The genetics and genomics of linear type traits in Irish beef cattle

A Thesis Presented for the Degree of

Doctor of Philosophy

by

Jennifer Doyle

Department of Science

Waterford Institute of Technology

[1]*Dept. of Science,*
Waterford Institute of Technology,
Cork Rd,
Waterford,
Co. Waterford

Agriculture and Food Development Authority

[2]*Animal and Grassland Research and Innovation Centre,*
Teagasc,
Moorepark,
Fermoy, Co. Cork

Supervisors

[1]Dr. Siobhán Walsh

[2]Prof. Donagh Berry

Submitted to Waterford Institute of Technology, December 2020

Table of Contents

List of Tables

List of Figures

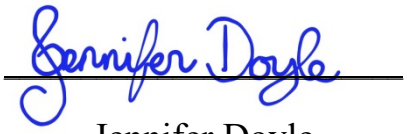
Abbreviations

AA	Angus
BCS	Body condition score
BL	Back length
BLUP	Best linear unbiased predictions
bp	Base pair
BTA	*Bos Taurus* autosome
CD	Chest depth
CH	Charolais
CHR	Chromosome
CSO	Central Statistics Office
CW	Chest width
DAFM	Department of Agriculture Food and the Marine
DBI	Dairy beef index
DHQ	Development of hind quarter
DIT	Development of inner thigh
DL	Development of loin
DOC	Docility
EBI	Economic breeding index
EBV	Estimated breeding value
FL-FV	Fore leg front view
GAO	Gross agricultural output
GO	Gene Ontology
GRM	Genomic relationship matrix
GWAS	Genome wide association study
HD	High density
HE	Hereford

HF	Holstein Friesian
HL-RV	Hind leg rear view
HL-SV	Hind leg side view
HMM	Hidden Markov model
HW	Hip width
HWE	Hardy Weinberg equilibrium
ICBF	Irish Cattle Breeding Federation
KEGG	Kyoto Encyclopaedia of Genes and Genomes
LD	Linkage disequilibrium
LM	Limousin
LOCO	Locomotion
MAF	Minor allele frequency
Mb	Megabase
OSW	Overlapping sliding window
PAR	Pseudo-autosomal region
PL	Pelvic length
QTL	Quantitative trait loci
RBI	Relative breeding index
SI	Simmental
SNP	Single nucleotide polymorphism
TW	Thigh width
WBW	Width behind wither
WGS	Whole genome sequence
WH	Wither height
WOW	Width of wither

Statement of Original Authorship

I hereby certify that the submitted work is my own work, was completed while registered as a candidate for the degree stated on the Title Page, and I have not obtained a degree elsewhere on the basis of the research presented in this submitted work.



Jennifer Doyle

Acknowledgements

I would like to acknowledge the opportunity presented to me by Teagasc and Waterford Institute of Technology to carry out this research. I would also like to acknowledge Science Foundation Ireland for providing the funding for this research to be completed. In addition, I wish to thank the Irish Cattle Breeding Federation for supplying all the data used in this thesis. To my supervisors Dr. Siobhán Walsh and Prof. Donagh Berry, thank you both for all your support, guidance, and advice over the last 4 years and for sharing your vast knowledge, skills, and expertise with me. I would like to especially thank Prof. Donagh Berry for always having an open door and for being so patient when things didn't always go to plan - your words of encouragement and your dedication and enthusiasm for all things genetics always managed to get me back on track. A special thank you to Dr. Deirdre Purfield and Dr. Tara Carthy, who not only contributed to this PhD as co-authors but also provided so much help and support over the years. I would also like to thank my other co-authors, Prof. Roel Veerkamp of Wageningen University and Dr. Ross Evans of the Irish Cattle Breeding Federation. To the Moorepark genetics team, thank you for all the laughs (and even the tears!), the teatime chats and the office debates made this all the more enjoyable. To my fellow MPK nerd, Rachel, the random chats, rants (and later voice notes) about Superheroes, Harry Potter, GOT, and just life in general were always much appreciated!

To my amazing friends, Shauna, Jamie, Cian, & Keith, thank you for always being there for me and for always believing in me even when I didn't believe in myself. To all the girls in the yard and of course to my pony, Casanova, thank you for providing a much-needed distraction with all our horsey adventures. Last but not least, I want to thank my family for all their encouragement and support throughout this journey. A special mention and thanks to Ollie, for giving me his toy bull that is used as the model throughout this thesis. To my brother, AJ, thank you for always challenging me and for keeping me grounded. To Allie, our random kitchen chats were always a welcome distraction. Thank you to my Grandparents, Jack & Lilzer, for your endless love and support even if you don't quite get what I actually do. To my parents, to whom this thesis is dedicated, you are the most resilient and patient people I know, this process hasn't been easy, but you have always believed in me and been behind me every step of the way. Thank you for all the love and sacrifices you have made for me, not only during my PhD, but throughout my entire life. None of this would have been possible without you.

Abstract for thesis titled "The genetics and genomics of linear type traits in beef cattle" by Jennifer Doyle

Irish beef genetic evaluations are currently undertaken using a multi-breed population; thus estimated breeding values for all beef animals are comparable regardless of breed. The two indexes published on Irish beef cattle, the Replacement index and Terminal index, both include carcass traits. These traits, however, are only measured after the animal is slaughtered. Linear type traits are measured on young live animals and are strongly correlated with carcass merit. The value in linear type traits is the ability to select for more morphologically superior carcasses, even for the same carcass weight. The objectives of this thesis were to: 1) determine if the genetic architecture of 5 muscular and 5 skeletal linear type traits differ by breed and/or sex with the aim of improving the accuracy of multi-breed beef genetic evaluations, using linear type traits as an example, and 2) to detect quantitative trait loci (QTL) associated with the linear type traits. Of particular interest was if detected QTLs overlapped both among traits and among breeds. Data used consisted of phenotypic data on 198,351 animals and imputed whole genome sequence data on 23,943 animals from 5 beef breeds and the Holstein-Friesian dairy breed. The heritability estimates and pairwise genetic correlations among the linear traits estimated within breed were similar to the respective statistics across the 3 continental breeds with the same phenomenon observed when comparing the two British breeds (i.e. Angus and Hereford). The majority of the QTL identified as being associated with the linear type traits were both trait- and breed-specific, with only some overlap in the QTLs occurring between the Charolais and Limousin for the muscular traits, while for the skeletal traits there was commonalities between the Angus and Limousin as well as between the Angus and Holstein-Friesian. While sexual dimorphism was evident at a genome level, only 1% of SNPs tested exhibited it; this was consistent with the near unity genetic correlations between the same linear type trait in both sexes estimated using mixed models. In summary, considering the continental beef breeds separately to the British beef breeds in genetic evaluations may improve the accuracy of these evaluations; however, it is unlikely that the consideration of each sex separately will impact the accuracy of selection. Furthermore, including the linear type traits in multi-trait genetic evaluations alongside (more granular) carcass traits may enable the breeding of morphologically different animals in the future with a more valuable carcass, even for the same carcass weight.

Chapter 1

Introduction & Review of the Literature

1.1 Introduction

The objective of a genetic evaluation is to estimate an animal's genetic potential for a particular trait, or set of traits, by taking into account all relevant performance data and disentangling the genetic effects from the environmental influences. The resulting estimated breeding value (EBV) provides breeders with a number to base selection decisions on. In Ireland, beef cattle are evaluated using two multi-breed indexes, the Replacement Index and the Terminal Index, in which all breeds are assumed to be genetically similar. While the breeding objectives catered to by these indexes are very different, both put some emphasis on carcass traits, including carcass weight, carcass conformation, and carcass fat, that can only be measured after an animal is slaughtered. Linear type traits are measured on the young, live animals and are known to be moderately to strongly genetically correlated with carcass merit. Therefore, it is possible that utilising these traits in a genetic evaluation might provide a more accurate way to predict carcass merit while also breaking the carcass down into more granular traits such as length, height, width etc. that may add value to the carcass through manipulation of the morphology of an animal to provide more of the higher value cuts of meat.

