Selection for complex traits leaves little or no classic signatures of selection

Kathryn E Kemper1*, Sarah J Saxton1,2, Sunduimijid Bolormaa3, Benjamin J Hayes3,4,5 and Michael E Goddard1,3

Abstract

Background: Selection signatures aim to identify genomic regions underlying recent adaptations in populations. However, the effects of selection in the genome are difficult to distinguish from random processes, such as genetic drift. Often associations between selection signatures and selected variants for complex traits is assumed even though this is rarely (if ever) tested. In this paper, we use 8 breeds of domestic cattle under strong artificial selection to investigate if selection signatures are co-located in genomic regions which are likely to be under selection.

Results: Our approaches to identify selection signatures (haplotype heterozygosity, integrated haplotype score and $F_{ST}$) identified strong and recent selection near many loci with mutations affecting simple traits under strong selection, such as coat colour. However, there was little evidence for a genome-wide association between strong selection signatures and regions affecting complex traits under selection, such as milk yield in dairy cattle. Even identifying selection signatures near some major loci was hindered by factors including allelic heterogeneity, selection for ancestral alleles and interactions with nearby selected loci.

Conclusions: Selection signatures detect loci with large effects under strong selection. However, the methodology is often assumed to also detect loci affecting complex traits where the selection pressure at an individual locus is weak. We present empirical evidence to suggest little discernible ‘selection signature’ for complex traits in the genome of dairy cattle despite very strong and recent artificial selection.

Background

Evolutionary change in a population, in response to a change in environment, consists of an increase in the frequency of favourable mutations. If the mutation was recent and the selection is strong, all alleles on the same chromosome segment as the mutant allele will increase in frequency by hitchhiking, generating a characteristic selection sweep or selection signature [1]. On the other hand, if selection at individual loci is weak or if the mutation is old, and therefore part of the standing variation when selection commences, little evidence of the selection may be left in the genome e.g. [2]. Many statistics have been proposed to detect signatures of selection but they all suffer from a severe problem – the distribution of the statistic under the null hypothesis of no selection is usually unknown. This is because the distribution depends on the demography of the population, including changes in effective population size and migration, which are difficult to define. Consequently, no formal test that a statistic comes from the null distribution is possible. Generally, the most extreme values of the statistic are simply assumed to be due to selection and there have been many papers claiming to find evidence for signatures of selection. The evidence for selection sweeps at a small number of loci, such as for lactase persistence in humans [3] and skin wrinkling in Shar-Pei dogs [4], is well documented and convincing, but in other cases it is hard to evaluate the strength of evidence. Certainly the evidence and persuasiveness of authors advocating adaptation via standing polymorphisms is increasing [5-7] and the influential paradigm of ‘hard sweep’ selection signatures is beginning to lose favour as the primary mechanism of adaptation [1].

In this study we have taken a different approach – we study sites in the genome at which we know selection has occurred to see if a signature of selection has been left behind. By studying a variety of selected loci, we are able to describe when a selection signature is generated and when it is not. Domestic cattle have been under
quite strong, recent and well documented selection for several traits and hence their genomes should contain evidence of this selection. We use 8 domestic Bos taurus cattle breeds and three types of loci which have been under selection: type 1 loci are genes that are part of the definition of a breed, such as absence of horns and coat colour; type 2 loci have a large effect on quantitative traits, such as stature and milk yield, and type 3 loci are quantitative trait loci (QTL) for milk production traits in dairy cattle. We consider two statistics that indicate selection signatures within a breed and $F_{ST}$ (which indicates a difference between breeds in a segment of the genome that could be caused by different selection histories between the breeds). Our results show clear signatures of selection when intense selection has been applied to a single locus because it causes a trait defining the breed such as coat colour. However, we find weak evidence for selection signatures at regions of the genome associated with complex traits under selection. This paper calls into question the reliability of selection signatures to identify mutations affecting complex traits under selection and provides empirical evidence for the ability to generate substantial genetic change between populations in complex traits without clear evidence for classic selection signatures.

Results

Measures of selection

The dataset consists of 23,641 domestic cattle with > 610,123 (real or imputed) genome-wide autosomal SNP from 8 Bos taurus breeds. Breeds were of European origin and have had previous, recent selection for milk (Holstein, Jersey) or meat (Angus, Charolais, Hereford, Limousin, Murray Grey, Shorthorn) production. There were between 61 (Limousin) and 13,501 (Holstein) animals genotyped per breed.

Three statistics were calculated to test for evidence of selection: a modified version of Depaulis-Veuille’s H-test (referred to as haplotype homozygosity, $HAPH$) [8], the integrated haplotype score ($|iHS|$) [4], and Wright’s measure of population differentiation ($F_{ST}$). The measure of haplotype homozygosity ($HAPH$) measures selection within breed and is defined as the variance of haplotypes frequencies at a particular position in the genome, i.e. $\Sigma (p_i - \frac{1}{N})^2$, where $p_i$ is the (within breed) frequency of the $i$th haplotype and $N$ is the total number of haplotypes at the position. The haplotypes consist of 30 or 31 consecutive SNPs. This statistic is high if one or more haplotypes are at high frequency while most haplotypes exist at low frequency. Similarly, $|iHS|$ identifies within breed selection and SNP where one allele is found on one or few long haplotypes whereas the other allele is associated with many haplotypes. Both $HAPH$ and $|iHS|$ are efficient for identification of sweeps which have not yet reached fixation, an essential feature for an association with type 3 loci (i.e. genomic regions with segregating mutations for complex traits under selection). In contrast, the $F_{ST}$ measurement is most efficient when there are large allele frequency differences between pairs of breeds. Selection is indicated by high values of $F_{ST}$ near the selected mutations because, for example, a population in which selection has taken place is expected to differ from other populations (that have not undergone the same selection) in the allele frequency for markers near the mutation.

The 3 measures of selection were calculated in 250 kb windows across the genome, where the value for each window was the mean $HAPH$, the maximum observed $|iHS|$ or the average between breed $F_{ST}$. To correct for average differences within and between breeds for $HAPH$ and $F_{ST}$ the values are standardised by dividing the window value by the mean value for all windows. Consequently the standardised estimates of selection have a mean of 1. $|iHS|$ was calculated following [9], and is thus standardised such that $|iHS|$ can be interpreted as standard deviations from the mean. The estimates (per window) of $HAPH$, $|iHS|$ and breed comparisons for $F_{ST}$ are given in Additional file 1 (where Additional file 2 provides definitions of the columns for Additional file 1). We examined the 5% of the genome with the strongest evidence for selection.

Breed-defining loci (type 1) and large effect QTL (type 2) were identified from the literature and the Online Inheritance in Animals database [10]. For type 3 loci, we used the Holstein and Jersey breeds to identify QTL regions in the genome for milk production traits using the ‘genomic selection’ methodology [11]. These two breeds have been under strong selection for milk production for at least the last 100 years [12] and especially since the 1970’s (Additional file 3: Figure S1-S3). In genomic selection, the prediction of genetic merit is a linear regression in which each SNP genotype is multiplied by the estimated effect of a SNP and summed to yield an estimated breeding value (EBV) for the animal. In our case, we want to attach variation in the trait to each chromosome segment. Thus we estimated the effect of each SNP using the genomic selection methodology and then calculated the variance across animals for a local 250 kb EBV e.g. [13]. The 5% of windows with the highest variance were considered to have QTL and defined as type 3 loci.

Breed-defining loci often showed selection signatures

There were 5 loci that control phenotypes which are characteristics of the breed. These loci are polled (i.e. absence of horns) and 4 loci (MC1-R, PMEL, KIT and KITLG) that determine coat colour (Table 1). Most of these loci (including POLLED, MC1-R, KIT and PMEL) have previously been reported as under selection e.g.
and we find evidence for all loci of within breed selection using HAPH (Table 2).

There is evidence for more than one selected mutation at each of the type 1 loci. This evidence includes selection within 2 or more breeds but large $F_{ST}$ between these selected breeds as well as between each selected breed and the breeds not selected at this gene. For example, near POLLED we found within breed selection signatures (i.e. top 5% of window HAPH values) for Limousin, Charolais, Angus, Holstein, Hereford, Murray Grey and Shorthorn and across-breed differentiation (i.e. top 5% of $F_{ST}$ values) for Holstein with Angus, Murray Grey and Limousin (Figure 1). This is consistent with the 2 different reported mutations for POLLED [18,19], where the $P_{C}$ allele segregates in Angus, Charolais, Limousin and Hereford and the $P_{F}$ allele segregates in Holstein. Selection signatures near POLLED in Western European cattle are also thought to pre-date $P_{c}$ mutation [18], indicating the possibility of further (as yet undescribed) alleles. We also propose allelic heterogeneity for PMEL in Charolais and Murray Grey cattle, where both breeds show strong within breed selection using HAPH but a large value of $F_{ST}$ between them (Additional file 3: Figure S5). Different PMEL mutations are known to

| Locus | Location | Description |
|-------|----------|-------------|
| POLLED | BTA1 1.71 Mbp | Determines the presence and absence of horns. Two identified alleles: $P_{C}$ (Celtic-origin) a 212 bp insertion-deletion at 1.706 Mbp; and $P_{F}$ (Holstein Friesian-origin) which segregates as a 260 kb haplotype (from 1.649 – 1.989 Mbp) in Holstein and Jersey [18,19]. No known associated gene. Most domestic cattle are horned but Angus and Murray Grey breeds are exclusively polled and the POLLED locus segregates in other breeds. |
| MC1-R | BTA18 14.75 Mbp | The main determinant of coat colour in cattle [20]. Two identified alleles: $E^{D}$ (p.L99P) which produces a black coat; and $e$ (inducing a premature stop codon) which is recessive produces a red coat when homozygous [21]. |
| PMEL | BTA5 57.67 Mbp | Coat colour dilution mutation (c.64G > A) identified in Charolais [22]. Different PMEL mutations segregate in Highland and Charolais cattle [23]. |
| KIT | BTA6 71.85 Mbp | Locus associated with piebald colour in Hereford [24] and degree of white-spotting in Holstein [25]. No known causative mutations but the different coat colour patterns in these breeds, suggests different KIT mutations. |
| KITLG | BTA5 18.34 Mbp | A SNP mutation (p.A193D) identified in Shorthorn and Belgian Blue as causative for the roan phenotype [26]. KITLG is also associated with pigmentation surrounding the eyes in Fleckvieh cattle [27]. |

Table 1 Description of (type 1) loci with a large effects on breed-defining traits, such as coat colour, in in domestic cattle and likely to be segregating in our populations

| Locus | Location | Description |
|-------|----------|-------------|
| POLLED | BTA1 1.71 Mbp | Determines the presence and absence of horns. Two identified alleles: $P_{C}$ (Celtic-origin) a 212 bp insertion-deletion at 1.706 Mbp; and $P_{F}$ (Holstein Friesian-origin) which segregates as a 260 kb haplotype (from 1.649 – 1.989 Mbp) in Holstein and Jersey [18,19]. No known associated gene. Most domestic cattle are horned but Angus and Murray Grey breeds are exclusively polled and the POLLED locus segregates in other breeds. |
| MC1-R | BTA18 14.75 Mbp | The main determinant of coat colour in cattle [20]. Two identified alleles: $E^{D}$ (p.L99P) which produces a black coat; and $e$ (inducing a premature stop codon) which is recessive produces a red coat when homozygous [21]. |
| PMEL | BTA5 57.67 Mbp | Coat colour dilution mutation (c.64G > A) identified in Charolais [22]. Different PMEL mutations segregate in Highland and Charolais cattle [23]. |
| KIT | BTA6 71.85 Mbp | Locus associated with piebald colour in Hereford [24] and degree of white-spotting in Holstein [25]. No known causative mutations but the different coat colour patterns in these breeds, suggests different KIT mutations. |
| KITLG | BTA5 18.34 Mbp | A SNP mutation (p.A193D) identified in Shorthorn and Belgian Blue as causative for the roan phenotype [26]. KITLG is also associated with pigmentation surrounding the eyes in Fleckvieh cattle [27]. |

Table 2 Evidence for within and between breed selection at breed-defining (type 1) loci

| Locus | Within breed | Evidence for selection | Differentiation between breeds |
|-------|--------------|------------------------|--------------------------------|
| POLLED | Angus$^1$$^2$, Charolais$^1$$^2$, Holstein$^1$$^2$, Limousin$^1$$^2$, Hereford$^1$, Shorthorn$^1$ | 1. Holstein with Angus, Murray Grey and Limousin | Figure 1 |
| MC1-R | Limousin$^1$$^2$, Charolais$^1$, Angus$^1$, Holstein$^1$, Murray Grey$^1$ | 1. Breeds with black ($E^{D}$) allele (Holstein, Angus, Murray Grey) with breeds with recessive red (e) allele (Charolais, Limousin, Shorthorn, Hereford) 2. Jersey (E$^{D}$ allele) with all other breeds, except Hereford | Additional file 3: Figure S4 |
| PMEL | Charolais$^1$$^2$, Angus$^2$, Murray Grey$^1$ | 1. Charolais with all other breeds 2. Murray Grey with all breeds, excluding Jersey 3. Shorthorn and Jersey | Additional file 3: Figure S5 |
| KIT | Hereford$^1$$^2$, Holstein$^1$ | 1. Hereford with all other breeds. 2. Holstein with all breeds, except Jersey 3. Shorthorn with all breeds, except Jersey 4. Jersey with Angus, Charolais and Limousin | Additional file 3: Figure S6 |
| KITLG | Hereford$^1$ | 1. Hereford will all other breeds, except Murray Grey 2. Murray Grey and Charolais with each other, and with Holstein, Angus and Limousin 3. Shorthorn with Augus | Additional file 3: Figure S7 |

$^1$windows encompassing loci and identified in the top 5% of within or between breed measures of selection. Measures of selection were $^2$haplotype homozygosity (HAPH), $^3$integrated haplotype score (|iHS|) and $^4$F$_{ST}$. |
segregate in Charolais and Scottish Highland cattle [23], and here it appears the Charolais mutation is also different to a PMEL mutation in Murray Grey.

The observed frequency of the selected haplotype played an important role in determining the ability of the three test statistic to indicate selection. At POLLED, for example, neither HAPH nor |iHS| indicated evidence of within breed selection in Murray Grey despite all animals of this breed being polled. This is because this region is homozygous in Murray Grey and neither of these statistics indicates selection in homozygous regions, being either undefined (|iHS|) or with values close to zero (HAPH). Further at PMEL, long selected haplotypes were indicated by HAPH and $F_{ST}$ in Murray Grey but there was no |iHS| selection signature near the locus. The results show that $F_{ST}$ is most efficient when the region is near fixation (homozygous) in alternate breeds, |iHS| is most efficient for intermediate frequency (or segregating) variants [9] and HAPH is midway between the two measures.

The mode of action and favoured phenotype also determined if loci indicated selection. In Shorthorn, for example, there was no within breed selection signature near KITLG despite a roan coat (where white hairs are intermingled with coloured hairs) being a characteristic of this breed [26]. This can be explained by balancing selection, where heterozygotes express the roan phenotype and homozygotes have either a solid coloured or white coat, which would not be efficiently detected by any method. There was also evidence for a within breed selection near KITLG in Hereford. Herefords do not have a roan phenotype and, considering results in Fleckvieh cattle [27], this may indicate that a KITLG mutation contributes to the characteristic white spotting pattern seen in Hereford and Fleckvieh.

Selection at known loci affecting quantitative traits
There were 5 type 2 loci chosen which had large effects mutations on stature (PLAG1), milk production (DGAT1,
GHR, ABCG2) and muscle mass (MSTN) (Table 3). These loci were examined for the presence of selection signatures and, for DGAT1, GHR and ABCG2, to confirm their effect on milk production (Table 4). Selection signatures indicating selection in dairy, as compared to beef, breeds have previously been reported for GHR and ABCG2 [14,28], while other loci (PLAG1, DGAT1 and MSTN) have previous reported selection signatures e.g. [17,29,30].

We find evidence for selection signatures near all type 2 loci, but the evidence had greater ambiguity than for the breed-defining (type 1) loci in most cases. The notable exception was at MSTN, where there was clear evidence of recent and strong selection in the Limousin breed (Table 4, Additional file 3: Figure S11). The other loci showed more ambiguous patterns of selection. In the case of ABCG2 and GHR, this was likely to be because selection signatures were affected by several mutations in the region. For example, near ABCG2 there is a strong selection signature in Charolais, probably due to selection at the LCORL or NCPAG locus [17,42], and there appears to be several QTL for milk production traits in BTA20 near GHR [43]. In other cases, such as PLAG1, a more complex pattern of selection arises (Figure 2). For instance, Limousin differ from other breeds for most windows in the region except a window centred near LYN and incorporating PLAG1. Limousin seem to have the same haplotype as other breeds in the immediate LYN-PLAG1 region but different in the surrounding region. This could be explained if the mutation was introduced into Limousin from another breed and one hybrid haplotype became the common ancestor for most Limousin haplotypes in the region.

Aligning selection signatures and QTL in dairy cattle was also not always straightforward. Sometimes this was because alleles did not segregate within the dairy breeds and sometimes because recent selection was for the ancestral (rather than the derived) allele. For example, there was no stature QTL for Holstein or Jersey near PLAG1 because Jerseys have a high frequency of the ancestral allele and Holstein have a high frequency of the (proposed) mutant allele [31]. Further, our QTL results confirm the segregation of the DGAT1 mutation in both dairy breeds (Jersey and Holstein) but DGAT1 showed within breed selection signatures only in the beef breeds. It is possible that selection some time ago for the mutant allele (in both dairy and beef cattle) because it increased milk volume but more recent selection in Jersey and Holstein has been for the ancestral allele because it increases milk fat. Thus the recent selection in dairy breeds is not detected within either Jerseys or Holsteins because selection has been for the ancestral allele which is likely to be carried on a variety of haplotype backgrounds and so is unlikely to show a discernable selection signature.

Has selection for milk production left selection signatures in dairy cattle? Type 3 loci are regions of the genome which show genetic variation in Holstein and Jersey cattle for 7 different production traits (fat, milk and protein yield; stature; fertility; and percentage of fat and protein in milk). Most of these traits have been under strong recent selection (Additional file 3: Figure S1-S3). We used a chi-squared test to investigate if there was greater overlap, than expected by chance, between the windows identified as containing QTL (i.e. type 3 loci, top 5% of windows with the highest variance) and windows identified with selection signatures (i.e. top 5% of HAPh, |iHS| or FST values). The within breed measures of selection (HAPh, |iHS|) assess haplotype frequencies and should be efficient at detecting on-going recent selection while, in contrast, high FST between dairy by beef breeds will identify areas

Table 3 Description of (type 2) loci with large effects on complex traits under selection in domestic cattle, such as milk and meat yield, and likely to be segregating in our populations

| Locus | Location | Description |
|-------|----------|-------------|
| PLAG1 | BTA14 25.00 Mbp | Region affecting many traits, including stature (31) and fertility (29). Originally identified in Jersey-Holstein cross, Jersey are thought to be near fixation for the ancestral allele while Holstein and other breeds are near fixation for the alternate allele (29,32). |
| DGAT1 | BTA14 1.80 Mbp | Dinucleotide substitution causing a lysine to alanine substitution (p.K232A) (33), where the mutant A allele decreases fat yield, and increases protein yield and milk volume (34,35). The mutant DGAT1 allele is at high frequency or fixed in Hereford, Angus and Charolais; and at lower frequencies in Holstein and Jersey (35). |
| GHR | BTA20 32.05 Mbp | A SNP mutation causing a missense phenylalanine to tyrosine substitution (p.F279Y). Effects on milk volume and composition (36). |
| ABCG2 | BTA6 37.97 Mbp | A SNP mutation causes a missense tyrosine to serine (p.Y581S) mutation which increases milk yield and decreases milk solids (37). Identified in Israeli Holsteins where the frequency of the ABCG2 allele had increased in response to selection for milk yield and then decreased when selection changed to focus on increased milk solids (37). The ABCG2 allele is at low frequency (< 10%) in US and German Holsteins, Angus, British Frisian, Charolais and Hereford (38). |
| MSTN | BTA2 6.22 Mbp | A negative regulator of muscle development, multiple mutations have been described that cause ‘double muscling’ or extreme muscular hypertrophy (32,39,40). In Limousin, a mutation associated with a mild increase in muscling, F94L, has been identified (41). |
of the genome where there is differentiation between dairy and beef breeds, but not within either group. Overall, there was a relatively weak association between QTL and selection signatures (Table 5). There was evidence for an association between \(|iHS|\) and QTL for protein yield in Holstein and between \(|iHS|\) and QTL for stature in Jersey (\(P < 0.05\), Bonferroni corrected). There were 1.6 and 1.8 times the number of windows with QTL and high \(|iHS|\) than expected by chance. There was no association between selection as measured by \(HAP\) or dairy-beef \(F_{ST}\) and any traits. This is despite the strong correlation between \(|iHS|\) and \(HAP\), where 2.8 and 5 times more windows were identified in the top 5% of \(HAP\) and \(|iHS|\) than expected by chance (for Holstein and Jersey respectively). Increasing the proportion of the genome considered to contain QTL and showing selection signatures did lead to a weak association between selection signatures and QTL. For example, the number of windows in top 20% for \(|iHS|\) and QTL variance was about 1.15 times the number expected by chance for all traits, with the exception of fat and protein percentage in milk for Jersey. This weak association was nevertheless significant (\(P < 0.05\), Bonferroni corrected). Thus our data supports weak selection across many loci for most production traits.

Windows with high \(F_{ST}\) values between beef and dairy breeds were not enriched for QTL affecting production traits (Table 5) even when the proportion of the genome considered was increased to 20%. Thus despite many generations of selection for increased milk production in dairy cattle, we do not find big differences in allele frequency between beef and dairy breeds near QTL for milk production. This may indicate that genetic drift between beef and dairy breeds is greater than the effects of selection. Our finding are in contrast to other studies [28], which used fewer SNP and fewer breeds than in the current analysis. However, windows containing QTL in Holstein were significantly over-represented (by 1.8 - 2.1 times) in the windows with QTL for the same trait in Jersey (Bonferroni corrected; \(P < 0.05\)), for all traits except fertility. Thus at least some QTL appear to segregate in both breeds. If the same alleles segregate in both breeds, this implies that either the polymorphisms existed since before the breeds diverged or it may be the result of admixture among our dairy cattle populations. Given that some QTL segregate across breeds, it is perhaps surprising that selection has not caused both dairy breeds to differ from the beef breeds as measured by \(F_{ST}\).

Novel regions with strong selection sweeps in the genome

It is possible that selection has operated for traits other than those reported in Table 5 so we considered the
Evidence of selection was indicated by extreme (top 5%) values for haplotype homozygosity (HAPH), the integrated haplotype score (|iHS|) and Wright's measure of population differentiation (FST).

Traits analysed for type 3 loci are: fat yield (FAT, kg per lactation), milk yield (MILK, litres per lactation), protein yield (PROT, kg per lactation), stature (STAT), fertility (FERT, calving interval), FPC (fat percentage in milk) and PPC (protein percentage in milk).

**Table 5** Association between measures of selection and genome-wide quantitative trait loci (i.e. type 3 loci) in Holstein and Jersey cattle

|                 | FAT | MILK | PROT | STAT | FERT | FPC | PPC |
|-----------------|-----|------|------|------|------|-----|-----|
| (a) QTL Holstein| HAPH Holstein | 31.4 | 32.8 | 35.6 | 34.0 | 31.0 | 30.6 | 30.8 |
| (b) QTL Jersey  | HAPH Jersey | 21.4 | 25.2 | 29.2 | 22.6 | 20.4 | 16.6 | 19.8 |
| (c) QTL Holstein| | 40.0 | 39.0 | 47.0<sup>*</sup> | 40.2 | 39.6 | 36.0 | 34.6 |
| (d) QTL Jersey  | | 31.8 | 36.4 | 35.0 | 43.0<sup>*</sup> | 34.2 | 28.6 | 21.7 |
| (e) QTL Holstein or Jersey | F<sub>ST</sub> Dairy vs. Beef | 55.2 | 47.0 | 48.0 | 42.6 | 44.0 | 44.0 | 45.8 |
| (f) QTL Holstein| QTL Jersey | 46.0<sup>*</sup> | 47.6<sup>*</sup> | 47.6<sup>*</sup> | 51.6<sup>*</sup> | 34.2 | 50.4<sup>*</sup> | 55.2<sup>*</sup> |

<sup>Chi-squared test P < 0.05, Bonferroni corrected P-value.</sup>

Values are the average number of windows showing both selection and type 3 loci for production traits in either Holstein or Jersey cattle (a-e) across 5 sets of 250 kb windows. Also shown is the number of overlapping windows with type 3 loci in both Holstein and Jersey (f). There are approximately 32 (a-d, f) and 46 (e) windows expected by chance. Additional file 3: Tables S1-S3 contain the full chi-squared tests.
overall prevalence of strong selection signatures in the genomes for the 8 cattle breeds. Based on long regions of high \textit{HAPH}, there were a total of 190 regions which contained windows from the top 5\% of within breed selected windows and were greater than 2 Mbp in length (Additional file 3: Figure S12) and 25 cases where sweeps were > 5 Mbp (Table 6).

Six of the 25 long regions of high \textit{HAPH} could be ascribed to the type 1 and type 2 loci. The strong selection sweep on BTA13 in Shorthorn contains the agouti (\textit{ASIP}) locus (Table 6), which is known to affect coat colour in several species [20]. However, phenotypic expression of \textit{ASIP} requires an agouti-susceptible allele at \textit{MC1-R}, such as the wild-type \textit{E}\textsuperscript{+} allele found in Jerseys [44]. Thus most of our other breeds will not show a coat colour phenotype from \textit{ASIP} mutations. There seems to be a selected mutation specific to British breeds (i.e. Shorthorn, Angus, Murray Grey and Hereford; Additional file 3: Figure S13) and, although \textit{ASIP} mutations are unlikely to affect coat colour in these cattle, the locus may have affected coat colour in ancestors without the \textit{MC1-R} mutation or the mutation may affected other traits such as fatness and homeostasis [45].

Other strong selection sweeps for several breeds were located on BTA 16 (41 – 47 Mbp) and BTA 7 (42 – 47 Mbp) (Table 6). However, unlike the \textit{ASIP} region, \textit{FST} in these two regions did not indicate clear differentiation patterns between the breeds and breeds within the selected group frequently differed from each other. The selected region on BTA7 was particularly gene dense and includes, among others, 23 olfactory receptor loci. Interestingly, this region was also identified in an independent study of Fleckvieh cattle [46]. The large sweep identified in Shorthorn on BTA3 (69.75 – 88.4 Mbp) contains \textit{LEPR} (leptin receptor, 80.1 Mbp) which has been reported to be associated with multiple growth and fatness traits in beef cattle [47]. The longest identified selected region in Holstein, where we had the largest number of genotyped animals (n = 13,501), was on BTA26. In a region also supported by a high \textit{iHS} value, a promising candidate is \textit{FGF8} (fibroblast growth factor 8 (androgen-induced)) (Additional file 3: Figure S14). There is functional evidence for the involvement of \textit{FGF8} in lactation, as it has been found to be highly expressed in lactating (human) breast tissue and milk [48]. The selection signature on BTA3 was also identified by Stella \textit{et al.} [15]. The region contains \textit{SLC35A3} (solute carrier family 35 (UDP-GlcA) transporter, member A3; at 43.4 Mbp) which is the gene at which a recessive lethal mutation causes complex vertebral malformations (CVM) in Holstein cattle [49]. A lethal recessive mutation would not cause the type of selection signature detected here but selection at a nearby linked locus could explain why the mutation in \textit{SLC35A3} has drifted to high frequency.

Some of the long selection sweeps reported in Table 6 could be the result of random processes, such as genetic drift or demographic changes, rather than selection. However, we find that strong selection (or ‘hard’) sweeps are relatively rare in our 8 breeds of cattle. This is despite strong, recent selection for numerous traits and particularly for milk production traits in our dairy breeds. Thus one can conclude the substantial genetic improvement in milk yield in dairy cattle has not generated many clear signatures of selection.

### Table 6 Genomic regions with evidence of recent selection using haplotype homozygosity

| Breed     | BTA | Sweep location & size (Mbp) | Type 1 & 2 loci |
|-----------|-----|-----------------------------|-----------------|
|           |     | Beginning | End | Length |                        |
| Limousin  | 2   | 0         | 13.85 | 13.85 | \textit{MSTN}           |
| Hereford  | 2   | 68.85     | 74.95 | 6.1   | \textit{PMEL}           |
| Jersey    | 3   | 38.15     | 47.8  | 9.65  | \textit{PMEL}           |
| Jersey    | 3   | 50.95     | 57.7  | 6.75  | \textit{PMEL}           |
| Shorthorn | 3   | 69.75     | 88.4  | 18.65 | \textit{PMEL}           |
| Angus     | 3   | 89.6      | 94.65 | 5.05  | \textit{PMEL}           |
| Shorthorn | 4   | 67.15     | 73    | 5.85  | \textit{PMEL}           |
| Murray Grey | 5   | 40.65     | 61.8  | 21.15 | \textit{PMEL}           |
| Charolais | 5   | 52.8      | 64.75 | 11.95 | \textit{PMEL}           |
| Hereford  | 6   | 67.85     | 79.35 | 11.5  | \textit{PMEL}           |
| Jersey    | 7   | 36.3      | 48.45 | 12.15 | \textit{PMEL}           |
| Angus     | 7   | 42.3      | 47.75 | 5.45  | \textit{PMEL}           |
| Shorthorn | 11  | 34.1      | 40.65 | 6.55  | \textit{PMEL}           |
| Shorthorn | 13  | 57.45     | 66.45 | 9     | \textit{PMEL}           |
| Charolais | 14  | 19.75     | 29.55 | 9.8   | \textit{PMEL}           |
| Angus     | 16  | 38.5      | 47.75 | 9.25  | \textit{PMEL}           |
| Shorthorn | 16  | 39.65     | 48.85 | 9.2   | \textit{PMEL}           |
| Holstein  | 16  | 40.1      | 47.05 | 6.95  | \textit{PMEL}           |
| Charolais | 16  | 41.45     | 46.9  | 5.45  | \textit{PMEL}           |
| Jersey    | 20  | 1.5       | 7.1   | 5.6   | \textit{PMEL}           |
| Jersey    | 20  | 22.8      | 29    | 6.2   | \textit{PMEL}           |
| Holstein  | 20  | 29.85     | 34.9  | 5.05  | \textit{PMEL}           |
| Murray Grey | 22  | 33.2      | 39.45 | 6.25  | \textit{PMEL}           |
| Murray Grey | 24  | 22.35     | 29.35 | 7     | \textit{PMEL}           |
| Holstein  | 26  | 17.6      | 24.3  | 6.7   | \textit{PMEL}           |

Discussion

We searched for selection signatures at locations in the genome which were likely to be under selection using dense SNP genotypes in the genomes of 8 domestic \textit{B. taurus} cattle breeds. The evidence is consistent with one or more mutant alleles having been selected to high frequency in some of the eight breeds for some of loci we investigated. Consistent with a ‘hard sweep’ model of
selection, the breeds carrying the mutant allele show a common long haplotype (indicated by high values of $HAPH$) and a large genetic distance ($F_{ST}$) from the breeds carrying the ancestral allele or a different mutant allele in the region. We clearly observed this type of selection pattern at $PMEL$ and $MSTN$. However, selection signatures at loci with a large effect on complex traits under selection (type 2 loci) were weaker, and almost absent for most QTL for traits under selection (type 3 loci). How can these results be explained?

A classic ‘hard sweep’ is expected when the environment changes such that a mutation that would previously been detrimental becomes favourable. Typically there is a lag and then the frequency of the favoured allele increases slowly until it reaches a modest frequency after which it is swept quickly to fixation. This is the pattern seen, for instance, in insecticide resistance [50]. Our data on $POLLED$, $MC1-R$, $KIT$, $KITLG$, $PMEL$, $PLAG1$ and $MSTN$ are consistent with this explanation although here the changed ‘environment’ is one in which cattle owners control which animals will be allowed to breed. The selected mutations were probably deleterious in the wild and this natural selection may still operate in domestic cattle along with the artificial selection applied by cattle owners. Therefore to drive a mutation rapidly to high frequency, artificial selection must be strong and natural selection weak. This combination is likely for some coat colour mutations – if a breed is defined to be red, then selection for a red mutation will be very strong while natural selection against the mutation may be weak, particularly if natural selection was related to environmental factors that have been reduced through the process of domestication (i.e. camouflaged from predators).

On the other hand, mutations with a large effect on growth, reproduction or milk production are likely to have detrimental side effects even under domestication. Pleiotropy is commonly observed for large-effect mutations, such as $PLAG1$ affecting fertility and stature [29] or $DGAT1$ affecting both milk volume and solids (fat and protein) [33], and it is unlikely that the overall effect of a particular mutation would always be weak. Consequently, few mutations affecting these types of traits will be driven rapidly to high frequency and leave a clear selection signature. Occasionally large-effect mutations with small or inconspicuous pleiotropic effects are observed as under strong selection. We observed strong selection in Limousin at $MSTN$ and there is strong, recent selection near the $PLAG1$ region in Brahman cattle despite its negative effects on fertility [29].

Thus the results for type 1, 2 and 3 loci are best reconciled by considering the selection on each locus. Selection for simple (monogenic) traits applies strong selection pressure to a mutation and the results are consistent with a ‘hard sweep’ model of selection. However, complex traits in our data were not associated with classic selection signatures and ‘hard sweeps’ are relatively rare despite the recent selection for milk traits in our dairy cattle. This suggests the selection response is caused by weak selection at many sites across the genome, probably for previously segregating variants. Weak selection is expected since each QTL has a small effect the on phenotype e.g. [51,52]. Since there are many loci, each with small effect, selection will not change the allele frequency rapidly and there will be little evidence of a selection sweep. Small changes to allele frequencies at many loci can combine to make large changes to a phenotype, consistent with the large selection response observed for the complex traits in our data. The ability to detect selection sweeps would be further hampered if selection was conducted on genetic variants already segregating in the population. Innan & Kim [53], for example, find the initial frequency of the selected alleles to be one of the primary determinants for the ability to detect a selection event using classic selection signatures.

The explanation of weak selection on old genetic variation for complex traits, although speculative, is supported by other evidence. One key and consistent observation in support of selection on standing variants is the rapid and immediate response to selection observed for most (if not all) heritable characters in domestic and experimental populations [54]. This supports frequency changes to mutations already segregating in the population because, given the rapid response, there is insufficient time for accumulation of new favourable mutations. The selection response does not usually show an acceleration, as seen with insecticide resistance, but is approximately linear and can be predicted from estimates of the genetic variance prior to selection. Nor does the selection response diminish and reach a plateau e.g. [55], except in small populations, indicating that few genes of large effect have reached fixation. Historically, debate on the mutations underlying the response to selection was divided by strong selection at a few loci or relatively weak selection at many loci. However in Holstein, for example, there has been large increases in milk production with very few ‘hard sweeps’ observed in the genome and few observations of large-effect QTL.

Although we show that most selection for complex traits does not leave a classic signature of selection, we do not imply that selection does not change the allele frequency at sites causing variation in complex traits. Turchin et al. [56] show that mutations affecting human height have been subject to selection because, at many loci, the alleles for increased height have higher frequency in northern than in southern Europe. However, Turchin et al. present no evidence that a selection signature could be discerned if the sites associated with variation in height were not already known. In human
height and in cattle milk yield, selection has no doubt changed allele frequencies at causal loci but not enough to leave a selection signature that is recognisable in the absence of prior knowledge of loci associated with height or milk yield or indeed most complex traits. An implication of this conclusion is that searching for classic selection signatures is not a powerful method to map genes for complex traits even if the traits have been under selection.

Identification of genomic regions under selection for complex traits requires approaches more sensitive to detect subtle changes in allele frequencies over time and with greater flexibility to detect selection on segregating variants. At least in domestic animals, the explicit use of the pedigree structure in may be more appropriate to detect genomic regions responsible for recent selection e.g. [57,58]. We did find a weak association between selection signatures (|iHS|) and QTL for milk production traits by considering 20% of the genome. However, finding such a weak association over such a large part of the genome is not very useful in practice. This weak association occurred despite the advantages of using genomic selection methodologies to identify QTL [11]. For example, compared to single SNP regressions, our approach to identify QTL can capture a higher proportion of the genetic variance [52] and has an improved ability to account for population stratification [59].

The detection of clear selection signatures is compromised by a number of other factors that are illustrated by the individual loci that we examined. There are many traits subject to natural and artificial selection and many genes affect each trait. Therefore the genome contains many possible sites of selection and this complicates the interpretation of the data. For instance, we examined the region surrounding ABCG2 but may well have detected selection at NCAPG-LCORL. The large number of loci segregating for many traits possibly also leads to complex results on BTA20 where there are >1 QTL for milk production [43]. Also multiple alleles at a locus under selection seems to be common and could cloud the interpretation. We found or confirmed multiple alleles at POLLED, MCI-R, KIT, KITLG and PMEL. Migration or introgression of a selected mutation from one breed to another leaves an unusual selection signature as shown by PLAG1 in Limousin where $F_{ST}$ between Limousin and other breeds is high except at the position of the selected mutation. This pattern is expected if the common ancestor of all PLAG1 mutant alleles in Limousin is a Limousin haplotype that differs except at the PLAG1 mutation from haplotypes in other breeds carrying the same mutation. In the case of DGAT1 there has been recent selection for the ancestral allele after possible earlier selection for the mutant. Thus many of the small sample of genes studied display properties that complicate the interpretation of the data and decrease our ability to find clear evidence of classic selection signatures.

**Conclusions**

We conclude that the conditions that give rise to a clear selection signatures (i.e. strong selection for a mutation that would previously have been detrimental) are rare. More usually the response to selection is based on small frequency changes at many loci that were already polymorphic in the population before selection began. Consequently, many of the claims for identifying loci affecting complex traits using selection signatures must be treated with caution.

**Methods**

**Overview**

We obtained real and imputed Illumina Bovine high-density genotypes from 8 cattle selected primarily for dairy or beef production (dairy breeds: Holstein, Jersey; Beef breeds: Angus, Charolais, Limousin, Hereford, Murray Grey, Shorthorn). Sliding windows of 250 kb were constructed across the genome, where each 250 kb length was separated by 50 kb. A window size of 250 kb was chosen because its approximate time to coalescence is 2,000 years (i.e. 1/0.0025 Morgan = 400 generations or 2,000 years assuming 5 years per generation; following [60]), which should represent chromosome segments segregating in domesticated cattle prior to breed formation. For each window, we calculated statistics which would identify within breed selection (i.e. HAPH and |iHS| defined below), computed the divergence between the breeds using Wright's $F_{ST}$ and calculated the variance in genomic estimated breeding values (GBEV) for Jersey and Holstein breeds for dairy traits (milk, fat and protein yield; fat and protein concentration; stature and fertility). We tested for over-representation of the top 5% of windows with selection signatures (within either Holstein or Jersey, and across dairy and beef breeds) that were also in the top 5% of windows for genetic variance in dairy traits. The significance of this over-representation was assessed by a chi-squared test on a 2x2 contingency table. The 3 selection statistics and annotated genomic features for each 250 kb window are contained in Additional file 1.

**Genotype data**

Datasets from dairy and beef cattle were available for analysis. We analysed only autosomal SNP. The dairy dataset consisted of 616,350 SNP for 13,501 Holstein and 5240 Jersey animals. The beef dataset consisted of 692,527 SNP for 2510 Angus, 463 Charolais, 744 Hereford, 61 Limousin, 254 Murray Grey and 868 Shorthorn cattle. Genotype quality control and imputation methods for the dairy data are described by Erbe et al. [61] and Bolormaa et al. [62] describes the beef data.
Within breed selection – haplotype homozygosity (HAPH)

Haplotype segments were constructed for dairy and beef datasets using phased data from Beagle [63] and non-overlapping segments of 30 or 31 SNP. For each chromosome segment we calculated a modified version of Depaulis-Veuille’s H-test [8], referred to as HAPH, where

\[ HAPH = \sum (p_i - q_i)^2, \]

where \( p_i \) is the (within breed) frequency of the \( i \)th haplotype and \( N \) is the total number of haplotypes observed for the breed at the position. Chromosome segments were allocated to 250 kb windows in which their mid-point fell and the average calculated for each 250 kb window. HAPH was then standardized by dividing this value by the breed average over all windows ‘Hard sweeps’ (i.e. Table 6) were identified by windows in the top 5% of HAPH values and separated by less than 1 Mb.

Differentiation between breeds – calculation of \( F_{ST} \) for each breed by breed comparison

Wright’s measure of population differentiation (\( F_{ST} \)) was calculated for each breed combination (i.e. 8 breeds = 28 comparisons) using a common set of 610,123 SNP. The average \( F_{ST} \) was calculated in each 250 kb window following Weir & Cockerham [66] as:

\[ F_{ST} = \frac{\sum (p_{ij}^2 - (\bar{p}_j)^2)}{\sum \bar{p}_j (1 - \bar{p}_j)} \]

where \( j \) is each SNP in the 250 kb window, \( p_{ij} \) is the allele frequency for breed \( i \) at SNP \( j \), and \( \bar{p}_j \) is the mean allele frequency of the breeds at SNP \( j \). On average there were 60 SNP per window (range: 1 to 173 SNP; SD: 22 SNP).

To find windows where dairy breeds differed most from beef breeds the \( F_{ST} \) values between pairs of breeds where one was a dairy breed and one was a beef breed (e.g. Holstein with Angus) were compared to \( F_{ST} \) values between breeds where both were either dairy (Holstein with Jersey) or type 1 and type 2 loci beef breeds (e.g. Angus with Charolais). \( F_{ST} \) values for a window were divided by the mean \( F_{ST} \) over all windows for that pair of breeds and then compared using a one-sided non-parametric Mann–Whitney U test.

Variance in GEBV for milk production traits

Phenotypes and genotypes were obtained from the Australian Dairy Herd Improvement Scheme (ADHIS) for 3,391 Holstein and 1,014 Jersey bulls. Bull genotypes were a subset of animals used to detect the selection signatures. The effect of each SNP was estimated using BayesR, using the same process as Erbe et al. [61], which simultaneously estimates the mean, a polygenic effect and the effects of all SNP. Separate analysis were conducted for each trait by breed combination, where each analysis used 50,000 iterations (30,000 discarded as burn in) and SNP effects were the mean of 5 replicate chains. For each trait we estimated the genetic value of each 250 kb window in each animal (its local GEBV) by \( X \hat{b} \) (i.e. \( X \) is a matrix of genotypes, and \( \hat{b} \) is the estimated SNP effect from BayesR). The variance across animals of GEBVs at a window indicates the windows contribution to genetic variance for that trait. The windows with the top 5% of values for this variance for each breed by trait combination were assumed to contain putative QTL.

Genomic annotations and selection of type 1 and type 2 loci

The locations of genomic features were downloaded using BioMart [67] on 15th March 2013. Genes were mapped to each 250-kb window using their gene start and stop positions using their Ensemble ID and associated gene name (when available). All map positions of SNP and genomic features used UMD3. The loci used as type 1 and type 2
loci were a selection of loci available from the literature, including some identified from the Online Inheritance in Animals [10] database.

Testing for over-representation of selection signatures with QTL for production traits
The top 5% of windows for \textit{HAPPH}, \textit{|iHS|} and the dairy by beef \textit{FST} \textit{c} test were deemed to indicate evidence of selection. A chi-squared test with 1 df was used to determine if the number of windows which ranked in the top 5% for the indicator of selection and the top 5% for the variance in GEBV for the production trait was more than expected by chance. The chi-squared test used the average of 5 non-overlapping sets of windows by dividing the actual number of overlapping windows by 5 (i.e. the number of times each segment of the genome was counted in a window). For the dairy by beef breed comparison, windows were counted if they were in the top 5% of windows for GEBV variance in either Holstein or Jersey.

Ethics statement
No animal experiments were performed specifically for this manuscript. Where data were obtained from existing sources, references for these experiments are provided.

Additional files
Additional file 1: This file contains the estimated haplotype homozygosity (\textit{HAPPH}), the integrated haplotype score (\textit{iHS}) and pairwise breed comparisons for Wright’s measure of population differentiation (\textit{FST}) at all 250 kb windows. The data columns are defined in Additional file 3. The file is compressed with gzip.

Additional file 2: This file contains the column headers for Additional file 1.

Additional file 3: This file contains Figures S1-S14 and Tables S1-S3.

Competing interests
The authors declare that they have no competing interests.

Authors’ contributions
KEK carried out the analysis and wrote the first draft of the manuscript. MEG conceived the study and reviewed the draft manuscript. KEK, SJS and MEG contributed to the study design. SB and BJH collated and imputed the genotype datasets. All authors have read and approved the final manuscript.

Acknowledgements
Authors thank the Beef Cooperative Research Centre for Beef Genetic Technologies and Dairy Futures Co-operative Research Centre for funding and the provision of data to conduct this research. We thank Gert Nieuwhof from the Australian Dairy Hered Improvement Scheme (ADHIS) for providing the data to create genetic trends in dairy cattle (Additional file 3: Figures S1-S3). This research was supported under Australian Research Council’s Discovery Projects funding scheme (project DP1093502). The views expressed herein are those of the authors and are not necessarily those of the Australian Research Council.

Author details
1Department of Agriculture and Food Systems, University of Melbourne, Parkville 3052, Australia. 2Australian Dairy Herd Improvement Scheme, 22 William Street, Melbourne 3000, Australia. 3Department of Environment and Primary Industries, Agribio, Bundoola 3086, Australia. 4La Trobe University, Bundoola 3086, Australia. 5Dairy Futures Co-operative Research Centre, Bundoola 3086, Australia.

Received: 3 January 2014 Accepted: 20 March 2014

References
1. Maynard-Smith JM, Haigh J: The hitch-hiking effect of a favourable gene. Genet Res 1974, 23(1):23-35.
2. Hermisson J, Pennings PS: Soft sweeps: molecular population genetics of adaptation from standing genetic variation. Genetics 2005, 169(4):2335–2352.
3. Bersaglieri T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, Drake JA, Rhodes M, Reich DE, Hirschhorn JN: Genetic signatures of strong recent positive selection at the lactase gene. Am J Hum Genet 2004, 74(6):1111–1120.
4. Akey JM, Ruhe AL, Akey DT, Wong AK, Connelly CF, Madey J, Nicholas TJ, Neff MW: Tracking footprints of artificial selection in the dog genome. Proc Natl Acad Sci U S A 2010, 107(3):1160–1165.
5. Fu W, Akey JM: Selection and adaptation in the human genome. Annu Rev Hum Genomics Hum Genet 2013, 14:467–489.
6. Pritchard JK, Pickrell JK, Coop G: The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation.Curr Biol 2010, 20(4):R208–R215.
7. Hernandez RD, Kelley JL, Elyashiv E, Melton SC, Auton A, McVean G, Project G, Sella G, Przeworski M: Classic selective sweeps were rare in recent human evolution. Science 2011, 331(6019):920–924.
8. Depaulis F, Veuille M: Neutrality tests based on the distribution of haplotypes under an infinite-site model. Mol Biol Evol 1998, 15(12):1788–1790.
9. Voight BF, Kudaravalli S, Wen X, Pritchard JK: A map of recent positive selection in the human genome. PLoS Biol 2006, 4(3):e72.
10. Online Mendelian Inheritance in Animals. [http://omia.angis.org.au/]
11. Meuwissen THE, Hayes BJ, Goddard ME: Prediction of total genetic value using genome-wide dense marker maps. Genetics 2001, 157(4):1819–1829.
12. Hodges J: Jubilee history of the European Association for animal production. 1949–1999, Jevst Prod Sci 1999, 60(2–3):105–168.
13. Hernandez BD, Bovenhuis H, Janss LL, Schrooten C, Visser GG, Goudena ME: The genome response to artificial selection: a case study in dairy cattle. PLoS ONE 2011, 6(2):e14726.
14. Flori L, Fritz S, Jaffrezic F, Boussaha M, Gut I, Heath s, Foulley JL, Gaulier M: The genome response to artificial selection: a case study in dairy cattle. PLoS ONE 2009, 4(8):e6595.
15. Stella A, Ajmone-Marsan P, Lazzari B, Boettcher P: Identification of selection signatures in cattle breeds selected for dairy production. Genetics 2010, 185(4):1451–1461.
16. Fontanesi L, Tazzoli M, Russo V, Beever J: Genetic heterogeneity at the bovine KIT gene in cattle breeds carrying different putative alleles at the spotting locus. Anim Genet 2010, 41(3):295–303.
17. 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):758–762.
18. Allain-Bonnet A, Grohs C, Medugorac I, Krebs S, Djari A, Graf A, Fritz S, Seichert D, Baur A, Russ I, Bouret S, Rothammer S, Wahlberg P, Esquénola D, Hoze C, Boussaha M, Weiss B, Thépot D, Fouilloux M-N, Rossignol M-N, van Marle-Köster E, Hreiðarsdóttir GE, Barbey S, Dozias D, Cobo E, Reversé P, Catsos O, Marchand JL, Souls P, Roy P, et al: Novel insights into the bovine polled phenotype and horn ontogenesis in Bovidae. PLoS ONE 2013, 8(5):e63512.
19. Medugorac I, Seichert D, Graf A, Russ I, Blum H, Göpel KH, Rothammer S, Förster M, Krebs S: Bovine polledness – An autosomal dominant trait with allelic heterogeneity. PLoS ONE 2012, 7(6):e39477.
20. Schmutz SM: Genetics of Coat Color in Cattle. In Bovine Genomics. Edited by Womack JE. Iowa: John Wiley & Sons; 2012.
21. Kunglhand H, Vage DI, Fernández-Raya L, Adalsteinsson S, Lien S: The role of melanocyte-stimulating hormone (\textit{MSH}) receptor in bovine coat color determination. Mamm Genome 1995, 6(10):636–639.
22. Gutierrez-Gil B, Wiener P, Williams J: Genetic effects on coat colour in cattle: dilution of eumelanin and phaeomelanin pigments in an \textit{F2}Backcross Choralsix Holstein population. BMC Genet 2007, 8:156.
23. Schmutz SM, Dreger DJ: Interaction of \textit{MC1R} and \textit{PMEL} alleles on solid coat colors in Highland cattle. Anim Genet 2013, 44(1):9–13.

http://www.biomedcentral.com/1471-2164/15/246
24. Grosz M, MacNeil M: Brief communication. The ‘spotted’ locus maps to bovine chromosome 6 in Hereford-cross population. J Hered 1999, 90(1):233–236.

25. Hayes BJ, Pryce J, Chamberlain AJ, Bowman PJ, Goddard ME: Genetic architecture of complex traits and accuracy of genomic prediction: coat colour, milk-fat percentage, and type in Holstein cattle as contrasting model traits. PLoS Genet 2010, 6(10):e1001159.

26. Setz J, Schmutz SM, Thue TD, Buchanan FC: A missense mutation in the bovine MGF gene is associated with the roan phenotype in Belgian Blue and Shorthorn cattle. Mamm Genome 1999, 10(7):710–712.

27. Pausch H, Wang X, Jung S, Krogmeier D, Edel C, Cremer R, Gotz K-U, Fries H: Identification of QTL for UV-protective eye area pigmentation in cattle by progeny phenotyping and genome-wide association analysis. PLoS ONE 2012, 7(3):e36396.

28. Hayes BJ, Chamberlain AJ, Maceachern S, Savin K, McPartlan HC, MacLeod I, Sethuraman L, Goddard ME: A genome map of divergent artificial selection between Bos taurus and Bos indicus beef cattle. Anim Genet 2009, 40(2):164–184.

29. Fortes MRS, Kemper KE, Sasazuki S, Rerenter A, Pryce JE, Barendse W, Brunch R, McCulloch R, Harrison B, Bolormaa S, Zhang YD, Hawken RJ, Goddard ME, Lehnert SA: Evidence for pleiotropy and recent selection in the PLAG1 region in Australian beef cattle. Anim Genet 2013, 44(6):636–647.

30. Grisart B, Farnir F, Larim M, Cambisano N, Kim JJ, Kivaz S, Mei Y, Simon P, Freire JM, Georges M, Coppieters W: Genetic and functional demonstration of the causality of the DGAT1 K322A mutation in the determination of the BT/AD QTIs affecting milk yield and composition. Proc Natl Acad Sci U S A 2004, 101(29):10243–10248.

31. Karim L, Takeda H, Lin L, Druet T, Arias JAC, Baurain D, Cambisano N, Davis SR, Farnir F, Grisart B, Harris BL, Keehan MD, Littlejohn MD, Spelman RJ, Georges M, Coppieters W: Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature. Nat Genet 2011, 43(5):405–413.

32. Karim L, Coppieters W, Grobet L, Georges M, Valentini A: Convenient genotyping of six myostatin mutations causing double-muscling in cattle using a multiplex oligonucleotide ligation assay. Anim Genet 2000, 31(6):396–399.

33. Coppieters W, Farnir F, Larim M, Ford L, Berzi P, Cambisano N, Mihi M, Reid S, Simon P, Spelman R, Georges M, Snell R: 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, 12:222–231.

34. Spelman RJ, Ford CA, McElhinney P, Gregory GC, Snell RG: Characterization of the DGAT1 gene in the New Zealand dairy population. J Dairy Sci 2002, 85(12):3145–3157.

35. Keper RE, Winkler A, Fries R, Erhardt G: DGAT1 polymorphism in Bos indicus and Bos taurus cattle breeds. J Dairy Res 2004, 71(2):182–187.

36. Blott S, Kim JJ, Moisio S, Schmidt-Kuntzel A, Cornet A, Berzi P, Cambisano N, Ford C, Grisart B, Johnson D, Larim M, Simon P, Snell R, Spelman R, Wong J, Vilki J, Georgis M, Farnir F, Coppieters W: Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. Genetics 2003, 163:253–266.

37. Cohen-Zinder M, Seroussi E, Larkin DM, Loor JJ, Wind AEv, Lee JH, Drakeley J, Band MR, Hernandez AG, Shani A, Levin HA, Weller JR, Ron M: Identification of a missense mutation in the bovine ABC2BC gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. Genome Res 2005, 15(7):936–944.

38. Ron M, Cohen-Zinder M, Peter C, Weller JR, Erhardt G: Short communication: a polymorphism in ABC2BC in Bos indicus and Bos taurus cattle breeds. J Dairy Sci 2006, 89(12):4921–4923.

39. McPherson AM, Lee S-D: Double muscling in cattle due to mutations in the myostatin gene. Proc Natl Acad Sci U S A 1997, 94(23):12457–12461.

40. Grobet L, Poncetel D, Royo L, Brouwers B, Piroit D, Michaux C, Menisier F, Zanotti M, Durner S, Georges M: Molecular definition of an allelic series of mutations disrupting the myostatin function and causing double-muscling in cattle. Mammm Genome 1998, 9(3):210–213.

41. Esmaillzadeh AK, Bottema CDK, Sellick GS, Verbyla AP, Morris CA, Cullen NG, Pitchford WS: Effects of the myostatin F94L substitution on beef traits. J Anim Sci 2008, 86(Suppl 1):1038–1046.

42. Setoguchi K, Watanabe T, Welkard R, Albeckt E, Kuhn C, Kinosita A, Sugimoto Y, Takasuga A: The SNP c.1326 T > G in the non-SMC condensin I complex, subunit G (NCAPG) gene encoding a p.Ile442Met variant is associated with an increase in body frame size at puberty in cattle. Anim Genet 2011, 42(6):650–655.

43. Chamberlain AJ, Hayes BJ, Savin K, Bolormaa S, McPartlan HC, Bowman PJ, van der Jagt C, MacSachern S, Goddard ME: Validation of single nucleotide polymorphisms associated with milk production traits in dairy cattle. J Dairy Sci 2012, 95(2):864–875.

44. Lu D, Willard D, Patel IR, Kadwell S, Overton L, Kost T, Luther M, Chen W, Woychik RP, Wilkinson WC, Cone RD: Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor. Nature 1994, 371(6500):799–802.

45. Kleibig ML, Wilkinson JE, Geisler SG, Woychik RP: Ectopic expression of the agouti gene in transgenic mice causes obesity, features of type II diabetes and yellow fur. Proc Natl Acad Sci U S A 1995, 92(11):4728–4732.

46. Qianbi P, Pausch H, Jansen S, Sorel M, Strom TM, Fries R, Nielsen R, Simianer H: Classic selective sweeps revealed by massive sequencing in cattle. PLoS Genet 2014, 10(2):e1004148.

47. Bolormaa S, Pryce JE, Reverter A, Zhang Y, Barendse W, Kemper K, Tier B, Savin K, Hayes BJ, Goddard ME: A multi-trait meta-analysis for detecting pleiotropic polymorphisms for stature fatness and reproduction in beef cattle. PLoS Genet 2014, 10(3):e1004198.

48. Zammit C, Coope R, Gomm JJ, Shousha S, Johnston CL, Coombes RC: Fibroblast growth factor 8 is expressed at higher levels in lactating human breast and in breast cancer. Br J Cancer 2002, 86(1):109–113.

49. Thomson B, Horn P, Paulitz F, Bendixen E, Petersen AH, Hofm L, Einesen V, Agerholm JS, Ambjerg J, Bendixen CA: A missense mutation in the bovine SLC35A3 gene, encoding a UDP-N-acetylgalcosamine transporter, causes complex vertebral malformation. Genome Res 2006, 16(11):197–105.

50. Lynd A, Weetman D, Baroba S, Eggyr Yasovan A, Mitchell S, Pinto J, Hastings D, Donnelly MJ: Field, genetic, and modeling approaches show strong positive selection acting upon an insecticide resistance mutation in Anopheles gambiae. Mol Biol Evol 2010, 27(3):1117–1125.

51. Cole JB, van Radem PN, O’Connell JR, van Tassell CP, Sonstegard TS, Schnabel RD, Taylor JF, Wiggins GR: Distribution and location of genetic effects for dairy traits. J Dairy Sci 2009, 92(6):2931–2946.

52. Yang J, Benjamini B, McEvoy BP, Gordon S, Henders A, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME, Vischer PM: Common SNPs explain a large proportion of the heritability for human height. Nat Genet 2010, 42(5):565–569.

53. Innan H, Kim Y: Pattern of polymorphism after strong artificial selection in a domestication event. Proc Natl Acad Sci U S A 2004, 101(29):10667–10672.

54. Hill WG, Caballero A: Artificial selection experiments. Annu Rev Ecol Syst 1992, 23:287–310.

55. Brotherstone S, Goddard M: Artificial selection and maintenance of genetic variance in the global dairy cow population. Philos Trans R Soc Lond B Biol Sci 2005, 360(1459):1479–1488.

56. Turchin MC, Chiang CW, Palmer CD, Sankararaman S, Reich D, Hirschhorn JN: Evidence of widespread selection on standing variation in Europe at height-associated SNPs. Nat Genet 2012, 44(9):1015–1019.

57. Larkin DM, Daetwyler HD, Hernandez AG, Wright CL, Hetrick LA, O’Connell JR, Greenwood PL, Arthur PF, Goddard ME: Improved accuracy of genomic predictions within and between dairy cattle breeds with imputed high-density single nucleotide polymorphism panels. J Dairy Sci 2012, 95(7):4114–4129.
62. Bolormaa S, Pryce JE, Kemper K, Savin K, Hayes BJ, Barendse W, Zhang Y, Reich CM, Mason BA, Bunch RJ, Harrison BE, Reverter A, Herd RM, Tier B, Graser HU, Goddard ME: Accuracy of prediction of genomic breeding values for residual feed intake, carcass and meat quality traits in Bos taurus, Bos indicus and composite beef cattle. J Anim Sci 2013, 91(7):3088–3104.

63. Browning SR, Browning BL: Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet 2007, 81(5):1084–1097.

64. Gautier M, Vitalis R: rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. Bioinformatics 2012, 28(8):1176–1177.

65. R: A language and environment for statistical computing. [http://www.R-project.org/]

66. Weir BS, Cockerham CC: Estimating F-statistics for the analysis of population structure. Evolution 1984, 38:1358–1370.

67. Flicek P, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, Fitzgerald S, Gil L, Gordon L, Hendriks M, Houlier T, Johnson N, Kühner AK, Keefe D, Keenan S, Kinsella R, Komorowska M, Koscielny G, Kulesha E, Larsson P, Longden I, McLaren W, Muffato M, Overduin B, Pignatelli M, Pritchard B, Riat HS, et al: Ensembl 2012. Nucleic Acids Res 2012, 40(D1):D84–D90.

doi:10.1186/1471-2164-15-246
Cite this article as: Kemper et al.: Selection for complex traits leaves little or no classic signatures of selection. BMC Genomics 2014 15:246.
