

Supplementary Materials and Methods to

The evolution of tropical adaptation: Comparing taurine and zebu cattle

Eva KF Chan¹, Shivashankar H Nagaraj and Antonio Reverter²

Cooperative Research Centre for Beef Genetic Technologies

CSIRO Livestock Industries, Queensland Bioscience Precinct

306 Carmody Rd., St. Lucia, Brisbane, Queensland 4067, Australia

Supplementary Methods

Sensitivity analysis to determine if sample size and/or SNP-discovery breed affects % monomorphism

Brahman, the breed with the largest sample size (70), consistently showed low levels of monomorphism to both Brahman and taurine SNP. To test whether Brahman is indeed more variable (more polymorphic at both Brahman and taurine SNP) or whether our observation was an artefact of imbalanced sample size, a sensitivity analysis was performed. Simulated datasets were generated by re-sampling, with replacement, from the 70 Brahman animals to make up 12 sample sizes, corresponding to the 12 existing sample sizes (between 4 and 70).

Results (Supplementary Figure 2) suggest the extent of monomorphism, from both real and simulated Brahman-only samples, are dependent on sample size. That is, a smaller sample size of Brahman showed comparable amount of monomorphic SNP to another breed of equal size. This is confirmed with a similar analysis using Holstein animals: 12 simulated samples were generated from re-sampling of the 52 Holstein animals (Supplementary Figure 3). However, in both cases, it is clear that Brahman more variant: Brahman animals for more polymorphic for Taurine SNP while Holstein animals are less polymorphic for Brahman SNP.

More STRUCTURE Analyses

In addition to using the 7,821 autosomal SNP, various subsets of SNP were examined for their ability to predict the population substructure of our cattle sample. We used the sets of 91 private SNP, the 14 SNP representing the 14 compound diplotypes, and 12 SNP representing the 12 regions of positive selection to re-estimate ancestral populations of the 13 breeds for $K = 2$ (Supplemental Table 5 and Supplementary Figure 7). Using private SNP and compound diplotype SNP, the two clusters corresponded to the two breed types, with Brahmans having $> 97\%$ Zebu ancestry on average and each of the 10 Taurine breeds having, on average, $> 93\%$ Taurine Taurine ancestry using either subset of SNP. The major difference is in the estimates for the two composite breeds. Compared to estimates using all autosomal SNP, estimates of Zebu ancestry for the two composite breeds had larger standard deviations (> 0.12)

using either of these subsets. Conversely, the 12 EHH SNP performed much more poorly in distinguishing the two breed types.

To determine whether the 14 compound and 91 private diplotype SNP are genuinely effective in distinguishing between Zebu and Taurine cattle, we repeated the analysis with five sets of 91 and five sets of 14 randomly selected SNP (Supplementary Table 5). Our new estimate of Zebu and Taurine ancestries using 91 random subsets of SNP were comparable to that using all or 91 private SNP. Conversely, our estimates using only 14 random SNP were poor, suggesting while we were able to distinguish between the two breed types with only 91 random SNP, power was reduced with only 14 random SNP.

We also repeated the analysis using, separately, the set of 78 SNP private in Brahman (fixed in Taurine animals) and the set of 13 SNP private in Taurine animals (fixed in Brahman). These two analyses clearly demonstrate the nature of using private SNP for estimating cattle ancestry: prediction is highly accurate for the breed type for which the SNP are fixed, while the prediction accuracy of the alternate breed type for which the SNP are polymorphic is dependent on the number of SNP used. This is evident from results obtained using the set of 13 SNP that are polymorphic in all Taurine breeds but monomorphic in Brahman and a random subset of 13 SNP that are private in Brahman (Supplementary Table 2). Using the 13 SNP private in Brahman predicted 100% Taurine ancestry for all Taurine samples because these 13 SNP are uninformative in these animals. Similarly, any set of 13 SNP private in Taurine predicted 100% Zebu ancestry for all Brahman animals; the small proportion of Taurine ancestry predicted for Brahman animals was due to missing data at these SNP for these animals. Conversely, prediction accuracy of the alternate breed type using monomorphic SNP is affected by the number of SNP. Prediction variability of Taurine ancestry for the Taurine breeds were similar using either the set of 13 SNP private in Taurine animals or a random subset of 13 SNP. The absence of expected increase in prediction error for Zebu animals using a subset of 13 SNP private for Brahman is likely due to it being the only representative Zebu in this study.

The negative correlation between prediction variability and the number of SNP used in the STRUCTURE analyses is further evident in the variable estimates of Zebu

proportions for Belmont Red and Santa Gertrudis using different numbers of randomly chosen SNP. Despite this, it is clear that prediction error is lower with Zebu-Taurine distinguishing SNP compared to using the 12 EHH-representative SNP or random SNP sets of the same size. Correlations of predicted Zebu proportions (Supplementary Table 5) for Belmont Red animals between using all autosomal SNP and either the subset of 91 private SNP (0.70) or the 14 compound diplotyperepresentative SNP (0.54) were higher than those for Santa Gertrudis (0.64 for private SNP and 0.42 for diplotyperepresentative SNP). This is similarly true for the subset of 78 Brahman private SNP or any random SNP subsets, but not for the subset of 13 Taurine private SNP. This is not surprising as these 13 SNP are uninformative in Zebu and as Belmont Red animals are Taurine-Sanga composites with minimal to no Brahman ancestry (Burrow & Corbet 2000) we would expect poorer predictive ability for these animals.

Supplementary Tables

Supplementary Table 1 F_{ST} estimated for between all 13 breeds, between taurine and zebu animals, and between taurine breeds using all SNPs, Brahman (BRM) SNPs, taurine SNPs (include HOL, ANG, & LMS SNPs), Holstein (HOL) SNPs, Angus (ANG) SNPs, Limousine (LMS) SNPs, and SNPs from the IBISS. These SNP sets may include or exclude SNPs that are not mapped onto Btau4.0 assembly or are X-linked.

	Between 13 breeds	Between taurine and zebu	Between taurine breeds
All SNPs	18.2	22.8	12.1
- (ChrX + ChrU)	18.2	22.8	12.2
BRM SNPs	30.4	50.6	9.5
- (ChrX + ChrU)	30.9	51.4	9.6
Taurine SNPs	17.9	22.1	12.2
- (ChrX + ChrU)	18.0	22.0	12.3
HOL SNPs	17.9	22.1	12.2
- (ChrX + ChrU)	18.0	22.1	12.2
ANG SNPs	17.8	21.9	12.2
- (ChrX + ChrU)	17.8	21.7	11.9
LMS SNPs	18.8	24.6	11.8
- (ChrX + ChrU)	18.0	22.4	12.1
cSNPs	18.3	22.9	12.0
- (ChrX + ChrU)	18.4	23.1	12.3

Supplementary Table 2 SNPs that are private in either Brahman (zebu) or taurine breeds.

Chromosome and coordinates correspond to the Btau4.0 assembly. A SNP is private in zebu (Z) if it has $MAF \geq 0.05$ in the Brahman animals and monomorphic in all taurine animals and it is private in taurine (T) if it has $MAF \geq 0.05$ in all taurine animals and monomorphic in Brahman animals. * in the last two columns indicate whether the SNP is polymorphic in Belmont Red (BEL) or Santa Gertrudis (SGT). BREED: SNP origin (BRM: Brahman, HOL: Holstein, ANG: Angus, LMS: Limousin); molecular type: location of the SNP {genomic, cDNA}; -: information unknown.

Chrom: coordinate	Allele	Breed	Molecular type	Private	Poly mor phic in BEL	Poly mor phic in SGT
1:96,826,638	G/T	HOL	genomic	T	*	*
1:106,968,102	C/T	HOL	genomic	T	*	*
2:16,862,013	A/T	BRM	genomic	Z		
2:49,761,603	C/T	BRM	genomic	Z	*	*
2:109,759,965	NA	-	genomic	Z		*
3:84,301,109	C/G	BRM	genomic	Z		*
3:91,336,440	C/T	HOL	genomic	Z		*
3:103,727,108	NA	-	genomic	Z		*
3:114,618,417	A/G	HOL	genomic	Z	*	*
3:114,996,176	NA	-	genomic	Z		
4:115,438,525	NA	-	genomic	Z		
4:115,438,679	G/T	BRM	genomic	Z		
5:57,066,312	C/T	ANG	genomic	T	*	
5:67,210,397	C/T	BRM	genomic	Z		*
5:85,279,449	A/G	LIM	genomic	T		*
5:91,723,622	C/G	HOL	genomic	T	*	*
5:111,253,936	A/G	BRM	genomic	Z	*	*
5:111,253,953	G/T	BRM	genomic	Z	*	*
5:115,318,910	A/G	BRM	genomic	Z	*	*
5:120,717,784	A/C	BRM	genomic	Z	*	*
5:120,717,804	G/T	BRM	genomic	Z	*	*
5:120,718,079	A/G	BRM	genomic	Z	*	*
6:32,154,206	A/G	HOL	genomic	Z		*
6:32,154,207	G/T	HOL	genomic	Z		*
6:35,955,387	A/G	BRM	genomic	Z	*	*
6:83,996,498	A/C	BRM	genomic	Z		
6:105,500,001	C/T	BRM	genomic	Z		*
6:105,500,454	C/T	BRM	genomic	Z	*	*
6:108,535,543	NA	-	genomic	Z		*
8:17,016,098	A/G	BRM	genomic	Z	*	*
8:17,016,171	C/T	BRM	genomic	Z	*	*
8:17,016,654	C/T	BRM	genomic	Z	*	*
8:32,446,240	C/T	BRM	genomic	Z	*	*
8:45,507,466	C/T	LIM	genomic	T	*	*
8:52,301,026	A/G	HOL	genomic	Z	*	*
8:99,635,592	C/G	BRM	genomic	Z		*

8:99,635,923	A/G	BRM	genomic	Z		*
8:99,635,999	A/G	BRM	genomic	Z	*	*
9:53,993,831	NA	-	genomic	Z	*	*
10:4,751,255	A/G	-	cDNA	Z	*	*
10:14,420,918	A/G	HOL	genomic	Z	*	*
10:35,991,172	C/T	-	cDNA	T	*	
10:53,139,971	C/T	BRM	genomic	Z	*	*
10:75,750,792	A/G	BRM	genomic	Z	*	*
10:92,605,269	A/T	BRM	genomic	Z		*
11:86,777,504	A/G	BRM	genomic	Z	*	*
11:86,777,914	A/G	BRM	genomic	Z	*	*
12:30,221,998	C/T	HOL	genomic	Z	*	
13:9,045,489	C/T	BRM	genomic	Z		
13:9,045,586	A/C	BRM	genomic	Z	*	
13:17,829,789	C/T	BRM	genomic	Z		
13:22,163,907	C/T	BRM	genomic	Z		*
13:38,510,345	A/T	BRM	genomic	Z		
14:21,007,395	A/T	ANG	genomic	Z		
14:38,035,923	C/T	HOL	genomic	Z	*	*
14:42,269,664	C/T	-	genomic	T	*	*
14:42,855,060	A/G	HOL	genomic	T	*	*
15:65,064,596	G/T	HOL	genomic	Z	*	*
15:72,078,351	G/T	HOL	genomic	Z		
15:72,078,353	C/T	HOL	genomic	Z		
16:30,575,585	C/T	BRM	genomic	Z	*	*
16:30,575,950	A/T	BRM	genomic	Z	*	*
16:31,068,223	C/T	HOL	genomic	Z	*	*
16:63,286,665	A/T	BRM	genomic	Z		
16:63,286,751	C/T	BRM	genomic	Z	*	*
17:11,293,757	C/T	BRM	genomic	Z	*	*
17:64,123,168	C/G	BRM	genomic	Z		*
18:45,430,001	G/T	HOL	genomic	Z		
19:53,525,613	A/G	BRM	genomic	Z	*	*
19:53,525,789	A/G	BRM	genomic	Z		*
20:2,739,448	A/G	BRM	genomic	Z	*	*
20:56,713,121	G/T	HOL	genomic	T		*
20:68,501,198	A/C	BRM	genomic	Z	*	*
21:15,643,298	C/T	HOL	genomic	T		
21:15,643,800	A/G	HOL	genomic	T		
21:59,555,547	C/T	BRM	genomic	Z	*	
22:21,357,235	C/T	BRM	genomic	Z	*	*
22:21,357,550	A/T	BRM	genomic	Z	*	*
22:40,139,571	C/T	BRM	genomic	Z	*	*
24:4,329,833	A/G	BRM	genomic	Z		
24:16,734,115	A/G	BRM	genomic	Z	*	*
24:16,734,137	C/T	BRM	genomic	Z	*	*
24:16,734,180	C/T	BRM	genomic	Z	*	*
24:16,734,288	C/T	BRM	genomic	Z	*	*
24:42,951,913	NA	-	genomic	Z	*	
26:43,774,969	C/T	HOL	genomic	Z		*

29:15,794,032	C/T	BRM	genomic	Z	*	*
29:17,109,628	A/G	ANG	genomic	Z	*	*
X:1,231,131	G/T	BRM	genomic	Z	*	
X:2,124,730	A/G	BRM	genomic	Z		
-	C/T	ANG	genomic	T	*	*

Supplementary Table 3 SNPs within the 14 compound diplotypes. The representative SNP for the diplotype is the SNP within a diplotype with the largest |z|, indicated by **. Also see Supplementary Table 2 for legend. The Genes column list genes or transcripts annotated to the corresponding regions.

Chromosome: coordinate	Allele	Breed	Molecular type	max z	Genes
1:89,677,287	A/G	HOL	genomic		
1:89,938,724	A/C	ANG	genomic		
1:90,233,258	A/G	HOL	genomic		
1:91,730,436	A/G	HOL	genomic		
1:92,010,304	C/T	HOL	genomic		
1:92,010,421	C/G	HOL	genomic		
1:92,319,188	C/G	HOL	genomic		<i>FUNDC1 GNB4</i>
1:92,433,978	A/C	HOL	genomic		<i>KCNMB3</i>
1:92,434,645	A/C	BRM	genomic		<i>LOC515180</i>
1:92,669,977	A/G	ANG	genomic		<i>LOC539609</i>
1:92,823,720	A/G	HOL	genomic		<i>LOC616200</i>
1:94,369,684	A/C	LMS	genomic		<i>LOC784483</i>
1:94,369,865	C/T	LMS	genomic		<i>NAALADL2</i>
1:94,370,042	C/G	LMS	genomic		<i>PIK3CA</i>
1:94,470,191	A/G	HOL	genomic		<i>TBL1XR1</i>
1:94,697,347	A/G	HOL	genomic	**	<i>ZMAT3 ZNF639</i>
1:94,958,245	C/T	HOL	genomic		
1:95,379,006	A/T	HOL	genomic		
1:95,456,257	G/T	ANG	genomic		
1:95,456,444	A/G	ANG	genomic		
1:95,456,482	A/G	ANG	genomic		
1:95,456,527	C/T	ANG	genomic		
3:98,789,043	A/G	HOL	genomic		
3:99,078,080	C/T	HOL	genomic		
3:99,078,499	C/T	HOL	genomic		<i>BTF3L4 C3H1orf123</i>
3:99,380,438	A/G	HOL	genomic		<i>C3H1orf185 CC2D1B</i>
3:99,467,830	A/G	ANG	genomic		<i>CDKN2C CPT2 DIO1</i>
3:100,811,876	G/T	HOL	genomic		<i>DMRTA2 DMRTB1</i>
3:102,131,007	C/T	-	cDNA		<i>ECHDC2 ELAVL4</i>
3:102,812,360	C/T	LMS	genomic		<i>EPS15 FAF1 GLIS1</i>
3:102,812,458	C/T	LMS	genomic		<i>GPX7 KTI12 LDLRAD1</i>
3:102,962,479	C/T	LMS	genomic		<i>LOC100140166</i>
3:102,992,977	C/T	HOL	genomic		<i>LOC520518</i>
3:103,040,223	A/C	HOL	genomic		<i>LOC523160</i>
3:103,365,892	C/T	ANG	genomic		<i>LOC527172</i>
3:103,727,108	-	-	genomic		<i>LOC535329</i>
3:103,727,190	A/G	HOL	genomic		<i>LOC613487</i>
3:103,727,204	C/G	HOL	genomic		<i>LOC616319</i>
3:104,179,493	A/C	HOL	genomic		<i>LOC781064</i>
3:104,310,886	A/C	HOL	genomic		<i>LOC783467</i>
3:104,508,445	A/G	ANG	genomic		<i>LOC787024</i>
3:104,551,311	G/T	HOL	genomic		<i>LOC787775</i>
3:104,551,588	A/G	HOL	genomic	**	<i>LOC790544 LRP8</i>
3:104,551,720	A/G	HOL	genomic		<i>LRRC42 MAGOH</i>
3:104,564,433	C/T	HOL	genomic		<i>NDUFA12 NRD1</i>
					<i>ORC1L OSBP19 PODN</i>
					<i>PRPF38A RAB3B</i>
					<i>RNF11 SCP2 SLC1A7</i>
					<i>TMEM48 TXNDC12</i>
					<i>YIPF1 ZCCHC11</i>
					<i>ZFYVE9 ZYG11A</i>
					<i>ZYG11B</i>

Chromosome: coordinate	Allele	Breed	Molecular type	max z	Genes
3:104,616,099	C/T	LMS	genomic		
4:47,395,192	A/T	HOL	genomic		<i>ASZ1 ATXN7L1</i>
4:47,395,262	A/G	HOL	genomic		<i>BCAP29</i>
4:47,788,346	A/C	HOL	genomic		<i>C4H7orf53</i>
4:49,066,548	C/G	HOL	genomic		<i>CAPZA2 CAV1</i>
4:49,066,591	A/G	HOL	genomic		<i>CAV2 CBLL1</i>
4:49,419,706	C/T	ANG	genomic		<i>CFTR COG5 DLD</i>
4:51,115,663	A/G	HOL	genomic		<i>DUS4L EPDR1</i>
4:51,115,714	A/G	HOL	genomic		<i>FOXP2 GPR22</i>
4:51,115,738	C/G	HOL	genomic		<i>GPR85 HBP1</i>
4:52,437,363	A/G	HOL	genomic		<i>IFRD1 LAMB1</i>
4:52,897,172	A/C	HOL	genomic		<i>LHFPL3</i>
4:52,897,683	A/G	HOL	genomic		<i>LOC100140059</i>
4:53,718,968	A/G	HOL	genomic		<i>LOC100140447</i>
4:53,972,254	C/T	ANG	genomic		<i>LOC530341</i>
4:54,568,493	A/T	ANG	genomic	**	<i>LOC534227</i>
4:54,727,243	A/G	HOL	genomic		<i>LOC538658</i>
4:54,727,354	A/G	HOL	genomic		<i>LOC615671</i>
4:54,727,783	C/T	HOL	genomic		<i>LOC781154</i>
4:54,775,262	A/G	ANG	genomic		<i>LOC784327</i>
4:54,777,102	A/T	ANG	genomic		<i>LOC784535</i>
4:56,448,725	A/G	HOL	genomic		<i>LOC788400</i>
4:58,473,086	C/G	HOL	genomic		<i>LOC788400</i>
4:58,473,419	C/T	HOL	genomic		<i>LOC789240</i>
4:58,491,657	A/G	HOL	genomic		<i>LRRN3 MDFIC</i>
4:58,545,752	A/G	HOL	genomic		<i>MET MLL5</i>
4:58,779,526	A/G	HOL	genomic		<i>NRCAM PBEF1</i>
4:58,992,524	C/T	HOL	genomic		<i>PIK3CG PNPLA8</i>
4:59,252,723	A/G	HOL	genomic		<i>PPP1R3A</i>
4:59,793,285	C/T	HOL	genomic		<i>PRKAR2B PUS7</i>
4:59,933,196	A/C	HOL	genomic		<i>RINT1 SFRP4</i>
5:64,474,202	C/T	HOL	genomic		<i>SLC26A3 SLC26A4</i>
5:64,565,964	C/T	HOL	genomic		<i>SRPK2 ST7</i>
5:64,856,924	C/T	ANG	genomic		<i>STARD3NL SYPL1</i>
5:65,375,214	C/T	ANG	genomic		<i>TES TFEC</i>
5:65,557,646	-	-	genomic	**	<i>TMEM168</i>
5:65,979,754	C/T	HOL	genomic		<i>TXNDC3 WNT2</i>
5:66,753,942	C/T	ANG	genomic		<i>ZNF277</i>
5:66,843,802	A/G	ANG	genomic		<i>ACTR6 AMDHD1</i>
5:67,210,397	C/T	BRM	genomic		<i>ANKS1B ANO4 ARL1</i>
5:67,335,731	A/G	HOL	genomic		<i>ASCL1 CCDC53</i>
5:67,338,758	C/T	HOL	genomic		<i>CD38 CHPT1 DRAM</i>
5:67,338,899	A/G	HOL	genomic		<i>ELK3 GAS2L3</i>
5:67,700,904	A/C	ANG	genomic		<i>GLT8D2 GNPTAB</i>
5:67,701,052	C/T	ANG	genomic		<i>HAL HCFC2 IGFI</i>
5:68,957,930	C/T	ANG	genomic		<i>IKIP LOC100138094</i>
5:69,087,624	G/T	ANG	genomic		<i>LOC100139879</i>
					<i>LOC100140289</i>
					<i>LOC100190887</i>
					<i>LOC510487</i>
					<i>LOC513218</i>
					<i>LOC516896</i>
					<i>LOC537782</i>
					<i>LOC615510</i>
					<i>LOC615576</i>
					<i>LOC781331</i>
					<i>LOC781730</i>
					<i>LOC781832</i>
					<i>LOC782673</i>

Chromosome: coordinate	Allele	Breed	Molecular type	max z	Genes
5:69,384,361	A/G	HOL	genomic		<i>LOC785006</i>
5:69,384,405	G/T	HOL	genomic		<i>LOC787705</i>
5:69,845,095	A/G	HOL	genomic		
5:70,277,350	A/C	HOL	genomic		<i>LOC789016</i>
5:70,280,997	C/T	HOL	genomic		<i>LOC790861 LTA4H</i>
5:70,361,669	C/T	HOL	genomic		<i>MYBPC1 NEDD1</i>
5:70,361,669	C/T	HOL	genomic		<i>NFYB NR1H4</i>
5:71,150,007	C/T	HOL	genomic		<i>NT5DC3 NTN4 PAH</i>
5:71,427,362	C/T	HOL	genomic		<i>PCTK2 PMCH</i>
5:71,545,001	A/T	HOL	genomic		<i>SCYL2 SLC17A8</i>
5:72,588,270	A/G	HOL	genomic		<i>SLC25A3 SLC5A8</i>
5:72,588,406	C/T	HOL	genomic		<i>SPIC STAB2 SYCP3</i>
5:72,940,885	A/G	HOL	genomic		<i>TDG TMPO TRAI</i> <i>UHRF1BP1L hare</i>
5:110,824,986	C/T	-	genomic		<i>ADIPOR2 AKAP3 ANKRD54</i>
5:111,253,936	A/G	BRM	genomic		<i>APOBEC3A APOBEC3B ATF4</i>
5:111,253,953	G/T	BRM	genomic		<i>ATP6V1E1 BAIAP2L2 BCL2L13</i>
5:112,160,884	C/G	HOL	genomic		<i>BID CSH12ORF4 CSH12ORF5</i>
5:112,340,738	G/T	-	cDNA		<i>CSH12orf32 CSH22ORF23</i>
5:113,734,415	C/T	HOL	genomic		<i>CACNA1C CACNA1I CBX6</i>
5:114,425,930	C/T	HOL	genomic		<i>CBX7 CBY1 CCDC77 CCND2</i>
5:114,618,071	A/G	-	cDNA		<i>CD9 CDC42EP1 CECRS5</i>
5:114,748,498	A/T	HOL	genomic		<i>CSNK1E DCP1B DDX17 DMC1</i>
5:114,748,594	A/G	HOL	genomic		<i>DNAJC19 DNAL4 EIF3L</i>
5:114,827,047	C/G	HOL	genomic		<i>ENTHD1 ERC1 FBXL14 FGF23</i>
5:114,827,201	A/G	HOL	genomic		<i>FGF6 FKBP4 FOXM1 GALNT8</i>
5:115,318,910	A/G	BRM	genomic	**	<i>GALR3 GCAT GGA1 GRAP2</i>
5:115,319,293	C/T	BRM	genomic		<i>GTPBP1 H1FO ILI7RA ITFG2</i>
5:115,900,315	A/G	HOL	genomic		<i>JARID1A JOSD1 KCNA1</i>
5:115,900,606	A/G	HOL	genomic		<i>KCNA5 KCNA6 KDELR3</i>
5:115,900,760	C/T	HOL	genomic		<i>LGALS1 LGALS3</i>
5:116,347,366	A/G	ANG	genomic		<i>LOC100138973 LOC100139283</i>
5:117,866,771	C/T	ANG	genomic		<i>LOC100141275 LOC506293</i>
5:118,645,507	A/G	HOL	genomic		<i>LOC512529 LOC517202</i>
5:118,682,287	G/T	HOL	genomic		<i>LOC525377 LOC527026</i>
6:33,989,255	A/G	HOL	genomic		<i>LOC531276 LOC532244</i>
6:34,388,909	C/G	HOL	genomic		<i>LOC533602 LOC535353</i>
6:34,389,038	C/T	HOL	genomic		<i>LOC538476 LOC540132</i>
6:34,389,115	A/G	HOL	genomic		<i>LOC613306 LOC617805</i>
6:34,389,417	C/T	HOL	genomic		<i>LOC782076 LOC782484</i>
6:34,391,551	C/T	HOL	genomic		<i>LOC784075 LOC785671</i>
6:35,462,604	A/G	HOL	genomic		<i>LOC788054 LOC788643</i>
6:35,462,919	G/T	HOL	genomic		<i>LOC788738 MAFF MAP3K7IP1</i>
6:35,815,786	C/G	HOL	genomic		<i>MGAT3 MICALL1 NDUFA9</i>
6:35,955,387	A/G	BRM	genomic	**	<i>NINJ2 NOL12 NRIP2 NTF3</i>
6:35,955,585	C/T	BRM	genomic		<i>PARP11 PDGFB PDXP PEX26</i>
6:35,955,623	C/T	BRM	genomic		<i>PICK1 PLA2G6 POLR2F</i>
6:36,518,779	A/G	HOL	genomic		<i>PRMT8 RAD51AP1 RAD52</i>
					<i>RPL27A RPS19BP1 Rpl3</i>
					<i>SLC16A8 SLC6A12 SLC6A13</i>
					<i>SOX10 SYNGR1 TEAD4</i>
					<i>TMEM184B TNFRSF1A</i>
					<i>TMM22 TRIBP TSPAN11</i>
					<i>TUBA8 TULP3 UNC84B VWF</i> <i>WNK1 WNT5B</i>
					<i>ABCG2</i>
					<i>FAM13A1</i>
					<i>FARSB HERC3</i>
					<i>HERC5 HERC6</i>
					<i>LOC509118</i>
					<i>LOC616908</i>
					<i>LOC781421</i>
					<i>MMRN1</i>
				**	<i>NAP1L5 PIGY</i>
					<i>PKD2 PPM1K</i>
					<i>SNCA SPP1</i>

Chromosome: coordinate	Allele	Breed	Molecular type	max z	Genes
6:36,786,886	A/G	ANG	genomic		
6:36,787,039	C/T	ANG	genomic		
6:36,978,345	A/G	HOL	genomic		
6:37,432,832	A/G	HOL	genomic		
6:37,433,107	A/G	HOL	genomic		
6:37,433,382	A/G	HOL	genomic		
6:37,608,440	A/G	ANG	genomic		
6:37,608,843	A/G	ANG	genomic		
6:107,912,894	C/T	HOL	genomic		
6:108,219,700	A/G	HOL	genomic		
6:108,535,543	-	-	genomic		
6:108,554,269	C/T	HOL	genomic		
6:109,596,223	A/G	HOL	genomic		
6:109,619,480	G/T	HOL	genomic		
6:110,775,912	A/G	HOL	genomic		<i>BST1 C1QTNF7</i>
6:111,541,598	C/T	HOL	genomic		<i>CC2D2A CD38</i>
6:111,703,787	A/C	HOL	genomic		<i>CPEB2 CYTL1</i>
6:112,708,393	A/G	HOL	genomic		<i>FBXL5 HS3ST1</i>
6:112,708,414	C/T	HOL	genomic		<i>LOC281812</i>
6:112,940,209	A/G	HOL	genomic		<i>LOC508527</i>
6:113,022,661	C/T	ANG	genomic		<i>LOC614799</i>
6:113,135,180	A/G	HOL	genomic		<i>LOC783341</i>
6:113,748,148	A/G	ANG	genomic		<i>LOC786062</i>
6:113,779,026	C/T	ANG	genomic		<i>LOC789276</i>
6:113,830,597	C/T	HOL	genomic		<i>LYAR MSX1</i>
6:113,830,608	A/G	HOL	genomic		<i>NKX3-2 NSG1</i>
6:114,817,706	G/T	HOL	genomic	**	<i>PROM1 RAB28</i>
6:115,253,699	A/G	HOL	genomic		<i>STK32B STX18</i>
6:115,573,567	A/G	HOL	genomic		<i>TAPT1</i>
6:115,690,439	C/T	LMS	genomic		<i>TMEM128</i>
6:115,690,569	A/G	LMS	genomic		<i>WDR1 ZNF509</i>
6:115,690,633	C/T	LMS	genomic		<i>ZNF518B</i>
6:115,690,674	C/T	LMS	genomic		
6:115,690,854	C/T	LMS	genomic		
6:115,693,714	A/T	LMS	genomic		
6:115,693,956	A/G	LMS	genomic		
6:116,092,430	A/C	ANG	genomic		
6:116,092,711	A/G	ANG	genomic		
8:40,453,120	C/T	HOL	genomic		<i>AK3L1</i>
8:40,456,636	C/T	HOL	genomic		<i>C8H9ORF46</i>
8:41,027,507	C/T	HOL	genomic		<i>CBWD2 CD274</i>
8:41,586,590	G/T	HOL	genomic		<i>CDC37L1 DMRT1</i>
8:41,791,628	A/G	HOL	genomic		<i>DMRT2 FAM122A</i>
8:41,791,805	C/G	HOL	genomic		<i>FOXD4L1 IL33</i>
8:41,794,680	A/G	HOL	genomic		<i>INSL6 JAK2</i>
8:42,045,667	A/T	HOL	genomic		<i>KANK1 KIAA0020</i>
8:42,236,488	A/C	ANG	genomic		<i>LOC100139009</i> <i>LOC506005</i>

Chromosome: coordinate	Allele	Breed	Molecular type	max z	Genes
8:42,424,378	A/G	ANG	genomic		<i>LOC506833</i>
8:42,424,960	A/G	ANG	genomic		<i>LOC513962</i>
8:42,492,165	C/T	HOL	genomic	**	<i>LOC524909</i>
8:42,698,845	G/T	HOL	genomic		<i>LOC531778</i>
8:42,699,002	A/G	HOL	genomic		<i>LOC541094</i>
8:42,736,163	A/G	HOL	genomic		<i>LOC613788</i>
8:44,047,255	G/T	HOL	genomic		<i>LOC614257</i>
8:44,667,098	A/G	HOL	genomic		<i>LOC616937</i>
8:45,269,167	A/G	HOL	genomic		
8:45,330,738	C/T	HOL	genomic		<i>LOC617094</i>
8:45,331,204	A/G	HOL	genomic		<i>LOC781462</i>
8:45,507,466	C/T	LMS	genomic		<i>LOC784342</i>
8:46,356,511	C/G	HOL	genomic		<i>MGC133950 MLANA</i>
8:46,409,513	A/C	HOL	genomic		<i>PDCD1LG2 PGM5</i>
8:46,412,912	A/G	HOL	genomic		<i>PPAPDC2 RANBP6</i>
8:46,785,406	A/G	-	genomic		<i>RCL1 RFX3 SLC1A1</i>
8:46,807,443	A/C	HOL	genomic		<i>SMARCA2 TIMM8A</i>
8:47,260,880	G/T	HOL	genomic		<i>TPD52L3 TSPAN3</i>
8:47,263,964	A/G	HOL	genomic		<i>vldlr</i>
10:28,258,027	A/C	-	genomic		
10:28,303,654	C/G	HOL	genomic		
10:30,750,446	A/C	HOL	genomic		
10:31,457,448	C/G	HOL	genomic		
10:31,457,482	C/T	HOL	genomic		<i>ACTC1 AQR ARHGAP11A</i>
10:32,326,154	C/T	HOL	genomic		<i>ATPBD4 BAHD1 BMF</i>
10:32,654,984	C/T	HOL	genomic		<i>BUB1B C10H15orf23</i>
10:33,118,457	A/G	HOL	genomic	**	<i>C10H15orf41 CASC5</i>
10:33,312,366	A/G	HOL	genomic		<i>CCDC32 CHAC1 CHP</i>
10:33,429,681	A/G	HOL	genomic		<i>CHST14 DISP2 DLL4</i>
10:33,477,339	C/T	ANG	genomic		<i>DNAJC17 EHD4 EXDL1</i>
10:33,887,032	A/G	HOL	genomic		<i>FAM82C FMN1 FSIP1</i>
10:34,284,667	C/G	HOL	genomic		<i>GCHFR GJD2 GPR176</i>
10:34,285,266	C/T	HOL	genomic		<i>GREM1 INOC1 ITPKA IVD</i>
10:35,014,358	C/T	HOL	genomic		<i>JMJD7-PLA2G4B</i>
10:35,072,460	A/C	HOL	genomic		<i>LOC100137087</i>
10:35,928,246	A/G	HOL	genomic		<i>LOC100140532</i>
10:35,928,330	C/T	HOL	genomic		<i>LOC504451 LOC508888</i>
10:35,928,854	A/G	HOL	genomic		<i>LOC513496 LOC513829</i>
10:36,557,397	C/T	HOL	genomic		<i>LOC531014 LOC532711</i>
10:36,849,947	A/G	-	cDNA		<i>LOC616127 LOC786537</i>
10:37,391,016	C/T	HOL	genomic		<i>LOC787480 LOC787676</i>
10:37,391,101	A/G	HOL	genomic		<i>LOC787749 LOC790007</i>
13:24,860,177	C/T	HOL	genomic	**	<i>MAPKBP1 MEIS2 MGA</i>
13:25,009,303	A/G	HOL	genomic		<i>NDUF4F1 NUSAP1 OIP5</i>
13:25,009,332	C/T	HOL	genomic		<i>PAK6 PLA2G4E</i>
13:25,009,472	A/G	HOL	genomic		<i>PPP1R14D RAD51</i>
					<i>RASGRP1 RHOV RPAP1</i>
					<i>RPUSD2 RTF1 SCG5</i>
					<i>SPINT1 SPRED1 SRP14</i>
					<i>SUMO1 THBS1 TMC05A</i>
					<i>TMC05B TYRO3 VPS18</i>
					<i>ZFYVE19</i>
					<i>ACBD7</i>
					<i>ARHGAP21</i>
					<i>ARMETL1</i>
					<i>C13H10ORF38</i>

Chromosome: coordinate	Allele	Breed	Molecular type	max z	Genes
13:25,009,879	C/T	HOL	genomic		<i>C13H10ORF97</i>
13:25,009,905	A/G	HOL	genomic		<i>DCLRE1C</i>
13:25,319,337	A/C	HOL	genomic		<i>FAM107B</i>
13:26,070,157	A/G	HOL	genomic		<i>FRMD4A FZD8</i>
13:26,448,828	C/T	LMS	genomic		<i>GAD2 GJD4</i>
13:26,448,943	-	-	genomic		<i>HSPA14</i>
13:26,449,087	C/G	LMS	genomic		<i>LOC504404</i>
13:26,449,195	G/T	LMS	genomic		<i>LOC511976</i>
13:26,449,315	A/G	LMS	genomic		<i>LOC516469</i>
13:26,898,302	A/G	HOL	genomic		<i>LOC538123</i>
13:28,264,617	C/T	HOL	genomic		
13:28,264,918	-	-	genomic		
13:28,622,073	C/T	HOL	genomic		<i>LOC615924</i>
13:28,850,461	C/T	ANG	genomic		<i>MCM10 MEIG1</i>
13:28,850,534	C/T	ANG	genomic		<i>MYO3A OLAH</i>
13:28,925,713	A/T	LMS	genomic		<i>OPTN PHYH</i>
13:29,247,337	A/G	HOL	genomic		<i>PRPF18 SEPHS1</i>
13:29,652,206	C/T	HOL	genomic		<i>SUV39H2</i>
13:30,344,247	A/G	ANG	genomic		<i>THNSL1</i>
13:30,344,308	A/G	ANG	genomic		
13:30,344,377	A/G	ANG	genomic		
15:61,633,521	C/T	HOL	genomic		
15:62,330,889	A/G	HOL	genomic		<i>ABTB2 APIP</i>
15:62,331,553	C/T	HOL	genomic		<i>C15H11orf74</i>
15:62,525,217	C/G	HOL	genomic		<i>CAPRIN1 CAT</i>
15:63,320,272	C/T	HOL	genomic		<i>CCDC73 CD44</i>
15:63,498,829	C/T	HOL	genomic		<i>CD59 COMMD9</i>
15:63,498,921	C/T	HOL	genomic		<i>CSTF3 DCDC1</i>
15:64,195,562	A/G	ANG	genomic		<i>DEPDC7</i>
15:64,195,635	A/G	ANG	genomic		<i>DKFZP586H2123</i>
15:64,195,762	A/G	ANG	genomic		<i>DNAJC24 EHF</i>
15:64,408,503	C/G	-	genomic		<i>EIF3M ELF5 ELP4</i>
15:65,064,596	G/T	HOL	genomic		<i>FBXO3 FLJ14213</i>
15:65,080,384	C/T	HOL	genomic	**	<i>HIPK3 IMMP1L</i>
15:65,080,495	C/T	HOL	genomic		<i>LDLRAD3 LMO2</i>
15:65,207,827	A/C	HOL	genomic		<i>LOC520044</i>
15:65,208,066	A/C	HOL	genomic		<i>LOC537018</i>
15:65,208,287	A/T	HOL	genomic		<i>LOC616520</i>
15:65,211,662	C/T	HOL	genomic		<i>LOC782183</i>
15:65,211,879	A/G	HOL	genomic		<i>LOC782209</i>
15:65,216,660	C/G	-	cDNA		<i>LOC789207</i>
15:65,451,128	-	-	cDNA		<i>MGC152383</i>
15:66,268,073	C/T	HOL	genomic		<i>NAT10 PAX6</i>
15:66,806,726	A/G	ANG	genomic		<i>PDHX PRRG4</i>
15:67,558,382	C/G	HOL	genomic		<i>QSER1 RAG1</i>
16:35,204,826	C/G	HOL	genomic		<i>RAG2 SLC1A2</i>
16:35,286,684	C/T	HOL	genomic	**	<i>TCP11L1 WT1</i>
					<i>ACOT7 AGTRAP</i>
					<i>ANGPTL7 APITD1 CA6</i>
					<i>CAMTA1 CASZ1 CHD5</i>

Chromosome: coordinate	Allele	Breed	Molecular type	max z	Genes
16:35,758,286	A/G	-	cDNA		<i>CLCN6 CLSTN1 CORT</i>
16:36,760,274	C/G	HOL	genomic		<i>CTNNBIP1 DFFA DHRS3</i>
16:38,072,091	C/T	HOL	genomic		<i>DNAJC11 ENO1 ERRF11</i>
16:38,698,090	C/T	ANG	genomic		<i>EXOSC10 FASLG FBXO2</i>
16:38,698,447	C/G	ANG	genomic		<i>FBXO44 FBXO6 GPR153</i>
16:40,040,569	C/T	ANG	genomic		<i>GPR157 H6PD HES2 HES3</i>
16:40,789,173	C/T	HOL	genomic		<i>KCNAB2 KIAA2013</i>
16:40,789,408	A/C	HOL	genomic		<i>KLHL21 LOC100139219</i>
16:41,483,232	C/T	HOL	genomic		<i>LOC100140094</i>
16:43,246,647	C/T	HOL	genomic		<i>LOC100140347</i>
16:43,490,232	G/T	HOL	genomic		<i>LOC506110 LOC506315</i>
16:43,490,316	A/G	HOL	genomic		<i>LOC507211 LOC513399</i>
16:44,202,741	G/T	HOL	genomic		<i>LOC520588 LOC521303</i>
16:44,202,760	A/C	HOL	genomic		<i>LOC534574 LOC535394</i>
16:44,202,835	C/T	HOL	genomic		<i>LOC539821 LOC614193</i>
16:44,203,082	A/G	HOL	genomic		<i>LOC614690 LOC616962</i>
16:44,205,126	A/G	HOL	genomic		<i>LOC768081 LOC780990</i>
16:44,205,288	C/G	HOL	genomic		<i>LOC781542 LOC782560</i>
16:44,614,076	A/C	BRM	genomic		<i>LOC784352 LOC784385</i>
16:44,984,159	G/T	HOL	genomic		<i>LOC784388 LOC785548</i>
16:45,207,489	C/T	ANG	genomic		<i>LOC785839 LOC785901</i>
16:45,207,652	A/G	ANG	genomic		<i>LOC786597 LOC786885</i>
22:5,778,362	G/T	ANG	genomic		<i>LOC787623 LZIC MAD2L2</i>
22:5,778,643	C/T	ANG	genomic	**	<i>MASP2 METTL13 MTHFR</i>
22:5,778,766	A/C	ANG	genomic		<i>MYOC NMNAT1 NOL9</i>
22:6,000,191	C/T	HOL	genomic		<i>NPHP4</i>
22:6,075,586	G/T	HOL	genomic		<i>NPM1 NPPA NPPB PARK7</i>
22:6,077,987	A/G	HOL	genomic		<i>PEX14 PGD PHF13 PIGC</i>
22:6,121,275	A/G	-	genomic		<i>PIK3CD PLEKHG5</i>
22:6,126,354	C/T	HOL	genomic		<i>PLOD1 PTCHD2 RNF207</i>
22:6,129,390	A/T	HOL	genomic		<i>RPL7 SCYL1BP1</i>
22:6,463,647	C/T	HOL	genomic		<i>SLC25A33 SLC2A5</i>
22:6,945,152	C/G	-	cDNA		<i>SLC45A1 SPSB1 SRM</i>
22:7,020,282	C/T	HOL	genomic		<i>TARDBP TASIR1 THAP3</i>
22:8,379,541	A/T	ANG	genomic		<i>TMEM201 TNFRSF1B</i>
22:8,880,002	A/G	HOL	genomic		<i>TNFRSF25 TNFRSF8</i>
22:8,929,082	A/C	ANG	genomic		<i>TNFRSF9 TNFSF4 UBE4B</i>
22:9,463,695	C/G	HOL	genomic		<i>UBIAD1 VAMP3 VAMP4</i>
22:9,464,019	C/T	HOL	genomic		<i>VPS13D ZBTB48</i>
22:9,464,039	C/T	HOL	genomic		
22:9,799,149	A/T	HOL	genomic		<i>ARPP-21 CCR4</i>
22:10,237,567	A/G	HOL	genomic		<i>CLASP2 CMTM6</i>
22:10,312,098	G/T	HOL	genomic		<i>CMTM7 CMTM8</i>
22:10,312,172	C/T	HOL	genomic		<i>CNOT10 CRTAP</i>
22:10,312,654	A/G	HOL	genomic		<i>DYNC1L11</i>
22:10,729,696	C/G	-	cDNA		<i>EPM2AIP1</i>
X:48,746,914	-	-	genomic		<i>FBXL2 GLB1</i>
					<i>GPD1L</i>
					<i>LOC100140633</i>
					<i>LOC507708</i>
					<i>LOC514651</i>
					<i>LOC785419</i>
					<i>LOC785477</i>
					<i>LOC789829</i>
					<i>LRRFIP2 MLH1</i>
					<i>PDCD6IP STAC</i>
					<i>STT3B TMPPE</i>
					<i>TRIM71</i>
					<i>AKAP4 ALAS2 APEX2 AR</i>

Chromosome: coordinate	Allele	Breed	Molecular type	max z	Genes
X:50,269,536	A/G	HOL	genomic		ARAF ARHGEF9 ARR3 ASB12 CACNA1F
X:53,002,581	C/T	HOL	genomic		CCDC120 CCDC22 CFP
X:55,026,211	C/T	ANG	genomic	**	CLCN5 CSNK1B CXCR3 CXORF36 DGAT2L3
X:55,026,361	-	-	genomic		DGAT2L6 DLG3 EBP EFNB1 ELK1 ERAS
X:55,259,065	-	-	genomic		FO XO4 FOXP3 FTSJ1
X:61,365,643	C/T	LMS	genomic		FUNDC1 GATA1 GPPD2 GJB1 GNL3L GPKOW
X:62,721,733	C/T	BRM	genomic		GPRI73 GRIPAP1 GSPT2 HDAC6 HSD17B10
X:62,811,808	A/G	HOL	genomic		HUWE1 IL2RG IQSEC2 ITGB1BP2 JARID1C
X:63,059,209	C/T	-	cDNA		KCND1 KIF4A LASIL LOC100138330
X:64,139,050	A/G	ANG	genomic		LOC100138998
X:64,464,505	A/G	HOL	genomic		LOC100139177
X:64,682,701	G/T	ANG	genomic		LOC100140102
X:65,004,047	C/G	BRM	genomic		LOC100140899
X:66,004,226	C/T	HOL	genomic		LOC100141016
X:66,007,521	C/T	HOL	genomic		LOC509020 LOC511713 LOC520057 LOC522428
X:67,093,293	G/T	ANG	genomic		LOC523373 LOC531038 LOC531674 LOC531957
X:67,093,329	C/G	ANG	genomic		LOC538185 LOC538872 LOC539029 LOC539973
X:67,093,372	A/G	ANG	genomic		LOC541276 LOC613364 LOC613370 LOC613728
X:67,093,673	A/C	ANG	genomic		LOC615663 LOC617166 LOC781001 LOC781643
X:67,093,709	C/T	ANG	genomic		LOC782639 LOC783214 LOC783297 LOC784528
					LOC785008 LOC785085 LOC785177 LOC785244 LOC785842 LOC786555 LOC787130 LOC789031 LOC790346 LOC790827
					MAGED2 MAGEH1 MAOA MAOB MED12 MGC128691 MGC133640 MID1IP1 MSN MTMR8 NDP NLGN3 NONO NUDT10 NUDT11 NXF3 OTC OTUD5 P2RY4 PAGE4 PCSK1N PCTK1 PDZD11 PKFBI PHF8 PIM2 PJA1 PLP2 PORCN PPP1R3F PQBP1 PRAF2 PRICKLE3 RBM10 RBM3 RIBC1 RNF219 RPGR SLC35A2 SLC7A3 SMC1A SNX12 SRPX STARD8 SYN1 SYP SYTL5 TAF1 TBC1D25 TEX11 TFE3 TIMM17B TIMP1 TMEM47 TSPAN7 TSPYL2 UBA1 USP11 USP27X UTX UXT VSIG4 WDR13 WDR45 YIPF6 ZC3H12B ZC4H2 ZMYM3
X:67,093,794	C/T	ANG	genomic		

Table 4 Extended Haplotype Homozygosity Regions with significant evidence for recent positive selection between Zebu and Taurine cattle. * indicate regions where selection is in the direction of Zebu. The Genes column list annotated genes and transcripts within the corresponding regions.

Number of SNP	Chromosome: Interval (Mb)	Genes
6	5: 15.7 – 18.5	<i>ALX1 LOC781108 LRR1Q1 MGAT4C NTS PAMCI PRKRIP1 SLC6A15 TSPAN19</i>
8	10: 9.3 – 11.5	<i>ARSB BHMT DMGDH ERO1L GNP NAT1 GPR137C HOMER1 LOC528498 LOC533581 LOC618795 MTX3 PAPD4 PSMC6 SERINC5 STYX THBS4</i>
2*	10: 81.5 – 82.9	<i>ARG2 ATP6V1D C10H14orf83 EIF2S1 LOC614449 LOC781580 LOC781894 LOC783372 LOC785061 PIGH PLEK2 PLEKHH1 RAD51L1 RDH11 RDH12 VTI1B ZFYVE26</i>
2	11: 28.1 – 28.1	<i>None found</i>
4	13: 70.8 – 72.0	<i>LOC614378</i>
10	18: 14.9 – 21.1	<i>ABCC11 ADCY7 BRD7 CHD9 CYLD HEATR3 LOC100137797 LOC516179 LOC528068 LOC533093 LOC781014 LOC786252 LOC789672 LOC789692 LOC789724 LOC789870 LOC789894 LOC789927 LONP2 N4BP1 NKD1 NOD2 PHKB SALL1 SIAH1 SNX20 TMEM188 TOX3 ZNF423</i>
5*	19: 42.2 – 44.2	<i>ACLY ATP6V0A1 CCDC56 CCR10 CNP CNTNAP1 DNAJC7 EIF1 EZH1 FKBP10 GAST GHDC HCRT HSD17B1 HSPB9 JUP KLHL10 KLHL11 KRT10 KRT12 KRT15 KRT17 KRT19 KRT20 KRT23 KRT25 KRT26 KRT27 KRT31 KRT32 KRT33A KRT33B KRT34 KRT35 KRT36 KRT39 KRTAPI-1 KRTAP3-3 KRTAP4-7 KRTAP9-2 KRTAP9-4 LOC100137846 LOC100138157 LOC100138622 LOC100138853 LOC100138974 LOC100139067 LOC100139290 LOC100139304 LOC100139539 LOC100139727 LOC100139742 LOC100140245 LOC100140417 LOC100140506 LOC100140799 LOC100140870 LOC100141097 LOC100141108 LOC508529 LOC510062 LOC514420 LOC514507 LOC514812 LOC515000 LOC535907 LOC538611 LOC540285 LOC614728 LOC618422 LOC618436 LOC618491 LOC618495 LOC618938 LOC777598 LOC785923 LOC785989 LOC786727 LOC786827 LOC786968 LOC787032 LOC787051 LOC787090 LOC787110 LOC787132 LOC787225 LOC787429 LOC787447 LOC788110 LOC788124 LOC788170 LOC788186 LOC788217 LOC788228 LOC788269 LOC788272 LOC788284 LOC788650 LOC788727 LOC788823 LOC789163 LOC789362 LOC790321 LOC790883 MLX NAGLU NKIRAS2 NT5C3L PLEKHH3 PTRF RAB5C RAMP2 STAT3 STAT5A STAT5B TTC25 TUBG1 TUBG2 VPS25</i>
13*	21: 24.2 – 30.7	<i>ACSBG1 ANKRD34C APBA2 ARNT2 BCL2A1 BNC1 BTBD1 C21H15ORF26 CHRNA7 CTSH FAH FAM108C1 IDH3A IL16 KLF13 LOC100138339 LOC504694 LOC519047 LOC520039 LOC526431 LOC530472 LOC541059 LOC617373 LOC617531 LOC783145 LOC784070 LOC784266 LOC784556 LOC789855 LOC789940 LOC789946 MCEE MESDC1 MESDC2 MORF4L1 MPHOSPH10 MTHFS MTMR10 MTMR15 NDNL2 PCSK6 SH3GL3 SNRPA1 STARD5 TARSL2 TBC1D2B TJP1 TM2D3 TM6SF1 TMED3 TRPM1 ZFAND6</i>
4	22:20.8 – 21.9	<i>None found</i>
16*	22: 46.1 – 56.7	<i>ABHD14A ABHD14B ACTR8 ACY1 ALAS1 ALS2CL AMT APEH ARIH2 ARMET ATP2B2 ATRIP C22H3ORF18 C22H3orf60 CACNA1D CACNA2D2 CAMKV CAMP CATHL1 CATHL2 CATHL4 CATHL5 CATHL6 CCDC12 CCDC36 CCDC51 CCDC71 CCR1 CCR2 CCR3 CCR5 CCR9 CCRL2 CDC25A CDCP1 CELSR3 CHDH CISH CLEC3B CSPG5 CXCR6 CYB561D2 DAG1 DCPIA DHX30 DUSP7 EEF1A1 EPRS GHRL GLT8D1 GLYCTK GMPPB GNAI2 GNAT1 GNL3 GPR62 GPX1 GRM2 HEMK1 HYAL1 HYAL2 HYAL3 IFRD2 ILI7RB IMPDH2 IP6K1 IP6K2 IQCF1 IQCF2 IQCF5 ITIH1 ITIH3 ITIH4 KIF9 KLHDC8B KLHL18 LAMB2 LARS2 LOC100125949 LOC100138407 LOC100138572 LOC100139030 LOC100139031 LOC100139246 LOC100139678 LOC100139910 LOC100140621 LOC100141058 LOC100141067 LOC507133 LOC507426 LOC508041 LOC509924 LOC511069 LOC514296 LOC515009 LOC515991 LOC517392 LOC518845 LOC519644 LOC522045 LOC522434 LOC522724 LOC524080 LOC524692 LOC526885 LOC529196 LOC530599 LOC537884 LOC541079 LOC614114 LOC616365 LOC616410 LOC616532 LOC616798 LOC618550 LOC782900 LOC782954 LOC783185 LOC783221 LOC783987 LOC784021 LOC784204 LOC784245 LOC785460 LOC785746 LOC785925 LOC786337 LOC786363 LOC786520 LOC786618 LOC786887 LOC787237 LOC787673 LOC787697 LOC787776 LOC788093 LOC788112 LOC788383 LOC788394 LOC789065 LOC789098 LOC790172 LRRC2 LTF LZTFL1 MAP1LC3B MAP4 MAPKAPK3 MGC137053 MGC148581 MON1A MST1 MSTIR MUSTN1 MYL3 NBEAL2 NCKIPSD NEK4 NICN1 NME6 NRADD P4HTM PBI PCBPA PFKFB4 PPM1M PRKAR2A PRKCD PTH1R PTPN23 QRICH1 RASSF1 RBM15B RBM5 RFT1 RHOA RNF123 RPL29 RRP9 RTP3 SACMIL SCAP SEC13 SELK SEMA3B SEMA3F SETD2 SFMBT1 SHISA5 SLC25A20 SLC26A6 SLC38A3 SLC6A1 SLC6A20 SPCS1 STAB1 TCTA TDGF1 TESSP2 TESSP5 TEX264 TKT TLR9 TMEM103 TMEM110 TMEM115 TMEM158 TMEM42 TMEM89 TMIE TRAIR TREX1 TSG101 TUSC4 TWF2 UBA7 UQCRC1 USP19 USP4 VPRBP WDR51A WDR6 WDR82P1 WNT5A XCR1 ZDHHC3 ZMYND10</i>

7	X: 1.2 – 7.0	AIFM1 APLN BCORL1 ELF4 GRIA3 LOC510101 LOC535439 LOC616656 LOC784703 LOC784970 LOC785262 MGC160048 OCRL SASH3 THOC2 UTP14A XPNPEP2 ZDHHC9 ZNF280C
47	X: 39.5 – 73.5	5hr2c ABCB7 ACOT9 ACRC AKAP4 ALAS2 APEX2 APOO AR ARAF ARHGEF9 ARR3 ASB12 ATRX BRWD3 CACNA1F CCDC120 CCDC22 CFP CHIC1 CHM CLCN5 CSNK1B CXCR3 CXORF36 DGAT2L3 DGAT2L6 DLG3 EBP EFNBI EIF2S3 ELK1 ERAS ERCC6L FGF16 FOXO4 FOXP3 FTSJ1 FUNDC1 GATA1 GDPD2 GJB1 GNL3L GPKOW GPR173 GRIPAP1 GSPT2 H11CXORF26 HDAC6 HSD17B10 HUWE1 IL2RG IQSEC2 ITGB1BP2 JARID1C KCND1 KIF4A KLHL15 LAS1L LDHA LOC100137994 LOC100138330 LOC100138998 LOC100139177 LOC100140102 LOC100140899 LOC100141016 LOC509020 LOC511713 LOC512493 LOC513925 LOC519208 LOC520057 LOC521086 LOC521092 LOC522428 LOC522603 LOC522938 LOC523373 LOC524097 LOC529211 LOC530295 LOC531038 LOC531674 LOC531957 LOC535872 LOC537655 LOC538185 LOC538872 LOC539029 LOC539973 LOC541276 LOC613364 LOC613370 LOC613728 LOC613970 LOC614207 LOC614223 LOC614449 LOC614855 LOC615291 LOC615663 LOC616695 LOC617166 LOC618364 LOC781001 LOC781580 LOC781643 LOC781736 LOC781894 LOC782007 LOC782010 LOC782293 LOC782514 LOC782639 LOC782703 LOC782882 LOC783214 LOC783297 LOC783300 LOC783372 LOC783975 LOC784097 LOC784438 LOC784528 LOC784572 LOC784663 LOC785008 LOC785061 LOC785085 LOC785177 LOC785244 LOC785599 LOC785842 LOC786555 LOC786944 LOC786985 LOC787130 LOC789031 LOC789236 LOC790346 LOC790827 MAGEB18 MAGED2 MAGEH1 MAOA MAOB MED12 MGC128691 MGC133640 MGC140080 MGC152340 MGC165780 MID1IP1 MSN MTMR8 NDP NDUFV2 NHSL2 NLGN3 NONO NSBP1 NUDT10 NUDT11 NXF3 OGT OTC OTUD5 P2RY4 PAGE4 PCSK1N PCK1 PCYT1B PDK3 PDZD11 PFKFB1 PHF8 PIM2 PIN4 PJA1 PLP2 PLS3 POLA1 PORCN POU3F4 PPP1R3F PQBP1 PRAF2 PRDX4 PRICKLE3 RBM10 RBM3 RIBC1 RNF12 RNF219 RPGR RPS3A RPS4Y2 SAT1 SH3BGR1 SLC35A2 SLC7A3 SMC1A SNX12 SRPX STARD8 SYN1 SYP SYTL5 TAF1 TBC1D25 TEX11 TFE3 TIMM17B TIMP1 TMEM47 TSPAN7 TSPYL2 UBA1 UPRT USP11 USP27X UTX UXT VSIG4 WDR13 WDR45 XIST YIPF6 ZC3H12B ZC4H2 ZDHHC15 ZFX ZMYM3

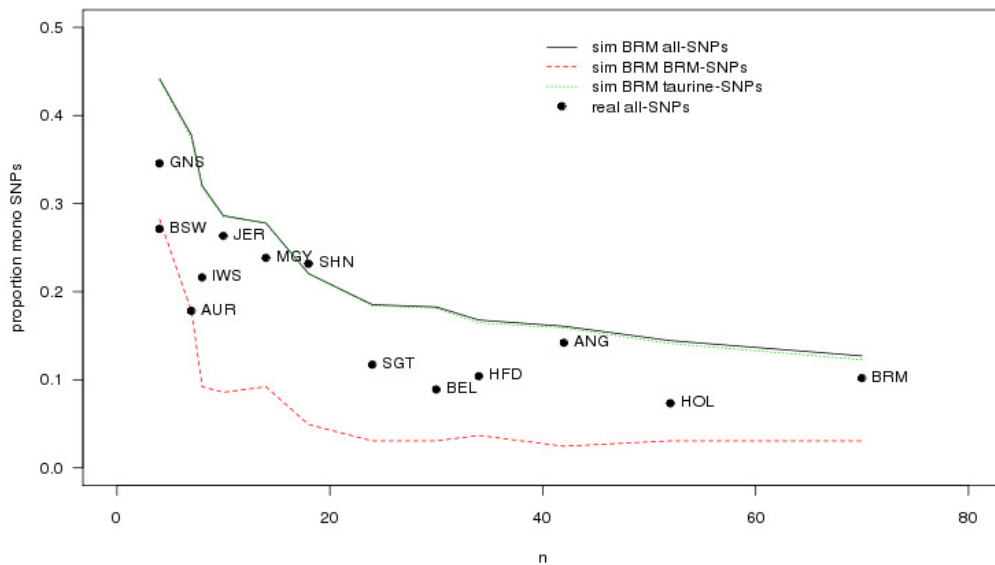
Supplementary Table 5 Proportion of Zebu or Taurine ancestry as estimated by STRUCTURE for K=2 and correlations of estimated Zebu proportions between those from using all autosomal SNP and the corresponding SNP subset (row). Shown are averages and standard deviations of five replicate runs; in cases of random subsets of SNP, averages and standard deviations were averaged over the five random subsets of SNP, and results for the 13 SNP private in Brahman are averages of two samplings of 13 of the 78 SNP.

	Brahman with zebu ancestry	Taurine with Taurine ancestry	Belmont Red with zebu ancestry		Santa Gertrudis with zebu ancestry	
All autosomal SNP	0.92 (±0.05)	0.971 (±0.008)	0.34 (±0.08)	Correlation to all autosomal SNP	0.37 (±0.05)	Correlation to all autosomal SNP
All SNP	0.91 (±0.05)	0.971 (±0.008)	0.34 (±0.08)	1.00**	0.37 (±0.05)	1.00**
91 private SNP	0.98 (±0.04)	0.999 (±0.000)	0.26 (±0.12)	0.70**	0.46 (±0.12)	0.64**
78 private in Brahman	1.00 (±0.06)	0.999 (±0.000)	0.25 (±0.12)	0.73**	0.44 (±0.11)	0.64**
13 private in Taurine breeds	0.96 (±0.00)	0.798 (±0.147)	0.42 (±0.22)	0.19	0.57 (±0.24)	0.23
13 random private in Brahman	0.95 (±0.09)	0.993 (±0.001)	0.30 (±0.20)	0.39	0.53 (±0.24)	0.37
14 diplotype SNP	0.97 (±0.04)	0.962 (±0.031)	0.29 (±0.17)	0.54*	0.33 (±0.17)	0.42*
12 EHH SNP	0.80 (±0.12)	0.65 (±0.118)	0.53 (±0.16)	-0.01	0.62 (±0.15)	-0.07
91 random SNP	0.95 (±0.06)	0.948 (±0.042)	0.37 (±0.14)	0.80**	0.40 (±0.13)	0.73**
78 random SNP	0.94 (±0.06)	0.934 (±0.051)	0.40 (±0.16)	0.87**	0.37 (±0.13)	0.61*
13 random SNP	0.86 (±0.12)	0.727 (±0.154)	0.42 (±0.21)	0.50*	0.46 (±0.23)	0.29
14 random SNP	0.81 (±0.09)	0.717 (±0.114)	0.49 (±0.21)	0.34	0.55 (±0.20)	0.23

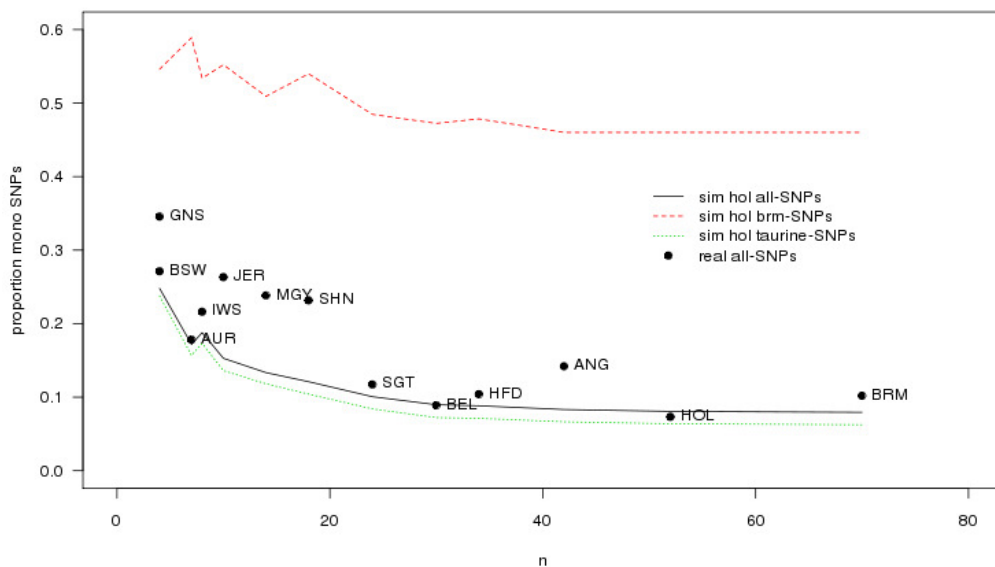
12 random SNP	0.84 (± 0.14)	0.711 (± 0.159)	0.31 (± 0.11)	0.65**	0.58 (± 0.26)	0.29
----------------------	---------------------	-----------------------	---------------------	--------	---------------------	------

*P<0.05; ** P<0.001

Supplementary Figures



Supplementary Figure 1 Sensitivity analysis of sample size, n, on proportion of monomorphic SNP. Solid dots mark the actual monomorphic proportion for the 13 breeds (as in Supplementary Figure 1). Lines show the decline of monomorphic proportions within increasing sample size from the simulated dataset generated using only Brahman animals; solid black line correspond to using all SNP, green dotted line correspond to taurine SNP, and red dashed line to Brahman-discovered SNP.



Supplementary Figure 2 Sensitivity analysis of sample size, n, on percentage of monomorphic SNP. Solid dots mark the actual monomorphic proportion for the 13 breeds (as in Supplementary Figure 1). Lines show decline of monomorphic proportions with increasing sample size from the simulated dataset generated using only Holstein animals; solid black line correspond to using all SNP, green dotted line correspond to taurine SNP, and red dashed line to Brahman-discovered SNP.

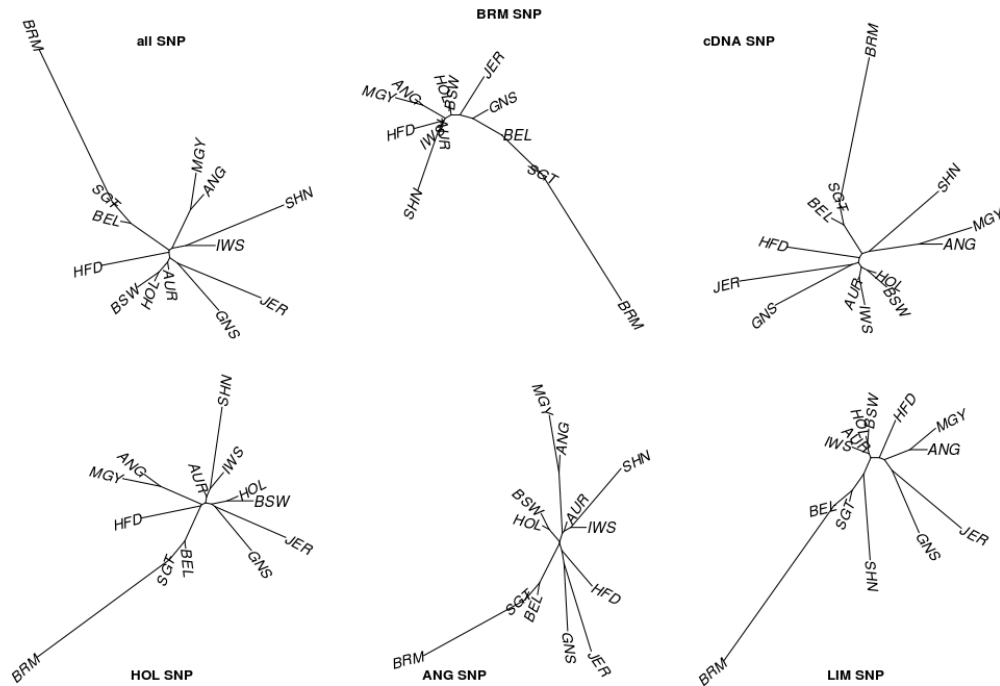
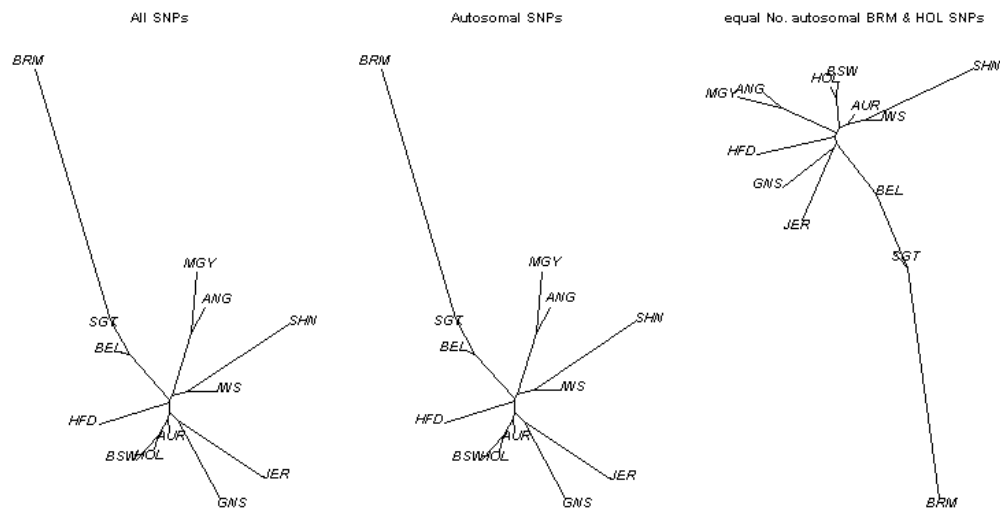
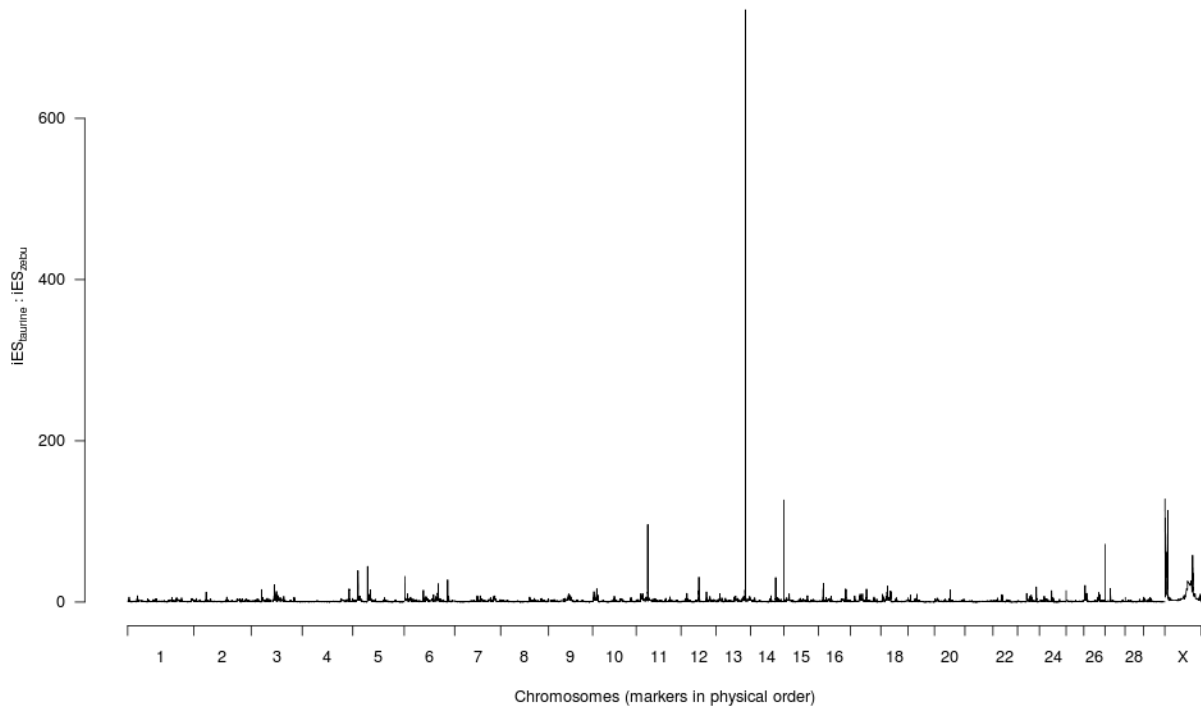


Figure 3 Neighbour Joining Trees showing the relationship between all 13 breeds. Trees were constructed using F_{ST} values estimated for each breed-pair using all 8,427 SNPs (top left), the 163 Brahman SNP (top middle), 775 cDNA SNP (top right), 6,032 Holstein SNP (bottom left), 1,213 Angus SNP (bottom middle), and 244 Limousin SNP (bottom right)

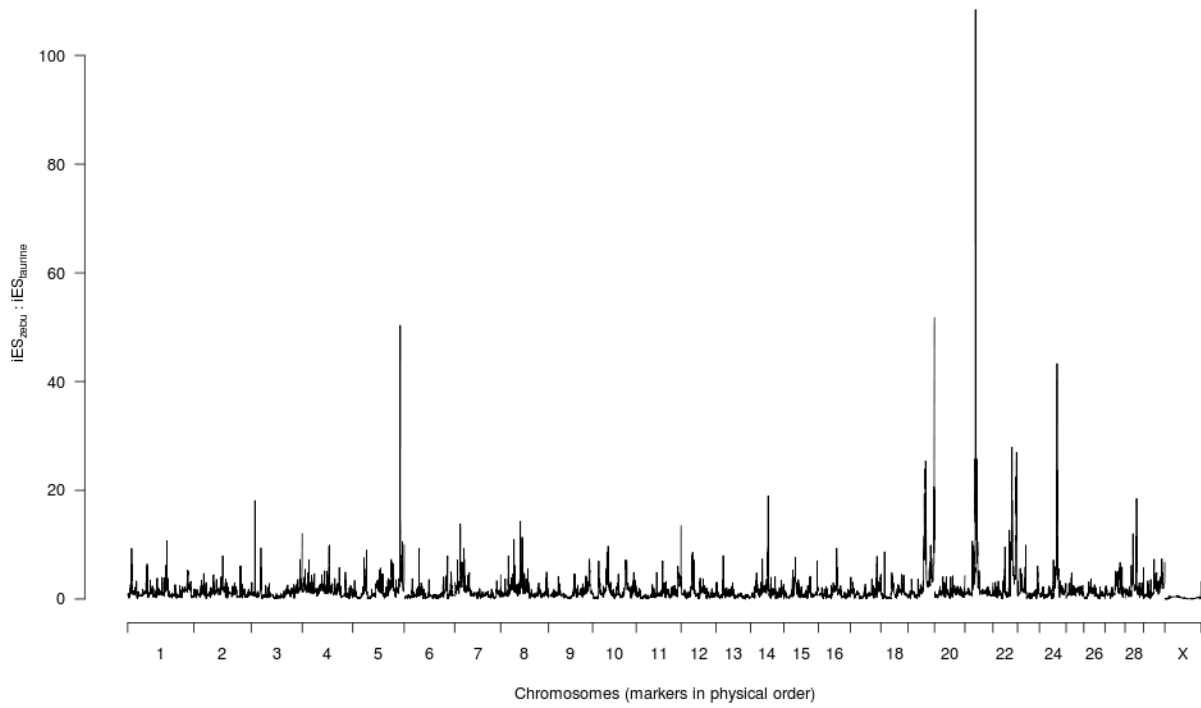


Supplementary Figure 4 Neighbouring-Joining Tree of all 13 breeds constructed using F_{ST} values estimated for each breed-pair using all 8,427 SNPs (left), 7,821 autosomal SNPs (middle), and 302 SNPs of which half were Brahman derived and half were random samples of Holstein derived SNPs (right).

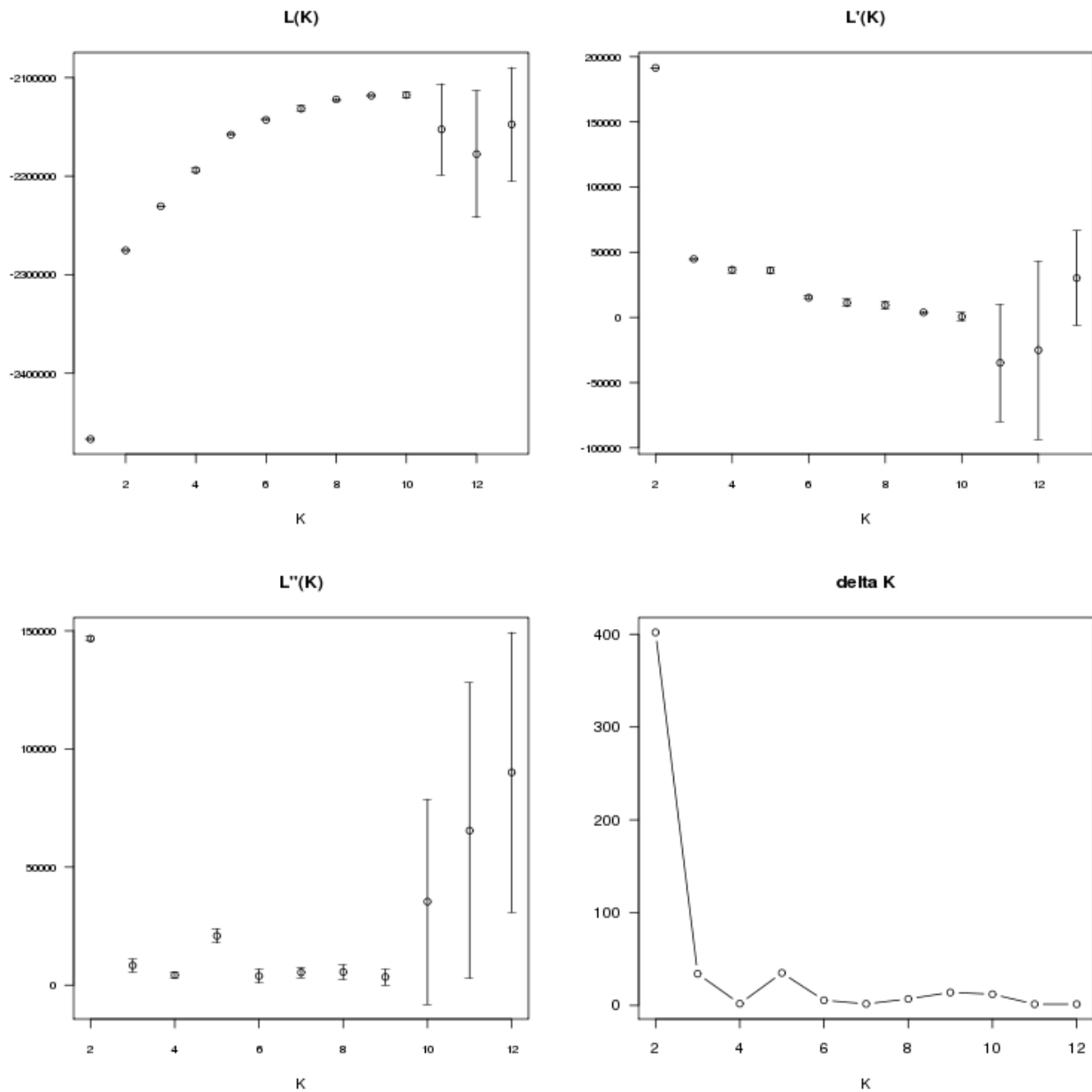
positive selection in taurine compare to zebu



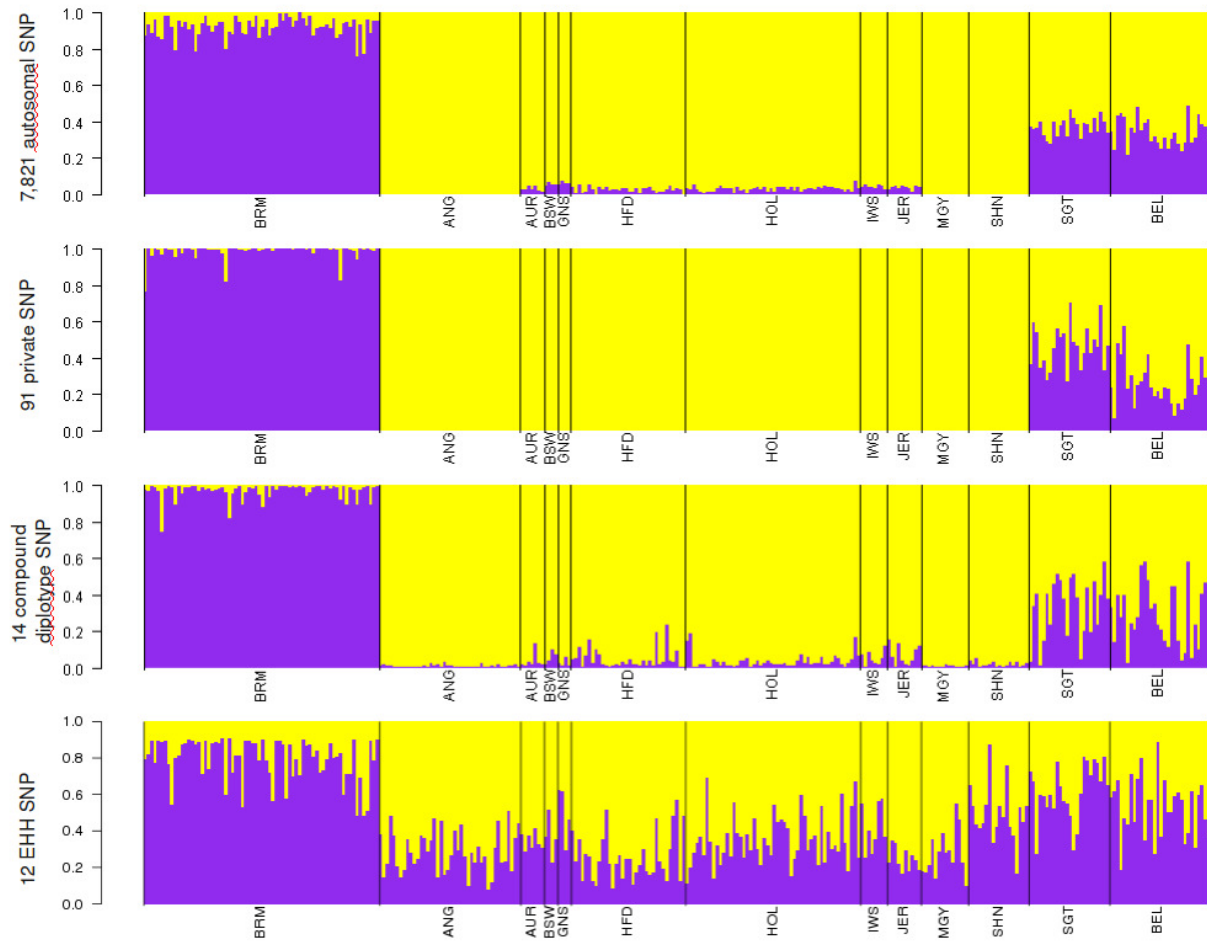
positive selection in zebu compare to taurine



Supplementary Figure 5 Comparison of extended haplotype homozygosity between taurine and zebu cattle breeds. Shown is the ratio of integrated EHHS of taurine to zebu (A) and zebu to taurine (B).



Supplementary Figure 6. Relationships between K and STRUCTURE's estimated log likelihood of the data, $L(K)$, (top left), the first order rate of change of the likelihood with respect to K, $L'(K)$, (top right), the second order rate of change of $L(K)$, $L''(K)$, (bottom left), and the ΔK statistic (bottom right). These plots correspond to those reported in Fig.2 of Evanno, Regnaut, and Goudet (2005).



Supplementary Figure 7 STRUCTURE predictions of the proportions of two ancestral populations ($K=2$) for 317 individuals belonging to 13 breeds using (from top to bottom): 7,821 autosomal SNP, 91 private SNP, 14 SNP representing the 14 compound diplotype regions, and 12 SNP representing the 12 regions of recent positive selection. In all cases, averages of five MCMC replicate runs are shown, and all individuals are ordered identically along the x-axis. Breed acronyms are as follows: BRM = Brahman; SGT = Santa Gertrudis; BEL = Belmont Red; HFD = Hereford; BSW = Brown Swiss; HOL = Holstein; AUR = Australian Red; GNS = Guernsey; JER = Jersey; IWS = Illawarra Shorthorn; SHN = Shorthorn; ANG = Angus; and MGY = Murray Grey