The overall objective of the present thesis was to determine if the genetic architecture of five beef breeds and the different sexes differed by breed and/or sex with the aim to improve the accuracy of multi-breed beef genetic evaluation, using linear type traits as an example. Chapter 1 summarises the current Irish genetic evaluations and summarises the literature of both the genetics and genomics of linear type traits, as well as the phenomenon of sexual dimorphism in cattle. The genetic parameters for each of the linear type traits within each of the breeds, and also within the sexes in the Charolais and Limousin breeds are quantified in Chapter 2. Chapter 3 and Chapter 4 are

both genome wide association studies where the aim was to identify genomic regions associated with the muscular (Chapter 3) and skeletal (Chapter 4) type traits and to evaluate whether these regions are common across breeds. These chapters also identify genomic regions associated with both the linear type traits and the carcass traits that could potentially be used to alter the morphology of an animal to increase the value of the cuts obtained from a carcass. In Chapter 5, the genomic regions associated with the type traits in each sex are identified and the extent of sexual dimorphism is quantified across the entire cattle genome. Overall, considering each trait as a separate trait within each sex is unlikely to improve the accuracy of genetic evaluations in beef cattle. However, considering the continental beef breeds separately to the British beef breeds may increase the accuracy of any future evaluations and/or genomic predictions. The thesis is summarised in Chapter 6, alongside the implications and conclusions of the presented body of work.

1.2 Irish Agriculture

In 2018, Gross Agricultural Output (GAO) was valued at €8.65 billion to the Irish economy (DAFM, 2019). Beef and dairy are the largest agri-food industries accounting for 38.8% and 29.5% of GAO, respectively. Of the 6.6 million cattle in Ireland in 2018, 1.37 million were dairy cows and 0.98 million were beef cows (CSO, 2019). It is estimated that Ireland is 650% self-sufficient in beef, making it the fourth largest net exporter of beef globally. In 2017, beef exports were estimated to be valued at more than €2.5 billion (Hennessy, 2018).

1.2.1 Irish Cattle Breeding Federation

The Irish Cattle Breeding Federation (ICBF) which was established in 1998 is a non-profit organisation in charge of providing cattle breeding information and services to the Irish beef and dairy industries. The main aim of the ICBF is to benefit farmers and the wider community through genetics, specifically via genetic gain, whereby the cattle identified to become the parents of the next generation of cattle are predicted to be the most genetically superior animals. In order to identify these genetically superior animals, ICBF maintains the national Irish cattle breeding database (Figure 1.1) which contains information on all cattle registered nationally. This database contains information from both the Irish beef and dairy herds including pedigree information, animal events data (e.g., calving, inseminations, health records), milk records, abattoir data, animal auction data, information gleaned from dairy and beef genetic evaluations and animal genotypes (Evans et al., 2007).

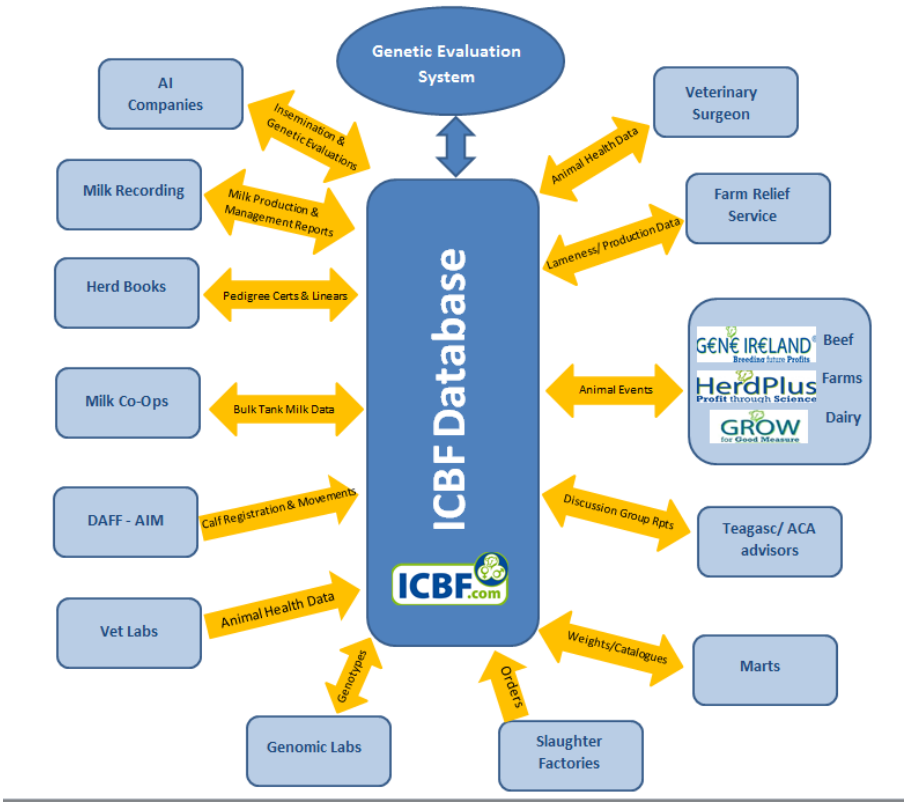


Figure 1.1 The inputs and outputs from the ICBF database. (Source: www.icbf.com)

A number of breeding indexes have been developed by the ICBF to reflect the different breeding objectives of the beef and dairy industries (Amer et al., 2001; Berry et al., 2007; Cromie et al., 2016; Berry et al., 2019). The national indexes include the Economic Breeding Index (EBI) for dairy cattle, the Replacement and Terminal Indexes for beef cattle, and the recently introduced Dairy-Beef Index for the selection of beef bulls for use in the dairy herd.

1.2.1.1 The Economic Breeding Index

The EBI was introduced in 2001 to replace the Relative Breeding Index (RBI) which was based solely on milk production (Berry et al., 2007). The EBI aims to help farmers identify the most profitable dairy bulls and cows to become parents of the next generation by providing a single figure profit index. This index is comprised of seven sub-indexes which include milk production, fertility, calving performance, beef merit, cow maintenance, cow management, and health. The evolution of the relative weighting of these sub-indexes is outlined in Figure 1.2.

