Supporting Information

Contents

Sequencing depth strategy

We determined the probability of detecting an allele if heterozygous or homozygous in each of two first-order relatives given a combined total of 18x coverage, and defined the probability of detecting alleles shared between these relatives at different sequencing depths. Alleles that were heterozygous in one individual and homozygous in the other were informative for reconstruction of the shared chromosomal haplotypes.

The probability of detecting an informative SNP for the haplotype reconstruction (heterozygous in one bull, homozygous in the other) at 6x Chief/12x Mark coverage was determined to be ~85%. We considered cases where one bull was homozygous and one heterozygous at a particular SNP and assumed that either bull was equally likely to be the heterozygous one. Then we assumed that the number of reads for Mark at a particular site was Poisson distributed with mean = 13.5 and calculated the probability that both alleles were observed at least 3 times. This is 93%. The equivalent calculation for Chief was the probability of observing each allele at least twice given an average read depth of 7.3 times which was 77%. Therefore the average across the two bulls was 85%. The different X chromosomes of Mark and Chief (inherited from different mothers) would be covered at .5x of their respective sequencing depths whereas the shared Y chromosome would be sequenced at ~9x. Because of the different maternal origins of the X (not part of the shared genome of the two bulls) we did not consider the X chromosome for analysis. The Y chromosome also was not analyzed because a complete Y assembly was not available at the time the sequence analysis was performed.

Library preparation

Briefly, 5 µg of the DNA sample from each bull were nebulized for 1 min and fragments of 300-500 bp were selected using a gel cut method. The DNA quality was assessed on a BioAnalyzer DNA 7500 LabChip followed by fragment-end polishing and adaptor ligation. Small fragments were removed after incubation of the DNA with the AMPure beads. Large fragments of the adaptor-ligated DNA were immobilized using library immobilization beads and a single-stranded DNA library was generated in NaOH (10 N) solution. After quality assessment and quantification, the library was diluted to 1×10^{8} molecules/µl in TE Buffer and stored at -25^{0}C until used.

Alignment and mapping

Reads from the same tier of the same run that had the same size ±2 bp and the same alignment coordinates were considered as duplicate reads produced from the rare accidental inclusion of two beads within a single emulsion droplet. Only one of such reads was used for the identification of nucleotide polymorphisms.

For all reads considered as mapped to chromosomes we generated sff files. A total of 30 sff files were generated for each bull. These files were used to generate the chromosome-specific multi-sequence FASTA and FASTA quality files for each animal using PyroBayes v. 0.9 software (1). We used a feature of PyroBayes that permitted adding an extension to each sequence name to distinguish reads originating from the two animals. After adding this extension, the FASTA and FASTA quality files that contained sequences mapped to the same cattle chromosome were merged generating a single set of 30 FASTA and 30 FASTA quality files. We used MosaikBuild v. 0.9 to generate the reference database files from individual cattle chromosome FASTA files (Btau4.0) and merged bull sequence read files. A hash size of 15 (-hs 15) recommended for usage with Titanium 454 reads was used to generate reference chromosome database files. Each reference chromosome file was aligned with the corresponding read database file using MosaikAligner with the following criteria (-hs 15, -mmp 0.05, -act 40, -mhp 100), which allowed the alignment of reads to contain ≤ 5% mismatches with the reference sequence and an initial alignment seed of 40 bp. For alignments with multiple position seeds, 100 random positions were used to build the alignments. Resulting alignment files were parsed by the MosaikSort and MosaikAssembler programs to generate a multi-sequence reference-based combined assembly of Mark and Chief chromosomes. Only the reads with unique chromosomal assignments from the MosaikAligner output were considered for the SNP detection.

A total of 1,473,596 reads had no matches in Btau4.0 when aligned by Newbler or BLAT (Table S2). Of these, we removed 474,164 reads with highly repetitive sequences and assembled the remaining 992,432 reads into 48,702 contigs with N50 of 913 bp using the "*de novo, accurate"* options of the MIRA3.0 next-generation sequence assembler (2).

All 25,365 contigs ≥500 bp were remapped against Btau4.0. None of these contigs could be mapped to the Btau4.0 chromosome assemblies; however, nine contigs <800 bp mapped to Btau4.0 unassigned contigs with >90% sequence identity and over 90% of the contig length covered in the alignment. To investigate if unmapped contigs mostly represent sequences omitted from the Btau4.0 assembly but still present in the sequenced reference genome we then aligned all contigs ≥500 bp long to the University of Maryland cattle genome assembly UMD 3.0 (3); 20,383 (80.3%) of the contigs containing reads originally unassigned in Btau4.0 were successfully mapped to this assembly. Because Btau 4.0 and UMD3.0 assemblies were built from the same set of raw sequence reads, the majority of our unmapped contigs represent sequences lost in the generation of Btau4.0.

Allele calling and SNP filtering

To identify putative SNP loci that are homozygous in one or both bulls, but with an allele different from the reference genome sequence, we parsed the initial gigaBayes output using the following criteria: 3x minimum read coverage for a new allele with at least 2 reads aligned in opposite orientation (forward, reverse), minimum allele sequence quality ≥20, and the polymorphism could not be located within homopolymeric regions (≥5 tandem uninterrupted repeats of the same nucleotide) within the reference sequence or assembled reads.

To identify putative SNPs that have two alleles in the sequence reads originating from at least one sequenced animal, the initial gigaBayes file was parsed to exclude polymorphisms in homopolymeric regions (≥5 bp tandem single base repeats), regions containing reads aligned only in one orientation for the allele that differs from the reference genome sequence, reads with sequence quality <20 at a SNP position, and with the minor allele coverage ratio to the total coverage at the SNP position ≤0.2. We eliminated SNPs that had coverage higher than 98% of all SNPs in each bull (≥31x coverage in Mark and ≥20x coverage in Chief) because these SNPs are likely to be located in duplicated chromosomal regions and be the results of misalignment. The rest of the *in silico* SNPs were considered as heterozygous in at least one individual if two alleles were present at the same position in the reads of Mark, Chief or both bulls. Additional filtering criteria were applied to the last set to identify *in silico* SNPs with each allele covered by at least three sequence reads in Mark and by at least two reads in Chief.

Linear model for allele frequency difference in Chief descendants and non-descendants.

The following model was fitted individually for all loci that were not homozygous in Chief (N= 35,036): $y_{ij} = mu + desc_i + ch_j + desc_i \times ch_j$, where y_{ij} was a variable coded 0, 1, or 2 for each individual i indicating how many copies of Chief chromosome j were tracked yielding two records per individual (one per Chief chromosome), $desc_i$ was coded 1 for descendants and 2 for non-descendants, and ch_j was 1 and 2 for chromosome 1 or 2. Significance of the interaction term $desc_i \times ch_j$ indicates a locus where one of the alleles identical to Chief was at significantly different frequency to its compliment in his descendants but this was not the case in non-descendants, or vice versa. Such loci are likely to have experienced selective pressure (Figure S7, Table S8).

Table S1. Mapping statistics.

Statistics	Mark	Chief
Mapped by Newbler		
reads (M)	57.9	31.8
base pairs (Gbp)	20.8	11.0
Mapped by BLAT		
reads (M)	17.2	10.0
base pairs (Gbp)	7.2	4.1
Mapped to chromosomes		
reads (M)	71.4	37.7
base pairs (Gbp)	26.6	13.7

Table S2. *De novo* assembly and mapping of reads with no Newbler or BLAT hits against Btau4.0

Statistic	No.
No-hit reads from Mark	906,545
No-hit reads from Chief	567,051
Total no-hit reads	1,473,596
Assembled reads (MIRA 3.0)*	992,432
No. singlets	382
No. contigs (≥2 reads)	48,702
combined length	32.1 Mbp
N50	913 bp
No. contigs ≥500 bp	25,365 (100.0%)
combined length	22.9 Mbp
No. contigs mapped to Btau 4.0**	9 (0.04%)
No. contigs mapped to UMD3.0**	20,383 (80.3%)

*Non-repetitive reads that could be used for *de novo* assembly

**Mapping criteria by BLAT were >90% identity and >90% of a contig length covered in the alignment

Table S3. SNP detection and filtering statistics

Filter	"Homozygous" in Mark and Chief reads	"Heterozygous" in Mark and Chief reads (autosomal)
Initial SNP set	1,851,126 (100.0%)	10,583,734 (100.0%)
Homopolymer regions (≥5bp)	1,672,815 (90.4%)	6,992,325 (66.1%)
SNPs with reads in a single direction	1,247,978 (67.4%)	2,787,407 (26.3%)
3x minimum allele coverage*	1,207,103 (65.2%)	NA
3x coverage per allele, Mark	NA**	1,311,454 (12.4%)
2x coverage per allele, Chief	NA	818,065 (7.7%)
Upper coverage threshold		
30x coverage Mark	NA	1,286,004 (12.2%)
21x coverage Chief	NA	802,559 (7.6%)
Allele-to-coverage ratio threshold (>0.2)		
Mark	NA	1,243,113 (11.7%)
Chief	NA	757,266 (7.2%)
Heterozygous in at least one bull	NA	1,356,094 (12.8%)
Heterozygous in both bulls	NA	265,707 (2.5%)
Homozygous for a different allele	NA	117,908 (1.1%)
Haplotype phase reconstructed***	NA	972,479 (9.2%)

