RESEARCH ARTICLE

A Meta-Assembly of Selection Signatures in Cattle

Imtiaz A. S. Randhawa*, Mehar S. Khatkar, Peter C. Thomson, Herman W. Raadsma

Reprogen - Animal Bioscience Group, Faculty of Veterinary Science, The University of Sydney, 425 Werombi Road, Camden, 2570, NSW, Australia

* imtiaz.randhawa@sydney.edu.au

Abstract

Since domestication, significant genetic improvement has been achieved for many traits of commercial importance in cattle, including adaptation, appearance and production. In response to such intense selection pressures, the bovine genome has undergone changes at the underlying regions of functional genetic variants, which are termed "selection signatures". This article reviews 64 recent (2009–2015) investigations testing genomic diversity for departure from neutrality in worldwide cattle populations. In particular, we constructed a meta-assembly of 16,158 selection signatures for individual breeds and their archetype groups (European, African, Zebu and composite) from 56 genome-wide scans representing 70,743 animals of 90 pure and crossbred cattle breeds. Meta-selection-scores (MSS) were computed by combining published results at every given locus, within a sliding window span. MSS were adjusted for common samples across studies and were weighted for significance thresholds across and within studies. Published selection signatures show extensive coverage across the bovine genome, however, the meta-assembly provides a consensus profile of 263 genomic regions of which 141 were unique (113 were breed-specific) and 122 were shared across cattle archetypes. The most prominent peaks of MSS represent regions under selection across multiple populations and harboured genes of known major effects (coat color, polledness and muscle hypertrophy) and genes known to influence polygenic traits (stature, adaptation, feed efficiency, immunity, behaviour, reproduction, beef and dairy production). As the first meta-assembly of selection signatures, it offers novel insights about the hotspots of selective sweeps in the bovine genome, and this method could equally be applied to other species.

Introduction

Modern domestic species are a result of selective breeding for many traits of economic and adaptive importance since domestication [1–4]. The footprints of selective breeding on genomic architecture can now be characterized with the development of molecular genomic and advanced computational biology tools [5–7]. Although the contribution of natural (adaptation) and artificial (domestication and subsequent directional selection) remains challenging to

differentiate, the rapid expansion of genomic data generated from global sequencing and genotyping projects are providing greater insights of selection on genomes of domestic species [8].

History of genetic changes in the bovine genome to adapt to new circumstances dates to about 10,000 years ago. Domestication followed by spatial dispersion due to human migration, has resulted in the population of many breeds of cattle across the world, exposing animals to new environments and husbandry practices [2, 4]. Therefore, long-term selection pressures have operated on the genomic regions that control traits for adaptive fitness. More recently, selection for various unique morphological traits during the development of specialized breeds (for example; coat colors, presence of horns, etc.) have left their selective signatures in the genome.

Modern cattle breeds are the most dominant and well-developed domesticated ruminant. They are raised for production of dairy, meat, leather and as draft animals [9, 10]. Dairying has underpinned the human cultural revolution as an important food source [11] along with influencing the human genome for selective adaptation to tolerate lactose in response to domestication of ruminants [12–15]. During the past 50 years the worldwide cattle population has increased approximately 50%, whereas, the production (milk, meat and hides) of this population has increased up to 100% (S1 Fig). This represents a remarkable improvement in both genetic value and husbandry practices. As such, the bovine genome has been exposed to intense selective forces for commercially important traits in many breeds.

Through different civilizations, humans have controlled breeding in cattle around the world that resulted in emergence of many contemporary breeds [1, 2, 4, 16–19]. It is thought that the progenitors (aurochs: *Bos primigenius*) of modern cattle were initially domesticated in the Euphrates Valley during the Neolithic era and then, within the next 2,000 years, they spread in the Fertile Crescent and further dispersed in the Mediterranean Basin [18]. It is presumed that subspecies of aurochs developed into various cattle types, and that the spatial distributions of multiple independent domestication events included different regions of the Fertile Crescent and Indus Valley and, possibly in Africa [2, 20–22]. Consequently, there are three distinct domestic cattle lineages, i.e., European *Bos taurus*, African *Bos taurus* and *Bos indicus* (Zebu), which represent all contemporary cattle breeds.

Recent investigations concluded that the taurine and indicine cattle diverged long before domestication, and an early separation in the ancestry of European, East Asian and African cattle breeds has resulted in widespread geographical distributions of taurine breeds [23–26]. Moreover, ancient and recent admixture between African taurine and indicine cattle has also been recorded [24, 25, 27, 28]. Present cattle breeds are believed to be remnants of a much larger cohort of ancient populations that once existed in isolation across various parts of the world. Bottlenecks due to domestication—breed formation and recent selection—have resulted in reduced effective population size of various geographically isolated populations [23].

Access to bovine genome sequence assemblies (S2 Fig) and high-density genotyping panels (S3 Fig) have provided researchers with remarkable resources to study the effects of domestication and selection on the architecture of the bovine genome. Subsequent investigations have revolutionized our understanding of mammalian evolution, domestication and devised strategies for enhancing genetic improvement for dairy and beef production [23, 29]. Organized breeding systems that have large well characterized cattle populations also provide pivotal resources for the discovery of genes contributing to complex traits such as milk production, fertility, muscle formation, energy partitioning and disease resistance [10, 12, 30, 31]. Approaches that integrate quantitative trait loci (QTLs) [32] mapping studies, genome wide association studies (GWAS) and selection signatures have led the way to fine-map and localize functional mutations of many genes contributing to phenotypic diversity in cattle [3].

Accurate identification of chromosomal regions that contain selection signatures is important in our understanding of the underlying genetic variation contributing to phenotypic

diversity [3, 33]. Cattle have been a model animal for such studies with almost 1,000 breeds [34] reflecting genetic diversity underlying domestication, adaptation, appearance and production. This study reviews a set of 64 studies that have investigated selection signatures in worldwide populations of cattle breeds. We constructed a meta-assembly of the 56 genome-wide scans to highlight the hotspots of positive selection within cattle, and discuss the historical selection events and the role of underlying genes controlling the economically important traits. We also highlight the limitations of the available bovine genomic resources and implications of using different methodologies.

Methods

Meta-assembly of selection signatures in cattle

Many genome-wide scans have investigated unique patterns of genetic polymorphisms in the cattle genome to identify selection signatures (Table 1, S1 Table). These published studies used different SNP genotype datasets and various tests to detect such selection signatures (Table 2) across demographically diverse populations (S2 Table). These investigations provide insights about the historical selection in cattle but provide little information on consensus of selection signatures across the cattle genome. Recently, Gutiérrez-Gil et al. [35] reviewed and compared the genomic regions under selection across European cattle breeds, published in 21 studies. However, this study is the first attempt to comprehend the published results of 56 studies on selection signatures within each cattle type (African, European, Zebu and Composite) by developing a novel meta-assembly approach extending on the method used in Khatkar et al. [30], Khatkar et al. [31]. Construction of the meta-assembly was challenging given the heterogeneity of published investigations on selection signatures, hence several critical measures were adopted to minimize sources of potential bias in using available information.

Construction of meta-assembly

Data collection. A total of 64 publications were identified which reported selection signatures in cattle, out of these, 56 genome-wide scans of selection signatures in cattle (Table 1) were compiled for the computation of meta-scores. The remaining eight studies (S1 Table) investigated selective sweeps on targeted regions of a few chromosomes or used other genomic markers (e.g., microsatellites), and were not included in the analyses to avoid any bias from single locus results [36] and could have overestimated the strength of selection in some regions, therefore distorting the meta-scores.

Many different approaches were used in the genome-wide studies to test the departure from neutrality. Most were based on estimates of population allele differentiation and haplotype homozygosity (Table 2). For the purpose of this review, selection tests were considered independent of each other and no adjustments were made for the perceived relationships across various selection tests. This position was taken considering that all tests within a study were mostly independent as presumed by those authors. Additional parameters that may cause inaccuracies or bias in the meta-assembly, such as differing bovine genome-maps, shared relationships among sample population, and significance threshold of reported signatures were included in the adjustment of the meta-assembly scores.

Mapping results onto UMD3.1 bovine assembly. Published results have been presented on different bovine genome assemblies, including Btau 3.1, Btau 4.0 and UMD3.1 (Table 1), which have variable genome coverage (S2 Fig). In order to find alignments of detected selection signatures across studies, results from Btau 3.1 and Btau 4.0 were mapped to the latest UMD3.1 bovine genome assembly. Studies which used different SNP genotyping panels (Table 1, S3 Fig) were updated using the common or nearest marker position for the start, end

Table 1. Summary of genome-wide scans on Bovine selection signatures.

No.	Study	SNPs (assay)[1]	Genome assembly[2]	Breeds (samples)[3]	Selection Tests[4]	Traits reported[5]
1	MacEachern et al. [48]	9,323 (10K)	Btau3.1	2 (713)	Fay & Wu's H, F_{ST}, ΔDAF	Body composition, Carcass yield
2	Hayes et al. [49]	9,323 (10K)	Btau3.1	4 (774)	AFD, iHS	Milk and Meat production
3	Gibbs et al. [23]	37,470 (BHMC)	Btau3.1	19 (497)	F_{ST}, iHS, CLR	Domestication traits: Behaviour, Immunity (MHC), Feed efficiency, Double Muscling, Milk yield and composition, Intramuscular fat content
4	Barendse et al. [121]	8,859 (10K) and 32,470 (BHMC)	Btau4.0	21 (385) and 19 (497)	F_{ST}	Residual feed intake, Beef yield (Intramuscular fatness)
5	Gautier et al. [61]	36,320 (50K)	Btau4.0	11 (437)	F_{ST} (BF)	Adaptation (pathogens and climate), Trypanosomiasis tolerance, Immune response, Nervous system, Skin and hair properties
6	Flori et al. [106]	41,777 (50K)	Btau4.0	3 (2803)	F_{ST}	Milk production, Reproduction, Coat color (Body coloration)
7	Chan et al. [60]	7,956 (10K)	Btau4.0	13 (317)	F_{ST}, EHH	Tropical adaptation: Tick resistance, Heat resistance, Immune system
8	Stella et al. [85]	32,689 (BHMC)	Btau4.0	19 (497)	CLL	Polledness, Coat color (Black, Piebald), Dairy production, Reproduction
9	Qanbari et al. [122]	40,854 (50K)	Btau4.0	1 (810)	EHH, REHH	Milk yield and composition, Reproduction, Behaviour, Dairy quality
10	Gautier and Naves [46]	44,057 (50K)	Btau4.0	22 (725)	iHS, Rsb	Reproduction, Metabolism, Immunity
11	Noyes et al. [125]*	21,034 (BHMC)	Btau4.0	17 (86)	XP-EHH, Fay & Wu's H	Trypanosomiasis tolerance
12	Qanbari et al. [70]	~42,600 (50K)	Btau4.0	12 (3,876)	F_{ST}, iHS	Reproduction (fertility), Muscle formation, Feed efficiency, Productive life
13	Glick et al. [156]	41,812 (50K)	Btau4.0	1 (912)	REHH	Milk (fat and protein) production and composition, Female fertility, Somatic cell score, Herd life
14	Hosokawa et al. [142]	40,635 (50K)	Btau4.0	2 (100)	SWAD	Carcass quality and yield, Milk (fat and protein) quality and yield
15	Larkin et al. [143]	52,150 (50K)	Btau4.0	1 (94)	AFD	Milk production, Fertility, Disease resistance
16	Schwarzenbacher et al. [157]*	34,851 (50K)	Btau4.0	1 (287)	iHS	Milk protein yield
17	Flori et al. [88]	47,365 (50K)	UMD3.1	19 (623)	F_{ST}, iHS, Rsb	Polledness, Thermotolerance
18	Pan et al. [158]	40,130 (50K)	Btau4.0	1 (2106)	EHH, REHH	Milk production, Reproduction, Immune system, Growth
19	Pintus et al. [95]	42,514 (50K)	Btau4.0	2 (1,113)	F_{ST}	Immune response, Production, Reproduction, Metabolism, Double muscles, Coat color
20	Boitard and Rocha [94]	35,554 (50K)	UMD3.1	1 (30)	HMM-SFS	Double Muscles, Body weight
21	Lee et al. [149]	15,125,420 (Sequencing)	UMD3.1	1 (12)	SFS (CLR)	Adaptive immunity, Domestication (BPV virus resistance)
22	Liao et al. [126]	9,990,733 (Sequencing)	UMD3.1	1 (14)	ZHp	Tropical adaptation: Resistance to hot temperature and tropical diseases (pathogens and parasites)
23	Lim et al. [159]	4522 (10K)	Btau3.1	1 (266)	iES	Beef production (intramuscular fat or marbling)
24	Utsunomiya et al. [141]	281,994 (800K)	UMD3.1	4 (136)	Meta-SS	Reproduction, Dairy and Meat production
25	Ramey et al. [62]	52,942 (50K) and 2,575,339 (AFFXB1P)	UMD3.1	15 (6,431)	Contiguously Low MAF	Behaviour, Polledness, Reproduction, Immune response, Ear floppiness, Beef production, Stature, Coat color
26	Druet et al. [86]	725,293 (800K)	UMD3.1	12 (674)	HMM-SFS	Polledness, Double muscle, Stature, Coat color
27	Porto-Neto et al. [103]	768,506 (800K)	UMD3.1	13 (505)	F_{ST}	Divergent selection between Zebu and Taurine cattle (eg., Stature, Fertility, Production, Immunity)
28	Rothammer et al. [80]	47,651 (50K)	UMD3.1	10 (430)	XP-EHH	Polledness, Double muscle, Coat color, Dairy and beef production traits
29	Kim et al. [139]	41,951 (50K)	UMD3.1	1 (2,087)	iHS	Dairy production (milk, fat, protein yields)
30	Mancini et al. [104]	29,848 (50K)	Btau4.0	5 (2,797)	F_{ST}	Reproduction, Lipid metabolism, Calving ease, Coat color, Stature, Dairy production
31	Edea et al. [138]	29,736 (50K)	UMD3.1	5 (47)	F_{ST}	Metabolism, Hypoxia response, Heat stress
32	Fan et al. [69]	39,094 (50K)	UMD3.1	1 (942)	EHH, REHH	Beef production and meat quality (carcass traits)
33	Randhawa et al. [50]	38,610 (50K)	UMD3.1	21 (375)	CSS	Polledness, Double muscle, Stature, Coat colors, Reproduction, Immunity
34	Porto-Neto et al. [28]	680,000 (800K)	UMD3.1	9 (1,842)	F_{ST}	Immunity, Reproduction, Metabolism
35	Kemper et al. [123]	610,123 (7K, 50K, 800K)	UMD3.1	8 (23,641)	F_{ST}, iHS, HAPH	Polledness, Coat colors, Stature, Milk production, Double muscle

(Continued)

Table 1. (*Continued*)

No.	Study	SNPs (assay)[1]	Genome assembly[2]	Breeds (samples)[3]	Selection Tests[4]	Traits reported[5]
36	Perez OBrien et al. [140]	~ 510,320 (800K)	UMD3.1	4 (235)	VarLD	Adaptation, Dairy and Beef production, Reproduction (Male fertility), Ear morphology
37	Qanbari et al. [105]	15,182,131 (Sequencing)	UMD3.1	1 (43)	iHS, CLR	Domestication traits: Coat colors, Neuro-behaviour, Immune system, Sensory perception
38	Lee et al. [128]	17,666,906 (Sequencing)	UMD3.1	2 (21)	XP-EHH, XP-CLR	Milk proteins and production, Coat colors
39	Kim and Rothschild [160]	Not given (50K)	UMD3.1	1 (162)	iHS	Adaptation
40	Ryu and Lee [161]	11,799 (50K)	Btau4.2	1 (71)	CLL	Meat quality, Marbling score
41	Somavilla et al. [162]	560,565 (800K)	UMD3.1	1 (789)	REHH	Growth, Feed efficiency, Meat quality, Fatty acid profiles, Reproductive traits, Immunity
42	Xu et al. [163]	710,681 (800K)	UMD3.1	5 (169)	di (F_{ST})	Milk yield and composition, Body size (Stature), Coat colors, Development, Disease resistance, Skeletal muscle formation, Mammary gland, Immune response, Nervous system
43	Bomba et al. [164]	44,271 (50K)	Btau4.0	5 (2813)	REHH	Milk and meat production
44	Gurgul et al. [165]	43,315(50K)	UMD3.1	2 (708)	SWAD	Milk production, Developmental process, Immune system, Fertility, Growth, Coat colors
45	Kim et al. [166]	37,154 (50K)	UMD3.1	1 (1602)	iHS, Rsb	Dairy production, Disease resistance, Immune response, Metabolism
46	Randhawa et al. [51]	38,033 (50K)	UMD3.1	28 (494)	CSS	Stature, Growth rate
47	Flori et al. [167]	38,100 (50K)	UMD3.1	1 (203)	iHS, Rsb	Adaptive admixture, Coat colors,
48	Bahbahani et al. [168]	46,171 (50K)	UMD3.1	1 (425)	F_{ST}, iHS, Rsb	Immunity, Reproduction, Development, Heat tolerance
49	Choi et al. [169]	17,936,399 (Sequencing)	UMD3.1	2 (20)	ZHp	Meat production (Intra-muscular fat)
50	Gurgul et al. [170]	40,378 (50K)	UMD3.1	1 (201)	REHH	Skeletal muscle growth, Meat quality
51	Kasarda et al. [171]	30,538 (50K)	UMD3.1	1 (40)	F_{ST}, iHS	Feed efficiency, Coat colors, Growth
52	Kim et al. [172]	45,632 (50K)	UMD3.1	1 (586)	iHS	MHC, Gastrointestinal nematode resistance
53	Li and Kim [173]	35,968 (50K)	Btau4.0	1 (547)	REHH	Meat and carcass production
54	Makina et al. [174]	21,290–45,657 (50K)	UMD3.1	6 (247)	F_{ST}, Contiguously Low MAF	Tropical adaptation, Nervous system, Immune response, Production, Reproductive performance
55	Sorbolini et al. [175]	43,009 (50K)	UMD3.1	2 (774)	F_{ST}, VarLD	Double muscling, Immunity, Behaviour, Reproduction, Metabolism, Bone morphogenesis
56	Zhao et al. [176]	705,234 (800K)	UMD3.1	7 (3122)	F_{ST}, iHS	Milk production, Reproduction, Body size, Muscle formation, Coat colors

[1] Genotyping SNPChip assays; **10K**: MegAllele™ Genotyping Bovine 10K SNP Panel, features almost 10,000 SNPs randomly (not very evenly spaced) spanning the bovine genome. **BHMC**: The International Bovine HapMap Consortium (genome-wide panel of 37,470 SNPs. **50K**: Illumina BovineSNP50 Genotyping BeadChip, features more than 54,000 (version 1: 54,001 and version 2: 54,609) evenly spaced SNP probes (~51.5kb) spanning the bovine genome. **800K**: Illumina BovineHD BeadChip, features 786,799 SNPs that uniformly span the entire bovine genome. **AFFXB1P**: Affymetrix Axiom(™) Genome-Wide BOS 1 Array Plate, comprise 648,855 SNPs representing genetic diversity of ~3 (out of 46) million SNPs across breeds that uniformly locate after each 1 kb along the bovine genome and 2.5 million (2,575,339) SNPs from the Affymetrix SNP validation data were used in the reported study. Chromosome-wise SNP density of each genotyping assay is shown in S2 Fig.

[2] Bovine Genome assemblies that were according released as; **Btau3.1** in August 2006, **Btau4.0** in October 2007, **UMD3.1** in December 2009.

[3] Breed names and breed-wise sampling and genotyping information is provided in Table 2.

[4] Complete names and description of selection tests are provided in S2 Table.

[5] Summary of major findings related to various traits of economic importance from each study.

* Two studies have not provided complete (extended) list of significant regions, however, they performed genome-wide scans, therefore, the significant regions from these studies have been included in the computation of meta-scores. Additional studies that used targeted sequencing, microsatellites and parse genotypes on few chromosomes are provided in S1 Table.

doi:10.1371/journal.pone.0153013.t001

Table 2. A list of selection tests used in published studies on Bovine selection signatures.

Test	Description	Bovine references
Fay & Wu's H	**Fay and Wu's H test**: Detects an excess of high (compared to intermediate) frequency variants that are likely to have been influenced by positive selection [177].	[48, 125, 178]
AFD	**Allele Frequency Difference**: Detects positive selection as the difference between allele frequencies of two populations, specifically used as Sliding Window Average Difference (**SWAD**) for a set of adjacent SNPs by averaging the absolute values of AFDs calculated along the genome.	[49, 142]
F_{ST}	**Fixation Index** (Population differentiations): Detects both newly arising and pre-existing variation under selection by measuring the allelic diversity between populations versus within population [179–182].	[23, 28, 48, 50, 60, 70, 87, 95, 103, 104, 106, 121, 123, 138, 168, 171, 174–176]
di	**Divergence** (locus specific divergence in breed i): Detects high levels of population structure in loci of breed i by standardizing F_{ST} between breed i and other j breeds, by using all pairwise combinations of genome-wide F_{ST} [183].	[163]
BF	**Bays Factor**: Detects divergence selection from Bayesian binomial frameworks for loci that show concordant differences in allele frequencies across populations (such as F_{ST}) with respect to specific aspects of the selective pressures [184, 185].	[61]
CLR	**Composite Likelihood Ratios**: Detects selective sweeps by modelling the spatial (chromosome-wise) distributions of allele frequency under the selection versus neutrality within a population, in addition, taking care for ascertainment bias, recombination rate and demography [186]. Recently, XP-CLR is used for across population analyses.	[23, 105, 128]
CLL	**Composite Log Likelihood**: Detects positively selected regions of the genome by comparing the frequencies of major (common) alleles for a set of contiguous loci between the samples of a unique sub-population and a larger panel of diverse populations.	[85]
HMM-SFS	**Hidden Morkov Model—Site Frequency Spectrum**: Detects positive selection within a population at regions of reduced heterozygosity by modeling the correlation structure between linked sites that uses site frequency spectrum and the spatial pattern of diversity among the sequence or polymorphism [56].	[86, 94]
Low MAF	**Low Minor Allele Frequency**: Detects complete selective sweeps within the population considering the cluster of adjoining loci carrying very low MAF (< 0.01).	[62, 174]
ΔDAF	**Change in Derived Allele Frequency**: Detects positively selected new causal mutations as the difference in the derived (non-ancestral) allele frequencies between populations [150].	[48, 50, 141]
ZHp	**Z-transformed Heterozygosity Value**: Detects selective sweeps by counting alleles in a sliding window centered on a candidate SNP, then calculates heterozygosity scores (Hp) from the pool of samples from within a population (breed) and extreme (negative) Z-transformed Hp values represent reduction in heterozygosity in the candidate regions [187].	[126, 141, 169]
VarLD	**Variation in Linkage Disequilibrium**: Detects candidate regions under positive selection by comparing genome-wide LD variation between populations [188] and it is implemented in the varLD program [189].	[140, 175]
HAPH	**Haplotype Homozygosity**: Detects strong positive selection (hard sweeps) by comparing the frequency of the core haplotype against the total number of haplotypes observed within the breed (population) that implements the neutrality tests based on the distribution of haplotypes under an infinite-site model [64, 65, 103, 140, 155, 190].	[123]
EHH	**Extended Haplotype Homozygosity**: Detects positively selected regions carrying frequent haplotypes with unusually high long-range LD patterns within a population [191, 192].	[60, 69, 93, 122, 193]
REHH	**Relative EHH**: Detects evidence of recent selection on relatively high frequency haplotypes within a population by comparing the **EHH** of the tested core haplotype to that of other core haplotypes present at a locus to correct for local variation in recombination rates [191].	[69, 122, 156, 170, 173]
XP-EHH	**Across Population EHH**: Detects selective sweeps by comparing **EHH** across populations in which selected alleles (at core haplotypes) have risen to near fixation in one but not all populations [194].	[50, 80, 125, 128]
iHS	**Integrated Haplotype-homozygosity Score**: Detects evidence of recent positive selection at a locus based on the differential levels of LD surrounding a positively selected (derived) allele (at intermediate frequencies) compared to the background (ancestral) allele at the same position within a population [73].	[23, 46, 49, 70, 88, 105, 123, 139, 141, 157, 167, 168, 171, 172, 176, 195]
iES	**Integrated extended haplotype homozygosity at SNP site**: Detects recent positive selection by finding lower levels of EHH decay, by estimating locus-wise (overall rather than a particular allele) haplotypic homozygosity over a two-way distance each SNP site, using a counting algorithm implemented to genotypic (non-phased) data at within a population [196].	[159]
Rsb	**Across Population iES**: Detects recent selection on completely or nearly fixed selective sweeps by comparing the single locus **iES** associated with the same site and genomic region across populations [196]. Rsb and XP-EHH are based on similar assumptions to target haplotype decay, so they can be substituted.	[46, 88, 141, 167, 168]
Meta-SS	**Meta-analysis of Selection Signals**: Detects evidence of recent positive selection in common variations by combining P-values obtained from Gaussian cumulative distribution function of **ΔDAF**, **Rsb**, **iHS** and **ZHp** tests.	[141]
CSS	**Composite Selection Signal**: Detects positively selected genomics regions carrying highly differentiated loci and underlying variants hauling excess haplotype in the samples of a target population versus phenotypically contrasting populations using the rank distribution approach to unify the multiple pieces of selection evidence from F_{ST}, **ΔDAF** and **XP-EHH** tests.	[50, 51]

doi:10.1371/journal.pone.0153013.t002

and middle (or top score location, if given) of a selection signature using common reference markers across panels.

Breeds and groups of breed-types. All breeds and crossbred cattle from the study population can be divided into four breed-type groups, according to their geographical origins, namely; European (*Bos taurus*), African (*Bos taurus*), Zebu (*Bos indicus*) and composite (*Bos*

taurus × *Bos indicus*) and are detailed in S2 Table. Since the historical events of selection in these groups have been unique, the data from these four groups were organized and analysed separately. Further information on the origin of breeds, sample size, SNP density, bovine genome assembly, selection tests are also shown in S2 Table.

The populations listed in S2 Table are considered input datasets and the output (results) have been presented according to populations detailed in S3 Table. Based on the input datasets, the populations of 86 breeds have been used either as breed-wise, group-wise (grouped with other breeds) or as a reference panel (S3 Table). For example, Holstein has been investigated in 33 studies, where 22 studies presented results for the breed, seven studies presented Holstein results grouped with other populations and four studies have used Holstein samples as a reference panel to investigate selection signatures in other breeds (S3 Table). Overall, breed-wise results were available from 53 cattle breeds representing 1 to 22 studies. Moreover, 72 breeds have been presented as group-wise results consisted of 1 to 7 studies, whereas 33 breeds were used as various reference panels consisted of 1 to 4 studies (S3 Table).

The published results complied as breed-wise and group-wise selection signatures were used to compute meta-scores. For each position of published selection signature, two adjustment factors were computed: 1) to adjust for the use of repeated DNA samples across studies and 2) to account for the significance thresholds used for the top SNPs within each study. These adjustment factors are termed as "DNA score" and "SNP score", respectively, and are described in detail below.

Common usage of DNA samples across studies. It was noted that several breed-wise genotyping datasets were re-analyzed across different studies using different analytical tools. For example, DNA samples (breeds) from the Bovine Hap Map Consortium (BHMC) and subsequently genotyped with bovine Illumina BovineSNP50 (50K) and BovineHD (800K) BeadChip assays, have been used repeatedly in several studies (S2 Table). Furthermore, several studies have used new samples combined with samples used previously. Therefore, for every breed, which was investigated in multiple studies, the proportion of common individuals between each pair of *i* and *j* study was computed to account for common usage of DNA samples.

If n_i, n_j are the total number of DNA samples and n_{ij} is the number of common DNA samples between each pair of studies *i* and *j*, and there are *k* studies on a particular group/breed *b*, then the DNA score d_i can be computed for breed *b* of study *i* as follows:

$$d_i = \frac{1}{\sum_{j=1}^{k} \left(\frac{n_{ij}}{(n_i + n_j) - n_{ij}} \right)}$$

It is noteworthy that not all studies provided results for individual breeds; rather a selection signature was reported for a group of multiple breeds with a common defining attribute, e.g., polled, dairy, African or Zebu breeds. In those cases, common usage of the total number of DNA samples across multiple studies was computed as the DNA score using the above formula. Similarly, the DNA score for the subset of breeds represented in a particular group was computed e.g., a group represented by European breeds (Fig 1). Moreover, DNA scores d_i were re-computed for each analysis using the individual breed for which $k > 1$ published results of selection signatures were investigated.

SNP score (*s*). The different significance thresholds used across studies to declare significant selection signatures (regions or SNPs under selection) could distort the signal under a meta-assembly. Hence, a SNP score s_i, was assigned to provide an approximately equal contribution from each investigation (i.e., study *i*). Across all studies the range of significance thresholds vary from the top 5% to top 0.1% using various distributions of results, such as, *p*-values, *q*-values, percentiles etc. A simplistic approach was implemented to define *s*-values, so
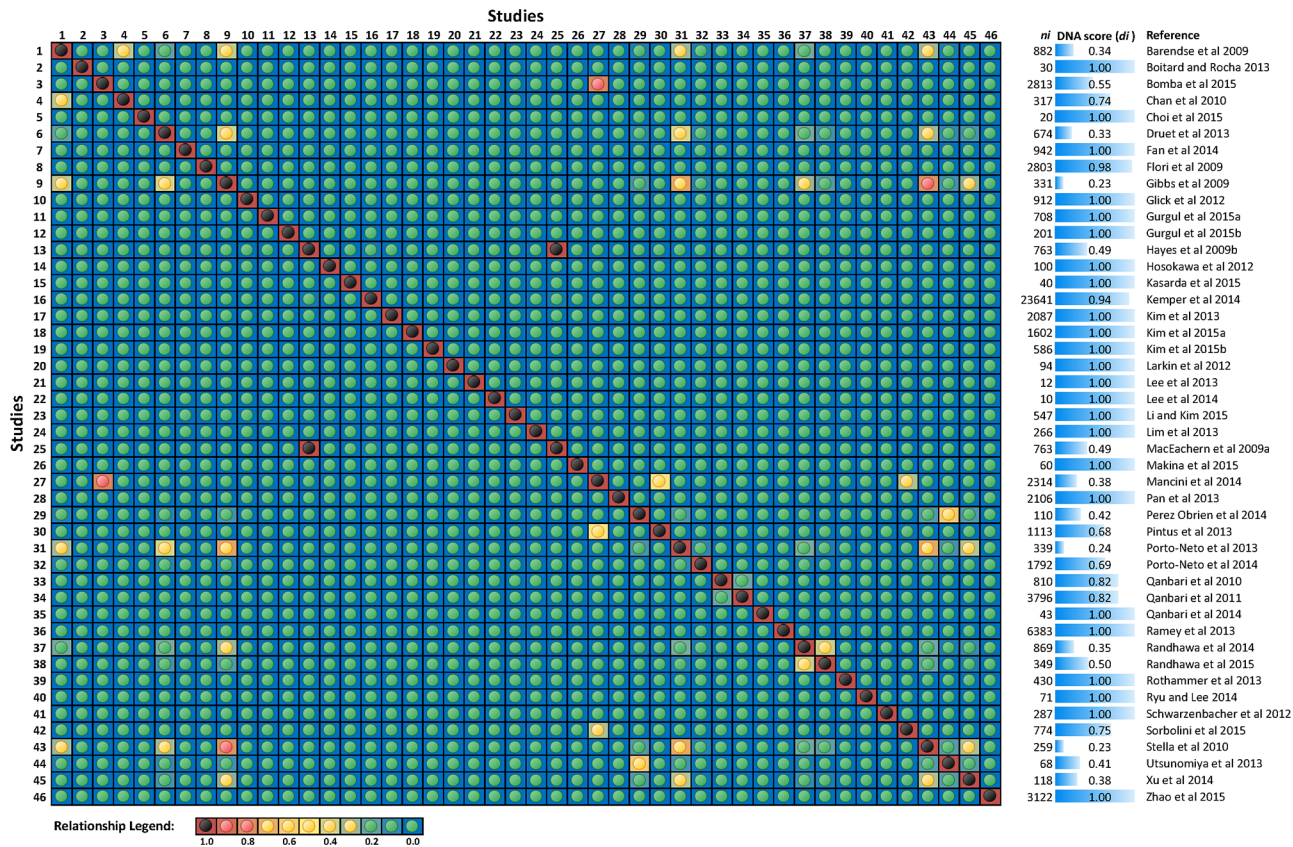
Fig 1. Plot of relationship matrix and DNA score (d_i) weighting computed from 46 studies that published selection signature using European cattle. n_i shows total number of DNA samples from European breed(s) from a particular study. The relationship score was computed as proportion of common samples between each pair of studies and it range between 0 and 1.

doi:10.1371/journal.pone.0153013.g001

that the set of published selection signatures within the top 5%, 1% and 0.1% thresholds were weighted as $s = 1$, $s = 2$ and $s = 3$, respectively.

Weighted selection signatures (SS). The DNA score (d_i) and SNP score (s_i) values were used to assign a weighted selection signature (SS) score at each unique position. For instance, for each breed-wise or group-wise result from any study i, at a locus l, the SS_{i_l} score was equal to $d_i \times s_i$.

Computation of meta-selection-score (MSS) across studies. Finally, the meta-score of selection signatures, termed as "meta-selection-score" (MSS) at each unique reported position, locus l, was computed as the sum of SS value within a sliding window (2 or 5 Mb) from multiple studies:

$$MSS_l = \sum_{i=1}^{k} SS_{i_l}$$

If a study has reported multiple positions within the window then only highest SS was included. It is noted that the span of the published selection signatures ranged from a single bp to 46.10 Mb (S4 Fig). For the selection signature reported as spans, the span overlapping with the sliding window were included in the computation of MSS.

A weighted SS from an individual study (for a breed or group) can contribute a maximum value of 3, when the DNA samples were completely independent ($d_i = 1$) and the significance

threshold was in the top 0.1% ($s_i = 3$). Notably, multiple tests of selection within a study were standardized for repeat sampling relationship in d_i and for their individual thresholds of significance in s_i. Hence, a MSS value above 3 represents a selection signature detected more than once (i.e., independently validated).

A higher magnitude of the MSS at a genomic position shows the consistency of a selection signature detected across studies within a breed or within a group. All peaks above the validation value of MSS > 3 were used to report the validated genomic regions under selection for each population, separately for the groups and breeds. The first and last positions of each peak were used to define the boundaries of putative regions. In general, all peaks were clearly separated from each other within the groups and breeds. However, because of the high density and large number of published signatures of selection for the European group and extensively studied breeds (Holstein and Simmental), several consecutive peaks were overlapping at MSS > 3. Therefore, the boundaries of genomic regions under selection in European, Holstein and Simmental meta-assemblies were defined where consecutive peaks intersect each other, i.e., the span of lowest MSS value(s) between two consecutive peaks.

Sliding-window spans for genome-wide MSS. S5 Fig compares a detailed analysis of the most important MSS peaks localizing classic selective sweeps in the European, Angus and Holstein meta-assembly by using 5 Mb and 2 Mb sliding window spans. The smaller (2 Mb) sliding window spans provides a narrow MSS peak around the candidate gene regions. However, the power of the meta-assembly analyses to fine map the hotspots of selection depends upon the number of available k studies. Hence, using a smaller window size can have profound effect at the genomic regions harbouring putative sweeps investigated in fewer studies. For example, out of the 25 cattle breeds analysed, the magnitude of MSS using 5 Mb spans in the genome-wide distribution of eight breeds have not been found above the minimum validation value (MSS = 3). Using the smaller window spans (1–2 Mb) can eliminate further breeds that have been investigated in fewer studies and show prominent (MSS > 3) peaks using larger (5 Mb) spans. Hence, the size of the sliding window for the computation of MSS_l was set to 5 Mb, i.e., ~ 2.5 Mb on each side of a locus l.

Results
MSS maps of cattle

MSS were computed for all 4 groups and 25 breeds where published results were available from multiple studies ($k \geq 2$). Moreover, the selection signatures from an additional 28 breeds (S6 Fig), where only a single study ($k = 1$) was available (S3 Table), contributed to the group-wise MSS only. In total 16,158 (European = 13,640, Zebu = 1,246, African = 1,112 and Composite = 160) individual selection signature scores contributed to the meta-assembly.

MSS maps of cattle groups. Meta-assemblies of European, Zebu, African and Composite groups were constructed from 46, 12, 9 and 8 studies, respectively (Fig 2). A comparison of MSS, gene density and individual published signatures of selection (showing their origin from studies, breeds, selection tests, SNP panels and significance thresholds) is presented for each chromosome of European *Bos taurus* group (S7 Fig). Overall peaks with extreme MSS in a meta-assembly map identify the hotspots of positive selection within cattle groups (Fig 2).

In the European group, the meta-assembly shows that significant signatures of selection have been detected across the whole genome. To define genome-wide coverage of published signatures of selection, the bovine genome was divided into 1 Mb non-overlapping windows. More than 90% of the 1 Mb windows contained a published signature of selection, and a maximum gap of 4.47 Mb between the consecutive genomic regions under selection was observed on bovine autosome (BTA) 1. The top 5% of MSS (MSS $\geq$ 75) of the meta-assembly were
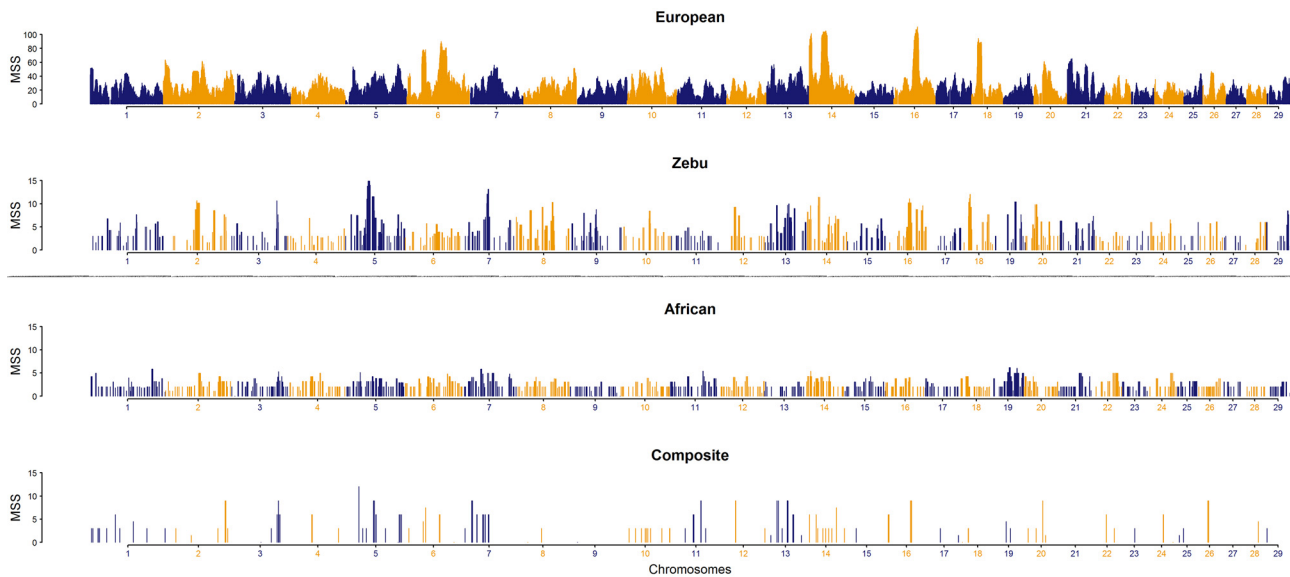
Fig 2. Meta-assembly of selection signatures in four groups within European, Zebu, African and Composite breeds of cattle.

located on BTA-6, BTA-14, BTA-16 and BTA-18. In addition, most of the chromosomes harbour prominent peaks within the top 25% (MSS $\geq$ 35), except BTA-15, BTA-23 and BTA-28, which have their top peaks above 50% (MSS $\geq$ 25).

On the other hand, because of limited number of investigations and breeds of Zebu, African and composite cattle, the genome coverage in these groups was limited relative to European *Bos taurus* investigations. In the Zebu group, major MSS peaks in the top 5% (MSS $\geq$ 11) were located on BTA-5, BTA-7, BTA-14, BTA-16 and BTA-18. In the African group, the MSS peaks in the top 5% (MSS $\geq$ 5.2) were located on BTA-1, BTA-3, BTA-7, BTA-11, BTA-14 and BTA-19. In the composite group, the MSS peaks in the top 5% (MSS $\geq$ 9.0) were located on BTA-2, BTA-3, BTA-5, BTA-7, BTA-11, BTA-12, BTA-13, BTA-16, BTA-20 and BTA-28.

In total, 439 prominent peaks of validated genomic regions (i.e., clusters of MSS > 3) were detected genome-wide for the four archetype groups (S4 Table), distributed across European = 173, African = 116, Zebu = 120 and Composite = 30 respectively (Table A in S1 File). Out of those, 141 validated genomic regions under selection were found putatively unique for each group viz. European = 64, African = 37, Zebu = 36 and Composite = 4. On the other hand, 298 peaks from the four archetype groups (European = 109, African = 79, Zebu = 84 and Composite = 26) shared at 122 genomic regions (Table A in S1 File). The 298 co-aligning peaks (122 shared regions) were common between at least two groups, such that; 36 peaks were shared at nine regions between all four groups, 108 peaks at 36 regions were shared between any three groups and 154 peaks at 77 regions were shared between any two groups (Table B in S1 File). Overall, there were 263 hotspots of positive selection at prominent peak scores found within (*n* = 141) or shared across (*n* = 122) the four cattle archetypes ranging in size from 0.1 to 12.34 Mb.

MSS maps of cattle breeds. Fig 3 shows the meta-assembly of selection signatures across 13 breeds, each constructed from at least *k* = 5 studies, where as MSS for a second cohort of 12 breeds which were represented in less than five (*k* = 2 to 4) are shown in Fig 4. The meta-assemblies of these 25 breeds have been constructed using signatures of selection from 2 to 22 studies (S3 Table).
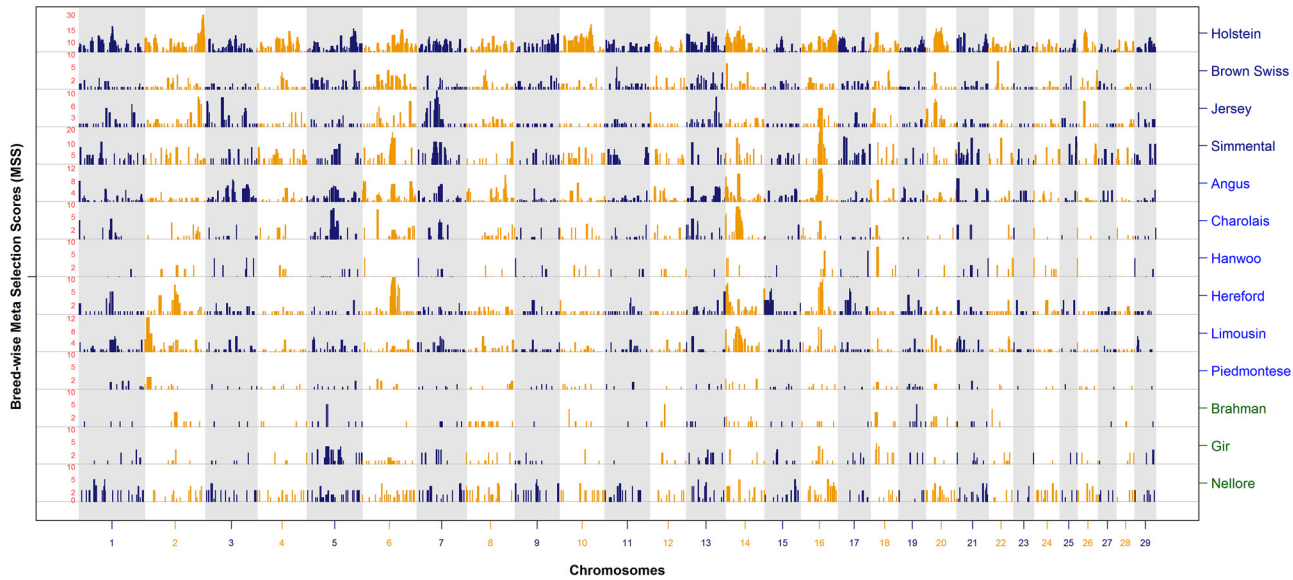
Fig 3. Meta-assembly of selection signatures of Holstein, Brown Swiss, Jersey, Simmental (Fleckvieh), Angus, Charolais, Hanwoo, Hereford, Limousin, Piedmontese, Brahman, Gir and Nellore cattle.

doi:10.1371/journal.pone.0153013.g003

The breed-specific meta-assemblies show that several genomic regions have been consistently detected within and across breeds. However, only 16 of the 25 breeds showed validation (MSS > 3) of selection signatures across independent studies. Lack of consistency was seen in the breeds with limited studies ($k \leq 3$); for example, Guernsey, Italian Brown, Murray Grey, NDama, Norwegian Red, Santa Gertrudis, Sheko and Shorthorn (S3 Table). There are some exceptions for Belgian Blue ($k = 4$), Braunvieh ($k = 3$) and Romagnola ($k = 2$). Several signatures of selection at the lower levels of the significance threshold did not achieve validation (MSS > 3) in some breeds (Fig 4).
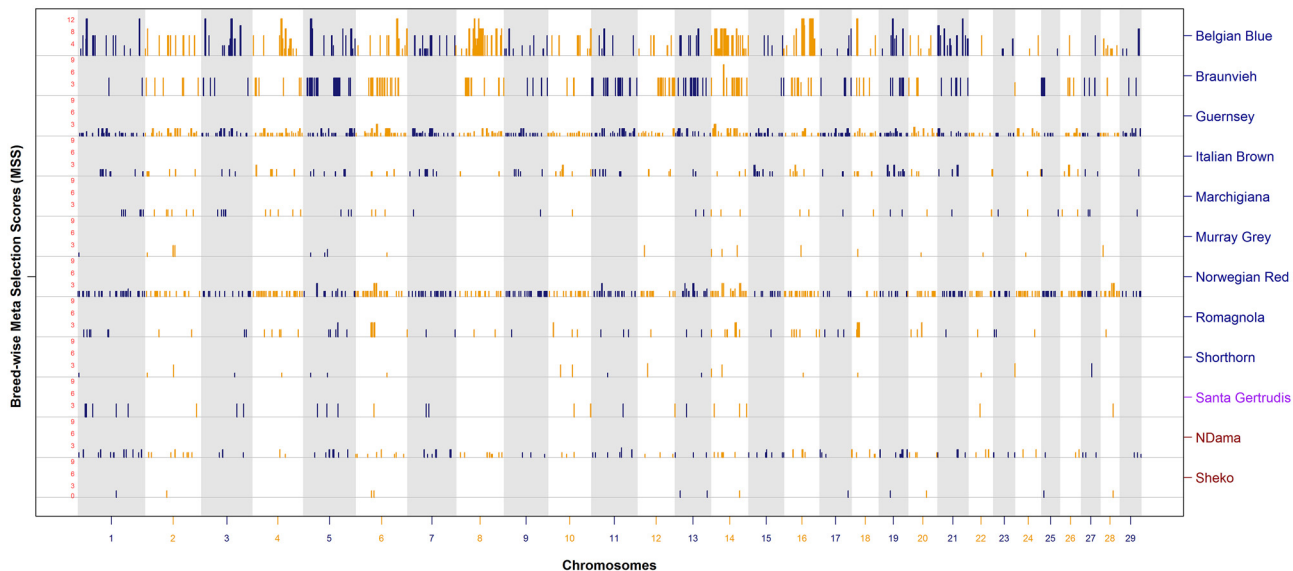
Fig 4. Meta-assembly of selection signatures of Belgian Blue, Braunvieh, Guernsey, Italian Brown, Marchigiana, Murray Grey, Norwegian Red, Romagnola, Shorthorn, Santa Gertrudis, NDama and Sheko.

doi:10.1371/journal.pone.0153013.g004

Overall, 854 prominent peaks of validated MSS were detected for 16 breeds. The complete list of validated MSS peaks, their location and the genes underlying each region is provided in S5 Table, and their unique or shared localization is summarized in Tables A and B in S2 File. In summary, out of the 854 peaks located at 341 hotspots under selection, 113 peaks were found unique within 15 breeds and 741 MSS peaks were shared across 2 to 12 breeds at 228 hotspots (Table A in S2 File). Out of the 228 shared regions, 170 (494 peaks) were found across multiple breeds within European archetype, whereas, 58 (247 peaks) were found shared between breeds of the European and Zebu archetypes (Table B in S2 File). No unique or shared validated MSS peaks were found in African or composite breeds. A direct comparison with the subset of European breeds reviewed earlier [35] was not practical as we report validated MSS and earlier study presented unprocessed list of selection signatures as core selective sweep regions.

Genomic regions under selection across groups and breeds. Fig 5 illustrates the comparison between the prominent regions across the 4 cattle archetypes and 16 cattle breeds. In cattle groups, the nine shared validated MSS peaks across all four archetypes were located on BTA-4, BTA-5, BTA-7, BTA-12, BTA-14, BTA-16 and BTA-19. In cattle breeds, most of the shared signatures of selection, in four or more breeds, were located on all but BTA-23, BTA-24, BTA-25 and BTA-28. Common genomic regions across various populations reveal historical selection shared between those cattle breeds, most likely due to ancestral (within archetypes), geographical and/or commercial similarities. Unique genomic regions under selection, which are private to a single breed, can be responsible in shaping particular characteristics of the population resulting in the origin and maintenance of that breed. Several of the important genes known for their role in some major traits have been labelled (Fig 5) and complete list of genes underlying each region within groups and breeds are provided in S4 and S5 Tables, respectively. Nevertheless, the number of genes within each hotspot of selection varied between 0 to 382 due to the variable span of MSS and gene density in the bovine genome (Fig 5).

Comparison of meta-assembly and gene density. To align the prominent regions with gene density, the number of genes within 1 Mb spans were counted along the genome. On
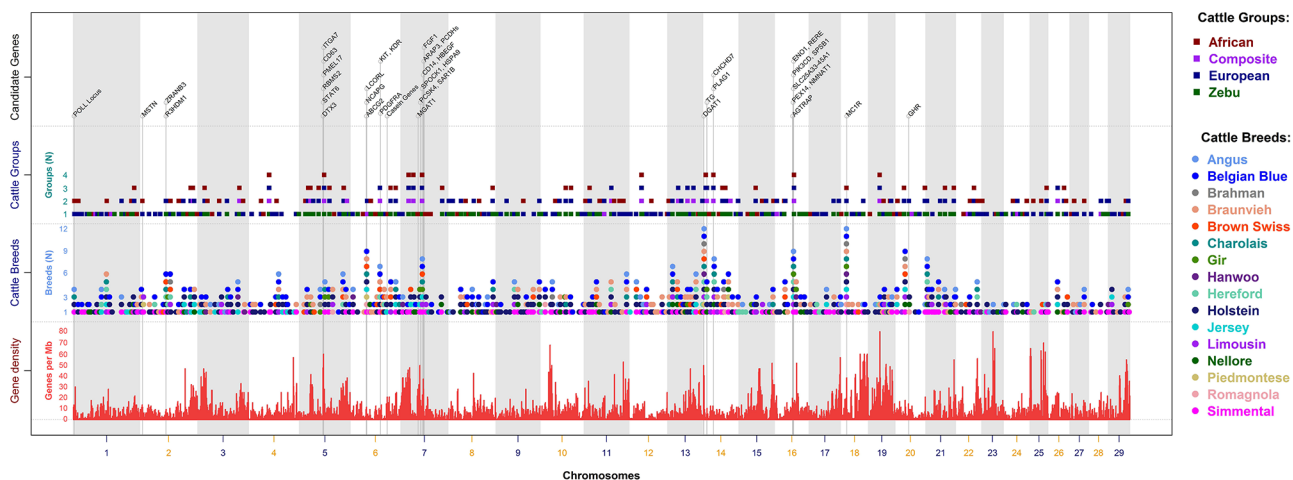
Fig 5. Map of selection signature hotspots captured in the meta-assembly of cattle breeds and groups. Middle Panels labelled as "Cattle Breeds" and "Cattle Groups", show the location of prominent regions in the cattle breeds and groups, respectively represented with the unique colours as shown in the legends. The clustered dots, within a locus, located on top of each other represent shared selection signatures across the breeds and groups, each of which has been validated in multiple investigations. Lower Panel labelled as "Gene density" shows distribution of bovine genes on each chromosome that ranges 0–80 genes/Mb (S7 Fig shows high-resolution comparison of MSS and genes per Mb in European). Top Panel labelled as "Candidate Genes" shows genomic locations of a few major genes underlying the outstanding peaks representing classic selective sweep regions in the meta-assemblies. Complete list of prominent regions, localized top MSS and underlying genes within the groups and breeds of cattle are respectively shown in S4 and S5 Tables.

doi:10.1371/journal.pone.0153013.g005

average, bovine genome contains approximately 9 genes per Mb and 7.5% of all 1 Mb regions are without any genes. A chromosome-wise comparison of meta-assembly and gene density for the European (S7 Fig) shows that signatures of selection have been detected in gene-dense as well as gene-poor regions. Using the spans of prominent peaks, we also compared genome-wide average gene density against the number of genes in each prominent region (S4 and S5 Tables). Many hotspots of selection localized in genomic regions with medium-to-high gene density. However, some prominent peaks within low-to-medium density genic regions indicate that selection targeted a few or single genes of major effects, for example, *MSTN* (BTA-2), *ABCG2* (BTA-6), *NCAPG-LCORL* (BTA-6), *PLAG1-CHCHD7* (BTA-14), and *GHR* (BTA-20). Inferring the candidate genes underlying signatures of selection at high gene-density regions is challenging. Patterns of genetic diversity implicating strong selective sweeps may have been generated by a cumulative effect of selection acting on multiple genes, for example, the regions identified in multiple populations at chromosome 5, 7 and 16 (Fig 5).

Discussion

Insights from meta-assembly

This study implements a novel approach to infer hotspots of positive selection in the bovine genome by using diverse set of published selection signatures across worldwide cattle breeds. The approach of computing meta-scores and construction of a meta-assembly map of selection signatures can be widely used where conventional meta-analysis methods cannot be applied in scenarios similar to this study. The meta-assembly highlights several genomic regions of variable gene density where positive selection has been replicated within and across breed-populations. In general, MSS across multiple breeds show that selection has acted on candidate genes; however, the span of selection is generally wide because of LD and hitchhiking in the target regions. The hotspots of selection could be linked to various biological functions enriched across those breeds, especially related to adaptation (disease, climate, feed resources), appearance (polledness, coat colors) and production (milk, meat, fertility) traits, each of which have economic importance in various environments and production systems. Several regions harbouring candidate genes controlling major traits, e.g., coat color (*MC1R*, *KIT*), stature (*NCAPG*, *LCORL*, *PLAG1*) and milk production (*ABCG2*, *DGAT1*, *GHR*) are consistently identified under strong selection in multiple breeds of cattle (Fig 5). There are also clearly strong selection signatures at some genomic regions, e.g., chromosome 7 and 16, where the associations for QTLs and functional variants of multiple traits have been localized. Overall, the genomic regions underlying prominent peaks of meta-scores provide the evidence of QTLs for the traits being constrained in cattle populations. Limited efforts for gene-networks and functional analyses for the selectively differentiated regions across cattle breeds warrant extensive resources in future.

Several pitfalls in construction of the meta-assembly originate from the limitation of individual studies and tests of selection. Ascertainment bias and introgression in the Zebu and African breeds resulted in lack of Zebu and African specific selection signatures. Using F_{ST} or similar tests of selection, which lack directionality to point the population under selection, can lead to misleading interpretations about selection signatures. Another common reason of finding coinciding genomic regions under selection relates to the use of reference populations across the European, African and Zebu breeds. Moreover, most of the breeds tested for selection in non-European cattle, for example Zebu, are not representatives of overall breed-type. The investigated breeds (Brahman, Nellore, Gir etc.) might have experienced bottlenecks because the samples used in selection signature scans are from imported herds into non-native countries that started from a few hundred founder animals, rather than representatives of the

natural populations. To date, limited investigations on African and Zebu breeds are available; therefore, conclusions are limited since these studies are mostly underpowered for these two archetypes as compared to the European breeds.

The tests of selection, which were implemented in the cattle studies, have been established and validated with human genomic data. Therefore, some of the intrinsic problems can relate to the differences in genetic architecture and demographic history of the two species. Establishing (fine-tuning) these tests considering the demographic history of cattle, can fine scale the regions under selection [37, 38]. Nonetheless, there are several common phenomena of detecting selection signature between cattle and human. The span of the regions under selection is generally wide; however, most of the tests of selection, based on the extensively used outlier approaches, detect only a section of the candidate region. As there can be many proxy SNPs around the causal variants, the overlap between various genome-wide selection scans that are in common is rare in both species. Low replication rate may point to high incidence of false positives, hence, validation data sets are required to exclude the rate of false positives, along with accounting for all the challenges causing confounding effects that generate false positive results. Notably, the bovine genome appears to be completely covered with selection signatures especially in the group of European breeds.

Implementation of new approaches, such as a meta-assembly, can provide the insights to resolve the false and true positive regions under selection. The magnitude of the meta-score depends upon the span of sliding windows and the number of studies reporting the co-located signals. Hence, given sufficiently available studies, the larger and smaller window spans, respectively, can elucidate the overall broad and fine-mapped functional regions for each selection signature. The meta-assembly approach can potentially be improved by incorporating additional measures, such the effects of sample size, as they become available from future research.

Challenges in identification of selection signatures

Critical factors that may influence the accuracy of individual studies and consequently the meta-assembly, are confounding effects of genomic structure within each breed (genetic drift, demographic effects, recombination rate, and age and type of selection) and methodological biases (sample size, ascertainment schemes, density and distribution of SNPs, power of the tests of selection, null distribution for significance threshold and selection models). Besides selection pressure, different demographic (non-selective) forces such as population size and gene flow can generate the patterns of population's genetic diversity [16]. Moreover, distinguishing the effect for selection and demography of a population is challenging [39]. Hence, the population history of current breeds may incur convolution to the investigation of genomic history, i.e., patterns of genetic diversity [40–42]. In human, different demographic scenarios mimicked identity of hitchhiking patterns, which shows the importance of estimating underlying demographic model to accurately implicate the genomic regions for a selection event [43]. In cattle, geographically isolated evolution and domestication, introgression, co-ancestry, admixture, migration, bottlenecks, inbreeding, and variable population sizes intricate the insights of cattle history [18, 23–25, 27, 44–49]. Incorporating this knowledge into selection models could help differentiate the true positive selection from that of confounding patterns of genetic drift arising from non-selective forces of primeval demography. For example, combining different populations/breeds based on similar phenotype can help to detect trait specific signatures of selection [50, 51].

The quality of genetic polymorphism data also affects the population genetic inferences [52–57]. In cattle, ascertainment bias has occurred for the African and Zebu cattle because the development of various current genotyping assays (e.g., Illumina SNP50 BeadChip) primarily

based on European breeds [58]. In addition to under-representation of rare variants in geno-typed samples of the non-European cattle breeds [59], the consequences of ascertainment bias include limited SNP density. Multiple studies have reported limited set of polymorphic SNPs in the African and Zebu breeds as compared to the European data [24, 25, 46, 48, 50, 60–62]. Genome-wide low density and inconsistent distribution of genotypic panels may produce false results and can have a strong influence on the population genetic procedures being used to esti-mate selection signatures. Recently, Heslot et al. [63] suggests that increasing the marker den-sity may reduce the effects of ascertainment bias and increase the accuracy of genetic diversity estimates. Hence, appropriate approaches should be implemented for the individual study data to correct the ascertainment bias.

In general, parameters of selection are inferred using estimates based on site frequency spec-trum, population differentiation and haplotype length (extend of LD) [43, 64, 65]. Several approaches implement these estimators by comparing genetic diversity within or across popu-lation and across species to detect various types (modes) and age (time-frame) of selection [7]. Two major modes of selection (positive and balancing) can be easily distinguished by most of the basic estimates of selection [66, 67], which have been extensively implemented in cattle (Table 2). Nevertheless, there are several challenges related to various tests [36, 68], and use of inappropriate approaches provide misleading conclusions [67]. Notably, because of the varying levels of statistical power, consistency of results across various tests of selection using common datasets is limited [43, 69]. For example, the genotype (F_{ST}) and haplotypes (iHS) based esti-mates detected selection signatures at different genomic regions in cattle [70]. The single marker (genotype) based estimates, in general, are constrained being sensitive to ascertainment bias and have several limitations [71]. On the other hand, haplotypes are based on patterns at multiple SNPs and, therefore, haplotype based estimates of selection are less sensitive to ascer-tainment schemes of SNP discovery for the genome-wide panels.

Tests of selection are generally criticized for their limitation of sample size and approaches to declare significant (outlier) loci. Adequacy of the sample size can be established by the effec-tive population size and SNP density that can capture maximum genetic diversity for the popu-lation under investigation. Various investigations have provided the range of sample size to be sufficient to detect regions under positive selection for beneficial traits being carried out by applying various measures including allele and haplotype frequencies and extent of LD [72, 73]. These studies suggest that some of the commonly used selection tests (F_{ST}, XP-EHH) have reasonable power for the available genetic resources of cattle. However, the breeds genotyped by low-density SNP panels are likely to provide less reliable results. Recently available high-density genotypic data, albeit on limited sample size, enable high power to the selection tests, such as F_{ST} [74], however, substantial decline in power still occurs for rare genetic variants [75, 76]. Approaches of combining different tests into a single index, such as CSS [50], can over-come some of the limitations of individual tests.

Failure to detect truly positive signatures of selection often relates to the approaches used to declare the significance level, especially in the absence of null distributions for the selection tests. Generally, estimates of across-population differences in allele and haplotype frequency imply selection using the so-called outlier loci using an arbitrary cut-off [42, 77], however, the lack of empirical evidence supporting this assumption spawn criticism [36, 78]. The genomic regions experiencing mild selective pressure are not detectable by the outlier approaches [68]. On the other hand, demographic effects and the frequent occurrence of cold spots of recombi-nation in the genome can dominate the outlier distribution and confound selection [65, 79]. The complex demography and evolutionary diversity of cattle still pose a challenge to be dealt to control false positive selection signatures [80].

Challenges related to meta-assembly

The biases associated with the challenges of individual selection signature investigation have direct impact upon the computation of meta-scores. For example, several common regions identified between the European, African and Zebu breeds, shown at prominent peaks in the breed-type groups (Fig 2), indicate the impact of reference populations being used. It appears that most of these common regions have originated from the European breeds, however, because of non-directional presentation of the results from selection tests, such as F_{ST}, have declared those regions significant in non-European breeds as well. The results from such tests of selection should be interpreted with care and within the context by requiring the evidence of direction of selection. Moreover, the selection tests (F_{ST}, VarLD etc.) that lack the direction of selection should be implemented within the breed-types and analyses performed using breeds across the breed-types should be interpreted as demographic rather than selection signatures. The potential sources of additional bias may be due to factors including; across study variation in the sample sizes, SNP chip assay, data quality control filters, SNP density, selection tests, post-analyses transformation of result (smoothing, clustering etc.), lack of the magnitude of individual selection scores, and provision of full or partial lists of significant regions. With the available information about published selection signatures, it is not possible to account for these additional factors. Nevertheless, this meta-assembly provides a framework for alleviating some of the limitations of individual study and provides consensus regions under selection pressure. Future investigations would account for the pitfalls outlined above and provided the required information, subsequent methodological improvements to the computation of meta-scores are possible.

Signatures of historical selection in cattle

This article, to our knowledge, is the first meta-assembly detailing a consensus profile of genomic regions influenced by the historical selection events. The meta-assembly shows some very strong and congruent signatures of selection in multiple breeds for historically selected traits, such as polledness, double muscling, coat colors and bovine stature. Several additional genomic regions harbour genes for traits with complex selection history, such as adaptation, reproduction, growth, and high performance of dairy and beef production. Interestingly, some genomic regions have been implicated for selective constraints on multiple genes of different traits. The candidate genes underlying strong selective sweeps, which are depicted by prominent peaks in the meta-assembly maps, are annotated in Fig 5. This section presents an overview of the traits and genes that have experienced long-term selective breeding.

Genes and traits linked to strong selective sweeps in cattle. The *POLL* locus at the high gene density proximal end of BTA-1 is associated with horn development [81–84]. The strong positive selection in multiple breeds for naturally polled animals have swept the functional mutation in the region to fixation in several European breeds [50, 58, 62, 80, 83, 85–88]. In Zebu, the complex nature of inheritance suggests additional loci may be involved in horn development, however, a diagnostic marker at the *POLL* locus shows its dominance effect for polledness [84, 89].

The bovine myostatin (*MSTN*, AKA *GDF8* i.e., growth differentiation factor 8) gene on BTA-2 controls the muscular hypertrophy (double muscle) phenotype and strong selective pressure has escalated the allele frequencies of functional mutations because of its economical imporatnce in beef breeds [16, 90–93]. Hence, classic selective sweeps are extensviely found harbouring the *MSTN* gene in several genome-wide scans in beef cattle [23, 50, 62, 86, 93–95].

The *KIT* (on BTA-6) and *MC1R* (on BTA-18) genes stipulate breed-specific colors in a polygenic inheritance pattern in cattle [96–98] and other species [3, 99–102]. The *KIT* locus has been under strong selection to control white-spotting in multiple cattle breeds [62, 80, 85, 86,

103–105]. Whereas, the *MC1R* locus strongly selected for coat color patterns that mediates the dark (black or brown) and light (yellow or red) shades via melanogenesis in cattle [62, 80, 85, 86, 95, 103, 106, 107]. The *KIT* neighbouring (*PDGFRA* and *KDR*) genes have also been considered under selection for their role in reproduction [46, 95, 106].

The *NCAPG-LCORL* (BTA-6) and *PLAG1-CHCHD7* (BTA-14) gene regions harbour functional variant for polygenic trait of stature, which have very high heritability [51, 108, 109]. Historical evidence shows that bovine ancesters (aurochs) had very large stature, and that initial selection decreased the height during the Middle Ages [18], and then reverse selection occured after the 17th century to increase bovine stature [2]. Several variants at *NCAPG-LCORL* and *PLAG1-CHCHD7* loci have been inferred as candidates for skeletal, carcase, growth and height related traits in cattle [110–116] and other mammal species [112, 117–120]. Selection signatures have been consistently localized at the *NCAPG-LCORL* region [23, 49, 50, 60, 80, 85, 86, 95, 103, 104, 106, 121–123].

Genes underlying traits with complex selection histories. Signatures of selection for adaptation are mainly attributed to tolerance in new climates, feed resources and resistance to different disease agents in various cattle breeds. The changes in genetic aspects of behavioural control for new adaptations, from survive to thrive, and using available resources have also been detected under positive selection in several populations of African and European taurines [23, 61, 62, 85, 122]. Hence, the low gene density genomic regions harbouring genes (e.g., *R3HDMI, ZRANB3*; BTA-2) for feed efficiency have shown the impact of selective forces [23, 46, 70, 95, 121]. Additional unique regions were also found under selection for controlling several physiological functions related to tropical climatic and pathogenic adaption in African and Zebu breeds. For example, the *slick* hair coat locus on BTA-20 [124] for thermotolerance in Senepol [88], Trypanotolerance related trait in African [47, 61, 125] and, tick and heat resistance genes in Zebu [60, 126].

A strong selective sweep at a region on BTA-14 harbouring *PLAG1* gene, associated with stature, is also associated with fertility traits [31, 103]. Similarly, breeds selected for high beef production also highlight the underlying selection for fertility traits, such as gamete generation, embryo development, and spermatogenesis [70]. Such selective sweeps are results of commercialization of farm animals in the recent past by implementation of genetic improvement programs for the core production traits to achieve maximum performance and to minimize the non-productive life spans that generally relates to reproductive inefficiencies.

Candidate genes on BTA-6 (*ABCG2*, Casein cluster), BTA-7 (*SAR1B, HBEGF*), BTA-14 (*DGAT1*), BTA-16 (*AGTRAP, KIF1B*) and BTA-20 (*GHR*) have been consistently localized under strong selective sweeps in multiple dairy breeds (Fig 5). Identified selection signatures in cattle show that selective forces operated on genetic architecture controlling the physiological and anatomical structure of mammary glands and quantity and quality of various milk components. Dairy production has been attributed to the cultural revolution providing an important source of human food [11] along with influencing the human genome for selective adaptation for tolerance to lactose in response to domestication of cattle [12–14, 127]. The selective pressures on these genes vary depending upon their functional importance in cattle. For instance, the casein genes (*CSN1S1, CSN2, CSN1S2* and *CSN3*) perform major role in milk protein quality, however, only limited genetic evidence for positive selection has been presented [122, 128]. On the other hand, *ABCG2, DGAT1* and *GHR* consistently found under selection signatures and investigated in further details. The *ABCG2* has been found involved in milk yield and composition [129–132] and a differential expression found it as lactation regulator [133]. Similarly, for dairy production, the *DGAT1* [134, 135] and *GHR* [80, 106, 136] are also strong candidate genes with major effect on milk yield and composition and the regional QTLs and strong selection signatures have been found coincided in multiple cattle breeds [30, 31].

TG (BTA-14) gene ([Fig 5](#))–responsible for meat tenderness or intra-muscular fat distribution [137]–has been captured underlying the signatures of selection in multiple breeds and groups of cattle [49, 61, 69, 80, 86, 103, 126, 138–140]. Signatures of selection of some beef-specific traits—vital in shaping the beef breeds—include intra-muscular fat content [23, 80, 94, 121], muscle formation [62, 70, 86, 93–95], body composition and carcass yield [48, 80, 141, 142]. Feed efficiency (as discussed above) has been associated with intensive selective breeding for beef production and signatures of population differentiation have been detected in several breeds [23, 46, 70, 95, 121].

Genomic regions selected for multiple traits.   The color (*KIT*) and dairy (*ABCG2*) associated genes co-locate with other genes related to reproduction and bovine stature, respectively on BTA-6 [123]. In addition, there are gene-rich regions on BTA-7 (41–55 Mb) and BTA-16 (40–50 Mb) harbouring several tightly clustered MSS for multiple traits, thus provide important information about the bovine genome. In these regions, the significant signatures have been implicated for different traits and explained to arise from multiple events of selective events.

On BTA-7, the region comprise of several candidate genes between *MGAT1* and *FGF1* ([Fig 5](#)). Both of these genes implicated with reproduction traits, due to their role in fertilization and subsequent embryonic development and growth [70]. Additional genes, *PCSK4* [140] and *SPOCK1* [23, 103] perform key functions in fertility, also considered important candidates of reproduction traits. On the other hand, *SAR1B* [143] and *HBEGF* [85], underlying a selective sweep located in the centre of the region identified in high producing dairy breeds ([Fig 5](#)), and the gene functions include milk production and disease resistance. A set of additional genes, including *HSPA9*, *CD14*, *ARAP3* (*CENTD3*) and multiple members of *PCDH*, in this region are also implicated in multiple investigations for their range of functions related to immune response [46, 103, 140, 141].

On BTA-16, the *AGTRAP* gene is involved in the functioning of mammary glands and has been implicated for dairy production [85]. Similarly, at the nearby location, the *KIF1B* gene was identified under strong selection in dairy Holstein cattle [106]. At the closest flanking region, *NMNAT1* [70] and *RERE* [62] genes were localized as candidates of positive selection for embryonic growth and reproductive development. In addition, highly differentiated loci and extended haplotype homozygosity underlain *SLC25A33* and *SLC45A1* genes in the region were characterized for their important role in immunity related to tropical adaptation [60]. Similarly, selective sweeps localizing *PIK3CD* and *SPSB1* genes were also implicated for recent selection pressures that underlies immune response and immune regulation, respectively [62].

Prospects for future studies

Building on genome-wide selection scans needs to address the confounding factors related to demographic and sampling effects. Controlling for potential confounding factors in the data acquisition techniques and statistical designs are necessary for making accurate interpretations and implications. The ascertainment bias can be overcome with dense and uniform distributed SNPs available in the recently developed high-density SNP-chip assays. Utility of the several current datasets can be further extended by combining them [144]. In addition, imputation from lower to higher density SNP panels can be performed by using several approaches [145–147], as implemented by Kemper et al. [123] to increase sample size and SNP density. More recently, the genome sequencing and ultra-high density SNP panels available at affordable cost provide opportunities to perform extensive scans for genome-wide selection signatures [105, 126, 148, 149], and collection of these data will be useful for across breed investigations. Genome-wide scans of selection signatures in additional breeds of African and Zebu archetypes

are required to characterize the archetype specific genomic regions due to their unique selection history and diverse geographical background.

New selection tests and robust approaches are emerging to deal with potential challenges, hence, care should be invested while choosing applicable statistical tests with good power under the given selection models. A very recent attempt includes combing several (semi) independent estimates that increase power and resolution [50, 141, 150]. To that end, additional cattle genomic resources, e.g., availability of ancestral and derived allelic phasing from outgroups and available datasets [25, 58, 151], further enable the robust estimates. Moreover, some of the across-breed estimates of selection can also be estimated better by using a reference panel from a pool of multiple breeds to generate a set of neutral genomes. Establishment of approaches to define significance thresholds and null distribution of the statistical tests being used are required to minimize the false positives. In addition, development of selection tests to detect genomic regions underlying complex traits (such as Randhawa et al. [51]) is required for identification of classic selective sweeps in cattle.

Recent investigations of the structural variation in cattle genome suggest that selective forces, in addition to the genotypic and haplotypic patterns, operate on the copy number variation (CNV) in candidate genes and can be helpful to characterize the effects of domestication, breed formation and artificial selection [152–155]. Hence, exploration of additional genomic features, such as CNVs as alternative targets of selection, can further help elucidate the prevalence of selection in the cattle genome [64, 65, 103, 140, 155]. Nonetheless, the role of other types of selection (e.g., standing, balancing) is also warranted to quantify the role of evolutionary forces in the genetic and phenotypic diversity of cattle. In addition, comparative analyses of meta-assembly of selection signatures against QTL and GWAS database can further highlight the genomic regions under selection for traits of economic and biological importance in cattle.

The ultimate objectives of genomic scans for evolving patterns of genetic diversity are to detect causative variants and its functional relevance to particular traits. Use of the high-throughput assays for generating dense genotypes and genomic sequences will aid the fine mapping of candidate variants. Finally, functional analysis of the detected variants in the regions and genes under hotspots of positive selection will be an active area of future research to understand the biological significance of molecular variations in adaptation, appearance and production in cattle.

Supporting Information

S1 Table. A summary of published studies using different polymorphism panels and partial scans on Bovine selection signatures.
(PDF)

S2 Table. Breed-wise data information about breed type, land of breed origin, country of sampling, DNA samples, SNPchip, SNPs, bovine assembly and selection tests for each published study. Breeds shown in **bold** have been used across multiple studies.
(PDF)

S3 Table. List of cattle breeds categorized for the number of available study-wise results from genome-wide scans of selection signatures.
(PDF)

S4 Table. List of regions under selection, validated across multiple studies, based on Meta-assembly of four groups of cattle.
(XLSX)

S5 Table. List of regions under selection, validated across multiple studies, based on Meta-assembly of 16 breeds of cattle.
(XLSX)

S1 File. Summaries of S4 Table, such that; Table A: Summary of 439 validated meta-selection-score (MSS) peaks found unique within and shared across the four groups of cattle. Table B: Summary of 298 validated regions shared across multiple archetypes at 122 hotspots of selection regions.
(XLSX)

S2 File. Summaries of S5 Table, such that; Table A: Summary of 854 validated meta-selection-score (MSS) peaks found unique within and shared across the 16 breeds of cattle. Table B: Summary of 741 validated regions shared across multiple breeds (within and/or across archetypes) at 228 hotspots of selection regions.
(XLSX)

S1 Fig. Bovine population (A), milk production (B), beef production (C) and hide production (D) in various top 20 countries of the World in 2010–11 (pie charts) and population trends in the past 50 years (trend lines).
(PDF)

S2 Fig. Chromosome-wise length (million base-pairs or Mbp) of bovine genome assemblies (Btau3.1, Btau4.0, Btau4.6 and UMD3.1).
(PDF)

S3 Fig. Chromosome-wise (x-axis) SNP density (y-axis) of various genotyping Bovine SNPchip assays; **A**: 10K, **B**: BHMC (mapped on Btau 3.1 and Btau 4.0), **C**: 50K (from two version, v1 and v2) and **D**: Illumina's 800K and Affymetrix AFFXB1P (~700K but features for ~2.5 million SNPs).
(PDF)

S4 Fig. Distribution of the size (Mb) of selection signature regions of cattle published in 56 studies.
(TIFF)

S5 Fig. Meta-selection-scores (MSS) in European group, Holstein and Angus highlighting classic selective sweeps on various chromosomes. The left (A, C, E, G) and right (B, D, F, H) panels show chromosome-wise MSS using 5 Mb and 2 Mb sliding window spans, respectively. **A-B**: Chromosome 6 of European group highlighted at *ABCG2*, *LAP3*, *NCAPG* and *LCORL* gene located between 37.95–39.00 Mbp (blue bar) and at *PDGFRA* and *KIT* genes located between 71.37–71.42 Mbp (pink bar). **C-D**: Chromosome 14 of European group highlighted at *DGAT1* (1.69–1.96 Mbp; blue bar), *TG* (9.26–9.51 Mbp; green bar) and *PLAG1-CHCHD7* (25.00–25.06 Mbp; pink bar) region. **E-F**: Chromosome 16 of Angus highlighting span (44.45–45.88 Mbp) between *NMNAT1* to *RERE* genes (blue bar). **G-H**: Chromosome 20 of Holstein highlighting *GHR* region at 31.89–32.07 Mb (blue bar).
(PNG)

S6 Fig. Genome-wide distribution of selection signatures of 28 breeds (Anatolian Black, Belmont Red, Blonde dAquitaine, Finnish Ayrshire, Franken Gelbvieh, Galloway, Illyrian Mountain Busa, Japanese Black, Korean, Murnau-Werdenfelser, Normande, Pinzgauer, Polish Red, Red Angus, Salers, Wagyu, Yanbian, Afrikaner, Bonsmara, Borgou, Drakensberger, Nguni, East African Shorthorn Zebu, Guzera, Beef Master, Creole, Kenyan crossbred, Senepol), each of which have results available from a single published study. Hence,

Meta-score cannot be computed for these breeds, however, these results have been used accordingly to the particular breed-type in various group-wise Meta-assemblies of European, Zebu, African and Composite breeds.
(TIFF)

S7 Fig. Comparison of Meta-assembly and signatures of selection regions on 29 bovine autosomes (BTA) identified across 46 studies in European breeds of cattle. Top panel shows Meta-selection-scores (MSS) computed for the group of all European breeds and provides a comparison of MSS and gene density (genes per Mb) distribution on each chromosome. Lower panel shows location of published selection signature regions. Symbol and colour of each dot point represent selection test (symbols in the bottom legend) and candidate breed (European breed coloured legend on the left), respectively. In lower panel, the central or top (if given) score region is shown with a dot point and spans of extended regions are shown with a horizontal solid-line. Labels on right-hand side of lower panel shows the SNP genotyping panel used in the particular study and the red, orange and pink colours of each SNP panel represent study-wise threshold of top 0.1%, 1% and 5%, respectively.
(PDF)

Author Contributions

Conceived and designed the experiments: IASR MSK PCT HWR. Performed the experiments: IASR. Analyzed the data: IASR MSK PCT HWR. Wrote the paper: IASR MSK PCT HWR.

References

1. Zeder MA, Emshwiller E, Smith BD, Bradley DG. Documenting domestication: the intersection of genetics and archaeology. Trends Genet. 2006; 22(3):139–55. doi: 10.1016/j.tig.2006.01.007 PMID: 16458995

2. Ajmone-Marsan P, Garcia JF, Lenstra JA. On the origin of cattle: How aurochs became cattle and colonized the world. Evolutionary Anthropology: Issues, News, and Reviews. 2010; 19(4):148–57. doi: 10.1002/evan.20267

3. Andersson L, Georges M. Domestic-animal genomics: deciphering the genetics of complex traits. Nat Rev Genet. 2004; 5(3):202–12. PMID: 14970822

4. Mirkena T, Duguma G, Haile A, Tibbo M, Okeyo AM, Wurzinger M, et al. Genetics of adaptation in domestic farm animals: A review. Livestock Science. 2010; 132(1–3):1–12.

5. Bruford MW, Bradley DG, Luikart G. DNA markers reveal the complexity of livestock domestication. Nat Rev Genet. 2003; 4:900–10. PMID: 14634637

6. Beja-Pereira A, Caramelli D, Lalueza-Fox C, Vernesi C, Ferrand N, Casoli A, et al. The origin of European cattle: Evidence from modern and ancient DNA. Proc Natl Acad Sci. 2006; 103(21):8113–8. doi: 10.1073/pnas.0509210103 PMID: 16690747

7. Oleksyk TK, Smith MW, O'Brien SJ. Genome-wide scans for footprints of natural selection. Phil Trans R Soc B. 2010; 365:185–205. doi: 10.1098/rstb.2009.0219 PMID: 20008396

8. Hill WG. Applications of Population Genetics to Animal Breeding, from Wright, Fisher and Lush to Genomic Prediction. Genetics. 2014; 196(1):1–16. doi: 10.1534/genetics.112.147850 PMID: 24395822

9. MacEachern S, McEwan J, McCulloch A, Mather A, Savin K, Goddard M. Molecular evolution of the Bovini tribe (Bovidae, Bovinae): Is there evidence of rapid evolution or reduced selective constraint in domestic cattle? BMC Genomics. 2009; 10:179. doi: 10.1186/1471-2164-10-179 PMID: 19393048

10. Taberlet P, Coissac E, Pansu J, Pompanon F. Conservation genetics of cattle, sheep, and goats. C R Biol. 2011; 334(3):247–54.

11. Capuco AV, Akers RM. The origin and evolution of lactation. J Biol. 2009; 8(37).

12. Tellam RL, Lemay DG, Tassell CPV, Lewin HA, Worley KC, Elsik CG. Unlocking the bovine genome. BMC Genomics. 2009; 10:193. doi: 10.1186/1471-2164-10-193 PMID: 19393070

13. Lemay DG, Rijnkels M, German JB. Lessons from the bovine genome: Implications for human nutrition and research. The Journal of Nutrition. 2009; 139(7):1271–2. doi: 10.3945/jn.109.107656 PMID: 19420348

14. Lemay DG, Lynn DJ, Martin WF, Neville MC, Casey TM, Rincon G, et al. The bovine lactation genome: insights into the evolution of mammalian milk. Genome Biology. 2009; 10(4):R43. doi: 10.1186/gb-2009-10-4-r43 PMID: 19393040

15. Laland KN, Odling-Smee J, Myles S. How culture shaped the human genome: bringing genetics and the human sciences together. Nat Rev Genet. 2010; 11(2):137–48. doi: 10.1038/nrg2734 PMID: 20084086

16. Wiener P, Wilkinson S. Deciphering the genetic basis of animal domestication. Proceedings of the Royal Society B: Biological Sciences. 2011; 278(1722):3161–70. doi: 10.1098/rspb.2011.1376 PMID: 21885467

17. Bradley DG, Loftus RT, Cunningham P, Machugh DE. Genetics and domestic cattle origins. Evolutionary Anthropology. 1998; 6(3):79–86.

18. Zeder MA. Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. Proc Natl Acad Sci. 2008; 105(33):11597–604. doi: 10.1073/pnas.0801317105 PMID: 18697943

19. Diamond J. Evolution, consequences and future of plant and animal domestication. Nature. 2002; 418 (6898):700–7. PMID: 12167878

20. Edwards CJ, Ginja C, Kantanen J, Pérez-Pardal L, Tresset A, Stock F, et al. Dual origins of dairy cattle farming—evidence from a comprehensive survey of European Y-chromosomal variation. PLoS ONE. 2011; 6(1):e15922. doi: 10.1371/journal.pone.0015922 PMID: 21253012

21. Loftus RT, MacHugh DE, Bradley DG, Sharp PM, Cunningham P. Evidence for two independent domestications of cattle. Proc Natl Acad Sci. 1994; 91(7):2757–61. PMID: 8146187

22. Chen S, Lin B-Z, Baig M, Mitra B, Lopes RJ, Santos AM, et al. Zebu cattle are an exclusive legacy of the South Asia Neolithic. Mol Biol Evol. 2010; 27(1):1–6. doi: 10.1093/molbev/msp213 PMID: 19770222

23. Gibbs RA, Taylor JF, Van Tassell CP, Barendse W, Eversole KA, Gill CA, et al. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. Science. 2009; 324:528–32. doi: 10.1126/science.1167936 PMID: 19390050

24. Gautier M, Laloë D, Moazami-Goudarzi K. Insights into the genetic history of French cattle from dense SNP data on 47 Worldwide breeds. PLoS ONE. 2010; 5(9):e13038. doi: 10.1371/journal.pone.0013038 PMID: 20927341

25. Decker JE, Pires JC, Conant GC, McKay SD, Heaton MP, Chen K, et al. Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. Proc Natl Acad Sci. 2009; 106 (44):18644–9.

26. Bollongino R, Burger J, Powell A, Mashkour M, Vigne J-D, Thomas MG. Modern Taurine cattle descended from small number of Near-Eastern founders. Mol Biol Evol. 2012; 29(9):2101–4. doi: 10.1093/molbev/mss092 PMID: 22422765

27. Stock F, Gifford-Gonzalez D. Genetics and African cattle domestication. Afr Archaeol Rev. 2013; 30 (1):51–72. doi: 10.1007/s10437-013-9131-6

28. Porto-Neto LR, Lee SH, Sonstegard TS, Van Tassell CP, Lee HK, Gibson JP, et al. Genome-wide detection of signatures of selection in Korean Hanwoo cattle. Anim Genet. 2014; 45(2):180–90. doi: 10.1111/age.12119 PMID: 24494817

29. Burt DW. The cattle genome reveals its secrets. J Biol. 2009; 8:36. doi: 10.1186/jbiol137 PMID: 19439025

30. Khatkar M, Thomson P, Tammen I, Raadsma H. Quantitative trait loci mapping in dairy cattle: review and meta-analysis. Genet Sel Evol. 2004; 36(2):163–90. PMID: 15040897

31. Khatkar MS, Randhawa IAS, Raadsma HW. Meta-assembly of genomic regions and variants associated with female reproductive efficiency in cattle. Livestock Science. 2014; 166:144–57. http://dx.doi.org/10.1016/j.livsci.2014.05.015.

32. Goddard ME, Hayes BJ. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat Rev Genet. 2009; 10:381–91. doi: 10.1038/nrg2575 PMID: 19448663

33. Biswas S, Akey JM. Genomic insights into positive selection. Trends Genet. 2006; 22(8):437–46. http://dx.doi.org/10.1016/j.tig.2006.06.005. PMID: 16808986

34. FAO. The state of the world's animal genetics resources for food and agriculture. edited by Barbara Rischkowsky and Dafydd Pilling. Rome 2007. p. ftp://ftp.fao.org/docrep/fao/010/a1250e/ae.pdf

35. Gutiérrez-Gil B, Arranz JJ, Wiener P. An interpretive review of selective sweep studies in Bos taurus cattle populations: identification of unique and shared selection signals across breeds. Frontiers in genetics. 2015; 6. doi: 10.3389/fgene.2015.00167

36. Nei M, Suzuki Y, Nozawa M. The neutral theory of molecular evolution in the genomic era. Annual Review of Genomics and Human Genetics. 2010; 11:265–89. doi: 10.1146/annurev-genom-082908-150129 PMID: 20565254

37. Lohmueller KE. The impact of population demography and selection on the genetic architecture of complex traits. PLoS Genet. 2014; 10(5):e1004379. doi: 10.1371/journal.pgen.1004379 PMID: 24875776

38. MacLeod IM, Hayes BJ, Goddard ME. The effects of demography and long term selection on the accuracy of genomic prediction with sequence data. Genetics. 2014; 198(4):1671–84. doi: 10.1534/genetics.114.168344 PMID: 25233989

39. Thornton KR, Jensen JD, Becquet C, Andolfatto P. Progress and prospects in mapping recent selection in the genome. Heredity. 2007; 98(6):340–8. PMID: 17473869

40. Fagny M, Patin E, Enard D, Barreiro LB, Quintana-Murci L, Laval G. Exploring the occurrence of classic selective sweeps in humans using whole-genome sequencing datasets. Mol Biol Evol. 2014; 31 (7):1850–68. doi: 10.1093/molbev/msu118 PMID: 24694833

41. Granka JM, Henn BM, Gignoux CR, Kidd JM, Bustamante CD, Feldman MW. Limited evidence for classic selective sweeps in African populations. Genetics. 2012; 192(3):1049–64. doi: 10.1534/genetics.112.144071 PMID: 22960214

42. Kelley JL, Madeoy J, Calhoun JC, Swanson W, Akey JM. Genomic signatures of positive selection in humans and the limits of outlier approaches. Genome Res. 2006; 16(8):980–9. doi: 10.1101/gr.5157306 PMID: 16825663

43. Crisci JL, Poh Y-P, Bean A, Simkin A, Jensen JD. Recent progress in polymorphism-based population genetic inference. J Hered. 2012; 103(2):287–96. doi: 10.1093/jhered/esr128 PMID: 22246406

44. McTavish EJ, Decker JE, Schnabel RD, Taylor JF, Hillis DM. New World cattle show ancestry from multiple independent domestication events. Proc Natl Acad Sci. 2013; 110(15):E1398–E406. doi: 10.1073/pnas.1303367110 PMID: 23530234

45. Hanotte O, Tawah CL, Bradley DG, Okomo M, Verjee Y, Ochieng J, et al. Geographic distribution and frequency of a taurine Bos taurus and an indicine Bos indicus Y specific allele amongst sub-Saharan African cattle breeds. Mol Ecol. 2000; 9(4):387–96. doi: 10.1046/j.1365-294x.2000.00858.x PMID: 10736042

46. Gautier M, Naves M. Footprints of selection in the ancestral admixture of a New World Creole cattle breed. Mol Ecol. 2011; 20(15):3128–43. doi: 10.1111/j.1365-294X.2011.05163.x PMID: 21689193

47. Dayo G-K, Thevenon S, Berthier D, Moazami-Goudarzi K, Denis C, Cuny G, et al. Detection of selection signatures within candidate regions underlying trypanotolerance in outbred cattle populations. Mol Ecol. 2009; 18:1801–13. doi: 10.1111/j.1365-294X.2009.04141.x PMID: 19302350

48. MacEachern S, Hayes B, McEwan J, Goddard M. An examination of positive selection and changing effective population size in Angus and Holstein cattle populations (Bos taurus) using a high density SNP genotyping platform and the contribution of ancient polymorphism to genomic diversity in Domestic cattle. BMC Genomics. 2009; 10:181. doi: 10.1186/1471-2164-10-181 PMID: 19393053

49. Hayes BJ, Chamberlain AJ, Maceachern S, Savin K, McPartlan H, MacLeod I, et al. A genome map of divergent artificial selection between Bos taurus dairy cattle and Bos taurus beef cattle. Anim Genet. 2009; 40(2):176–84. doi: 10.1111/j.1365-2052.2008.01815.x PMID: 19067671

50. Randhawa IAS, Khatkar MS, Thomson PC, Raadsma HW. Composite selection signals can localize the trait specific genomic regions in multi-breed populations of cattle and sheep. BMC Genet. 2014; 15(1):34.

51. Randhawa IAS, Khatkar MS, Thomson PC, Raadsma HW. Composite selection signals for complex traits exemplified through bovine stature using multibreed cohorts of European and African Bos taurus. G3: Genes|Genomes|Genetics. 2015; 5(7):1391–401. doi: 10.1534/g3.115.017772 PMID: 25931611

52. Nielsen R. Population genetic analysis of ascertained SNP data. Human Genomics. 2004; 1(3):218–24. PMID: 15588481

53. Nicholas FW. Discovery, validation and delivery of DNA markers. Australian Journal of Experimental Agriculture. 2006; 46:155–8.

54. Ayers KL, Sabatti C, Lange K. A dictionary model for haplotyping, genotype calling, and association testing. Genet Epidemiol. 2007; 31:672–83. doi: 10.1002/gepi.20232 PMID: 17487885

55. Achaz G. Testing for neutrality in samples with sequencing errors. Genetics. 2008; 179:1409–24. doi: 10.1534/genetics.107.082198 PMID: 18562660

56. Boitard S, Schlotterer C, Futschik A. Detecting selective sweeps: a new approach based on Hidden Markov models. Genetics. 2009; 181:1567–78. doi: 10.1534/genetics.108.100032 PMID: 19204373

57. Nielsen R, Signorovitch J. Correcting for ascertainment biases when analyzing SNP data: applications to the estimation of linkage disequilibrium. Theor Popul Biol. 2003; 63(3):245–55. PMID: 12689795

58. Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, et al. Development and characterization of a high density SNP genotyping assay for cattle. PLoS ONE. 2009; 4(4):e5350. doi: 10.1371/journal.pone.0005350 PMID: 19390634

59. Qanbari S, Simianer H. Mapping signatures of positive selection in the genome of livestock. Livestock Science. 2014; 166:133–43. http://dx.doi.org/10.1016/j.livsci.2014.05.003.

60. Chan EKF, Nagaraj SH, Reverter A. The evolution of tropical adaptation: comparing taurine and zebu cattle. Anim Genet. 2010; 41(5):467–77. doi: 10.1111/j.1365-2052.2010.02053.x PMID: 20477791

61. Gautier M, Flori L, Riebler A, Jaffrezic F, Laloe D, Gut I, et al. A whole genome Bayesian scan for adaptive genetic divergence in West African cattle. BMC Genomics. 2009; 10(1):550.

62. Ramey H, Decker J, McKay S, Rolf M, Schnabel R, Taylor J. Detection of selective sweeps in cattle using genome-wide SNP data. BMC Genomics. 2013; 14(1):382.

63. Heslot N, Rutkoski J, Poland J, Jannink J-L, Sorrells ME. Impact of Marker Ascertainment Bias on Genomic Selection Accuracy and Estimates of Genetic Diversity. PLoS ONE. 2013; 8(9):e74612. doi: 10.1371/journal.pone.0074612 PMID: 24040295

64. Kelley J, Swanson W. Positive selection in the human genome: from genome scans to biological significance. Annual Review of Genomics and Human Genetics. 2008; 9:143–60. doi: 10.1146/annurev.genom.9.081307.164411 PMID: 18505377

65. Vitti JJ, Grossman SR, Sabeti PC. Detecting natural selection in genomic data. Annu Rev Genet. 2013; 47(1):97–120. doi: 10.1146/annurev-genet-111212-133526

66. Hurst LD. Fundamental concepts in genetics: Genetics and the understanding of selection. Nat Rev Genet. 2009; 10(2):83–93. doi: 10.1038/nrg2506 PMID: 19119264

67. Hohenlohe PA, Phillips PC, Cresko WA. Using population genomics to detect selection in natural populations: Key concepts and methodological considerations. Int J Plant Sci. 2010; 171(9):1059–71. doi: 10.1086/656306 PMID: 21218185

68. Novembre J, Di Rienzo A. Spatial patterns of variation due to natural selection in humans. Nat Rev Genet. 2009; 10(11):745–55. doi: 10.1038/nrg2632 PMID: 19823195

69. Fan H, Wu Y, Qi X, Zhang J, Li J, Gao X, et al. Genome-wide detection of selective signatures in Simmental cattle. J Appl Genet. 2014; 55:343–51. Epub 2014/03/13. doi: 10.1007/s13353-014-0200-6 PMID: 24619661

70. Qanbari S, Gianola D, Hayes B, Schenkel F, Miller S, Moore S, et al. Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. BMC Genomics. 2011; 12(1):318.

71. Wray NR, Yang J, Hayes BJ, Price AL, Goddard ME, Visscher PM. Pitfalls of predicting complex traits from SNPs. Nat Rev Genet. 2013; 14(7):507–15. doi: 10.1038/nrg3457 PMID: 23774735

72. Pickrell J, Coop G, Novembre J, Kudaravalli S, Li J, Absher D, et al. Signals of recent positive selection in a worldwide sample of human populations. Genome Res. 2009; 19:826–37. doi: 10.1101/gr.087577.108 PMID: 19307593

73. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. PLoS Biol. 2006; 4(3):e72. doi: 10.1371/journal.pbio.0040072 PMID: 16494531

74. Willing E-M, Dreyer C, van Oosterhout C. Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. PLoS ONE. 2012; 7(8):e42649. doi: 10.1371/journal.pone.0042649 PMID: 22905157

75. Bhatia G, Patterson NJ, Sankararaman S, Price AL. Estimating and interpreting Fst: the impact of rare variants. Genome Res. 2013; 23(9):1514–21. doi: 10.1101/gr.154831.113 PMID: 23861382

76. Zeggini E, Rayner W, Morris AP, Hattersley AT, Walker M, Hitman GA, et al. An evaluation of HapMap sample size and tagging SNP performance in large-scale empirical and simulated data sets. Nat Genet. 2005; 37(12):1320–2. doi: 10.1038/ng1670 PMID: 16258542

77. Hancock AM, Di Rienzo A. Detecting the genetic signature of natural selection in human populations: Models, methods, and data. Annual Review of Anthropology. 2008; 37:197–217. PMID: 20622977

78. Nosil P, Feder JL. Genomic divergence during speciation: causes and consequences. Phil Trans R Soc B. 2012; 367(1587):332–42. doi: 10.1098/rstb.2011.0263 PMID: 22201163

79. O'Reilly PF, Birney E, Balding DJ. Confounding between recombination and selection, and the Ped/Pop method for detecting selection. Genome Res. 2008; 18(8):1304–13. doi: 10.1101/gr.067181.107 PMID: 18617692

80. Rothammer S, Seichter D, Forster M, Medugorac I. A genome-wide scan for signatures of differential artificial selection in ten cattle breeds. BMC Genomics. 2013; 14(1):908.

81. Medugorac I, Seichter D, Graf A, Russ I, Blum H, Göpel KH, et al. Bovine polledness—an autosomal dominant trait with allelic heterogeneity. PLoS ONE. 2012; 7(6):e39477. doi: 10.1371/journal.pone.0039477 PMID: 22737241

82. Allais-Bonnet A, Grohs C, Medugorac I, Krebs S, Djari A, Graf A, et al. Novel insights into the bovine polled phenotype and horn ontogenesis in Bovidae. PLoS ONE. 2013; 8(5):e63512. doi: 10.1371/journal.pone.0063512 PMID: 23717440

83. Seichter D, Russ I, Rothammer S, Eder J, Förster M, Medugorac I. SNP-based association mapping of the polled gene in divergent cattle breeds. Anim Genet. 2012; 43(5):595–8. doi: 10.1111/j.1365-2052.2011.02302.x PMID: 22497248

84. Mariasegaram M, Harrison BE, Bolton JA, Tier B, Henshall JM, Barendse W, et al. Fine-mapping the POLL locus in Brahman cattle yields the diagnostic marker CSAFG29. Anim Genet. 2012; 43(6):683–8. doi: 10.1111/j.1365-2052.2012.02336.x PMID: 22497221

85. Stella A, Ajmone-Marsan P, Lazzari B, Boettcher P. Identification of selection signatures in cattle breeds selected for dairy production. Genetics. 2010; 185:1451–61. doi: 10.1534/genetics.110.116111 PMID: 20479146

86. Druet T, Pérez-Pardal L, Charlier C, Gautier M. Identification of large selective sweeps associated with major genes in cattle. Anim Genet. 2013; 44(6):758–62. doi: 10.1111/age.12073 PMID: 23859468

87. Li M-H, Iso-Touru T, Lauren H, Kantanen J. A microsatellite-based analysis for the detection of selection on BTA1 and BTA20 in northern Eurasian cattle (Bos taurus) populations. Genet Sel Evol. 2010; 42(1):32.

88. Flori L, Gonzatti MI, Thevenon S, Chantal I, Pinto J, Berthier D, et al. A quasi-exclusive European ancestry in the Senepol tropical cattle breed highlights the importance of the slick locus in tropical adaptation. PLoS ONE. 2012; 7(5):e36133. doi: 10.1371/journal.pone.0036133 PMID: 22675421

89. Brenneman RA, Davis SK, Sanders JO, Burns BM, Wheeler TC, Turner JW, et al. The polled locus maps to BTA1 in a Bos indicus × Bos taurus cross. J Hered. 1996; 87(2):156–61. PMID: 8830095

90. Marchitelli C, Savarese M, Crisà A, Nardone A, Marsan P, Valentini A. Double muscling in Marchigiana beef breed is caused by a stop codon in the third exon of myostatin gene. Mamm Genome. 2003; 14(6):392–5. doi: 10.1007/s00335-002-2176-5 PMID: 12879361

91. Georges M. When less means more: Impact of myostatin in animal breeding. Immun, Endoc & Metab Agents in Med Chem. 2010; 10(4):240–8.

92. Stinckens A, Georges M, Buys N. Mutations in the Myostatin gene leading to hypermuscularity in mammals: indications for a similar mechanism in fish? Anim Genet. 2011; 42(3):229–34. doi: 10.1111/j.1365-2052.2010.02144.x PMID: 21175702

93. Wiener P, Gutierrez-Gil B. Assessment of selection mapping near the myostatin gene (GDF-8) in cattle. Anim Genet. 2009; 40:598–608. doi: 10.1111/j.1365-2052.2009.01886.x PMID: 19456316

94. Boitard S, Rocha D. Detection of signatures of selective sweeps in the Blonde d'Aquitaine cattle breed. Anim Genet. 2013; 44(5):579–83. doi: 10.1111/age.12042 PMID: 23647053

95. Pintus E, Sorbolini S, Albera A, Gaspa G, Dimauro C, Steri R, et al. Use of locally weighted scatterplot smoothing (LOWESS) regression to study selection signatures in Piedmontese and Italian Brown cattle breeds. Anim Genet. 2014; 45(1):1–11. doi: 10.1111/age.12076 PMID: 23889699

96. Guastella AM, Sorbolini S, Zuccaro A, Pintus E, Bordonaro S, Marletta D, et al. Melanocortin 1 receptor (MC1R) gene polymorphisms in three Italian cattle breeds. Animal Production Science. 2011; 51(11):1039–43. http://dx.doi.org/10.1071/AN11070.

97. Schmutz SM, Dreger DL. Interaction of MC1R and PMEL alleles on solid coat colors in Highland cattle. Anim Genet. 2013; 44(1):9–13. doi: 10.1111/j.1365-2052.2012.02361.x PMID: 22524257

98. Hulsman Hanna L, Sanders J, Riley D, Abbey C, Gill C. Identification of a major locus interacting with MC1R and modifying black coat color in an F2 Nellore-Angus population. Genet Sel Evol. 2014; 46(1):4.

99. Wong AK, Ruhe AL, Robertson KR, Loew ER, Williams DC, Neff MW. A de novo mutation in KIT causes white spotting in a subpopulation of German Shepherd dogs. Anim Genet. 2012; 44(3):305–10. doi: 10.1111/age.12006 PMID: 23134432

100. Fang M, Larson G, Soares Ribeiro H, Li N, Andersson L. Contrasting Mode of Evolution at a Coat Color Locus in Wild and Domestic Pigs. PLoS Genet. 2009; 5(1):e1000341.

101. Haase B, Signer-Hasler H, Binns MM, Obexer-Ruff G, Hauswirth R, Bellone RR, et al. Accumulating Mutations in Series of Haplotypes at the KIT and MITF Loci Are Major Determinants of White Markings in Franches-Montagnes Horses. PLoS ONE. 2013; 8(9):e75071. doi: 10.1371/journal.pone.0075071 PMID: 24098679

102. Hauswirth R, Jude R, Haase B, Bellone RR, Archer S, Holl H, et al. Novel variants in the KIT and PAX3 genes in horses with white-spotted coat colour phenotypes. Anim Genet. 2013; 44(6):763–5. doi: 10.1111/age.12057 PMID: 23659293

103. Porto-Neto L, Sonstegard T, Liu G, Bickhart D, Da Silva M, Machado M, et al. Genomic divergence of zebu and taurine cattle identified through high-density SNP genotyping. BMC Genomics. 2013; 14 (1):876.

104. Mancini G, Gargani M, Chillemi G, Nicolazzi E, Marsan PA, Valentini A, et al. Signatures of selection in five Italian cattle breeds detected by a 54K SNP panel. Mol Biol Rep. 2014; 41(2):957–65. doi: 10. 1007/s11033-013-2940-5 PMID: 24442315

105. Qanbari S, Pausch H, Jansen S, Somel M, Strom TM, Fries R, et al. Classic selective sweeps revealed by massive sequencing in cattle. PLoS Genet. 2014; 10(2):e1004148. doi: 10.1371/journal. pgen.1004148 PMID: 24586189

106. Flori L, Fritz S, Jaffrézic F, Boussaha M, Gut I, Heath S, et al. The genome response to artificial selection: A case study in dairy cattle. PLoS ONE. 2009; 4(8):e6595. doi: 10.1371/journal.pone.0006595 PMID: 19672461

107. Liu L, Harris B, Keehan M, Zhang Y. Genome scan of pigmentation traits in Friesian-Jersey crossbred cattle. J Genet Genomics. 2009; 36(11):661–6. doi: 10.1016/s1673-8527(08)60158-7 PMID: 19932462

108. Kemper KE, Goddard ME. Understanding and predicting complex traits: knowledge from cattle. Hum Mol Genet. 2012; 21(R1):R45–51. doi: 10.1093/hmg/dds332 PMID: 22899652

109. Lanktree MB, Guo Y, Murtaza M, Glessner JT, Bailey SD, Onland-Moret NC, et al. Meta-analysis of dense genecentric association studies reveals common and uncommon variants associated with height. Am J Hum Genet. 2011; 88(1):6–18. doi: 10.1016/j.ajhg.2010.11.007 PMID: 21194676

110. Lindholm-Perry A, Sexten A, Kuehn L, Smith T, King D, Shackelford S, et al. Association, effects and validation of polymorphisms within the NCAPG—LCORL locus located on BTA6 with feed intake, gain, meat and carcass traits in beef cattle. BMC Genet. 2011; 12(1):103.

111. Pryce JE, Hayes BJ, Bolormaa S, Goddard ME. Polymorphic regions affecting human height also control stature in cattle. Genetics. 2011; 187(3):981–4. doi: 10.1534/genetics.110.123943 PMID: 21212230

112. Hoshiba H, Setoguchi K, Watanabe T, Kinoshita A, Mizoshita K, Sugimoto Y, et al. Comparison of the effects explained by variations in the bovine PLAG1 and NCAPG genes on daily body weight gain, linear skeletal measurements and carcass traits in Japanese Black steers from a progeny testing program. Animal Science Journal. 2013; 84(7):529–34. doi: 10.1111/asj.12033 PMID: 23607392

113. Nishimura S, Watanabe T, Mizoshita K, Tatsuda K, Fujita T, Watanabe N, et al. Genome-wide association study identified three major QTL for carcass weight including the PLAG1-CHCHD7 QTN for stature in Japanese Black cattle. BMC Genet. 2012; 13(1):40.

114. Littlejohn M, Grala T, Sanders K, Walker C, Waghorn G, Macdonald K, et al. Genetic variation in PLAG1 associates with early life body weight and peripubertal weight and growth in Bos taurus. Anim Genet. 2012; 43(5):591–4. Epub 2012/04/14. doi: 10.1111/j.1365-2052.2011.02293.x PMID: 22497486

115. Karim L, Takeda H, Lin L, Druet T, Arias JAC, Baurain D, et al. Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature. Nat Genet. 2011; 43(5):405–13. doi: 10.1038/ng.814 PMID: 21516082

116. Fortes MRS, Kemper K, Sasazaki S, Reverter A, Pryce JE, Barendse W, et al. Evidence for pleiotropism and recent selection in the PLAG1 region in Australian Beef cattle. Anim Genet. 2013; 44(6):636–47. Epub 5 August 2013. doi: 10.1111/age.12075 PMID: 23909810

117. Tetens J, Widmann P, Kühn C, Thaller G. A genome-wide association study indicates LCORL/ NCAPG as a candidate locus for withers height in German Warmblood horses. Anim Genet. 2013; 44 (4):467–71. doi: 10.1111/age.12031 PMID: 23418885

118. Signer-Hasler H, Flury C, Haase B, Burger D, Simianer H, Leeb T, et al. A genome-wide association study reveals loci influencing height and other conformation traits in horses. PLoS ONE. 2012; 7(5): e37282. doi: 10.1371/journal.pone.0037282 PMID: 22615965

119. Rubin C-J, Megens H-J, Barrio AM, Maqbool K, Sayyab S, Schwochow D, et al. Strong signatures of selection in the domestic pig genome. Proc Natl Acad Sci. 2012; 109(48):19529–36. doi: 10.1073/ pnas.1217149109 PMID: 23151514

120. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature. 2010; 467(7317):832–8. doi: 10.1038/nature09410 PMID: 20881960

121. Barendse W, Harrison BE, Bunch RJ, Thomas MB, Turner LB. Genome wide signatures of positive selection: The comparison of independent samples and the identification of regions associated to traits. BMC Genomics. 2009; 10:178. doi: 10.1186/1471-2164-10-178 PMID: 19393047

122. Qanbari S, Pimentel ECG, Tetens J, Thaller G, Lichtner P, Sharifi AR, et al. A genome-wide scan for signatures of recent selection in Holstein cattle. Anim Genet. 2010; 41(4):377–89. doi: 10.1111/j.1365-2052.2009.02016.x PMID: 20096028

123. Kemper K, Saxton S, Bolormaa S, Hayes B, Goddard M. Selection for complex traits leaves little or no classic signatures of selection. BMC Genomics. 2014; 15(1):246.

124. Littlejohn MD, Henty KM, Tiplady K, Johnson T, Harland C, Lopdell T, et al. Functionally reciprocal mutations of the prolactin signalling pathway define hairy and slick cattle. Nature communications. 2014; 5:5861. Epub 2014/12/19. doi: 10.1038/ncomms6861 PMID: 25519203

125. Noyes H, Brass A, Obara I, Anderson S, Archibald AL, Bradley DG, et al. Genetic and expression analysis of cattle identifies candidate genes in pathways responding to Trypanosoma congolense infection. Proc Natl Acad Sci. 2011; 108(22):9304–9. doi: 10.1073/pnas.1013486108 PMID: 21593421

126. Liao X, Peng F, Forni S, McLaren D, Plastow G, Stothard P. Whole genome sequencing of Gir cattle for identifying polymorphisms and loci under selection. Genome. 2013; 56(10):592–8. doi: 10.1139/gen-2013-0082 PMID: 24237340

127. Tishkoff S, Reed F, Ranciaro A, Voight B, Babbitt C, Silverman J, et al. Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet. 2007; 39(1):31–40. PMID: 17159977

128. Lee H-J, Kim J, Lee T, Son JK, Yoon H-B, Baek K-S, et al. Deciphering the genetic blueprint behind Holstein milk proteins and production. Genome Biol Evol. 2014; 6:1366–74. doi: 10.1093/gbe/evu102 PMID: 24920005

129. Braunschweig MH. Mutations in the bovine ABCG2 and the ovine MSTN gene added to the few quantitative trait nucleotides identified in farm animals: a mini-review. J Appl Genet. 2010; 51(3):289–97. doi: 10.1007/bf03208858 PMID: 20720303

130. Cohen-Zinder M, Seroussi Eyal, Larkin Denis M., Loor Juan J., Everts-van der Wind Annelie, Lee Jun-Heon, et al. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. Genome Res. 2005; 15:936–44. PMID: 15998908

131. Lillehammer M, Goddard ME, Nilsen H, Sehested E, Olsen HG, Lien S, et al. Quantitative trait locus-by-environment interaction for milk yield traits on Bos taurus autosome 6. Genet Mol Biol. 2008; 179:1539–46.

132. Olsen H, Nilsen H, Hayes B, Berg P, Svendsen M, Lien S, et al. Genetic support for a quantitative trait nucleotide in the ABCG2 gene affecting milk composition of dairy cattle. BMC Genet. 2007; 8:32. PMID: 17584938

133. Sheehy PA, Riley LG, Raadsma HW, Williamson P, Wynn PC. A functional genomics approach to evaluate candidate genes located in a QTL interval for milk production traits on BTA6. Anim Genet. 2009; 40:492–8. doi: 10.1111/j.1365-2052.2009.01862.x PMID: 19392827

134. Schennink A, Stoop W, Visker M, Heck J, Bovenhuis H, van der Poel J, et al. DGAT1 underlies large genetic variation in milk-fat composition of dairy cows. Anim Genet. 2007; 38(5):467–73. PMID: 17894561

135. Bernard G, Coppieters W, Farnir F, Karim L, Ford C, Berzi P, et al. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. Genome Res. 2002; 12:222–31. PMID: 11827942

136. Banos G, Woolliams JA, Woodward BW, Forbes AB, Coffey MP. Impact of single nucleotide polymorphisms in leptin, leptin receptor, growth hormone receptor, and diacylglycerol acyltransferase (DGAT1) gene loci on milk production, feed, and body energy traits of UK dairy cows. J Dairy Sci. 2008; 91:3190–200. doi: 10.3168/jds.2007-0930 PMID: 18650297

137. Bennett GL, Shackelford SD, Wheeler TL, King DA, Casas E, Smith TPL. Selection for genetic markers in beef cattle reveals complex associations of thyroglobulin and casein1-S1 with carcass and meat traits. J Anim Sci. 2012; 91(2):565–71. doi: 10.2527/jas.2012-5454 PMID: 23148258

138. Edea Z, Dadi H, Kim SW, Park JH, Shin GH, Dessie T, et al. Linkage disequilibrium and genomic scan to detect selective loci in cattle populations adapted to different ecological conditions in Ethiopia. J Anim Breed Genet. 2014; 131(5):358–66. doi: 10.1111/jbg.12083 PMID: 24602159

139. Kim E-S, Cole JB, Huson H, Wiggans GR, Van Tassell CP, Crooker BA, et al. Effect of Artificial Selection on Runs of Homozygosity in U.S. Holstein Cattle. PLoS ONE. 2013; 8(11):e80813. doi: 10.1371/journal.pone.0080813 PMID: 24348915

140. Perez OBrien A, Utsunomiya Y, Meszaros G, Bickhart D, Liu G, Van Tassell C, et al. Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. Genet Sel Evol. 2014; 46(1):19.

141. Utsunomiya YT, Pérez O'Brien AM, Sonstegard TS, Van Tassell CP, do Carmo AS, Mészáros G, et al. Detecting loci under recent positive selection in dairy and beef cattle by combining different genome-wide scan methods. PLoS ONE. 2013; 8(5):e64280. doi: 10.1371/journal.pone.0064280 PMID: 23696874

142. Hosokawa D, Ishii A, Yamaji K, Sasazaki S, Oyama K, Mannen H. Identification of divergently selected regions between Japanese Black and Holstein cattle using bovine 50k SNP array. Animal Science Journal. 2012; 83(1):7–13. doi: 10.1111/j.1740-0929.2011.00927.x PMID: 22250733

143. Larkin DM, Daetwyler HD, Hernandez AG, Wright CL, Hetrick LA, Boucek L, et al. Whole-genome resequencing of two elite sires for the detection of haplotypes under selection in dairy cattle. Proc Natl Acad Sci. 2012; 109(20):7693–8. doi: 10.1073/pnas.1114546109 PMID: 22529356

144. Nicolazzi E, Picciolini M, Strozzi F, Schnabel R, Lawley C, Pirani A, et al. SNPchiMp: a database to disentangle the SNPchip jungle in bovine livestock. BMC Genomics. 2014; 15(1):123.

145. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am J Hum Genet. 2009; 84(2):210–23. doi: 10.1016/j.ajhg.2009.01.005 PMID: 19200528

146. Khatkar M, Moser G, Hayes B, Raadsma H. Strategies and utility of imputed SNP genotypes for genomic analysis in dairy cattle. BMC Genomics. 2012; 13(1):538.

147. Daetwyler HD, Wiggans GR, Hayes BJ, Woolliams JA, Goddard ME. Imputation of missing genotypes from sparse to high density using long-range phasing. Genetics. 2011; 189(1):317–27. doi: 10.1534/genetics.111.128082 PMID: 21705746

148. Daetwyler HD, Capitan A, Pausch H, Stothard P, van Binsbergen R, Brondum RF, et al. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nat Genet. 2014; 46(8):858–65. doi: 10.1038/ng.3034 PMID: 25017103

149. Lee T, Cho S, Seo KS, Chang J, Kim H, Yoon D. Genetic variants and signatures of selective sweep of Hanwoo population (Korean native cattle). BMB Reports. 2013; 46(7):346–51.

150. Grossman SR, Shylakhter I, Karlsson EK, Byrne EH, Morales S, Frieden G, et al. A composite of multiple signals distinguishes causal variants in regions of positive selection. Science. 2010; 327:883–6. doi: 10.1126/science.1183863 PMID: 20056855

151. Rocha D, Billerey C, Samson F, Boichard D, Boussaha M. Identification of the putative ancestral allele of bovine single-nucleotide polymorphisms. J Anim Breed Genet. 2014; 131(6):483–6. doi: 10.1111/jbg.12095 PMID: 24862839

152. Liu GE, Hou Y, Zhu B, Cardone MF, Jiang L, Cellamare A, et al. Analysis of copy number variations among diverse cattle breeds. Genome Res. 2010; 20(5):693–703. doi: 10.1101/gr.105403.110 PMID: 20212021

153. Hou Y, Liu G, Bickhart D, Cardone M, Wang K, Kim E-s, et al. Genomic characteristics of cattle copy number variations. BMC Genomics. 2011; 12(1):127.

154. Fadista J, Thomsen B, Holm L, Bendixen C. Copy number variation in the bovine genome. BMC Genomics. 2010; 11(1):284.

155. Choi J-W, Liao X, Stothard P, Chung W-H, Jeon H-J, Miller SP, et al. Whole-genome analyses of Korean native and Holstein cattle breeds by massively parallel sequencing. PLoS ONE. 2014; 9(7):e101127. doi: 10.1371/journal.pone.0101127 PMID: 24992012

156. Glick G, Shirak A, Uliel S, Zeron Y, Ezra E, Seroussi E, et al. Signatures of contemporary selection in the Israeli Holstein dairy cattle. Anim Genet. 2012; 43 Suppl 1:45–55. doi: 10.1111/j.1365-2052.2012.02348.x PMID: 22742502

157. Schwarzenbacher H, Dolezal M, Flisikowski K, Seefried F, Wurmser C, Schlotterer C, et al. Combining evidence of selection with association analysis increases power to detect regions influencing complex traits in dairy cattle. BMC Genomics. 2012; 13(1):48.

158. Pan D, Zhang S, Jiang J, Jiang L, Zhang Q, Liu J. Genome-wide detection of selective signature in Chinese Holstein. PLoS ONE. 2013; 8(3):e60440. doi: 10.1371/journal.pone.0060440 PMID: 23555972

159. Lim D, Gondro C, Park HS, Cho YM, Chai HH, Seong HH, et al. Identification of recently selected mutations driven by artificial selection in Hanwoo (Korean cattle). Asian-Australasian Journal of

Animal Sciences. 2013; 26(5):603–8. Epub 24 October 2012. http://dx.doi.org/10.5713/ajas.2012. 12456. doi: 10.5713/ajas.2012.12456 PMID: 25049829

160. Kim E-S, Rothschild MF. Genomic adaptation of admixed dairy cattle in East Africa. Frontiers in genetics. 2014; 5. doi: 10.3389/fgene.2014.00443

161. Ryu J, Lee C. Identification of contemporary selection signatures using composite log likelihood and their associations with marbling score in Korean cattle. Anim Genet. 2014; 45(6):765–70. doi: 10. 1111/age.12209 PMID: 25179770

162. Somavilla AL, Sonstegard TS, Higa RH, Rosa AN, Siqueira F, Silva LOC, et al. A genome-wide scan for selection signatures in Nellore cattle. Anim Genet. 2014; 45(6):771–81. doi: 10.1111/age.12210 PMID: 25183526

163. Xu L, Bickhart DM, Cole JB, Schroeder SG, Song J, Tassell CPV, et al. Genomic signatures reveal new evidences for selection of important traits in domestic cattle. Mol Biol Evol. 2014; 32(3):711–25. doi: 10.1093/molbev/msu333 PMID: 25431480

164. Bomba L, Nicolazzi E, Milanesi M, Negrini R, Mancini G, Biscarini F, et al. Relative extended haplotype homozygosity signals across breeds reveal dairy and beef specific signatures of selection. Genet Sel Evol. 2015; 47(1):25.

165. Gurgul A, Pawlina K, Frys-Żurek M, Bugno-Poniewierska M. Identification of differential selection traces in two Polish cattle breeds. Animal Science Journal. 2015; 86(1):17–24. doi: 10.1111/asj. 12242 PMID: 25124517

166. Kim E-S, Sonstegard T, Rothschild M. Recent artificial selection in U.S. Jersey cattle impacts autozygosity levels of specific genomic regions. BMC Genomics. 2015; 16(302). doi: 10.1186/s12864-015-1500-x

167. Flori L, Thevenon S, Dayo G-K, Senou M, Sylla S, Berthier D, et al. Adaptive admixture in the West African bovine hybrid zone: insight from the Borgou population. Mol Ecol. 2014; 23(13):3241–57. doi: 10.1111/mec.12816 PMID: 24888437

168. Bahbahani H, Clifford H, Wragg D, Mbole-Kariuki MN, Van Tassell C, Sonstegard T, et al. Signatures of positive selection in East African Shorthorn Zebu: A genome-wide single nucleotide polymorphism analysis. Scientific Reports. 2015; 5. doi: 10.1038/srep11729

169. Choi JW, Choi BH, Lee SH, Lee SS, Kim HC, Yu D, et al. Whole-Genome Resequencing Analysis of Hanwoo and Yanbian Cattle to Identify Genome-Wide SNPs and Signatures of Selection. Molecules and Cells. 2015; 38(5):466–73. Epub 2015/05/29. doi: 10.14348/molcells.2015.0019 PMID: 26018558

170. Gurgul A, Szmatoła T, Ropka-Molik K, Jasielczuk I, Pawlina K, Semik E, et al. Identification of genome-wide selection signatures in the Limousin beef cattle breed. J Anim Breed Genet. 2015:n/a–n/a. doi: 10.1111/jbg.12196

171. Kasarda R, Moravčíková N, Trakovická A, Mészáros G, Kadlečík O. Genome-wide selection signatures in Pinzgau cattle. Potravinarstvo. 2015; 9(1):268–74.

172. Kim E-S, Sonstegard TS, da Silva MV, Gasbarre LC, Van Tassell CP. Genome-wide scan of gastrointestinal nematode resistance in closed Angus population selected for minimized influence of MHC. PLoS One. 2015; 10(3):e0119380. Epub 2015/03/25. doi: 10.1371/journal.pone.0119380 PMID: 25803687

173. Li Y, Kim J-J. Effective population size and signatures of selection using bovine 50K SNP chips in Korean native cattle (Hanwoo). Evolutionary Bioinformatics. 2015; 11:143–53. doi: 10.4137/EBO. S24359

174. Makina S, Muchadeyi F, van Marle-Koster E, Taylor J, Makgahlela M, Maiwashe A. Genome-wide scan for selection signatures in six cattle breeds in South Africa. Genet Sel Evol. 2015; 47(1):92.

175. Sorbolini S, Marras G, Gaspa G, Dimauro C, Cellesi M, Valentini A, et al. Detection of selection signatures in Piemontese and Marchigiana cattle, two breeds with similar production aptitudes but different selection histories. Genet Sel Evol. 2015; 47(1):52.

176. Zhao F, McParland S, Kearney F, Du L, Berry D. Detection of selection signatures in dairy and beef cattle using high-density genomic information. Genet Sel Evol. 2015; 47(1):49.

177. Fay JC, Wu C-I. Hitchhiking under positive Darwinian selection. Genetics. 2000; 155:1405–13. PMID: 10880498

178. Freeman AR, Lynn DJ, Murray C, Bradley DG. Detecting the effects of selection at the population level in six bovine immune genes. BMC Genet. 2008; 9:62. doi: 10.1186/1471-2156-9-62 PMID: 18838007

179. Akey JM, Zhang G, Zhang K, Jin L, Shriver MD. Interrogating a high-density SNP map for signatures of natural selection. Genome Res. 2002; 12(12):1805–14. doi: 10.1101/gr.631202 PMID: 12466284

180. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution. 1984; 38(6):1358–70.

181. Nicholson G, Smith AV, Jónsson F, Gústafsson Ó, Stefánsson K, Donnelly P. Assessing population differentiation and isolation from single-nucleotide polymorphism data. Journal of the Royal Statistical Society Series B (Statistical Methodology). 2002; 64(4):695–715.

182. Weir BS, Hill WG. Estimating F-statistics. Annu Rev Genet. 2002; 36:721–50.

183. Akey JM, Ruhe AL, Akey DT, Wong AK, Connelly CF, Madeoy J, et al. Tracking footprints of artificial selection in the dog genome. Proc Natl Acad Sci. 2010; 107(3):1160–5. doi: 10.1073/pnas.0909918107 PMID: 20080661

184. Riebler A, Held L, Stephan W. Bayesian variable selection for detecting adaptive genomic differences among populations. Genetics. 2008; 178(3):1817–29. doi: 10.1534/genetics.107.081281 PMID: 18245358

185. Beaumont MA, Balding DJ. Identifying adaptive genetic divergence among populations from genome scans. Mol Ecol. 2004; 13(4):969–80. PMID: 15012769

186. Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C. Genomic scans for selective sweeps using SNP data. Genome Res. 2005; 15:1566–75. doi: 10.1101/gr.4252305 PMID: 16251466

187. Rubin C-J, Zody MC, Eriksson J, Meadows JRS, Sherwood E, Webster MT, et al. Whole-genome resequencing reveals loci under selection during chicken domestication. Nature. 2010; 464 (7288):587–91. doi: 10.1038/nature08832 PMID: 20220755

188. Teo YY, Fry AE, Bhattacharya K, Small KS, Kwiatkowski DP, Clark TG. Genome-wide comparisons of variation in linkage disequilibrium. Genome Res. 2009; 19(10):1849–60. doi: 10.1101/gr.092189.109 PMID: 19541915

189. Ong RT-H, Teo Y-Y. varLD: a program for quantifying variation in linkage disequilibrium patterns between populations. Bioinformatics. 2010; 26(9):1269–70. doi: 10.1093/bioinformatics/btq125 PMID: 20308177

190. Depaulis F, Veuille M. Neutrality tests based on the distribution of haplotypes under an infinite-site model. Mol Biol Evol. 1998; 15(12):1788–90. Epub 1999/01/23. PMID: 9917213

191. Sabeti PC, Reich DE, Higgins JM, Levine HZP, Richter DJ, Schaffner SF, et al. Detecting recent positive selection in the human genome from haplotype structure. Nature. 2002; 419(6909):832–7. PMID: 12397357

192. Mueller JC, Andreoli C. Plotting haplotype-specific linkage disequilibrium patterns by extended haplotype homozygosity. Bioinformatics. 2004; 20(5):786–7. doi: 10.1093/bioinformatics/btg481 PMID: 14764566

193. Prasad A, Schnabel RD, McKay SD, Murdoch B, Stothard P, Kolbehdari D, et al. Linkage disequilibrium and signatures of selection on chromosomes 19 and 29 in beef and dairy cattle. Anim Genet. 2008; 39:597–605. doi: 10.1111/j.1365-2052.2008.01772.x PMID: 18717667

194. Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, et al. Genome-wide detection and characterization of positive selection in human populations. Nature. 2007; 449:913–8. PMID: 17943131

195. Hayes BJ, Lien S, Nilsen H, Olsen HG, Berg P, Maceachern S, et al. The origin of selection signatures on bovine chromosome 6. Anim Genet. 2008; 39:105–11. doi: 10.1111/j.1365-2052.2007.01683.x PMID: 18241301

196. Tang K, Thornton K, Stoneking M. A new approach for using genome scans to detect recent positive selection in the human genome. PLoS Biol. 2007; 5(7):e171. PMID: 17579516