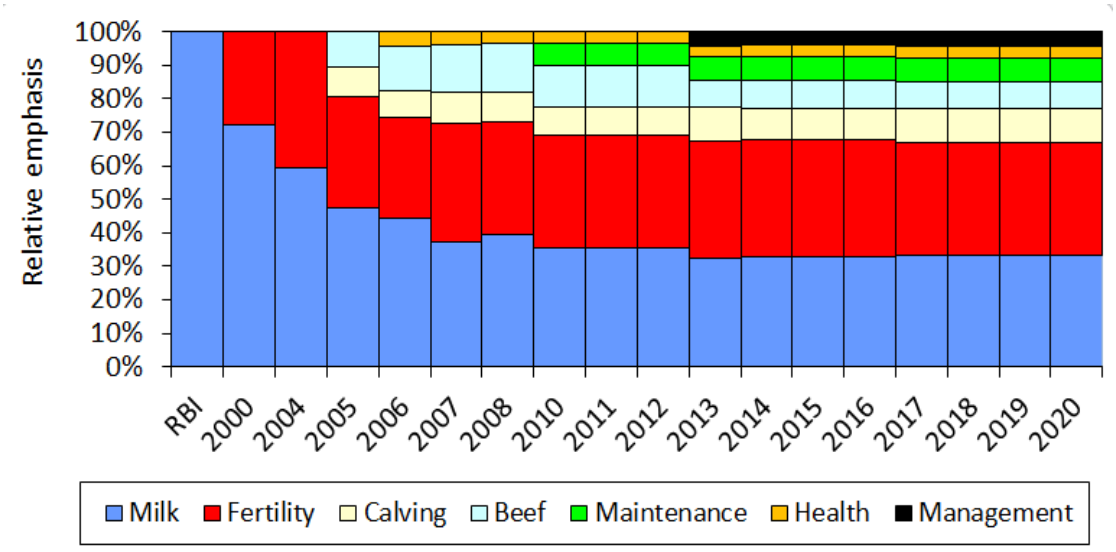
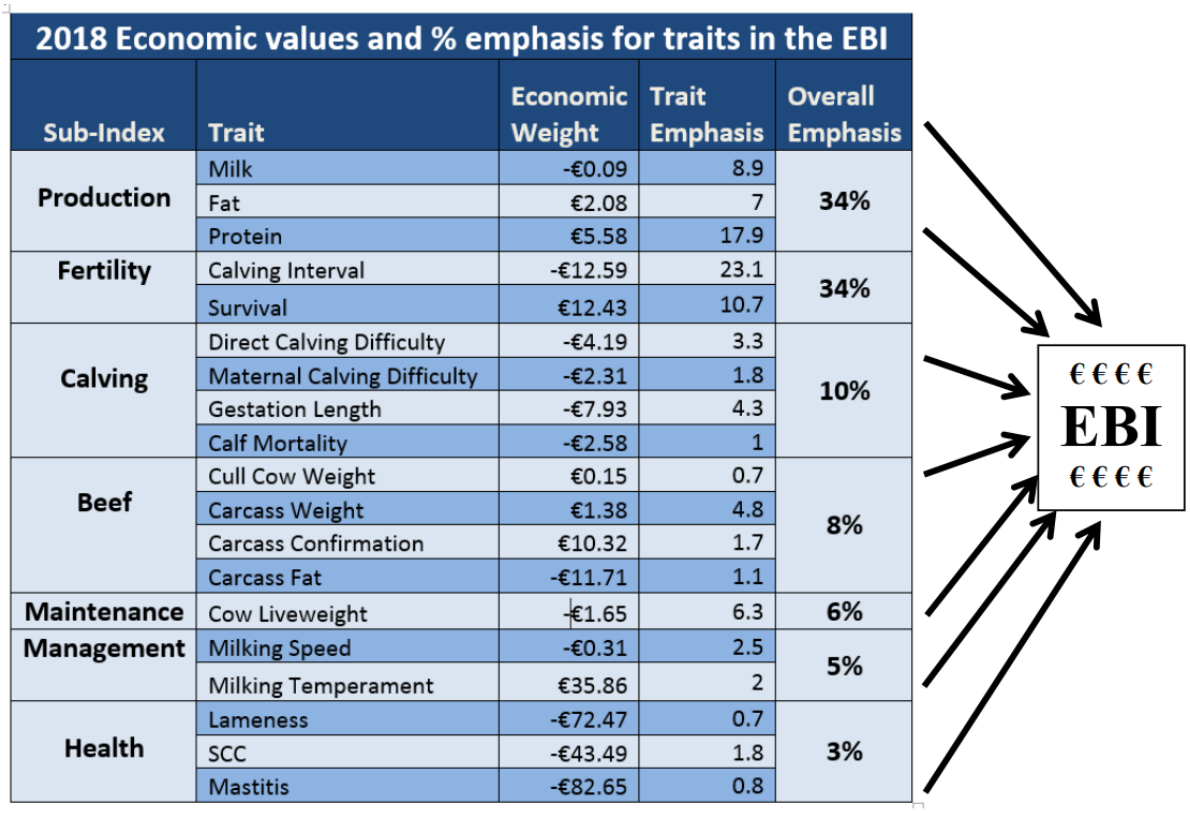


Figure 1.2 The evolution of the EBI over the last two decades showing the changing emphasis on various traits. (Source: www.icbf.com)

The weightings on each sub-index within the EBI are based on the perceived value of that sub-index to the farmer and wider dairy industry. Consequently, milk production and fertility currently account for 68% of the emphasis within the EBI (Figure 1.3). In recent years, the emphasis on calving traits has increased. The calving sub-index helps to identify easy calving bulls; easier calving leads to lower mortality rates in calves and cows during parturition. The latest inclusion to the EBI is the management sub-index which accounts for 5% of the total emphasis and describes the milking speed and milking temperament of the animal. Since the introduction of the EBI, it is estimated that Irish dairy farmers are now making an extra profit of approximately €280 per cow per lactation (www.ICBF.com).



2018 Economic values and % emphasis for traits in the EBI				
Sub-Index	Trait	Economic Weight	Trait Emphasis	Overall Emphasis
Production	Milk	-€0.09	8.9	34%
	Fat	€2.08	7	
	Protein	€5.58	17.9	
Fertility	Calving Interval	-€12.59	23.1	34%
	Survival	€12.43	10.7	
Calving	Direct Calving Difficulty	-€4.19	3.3	10%
	Maternal Calving Difficulty	-€2.31	1.8	
	Gestation Length	-€7.93	4.3	
	Calf Mortality	-€2.58	1	
Beef	Cull Cow Weight	€0.15	0.7	8%
	Carcass Weight	€1.38	4.8	
	Carcass Confirmation	€10.32	1.7	
	Carcass Fat	-€11.71	1.1	
Maintenance	Cow Liveweight	-€1.65	6.3	6%
Management	Milking Speed	-€0.31	2.5	5%
	Milking Temperament	€35.86	2	
Health	Lameness	-€72.47	0.7	3%
	SCC	-€43.49	1.8	
	Mastitis	-€82.65	0.8	

Figure 1.3 Traits, economic weights, trait emphasis and overall emphasis of each sub-index in the EBI. (Source: www.icbf.com)

1.2.1.2 Beef Indexes

The ICBF introduced two new beef cattle indices in 2012 to replace the single Suckler Beef Value Index; the indexes are the Replacement Index and the Terminal Index. As these names suggest, the Replacement Index is used to select the most suitable sires and females to produce females that will ultimately be low maintenance and profitable cows used to replace the current cows in the herd (ICBF, 2013; ICBF, 2019). The Replacement Index is made up of two sub-indexes; those relating to the cow and those relating to the calf, as outlined in Figure 1.4.

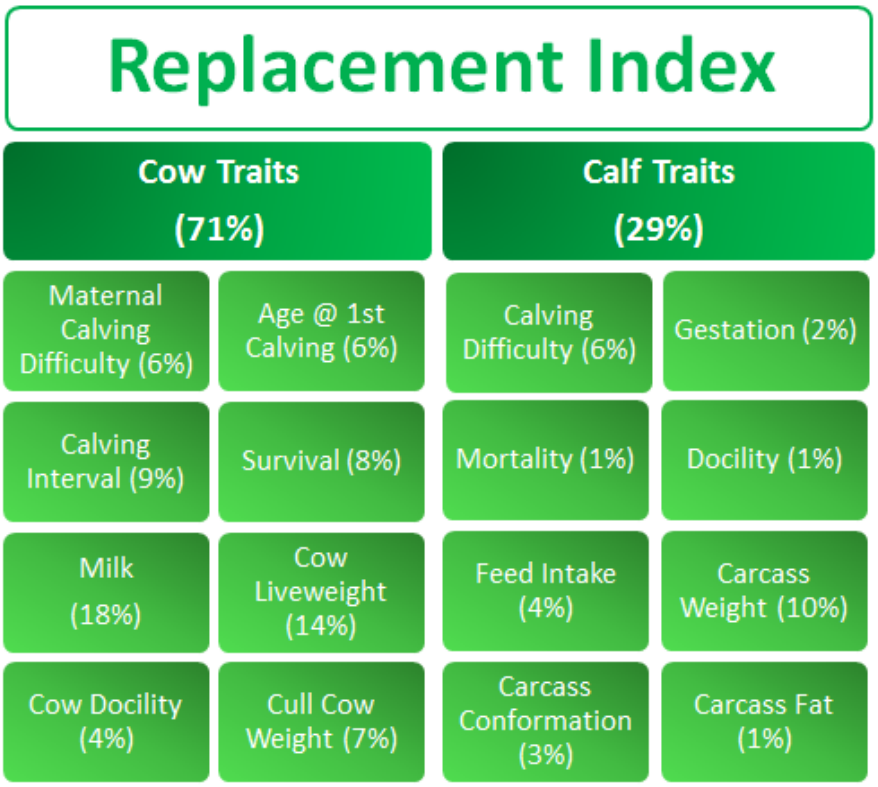


Figure 1.4 Relative emphasis of the traits that make up the beef maternal (replacement) index (ICBF, 2019).

The aim of the terminal index is to rank sires and dams to produce the next generation of profitable animals for slaughter (ICBF, 2013); therefore, the main principle of the terminal index is on low cost of production and higher value leading to

a higher profit for the farmer (ICBF, 2019). Sires and dams selected using the terminal index should have reduced costs associated with maintenance and growth and produce a higher value carcass. Consequently, carcass traits such as carcass conformation, weight, and fat, account for 57% of the emphasis within the terminal index (Figure 1.5).

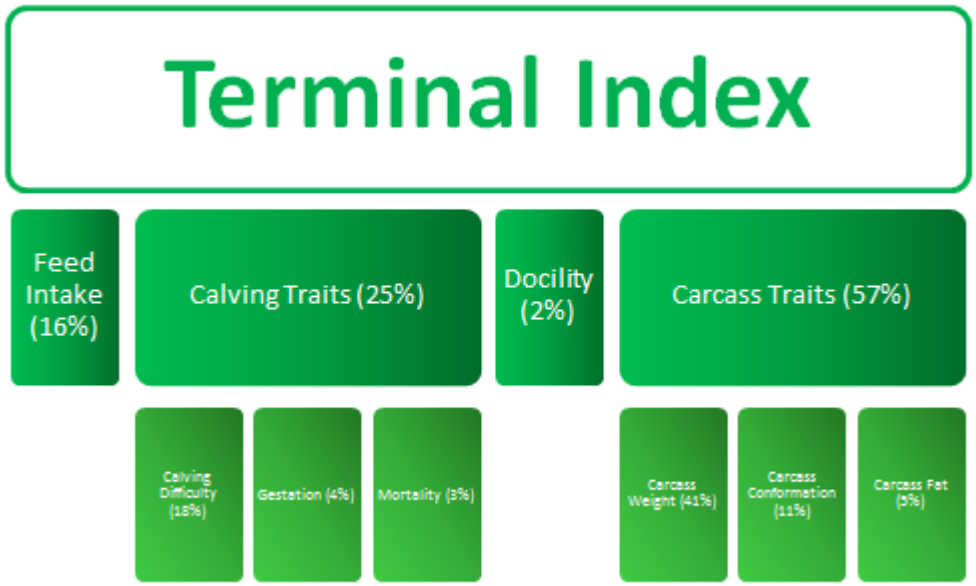


Figure 1.5 Relative emphasis of the traits that make up the beef terminal index (ICBF, 2019)

1.2.1.3 Dairy Beef Index

The Dairy Beef Index (DBI) is a tool used to promote the breeding of high quality beef cattle from the dairy herd that are more profitable as calves or at slaughter, but have minimal consequences on the calving difficulty or gestation length for the dairy female (Berry et al., 2019a). The main concern of dairy farmers using beef bulls on their females is the perceived increase in calving difficulty. To reflect this, calving traits account for 63% of the emphasis within the DBI (Figure 1.6). Carcass traits account for 24% of the emphasis within the DBI in order to generate a higher calf price and

maximise the quality of the carcass when using beef bulls with a high DBI on the dairy herd in comparison to choosing bulls based on calving traits alone.

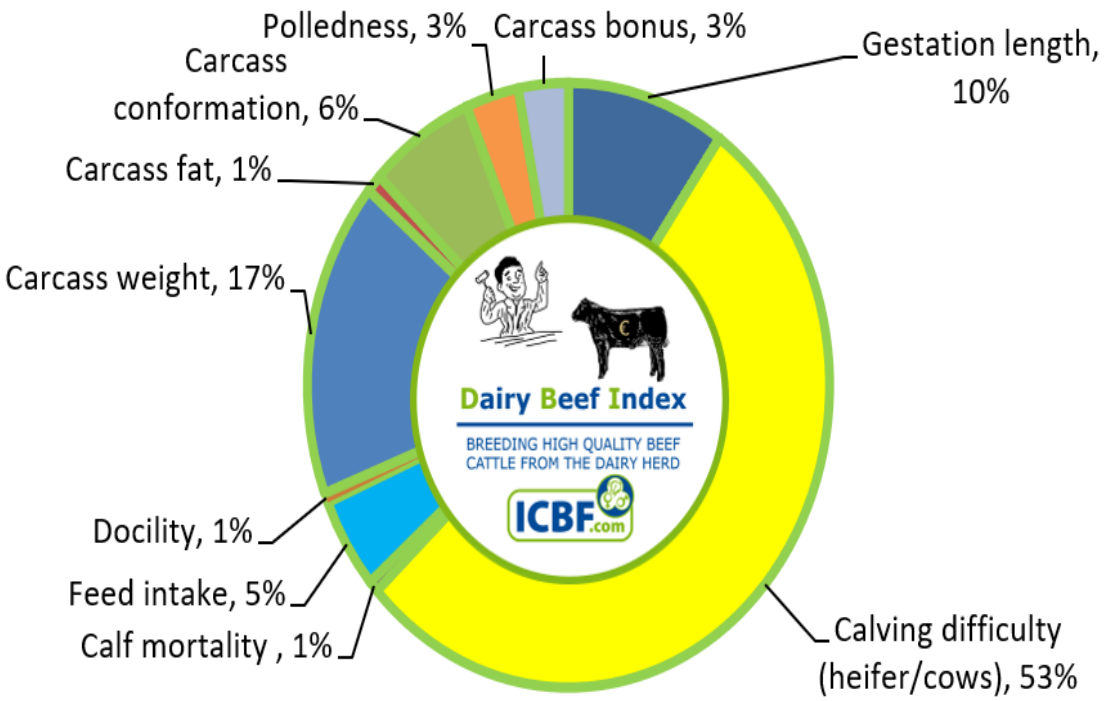


Figure 1.6 Relative emphasis of the traits included in the dairy beef index when it was first introduced in spring 2019. (Source: www.icbf.com)

1.3 Statistical & Genetic Methodology

1.3.1 Genetic Terminology

1.3.1.1 Phenotype

A phenotype is simply the observable characteristic (e.g., performance) of the animal. It may refer to the animal's appearance, conformation, behaviour, or development; for example height, width, and colour. A phenotype may be continuous or discrete. Continuous (quantitative) traits vary within a population to produce a range of values for that trait; examples of continuous traits include height and weight. Discrete

(qualitative) or categorical traits are either present or not in that animal, and include, for example whether an animal is a particular colour or not.

