	19	43,723	0.22	0.21
Guernsey	21	36,748	0.19	0.17
Hereford	24	42,132	0.22	0.23
Holstein	49	42,849	0.22	0.22
Jersey	23	35,346	0.18	0.15
Limousin	39	42,617	0.22	0.21
N'Dama	23	28,869	0.14	0.07
Norwegian Red	17	42,055	0.22	0.21
Piedmontese	21	41,912	0.22	0.21
Red Angus	10	42,388	0.21	0.20
Romagnola	21	38,524	0.20	0.19
Santa Gertrudis	21	41,783	0.22	0.21
Sheko	16	35,084	0.17	0.13
Overall	**434**	**47,168**	**0.24**	**0.25**

* MAF > 0.05

** Across all 54,609 loci

Product Summary

Developed in collaboration with leading bovine researchers, the BovineSNP50 BeadChip features more than 54,000 evenly spaced SNP probes that span the bovine genome. This 24-sample BeadChip presents a high-throughput, cost-effective solution for whole-genome studies in beef and dairy cattle.

References

1. ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/fasta
2. www.bovinehapmap.org
3. Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS (2005) A genome-wide scalable SNP genotyping assay using microarray technology. Genet 37(5): 549–554.
4. Steemers FJ, Weihua Chang W, Lee G, Barker DL,Shen R, et al. (2006) Whole-genome genotyping with the single-base extension assay. Nat Methods 3(1): 31–33.
5. ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Btaurus/snp/Btau20070913/README
6. Heaton MP, Keen JE, Clawson ML, Harhay GP, Bauer N, et al. (2005) Use of bovine single nucleotide polymorphism markers to verify sample tracking in beef processing. J Am Vet Med Assoc 226(8): 1311–1314.
7. Matukumalli LK ,Taylor JF, and Van Tassell CP. Personal communication.
8. Van Tassell CP, Smith TPL, Matukumalli LK, Taylor JF, Schnabel, RD, et al. (2008) Simultaneous SNP discovery and allele frequency estimation by high-throughput sequencing of reduced representation libraries. Nat Meth (accepted).
9. Khatkar MS, Zenger KR, Hobbs M, Hawken RJ, Cavanagh JAL, et al. (2007) A Primary Assembly of a Bovine Haplotype Block Map Based on a 15,036-Single-Nucleotide Polymorphism Panel Genotyped in Holstein–Friesian Cattle. Genetics 176(2): 763–772.

Ordering Information

Catalog No.	Product	Description
WG-450-2001	BovineSNP50 v2 Whole-Genome Genotyping Kit (for 48 samples)	Each package contains 2 BeadChips, along with reagents for amplifying, fragmenting, hybridizing, labeling, and detecting 48 DNA Samples
WG-450-2002	BovineSNP50 v2 Whole-Genome Genotyping Kit (for 288 samples)	Each package contains 12 BeadChips, along with reagents for amplifying, fragmenting, hybridizing, labeling, and detecting 288 DNA Samples
WG-450-2003	BovineSNP50 v2 Whole-Genome Genotyping Kit (for 1152 samples)	Each package contains 48 BeadChips, along with reagents for amplifying, fragmenting, hybridizing, labeling, and detecting 1152 DNA Samples

Each BovineSNP50 v2 DNA Analysis BeadChip can process 24 samples and analyze approximately 54,000 loci.

Illumina, Inc. • 9885 Towne Centre Drive, San Diego, CA 92121 USA • 1.800.809.4566 toll-free • 1.858.202.4566 tel • techsupport@illumina.com • illumina.com

FOR RESEARCH USE ONLY

© 2011 Illumina, Inc. All rights reserved.
Illumina, illumina*Dx*, Solexa, Making Sense Out of Life, Oligator, Sentrix, GoldenGate, GoldenGate Indexing, DASL, BeadArray, Array of Arrays, Infinium, BeadXpress, VeraCode, IntelliHyb, iSelect, CSPro, GenomeStudio, Genetic Energy, HiSeq, HiScan, TruSeq, Eco, and MiSeq are registered trademarks or trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners.
Pub. No. 370-2007-029 Current as of 06 January 2011

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGACG
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA