Genome-wide association and genotype by environment interactions for growth traits in U.S. Gelbvieh cattle

Johanna L. Smith 1, Miranda L. Wilson 1, Sara M. Nilson 2, Troy N. Rowan 2,3, David L. Oldeschulte 1, Robert D. Schnabel 2,3,4, Jared E. Decker 2,3,4 and Christopher M. Seabury 1*

Abstract

Background: Single nucleotide polymorphism (SNP) arrays have facilitated discovery of genetic markers associated with complex traits in domestic cattle; thereby enabling modern breeding and selection programs. Genome-wide association analyses (GWAA) for growth traits were conducted on 10,837 geographically diverse U.S. Gelbvieh cattle using a union set of 856,527 imputed SNPs. Birth weight (BW), weaning weight (WW), and yearling weight (YW) were analyzed using GEMMA and EMMAX (via imputed genotypes). Genotype-by-environment (GxE) interactions were also investigated.

Results: GEMMA and EMMAX produced moderate marker-based heritability estimates that were similar for BW (0.36–0.37, SE = 0.02–0.06), WW (0.27–0.29, SE = 0.01), and YW (0.39–0.41, SE = 0.01–0.02). GWAA using 856K imputed SNPs (GEMMA; EMMAX) revealed common positional candidate genes underlying pleiotropic QTL for Gelbvieh growth traits on BTA6, BTA7, BTA14, and BTA20. The estimated proportion of phenotypic variance explained (PVE) by the lead SNP defining these QTL (EMMAX) was larger and most similar for BW and YW, and smaller for WW. Collectively, GWAA (GEMMA; EMMAX) produced a highly concordant set of BW, WW, and YW QTL that met a nominal significance level (P ≤ 1e-05), with prioritization of common positional candidate genes; including genes previously associated with stature, feed efficiency, and growth traits (i.e., PLAG1, NCAPG, LCORL, ARRD1C3, STC2). Genotype-by-environment QTL were not consistent among traits at the nominal significance threshold (P ≤ 1e-05); although some shared QTL were apparent at less stringent significance thresholds (i.e., P ≤ 2e-05).

Conclusions: Pleiotropic QTL for growth traits were detected on BTA6, BTA7, BTA14, and BTA20 for U.S. Gelbvieh beef cattle. Seven QTL detected for Gelbvieh growth traits were also recently detected for feed efficiency and growth traits in U.S. Angus, SimAngus, and Hereford cattle. Marker-based heritability estimates and the detection of pleiotropic QTL segregating in multiple breeds support the implementation of multiple-breed genomic selection.

Keywords: GWAA, QTL, Genotype-by-environment interaction, Growth traits, Gelbvieh

Background

Growth traits are commonly recorded and used as selection criteria within modern beef cattle breeding programs and production systems; primarily because of their correlation with increased overall meat production and other economically important traits [1–4]. Some of the most commonly investigated growth traits include birth weight (BW), weaning weight (WW) and yearling weight (YW); with BW considered as both a production indicator, and a primary selection criterion for improving calving ease by reducing dystocia events [1, 2, 5–7]. Moreover, while previous studies have demonstrated that low estimated breeding values (EBVs) for BW are associated with reductions in both calf viability [6] and growth rates [5, 7], increased dystocia rates may also occur if sires with high EBVs for BW are used in conjunction with dams that possess small pelvic size. Therefore, modern beef breeding programs and production systems generally strive to increase calving ease, and...
maximize other growth-related traits such as WW and YW, particularly considering the known correlations between growth traits and other economically important carcass and reproductive traits [3, 5, 7].

Given the increasing economic importance of growth traits in beef cattle, a number of studies have sought to identify quantitative trait loci (QTL) influencing bovine body weight, growth, and aspects of stature, including both linkage studies and modern genome-wide association analyses [2, 8–13]. Several recent studies have also established moderate heritability estimates for bovine growth traits in U.S. beef cattle including BW, WW, and YW [14–17], with a number of relevant QTL and positional candidate genes identified to date, including orthologous genes that affect both human and bovine height [2, 18–22]. Notably, with the advent of the bovine genome assembly [23], the development of the Illumina Bovine SNP50 and 778K HD assays [23, 24], and more recently, the demonstrated ability to impute high density genotypes with high accuracy [25], an industry-supported research framework [26] has emerged supporting research frameworks that were developed for elucidating potential for eliciting economic gains resulting from the implementation of genomic selection [30].

**GWAA for BW, WW, and YW in U.S. Gelbvieh beef cattle**

The results of our 856K single-marker analyses for BW (GEMMA; EMMAX) [27–29] are shown in Fig. 1 and in Figure S1 (Additional File 1), with detailed summary data for QTL detected by GEMMA and EMMAX described in Table 2 and Table S1, respectively. A comparison of GEMMA and EMMAX results revealed a concordant set of QTL defined by lead SNPs (i.e., the most strongly associated SNP within a QTL region) which met a nominal significance threshold ($P \leq 1e-05$) [31] (Table 2, Table S1, Additional File 1, Additional File 2). Specifically, QTL signals for BW were detected on BTA6, BTA7, BTA14, and BTA20 across both analyses (Table 2, Table S1, Additional File 1), and included an array of positional candidate genes generally involved in diverse aspects of mammalian growth and development (i.e., CCSER1, ST18, RPI/XKR4, SLIT2, STC2, IBSP) as well as bovine growth (i.e., NCAPG, KCNIP4, PLAG1, and production traits (i.e., IMPAD1/FAM110B, HERC6/PPM1K) [2, 13, 14, 18, 21, 22, 30–60]. Interestingly, the lead SNP defining the BW QTL detected on BTA14 (14.25 Mb) was located in PLAG1, thereby further supporting the involvement of this gene in various aspects of bovine growth and stature across breeds [2, 14, 18, 21, 30, 32–34]. Additionally, all but one (i.e., NCAPG, exon 9) of the lead SNPs for the detected Gelbvieh BW QTL (GEMMA, EMMAX) were noncoding variants (Table 2, Table S1, Additional File 1). Genomic inflation factors and correlation coefficients for $P$-values obtained from all BW analyses are shown in Tables S2-S3 (Additional File 1).

Single-marker analyses (856K) for WW in U.S. Gelbvieh beef cattle (GEMMA; EMMAX) revealed several of the same QTL detected for BW (Table 3, Fig. 2, Table S4, Figure S2, Additional File 1), thus providing statistical support for pleiotropic QTL located on BTA6 (i.e., NCAPG, CCSER1, KCNIP4, HERC6/PPM1K, LOC782905/SLIT2, LOC100336621/LOC1049772717) as well as BTA14 (i.e., PLAG1, XKR4, IMPAD1/FAM110B). The lead SNPs for Gelbvieh BW and WW QTL detected on BTA20 (20.05 Mb) suggested proximal but independent causal mutations, thus implicating the potential involvement of at least three positional candidate genes (LOC104975192/STC2, ERGIC1). A detailed summary of lead and supporting SNPs for pleiotropic QTL is provided in Additional File 2. Beyond evidence for pleiotropy, four additional Gelbvieh WW QTL.
were also detected on BTA5 (5.60 Mb), BTA6 (6.31 Mb, 6.37 Mb) and BTA28 (28.37 Mb; Table 3, Figure 2, Table S4, Figure S2, Additional File 1). Among the additional QTL detected, several positional candidate genes have been implicated in aspects of development (\textit{UNC5C}, \textit{SNCA}/\textit{GPRIN3}) and immune function (\textit{SH2D4B})\cite{61–67}. An investigation of all lead SNPs for the detected Gelbvieh WW QTL revealed 13 noncoding variants and one nonsynonymous variant (Table 3, Table S4, Additional File 1). Genomic inflation factors and correlation coefficients for \( P \)-values obtained from all WW analyses are presented in Tables S2 and S3 (Additional File 1).

Consistent with our analyses of BW and WW, our single-marker analyses (856K) for YW in U.S. Gelbvieh beef cattle again revealed evidence for pleiotropic QTL located on BTA6 and BTA14 (Table 4, Fig. 3, Table S5, Figure S3, Additional File 1). Specifically, the results obtained from our analyses of BW, WW, and YW revealed some common QTL signals for all investigated traits on BTA6 (6.36 Mb, 6.38 Mb, 6.39 Mb, 6.41 Mb, 6.42 Mb) and BTA14 (14.24 Mb, 14.25 Mb, 14.26 Mb). Likewise, the lead SNPs defining these QTL also resulted in the prioritization of the same positional candidate genes on BTA6 (i.e., \textit{LCORL}, \textit{KCNIPI4}, \textit{HERC6}/\textit{PPM1K}, \textit{SLIT2}, \textit{CCSER1}) and BTA14 (i.e., \textit{PLAG1}, \textit{IMPA1}/\textit{FAM110B}, \textit{RP1/XKR4}). Together with pleiotropic signals on BTA6 and BTA14, eight additional YW QTL were also detected; including one QTL (7.93 Mb) that was also found to influence Gelbvieh BW (Table 4, Table S5, Additional File 1). Positional candidate genes for these QTL have been implicated in diverse aspects of growth and development as well as bovine production traits (i.e., \textit{SNCA}/\textit{GPRIN3}, \textit{SLIT2}, \textit{NSMAF}, \textit{LOC101905238}/\textit{ARRDC3}), bovine milk traits (i.e., \textit{PPARGC1A}), and chromatin modification (i.e., \textit{IWS1})\cite{68–71}. Relevant to YW, it should also be noted that several of the pleiotropic QTL detected for U.S. Gelbvieh in this study have also been detected for mid-test metabolic weight in U.S. SimAngus beef cattle (6.39 Mb, 14.24 Mb, 14.25 Mb, 14.26 Mb)\cite{30}. Moreover, Gelbvieh QTL (BW, YW) detected on BTA14 and BTA7 have also been detected for Angus residual feed intake (14.27 Mb), and Hereford average daily gain (7.93 Mb)\cite{30}. An investigation of all lead SNPs for the detected Gelbvieh YW QTL revealed 16 noncoding variants (Table 4, Table S5, Additional File 1). Genomic inflation factors and correlation coefficients for \( P \)-values obtained from all YW analyses are shown in Tables S2-S3 (Additional File 1).

\textbf{GxE GWAA for BW, WW, and YW in U.S. Gelbvieh beef cattle}

To investigate the potential for significant GxE interactions in relation to BW, WW, and YW in U.S. Gelbvieh beef cattle, we conducted six additional single-marker (856K) analyses using both GEMMA and EMMAX\cite{27–29}. For all

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
\textbf{Trait} & \textbf{GEMMA\textsuperscript{a}} & \textbf{GEMMA\textsuperscript{b}} & \textbf{GEMMA\textsuperscript{c}} & \textbf{EMMAX\textsuperscript{a}} & \textbf{EMMAX\textsuperscript{b}} & \textbf{EMMAX\textsuperscript{c}} \\
\hline
\textbf{BW} & 0.36 & 0.02 & 15.65 & 27.62 & 0.37 & 0.06 \\
\textbf{WW} & 0.27 & 0.01 & 712.07 & 1910.71 & 0.29 & 0.01 \\
\textbf{YW} & 0.39 & 0.02 & 2751.21 & 4242.85 & 0.41 & 0.01 \\
\hline
\end{tabular}
\caption{Variance component analysis with marker-based heritability estimates}
\end{table}

\textsuperscript{a} GEMMA chip heritability \cite{27}; EMMAX pseudo-heritability \cite{28, 29}.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Fig_1.png}
\caption{Birth weight (BW) QTL Manhattan plot with GEMMA \(-\log_{10} P\)-values. Lead and supporting SNPs for QTL represented at or above the blue line (\( P \leq 1e-05; -\log_{10} P\)-values \( \geq 5.00 \)) for \textit{n} = 10,837 U.S. Gelbvieh beef cattle. A summary of all markers passing the nominal significance threshold \cite{31} is presented in Table 2.}
\end{figure}
analyses, we included a variable for Gelbvieh geographic zone, which was generated via K-means clustering using thirty-year U.S. climate data, and treated as an interaction term (See Methods). Notably, a BW GxE QTL detected on BTA2 (2_32 Mb; lead SNP is intergenic) revealed multiple biologically relevant positional candidate genes, including GRB14, which has been shown to regulate insulin in mice [72], and FIGN, which has been associated with plasma folate levels in humans (Fig. 4, Table 5, Additional File 2) [73]. Importantly, maternal folate levels have been shown to influence human birthweight [74], and a role for insulin regulation in bovine feed efficiency and growth traits has also been described [30]. Beyond BTA2, BW GxE QTL were also detected on BTA17 (17_66 Mb) and BTA13 (13_67 Mb). Positional candidate genes for these QTL have been implicated in the removal of uracil residues from DNA and apoptosis (UNG) as well as human obesity (CTNNBL1) (Fig. 4, Table 5, Figure S4, Table S6, Additional File 1) [75, 76]. Examination of the lead SNPs for all GxE QTL detected for Gelbvieh BW (Table 5, Table S6, Additional File 1, Additional File 2) revealed three noncoding variants, which is suggestive of quantitative (i.e., regulatory) effects. Genomic inflation factors and correlation coefficients for P-values obtained from all GxE BW analyses are shown in Tables S2-S3 (Additional File 1).

Our analyses (GEMMA, EMMAX) to evaluate the potential for significant GxE interactions with respect to WW in U.S. Gelbvieh beef cattle produced evidence for one GxE QTL on BTA2 (2_18 Mb) which was only detected by GEMMA, and included relatively few supporting SNPs (P ≤ 1e-05, Table 6; Fig. 5, Figure S5, Additional File 1). The lead SNP defining this QTL was located in exon 304 of TTN, and encoded a nonsynonymous variant (Table 6, Fig. 5, Table S7, Figure S6, Additional File 1, Additional File 2). The

**Table 2** Summary of QTL detected by GEMMA for BW in U.S. Gelbvieh beef cattle

| Chr_Mb | MAF  | -log10 P-value | Supporting SNPs | Positional Candidate Genes | Lead SNP Location | Scientific Precedence [reference; organism; trait] |
|--------|------|----------------|----------------|-----------------------------|------------------|--------------------------------------------------|
| 14_25a | 0.398 | 29.56          | 41             | PLAG1                       | 3’UTR            | [2, 14, 18, 21, 30, 32–34]; Cattle; SimAngus mid-test metabolic weight association, carcass weight, stature, body weight and milk |
| 6_39a  | 0.293 | 23.71          | 140            | NCPG                        | Exon             | [18, 21, 30, 35–39]; Cattle, chicken; stature, calming ease and growth traits association, SimAngus mid-test metabolic weight association, fetal growth, carcass trait association, average daily gain and daily feed intake, muscle mass |
| 14_26a | 0.396 | 14.63          | 33             | IMPAD1, FAM110B             | Intergenic       | [30, 32, 34, 40]; Cattle; SimAngus mid-test metabolic weight association, carcass weight association, stature and body weight association, bone and cartilage system |
| 6_42a  | 0.186 | 9.66           | 9              | KCNP4                       | Intron           | [39, 41, 42]; Chicken, cattle, human; growth and muscle mass trait association, potassium channel activity |
| 14_24a | 0.244 | 8.93           | 35             | XKR4                        | Intron           | [2, 30, 43, 44]; Cattle; birth weight association, SimAngus mid-test metabolic weight association, growth trait association, feed intake and growth traits |
| 20_05a | 0.193 | 8.65           | 21             | LOC104975192, STC2          | Intergenic       | [30, 45]; Cattle, mouse; mid-test metabolic weight in Hereford and SimAngus, development and adult tissue maintenance, body size, related to postnatal growth |
| 7_93a  | 0.283 | 8.00           | 30             | ARRD3C3, LOC104972872      | Intergenic       | [14, 22, 30, 46]; Cattle; body and carcass weight association, calming ease, average daily gain in Hereford, growth and muscularity, birth weight, weaning weight, yearling weight, and ribeye area in Angus |
| 6_38a  | 0.053 | 7.90           | 23             | IBSP, LOC104972276         | Intergenic       | [13, 47–49]; Cattle, mouse, human; yearling weight association, bone formation and remodeling, cellular proliferation, milk-fat and protein association |
| 6_41a  | 0.407 | 7.25           | 5              | LOC782905, SLT2             | Intergenic       | [39, 49–53]; Cattle, chicken, human; milk fat and protein association, organ and muscle weight, development of central nervous system, tumor suppressor activity |
| 14_23a | 0.467 | 6.19           | 3              | ST18                        | Intron           | [54]; Human; regulation of apoptosis and inflammatory response |
| 6_34a  | 0.039 | 5.98           | 8              | LOC104972717, LOC326089     | Intergenic       | NA |
| 6_40a  | 0.304 | 5.25           | 2              | LCCRL, LOC782905            | Intergenic       | [18, 21, 37–39, 50, 55, 56]; Cattle, sheep; stature, muscle and organ growth, feed intake and gain association, growth and carcass traits, skeletal growth and muscle mass |

*a* Indicates QTL was detected in EMMAX analysis

*b* Indicates a predicted nonsynonymous mutation Ile→Met, exon 9

---

Our analyses (GEMMA, EMMAX) to evaluate the potential for significant GxE interactions with respect to YW in U.S. Gelbvieh beef cattle produced evidence for one GxE QTL on BTA2 (2_18 Mb) which was only detected by GEMMA, and included relatively few supporting SNPs (P ≤ 1e-05, Table 6; Fig. 5, Figure S5, Additional File 1). The lead SNP defining this QTL was located in exon 304 of TTN, and encoded a nonsynonymous variant (Table 6, Fig. 5, Table S7, Figure S6, Additional File 1, Additional File 2). The
signal on BTA17 (i.e., GEMMA lead SNP in Intron 4 of LOC101904475 and supporting SNPs) was replicated by EMMAX (Figure S6, Additional File 1); but at a less stringent significance threshold (i.e. $P < 6 \times 10^{-4}$). Notably, while the function of LOC101904475 remains unclear, LRAT is known to catalyze esterification of retinol (i.e., from Vitamin A) [80], and Vitamin A has been shown to promote growth in beef cattle as well as humans [81–83]. However, FGG is also an intriguing candidate, as fibrinogen has been shown to constrict blood vessels [84]. This vasoconstriction may alter the ability to cope with heat stress, but in the context of cattle production, the relationship between vasoconstriction and fescue toxicosis is perhaps more noteworthy. Fescue toxicosis is the result of ergot alkaloids produced by the endophytic fungus in fescue forage [85], especially the Kentucky 31 variety. One of the major symptoms of fescue toxicosis is vasoconstriction, thus variation in FGG expression levels may potentially alter cattle’s innate degree of vasoconstriction; perhaps further complicating both fescue toxicosis and heat

| Chr_Mb | MAF | $-\log_{10} P$-value | Supporting SNPs | Positional Candidate Genes | Lead SNP Location | Scientific Precedence [reference]; organism; trait |
|--------|-----|----------------------|----------------|-----------------------------|------------------|-------------------------------------------------|
| 6_39\textsuperscript{a} | 0.289 | 18.32 | 107 | NCAPG | Exon\textsuperscript{b} | [18, 21, 30, 35–39]; cattle, chicken; stature, calving ease and growth traits association, SimAngus mid-test metabolic weight association, fetal growth, carcass trait association, average daily gain and daily feed intake, muscle mass |
| 14_25\textsuperscript{a} | 0.398 | 10.69 | 2 | PLAG1 | 3'UTR | [2, 14, 18, 21, 30, 32–34]; cattle; SimAngus mid-test metabolic weight association, carcass weight, stature, body weight and milk |
| 5_60\textsuperscript{a} | 0.046 | 8.83 | 2 | LOC527216, LOC788998 | Intergenic | NA |
| 6_38\textsuperscript{a} | 0.214 | 7.95 | 29 | CCSE1 | Intron | [14, 60]; cattle, human; body and carcass weight association, regulator of mitosis |
| 14_26\textsuperscript{a} | 0.415 | 7.90 | 11 | IMPAD1, FAM1108 | Intergenic | [30, 32, 34, 40]; cattle; SimAngus mid-test metabolic weight association, carcass weight association, stature and body weight association, bone and cartilage system |
| 6_42\textsuperscript{a} | 0.340 | 7.77 | 3 | KCNIP4 | Intron | [39, 41, 42]; chicken, cattle, human; growth and muscle mass trait association, potassium channel activity |
| 6_38\textsuperscript{a} | 0.220 | 7.70 | 9 | HERC6, PPM1K | Intergenic | [49, 58, 59]; cattle; milk, fat, and protein yield, metabolic processes, feed efficiency association |
| 6_41\textsuperscript{a} | 0.238 | 6.46 | 4 | LOC782905, SLT2 | Intergenic | [39, 49–53]; cattle, chicken, human; milk fat and protein association, organ and muscle weight, development of central nervous system, tumor suppressor activity |
| 6_37\textsuperscript{a} | 0.325 | 5.97 | 5 | SNCA, GPRIN3 | Intergenic | [61–64]; human, goat, equine; neurological regulation, milk and meat associations, tendon tissue association |
| 6_34\textsuperscript{a} | 0.295 | 5.36 | 4 | LOC100336621, LOC104972717 | Intergenic | NA |

\textsuperscript{a} Indicates QTL was detected in EMMAX analysis

\textsuperscript{b} Indicates a predicted nonsynonymous mutation Ile$\rightarrow$Met, exon 9

Table 3: Summary of QTL detected by GEMMA for WW in U.S. Gelbvieh beef cattle

Fig. 2 Weaning weight (WW) QTL Manhattan plot with GEMMA $-\log_{10} P$-values. Lead and supporting SNPs for QTL represented at or above the blue line ($P \leq 1 \times 10^{-5}; -\log_{10} P$-values $\geq 5.00$) for $n = 10,837$ U.S. Gelbvieh beef cattle. A summary of all markers passing the nominal significance threshold [31] is presented in Table 3.
Table 4 Summary of QTL detected by GEMMA for YW in U.S. Gelbvieh beef cattle

| Chr_Mb | MAF  | -log10 P-value | Supporting SNPs | Positional Candidate Genes | Lead SNP Location | Scientific Precedence [reference]; organism; trait |
|--------|------|----------------|-----------------|--------------------------|-----------------|---------------------------------------------------|
| 6_39a  | 0.305| 20.81          | 103             | LCORL                    | Intron          | [18, 21, 30, 37–39, 55, 56]; Cattle, sheep; stature, SimAngus mid-test metabolic weight association, muscle and organ growth, feed intake and gain association, growth and carcass traits, skeletal growth and muscle mass |
| 14_25a | 0.399| 13.82          | 3               | PLAG1                    | 3'UTR           | [2, 14, 18, 21, 30, 32–34]; Cattle; SimAngus mid-test metabolic weight association, carcass weight, stature, body weight and milk |
| 6_38a  | 0.222| 11.00          | 20              | HERC6, PPM1K             | Intergenic      | [49, 58, 59]; Cattle; milk, fat, and protein yield, metabolic processes, feed efficiency association |
| 6_42a  | 0.344| 11.00          | 11              | KCNIP4                   | Intron          | [39, 41, 42]; Chicken, cattle, human; growth and muscle mass trait association, potassium channel activity |
| 6_37a  | 0.330| 10.12          | 8               | SNCA, GPRIN3             | Intergenic      | [61–64]; Human, goat, equine; neurological regulation, milk and meat associations, tendon tissue association |
| 5_60a  | 0.042| 9.62           | 2               | LOC527216, LOC78899B     | Intergenic      | NA |
| 6_41a  | 0.247| 8.44           | 6               | SLIT2                    | Intron          | [39, 49–53]; Cattle, chicken, human; milk fat and protein association, organ and muscle weight, development of central nervous system, tumor suppressor activity |
| 6_36a  | 0.227| 8.23           | 20              | CCSER1                   | Intron          | [14, 60]; Cattle, human; body and carcass weight association, regulator of mitosis |
| 14_26a | 0.357| 6.94           | 12              | IMPAD1, FAM110B          | Intergenic      | [30, 32, 34, 40]; Cattle; SimAngus mid-test metabolic weight association, carcass weight association, stature and body weight association, bone and cartilage system |
| 7_93a  | 0.286| 6.23           | 14              | LOC101905238, ARRDCC3    | Intergenic      | [14, 22, 30, 46]; Cattle; body and carcass weight association, calving ease, average daily gain in Hereford, growth and muscularity, birth weight, weaning weight, yearling weight, and ribeye area in Angus |
| 6_40a  | 0.109| 6.21           | 11              | LOC782995, SLIT2         | Intergenic      | [39, 49–53]; Cattle, chicken, human; milk fat and protein association, organ and muscle weight, development of central nervous system, tumor suppressor activity |
| 14_27a | 0.348| 6.04           | 6               | NSMAF                    | Intron          | [30, 68]; Cattle, human; Angus residual feed intake association, immune system response |
| 2_05   | 0.497| 5.15           | 3               | IWS1                     | Intron          | [69]; Human; chromatin modification, histone chaperone, maintenance of virus latency |

* Indicates QTL was detected in EMMAX analysis

**Fig. 3** Yearling weight (YW) QTL Manhattan plot with GEMMA -log10 P-values. Lead and supporting SNPs for QTL represented at or above the blue line (P ≤ 1e-05; -log10 P-values ≥ 5.00) for n = 10,837 U.S. Gelbvieh beef cattle. A summary of all markers passing the nominal significance threshold [31] is presented in Table 4.
stress. The other interesting positional candidate gene on BTA5 (PHF21B) is known to be involved in the modulation of stress responses, and the regulation of cellular division [86, 87].

Conclusions

Herein, we present evidence for pleiotropic QTL influencing BW, WW, and YW in U.S. Gelbvieh beef cattle, and further confirm the involvement of PLAG1 in various aspects of bovine growth and stature across breeds [2, 14, 18, 21, 30, 32–34]. Additionally, we also present compelling evidence for QTL segregating in multiple breeds; with at least seven U.S. Gelbvieh growth QTL that were also detected for feed efficiency and growth traits in U.S. Angus, SimAngus, and Hereford beef cattle [30]. Despite the involvement of major genes such as NCAPG, PLAG1 and LCORL, more of the phenotypic variance in Gelbvieh BW, WW, and YW was explained by many other genome-wide loci (See Additional File 1, Additional File 2). Moreover, we demonstrate that most of the Gelbvieh QTL are detectable by two different large-sample analyses (GEMMA; EMMAX). However, some discordant QTL detected by the GxE GWAA can also be attributed to differences in the model specifications for these analyses, as implemented by GEMMA and EMMAX (See Methods). While relatively few GxE QTL were detected, the identified GxE QTL harbor physiologically meaningful positional candidates. Moreover, the results of this study demonstrate that imputation to a union set of high-density SNPs (i.e., 856K) for use in large-sample analyses can be expected to facilitate future discoveries at a fraction of the cost associated with direct genotyping, which also underscores the present impact of genomic tools and resources developed by the domestic cattle research community.

Methods

Cattle phenotypes were received from the American Gelbvieh Association (pre-adjusted for age of animal [i.e. 205-day weight for WW] and age of dam as per breed association practice), and corresponding genotypes were transferred from their service provider Neogen GeneSeek. For GWAA analyses, the phenotypes were pre-adjusted for sex and contemporary group consisting of 5-digit breeder zip-code, birth year, and birth season (Spring, Summer, Fall, and Winter) using the mixed.solve() function from the rrBLUP package v4.4 [88] in R v3.3.3 [89]. To group individuals into discrete climate zones, K-means clustering was performed on three continuous climate variables. Thirty-year normal values for temperature, precipitation, and elevation were drawn from the PRISM climate dataset [90]. Each one km square of the continental United States was assigned to one of nine climate zones using K-means clustering implemented in the RStoolbox R package [91, 92]. The optimal number of zones was identified using the pamk function from the R package fpc [93]. Individuals were assigned to zones based

| Table 5 Summary of GxE QTL detected by GEMMA for BW in U.S. Gelbvieh beef cattle |
|-----------------|-----------------|-----------------|-----------------|
| Chr_Mb | MAF | -log10 P-value | Supporting SNPs | Positional Candidate Genes | Lead SNP Position | Scientific Precedence [reference]; organism; trait |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 2_32 | 0.105 | 6.25 | 2 | GRB14, FIGN | Intergenic | [72–74]; Mouse; human; insulin receptor related to growth and metabolism, folic acid association with impact on BW |
| 17_66 | 0.026 | 6.21 | 2 | UNG | Intron | [75]; Human; DNA maintenance |

Fig. 4 Birth weight genotype-by-environment (BW GxE) QTL. Manhattan plot with GEMMA -log10 P-values. Lead and supporting SNPs for QTL represented at or above the blue line (P ≤ 1e-05; −log10 P-values ≥ 5.00) for n = 10,837 U.S. Gelbvieh beef cattle. A summary of all markers passing the nominal significance threshold [31] is presented in Table 5.
Quality control was performed on genotypes for 13,166 Gelbvieh individuals using PLINK 1.9 [94]. Individuals with call rates < 0.90 were removed on an assay-by-assay basis (For assay information see Additional File 3). Variants with call rates < 0.90 or Hardy-Weinberg Equilibrium (HWE) \( P \)-values < 1e-20 were also removed. For this analysis, only autosomal chromosomes were analyzed. After filtering, genotypes for the 12,422 individuals that remained were merged using PLINK and then phased using EagleV2.4 [95]. Imputation was performed with IMPUTE2 [97] using the “merge_ref_panels” flag. This allowed the phased haplotypes for 315 individuals genotyped on the Illumina HD (Illumina, San Diego, CA) and 559 individuals genotyped on the GGP-F250 (GeneSeek, Lincoln, NE) to be recursively imputed and treated as reference haplotypes. These reference haplotypes were used to impute the remaining 11,598 low-density genotypes from various assays (Additional File 3) to the shared number of markers between the two high-density research chips. The resulting dataset consisted of 12,422 individuals with 856,527 markers each (UMD3.1). To account for uncertainty in imputation, IMPUTE2 reports dosage genotypes. Hard-called genotypes were inferred from dosages using PLINK. When making hard-calls, PLINK treats genotypes with uncertainty > 0.1 as missing. This resulted in a hard-called dataset of 856,527 variants, which includes genotypes set as missing. Prior to the execution of all GWAAAs (GEMMA; EMMAX), we filtered the Gelbvieh samples and all SNP loci as follows: Gelbvieh sample call rate filtering (<90% call rate excluded); thereafter SNP filtering by call rate (>15% missing excluded), MAF (<0.01 excluded), polymorphism (monomorphic SNPs excluded), and HWE (excludes SNPs with HWE \( P \)-values < 1e-50), which resulted in 618,735 SNPs. Additionally, prior to all GWAAAs (GEMMA; EMMAX) hard-called genotypes were numerically recoded as 0, 1, or 2, based on the incidence of the minor allele. Missing hard-called genotypes (i.e., that met our filtering criteria) were modeled as the SNP’s average value (0, 1, or 2) across all samples.

Using the numerically recoded hard-called genotypes and the adjusted Gelbvieh phenotypes, we employed GEMMA to conduct univariate linear mixed model GWAAAs where the general mixed model can be specified as

\[
y = W\alpha + x\beta + u + \epsilon;
\]

where \( y \) represents a \( n \)-vector of quantitative traits for \( n \)-individuals, \( W \) is an \( n \times c \) matrix of specified covariates (fixed effects) including a column of 1s, \( \alpha \) is a \( c \)-vector of the corresponding coefficients including the intercept, \( x \) represents an \( n \)-vector of SNP genotypes, \( \beta \) represents the effect size of the SNP, \( u \) is an \( n \)-vector of random effects, and \( \epsilon \) represents an \( n \)-vector of errors [27]. Moreover, it should also be noted that \( u \sim \]

---

**Table 6** Summary of GxE QTL detected by GEMMA for WW in U.S. Gelbvieh beef cattle

| Chr_Mb | MAF  | -log_{10} P-value | Supporting SNPs | Positional Candidate Genes | Lead SNP Location | Scientific Precedence [reference]; organism; trait |
|--------|------|-------------------|-----------------|-----------------------------|-------------------|-------------------------------------------------|
| 2_18   | 0.012| 5.22              | 2               | TTN                        | Exon*             | [77–79]; Rabbit, rat, human; aids in myofibrillar assembly, positioning of myosin filaments in muscle, coordinates multiple signaling pathways for gene activation, protein folding, quality control and degradation, heart disease relation |

* Indicates a predicted nonsynonymous mutation Arg➔Gln, exon 304

---

**Fig. 5** Weaning weight genotype-by-environment (WW GxE) QTL. Manhattan plot with GEMMA -log_{10} P-values. Lead and supporting SNPs for QTL represented at or above the blue line (\( P \leq 1e-05; -\log_{10} P \)-values \( \geq 5.00 \)) for \( n = 10,837 \) U.S. Gelbvieh beef cattle. A summary of all markers passing the nominal significance threshold [31] is presented in Table 6.
$MVN_n(0, \lambda r^{-1}K)$ and $e \sim MVN_n(0, \lambda r^{-1}I_n)$, where $MVN$ denotes multivariate normal distribution, $\lambda r^{-1}$ is the variance of the residual errors, $\lambda$ is the ratio between the two variance components, $K$ is a known $n \times n$ relatedness matrix, and $I_n$ represents an $n \times n$ identity matrix [27]. Using this general approach, GEMMA evaluated the alternative hypothesis for each SNP ($H_1: \beta \neq 0$) as compared to the null ($H_0: \beta = 0$) by performing a likelihood ratio test with maximum likelihood estimates ($-\text{lmm 2}$) as follows:

$$D_{\text{lr}} = 2 \log \left( \frac{\hat{\lambda}_1}{\hat{\lambda}_0} \right),$$

where $\hat{\lambda}_0$ and $\hat{\lambda}_1$ represent the maximum likelihood estimates for the null and the alternative models, respectively, where P-values come from a $\chi^2$ distribution, as previously described [27]. Herein, the only fixed-effect covariate specified for all GWAAAs was an environmental variable (geographic zone for each individual). For all GxE GWAAAs (-gxe command), the environmental variable (geographic zone for each individual) was treated as an interaction term, where the resulting P-values represent the significance of the genotype x environment interaction. Specifically, for the GxE GWAAAs in GEMMA, the model is specified as $y = Wa + x_{\text{snp}}\beta_{\text{snp}} + x_{\text{env}}\beta_{\text{env}} + x_{\text{snp}} \times x_{\text{env}}\beta_{\text{snp} \times \text{env}} + u + e$;

where $y$ represents a $n$-vector of quantitative traits for $n$-individuals, $W$ is an $n \times c$ matrix of specified covariates (fixed effects) including a column of 1s, $\alpha$ is a $c$-vector of the corresponding coefficients including the intercept, $x_{\text{snp}}$ represents an $n$-vector of SNP genotypes, $\beta_{\text{snp}}$ represents the effect size of the SNP, $x_{\text{env}}$ represents an $n$-vector of environmental covariates, $\beta_{\text{env}}$, represents the fixed effect of the environment, $\beta_{\text{snp} \times \text{env}}$, is the interaction between SNP genotype and environment, $u$ is an $n$-vector of random effects, and $e$ represents an $n$-vector of errors.

GEMMA evaluated the alternative hypothesis for each interaction ($H_1: \beta_{\text{snp} \times \text{env}} \neq 0$) as compared to the null ($H_0: \beta_{\text{snp} \times \text{env}} = 0$). Marker-based relatedness matrices ($G_i$) relating instances of the random effect specified to each of the growth phenotypes among all genotyped cattle were used to estimate the proportion of variance explained (PVE) by the hard-called genotypes in GEMMA, which is also commonly referred to as the “chip heritability” [27, 98]. For all investigated traits, single-marker P-values obtained from GEMMA ($-\text{lmm 2}$, -gxe) were used to generate Manhattan plots in R (manhattan command) and QTL were defined by $\geq 2$ SNP loci with MAF $\geq 0.01$ (i.e., a lead SNP plus at least one additional supporting SNP within 1 Mb) which also met a nominal significance threshold ($P \leq 1e-05$) [30, 31].

Using hard-called genotypes and the adjusted Gelbvieh phenotypes, we performed a second set of GWAAAs using a mixed linear model with variance component estimates, as

![Fig. 6 Yearling weight genotype-by-environment (YW GxE) QTL. Manhattan plot with GEMMA -log10 P-values. Lead and supporting SNPs for QTL represented at or above the blue line ($P \leq 1e-05$; -log10 P-values $\geq 5.00$) for $n = 10,837$ U.S. Gelbvieh beef cattle. A summary of all markers passing the nominal significance threshold [31] is presented in Table 7](image-url)
implemented by EMMAX [28–30, 99–101]. Briefly, the general mixed model used in this approach can be specified as: y = Xβ + Zu + ε, where y represents a n × 1 vector of phenotypes, X is a n × q matrix of fixed effects, β is a q × 1 vector representing the coefficients of fixed effects, and Z is a n × t matrix relating the random effect to the phenotypes of interest [30, 99–101]. Herein, we must assume that Var(u) = σ²uK and Var(ε) = σ²εI, such that Var(y) = σ²ZKZ’ + σ²εI, however, in this study Z represents the identity matrix I, and K represents a kinship matrix of all Gelbvieh samples with hard-called genotypes. Moreover, to solve the mixed model equations using a generalized least squares approach, we must estimate the variance components (σ²u and σ²ε) as previously described [28–30, 99, 100]. For this study, we estimated the variance components using the REML-based EMMA approach [29], with stratification accounted for and controlled using the genomic relationship matrix [25, 30], as computed from the Gelbvieh hard-called genotypes. Moreover, the only fixed-effect covariate specified for all GWAAAs was an environmental variable (geographic zone for each individual). For all EMMAX GxE GWAAAs utilizing hard-called genotypes, we used an implementation of EMMAX [29, 102] where interaction-term covariates may be specified; with the environmental variable (geographic zone for each individual) specified as the interaction term. The basis of this approach is rooted in full versus reduced model regression [99], where interaction-term covariates are included in the model as follows: each specified interaction-term covariate or predictor variable that may be interacting, and ϵ reduced as previously described [28–30, 99, 100]. Like the EMMAX method without interactions [28, 29], we approximate this by finding the variance components one using the parts of the finding that are independent of Xk as follows: y = XkβX + Xkβuc + uc + ϵ uc, where uc indicates the variance components. To estimate the variance components, we must again assume that Var(uc) = σ²uK and Var(εc) = σ²εI, such that Var(y) = σ²ZKZ’ + σ²εI. The EMMA technique can then be used to estimate the variance components σ²u and σ²ε as well as a matrix B (and its inverse) such that BB’ = H = σ²ZKZ’ + σ²εI. Thereafter, for every marker (k) we can compute (as an EMMAX-type approximation) the full and reduced models as: B’−1y = B’−1XkβX + B’−1Xkβuc + B’−1Xkβuc + B’−1Xkuc + B’−1uc and B’−1uc is assumed to be an error term proportional to the identity matrix, and as B’−1XkβX + B’−1Xkβuc + B’−1Xkβuc + B’−1uc reduced + ϵ reduced is assumed to be an error term proportional to the identity matrix. To estimate the significance of the full versus reduced model, an F-test was performed; with all analyses utilizing the EMMAX method [28, 29] (i.e., GWAAAs, GxE GWAAAs) produced and further evaluated by constructing Manhattan plots within SSV v8.8.2 (Golden Helix, Bozeman, MT). Moreover, while SSV explicitly computes the full model mentioned above and outputs all of its β values, it only performs an optimization of the reduced model computation, which is sufficient to determine the SSE of the reduced-model equation, and thereafter, estimate the full versus reduced model P-value via F-test. This optimization is used to solve: MB’−1y = MB’−1XkβX + ϵ MB where M = (I – QQ’), and Q is derived from performing the QR algorithm, as Q = B’−1[Xk1; Xk]. All Gelbvieh QTL were defined by ≥ 2 SNP loci with MAF ≥ 0.01 (i.e., a lead SNP plus at least one additional supporting SNP within 1 Mb) which also met a nominal significance threshold (P ≤ 1e-05) [30, 31], and all EMMAX marker-based pseudo-heritability estimates were produced as previously described [28–30, 99, 100].

Genomic inflation factors (λ) for all analyses (EMMA; EMMAX) were estimated from the observed F-values using genABEL [103], and the relationships between the observed P-values were estimated (GENMAA versus EMMAX) via correlation coefficients (i.e., Pearson, Spearman) in R v3.3.3 [89].

Supplementary information
Supplementary information accompanies this paper at https://doi.org/10.1186/s12864-019-6231-y.
Table S3. Correlation coefficients for GEMMA versus EMMAX P-values.

Table S4. Summary of QTL detected by EMMAX for WW in U.S. Gelbvieh cattle. Table S5. Summary of QTL detected by EMMAX for YW in U.S. Gelbvieh cattle. Table S6. Summary of GxE QTL detected by EMMAX for BW in U.S. Gelbvieh cattle. Table S7. Summary of GxE QTL detected by EMMAX for YW in U.S. Gelbvieh cattle.

Additional file 2. Summary of lead and supporting SNPs from analyses for BW, YY, BW GxW, WW GxW, and YW GxW using 778K imputed genotypes, including QC chromosome, base pair, and rounded Mb.

Additional file 3. Summary of SNP panels used in analyses, including number of SNPs and individuals available before and after filtering.

Abbreviations

BW: Birth Weight; GIWAA: Genome-wide association analysis; GxE: Genotype-by-environment interaction; QTL: Quantitative Trait Locus; WW: Weaning Weight; YY: Yearling Weight

Acknowledgements

We sincerely appreciate the American Gelbvieh Association and Gelbvieh breeders for sharing the genotypes and phenotypes used in this research.

Authors’ contributions

JED and CMS conceptualized and designed the research, RDS managed data collection, storage, and retrieval. TNR performed initial quality control, phasing and imputation. SMN pre-adjusted phenotypes. DLO provided programming support for data formatting and management. JLS, MLW, and CMS performed GIWAA. JLS, JED, and CMS interpreted results. JLS, JED, and CMS wrote initial version of the manuscript, which was edited by all authors.

Funding

This project was supported by Agriculture and Food Research Initiative Competitive Grant no. 2016-68004-24827 from the USDA National Institute of Food and Agriculture.

Availability of data and materials

Data are available for non-commercial use via data use agreement (DUA) with the American Gelbvieh Association.

Ethics approval and consent to participate

Prior to the planning and onset of the current study, data were collected under typical production practices and reported to the American Gelbvieh Association, for inclusion in an existing data repository; thus an animal care and use protocol is not necessary. Staff at the American Gelbvieh Association have read and approved the manuscript.

Consent for publication

Staff at the American Gelbvieh Association have provided the authors with consent to publish.

Competing interests

The authors declare no competing interests.

Author details

1Department of Veterinary Pathobiology, Texas A&M University, College Station 77843, USA. 2Division of Animal Sciences, University of Missouri, Columbia 65211, USA. 3Genetics Area Program, University of Missouri, Columbia 65211, USA. 4Informatics Institute, University of Missouri, Columbia 65211, USA.

Received: 6 May 2019 Accepted: 28 October 2019 Published online: 04 December 2019

References

1. Barwick SA, Henzell AL. Development successes and issues for the future in deriving and applying selection indexes for beef breeding. Aust J Exp Agric. 2005;45:923. https://doi.org/10.1071/EA05068.
2. Utsunomiya YT, do Carmo AS, Canhavheiro R, Neves HH, Matos MC, Zavarez LB, et al. Genome-wide association study for birth weight in Nellore cattle points to previously described orthologous genes affecting human and bovine height. BMC Genet. 2013;14:52. https://doi.org/10.1186/1471-2164-14-52.
3. Brito Lopes F, da Silva MC, Magnabosco CU, Goncalves Narciso M, Sainz RD. Selection indices and multivariate analysis show similar results in the evaluation of growth and carcass traits in beef cattle. PLoS One. 2016;11: e0147180. https://doi.org/10.1371/journal.pone.0147180.
4. Meirelles SLC, Mokry FB, Espasandin AC, Dias MAD, Baena MM, de A. Regitano LC. Genetic parameters for carcass traits and body weight using a Bayesian approach in the Canchim cattle. Genet Mol Res. 2016;15. https://doi.org/10.4238/gmr.15027471.
5. Bourdon RM, Brinks JS. Genetic, environmental and phenotypic associations among gestation length, birth weight, growth traits and age at first calving in beef cattle. J Anim Sci. 1982;55:543–55. https://doi.org/10.2527/1982.5553543x.
6. Erikkson S, Nåsholm A, Johansson K, Philipsson J. Genetic parameters for calving difficulty, stillbirth, and birth weight for Hereford and Charolais at first and later parturitions. J Anim Sci. 2004;82:375–83. https://doi.org/10.2527/2004-822375x.
7. Cook BR, Tess MW, Kress DD. Effects of selection strategies using heifer pelvic area and sire birth weight expected progeny difference on dystocia in first-calf heifers. J Anim Sci. 1993;71:602–7. https://doi.org/10.2527/1993.713602x.
8. Spelman RJ, Huisman AE, Singireddy SR, Coppieters RJ, Arranz J, Georges M, et al. Quantitative trait loci analysis on 17 nonproduction traits in the New Zealand dairy population. J Dairy Sci. 2000;83:370. https://doi.org/10.3168/jds.S0022-0302(00)74490-9.
9. Kneeland J, Li C, Basarab J, Snelling WM, Benkel B, Murdoch B, et al. Identification and fine mapping of quantitative trait loci for growth traits on bovine chromosomes 2, 6, 14, 19, 21, and 23 within one commercial line of Bos taurus. J Anim Sci. 2004;82:3405–14. https://doi.org/10.2527/2004-82123405x.
10. Malteca C, Weigel KA, Khatib H, Cowan M, Bagnato A. Whole-genome scan for quantitative trait loci associated with birth weight, gestation length and passive immune transfer in a Holstein · Jersey crossbred population. Anim Genet. 2008;27–34. https://doi.org/10.1111/j.1365-2052.2008.01793.x.
11. McClure MC, Morici NS, Schnabel RD, Kim JW, Yao P, Rolf MM, et al. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. Anim Genet. 2010;41:597–607. https://doi.org/10.1111/j.1365-2052.2010.02063.x.
12. Cole JB, Wiggans GR, Ma L, Sonstegard TS, Crooker BA, et al. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. BMC Genomics. 2011;12:408. https://doi.org/10.1186/1471-2164-12-408.
13. Santiago GI, Siqueira F, Cardoso FF, Regitano LCA, Ventura R, Sollero BP, et al. Genomewide association study for production and meat quality traits in Canchim beef cattle. J Anim Sci. 2017;95:S891–90. https://doi.org/10.2527/jas.2017.1570.
14. Saatchi M, Schnabel RD, Taylor JF, Garrick DJ. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. BMC Genomics. 2014;15.
15. Saatchi M, McClure MC, McKay SD, Rolf MM, Kim J, Decker JE, et al. Accuracies of genomic breeding values in American Angus beef cattle using K-means clustering for cross-validation. Genet Sel Evol. 2011;43.
16. Saatchi M, Schnabel RD, Rolf MW, Taylor JF, Garrick DJ. Accuracy of direct genomic breeding values for nationally evaluated traits in US Limousin and Simmental beef cattle. Genet Sel Evol. 2012;44.
17. Saatchi M, Garrick DJ, Tait RG, Mayes MS, Drewnoski M, Schoonmaker J, et al. Genome-wide association and prediction of direct genomic breeding values for composition of fatty acids in Angus beef cattle a. BMC Genomics. 2013;14.
18. Pryce JE, Hayes BJ, Bolormaa S, Goddard ME. Polygenic regions affecting human height also control stature in cattle. Genetics. 2011;187:981–4.
19. Guddbjartsson DF, Walters GB, Thorleifsson G, Stefansson H, Halldorsson BV, Zusmanovich P, et al. Many sequence variants affecting diversity of adult age at first calving in beef cattle. J Anim Sci. 1982;55:543–55. https://doi.org/10.2527/1982.5553543x.
20. Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, et al. Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet. 2008;40:609–15.
21. Pausch H, Flisikowski K, Jung S, Emmerling R, Edel C, Götz KU, et al. Genome-wide association study identifies two major loci affecting calving ease and growth-related traits in cattle. Genetics. 2011;187:289–97.
22. Abo-Ismail MK, Brito LF, Miller SP, Sargolzaei M, Grossi DA, Moore SS, et al. Genome-wide association studies and genomic prediction of breeding
values for calving performance and body conformation traits in Holstein cattle. Genet Sel Evol. 2017;49.

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

24. Rincon G, Weber KL, Van Eenennaam AL, Golden BL, Medrano JF. Hot topic: performance of bovine high-density genotyping platforms in Holsteins and jerseys. J Dairy Sci. 2011;94:6116–21. https://doi.org/10.3168/jds.2011-4764.

25. VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci. 2008;91:441–42. https://doi.org/10.3168/jds.2007-0980.

26. Decker JE. Agricultural genomics: commercial applications bring increased basic research power. PLoS Genet. 2015;11:e1005621. https://doi.org/10.1371/journal.pgen.1005621.

27. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2014;46:231–8. https://doi.org/10.1038/ng.3214.

28. Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, et al. An advanced intercross line. Anim Genet. 2017;48:295–305. https://doi.org/10.1111/age.12850.

29. Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S, Freimer NB, et al. Variance component model to account for sample structure in genome-wide association studies in structured populations. Nat Genet. 2012;44:238–50. https://doi.org/10.1038/ng.3214.

30. Seahury CM, Oldeschulte DL, Saatchi M, Breever JE, Decker JE, HallyYA, et al. Genomic association study for feed efficiency and growth traits in U.S. beef cattle. BMC Genomics. 2017;18:386. https://doi.org/10.1186/s12864-017-1374-y.

31. Wellcome Trust Case Control Consortium TWTC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007;447:661–78. https://doi.org/10.1038/nature05911.

32. Lee SH, Choi BH, Lim D, Gondo C, Cho YM, Dang CG, et al. Genome-wide association study identifies major loci for carcass weight on BTA14 in Hanwoo (Korean cattle). PLoS One. 2013;8.

33. Utsunomiya YI, Milanese M, Utsunomiya ATH, Torrecilla RBP, Kim ES, Costa MS, et al. A PLAG1 mutation contributed to stature recovery in modern cattle. Sci Rep. 2017;7.

34. Fink T, Tippldy K, Lopdell T, Johnson T, Snell RG, Spelman RJ, et al. Functional confirmation of PLAG1 as the candidate causative gene underlying major pleiotropic effects on body weight and milk characteristics. Sci Rep. 2017;7.

35. Eberlein A, Takasuga A, Setoguchi K, Pfuhl R, Flisikowski K, Fries R, et al. Segregation of an orthologous region on OAR6 to human and bovine genomic regions affecting height and weight. Genet Sel Evol. 2015;47.

36. Lindholm-Perry AK, Kuehn LA, Oliver WT, Sexten AK, Miles JR, Rempel SM, et al. Development and characterization of a high density SNP genotyping assay for cattle. PLoS One. 2009;4:e9160. https://doi.org/10.1371/journal.pone.0007984.

37. Martínez R, Gómez Y, Martínez-Roch JF. Genome-wide association study on longissimus muscle development and associated with growth traits in Equus caballus. Genet Mol Res. 2014;13:6420–32.

38. Liu Y, Duan X, Chen S, He H, Liu X. NCAPG is differentially expressed during allergic airway inflammation. Sci Rep. 2017;7.

39. Al-Mamun HA, Kwan P, Clark SA, Ferdosi MH, Tellam R, Gondro C. Genome-wide association study identifies loci and candidate genes for internal organ weights in simental beef cattle. Physiol Genomics. 2018;50:23–31.

40. Saunders J, Wisidagama DR, Morford T, Malone CS. Maximal expression of the evolutionarily conserved Slt2 promoter requires Sp1. Cell Mol Neurobiol. 2016;36:955–64.

41. Holmes G, Niwander L. Expression of slit-2 and slit-3 during chick development. Dev Dyn. 2001;222:301–7. https://doi.org/10.1002/dvdy.1182.

42. Dallof A, Da Silva NF, Vacaeva P, Minna JD, Bieche I, Maher ER, et al. SLIT2, a human homologue of the Drosophila Slit2 gene, has tumor suppressor activity and is frequently inactivated in lung and breast cancers. Cancer Res. 2002;62:5974–80. http://www.ncbi.nlm.nih.gov/pubmed/12384951.

43. Anderson KL, Di Cera E. The evolutionary conserved Slit2 gene promoter requires Sp1. Cell Mol Neurobiol. 2016;36:955–64.

44. Ghoul A, Al-Harthy M, El-Demerdash L, El-Dosoky T, El-Mansy A, et al. A genetic approach identifies an ancestral frameshift mutation in RP1 causing retinitis pigmentosa in Egyptian populations. Am J Med Genet A. 2017;173:370–77.

45. Kuemmerle JM, Theiss F, Okoniewski MJ, Weber FA, Hemmi S, Mirsaidi A, et al. Pathogenesis of Parkinson’s disease: a genetic and epigenetic approach. J Neurochem. 2017;142:1–13.
66. Sugimoto M, Gotoh Y, Kawahara T, Sugimoto Y. Molecular effects of polymorphism in the 3′UTR of unc-5 homolog C associated with conception rate in Holsteins. PLoS One. 2015;10.

67. Lapinpek ME, Oliver JA, Bodie JR, Marti F, Kimoto D. The T-cell-specific adapter protein family: Tad1, ALX, and SH2D3A/SH2D8B. Immuno Rev. 2009;223:240–54.

68. Orefi M, Asekkaj R, Bourkadi JE, El Aouad R, Sadiki K. New variant identified in major susceptibility locus to tuberculosis on chromosomal region 8q12–q13 in Moroccan population: A case control study. BMC Infect Dis. 2017;17.

69. Gèrard A, Ségéral E, Naughtin M, Abdouni A, Charmet M, Cheynier R, et al. The Integrase cofactor LEDGF/p75 associates with L1st and Sp6 for Postintegration silencing of HIV-1 gene expression in latently infected cells. Cell Host Microbe. 2015;17:107–17.

70. Khatib H, Zaitoun I, Wiebelsmaier-Finger J, Chang YM, Rosa GM. The association of bovine PPARC1A and OPN genes with milk composition in two independent Holstein cattle populations. J Dairy Sci. 2007;90:2966–70.

71. Weikard R, Kuhn C, Goldammer T, Freyer G, Schwerin M. The bovine PPARGC1A gene: molecular characterization and association of an SNP with variation of milk fat synthesis. Physiol Genomics. 2005;1:13.

72. Holt LJ, Siddle K. Gb30 and Gb31: enigmatic regulators of insulin action - and more! Biochem J. 2005;386:393–406.

73. Hazra A, Kraft P, Lazarus R, Chen C, Chock SJ, Jacques P, et al. Genome-wide significant predictors of metabolites in the one-carbon metabolism pathway. Hum Mol Genet. 2009;18:4677–87.

74. Iyengar L, Rajalakshmi K. Effect of folic acid supplement on birth weights of infants. Am J Obstet Gynecol. 1975:122:332–5.

75. Linke WA, Stockmeier MR, Ivemeyer M, Hosser H, Mundel P. Characterizing titin isoform switch in ischemic human heart disease. Circulation. 2009;119:857–64.

76. Liu Y-J, Liu X-G, Wang L, Dina C, Yan H, Liu J-F, et al. Genome-wide significant predictors of metabolites in the one-carbon metabolism pathway. Hum Mol Genet. 2009;18:4677–87.

77. Labeit S, Fautel M, Lakey A, Trinick J. Towards a molecular understanding of titin. EMBO J. 1992;11:1711–6.

78. Linke WA, Stockmeier MR, Ivermeyer M, Hossie H, Mundel P. Characterizing titin’s I-band lg domain region as an entropic spring. J Cell Sci. 1998:1567–74.

79. Neagoe C, Kulke M, Del Monte F, Gwathmey JK, De Tombe PP, Hajjar RJ, et al. Mechanisms underlying mammalian hybrid sterility in two feline interspecies models. Mol Biol Evol. 2015;32:2534–46. https://doi.org/10.1093/molbev/msv124.

80. Liu Y-J, Liu X-G, Wang L, Dina C, Yan H, Liu J-F, et al. Genome-wide significant predictors of metabolites in the one-carbon metabolism pathway. Hum Mol Genet. 2009;18:4677–87.

81. Liu B, Nie W, Fu X, de Avila JM, Ma Y, Zhu MJ, et al. Neonatal vitamin A injection promotes cattle muscle growth and increases oxidative muscle fiber development in Angus beef cattle. J Anim Sci Biotechnol. 2018;9.

82. Liu L, Gudas LJ. Disruption of the lecithin:retinol acyltransferase gene makes chicks susceptible to toxicosis on beef cattle productivity. J Anim Sci. 1995;73:889–98.

83. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics. 2011;27:2987–93. https://doi.org/10.1093/bioinformatics/btr509.

84. Lominadze D, Tsakadze N, Sen U, Falcone JC, D’Antuono P, Petrukh L, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016;48:1443–6. https://doi.org/10.1038/ng.3679.

85. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75. https://doi.org/10.1086/510795.

86. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75. https://doi.org/10.1086/510795.

87. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75. https://doi.org/10.1086/510795.

88. Endelman JB. Ridge regression and other kernels for genomic selection with R package lRBLUP. Plant Genome. 2011;4:250–5. https://doi.org/10.1093/tpg/spr001.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.