1.3.1.2 Genotype

Animal breeders traditionally used the broader sense definition of a genotype, where genotype refers to the overall genetic makeup of the animal i.e. it describes the animal's genome. Molecular geneticists, and more lately now animal breeders, are tending to use the narrow sense definition, where genotype refers to the particular allele an animal possesses at a particular locus or the entire complement of alleles.

1.3.1.3 Variance

In statistics, variance is a measurement of the spread between numbers in a dataset. In genetics, variance is a measure of the variation within a sample population i.e. the differences that exist among the individuals in that population. Variation of phenotypes within a population is due to differences in both genetic and environmental variation. Genetic variation across populations can be caused by variation in alleles and their effects and exists because of evolutionary forces such as selection, inbreeding, genetic drift or mutations. Once genetic variation exists for a trait then breeding for improvement is possible (Falconer, 1952). Environmental variation can be both permanent and temporary and reflects external factors including the herd the animal is in, the management of the animal/herd, or the feed the animal receives, etc.

1.3.1.4 Heritability

Heritability may be defined as the proportion of phenotypic variance between individuals that is attributable to genetic differences between individuals, or the strength of a relationship between an individuals observed performance and its true genetic merit

(after adjustment for environmental effects; (Lush, 1949; Berry et al., 2019b). Heritability is measured on a scale from 0 (not heritable) to 1 (highly heritable).

Animal breeders usually use heritability in the narrow sense (h^2) which is the proportion of phenotypic variance due to additive genetic variance. Heritability is calculated using the following equation:

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2}$$

where σ_a^2 is the additive genetic variance and σ_e^2 is the residual variance. Adding the additive variance to the residual variance gives the total phenotypic variance. In the broader sense, the calculation of heritability (H^2) also takes into account the non-additive genetic variation in both the numerator and denominator of the calculation.

1.3.1.5 Genetic Variation

Genetic variation is the term used to describe the cumulative effects of the variation in the DNA sequence in each animal's genome within a population. Genetic variation (σ_g^2) can also be presented as the genetic standard deviation, the square root of the genetic variation ($\sqrt{\sigma_g^2}$) or simply (σ_g). Genetic variation can be divided into 3 subcategories: additive variance (σ_a^2), dominance variance (σ_D^2), and epistatic variance (σ_I^2):

$$\sigma_g^2 = \sigma_a^2 + \sigma_D^2 + \sigma_I^2$$

Dominance variation results from the interaction between alternative alleles at the same locus while epistatic variation results from the interaction between alleles at different loci.

1.3.1.6 Genetic Correlation

The genetic correlation (r_g) describes the linear relationship between two variables due to the genetic influences on each variable. Genetic correlations vary from -1 (strong negative relationship) to 0 (no relationship) to +1 (strong positive relationship).

The genetic correlation within a population is calculated using the following equation:

$$r_g = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2 . \sigma_y^2}}$$

where σ_{xy} is the genetic (co)variance between traits x and y, and σ_x^2 and σ_y^2 are the additive genetic variation of x and y.

1.3.1.7 Genetic gain

Genetic gain is the predicted change in the mean value of a trait over time. The expected annual genetic gain is calculated using the following equation:

$$\Delta G = \frac{i \, . r . \sigma_a}{L}$$

where ΔG is the change in performance or genetic gain, i is the intensity of selection, r is the accuracy of selection, σ_a is the additive genetic standard deviation of the trait being investigated, and L is the generation interval.

Despite a low heritability for some traits, the accuracy of selection can still be high if a sufficient number of progeny records or genomic information is available. Therefore, genetic gain is possible for any trait that exhibits genetic variation within a population (Berry et al., 2011).

1.3.2 Genetic Evaluations

Genetic evaluations in cattle are routinely undertaken using linear models. The most common types of linear models include fixed effect models (not used in genetic evaluations), which contain only fixed effects, or mixed models (commonly used in genetic evaluations), which contain both fixed and random effects.

Fixed effects: Fixed effects can be either continuous or class variables and, if the latter, usually have a finite number of levels which can all be accounted for in the model. Examples include sex and age.

Random effects: Random effects usually have an infinite number of levels and the levels included in the model are a random sample from a larger population. Examples include individual and contemporary group.

1.3.2.1 Best Linear Unbiased Predictions (BLUP)

Best Linear Unbiased Predictions (BLUP) methodology was developed by Henderson (Henderson, 1949; Henderson, 1950; Henderson, 1975) where estimates of breeding values and fixed effects are simultaneously estimated within the framework of mixed models. Best Linear Unbiased Predictions incorporates information from relatives to generate an unbiased estimate of the genetic merit of an individual. The properties of BLUP are incorporated into the name:

Best: minimizes prediction error variance by maximizing the correlation between true and predicted breeding value.

Linear: predicted breeding values are linear functions of the observations.

Unbiased: estimates of fixed effects and of realized values for a random variable (e.g., breeding value) are unbiased.

Prediction: the true breeding value is predicted

The models used in BLUP have evolved over the decades from simple models, such as a sire model, to the more complex models including animal and multi-trait linear mixed models (Mrode, 2014).

1.3.2.2 Linear Mixed Models

The basic form of BLUP is the linear mixed model which is used to model phenotypic performance. It is expressed as:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{e}$$

where $\boldsymbol{y}$ is a n x 1 vector of observations, $\boldsymbol{\beta}$ is a p x 1 vector of fixed effects; where p is the number of levels for fixed effects, $\boldsymbol{u}$ is a q x 1 vector of random animal effects; where q is the number of levels of random effects, **X** is a design matrix which relates records to fixed effects and is of the order n x p and $\boldsymbol{Z}$ is a design matrix which relates records to random animal effects and is of the order n x q (Mrode, 2014).

1.3.2.3 Animal models

An animal model produces an estimated breeding value (EBV) for each animal via the incorporation of a pedigree cert which contains records on the animal itself and information from its relatives. The following matrix equation is used to calculate EBV from animal models:

$$\begin{bmatrix} \boldsymbol{X'X} & \boldsymbol{X'Z} \\ \boldsymbol{Z'X} & \boldsymbol{Z'Z + A^{-1}}\alpha \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{b}} \\ \widehat{\boldsymbol{a}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'y} \\ \boldsymbol{Z'y} \end{bmatrix}$$

where **X** is a design matrix which relates records to fixed effects, **Z** is a design matrix which relates records to random animal effects, **A** is the numerator relationship matrix, α is the ratio of the residual variance to the additive variance, $\widehat{\boldsymbol{b}}$ is the fixed effect solutions, $\widehat{\boldsymbol{a}}$ is the random effects solutions, **y** is the vector of phenotypic records (Mrode, 2014).

1.3.2.4 Multi-Trait Models

Multi-trait models use the genetic and phenotypic correlations among traits to generate EBVs with a greater reliability than that achieved through using other models. The following equation is used in multi-trait models:

$$\boldsymbol{y_i} = \boldsymbol{X_i b_i} + \boldsymbol{Z_i u_i} + \boldsymbol{e_i}$$

where $\mathbf{y}_i$ is a vector of phenotypic observations relating to trait *i*, $\mathbf{b}_i$ is the vector of fixed effects for trait *i,* $\boldsymbol{u_i}$ represents the random effects of the *i*th trait, $\mathbf{X_i}$ and $\mathbf{Z_i}$ are design matrices that relates records to the fixed and random effects, respectively, and $\boldsymbol{e_i}$ is the residual error.

Multi-trait models can be advantageous when used to generate EBVs for a lowly heritable trait. If one trait has a much higher heritability than the other trait in the model, the lower heritability trait tends to gain more in accuracy during a multi-trait analysis if a genetic correlation between traits exists. Multi-trait models are also very useful to account for selection in the data, assuming the trait upon which selection was practiced, is also included in the model.

1.3.2.5 Selection Index Theory

Selection index theory (Hazel, 1943) is used in most breeding programs to select for multiple traits simultaneously. This method combines information from multiple sources, such as information on the same trait from different relatives or different traits measured on the animal itself or its relatives, to predict the animal's EBV. Although BLUP also incorporates relationship information, selection index theory can be used to derive weights for each trait which determine the value of different traits from different sources, and can subsequently calculate the accuracy of, and response to, selection. The selection index is calculated as:

$$Selection\ Index = b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_nX_n$$

where X represents a trait in the index and b represents the index weight designed to maximise the response to selection. The index weights can be calculated in two ways:

1. $$\boldsymbol{b} = \boldsymbol{P}^{-1}\boldsymbol{G}\boldsymbol{a}$$

where **P** is the *n* x *n* phenotypic (co)variance matrix among the n selection criteria , **G** is the *n* x *m* genetic (co)variance matrix among the *n* selection criteria and *m* objective traits in the index, and **a** is a vector of economic values associated with each trait (Hazel, 1943). This calculation for b is used where weights are applied to phenotypes and the selection criteria are not the same as the goal traits in the index

2. $$\boldsymbol{b} = \boldsymbol{G_{11}}^{-1}\boldsymbol{G_{12}}\boldsymbol{a}$$

where $\boldsymbol{G_{11}}$ is an *n* x *n* matrix of the genetic co-variances among the *n* selection criteria, $\boldsymbol{G_{12}}$ represents an *n* x *m* matrix of the 'true' genetic co-variances between the *n* selection criteria and the *m* objective traits, and a is an *m* x 1 vector of economic values for all objective traits (Schneeberger et al., 1992). This calculation for b is used where weights are to be applied to EBVs and the selection criteria are not the same as the goal traits.

Selection index coefficients are rarely known with certainty due to errors in the co-variances and economic values used (Ochsner et al., 2017); thus, it is often important to determine the sensitivity of an index to fluctuations in these coefficients. One way to do this is to calculate the efficiency (E_u) of the index (Ochsner et al., 2017):

$$E_u = \frac{\boldsymbol{b'}_u\boldsymbol{G}_{12t}}{\sqrt{\boldsymbol{b'}_u\boldsymbol{G}_{11t}\boldsymbol{b}_u}}(\sqrt{\boldsymbol{b'}_t\boldsymbol{G}_{12t}})^{-1}$$

where $\mathbf{G}_{11t}$ represents an *n* x *n* matrix of the 'true' genetic co-variances among the *n* selection criteria, $\mathbf{G}_{12t}$ represents an *n* x *m* matrix of the 'true' genetic co-variances between the *n* selection criterion and the *m* objective traits, $\mathbf{b}_t$ is a vector of index

coefficients derived from the 'true' values, and $\mathbf{b}_u$ is a vector of used index coefficients. The used index coefficients are based on current belief while the 'true' values are assumed to be the optimum (Ochsner et al., 2017). This calculation can be useful for determining the impact of using co-variance components in an index of *n* traits in one breed to predict the performance of those traits in another breed, such as what could occur in a multi-breed index. The 'true' parameters would be calculated from the co-variance parameters of the traits from the breed whose performance was being predicted while the used parameters could be derived from the co-variance parameters of the traits in another breed.

1.4 Genotyping & Genomics

Genotyping is the process of detecting allelic variability that exists in an animal by examining its DNA sequence. The most common type of variation in DNA is the single nucleotide polymorphism (SNP), although structural variants such as deletions, duplications, insertions and copy number variants also exist.

1.4.1 Single Nucleotide Polymorphisms (SNPs)

A SNP is a single base-pair change in an individual's DNA (Figure 1.7) that occurs at a specific position. For example, a SNP may replace the nucleotide thymine (T) with the nucleotide cytosine (C) in a certain position in the DNA. Using this example, if more than 1% of a population carries the C, the least frequent allele, in place of the T then this locus can be classified as a SNP (Vignal et al., 2002). Single nucleotide polymorphisms are usually biallelic; there is usually only two possible alleles that may occur at a particular locus (Vignal et al., 2002).

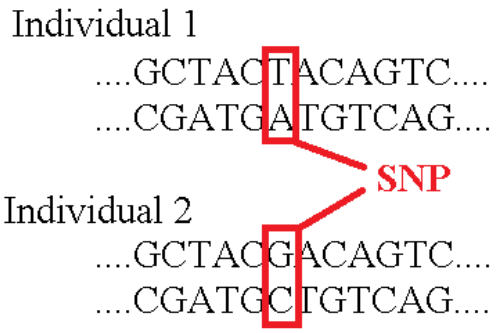


Figure 1.7 An example of a single nucleotide polymorphism (SNP; highlighted by the red box) in a DNA sequence

1.4.2 Genotype Panels

Numerous genotype panels are available to genotype cattle; the main difference between these panels is the density of SNPs but also which SNPs are included on the panel. Generally, the greater number of SNPs on a panel, the greater the cost. Higher costs can deter farmers from genotyping their animals; therefore, reducing the number of SNPs genotyped while retaining high coverage of the genome coupled with the process of imputation has been the focus of much research over the years (Boichard et al., 2012; Judge et al., 2016).

The genotype panels commonly used in Ireland are as follows:

- *Illumina 3k*. This chip was developed as a cost-effective method of genotyping large numbers of animals. The Illumina Bovine 3k beadchip tests for just 2,909 SNPs but aims to provide the same level of coverage as the Illumina SNP50 chip that genotypes 54,001 SNPs. The 3k panel takes advantage of linkage disequilibrium between markers on the SNP50 and enables accurate imputation to SNP50 density (Illumina, 2011).
- *Illumina LD.* The Illumina BovineLD beadchip was developed to support low cost genotyping that would enable accurate imputation to higher density

genotypes in dairy and beef cattle (Boichard et al., 2012). It tests for 6,909 SNPs that generally have a high minor allele frequency and are relatively uniformly spaced across the genome, except for at the ends of chromosomes where the density was increased.

- *Illumina SNP50*. This chip was one of the first developed by Illumina, in conjunction with the United States Department of Agriculture, the University of Missouri, and the University of Alberta (Matukumalli et al., 2009). This chip tested 54,001 SNPs in version 1, 54,609 SNPs in version 2 and now tests 53,714 SNPs in version 3. These SNPs were identified from a number of different sources, including the Bovine HapMap data set, Btau assembly SNPs, and whole genome shotgun reads; the SNPs are relatively evenly and strategically placed across the bovine genome.
- *Illumina HD.* The Illumina BovineHD beadchip is one of the most comprehensive cattle genotyping chip available testing for 777,962 SNPs across the entire bovine genome.
- *IDB SNP chip.* This chip is the most widely used for genotyping cattle in Ireland as it was custom made by researchers from ICBF and Teagasc (Mullen et al., 2013) to lower the cost of genotyping for Irish farmers. The SNPs tested on this panel have been documented as previously associated with a number of lethal and unwanted traits, beneficial traits, and performance traits (i.e., milk and meat traits). The IDB SNP chip is currently on Version 4 but is being regularly updated. The number of SNPs evaluated by each panel increased from Version 1 (17,137 SNPs) to Version 2 (18,004 SNPs) to Version 3 (53,450 SNPs), while the number of SNPs on Version 4 has decreased slightly to 52,580 SNPs, without any loss to coverage.

1.4.2 1000 Bull Genomes Consortium Project

The first whole genome assembly of the bovine genome was published in 2009 (Zimin et al., 2009). Two years later, the 1000 Bull Genomes Consortium Project was founded in partnership with many agricultural research centres worldwide, including Teagasc. The aim of the consortium was to 1) reduce the costs of generating the sequence of an animal by providing access to a large reference database of genetic variants to enable increasingly accurate imputation from genotype arrays to full sequence and 2) use these data to identify mutations that compromise animal health, welfare, and productivity (Hayes et al., 2012; Daetwyler et al., 2014). For the first run of the 1000 Bulls project, the database only contained 234 bulls from three breeds, Friesian, Fleckvieh, and Jersey, and identified 28.3 million genetic variants (Daetwyler et al., 2014). The latest run (7.0) contains sequence data on >3,800 cattle from >150 breeds of both *Bos Taurus* and *Bos Indicus* and had identified >150 million filtered genetic variants.

1.4.3 Imputation

Whole genome sequencing of thousands of individuals is costly. Thus, sequencing is not widely used in animal breeding programs. Through the use of imputation, the genotype information of an unobserved genotyped marker can be inferred from the genotype information of the low density genotype panels. Imputation methods are generally either 1) family-based which exploits linkage information between close relatives with a known pedigree, or 2) population-based which does not require pedigree information and therefore uses population linkage information (Li et al., 2009). Family-based imputation methods are the most intuitive as genotypes for a small number of genetic markers can be used to infer haplotypes which are identical-by-descent and shared between individuals of known relationship (Li et al., 2009). These methods are

reasonably accurate (Sargolzaei et al., 2014). Population-based methods assume that the individuals are unrelated. Population-based methods can identify close relationships between individuals by identifying haplotypes that are identical by descent; however, these haplotypes may be shorter than those identified in family-based methods (Li et al., 2009; Sargolzaei et al., 2014). Population-based methods of imputation can be highly accurate if both the number of markers and the number of reference individuals is large, although they can be computationally intensive (Sargolzaei et al., 2014).

1.4.3.1 Methods of Imputation

1) Hidden Markov Model

The basic theory of Markov models was published in a series of mathematical papers in the late 1960s and early 1970s (Baum and Petrie, 1966; Baum and Eagon, 1967; Baum et al., 1970) and has since been adopted as a method for genotype imputation by the genetics community. A Markov model generally describes a sequence of observable symbols that are controlled by the state the model is in (Rabiner, 1989). In comparison, Hidden Markov models (HMM), as used by geneticists (Hickey et al., 2015) are, as indicated by the name, hidden or unobservable (Rabiner, 1989).

A HMM is characterised by the number of states in the model - N, the number of distinct observation symbols - M, the state transition probability distribution – A = $\{a_{ij}\}$ and the observation symbol probability distribution in state j. In the HMM, there is a hidden sequence of states and the probability of observing a symbol at a specific position in a sequence depends only on the state at that specific position while the probability of a particular state at the position x+1 depends only on the state at the position x (Hickey et al., 2015). When the observed symbols appear in a known or partially-known sequence, the probabilities of the unknown symbols and the hidden

states can be calculated using a Forward-Backward algorithm (Rabiner, 1989; Hickey et al., 2015). It is this property of HMM of being able to predict unknown symbols in the Forward-Backward algorithm that is used in genotype imputation.

The HMM can be described simply by using the "urn and ball model" (Figure 1.8; Rabiner, 1989): if we assume that there are N large urns in a room and within each urn there are a large number of coloured balls of M distinct colours. An initial random urn is chosen, a ball is chosen at random from this urn, and its colour is recorded. The ball is then replaced in this initial urn and a new urn is selected and a new ball is picked from this urn and so on. This process generates a finite observation sequence of colours which can be modelled as the observable output of HMM where the specific urn corresponds to the state and for which the colour probability is defined for each state. The choice of urns is then dictated by the state transition matrix of the HMM.

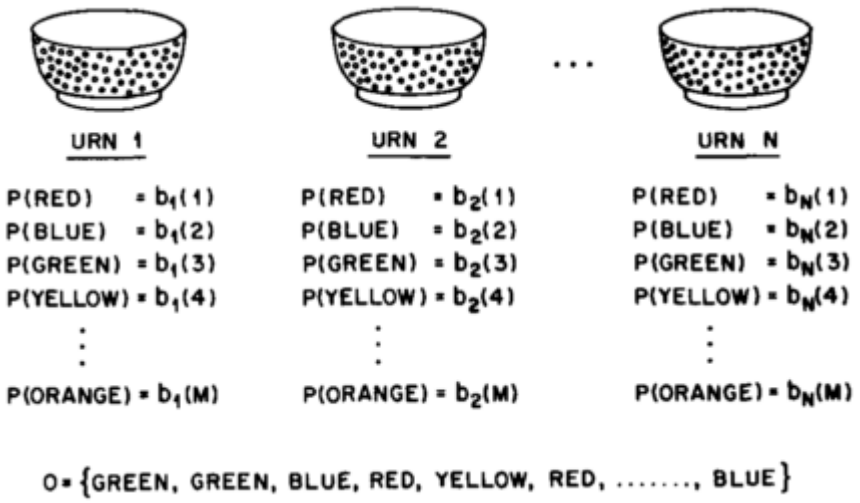


Figure 1.8 An N-state urn and ball model which illustrates the general case of a discrete symbol HMM (Rabiner, 1989).

In a genetic model, the observed marker alleles of an individual and its ancestors are the symbols, the identity of marker alleles by descent, i.e., which grandparent allele carried by the parents was transmitted to the individual, are the states, and recombination events are the transitions among the states (Hickey et al., 2015). When the same parent appears in multiple offspring pedigrees, all of the information from those offspring can be used to calculate an estimate of the phase for each marker in the parent (Nettelblad, 2012; Hickey et al., 2015).

2) Pedigree & population haplotyping

A haplotype is a set of alleles that tend to be inherited together. Thus, haplotyping is widely used in genotype imputation as a way of identifying regions of the genome that may be identical-by-descent. Haplotyping requires a reference set of genotyped individuals such as those available from the International HapMap Project in humans (Consortium, 2003) and from the 1000 Bull Genomes Project in cattle (Hayes et al., 2012).

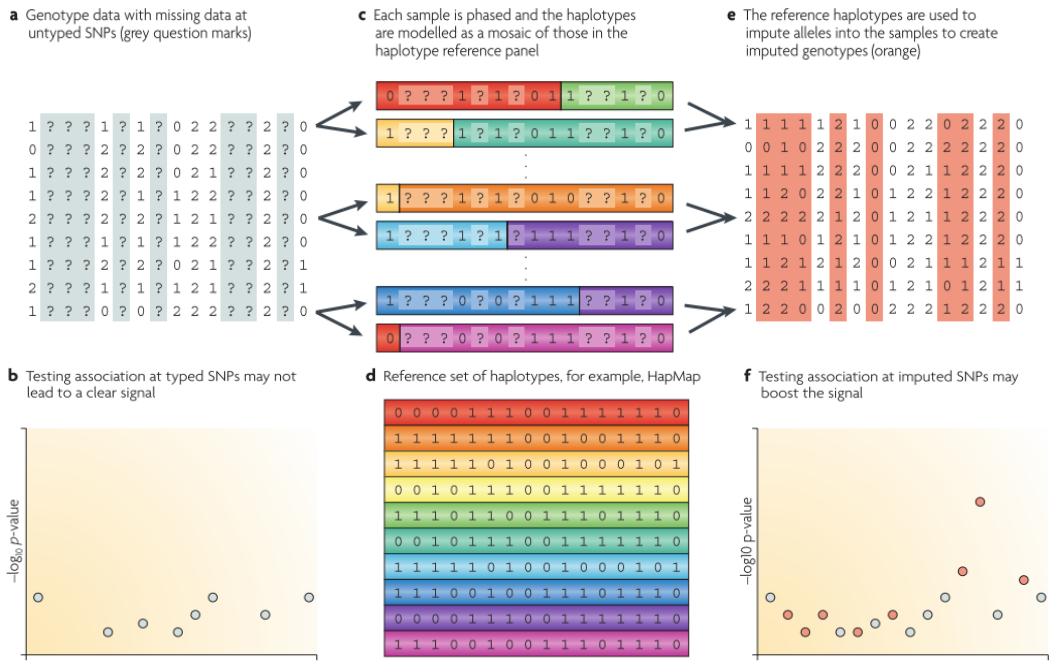


Figure 1.9 Haplotyping for genotype imputation (Marchini and Howie, 2010)

Figure 1.9 describes haplotype imputation in a sample of unrelated individuals (Marchini and Howie, 2010). Each individual has been genotyped, but some information is missing (Figure1.9a); the aim of imputation is to predict these missing genotypes. Firstly, each individual's genotype is phased. The haplotypes identified are then modelled as a mosaic of those in the reference population (Figure 1.9c). The haplotypes from the sample are matched to those in the reference population and the missing genotype information is filled in (Figure1.9e). In real life, the genotypes may be imputed with a level of uncertainty as the information in the sample population may match a number of haplotypes or parts of haplotypes in the reference population due to undergoing recombination so a probability distribution over all three possible genotypes (00, 01, 11) is produced.

In animal breeding, pedigree information stretching back a number of generations is often known. If the ancestors of the individual being haplotyped are known and have previously been genotyped, this can reduce the possible haplotypes of the individual from thousands to just the four haplotypes of the parents plus the potential crossovers from recombination events (VanRaden et al., 2015). Thus, known pedigrees of genotyped animals can increase the accuracy of haplotype imputation by increasing the probability of selecting the correct haplotype from the reference population.

3) Overlapping sliding window

The overlapping sliding window (OSW) approach to imputation exploits shared haplotypes between both close and distant relatives (Sargolzaei et al., 2014). This method assumes that all individuals are related to some degree (Sargolzaei et al., 2014). The missing genotype information is captured by moving "long windows" over a chromosome to identify the long haplotypes that are typically shared among close

relatives. These windows are then shortened by a constant factor after each sweep of the chromosome to capture shorter haplotypes that are typically shared with more distant relatives. For each window a library of haplotypes is built which can then be used for phasing and imputation within that window. Including pedigree information with the respective genotype information leads to more accurate genotype imputation and becomes more important the sparser the genotypes needing imputation are (Sargolzaei et al., 2014).

1.4.3.2 Factors Affecting Accuracy of Imputation

There are four main factors affecting the accuracy of imputation:

1) *The extent of relationship between animals*. Since imputation is largely based on pedigrees and shared haplotypes, the more closely related the animals to be imputed are to the animals in the reference population, the higher the accuracy of imputation will be. Previous studies have revealed that imputation from LD to HD was most accurate (0.97) when the true HD genotypes of all sires and 50% of the dams were included in the reference population (Judge et al., 2017) and that regardless of the depth of the LD panel, greater imputation accuracies were obtained when a larger fraction of the parents true HD genotypes were included in the reference population (Sollero et al., 2019).
2) *Breed composition of the test versus calibration dataset*. The breed composition of the animals to be imputed can greatly affect the accuracy of imputation. A study by Ventura et al. (2014) on the accuracy of imputation from 6,000 to 50,000 SNPs in purebred and crossbred beef cattle determined that the average accuracy of imputation for purebred animals ranged from 90.02 to 98.31% while the average accuracy of imputation for crossbred animals ranged from 54.15 to 97.53%. The accuracy of imputation of the crossbred animals increased when

the breed composition of the animals to be imputed was well represented in the reference population.

3) *Minor allele frequency (MAF).* Imputation of rare alleles is often very difficult as variants with a low MAF are either under selection to remove them from the population (Ventura et al., 2016), are recent mutations (Sargolzaei et al., 2014) and, according to Sargolzaei et al. (2014), are more likely to be identified after detecting long haplotypes, or are simply genotyping errors. The accuracy of imputation of rare alleles (MAF $\leq$ 0.05) varies from 58% in sheep (Ventura et al., 2016) to >80% in dairy cattle (Sargolzaei et al., 2014). It is thought that the difference in accuracy of imputing rare variants between different species is mainly due to the differences in population structure between the dairy and sheep populations. The more closely related the animals in the reference population are to the animals to be imputed, the higher the accuracy of imputation for these rare alleles will be.

4) *Marker Density*. The level of linkage disequilibrium between SNPs generally strengthens with increasing marker density (Hozé et al., 2013). Thus, it is widely agreed that it is easier to impute a genotype from a panel with more markers than it is to impute from a lower density SNP panel. During the development of a low-cost low-density genotyping panel in Ireland (Judge et al., 2016) it was discovered that the imputation accuracy improved at a diminishing rate as the marker density of the panel increased. Also, the variability in mean imputation accuracy per individual reduced as marker density increased (Judge et al., 2016).

1.4.4 Genome Wide Association Studies

Genome wide association studies (GWAS) are used to identify genomic regions associated with a phenotype by using high-throughput genotyping technologies to assay hundreds of thousands (now millions) of genetic markers (Pearson and Manolio, 2008). The primary goal of a GWAS is to identify candidate regions of the genome containing genes and gene-products that affect the phenotype of interest which can help in prioritising genes or genomic regions for further investigation (Stranger et al., 2011); this leads to a better understanding of the biology and the genetic architecture of that phenotype (Visscher et al., 2017). The information gleaned on the genetic architecture of a trait from a GWAS includes the number of potential loci underlying the genetic variation, and therefore the phenotypic variation, the distribution of the allele effect sizes, as well as suggesting whether epistasis or pleiotropy exists at that locus (Stranger et al., 2011).

Genome wide association studies have been commonly used to identify genetic risk factors for diseases in humans (Cantor et al., 2010). The first successful GWAS was of age-related macular degeneration in humans (Klein et al., 2005) and, since then, many other GWAS on human health have been used to develop new treatment and prevention strategies for complex diseases (Bush and Moore, 2012). In animal breeding, there is less emphasis on discovering genes and pathways associated with disease traits but greater emphasis on predicting genetic merit and subsequently phenotype using the results from a genome-based study, although both can be important (Goddard and Hayes, 2009). Traditionally, animal selection has been based on estimated breeding values calculated from a combination of phenotypic records and recoded pedigrees as well as the knowledge of the extent of underlying genetic and environmental variation. However, this process can be slow if the trait can only be measured later in life (e.g.

longevity), after death (e.g. carcass yield, meat quality) or even just in one sex (e.g. milk yield); more records are needed for lowly heritable traits. Therefore, knowledge of genetic variants associated with these traits would be advantageous as this would enable selection of the animals known to be carrying the favourable alleles for these traits in a targeted breeding approach (Goddard and Hayes, 2009).

The most common type of genetic variant studied is the SNP. Millions of SNPs exist in the cattle genome (Hayes et al., 2012), but only a small proportion are likely to affect a given phenotype. