

# Composite Selection Signals for Complex Traits Exemplified Through Bovine Stature Using Multibreed Cohorts of European and African *Bos taurus*

Imtiaz A. S. Randhawa,<sup>1</sup> Mehar S. Khatkar, Peter C. Thomson, and Herman W. Raadsma

Reprogen–Animal Bioscience Group, Faculty of Veterinary Science, The University of Sydney, Camden, 2570, New South Wales, Australia

**ABSTRACT** Understanding the evolution and molecular architecture of complex traits is important in domestic animals. Due to phenotypic selection, genomic regions develop unique patterns of genetic diversity called signatures of selection, which are challenging to detect, especially for complex polygenic traits. In this study, we applied the composite selection signals (CSS) method to investigate evidence of positive selection in a complex polygenic trait by examining stature in phenotypically diverse cattle comprising 47 European and 8 African *Bos taurus* breeds, utilizing a panel of 38,033 SNPs genotyped on 1106 animals. CSS were computed for phenotypic contrasts between multibreed cohorts of cattle by classifying the breeds according to their documented wither height to detect the candidate regions under selection. Using the CSS method, clusters of signatures of selection were detected at 26 regions (9 in European and 17 in African cohorts) on 13 bovine autosomes. Using comparative mapping information on human height, 30 candidate genes mapped at 12 selection regions (on 8 autosomes) could be linked to bovine stature diversity. Of these 12 candidate gene regions, three contained known genes (*i.e.*, *NCAPG-LCORL*, *FBP2-PTCH1*, and *PLAG1-CHCHD7*) related to bovine stature, and nine were not previously described in cattle (five in European and four in African cohorts). Overall, this study demonstrates the utility of CSS coupled with strategies of combining multibreed datasets in the identification and discovery of genomic regions underlying complex traits. Characterization of multiple signatures of selection and their underlying candidate genes will elucidate the polygenic nature of stature across cattle breeds.

## KEYWORDS

body size  
outbred  
populations  
polygenic traits  
signatures of  
selection  
stature genes

Genomes that have experienced selective pressure due to their functional role in phenotypic diversity contain unique patterns of genetic diversity called signatures of selection (Oleksyk *et al.* 2010). Recent progress in molecular genetics has provided powerful tools to generate

and analyze genetic polymorphism data (Lenstra *et al.* 2011; Hancock and Di Rienzo 2008; Crisci *et al.* 2012). Consequently, genome-wide scans to detect signatures of selection have been conducted in several species to understand the role of molecular variants in shaping their genetic diversity (Gibbs *et al.* 2009; Ramey *et al.* 2013; Kijas *et al.* 2012; Zhang *et al.* 2012; Petersen *et al.* 2013; Qanbari *et al.* 2014; Utsunomiya *et al.* 2013; Akey *et al.* 2010).

Footprints of sustained selection in livestock are generally diverse and complex. It has been a great challenge to differentiate between the neutrally (random drift) and rapidly (selection) evolving traits (Hohenlohe *et al.* 2010). In general, signatures of selection are identified from the patterns of DNA variation without recorded phenotypic information on populations undergoing investigation. Nonetheless, the signatures of selection detected between populations are often indicative of segregating regions of functional mutations underlying divergently selected quantitative traits (Hayes *et al.* 2008,

Copyright © 2015 Randhawa *et al.*

doi: 10.1534/g3.115.017772

Manuscript received August 2, 2014; accepted for publication April 28, 2015; published Early Online April 30, 2015.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supporting information is available online at [www.g3journal.org/lookup/suppl/doi:10.1534/g3.115.017772/-/DC1](http://www.g3journal.org/lookup/suppl/doi:10.1534/g3.115.017772/-/DC1)

<sup>1</sup>Corresponding author: Reprogen–Animal Bioscience Group, Faculty of Veterinary Science, The University of Sydney, 425 Werombi Road, Camden, 2570, NSW, Australia. E-mail: imtiaz.randhawa@sydney.edu.au

2009). In complex genomes such as mammalian species, the functional mutations and their systematic patterns of polymorphism exist in multiple locations. Identification of the trait-specific genomic regions requires comprehensive phenotypic and genomic data of large populations to detect causal effects of selection (Decker *et al.* 2012). Despite using known mapping approaches, it has often been difficult to localize historical selection events for complex traits (Kemper *et al.* 2014). A number of tests are available to detect candidate regions under selection, which often provide differing results from the same genomic dataset (Qanbari *et al.* 2011; Crisci *et al.* 2012). Combining several tests with complementary characteristics can improve the ability to localize the candidate regions under selection (Utsunomiya *et al.* 2013; Grossman *et al.* 2010).

Recently, we developed a composite selection signal (CSS) test (Randhawa *et al.* 2014), a weighted index of signatures of selection from multiple estimates, to identify genomic regions under selection in domestic species. CSS is a nonparametric procedure that uses fractional ranks of constituent tests and it does not depend on their distributional assumptions. Hence, as compared to the previous approaches of combining multiple selection tests that have limitations to incorporate only the statistics with *p*-values (Utsunomiya *et al.* 2013; Grossman *et al.* 2010), CSS has an appealing feature to combine the evidence of historical selection from any selection test. We demonstrated that CSS is a robust method to identify loci for across-population genetic diversity of genomic regions, linking them to target phenotypes of monogenic traits under selective pressure (Randhawa *et al.* 2014). CSS can be extended to complex traits. It is expected that co-selection on multiple genomic loci underlying a complex trait can be captured in the form of CSS by combining the signals detected from the patterns of population differentiation, derived allele frequency, and the extended haplotype homozygosity.

The quantitative trait loci (QTL) and signatures of selection mapped at the candidate regions suggested the role of underlying genes in diversity of stature across various breeds (Druet *et al.* 2013; Ramey *et al.* 2013; Karim *et al.* 2011; Nishimura *et al.* 2012; Kemper *et al.* 2014).

Diversity in bovine stature can potentially be used as a model to demonstrate the application of new approaches, such as CSS, for the detection of selection for complex traits in cattle, given that the trait can be easily measured and data are available in a wide range of worldwide breeds. Notably, stature is a trait strongly influenced by domestication (Karim *et al.* 2011) and has a known history of selective pressure over time, both negative and positive (Ajmone-Marsan *et al.* 2010; Fortes *et al.* 2013). Stature (or adult height) is a complex trait under polygenic control with high heritability in many mammalian species including cattle (Kemper and Goddard 2012; Lettre 2011; Lanktree *et al.* 2011). Genetic architecture of height has been extensively investigated to identify genes with major effects across the genome of multiple species, although the leading model has been the human genome (Lettre *et al.* 2008; Sanna *et al.* 2008; Shiura *et al.* 2009; Tetens *et al.* 2013; Deng *et al.* 2013; Lanktree *et al.* 2011). To date, only a few genes responsible for stature, including body size traits, have been reported in cattle from genome-wide association studies (GWAS) (Pryce *et al.* 2011; Fortes *et al.* 2013; Li *et al.* 2013; Hoshiba *et al.* 2013; Bolormaa *et al.* 2014).

Two major candidate gene regions, *NCAPG-LCORL* and *PLAG1-CHCHD7*, on chromosome 6 and chromosome 14, respectively, have been linked to stature in European breeds of cattle. The *NCAPG* and *LCORL* genes are strong candidates for growth and height-related traits in multiple species of mammals, including cattle (Gudbjartsson *et al.* 2008; Weedon *et al.* 2008; Soranzo *et al.* 2009; Lango Allen *et al.*

2010; Okada *et al.* 2010; Pryce *et al.* 2011; Sovio *et al.* 2009). This candidate gene region has frequently shown strong signatures of selection in multiple breeds of cattle (Barendse *et al.* 2009; Druet *et al.* 2013; Flori *et al.* 2009; Gibbs *et al.* 2009; Kemper *et al.* 2014; Mancini *et al.* 2014; Pintus *et al.* 2013; Porto-Neto *et al.* 2013; Rothhammer *et al.* 2013; Stella *et al.* 2010; Qanbari *et al.* 2014; Perez *et al.* 2014). Several polymorphic variants at this locus have been identified in GWAS and QTL studies for wither height in horse (Tetens *et al.* 2013; Signer-Hasler *et al.* 2012), and growth, skeletal, carcass, and stature traits in cattle (Lindholm-Perry *et al.* 2011; Pryce *et al.* 2011; Hoshiba *et al.* 2013; Nishimura *et al.* 2012; Bolormaa *et al.* 2014). The *PLAG1-CHCHD7* locus harbors nine candidate genes and has been associated with height in humans (Gudbjartsson *et al.* 2008; Soranzo *et al.* 2009; Okada *et al.* 2010; Lettre *et al.* 2008; Lango Allen *et al.* 2010) and stature in cattle (Nishimura *et al.* 2012; Pryce *et al.* 2011; Karim *et al.* 2011; Fortes *et al.* 2013; Bolormaa *et al.* 2014). This locus was also identified as a candidate region based on selective sweeps in several cattle breeds that have been under strong selection for body size (Druet *et al.* 2013; Ramey *et al.* 2013; Kemper *et al.* 2014). However, known genes account only for a small proportion of the phenotypic variation in bovine stature (Kemper and Goddard 2012). Comparative mapping provides additional evidence that some orthologous genes of human and cattle affect stature in both species (Pryce *et al.* 2011). Therefore, the comparative mapping of genes identified in human GWAS can be used as supporting evidence to elucidate additional candidate gene regions in cattle (Visscher and Goddard 2011).

Selective pressure on phenotypically similar populations may have acted on the same causal variants in analogous genes (Kijas *et al.* 2012; Stella *et al.* 2010; Pickrell *et al.* 2009). Therefore, a multibreed panel by combining phenotypically alike breeds, and comparing such multibreed groups for contrasting phenotypes can maximize the detectable signatures of selection for bovine stature (Barendse *et al.* 2009; Stella *et al.* 2010; Kemper and Goddard 2012; Visscher and Goddard 2011).

This study applies the CSS method to identify candidate genomic regions harboring genes responsible for bovine stature, a polygenic complex trait. Cohorts of contrasting phenotypes were investigated in multibreed panel comparisons using morphological information on wither height of 55 worldwide breeds of European and African *Bos taurus*. Finally, we compiled orthologous genes from several GWAS on human height and mapped them on the bovine genome assembly (UMD3.1) to explicate the significant genomic regions in cattle defined by CSS.

## MATERIALS AND METHODS

### Genotypic data

In total, 1106 animals of 55 breeds were genotyped with an Illumina BovineSNP50 chip assay (Supporting Information, Table S1, Table S2). Information regarding DNA samples and SNP genotypes were previously reported (Decker *et al.* 2009; Gautier *et al.* 2009, 2010). The duplicate samples across multiple datasets of cattle were identified and removed based on an IBS matrix computed using PLINK (Purcell *et al.* 2007). The IBS matrix estimates the genetic relationship based on genome-wide SNPs similarity. A quality control was implemented to retain SNPs that can be mapped on UMD3.1 bovine assembly and have minor allele frequency (MAF)  $\geq 0.01$  (Table S3). Ancestral and derived allelic phases of these SNPs were determined based on information on several closely related out-group species, including bison, buffalo, and yak (Decker *et al.* 2009; Matukumalli *et al.* 2009). Imputation of missing genotypes and haplotype phasing were performed with BEAGLE 3.3 (Browning and Browning 2009). Analyses

of signatures of selection for European and African cohorts were performed using polymorphic and phased SNPs within each cattle type.

### Phenotypic data

Bovine stature is generally defined as height at the withers; however, this information was not directly available on the genotyped animals of the selected breeds. Therefore, in this study, for a broader perspective of phenotypic characterization, wither heights of 47 European *Bos taurus* breeds (listed in Table S1) and 8 African *Bos taurus* breeds (Table S2) reported by 52 countries were acquired from the online database (DAD-IS 2014) of the Food and Agriculture Organization (FAO). FAO data for the bovine stature are available as the country-wise breed averages for male and female height at withers. Availability of breed averages for wither height ranged from a single country (*i.e.*, country of breed origin or country of DNA sampling) to many countries ( $n = 37$ ). Breed-wise as well as overall estimates of median, upper quartile, and lower quartile of wither height were computed to categorize breeds into phenotypically contrasting cohorts (see below). For this purpose, first, country-wise stature data for each breed were prepared from the averages of male and female withers heights (cm). Second, if a breed was represented by a single country, then the country's average stature for the breed was used to represent the breed; otherwise, the median stature from multiple countries (based on the FAO data available as country-wise average) was used.

### Multibreed cohorts of contrasting stature

The contrasting cohorts of multiple breeds, within European (Table S1) and African (Table S2) breed types, were defined for analyses following the expectation that identification of plausible signatures of selection for stature would require a relatively stringent separation of breeds and sufficient sample sizes in each cohort. Therefore, a stringent strategy was adopted to construct two cohorts of contrasting phenotypes, *i.e.*, the large and small stature cohorts (Figure S1). The individual breeds that have their median stature above the overall upper quartile and below the overall lower quartile were categorized into the large and the small cohorts, respectively. Moreover, the large cohort's breeds (represented by several countries) are further restricted to have their individual lower quartiles above the overall median. Similarly, the small cohort's breeds (represented by several countries) have their individual upper quartiles below the overall median. Separate analyses to identify signatures of selection were performed to compare the large and small cohorts of European and African breeds.

Previously, we have shown that cohort composition had minimal effect on the computation of composite selection tests when the breeds with small sample size ( $n < 10$  and  $n < 20$ ) were excluded (Randhawa *et al.* 2014). Further, in this study, we also assessed the effect of cohort composition comprising European breeds of cattle, specifically whether a few breeds with large sample size ( $n > 20$ ) can disproportionately affect the analyses. To this effect we performed additional analyses by limiting the sample size ( $n \leq 10$  and  $n \leq 20$ ) per breed. These results were compared to those from the original cohorts, where we included all available samples from each breed and the constituent and composite selection tests were computed by analyzing the European large and small cohorts against each other.

### Composite selection signal

The detailed methods of computing composite selection signal (CSS) are provided by Randhawa *et al.* (2014). Briefly, the CSS statistics were computed at each locus by combining three popular constituent se-

lection methods, namely, change in allele diversity by  $F_{ST}$  (Weir and Cockerham 1984) across-population extended haplotype homozygosity (XP-EHH) test (Sabeti *et al.* 2007), and change in derived allele frequencies ( $\Delta DAF$ ) test (Grossman *et al.* 2010).  $F_{ST}$  statistics were computed as the differentiation index between the large and small cohorts. XP-EHH and  $\Delta DAF$  statistics were computed for the large and small cohorts such that, by analyzing the large (as selected) cohort against the small (as nonselected) cohort and vice versa. The composite statistics (CSS) were computed as follows:

1. For each constituent method, test statistics were ranked (1, ...,  $n$ ) genome-wide on  $n$  SNPs.
2. Ranks were converted to fractional ranks ( $r'$ ) (between 0 and 1) by  $1/(n + 1)$  through  $n/(n + 1)$ .
3. Fractional ranks were converted to  $z$ -values as  $z = \Phi^{-1}(r')$  where  $\Phi^{-1}(\cdot)$  is the inverse normal cumulative distribution function (CDF).
4. Mean  $z$  scores were calculated by averaging  $z$ -values across all constituent tests at each SNP position and  $p$ -values were directly obtained from the distribution of means from a normal  $N(0, m^{-1})$  distribution where  $m$  is the number of constituent test statistics.
5. Logarithmic ( $-\log_{10}$  of  $p$ -values) of these mean  $z$ -values were declared as CSS and were plotted against the genomic positions to identify the significant selection signals.
6. To reduce spurious signals, the empirical CSS scores were smoothed by averaging values within 0.5 Mb sliding windows on both sides at each SNP. The top 0.1% of smoothed CSS scores was used to declare a significant SNP relative to the rest of the genome.

### Genomic regions and genes under selection

Following a similar approach of Gibbs *et al.* (2009) and Kijas *et al.* (2012), significant genomic regions were defined as those that harbor at least one significant SNP (top 0.1%) flanked by a set of five or more adjoining SNPs in the top 1%. The genomic positions of the significant regions were defined by the first and last of the top 1% SNP in the cluster. In addition, multiple clusters localized in close proximity ( $\sim 1$  Mb) were considered as a single significant genomic region.

The genes underlying the significant regions and surrounding locations ( $\sim 1$  Mb of cluster boundaries) were listed. These putative genes were searched for evidence of association or signatures of selection for stature within other mammalian species. For this purpose, the records of orthologous genes for the major GWAS reports on human height (Gudbjartsson *et al.* 2008; Lango Allen *et al.* 2010; Lanktree *et al.* 2011; Lettre *et al.* 2008; Okada *et al.* 2010; Weedon *et al.* 2008; Soranzo *et al.* 2009; Kim *et al.* 2010) were mapped on UMD3.1 bovine genome assembly by matching the gene names (Table S4).

In total, 268 orthologous genes were successfully mapped on the bovine autosomes (BTA) and were located in 134 genomic locations, termed as "candidate gene regions" here and defined by the consecutive orthologous genes located on a chromosome within 1 Mb distance from each other. The number of genes per region ranged from one to nine, with an average size of 355 kb. Hence, the 134 regions harboring genes for human stature covered 47.50 Mb (1.9%) of the bovine genome. In addition, after adding 1 Mb surrounding of each candidate gene region, these regions cover 7% of the bovine genome. Hence, finding a candidate gene carrying CSS region out of those 134 regions as expected by chance is equal to 7%.

To compare whether the proportion of identified regions linked to the bovine stature in various analyses was by chance, we performed a permutation test whereby the animals were randomly assigned to two cohorts under a no-selection model. Sample sizes of the permuted cohorts were equal to the samples sizes of real cohorts and samples from individual breeds within cohorts were not weighted. Five sets of independently permuted cohorts, separately for European and African breeds, were analyzed by CSS and the results were compared against the empirical cohorts.

Finally, signatures of selection identified by CSS in candidate gene regions were further classified as previously reported in cattle or novel to the species. Signatures of selection outside of the candidate gene regions were considered as regions of interest that may be putatively related to stature.

## RESULTS

### Genotypic data after quality control

Genotypic data passing quality control consisted of 38,033 SNPs that were filtered after combining the available data on various cattle breeds from Decker *et al.* (2009) and Gautier *et al.* (2009, 2010) (see Table S3 to access raw datasets). The retained SNPs had  $MAF \geq 0.01$ , known ancestral and derived alleles, and mapped on 29 autosomes covering 2,512.08 Mb of UMD3.1 bovine genome assembly. The mean (median) distance between consecutive SNPs on the bovine autosomes was 65.52 kb (64.54 kb), ranging from 2.22 kb to 1.44 Mb (Table S3). Finally, data subsets of 37,218 and 33,024 informative SNPs were analyzed to detect signatures of selection within European and African cattle, respectively.

### Classification of breeds into phenotypically contrasting cohorts

From the FAO data (DAD-IS 2014), the overall lower quartile (Q1), median, and upper quartile (Q3) for the combined values for wither height of 47 European breeds were 134 cm, 140 cm, and 143 cm, respectively (Figure 1A). In comparison, the median wither height for the European large (11 breeds) and European small (12 breeds) stature cohorts were 146.8 cm and 129.5 cm, respectively (Table S1, Figure 1A). The minimum median (145 cm of Charolais and Montbeliarde breeds) of the large cohort and the maximum median (133.5 cm of Finnish Ayrshire and Belted Galloway breeds) of the small cohort values were separated by 11.5 cm.

Similarly using the FAO data for the eight African breeds, overall Q1, median, and Q3 for wither heights calculated were 100 cm, 106 cm, and 115 cm, respectively (Figure 1B). The median wither heights for the African large (two breeds) and African small (three breeds) stature cohorts were 138.5 cm and 97.5 cm, respectively (Table S2). The separation of closest medians at the tails of the two African cohorts was 31.7 cm.

Overall, the distribution of breed-wise data (Figure 1) shows that implementing the stringent grouping strategy resulted in a clear separation (contrast) of wither height (stature) between the large and the small cohorts of the European and African categories.

### Significant CSS

In the analysis of the European and African cohorts, 38 and 33 SNPs represented the top 0.1%, respectively, based on ranks of smoothed CSS, and these were used to map the significant genomic regions. The number of top 0.1% SNPs within each significant region ranged from 1 to 20 (Figure S2). The number of significant regions ranged from four to nine (Table 1). The number of top SNPs per region and region

size was highest in the European cohorts and lowest in the African cohorts (Figure S2). Median (mean) size of the significant regions was 920 kb (1.20 Mb), ranging from 410 kb to 3.46 Mb. Moreover, merging two neighboring regions of European small cattle cohort on BTA-3 increased the maximum region length to 6.34 Mb.

### CSS in European cattle cohorts

Genome-wide distribution of the smoothed CSS ( $-\log_{10}p$ ) for the large and small cohorts of European cattle identified nine genomic regions carrying clusters of significant SNPs (Figure 2A). A complete list of putative regions identified in the two cohorts of European *Bos taurus* is provided in Table S5. In the large cohort, the selected regions were identified on BTA-3, BTA-6, BTA-14, and BTA-19. In the small cohort, the selected regions were located on BTA-3, BTA-5, BTA-7, BTA-10, and BTA-16.

Figure S4 shows a genome-wide comparison for the distribution of significant CSS peaks between the original results (Figure 2A) and three additional analyses—performed to evaluate composition of European cohorts given the variable sample sizes of cattle breeds. Overall, high concordance was observed between the significant peaks in cohort-wise analyses using all available individuals (original), with a maximum of 20 and 10 individuals per breed. For comparison, the significant regions were considered valid if found within the top 1% significant threshold. Notably, the significant regions detected in original analyses (Table S5) were found above the top 1%, with the only exception being a region on BTA-19 in the European large cohort being not detected with a maximum of 10 individuals per breed. Overall, the impact of varying the number of animals per breed had negligible impact on the CSS profile. Hence, we propose that the maximum number of animals should be used when possible.

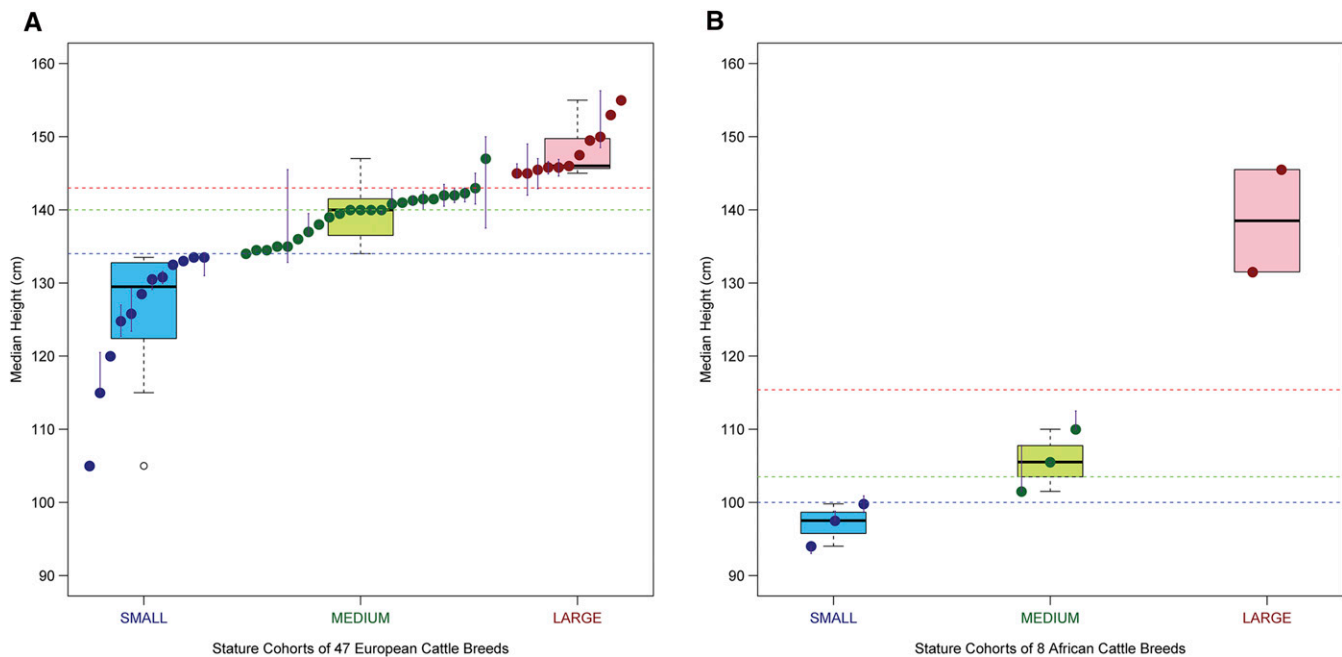
### CSS in African cattle cohorts

The Manhattan plot for African cohorts shows smoothed CSS results from the intercohort comparisons (Figure 2B). We observed that the genome-wide CSS peaks were less pronounced in African cohorts compared with European breeds. A complete list of 17 regions identified in the two cohorts of African *Bos taurus* is provided in Table S6. Overall, nine regions (on BTA-3, BTA-7, BTA-10, BTA-19, and BTA-22) were identified carrying selection signatures in the African large cohort and eight regions (on BTA-4, BTA-6, BTA-7, BTA-8, BTA-10, BTA-11, BTA-15, and BTA-19) were identified in the African small cohort.

### Genomic regions and candidate genes for stature

Table 1 shows a summary of the significant regions in each cohort with comparison to respective reference populations within the European and African categories. It also shows the number of regions harboring candidate genes associated with stature, which were found in human GWAS (Table S4). Overall, out of 9 and 17 significant genomic regions within the European and African cohorts, 7 (77.8%) and 5 (29.4%) were co-located with candidate gene regions that harbor stature-associated genes. The remaining 2 and 12 significant genomic regions in European and African cohorts, respectively, were identified outside the known candidate gene regions. Overall, 46% ( $n = 12$  out of 26) of all significant CSS were co-located with candidate gene regions, this proportion is much higher as compared to the 7% expected by chance alone as described above (Figure S3).

Figure 3 shows a summary comparison of empirical and permuted cohort results. Overall, 8.3% of the significant regions based on permuted CSS in five sets (*i.e.*, 20 permuted cohorts) contained stature-associated



**Figure 1** Distribution of the wither height (cm) of the 47 European (A) and eight African (B) *Bos taurus* breeds. Each colored dot represents the median of a breed's stature using FAO data from the multiple countries and the vertical error bars range between the upper and lower quartiles. Several breeds are represented from a single country; therefore, they do not have error bars. The horizontal dashed red, green, and blue lines, respectively, represent the overall upper quartile, median, and lower quartile of all breeds' data. The red, green, and blue dots represent cattle breeds categorized in the large, medium, and small stature cohorts, respectively (see Figure S1). Boxplots show distribution of breed median data for stature within each cohort.

loci (Figure 3) that was in agreement with the 7% expected by chance (Figure S3). On average, only 6.7% and 9.8% of significant regions of permuted CSS from European and African cohorts had stature-associated genes as compared to 77.8% and 29.4% of empirical regions, respectively.

The Circos image (Figure 4) represents all 26 significant genomic regions, names of important candidate genes for bovine stature (Table 2), and the locations of all the orthologous genes from human GWAS (Table S4). Table 2 shows the list of those 12 genomic regions that harbor 30 candidate genes, which have previously been associated with stature in human GWAS and other species. The results from the major contrasting cohorts, especially large and small cohorts of European *Bos taurus*, supported the strategy of finding candidate gene regions by grouping phenotypically alike breeds. All the significant signatures of selection (four out of four) identified in the European large cohort contained known candidate genes in multiple species (Table 2) within region 1 (gene: *DUSP23*), region 6 (*NCAPG*, *LCORL*), region 10 (*TGSI*, *LYN*, *RPS20*, *MOS*, *PLAG1*, *CHCHD7*, *SDR16C5*, *RDHE2*, *PENK*), and region 11 (*POLR2A*). However, in cattle, only regions 6 and 10 have previously been investigated in relation to stature. In the European small cohort, three out of the four significant genomic regions, *i.e.*, region 3 (*PKN2*), region 5 (*ATP5G2*, *ATF7*), and region 7 (*CAMLG*, *DDX46*, *TXNDC15*, *CATSPER3*, *PITX1*) harbor candidate genes that were previously reported to be associated with human height.

In the African cohorts, several significant signatures of selection regions did not contain known stature-associated candidate genes. Consequently, the percentage of selection regions with known candidate genes was lower, *i.e.*, 29.4%, which is almost one-third of that of European cohorts (77.8%). Despite that, this percentage in the African cohorts was almost three-times more than that of the likeli-

hood of occurrence by chance. Nevertheless, we identified five candidate gene regions (2, 4, 8, 9, and 12) in African cohorts that contain important genes for stature, as reported by multiple human GWAS (Table 2).

Finally, Figure S5 compares the distribution of CSS and its component tests ( $F_{ST}$ ,  $\Delta DAF$ , and XP-EHH) at the 12 candidate regions harboring stature-associated genes. These comparisons show that the CSS scores are not being derived from one particular component selection test. The results also suggest that the three component tests provide complementary evidence to each other at the candidate

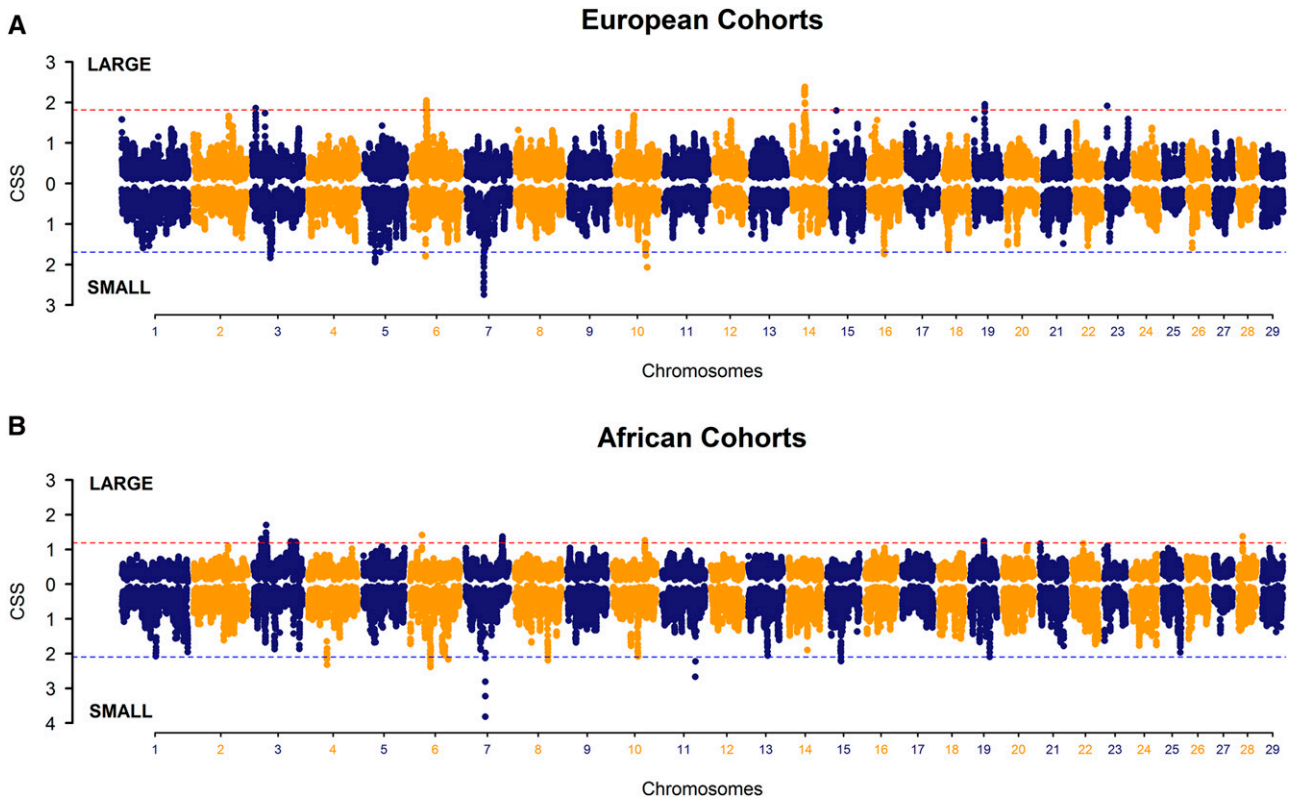
**Table 1** Summary of significant signatures of selection and the candidate regions containing stature-associated genes

Candidate Populations		Reference Population <sup>a</sup>		
Categories	Cohorts	Small	Large	Total Regions
<b>European</b>	Large	4 (4) <sup>b</sup>	—	9 (7)
	Small	—	5 (3)	
<b>African</b>	Large	9 (2)	—	17 (5)
	Small	—	8 (3)	
<b>Overall</b>				26 (12)

Complete CSS results are given in Figures and Supplementary Tables for European cohorts (Figure 2A, Table S5) and African cohorts (Figure 2B, Table S6). Significant signatures of selection (top 0.1% of CSS) and some candidate genes (complete list in Table 2) within each region are shown in Figure 4.

<sup>a</sup> Reference populations for candidate European cohorts were the contrasting European cohorts; similarly, African cohorts were compared against contrasting African cohorts.

<sup>b</sup> Each cell shows number of significant regions identified by analyzing the candidate population against the contrasting reference population. The numbers within parentheses are significant regions that harbor candidate genes associated with stature.



**Figure 2** Manhattan plots of genome-wide (smoothed) composite selection signals (CSS) in the large and small cohorts of (A) European and (B) African breeds. Each cohort was compared against the other as a reference population, as shown in each Manhattan plot for positive and negative CSS scores above or below the dashed lines. Dashed lines (red for positive CSS and blue for negative CSS) indicate the cut-off at the top 0.1% of the genome-wide smoothed CSS distribution.

regions. However, it is evident that many of these candidate regions would not have been detected with any of the component selection tests, particularly regions 1, 2, 5, 8, 9, and 12. These results provide overwhelming support for the CSS strategy, suggesting that it is a robust and efficient method for combining complementary evidence to detect the historical signatures of selection for complex traits, which otherwise cannot be detected by available tests of selection.

## DISCUSSION

### CSS for complex traits

The CSS approach demonstrates that the test is generally applicable for the identification of trait-specific genomic regions for complex traits by comparing phenotypically contrasting populations. The CSS test combines multiple pieces of evidence from the rank distribution of different selection tests in a weighted index of signatures of selection. The power of the test and its constituents ( $F_{ST}$ ,  $\Delta DAF$ , and XP-EHH) depends on the phenotypic and genetic divergence between the candidate and reference populations, as we demonstrated with simple binary monogenic (Randhawa *et al.* 2014) and complex polygenic traits (this study).

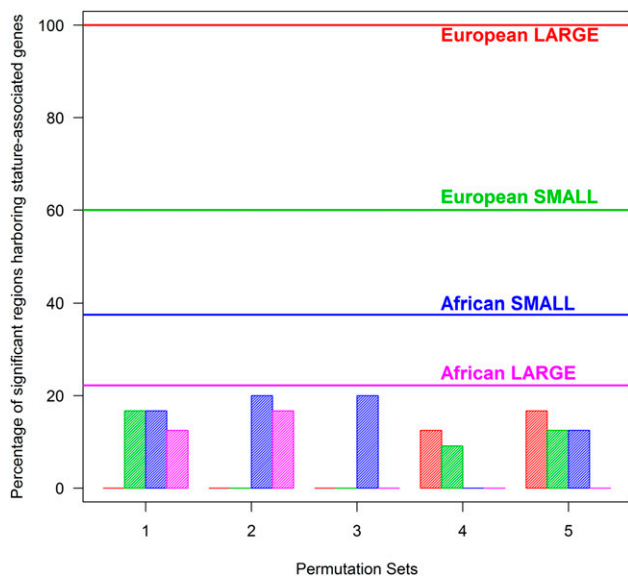
Use of multibreed cohorts has increased the discovery of trait-specific regions because of shared linkage disequilibrium between the causal mutations and neighboring SNPs (Kemper and Goddard 2012) arising from long-term historical selection. Contrasting patterns of genomic variation at the putative regions under selection across groups increases the likelihood of capturing the signatures of selection linked to stature. In our approach, signals of selection are amplified at

candidate regions across breeds within groups, whereas background noise (false-positive signals) is reduced in the rest of the genome. Such noise is usually expected from the demographic history of breed formation and random genetic drift (Seichter *et al.* 2012). Overall, application of the CSS method combined with grouping breeds according to their wither height identified candidate regions in the bovine genome for this complex trait.

Detection of signatures of selection is valuable in the discovery of potential genomic regions of functional mutations affecting quantitative traits (Hayes *et al.* 2008). Recent investigations suggest that, in the absence of classic selective sweeps, signatures of selection for complex traits are less likely to be detected by using individual selection tests on single breed data (Kemper *et al.* 2014). Our approach combines multiple selection tests and grouping of phenotypically alike breeds, and has been found to be highly efficient at detecting signatures of selection and identifying candidate gene regions for complex traits. This approach makes use of existing resources under long-term historical selection and provides a relatively inexpensive entry for more detailed follow-up studies in the genetic architecture of complex traits.

### Detection of signatures of selection for bovine stature

In this study, we used bovine wither height (stature) measures from an online resource, namely the FAO database (DAD-IS 2014), recorded as averages for both males and females across multiple countries. Height and body size have generally been considered important traits in cattle during domestication. During early stages of domestication, selection favored animals with lower stature compared to wild ancestors (Karim *et al.* 2011). However, during the recent past, strong



**Figure 3** Comparison of empirical and permuted CSS scores for localization of significant regions harboring stature-associated genes. Horizontal lines, labeled with cohorts (European large = 100%, European small = 60%, African small = 37.5%, and African large = 22.2%), indicate the percentage of significant genomic regions from each cohort that localized the stature associated genes. Vertical bars show the distribution of significant regions (%) harboring stature-associated genes in five sets of permutation cohorts that mimicked the empirical cohorts in matching colors.

selection for increased stature has been applied in selected breeds of modern cattle (Ajmone-Marsan *et al.* 2010).

We found 26 bovine genomic regions in European and African *Bos taurus* carrying distinctive signatures of selection. Of these, 12 (46%) contained 30 stature-associated candidate genes derived from comparative mapping. We showed that when cohorts of randomly permuted animals were tested, a much lower proportion (8.3%) of the significant regions contained candidate genes, which is consistent with random expectations (~7%). This suggests that our approach was powerful in capturing bovine genomic regions related to stature. Several of these should be considered novel because they were not previously reported in cattle.

It is noteworthy why different genomic regions are detected in the large and small cohorts within European and African cattle. The cohorts within each cattle type were compared against each other to compute the constituent ( $F_{ST}$ ,  $\Delta DAF$ , and XP-EHH) selection tests of CSS. Two selection tests ( $\Delta DAF$  and XP-EHH) provide directional signatures of selection. This implicates that different gene variants are regulating stature in large and small breeds of cattle (Visscher and Goddard 2011).

Several traits, such as birth weight, growth rate, adult weight, and stature, are found to be correlated with each other and have pleiotropic effects. Some of the genes in the validation regions (see below) have been previously found to be associated with multiple correlated traits. The FAO data on stature have been used to objectively classify cohorts of cattle, and the data were not adjusted for other traits given the nonavailability of FAO records for the correlated traits. Multibreed classification is expected to minimize the breed sampling effects; however, some of the candidate regions may not be directly reflecting selection of stature but could be confounded by some of the other correlated traits.

The significant CSS outside the candidate gene regions (Table S5 and Table S6) may indicate putative genes associated with height or targets of selection for traits other than stature (Utsunomiya *et al.* 2013; Druet *et al.* 2013; Ramey *et al.* 2013; Stella *et al.* 2010; Kemper *et al.* 2014; Perez *et al.* 2014; Porto-Neto *et al.* 2013; Rothhammer *et al.* 2013; Gautier *et al.* 2009). Moreover, several additional significant CSS regions in African cattle have not been reported previously, given the limited reports on selective sweep analyses in the African breeds.

### Validation of known stature-associated genomic regions in cattle

This study validated two important candidate gene regions (regions 6 and 10) linked to stature in European breeds of cattle. Region 6 harboring *NCAPG* and *LCORL* genes located on BTA-6 has frequently shown strong signatures of selection in multiple breeds of cattle (Barendse *et al.* 2009; Druet *et al.* 2013; Flori *et al.* 2009; Gibbs *et al.* 2009; Kemper *et al.* 2014; Mancini *et al.* 2014; Pintus *et al.* 2013; Porto-Neto *et al.* 2013; Rothhammer *et al.* 2013; Stella *et al.* 2010; Qanbari *et al.* 2014; Perez *et al.* 2014). Region 10 harboring *PLAG1* and *CHCHD7* genes on BTA-14 has been associated with height in humans (Gudbjartsson *et al.* 2008; Soranzo *et al.* 2009; Okada *et al.* 2010; Lettre *et al.* 2008; Lango Allen *et al.* 2010) and stature in cattle (Nishimura *et al.* 2012; Pryce *et al.* 2011; Karim *et al.* 2011; Fortes *et al.* 2013; Bolormaa *et al.* 2014). This locus was also identified as a candidate region based on selective sweeps in several cattle breeds that have been under strong selection for body size (Druet *et al.* 2013; Ramey *et al.* 2013; Kemper *et al.* 2014).

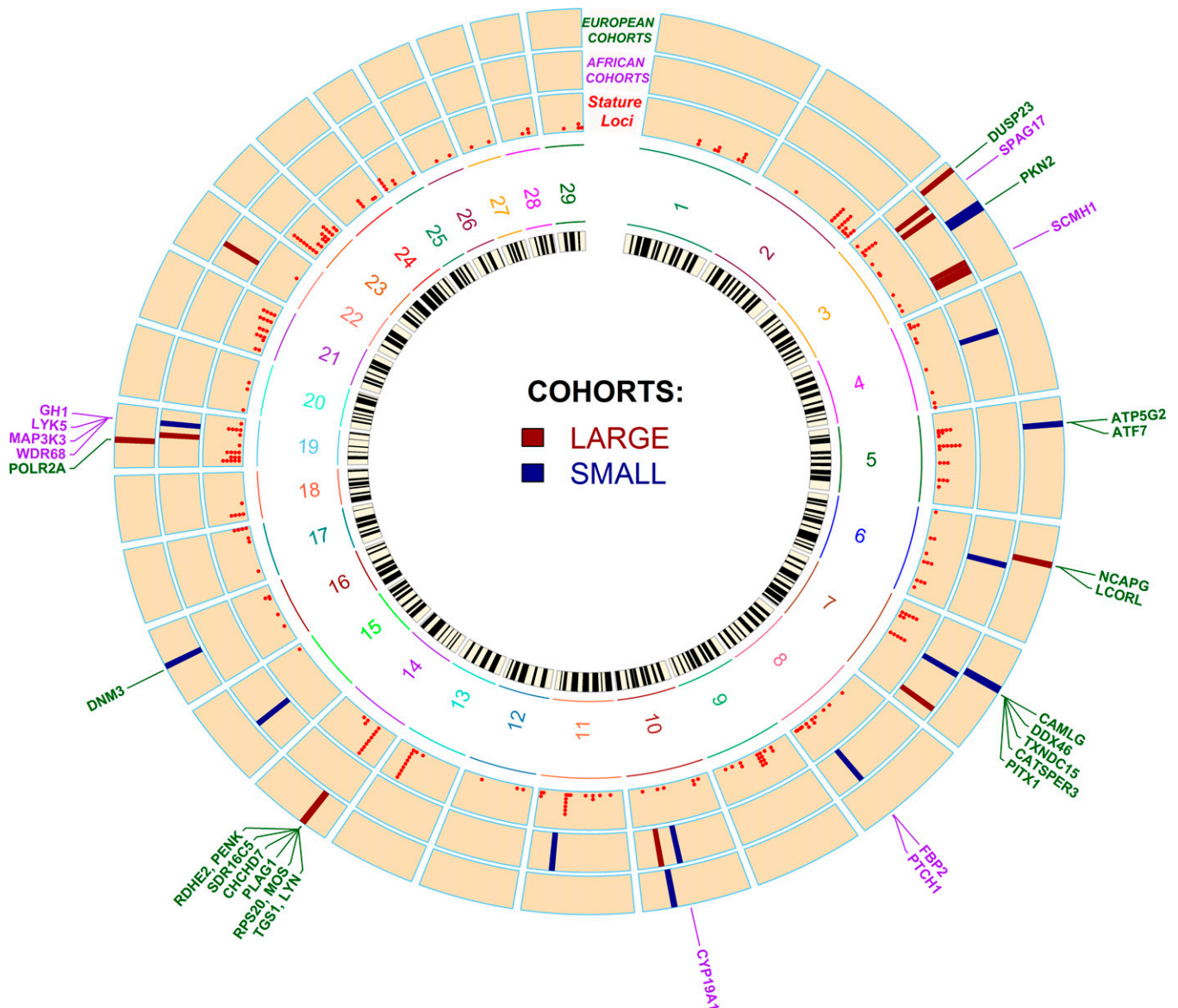
### Genomic regions associated with stature in European *Bos taurus*

Based on strong recent selection for stature in European *Bos taurus*, our comparison of the large cohort with the small cohort successfully localized height-related genes at all significant regions (regions 1, 6, 10, and 11). Regions 1 and 11 can be considered novel and may contain functional variants contributing to the genetic control of bovine stature, because they harbor *DUSP23* (BTA-3: 10.02-10.03 Mb) and *POLR2A* (BTA-19: 27.78-27.80) as candidate genes, respectively (Lanktree *et al.* 2011).

There can be some advantages of limiting body size in smaller breeds, for instance, in relation to ease of calving or resource-limited production systems. As noted above, early domestication favored selective breeding to reduce body size for ease of management (Karim *et al.* 2011; Ajmone-Marsan *et al.* 2010). With very few exceptions (e.g., Dexter), recent history of European *Bos taurus* breeds do not document selective breeding for smaller size. Interestingly, 60% of all significant regions can be considered novel in the small cohort and harbor stature-associated genes, namely *PKN2* at region 3 (Lanktree *et al.* 2011), *ATP5G2* and *ATF7* at region 5 (Okada *et al.* 2010), and *CAMLG*, *DDX46*, *TXNDC15*, *CATSPER3*, and *PITX1* at region 7 (Gudbjartsson *et al.* 2008). The functional role of these candidate genes in relation to stature is not well-documented.

### Genomic regions associated with stature in African *Bos taurus*

Five candidate regions (2, 4, 8, 9, 12) in the African cohorts harbor nine candidate genes associated with human height identified by GWAS (Gudbjartsson *et al.* 2008; Soranzo *et al.* 2009; Okada *et al.* 2010; Lanktree *et al.* 2011; Kim *et al.* 2010; Weedon *et al.* 2008). Previously, *FBP2* and *PTCH1* genes underlying region 8 (BTA-8) have been associated with stature in European dairy and beef breeds,



**Figure 4** Circos image showing localization of genome-wide significant regions and candidate genes across all cohorts. The inner circle shows a bovine chromosome (29 autosomes) ideogram. The first outer circle, labeled as *Stature Loci*, shows the location of 134 loci listed in Table S4, and each red dot represents a stature-associated candidate gene (clustered red dots, within a locus, located on top of each other represent multiple neighboring candidate genes implicated in the same or different GWAS studies). Two circles, in the middle, labeled as *EUROPEAN COHORTS* and *AFRICAN COHORTS* show cohort-wise genomic regions identified by top 0.1% CSS (legends in the center for cohort-wise colored bars). The outer track shows genomic locations of 30 candidate genes within the 12 significant signatures of selection linked to bovine stature. Gene colors indicate whether they were identified in European (green) or African (purple) cattle types.

respectively (Weedon *et al.* 2008; Kim *et al.* 2010; Lanktree *et al.* 2011; Pryce *et al.* 2011). In region 12 (BTA-19), the growth hormone genes (*GH1*, *GH2*) are strong candidates for dairy production traits (Hayes *et al.* 2009; Rothhammer *et al.* 2013; Mullen *et al.* 2010). Moreover, three genes (*WDR68*, *MAP3K3*, and *LYK5*) neighboring growth hormone genes have also been associated with human height in multiple GWAS reports (Gudbjartsson *et al.* 2008; Soranzo *et al.* 2009; Lanktree *et al.* 2011). The remaining three candidate regions, *i.e.*, regions 2, 4, and 9, harbor the stature-associated *SPAG17*, *SCMH1*, and *CYP19A1* genes, respectively (Weedon *et al.* 2008; Kim *et al.* 2010; Okada *et al.* 2010; Lanktree *et al.* 2011).

Overall, four of the five candidate regions identified in the African cohorts are novel for their association with bovine stature. This suggests that the strategy of combining CSS and multibreed panels showed utility in the African cohorts also, although the overall proportion of candidate gene regions detected was lower compared to European cohorts. High levels of *Bos indicus* admixture, smaller sample size, and ascertainment bias of the Illumina BovineSNP50 chip in the African *Bos taurus* breeds may have impacted the power to detect regions under selection (Randhawa *et al.* 2014; Decker *et al.* 2009; Gautier *et al.* 2010; Gautier and Naves 2011).



■ **Table 2 Significant selection regions that harbor candidate genes for stature identified by composite selection signals in all analyses of European and African cohorts**

Region	BTA	Location (Mb)	Top 0.1 (1)% SNPs	Candidate Genes	Candidate Population	Source of Discovery (References)
1	3	9.35–10.06	1 (15)	<i>DUSP23</i>	European large	Human (Lanktree <i>et al.</i> 2011)
2	3	24.00–24.83	1 (9)	<i>SPAG17</i>	African large	Human (Weedon <i>et al.</i> 2008; Kim <i>et al.</i> 2010)
3	3	50.62–56.96	5 (47)	<i>PKN2</i>	European small	Human (Lanktree <i>et al.</i> 2011)
4	3	105.09–106.96	11 (30)	<i>SCMH1</i>	African large	Human (Weedon <i>et al.</i> 2008)
5	5	26.88–27.50	4 (8)	<i>ATP5G2, ATF7</i>	European small	Human (Okada <i>et al.</i> 2010)
6	6	38.29–39.75	10 (29)	<i>NCAPG, LCORL</i>	European large	Human, cattle, and horse (Gudbjartsson <i>et al.</i> 2008; Weedon <i>et al.</i> 2008; Soranzo <i>et al.</i> 2009; Lango Allen <i>et al.</i> 2010; Okada <i>et al.</i> 2010; Pryce <i>et al.</i> 2011)
7	7	43.42–46.75	20 (36)	<i>CAMLG, DDX46, TXNDC15, CATSPER3, PITX1<sup>a</sup></i>	European small	Human (Gudbjartsson <i>et al.</i> 2008)
8	8	83.05–83.89	3 (13)	<i>FBP2, PTCH1</i>	African small	Human and cattle (Weedon <i>et al.</i> 2008; Kim <i>et al.</i> 2010; Lanktree <i>et al.</i> 2011; Pryce <i>et al.</i> 2011)
9	10	59.38–59.97	1 (12)	<i>CYP19A1</i>	African small	Human (Okada <i>et al.</i> 2010; Lanktree <i>et al.</i> 2011)
10	14	24.79–28.25	21 (66)	<i>TGS1, LYN, RPS20, MOS, PLAG1, CHCHD7, SDR16C5, RDHE2, PENK</i>	European large	Human, cattle, and horse (Gudbjartsson <i>et al.</i> 2008; Soranzo <i>et al.</i> 2009; Okada <i>et al.</i> 2010; Lango Allen <i>et al.</i> 2010; Lettre <i>et al.</i> 2008; Pryce <i>et al.</i> 2011)
11	19	27.64–28.63	6 (16)	<i>POLR2A</i>	European large	Human (Lanktree <i>et al.</i> 2011)
12	19	48.01–49.01	1 (14)	<i>WDR68, MAP3K3, LYK5, GH1</i>	African small	Human (Gudbjartsson <i>et al.</i> 2008; Soranzo <i>et al.</i> 2009; Lanktree <i>et al.</i> 2011)

<sup>a</sup> Candidate gene locus is localized slightly over 1 Mb in the surrounding of a significant region.

### Comparison of significant CSS regions between European and African *Bos taurus*

No candidate gene regions were found in common across the European and African cattle types. This may be due to the long-term differences in natural and production-oriented selective pressures in these diverse genetic archetypes (Edea *et al.* 2014; Gautier *et al.* 2009; Gautier and Naves 2011; Gibbs *et al.* 2009; Randhawa *et al.* 2014). Recent attempts to find the transferability of height-associated loci across human populations have shown comparable results (low consistency) for finding the same genetic variants when discoveries made in non-African ethnicities were compared in populations of African ancestry (Kang *et al.* 2010; Shriner *et al.* 2009). Our results indicate that there may be a persistent trend of dissimilarities for African genealogies across species, and some environmental factors would have shaped a comparatively different genetic architecture. We suggest that investigation of several other African breeds with sufficient sample sizes can further elaborate the nature of different candidate gene regions.

A region on BTA-13 harboring *UQCC* and *GDF5* genes, which we putatively linked to bovine stature (Randhawa *et al.* 2014), was also not detected in this study. The absence of BTA-13 genes (*UQCC* and *GDF5*) as the possible candidate for bovine stature in this study are due to the sensitivity of CSS (and its constituent tests of selection) to the sample composition because the cohort composition between the two papers is quite different regarding inclusion and exclusion of

several breeds. Moreover, another gene, *ASIP* (related to coat color), was also located at the BTA-13 peak, and *ASIP* may also be another plausible candidate of the signatures of selection detected in the previous study.

In comparison to previous findings in cattle, we note that some of the known genes and regions related to bovine stature were not detected here, mainly because they were identified in *Bos indicus* and composite breeds (Pryce *et al.* 2011). Moreover, genetic architecture of stature is not considered fully analogous in cattle and human; our results suggest that many common genes might be participating in the physiological control of stature in the two species (Pryce *et al.* 2011).

Characterization of biological function of candidate genes could help our understanding of mechanisms underlying stature diversity. Future research should focus on regional genomic sequencing (involving multiple genes) rather than single gene-centric approaches. In addition, gene networks and functional pathway analyses on subsets of genes identified by genome-wide association or selection scans can help uncover the biological processes in complex traits.

### CONCLUSION

This study demonstrates an approach to detect signatures of selection for complex traits. We successfully used CSS in the analysis of the genetic architecture of a complex polygenic trait using multibreed panels. Our results demonstrated the localization of genomic regions harboring candidate genes of known effect related to body

development, skeletal growth, and stature in cattle. We identified 26 regions with strong signatures of selection for stature, of which 12 regions contained known genes linked to height or stature and 9 could be considered novel in European and African *Bos taurus*. These results should be of interest for future investigations to characterize causal variants to explain their functional role in the diversity of bovine stature. Implementation of new tools, such as CSS, can resolve the selected genomic regions underlying the complex traits in livestock species. We suggest that the application of genome-wide selection scans and GWAS using larger resource populations with phenotypic information can further improve the fine-mapping of causal mutations controlling complex traits.

## LITERATURE CITED

- Ajmone-Marsan, P., J. F. Garcia, and J. A. Lenstra, 2010 On the origin of cattle: How aurochs became cattle and colonized the world. *Evol. Anthropol.* 19: 148–157.
- Akey, J. M., A. L. Ruhe, D. T. Akey, A. K. Wong, C. F. Connelly *et al.*, 2010 Tracking footprints of artificial selection in the dog genome. *Proc. Natl. Acad. Sci. USA* 107: 1160–1165.
- Barendse, W., B. E. Harrison, R. J. Bunch, M. B. Thomas, and L. B. Turner, 2009 Genome wide signatures of positive selection: The comparison of independent samples and the identification of regions associated to traits. *BMC Genomics* 10: 178.
- Bolormaa, S., J. E. Pryce, A. Reverter, Y. Zhang, W. Barendse *et al.*, 2014 A multi-trait, meta-analysis for detecting pleiotropic polymorphisms for stature, fatness and reproduction in beef cattle. *PLoS Genet.* 10: e1004198.
- Browning, B. L., and S. R. Browning, 2009 A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* 84: 210–223.
- Crisci, J. L., Y.-P. Poh, A. Bean, A. Simkin, and J. D. Jensen, 2012 Recent progress in polymorphism-based population genetic inference. *J. Hered.* 103: 287–296.
- DAD-IS, 2014 Domestic Animal Diversity Information System (DAD-IS), Food and Agriculture Organization of the United Nations. Available at: <http://www.fao.org/dad-is/>. Accessed: February 1, 2014.
- Decker, J., D. Vasco, S. McKay, M. McClure, M. Rolf *et al.*, 2012 A novel analytical method, Birth Date Selection Mapping, detects response of the Angus (*Bos taurus*) genome to selection on complex traits. *BMC Genomics* 13: 606.
- Decker, J. E., J. C. Pires, G. C. Conant, S. D. McKay, M. P. Heaton *et al.*, 2009 Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. *Proc. Natl. Acad. Sci. USA* 106: 18644–18649.
- Deng, F. Y., S. S. Dong, X. H. Xu, Y. J. Liu, Y. Z. Liu *et al.*, 2013 Genome-wide association study identified *UQCC* locus for spine bone size in humans. *Bone* 53: 129–133.
- Druet, T., L. Pérez-Pardal, C. Charlier, and M. Gautier, 2013 Identification of large selective sweeps associated with major genes in cattle. *Anim. Genet.* 44: 758–762.
- Edea, Z., H. Dadi, S. W. Kim, J. H. Park, G. H. Shin *et al.*, 2014 Linkage disequilibrium and genomic scan to detect selective loci in cattle populations adapted to different ecological conditions in Ethiopia. *J. Anim. Breed. Genet.* .10.1111/jbg.12083
- Flori, L., S. Fritz, F. Jaffrézic, M. Boussaha, I. Gut *et al.*, 2009 The genome response to artificial selection: A case study in dairy cattle. *PLoS ONE* 4: e6595.
- Fortes, M. R. S., K. Kemper, S. Sasazaki, A. Reverter, J. E. Pryce *et al.*, 2013 Evidence for pleiotropism and recent selection in the *PLAG1* region in Australian Beef cattle. *Anim. Genet.* 44: 636–647.
- Gautier, M., L. Flori, A. Riebler, F. Jaffrézic, D. Laloe *et al.*, 2009 A whole genome Bayesian scan for adaptive genetic divergence in West African cattle. *BMC Genomics* 10: 550.
- Gautier, M., D. Laloë, and K. Moazami-Goudarzi, 2010 Insights into the genetic history of French cattle from dense SNP data on 47 Worldwide breeds. *PLoS ONE* 5: e13038.
- Gautier, M., and M. Naves, 2011 Footprints of selection in the ancestral admixture of a New World Creole cattle breed. *Mol. Ecol.* 20: 3128–3143.
- Gibbs, R. A., J. F. Taylor, C. P. Van Tassel, W. Barendse, K. A. Eversole *et al.*, 2009 Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* 324: 528–532.
- Grossman, S. R., I. Shylakhter, E. K. Karlsson, E. H. Byrne, S. Morales *et al.*, 2010 A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science* 327: 883–886.
- Gudbjartsson, D. F., G. B. Walters, G. Thorleifsson, H. Stefansson, B. V. Halldorsson *et al.*, 2008 Many sequence variants affecting diversity of adult human height. *Nat. Genet.* 40: 609–615.
- Hancock, A. M., and A. Di Rienzo, 2008 Detecting the genetic signature of natural selection in human populations: Models, methods, and data. *Annu. Rev. Anthropol.* 37: 197–217.
- Hayes, B. J., S. Lien, H. Nilsen, H. G. Olsen, P. Berg *et al.*, 2008 The origin of selection signatures on bovine chromosome 6. *Anim. Genet.* 39: 105–111.
- Hayes, B. J., A. J. Chamberlain, S. Maceachern, K. Savin, H. McPartlan *et al.*, 2009 A genome map of divergent artificial selection between *Bos taurus* dairy cattle and *Bos taurus* beef cattle. *Anim. Genet.* 40: 176–184.
- Hohenlohe, P. A., P. C. Phillips, and W. A. Cresko, 2010 Using population genomics to detect selection in natural populations: Key concepts and methodological considerations. *Int. J. Plant Sci.* 171: 1059–1071.
- Hoshiba, H., K. Setoguchi, T. Watanabe, A. Kinoshita, K. Mizoshita *et al.*, 2013 Comparison of the effects explained by variations in the bovine *PLAG1* and *NCAPG* genes on daily body weight gain, linear skeletal measurements and carcass traits in Japanese Black steers from a progeny testing program. *Anim. Sci. J.* 84: 529–534.
- Kang, S. J., C. W. K. Chiang, C. D. Palmer, B. O. Tayo, G. Lettre *et al.*, 2010 Genome-wide association of anthropometric traits in African- and African-derived populations. *Hum. Mol. Genet.* 19: 2725–2738.
- Karim, L., H. Takeda, L. Lin, T. Druet, J. A. C. Arias *et al.*, 2011 Variants modulating the expression of a chromosome domain encompassing *PLAG1* influence bovine stature. *Nat. Genet.* 43: 405–413.
- Kemper, K. E., and M. E. Goddard, 2012 Understanding and predicting complex traits: knowledge from cattle. *Hum. Mol. Genet.* 21: R45–R51.
- Kemper, K., S. Saxton, S. Bolormaa, B. Hayes, and M. Goddard, 2014 Selection for complex traits leaves little or no classic signatures of selection. *BMC Genomics* 15: 246.
- Kijas, J. W., J. A. Lenstra, B. Hayes, S. Boitard, L. R. Porto Neto *et al.*, 2012 Genome-wide analysis of the World's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol.* 10: e1001258.
- Kim, J.-J., H.-I. Lee, T. Park, K. Kim, J.-E. Lee *et al.*, 2010 Identification of 15 loci influencing height in a Korean population. *J. Hum. Genet.* 55: 27–31.
- Lango Allen, H., K. Estrada, G. Lettre, S. I. Berndt, M. N. Weedon *et al.*, 2010 Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 467: 832–838.
- Lanktree, M. B., Y. Guo, M. Murtaza, J. T. Glessner, S. D. Bailey *et al.*, 2011 Meta-analysis of dense gene-centric association studies reveals common and uncommon variants associated with height. *Am. J. Hum. Genet.* 88: 6–18.
- Lenstra, J. A., L. F. Groeneveld, H. Eding, J. Kantanen, J. L. Williams *et al.*, 2011 Molecular tools and analytical approaches for the characterization of farm animal genetic diversity. *Anim. Genet.* 43: 483–502.
- Lettre, G., 2011 Recent progress in the study of the genetics of height. *Hum. Genet.* 129: 465–472.
- Lettre, G., A. U. Jackson, C. Gieger, F. R. Schumacher, S. I. Berndt *et al.*, 2008 Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat. Genet.* 40: 584–591.
- Li, M.-X., X.-M. Sun, L.-Z. Zhang, J. Wang, Y.-Z. Huang *et al.*, 2013 A novel c.-274C>G polymorphism in bovine *SIRT1* gene contributes to diminished promoter activity and is associated with increased body size. *Anim. Genet.* 44: 584–587.
- Lindholm-Perry, A., A. Sexten, L. Kuehn, T. Smith, D. King *et al.*, 2011 Association, effects and validation of polymorphisms within the

- NCAPG - LCORL locus located on BTA6 with feed intake, gain, meat and carcass traits in beef cattle. *BMC Genet.* 12: 103.
- Mancini, G., M. Gargani, G. Chillemi, E. Nicolazzi, P. A. Marsan *et al.*, 2014 Signatures of selection in five Italian cattle breeds detected by a 54K SNP panel. *Mol. Biol. Rep.* 41: 957–965.
- Matukumalli, L. K., C. T. Lawley, R. D. Schnabel, J. F. Taylor, M. F. Allan *et al.*, 2009 Development and characterization of a high density SNP genotyping assay for cattle. *PLoS ONE* 4: e5350.
- Mullen, M. P., D. P. Berry, D. J. Howard, M. G. Diskin, C. O. Lynch *et al.*, 2010 Associations between novel single nucleotide polymorphisms in the *Bos taurus* growth hormone gene and performance traits in Holstein-Friesian dairy cattle. *J. Dairy Sci.* 93: 5959–5969.
- Nishimura, S., T. Watanabe, K. Mizoshita, K. Tatsuda, T. Fujita *et al.*, 2012 Genome-wide association study identified three major QTL for carcass weight including the *PLAG1-CHCHD7* QTN for stature in Japanese Black cattle. *BMC Genet.* 13: 40.
- Okada, Y., Y. Kamatani, A. Takahashi, K. Matsuda, N. Hosono *et al.*, 2010 A genome-wide association study in 19 633 Japanese subjects identified *LHX3-QSOX2* and *IGF1* as adult height loci. *Hum. Mol. Genet.* 19: 2303–2312.
- Oleksyk, T. K., M. W. Smith, and S. J. O'Brien, 2010 Genome-wide scans for footprints of natural selection. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365: 185–205.
- Perez, O., A. Brien, Y. Utsunomiya, G. Meszaros, D. Bickhart *et al.*, 2014 Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. *Genet. Sel. Evol.* 46: 19.
- Petersen, J. L., J. R. Mickelson, A. K. Rendahl, S. J. Valberg, L. S. Andersson *et al.*, 2013 Genome-wide analysis reveals selection for important traits in domestic horse breeds. *PLoS Genet.* 9: e1003211.
- Pickrell, J., G. Coop, J. Novembre, S. Kudaravalli, J. Li *et al.*, 2009 Signals of recent positive selection in a worldwide sample of human populations. *Genome Res.* 19: 826–837.
- Pintus, E., S. Sorbolini, A. Albera, G. Gaspa, C. Dimauro *et al.*, 2013 Use of locally weighted scatterplot smoothing (LOWESS) regression to study selection signatures in Piedmontese and Italian Brown cattle breeds. *Anim. Genet.* 45: 1–11.
- Porto-Neto, L., T. Sonstegard, G. Liu, D. Bickhart, M. Da Silva *et al.*, 2013 Genomic divergence of zebu and taurine cattle identified through high-density SNP genotyping. *BMC Genomics* 14: 876.
- Pryce, J. E., B. J. Hayes, S. Bolormaa, and M. E. Goddard, 2011 Polymorphic regions affecting human height also control stature in cattle. *Genetics* 187: 981–984.
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira *et al.*, 2007 PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81: 559–575.
- Qanbari, S., D. Gianola, B. Hayes, F. Schenkel, S. Miller *et al.*, 2011 Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. *BMC Genomics* 12: 318.
- Qanbari, S., H. Pausch, S. Jansen, M. Somel, T. M. Strom *et al.*, 2014 Classic selective sweeps revealed by massive sequencing in cattle. *PLoS Genet.* 10: e1004148.
- Ramey, H., J. Decker, S. McKay, M. Rolf, R. Schnabel *et al.*, 2013 Detection of selective sweeps in cattle using genome-wide SNP data. *BMC Genomics* 14: 382.
- Randhawa, I. A. S., M. S. Khatkar, P. C. Thomson, and H. W. Raadsma, 2014 Composite selection signals can localize the trait specific genomic regions in multi-breed populations of cattle and sheep. *BMC Genet.* 15: 34.
- Rothhammer, S., D. Seichter, M. Forster, and I. Medugorac, 2013 A genome-wide scan for signatures of differential artificial selection in ten cattle breeds. *BMC Genomics* 14: 908.
- Sabeti, P. C., P. Varilly, B. Fry, J. Lohmueller, E. Hostetter *et al.*, 2007 Genome-wide detection and characterization of positive selection in human populations. *Nature* 449: 913–918.
- Sanna, S., A. U. Jackson, R. Nagaraja, C. J. Willer, W.-M. Chen *et al.*, 2008 Common variants in the *GDF5-UQCC* region are associated with variation in human height. *Nat. Genet.* 40: 198–203.
- Seichter, D., I. Russ, S. Rothhammer, J. Eder, M. Förster *et al.*, 2012 SNP-based association mapping of the polled gene in divergent cattle breeds. *Anim. Genet.* 43: 595–598.
- Shiura, H., K. Nakamura, T. Hikichi, T. Hino, K. Oda *et al.*, 2009 Paternal deletion of *Meg1/Grb10* DMR causes maternalization of the *Meg1/Grb10* cluster in mouse proximal chromosome 11 leading to severe pre- and postnatal growth retardation. *Hum. Mol. Genet.* 18: 1424–1438.
- Shriner, D., A. Adeyemo, N. P. Gerry, A. Herbert, G. Chen *et al.*, 2009 Transferability and fine-mapping of genome-wide associated loci for adult height across human populations. *PLoS ONE* 4: e8398.
- Signer-Hasler, H., C. Flury, B. Haase, D. Burger, H. Simianer *et al.*, 2012 A genome-wide association study reveals loci influencing height and other conformation traits in horses. *PLoS ONE* 7: e37282.
- Soranzo, N., F. Rivadeneira, U. Chinappen-Horsley, I. Malkina, J. B. Richards *et al.*, 2009 Meta-analysis of genome-wide scans for human adult stature identifies novel loci and associations with measures of skeletal frame size. *PLoS Genet.* 5: e1000445.
- Sovio, U., A. J. Bennett, I. Y. Millwood, J. Molitor, P. F. O'Reilly *et al.*, 2009 Genetic determinants of height growth assessed longitudinally from infancy to adulthood in the northern Finland birth cohort 1966. *PLoS Genet.* 5: e1000409.
- Stella, A., P. Ajmone-Marsan, B. Lazzari, and P. Boettcher, 2010 Identification of selection signatures in cattle breeds selected for dairy production. *Genetics* 185: 1451–1461.
- Tetens, J., P. Widmann, C. Kühn, and G. Thaller, 2013 A genome-wide association study indicates *LCORL/NCAPG* as a candidate locus for withers height in German Warmblood horses. *Anim. Genet.* 44: 467–471.
- Utsunomiya, Y. T., A. M. Pérez O'Brien, T. S. Sonstegard, C. P. Van Tassel, A. S. do Carmo *et al.*, 2013 Detecting loci under recent positive selection in dairy and beef cattle by combining different genome-wide scan methods. *PLoS ONE* 8: e64280.
- Visscher, P. M., and M. E. Goddard, 2011 Cattle gain stature. *Nat. Genet.* 43: 397–398.
- Weedon, M. N., H. Lango, C. M. Lindgren, C. Wallace, D. M. Evans *et al.*, 2008 Genome-wide association analysis identifies 20 loci that influence adult height. *Nat. Genet.* 40: 575–583.
- Weir, B. S., and C. C. Cockerham, 1984 Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.
- Zhang, H., S.-Z. Wang, Z.-P. Wang, Y. Da, N. Wang *et al.*, 2012 A genome-wide scan of selective sweeps in two broiler chicken lines divergently selected for abdominal fat content. *BMC Genomics* 13: 704.

Communicating editor: D. L. Bannasch