*Positions and alleles of filtered homozygous SNPs are available from:

http://www.bioinformaticsonline.com/sample.html

**NA; not applicable.

***Complete phasing of all Chief and Mark alleles is available from:

http://www.bioinformaticsonline.com/sample.html

Table S4. Comparison between genotyped and sequence-based SNPs

Set	Mark	Chief
SNP50 SNPs genotyped	54,001 (100.0%)	54,001 (100.0%)
Mapped to Btau4.0 chromosomes	52,329 (96.9%)	52,329 (96.9%)
Successful genotypes	52,138	52,155
Corrected genotypes*	44 (0.1%)	NA
Reconstructed genotypes	109 (0.2%)	NA
No. autosomal heterozygous SNP50 SNPs in the unfiltered sequence SNP set	14,747 (93.2%)	13,407 (86.8%)
No. SNP50 heterozygous SNPs in the sequence set	14,033 (95.2%)	10,219 (76.2%)
No. heterozygous SNPs in filtered sequence set	1,243,113	757,266
No. overlapping SNP50 heterozygous SNPs	8,009 (50.6%)	5,459 (35.3%)
Reported as homozygous in the sequence set	146 (1.8%)	801 (14.7%)
No. homozygous SNP50 SNPs	3,462	4,182
No. reported as heterozygous in the sequence set	52 (1.5%)	39 (0.9%)
Total no. of overlaps between sequence and SNP50	11,471	9,641
Total no. of inconsistent genotypes	198 (1.7%)	840 (8.7%)
Total no. of consistent genotypes	11,273 (98.3%)	8,801 (91.3%)

*These Mark genotypes contradicted segregation of the same SNPs in his genotyped offspring and were corrected to match offspring genotypes

Table S5. Agreement in allele phasing for haplotypes defined by 454 Sequencing and SNP50 genotyping

Min. total coverage	No. SNPs.	No. SNP50 SNPs	Agreement (%)
3x Mark, 2x Chief	972,479	6,109	92.6
3x Mark, 3x Chief	879,926	5,556	93.7
6x Mark, 6x Chief	512,545	3,378	96.7
7x Mark, 7x Chief	386,829	2,625	97.3

Table S6. Statistics on the length of tracked segments in descendents of Chief

All statistics in number of loci, except where indicated, convert to megabases by multiplying by mean interval of 0.0768 mb (i.e., 3000mb / 39047SNP = 0.0768)

All segments

chr	min	max	mean length of segments	mean length in Mb	Stand. Deviation
1	5	1146	155.65	11.95	148.17
2	2	1727	139.30	10.70	141.18
3	38	1589	180.65	13.87	166.62
4	41	1033	177.25	13.61	133.89
5	8	552	108.16	8.31	79.74
6	4	1396	161.33	12.39	139.97
7	17	1162	132.51	10.18	119.25
8	3	1641	185.88	14.28	195.38
9	42	1494	258.12	19.82	251.61
10	2	765	146.30	11.24	108.19
11	4	1184	139.44	10.71	127.31
12	27	1037	154.15	11.84	134.26
13	42	555	124.36	9.55	91.97
14	2	1318	188.42	14.47	199.51
15	42	1002	193.07	14.83	163.12
16	42	499	112.08	8.61	72.95
17	2	1113	141.52	10.87	125.10
18	41	662	126.73	9.73	97.96
19	4	1066	155.29	11.93	151.30
20	42	1063	198.38	15.24	163.28
21	9	1029	159.76	12.27	141.25
22	38	526	110.63	8.50	76.33
23	17	762	150.83	11.58	122.29
24	5	807	129.23	9.93	101.82
25	42	762	142.10	10.91	119.87
26	8	799	160.18	12.30	111.62
27	42	608	128.50	9.87	104.82
28	42	720	201.26	15.46	146.56
29	42	798	199.01	15.28	144.52
mean	22.59	993.62	157.24	12.08	133.79

Only segments longer than 30 loci inlcuded

chr	min	max	mean length of segments	mean length in Mb.	Stand. Deviation
1	33	1146	156.25	12.00	148.19
2	31	1727	145.68	11.19	141.61
3	38	1589	180.65	13.87	166.62
4	41	1033	177.25	13.61	133.89
5	31	552	114.69	8.81	78.24
6	32	1396	161.98	12.44	139.92
7	31	1162	132.99	10.21	119.28
8	31	1641	186.51	14.32	195.46
9	42	1494	258.12	19.82	251.61
10	41	765	146.44	11.25	108.14
11	32	1184	149.51	11.48	127.30
12	38	1037	154.26	11.85	134.26
13	42	555	124.36	9.55	91.97
14	31	1318	191.77	14.73	199.98
15	42	1002	193.07	14.83	163.12
16	42	499	112.08	8.61	72.95
17	31	1113	148.11	11.37	125.24
18	41	662	126.73	9.73	97.96
19	37	1066	155.44	11.94	151.30
20	42	1063	198.38	15.24	163.28
21	34	1029	161.30	12.39	141.22
22	38	526	110.63	8.50	76.33
23	32	762	151.01	11.60	122.28
24	34	807	131.15	10.07	101.49
25	42	762	142.10	10.91	119.87
26	34	799	160.77	12.35	111.46
27	42	608	128.50	9.87	104.82
28	42	720	201.26	15.46	146.56
29	42	798	199.01	15.28	144.52
mean	36.86	993.62	158.62	12.18	133.75

Table S7. Genes with SNPs in 49 selected regions

Region No.	chr	start	end	haplotype	No. intron SNPs	No. exon SNPs	No. 3'UTR SNPs	No. 5'UTR SNPs	No. down stream SNPs	No. up stream SNPs	No. ncRNA SNPs	No. Updown stream SNPs	Total gene related SNPs	Gene/Non-syn. SNP	No. Nonsyn SNPs	Gene/UTR5' SNPs	No. UTR5 SNPs	Gene/UTR3' SNPs	No. UTR3' SNPs	Highest significance value (SNP50)	Bp. of the SNP with the highest significance	Gene, closest to the highest significance SNP	Distance between the gene and the highest siginificance SNP
9	1	28926715	30828461	CH1	85	0	0	0	1	2	0	0	88		0		0		0	0.2689295	30036544	GBE1	394822
1	1	30843174	33095595	homozygous	0	0	0	0	0	0	0	0	0		0		0		0	0.4316797	32582468	VGLL3	2847864
26	1	33136013	33849425	CH2	0	0	0	0	0	0	0	0	0		0		0		0	0.3150566	33136013	VGLL3	2294319
27	1	78441352	80938504	CH2	258	8	0	0	1	1	0	0	268	LPP:NM_001192591:exon7:c.G910A:p.A304T	1		0		0	0.3720627	80938504	LPP	30140
2	1	80963023	86748745	homozygous	9	0	2	0	0	0	0	0	11		0		0	EIF4A2,EIF4A2,	2	0.3855527	81062374	LPP	154010
28	1	87774881	89771788	CH2	108	0	0	1	0	3	0	0	112		0	PEX5L	1		0	0.2615318	87774881	DNAJC19	318905
29	1	90839076	91721438	CH2	0	0	0	0	0	0	0	0	0		0		0		0	0.2506527	91187325	KCNMB2	384720
10	1	116539265	119740643	CH1	113	2	1	1	0	1	0	0	118	CLRN1:NM_001192101:exon2:c.C317T:p.T106I	1	GPR171	1	P2RY13,	1	0.2737163	118270530	P2RY12	3291
11	2	372648	4771836	CH1	15	2	5	0	1	0	0	0	23		0		0	IMP4,POLR2D, POLR2D,POLR2D, POLR2D,	5	0.2506527	3750890	HS6ST1	443106
12	2	19963404	25213239	CH1	415	6	6	2	1	10	0	0	440	CHRNA1:NM_176664:exon2:c.G52A:p.V18I, SCRN3:NM_001075547:exon5:c.C578T:p.T193M, ITGA6:NM_001109981:exon7:c.A1094G:p.K365R	3	KIAA1715,CHN1	2	CHRNA1,WIPF1, CDCA7,CDCA7, CDCA7,RAPGEF4	6	0.2615318	24418478	RAPGEF4	41418
30	2	131746089	131746089	CH2	0	0	0	0	0	0	0	0	0		0		0		0	0.2428198	131746089	LDLRAP1	25260
3	2	132134957	136766494	homozygous	9	4	0	0	0	0	0	0	13	TCEB3:NM_001102333:exon4:c.G624T:p.Q208H, TCEB3:NM_001102333:exon4:c.G490A:p.V164I	2		0		0	0.3798956	136680530	CAMK2N1	64799
31	2	136888949	136888949	CH2	0	0	0	0	0	0	0	0	0		0		0		0	0.3659704	136888949	PLA2G2F	147900
13	2	136946254	138931566	CH1	177	5	12	9	11	38	0	0	252	PLA2G2F:NM_001102522:exon1:c.G32T:p.G11V, PQLC2:NM_001080269:exon6:c.A592G:p.R198G, ALDH4A1:NM_001105646:exon6:c.A532G:p.M178V, AKR7A2:NM_001101949:exon2:c.C211T:p.Q71X	4	PLA2G2F,PLA2G2F, PLA2G5,PLA2G5, "LOC100125947,PLA2G2A", "LOC100125947,PLA2G2A", "LOC100125947,PLA2G2A", "LOC100125947,PLA2G2A", "LOC100125947,PLA2G2A"	9	PLA2G2F,PLA2G2F, PLA2G2F,PLA2G2F, PLA2G2F PLA2G2F,PLA2G2F, "PLA2G2D1,PLA2G2D3", "PLA2G2D1,PLA2G2D3", "PLA2G2D1,PLA2G2D3", "PLA2G2D1,PLA2G2D3", ALDH4A1	12	0.3803307	137256833	PLA2G2A	2457
32	4	19053103	22499221	CH2	22	2	0	2	0	0	0	0	26		0	ARL4A,ARL4A	2		0	0.2449956	22272736	ARL4A	697878
33	4	49699919	60292651	CH2	738	12	7	0	5	6	0	0	768	HBP1:NM_001046196:exon7:c.G761A:p.G254D, SLC26A3:NM_001083676:exon5:c.G466A:p.A156T, CFTR:NM_174018:exon11:c.A1402T:p.M468L, CFTR:NM_174018:exon9:c.G1185T:p.E395D,	4		0	HBP1,BCAP29, TXNDC3,STARD3NL, WNT2,WNT2, IMMP2L	7	0.3742385	53778230	MET	14946
14	4	100634712	102378316	CH1	269	6	0	0	8	2	0	0	285	LRGUK:NM_001192687:exon1:c.A591T:p.K20M, CALD1:NM_174258:exon4:c.C735G:p.F245L	2		0		0	0.2393386	101912397	AKR1B10	36605
34	5	79195620	85534666	CH2	337	20	11	1	19	5	0	0	393	NCF4:NM_001045983:exon10:c.G841C:p.E281Q, CSF2RB:NM_001192664:exon2:c.T41C:p.V14A, CSF2RB:NM_001192664:exon6:c.G595A:p.D199N, MPST:NM_001034291:exon2:c.C139T:p.R47C, MPST:NM_001034291:exon2:c.G385A:p.D129N, DNM1L:NM_001046494:exon7:c.T560C:p.V187A, LOC510651:NM_001101072:exon5:c.T4862A:p.L1621Q	7	MGC133880	1	RASD2,RASD2, RASD2,RASD2, LOC510193,LOC510193, LOC510193,TXN2, EIF3D,AMN1, AMN1	11	0.2528285	82792167	PKP2	95708

35	5	102991052	111840712	CH2	532	38	25	5	37	40	0	0	677	GPRC5A:NM_001034515:exon2:c.A268G:p.N90D, GPRC5A:NM_001034515:exon2:c.A256G:p.I86V, STYK1:NM_001143868:exon7:c.C856T:p.R286C, KLRA1:NM_174376:exon4:c.T250C:p.W84R, KLRA1:NM_174376:exon5:c.C503A:p.A168D, KLRA1:NM_174376:exon5:c.T536C:p.I179T, KLRA1:NM_174376:exon5:c.A550G:p.N184D, KLRA1:NM_174376:exon5:c.G557C:p.C186S, KLRA1:NM_174376:exon7:c.C697A:p.H233N, KLRA1:NM_174376:exon7:c.A783C:p.K261N, KLRJ1:NM_001002884:exon8:c.A707G:p.H236R, NKG2C:NM_001098163:exon1:c.C14T:p.T5M, NKG2C:NM_001098163:exon1:c.A32C:p.Q11P, NKG2C:NM_001098163:exon2:c.A86G:p.Q29R, NKG2C:NM_001098163:exon2:c.C134T:p.T45M, CLEC7A:NM_001031852:exon5:c.A589G:p.S197G, A2M:NM_001109795:exon8:c.G877C:p.E293Q, A2M:NM_001109795:exon14:c.A1678G:p.I560V, A2M:NM_001109795:exon14:c.C1694A:p.A565D, A2ML1:NM_001191301:exon1:c.C13T:p.L5F, MFAP5:NM_174386:exon2:c.G4A:p.A2T, AICDA:NM_001038682:exon2:c.C182T:p.S61F, WC1.3:NM_001190424:exon12:c.A2793C:p.R931S, WC1.3:NM_001190424:exon20:c.G4211T:p.G1404V	24	KLRA1,KLRA1, KLRA1,KLRA1, SCNN1A	5	HEBP1,HEBP1, HEBP1,GPRC5A, DDX47,DDX47, DDX47,APOLD1, APOLD1,STYK1, STYK1,STYK1, STYK1, MAGOHB,MAGOHB, MAGOHB,KLRA1, KLRA1,LOC618567, LOC618567,LOC618567, LOC618567,LOC618567, MRPL51,TAPBPL	25	0.4046997	105666728	TAS2R42	198542
36	5	112912610	115310536	CH2	118	2	2	0	2	0	0	1	125	TULP3:NM_001098992:exon4:c.G421A:p.A141T, C5H12orf32:NM_001075631:exon2:c.T95C:p.F32S	2		0	TULP3,C5H12orf32	2	0.2663185	114197492	CCDC77	4002
37	6	14655455	16338862	CH2	5	0	0	0	0	0	0	0	5		0		0		0	0.2402089	16145595	PITX2	108439
15	6	62196908	70336281	CH1	584	9	2	0	0	4	0	1	600	APBB2:NM_001076847:exon4:c.A541C:p.N181H, GABRG1:NM_001101250:exon1:c.T48G:p.N16K, CNGA1:NM_174278:exon8:c.C1682T:p.S561F	3		0	GABRG1,GABRG1	2	0.2989556	64023711	ATP8A1	38810
4	6	70403628	71660659	homozygous	3	0	0	0	0	0	0	0	3		0		0		0	0.3642298	71562444	SCFD2	3772
16	6	71909864	90075263	CH1	601	14	11	0	16	14	0	0	656	UGT2B10:NM_001046354:exon5:c.G1229A:p.G410E, SULT1B1:NM_001075823:exon6:c.G629A:p.R210Q, ODAM:NM_001080315:exon4:c.T160A:p.F54I, ODAM:NM_001080315:exon6:c.A520G:p.K174E, AMBN:NM_173988:exon7:c.T526C:p.S176P	5		0	SRP72,SRP72, SPINK2,SPINK2, TECRL,STAP1, SULT1E1, SULT1E1,SULT1E1, SULT1E1,HSTN	11	0.3629243	84174079	TECRL	922259
38	6	92124510	97644103	CH2	395	9	19	0	11	9	0	0	443	SCARB2:NM_001102153:exon6:c.G785A:p.R262K	1		0	RCHY1,RCHY1, THAP6,THAP6, CXCL10,CXCL10, SCARB2, SCARB2,SCARB2, SCARB2,SCARB2, SCARB2,SCARB2, SCARB2,CXCL13, CNOT6L,CNOT6L, CNOT6L,MRPL1	19	0.2650131	94601620	SCARB2	260299
39	7	37078577	43183117	CH2	194	14	5	1	11	19	0	0	244	CDHR2:NM_001192304:exon2:c.C82T:p.P28S, CDHR2:NM_001192304:exon7:c.G538A:p.V180I, CDHR2:NM_001192304:exon7:c.A584G:p.N195S, CDHR2:NM_001192304:exon18:c.A2636G:p.N879S, CDHR2:NM_001192304:exon29:c.C3733T:p.H1245Y, TRIM52:NM_001078024:exon1:c.A776T:p.D259V, MADCAM1:NM_001037821:exon3:c.A638G:p.H213R, MBD3:NM_001128505:exon5:c.T686C:p.V229A	8	RMND5B	1	TRIM52,TRIM52, SH3BP5L,MBD3, MBD3	5	0.259356	41428235	MGC137030	426299
17	7	40641856	43162929	CH1	52	3	3	0	4	2	0	0	64	MADCAM1:NM_001037821:exon3:c.A638G:p.H213R, MBD3:NM_001128505:exon5:c.T686C:p.V229A	2		0	SH3BP5L,MBD3,MBD3	3	0.3520453	42816139	MUM1	4426
40	7	44804506	45590074	CH2	38	1	6	1	3	3	0	0	52		0	C7H5orf24	1	C7H5ORF15,C7H5ORF15, VDAC1,VDAC1, SKP1,SAR1B	6	0.2328111	45321189	SAR1B	26

41	9	39369494	85474842	CH2	2165	33	23	1	42	30	0	0	2294	SLC22A16:NM_001076324:exon2:c.C281A:p.A94E, PREP:NM_174772:exon3:c.A235C:p.N79H, LOC782657:NM_001110094:exon1:c.A28T:p.S10C, MAP3K7:NM_001081595:exon10:c.A911G:p.H304R, SLC35A1:NM_001034637:exon2:c.G85A:p.V29I, C9H6orf163:NM_001079772:exon5:c.A614G:p.N205S, NT5E:NM_174129:exon8:c.A1526G:p.Q509R, NT5E:NM_174129:exon7:c.T1318C:p.Y440H, L3MBTL3:NM_001076871:exon7:c.T578C:p.V193A, L3MBTL3:NM_001076871:exon17:c.C1569A:p.D523E, ENPP3:NM_001075923:exon3:c.T218C:p.V73A, SLC2A12:NM_001011683:exon2:c.G1355A:p.R452Q, MMAB:NM_001079632:exon2:c.G666C:p.R222S, MMAB:NM_001079632:exon1:c.A356G:p.E119G, MMAB:NM_001079632:exon1:c.T310A:p.F104I	15	PLAGL1	1	DDO,CDC40, FIG4,NT5E, NT5E,NT5E, NT5E,NT5E, NT5E,ARG1, ARG1,ARG1, STX7,VNN2, VNN2,ALDH8A1, PDE7B,PEX7, PERP,PERP, HEBP2,HEBP2, SF3B5	23	0.3546562	79013299	REPS1	19940
42	9	86683885	106717246	CH2	2003	21	4	1	20	9	0	0	2058	LATS1:NM_001192866:exon2:c.G142T:p.A48S, RAET1G:NM_205785:exon1:c.G17T:p.R6I, ULBP17:NM_001168613:exon1:c.G17T:p.R6I, RAET1G:NM_205785:exon3:c.A386G:p.H129R, ULBP17:NM_001168613:exon3:c.A386G:p.H129R, RAET1G:NM_205785:exon3:c.C439G:p.Q147E, ULBP17:NM_001168613:exon3:c.C439G:p.Q147E, RAET1G:NM_205785:exon3:c.A440T:p.Q147L, ULBP17:NM_001168613:exon3:c.A440T:p.Q147L, RAET1G:NM_205785:exon3:c.G446T:p.S149I, ULBP17:NM_001168613:exon3:c.G446T:p.S149I, RAET1G:NM_205785:exon3:c.T461C:p.L154S, ULBP17:NM_001168613:exon3:c.T461C:p.L154S, RAET1G:NM_205785:exon3:c.G485A:p.G162D, ULBP17:NM_001168613:exon3:c.G485A:p.G162D, RAET1G:NM_205785:exon3:c.A489C:p.Q163H, ULBP17:NM_001168613:exon3:c.A489C:p.Q163H, RAET1G:NM_205785:exon3:c.T539C:p.M180T, ULBP17:NM_001168613:exon3:c.T539C:p.M180T, RAET1G:NM_205785:exon3:c.G540C:p.M180I, ULBP17:NM_001168613:exon3:c.G540C:p.M180I, RAET1G:NM_205785:exon3:c.G543C:p.E181D, ULBP17:NM_001168613:exon3:c.G543C:p.E181D, RAET1G:NM_205785:exon5:c.G818C:p.C273S, ULBP17:NM_001168613:exon5:c.G818C:p.C273S, RAET1G:NM_205785:exon5:c.G830A:p.C277Y, ULBP17:NM_001168613:exon5:c.G830A:p.C277Y, NOX3:NM_001191333:exon9:c.G932A:p.R311H, WTAP:NM_001113254:exon8:c.G739A:p.A247T, PLG:NM_173951:exon12:c.C1549A:p.P517T	17	RAET1G,ULBP17	1	PRPS1,WTAP, IGF2R,FGFR1OP	4	0.3259356	93337723	VIP	27142
43	10	49001707	61827419	CH2	595	10	13	1	13	8	3	0	643	CCNB2:NM_174264:exon3:c.G206A:p.S69N, RFX7:NM_001192820:exon9:c.T1750C:p.S584P, RFX7:NM_001192820:exon9:c.A2475T:p.E825D, MYO5C:NM_001192729:exon35:c.G4327A:p.D1443N, HDC:NM_001024551:exon2:c.G121A:p.A41T,	5	CYP19A1	1	ANXA2,ANXA2, MNS1,RFX7, CYP19A1,CYP19A1, CYP19A1, CYP19A1,CYP19A1, CYP19A1,CYP19A1, CYP19A1,TMOD3	13	0.3994778	52197725	ADAM10	3522
44	10	63224427	73324518	CH2	264	14	2	0	4	2	1	0	287	MYEF2:NM_001102085:exon13:c.C1277G:p.T426S, SPATA5L1:NM_001105625:exon7:c.A2083G:p.S695G, SPATA5L1:NM_001105625:exon7:c.G2050A:p.A684T, C10H14orf37:NM_001038060:exon2:c.G1150T:p.V384F, C10H14orf37:NM_001038060:exon2:c.G1027A:p.E343K, C10H14orf37:NM_001038060:exon2:c.T1022G:p.V341G	6		0	SEMA6D,SQRDL	2	0.3133159	68108176	BMP4	114062
5	11	19896588	20223876	homozygous	0	0	0	0	0	0	0	0	0		0		0		0	0.2885117	19896588	FEZ2	148275
18	11	21982224	22626344	CH1	29	0	0	1	0	1	0	0	31		0	MORN2	1		0	0.272846	22565107	MORN2	135767

19	13	29978704	33640015	CH1	609	16	2	0	8	2	0	0	637	CUBN:NM_001192575:exon66:c.G10753A:p.A3585T, CUBN:NM_001192575:exon64:c.A10277C:p.D3426A, CUBN:NM_001192575:exon64:c.A10243G:p.N3415D, CUBN:NM_001192575:exon61:c.G9813C:p.M3271I, CUBN:NM_001192575:exon49:c.A7565G:p.K2522R, CUBN:NM_001192575:exon27:c.C3860T:p.T1287I, CUBN:NM_001192575:exon10:c.G1039A:p.V347I, CUBN:NM_001192575:exon7:c.A667G:p.M223V, STAM:NM_001076842:exon14:c.G1396A:p.V466I, NSUN6:NM_001083670:exon8:c.G779T:p.G260V, ARHGAP12:NM_001102241:exon3:c.C353T:p.S118F	11		0	C13H10ORF97,ST8SIA6	2	0.2449956	30344377	C13H10ORF97	207721
6	14	2989275	4736994	homozygous	0	0	0	0	0	0	0	0	0		0		0		0	0.308094	4302229	LOC618755	283020
20	16	4274940	4718731	CH1	2	0	0	0	0	0	0	0	2		0		0		0	0.2637076	4718731	CFH	439650
21	16	5734127	8804868	CH1	0	1	0	0	0	0	0	0	1	R3HDM2:NM_001078034:exon1:c.C331A:p.Q111K,	1		0		0	0.2741514	7693002	R3HDM2	339647
45	16	60958039	64675532	CH2	346	14	12	0	7	1	0	0	380	DHX9:NM_174036:exon3:c.G247A:p.V83I, TSEN15:NM_001076312:exon4:c.C395T:p.S132L	2		0	GLUL,GLUL,GLUL, GLUL,GLUL,GLUL, GLUL,GLUL,GLUL, GLUL,SMG7,TSEN15	12	0.2645779	61058825	RGSL1	19358
22	16	71569808	75644283	CH1	326	12	1	1	4	9	0	0	353	LOC510385:NM_001105352:exon6:c.A806G:p.N269S, LOC511599:NM_001100341:exon2:c.T323G:p.F108C, LOC511599:NM_001100341:exon2:c.C387G:p.F129L, CD46:NM_183080:exon1:c.G98A:p.R33H, CRB1:NM_001192482:exon6:c.C1205T:p.S402L, CRB1:NM_001192482:exon9:c.A3694G:p.M1232V	6	HSD11B1	1	G0S2	1	0.3255004	73743439	CD34	32802
46	19	14279302	16900503	CH2	661	4	4	0	5	0	0	0	674	CCT6B:NM_001034642:exon11:c.C1250T:p.A417V,	1		0	FNDC8,CCL11,CCL11, CCL2	4	0.2715405	14700793	CCT6B	83596
47	21	3827974	8618090	CH2	82	1	2	0	1	0	0	0	86		0		0	MEF2A,MEF2A	2	0.2719756	6360252	TTC23	7723
23	21	65108809	68819863	CH1	199	4	0	1	13	5	0	3	225	TNFAIP2:NM_001191189:exon10:c.T1807G:p.S603A,	1	XRCC3	1		0	0.2576153	65990702	MIR379	10034
48	22	9665482	11980868	CH2	135	4	1	0	2	2	0	0	144	DLEC1:NM_001193198:exon19:c.A2786G:p.Q929R,	1		0	MYD88	1	0.2597911	11980868	ACVR2B	7208
7	24	32645084	44168238	homozygous	19	2	1	0	0	0	0	0	22	DLGAP1:NM_001192629:exon1:c.C214T:p.R72W,	1		0	RAB31	1	0.3559617	35443082	MIR1-2	136145
49	24	44255627	44255627	CH2	0	0	0	0	0	0	0	0	0		0		0		0	0.2872063	44255627	GNAL	12396
24	24	44306145	45601084	CH1	0	0	0	0	0	0	0	0	0		0		0		0	0.243255	44567434	CEP76	88452
8	24	45664478	47069405	homozygous	1	0	0	0	0	0	0	0	1		0		0		0	0.3220191	45664478	TNIP2	191481
25	24	47123716	49539373	CH1	57	0	1	0	2	1	0	0	61		0		0	IER3IP1	1	0.2789382	47123716	SIGLEC15	7118
Totals:					12570	293	183	29	252	229	4	5	13565		136								
												Grand total:	13079										

In yellow are intervals where allele frequencies differ in descendants and non-descendants of Chief (see Figure S7 for visualization and Table S8 for the list).

Table S8. Genes within 32 regions with the frequency of Chief alleles different between descendants and non-descendants

Region no.	chr	start	end	No. intron SNPs	No. exon SNPs	No. 3'UTR SNPs	No. 5'UTR SNPs	No. down stream SNPs	No. up stream SNPs	No. ncRNA SNPs	No. updown stream SNPs	Total gene related SNPs	Gene/Non-syn. SNPs	No. non-syn. SNPs	Gene/5'UTR SNPs	No. 5'UTR SNPs	Gene/3'UTR SNPs	No. 3'UTR SNPs	Highest Significance Value (SNP50)	Bp of the SNP with highest significance value	Gene, closest to the highest significance SNP	Distance between the gene and the highest siginificance SNP
1	1	67755812	76434294	686	10	7	4	4	14	0	0	725	CCDC14:NM_001105469:exon1:c.A775G:p.N259D, UMPS:NM_177508:exon3:c.C541T:p.R181C, ITGB5:NM_174679:exon13:c.C2089G:p.Q697E, BDH1:NM_001034600:exon5:c.C368T:p.A123V, CPN2:NM_001101232:exon2:c.T995C:p.V332A	5	TNK2,OSTALPHA,MFI2,BDH1	4	ITGB5,ITGB5,ITGB5,ITGB5,ITGB5,APOD,CPN2	7	0.000790396	68371871	DIRC2	74583
2	1	79972527	80938504	78	6	0	0	0	0	0	0	84	LPP:NM_001192591:exon7:c.G910A:p.A304T	1		0		0	0.000917886	79972527	TPRG1	68501
3	3	32874728	33402672	14	1	2	1	0	1	0	0	19	ST7L:NM_001083481:exon1:c.C74T:p.P25L	1	CTTNBP2NL	1	CAPZA1,CTTNBP2NL	2	0.000870123	33402672	CTTNBP2NL	8145
4	3	38753131	39339310	91	0	0	0	0	0	0	0	91		0		0		0	0.000874784	38753131	MGC139448	65598
5	4	49699919	60292651	738	12	7	0	5	6	0	0	768	HBP1:NM_001046196:exon7:c.G761A:p.G254D, SLC26A3:NM_001083676:exon5:c.G466A:p.A156T, CFTR:NM_174018:exon11:c.A1402T:p.M468L, CFTR:NM_174018:exon9:c.G1185T:p.E395D	4		0	HBP1,BCAP29,TXNDC3,STARD3NL,WNT2, WNT2,IMMP2L	7	0.000505471	50062728	PIK3CG	79565
6	4	80698120	83032965	17	0	1	0	0	0	0	0	18		0		0	C4H7orf25	1	0.000699175	83032965	LOC100125578	52100
7	5	36051712	36173811	0	0	0	0	0	0	0	0	0		0		0		0	7.69E-05	36173811	FAM113B	53698
8	5	73892593	77167629	255	1	5	1	4	8	0	0	274		0	MGC137014	1	TCP11L2,C5H12orf23,PWP1,FBXO7,MGC137014	5	0.00044741	73964925	LOC100270756	71498
9	5	92502082	94578774	249	1	0	0	0	1	0	0	251		0		0		0	0.000888919	94578774	CMAS	13022
10	5	104022455	107367506	210	26	20	4	32	38	0	0	330	GPRC5A:NM_001034515:exon2:c.A268G:p.N90D, GPRC5A:NM_001034515:exon2:c.A256G:p.I86V, STYK1:NM_001143868:exon7:c.C856T:p.R286C, KLRA1:NM_174376:exon4:c.T250C:p.W84R, KLRA1:NM_174376:exon5:c.C503A:p.A168D, KLRA1:NM_174376:exon5:c.T536C:p.I179T, KLRA1:NM_174376:exon5:c.A550G:p.N184D, KLRA1:NM_174376:exon5:c.G557C:p.C186S, KLRA1:NM_174376:exon7:c.C697A:p.H233N, KLRA1:NM_174376:exon7:c.A783C:p.K261N, KLRJ1:NM_001002884:exon8:c.A707G:p.H236R, NKG2C:NM_001098163:exon1:c.C14T:p.T5M, NKG2C:NM_001098163:exon1:c.A32C:p.Q11P, NKG2C:NM_001098163:exon2:c.A86G:p.Q29R, NKG2C:NM_001098163:exon2:c.C134T:p.T45M, CLEC7A:NM_001031852:exon5:c.A589G:p.S197G	16	KLRA1,KLRA1,KLRA1,KLRA1	4	GPRC5A,DDX47,DDX47,DDX47,APOLD1,APOLD1, STYK1,STYK1,STYK1,STYK1,MAGOHB,MAGOHB, MAGOHB,KLRA1,KLRA1,LOC618567,LOC618567, LOC618567,LOC618567,LOC618567,LOC618567	20	0.000820483	105288653	ETV6	58718
11	5	115473164	115509870	0	0	0	0	0	0	0	0	0		0		0		0	0.000608795	115509870	DCP1B	242872
12	5	120911371	121661107	64	1	0	0	3	3	0	0	71	PARVB:NM_001102299:exon4:c.C386T:p.P129L	1		0		0	0.000874784	121661107	PARVG	67179
13	5	122735170	125588039	384	6	3	0	6	1	1	0	401	LOC537366:NM_001076093:exon2:c.G20A:p.R7Q, TTC38:NM_001105259:exon7:c.G622T:p.A208S	2		0	FBLN1,PPARA,LOC537366	3	0.00082929	124715380	TBC1D22A	556275
14	6	41506939	42838773	28	0	0	0	2	3	0	0	33		0		0		0	0.000824875	41506939	KCNIP4	15976
15	6	54699009	57806018	0	0	0	0	0	0	0	0	0		0		0		0	0.000568133	54699009	PCDH7	2698188
16	6	62196908	67395437	435	8	2	0	0	2	0	0	447	APBB2:NM_001076847:exon4:c.A541C:p.N181H, GABRG1:NM_001101250:exon1:c.T48G:p.N16K	2		0	GABRG1,GABRG1	2	0.000670006	67244198	GABRA2	45275
17	6	90184757	97005042	431	11	20	0	14	13	0	0	489	COX18:NM_001082437:exon1:c.G56A:p.R19H, SCARB2:NM_001102153:exon6:c.G785A:p.R262K	2		0	IL8,RCHY1,RCHY1,THAP6,THAP6,CXCL10, CXCL10,SCARB2,SCARB2,SCARB2,SCARB2, SCARB2,SCARB2,SCARB2,SCARB2,CXCL13, CNOT6L,CNOT6L,CNOT6L,MRPL1	20	0.00089847	91417374	ALB	43790
18	9	99911918	103251233	1105	4	1	0	3	0	0	0	1113	PLG:NM_173951:exon12:c.C1549A:p.P517T	1		0	IGF2R	1	0.000932726	100524835	PLG	50901
19	10	13369708	15006348	36	0	0	0	1	2	0	0	39		0		0		0	0.00078619	15006348	FEM1B	71452
20	10	56586773	57638140	0	0	0	0	0	0	0	0	0		0		0		0	0.000903285	57638140	BANF1	720256
21	10	63023941	63201598	0	0	0	0	0	0	0	0	0		0		0		0	0.000571161	63170778	FBN1	162234
22	11	92077934	92609678	0	0	0	0	0	0	0	0	0		0		0		0	0.000484459	92466053	RSAD2	394071

23	16	60958039	64502613	293	14	12	0	7	1	0	0	327	DHX9:NM_174036:exon3:c.G247A:p.V83I, TSEN15:NM_001076312:exon4:c.C395T:p.S132L	2		0	GLUL,GLUL,GLUL,GLUL,GLUL,GLUL,GLUL,GLUL, GLUL,GLUL,SMG7,TSEN15	12	0.000908125	64502613	HMCN1	50137
24	16	73117624	74126360	8	1	0	0	0	2	0	0	11	CD46:NM_183080:exon1:c.G98A:p.R33H	1		0		0	0.000874784	74126360	F13B	124904
25	16	75441234	75566135	0	0	0	0	0	0	0	0	0		0		0		0	0.000521833	75566135	LHX9	348133
26	23	27078160	28890490	330	45	28	1	29	23	0	0	457	TNXB:NM_174703:exon2:c.C167T:p.A56V, TNXB:NM_174703:exon3:c.C500T:p.A167V, TNXB:NM_174703:exon5:c.G1249C:p.E417Q, G7C:NM_001075206:exon2:c.T25C:p.C9R, LY6G6F:NM_001076194:exon3:c.C392T:p.S131F, GPANK1:NM_001076468:exon3:c.C224T:p.T75M, TUBB2B:NM_001046549:exon4:c.A959G:p.Q320R, BOLA:NM_001040532:exon6:c.A970G:p.T324A, BOLA:NM_001038518:exon5:c.A995G:p.H332R, BOLA:NM_001040532:exon5:c.A710G:p.E237G, BOLA:NM_001040532:exon4:c.T546G:p.S182R, BOLA:NM_001038518:exon3:c.T571G:p.W191G, BOLA:NM_001040532:exon4:c.G535A:p.G179S, BOLA:NM_001038518:exon3:c.G560A:p.R187Q, BOLA:NM_001038518:exon3:c.C559G:p.R187G, BOLA:NM_001038518:exon3:c.C538T:p.R180C, BOLA:NM_001040532:exon4:c.G511C:p.D171H, BOLA:NM_001038518:exon3:c.G536C:p.R179T, BOLA:NM_001038518:exon3:c.G526A:p.E176K, BOLA:NM_001040532:exon4:c.G499A:p.V167M, BOLA:NM_001038518:exon3:c.G524A:p.G175D, TRIM40:NM_001103306:exon2:c.C316A:p.P106T	17	CLIC1	1	MGC151586,CLIC1,C6orf25,LY6G6C,LTA,LTA, ATP6V1G2,ATP6V1G2,ATP6V1G2,TCF19, TCF19,SFTA2,NRM,KIAA1949,KIAA1949,BOLA, BOLA,BOLA,BOLA,BOLA,BOLA,BOLA,BOLA,BOLA, BOLA,BOLA,BOLA,TRIM40	28	0.000842681	28890490	MOG	3064
27	23	31741848	32015190	1	0	0	0	1	2	0	0	4		0		0		0	0.000893682	32015190	H4	8438
28	23	34410374	34469180	0	0	0	0	0	0	0	0	0		0		0		0	0.00452179	34469180	PRP4	341198
29	24	26339920	26339920	0	0	0	0	0	0	0	0	0		0		0		0	0.000737449	26339920	RNF138	140089
30	24	44255627	44255627	0	0	0	0	0	0	0	0	0		0		0		0	0.00083373	44255627	GNAL	12396
31	24	47123716	49539373	57	0	1	0	2	1	0	0	61		0		0	IER3IP1,	1	0.00078619	49539373	SMAD2	128612
32	27	26021842	26545747	0	0	0	0	0	0	0	0	0		0		0		0	0.000710441	26149598	C27H8ORF79	94517
Total				5510	147	109	11	113	121	1	0	6013		55								
										Grand total:		5779										

In yellow are intervals overlapping with the 49 selected regions (see Table S7).

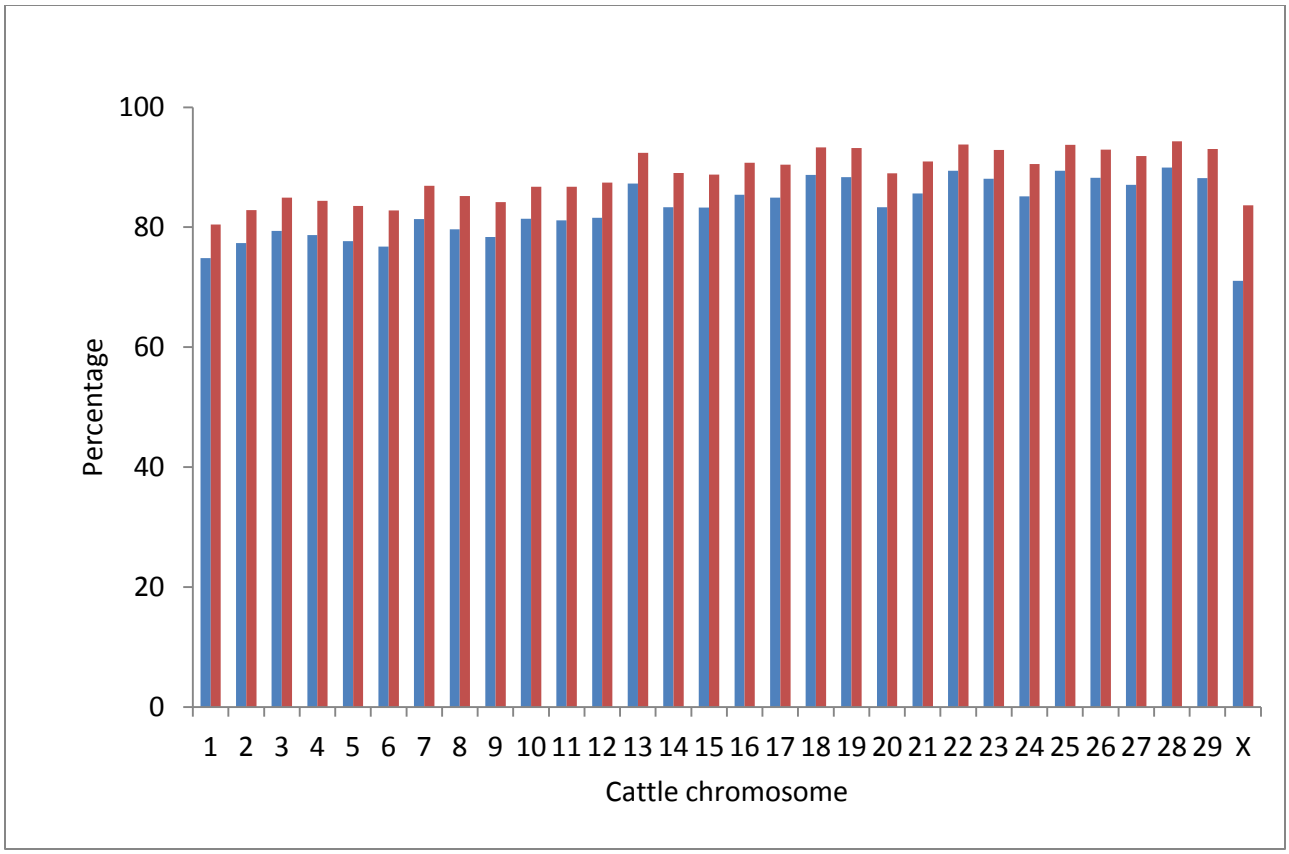

Figure S1. Coverage (percentage) of the reference chromosomes with Mark (red) and Chief (blue) reads for the intervals that had reads mapped by BLAT or Newbler.

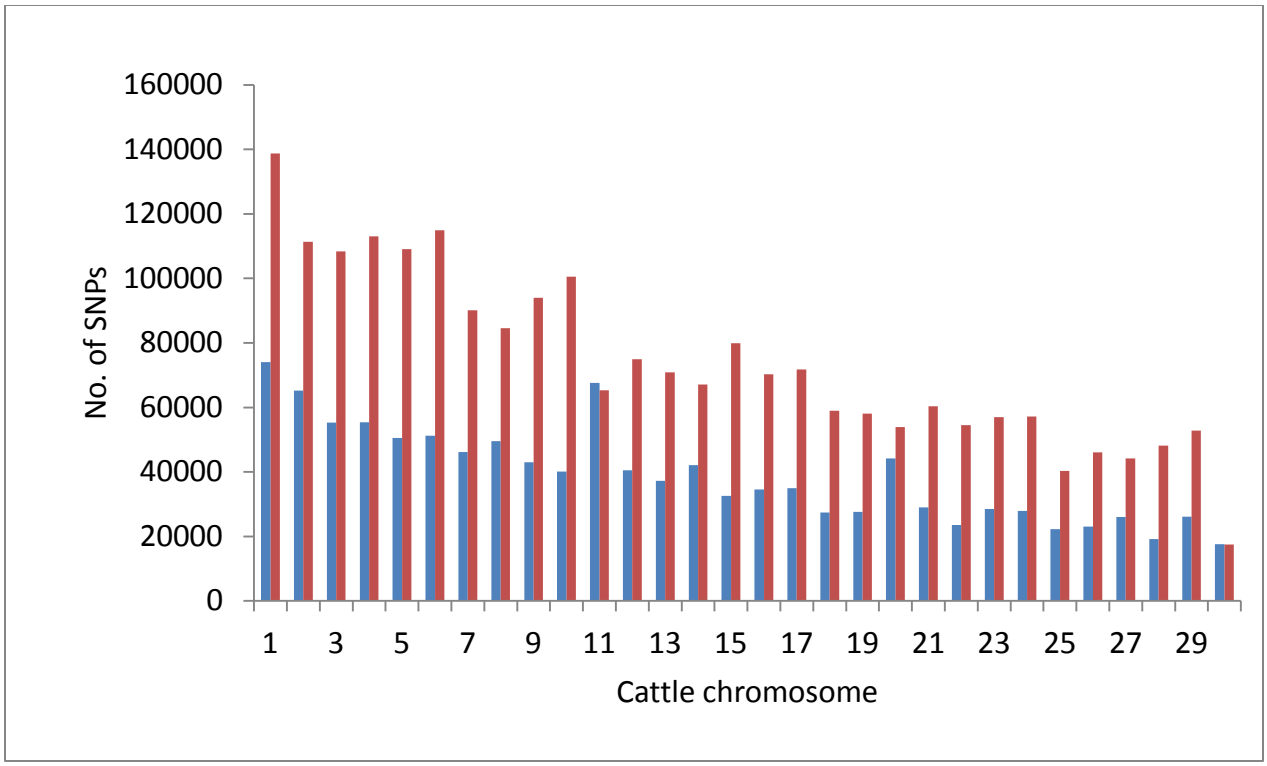

Figure S2. Distribution of high-quality Mark and Chief in silico SNPs across reference bovine chromosomes. In red is shown the distribution of SNPs heterozygous in at least one individual and in blue are SNPs that are homozygous in both sires for an allele different from the reference genome (Btau 4.0).

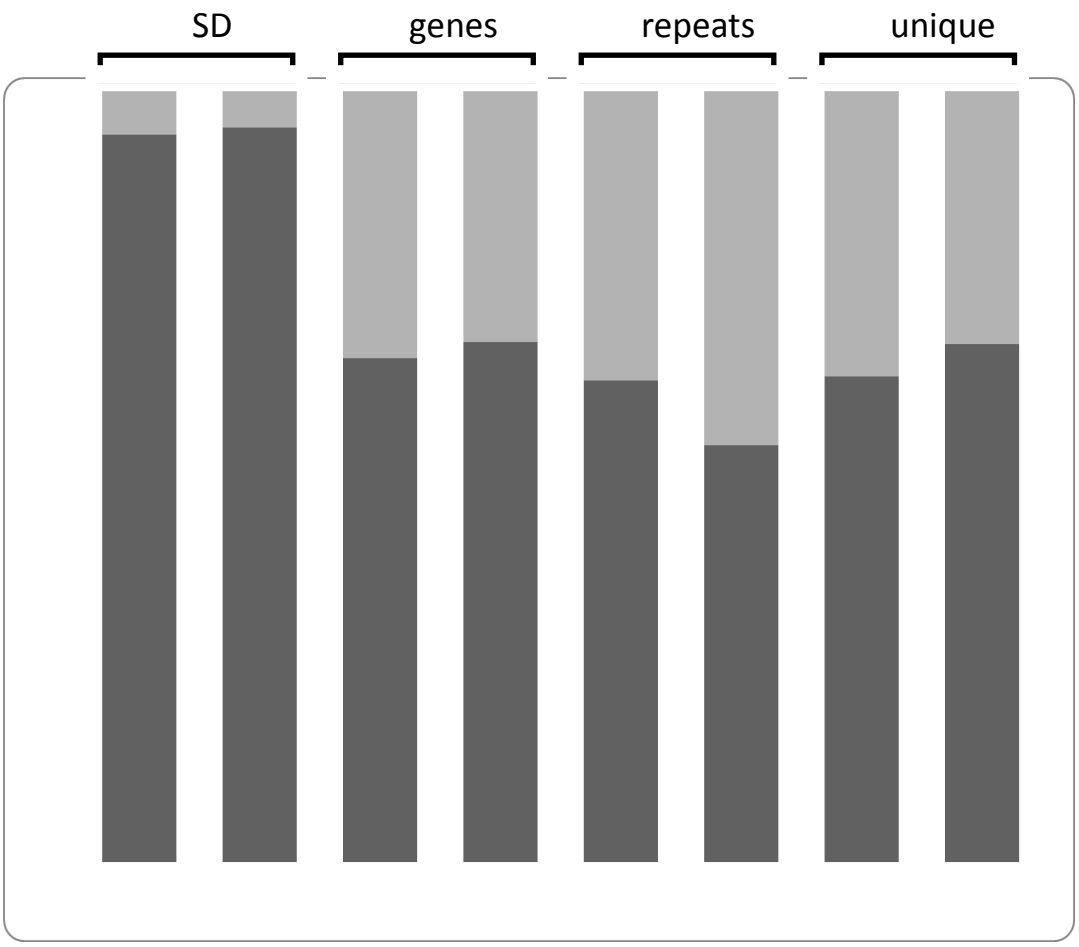

Figure S3. The reference Btau4.0 genome was divided into four classes: genes, unique, segmental duplications (SD), and repetitive elements. The GLEAN gene prediction collection defined the genes. RepeatMasker output defined the repetitive element category. The whole-genome SD from Liu et al., 2009 defined duplications. All remaining sequence was designated "unique." An overlap between the filtered set of SNPs homozygous in two bulls and SNPs heterozygous in the bulls was calculated for each class separately and shown (lighter gray) as the percentage of the total number of heterozygous and homozygous SNPs.

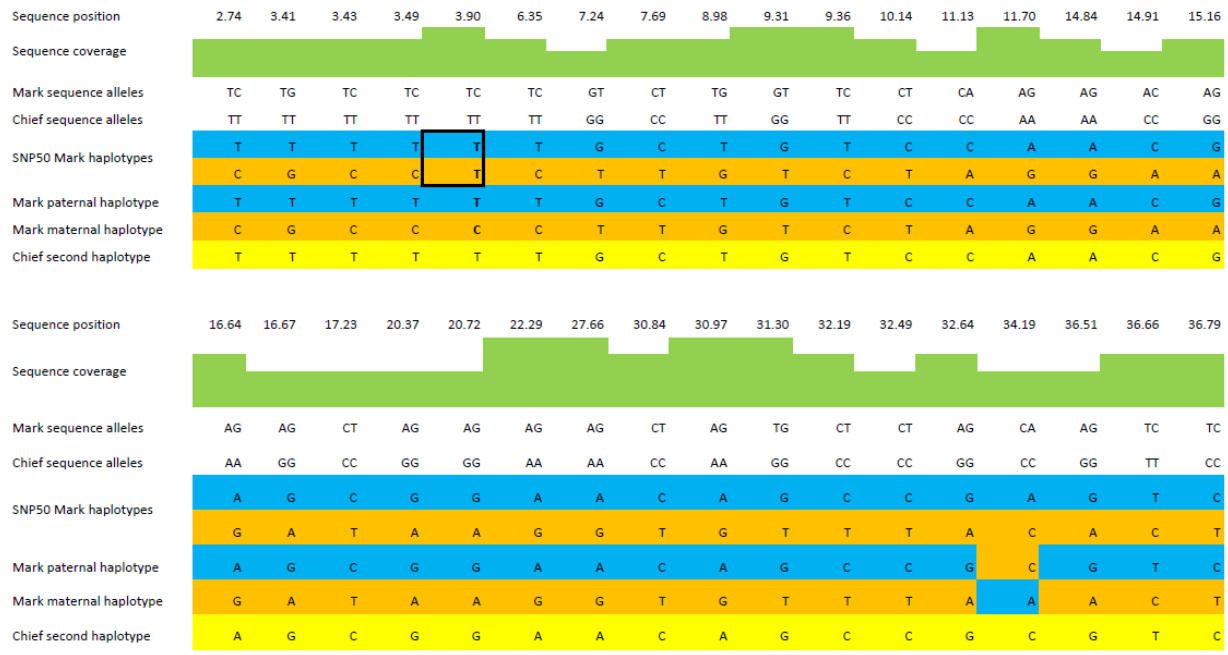

Figure S4. Reconstruction of Mark and Chief haplotypes from the sequence data and comparison to Mark haplotypes inferred from the genotyping of his 92 offspring. An example shows BTA2 position 2.74-36.79 Mbp. For simplicity only SNPs that overlap between the sequence data and the SNP50 Illumina Bovine array are shown. Coverage track indicates variation in coverage within the interval; sequence allele track shows the alleles that were detected in the filtered set of SNPs. Mark haplotypes are shown based on the genotyping reconstruction and the sequence data reconstruction. Differences between the two reconstructions are caused by a probable miss of alleles: one at 3.90 Mbp and another at 34.19 Mbp in Chief's sequence data. In both cases discrepancies between the sequence data and genotyping data result in differences between Mark haplotypes resolved by two independent methods**.**

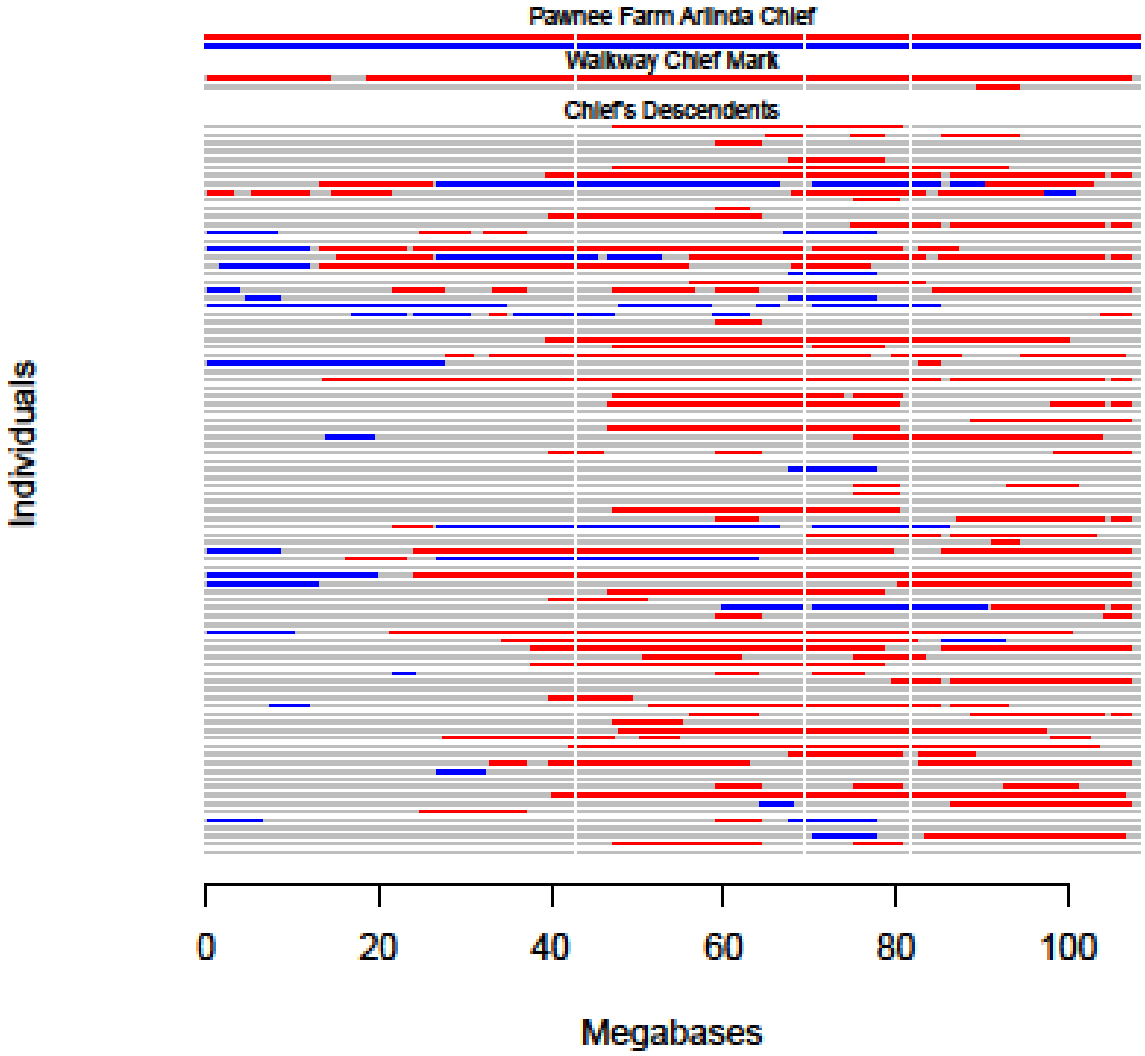

Figure S5. BTA9. Segments in common between Chief, his son Mark, and 50 of Chief's direct descendents in the current Holstein population (at least six generations removed) determined using the chromoPhase algorithm. Chief haplotype 1 segments are in red and the alternative haplotype is in blue. Each pair of lines represents alternative haplotypes of an individual. In gray are descendent haplotypes different from either of Chief's haplotype.

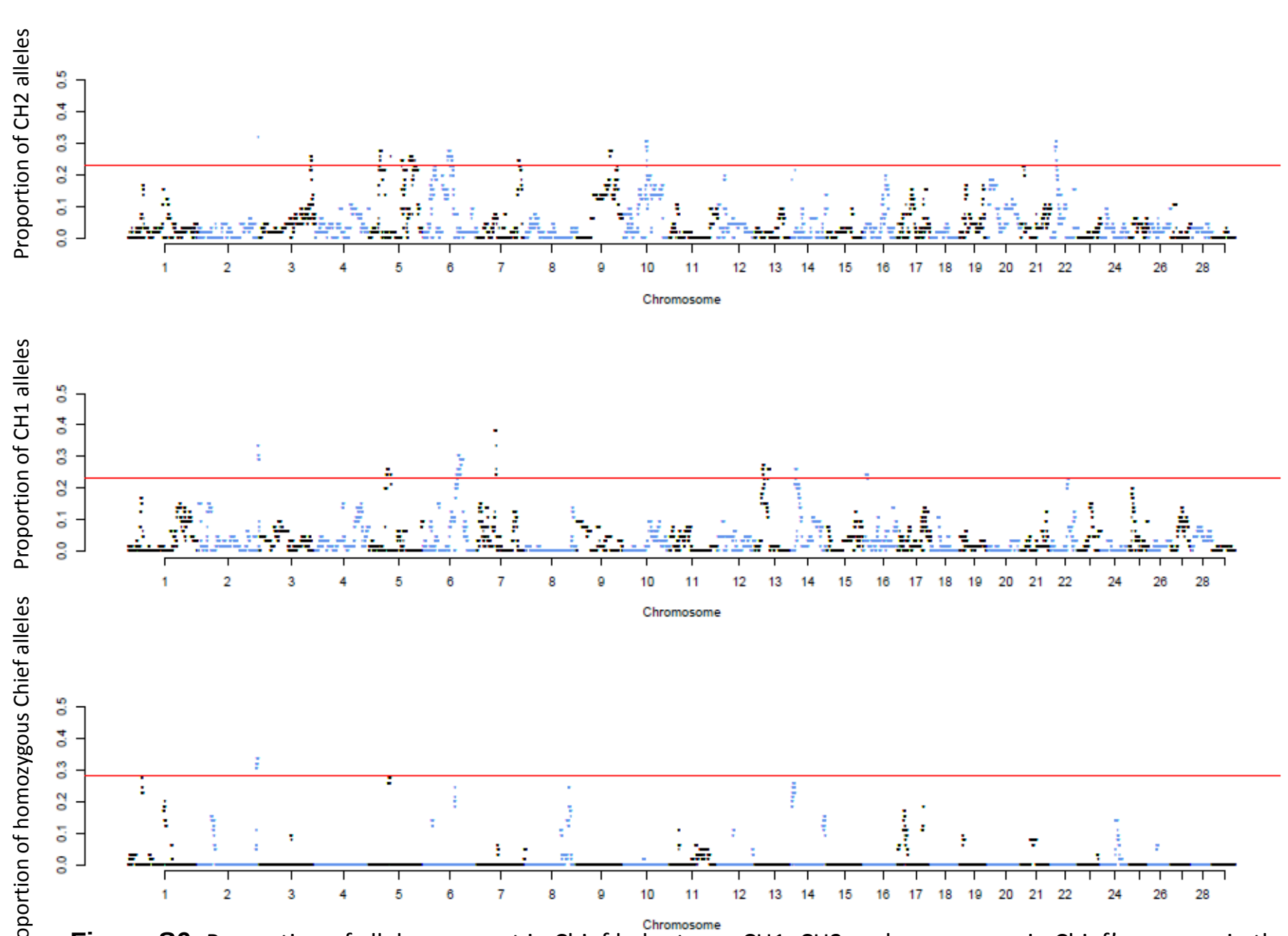

Figure S6. Proportion of alleles present in Chief halpotypes CH1, CH2, or homozygous in Chief's genome in the set of 33 Holstein individuals unrelated to Chief. Red line indicates a threshold at which the fraction of SNPs present in Chief genome could be explained by a random distribution in individuals unrelated to Chief.

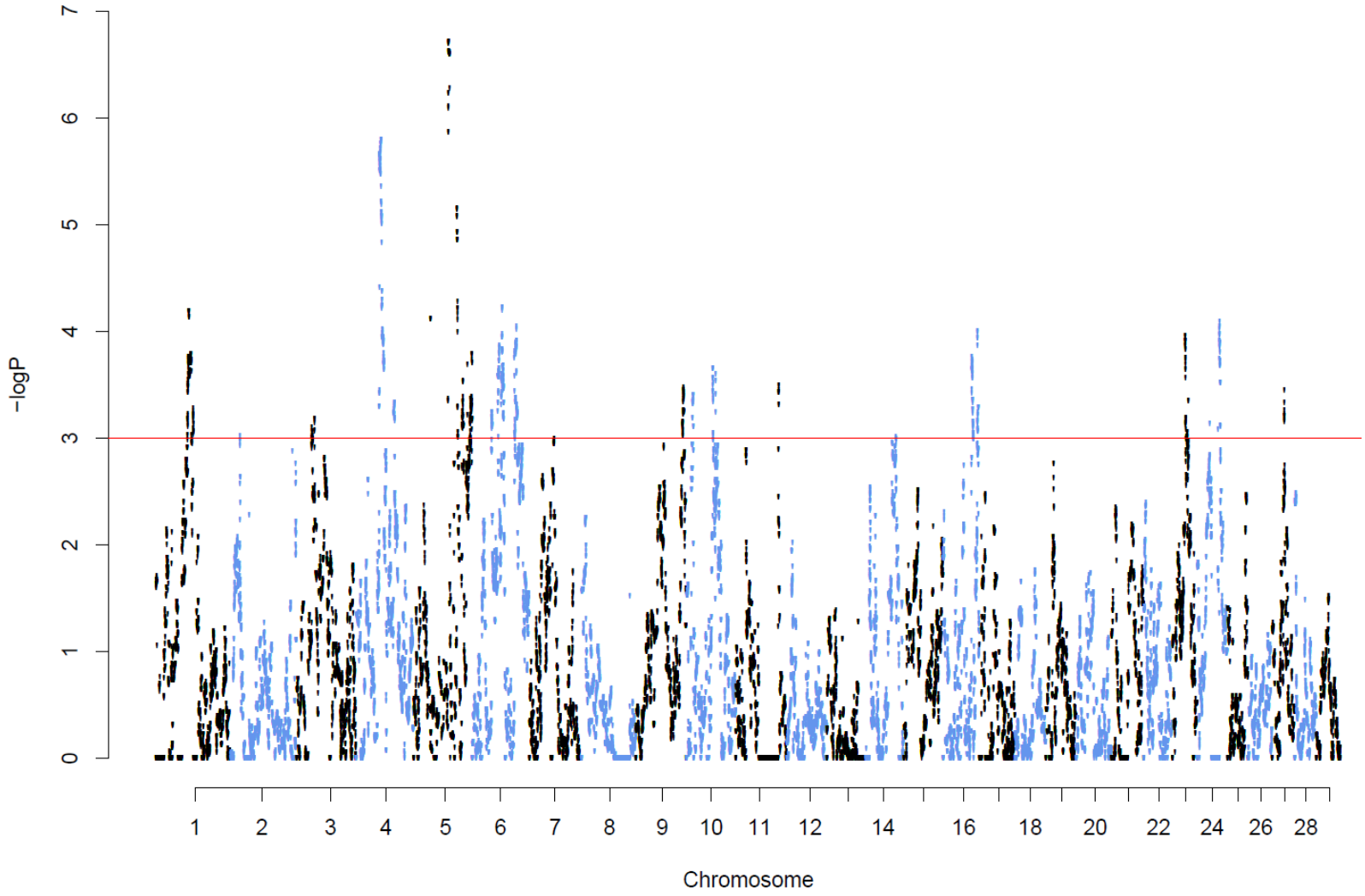

Figure S7. A plot of the $-\text{Log}_{10}$(P-values) of the interaction term from the linear model for allele frequency difference between Chief descendants and non-descendants.

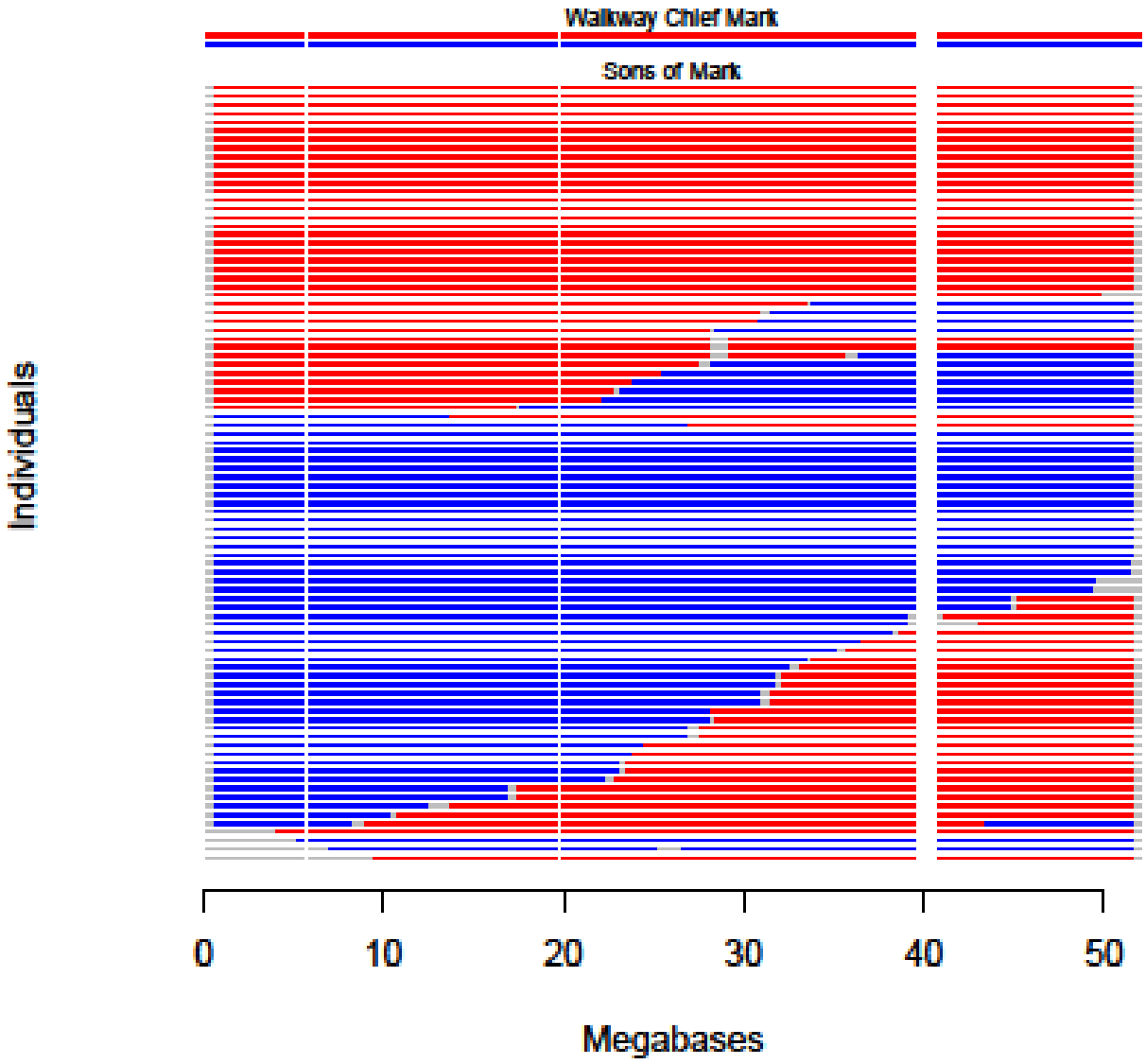

Figure S8. BTA29. Segments in common between Mark, and 92 of his sons' paternally-inherited chromosomes. Mark's paternal (Chief's) segments are in red and Mark's maternal segments are in blue. Each line represents a paternally-inherited haplotype of a son.

References

1. Marth GT*, et al.* (1999) A general approach to single-nucleotide polymorphism discovery. *Nat Genet* 23:452-456.
2. Chevreux B*, et al.* (2004) Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res* 14:1147-1159.
3. Zimin AV*, et al.* (2009) A whole-genome assembly of the domestic cow, Bos taurus. *Genome Biol* 10:R42.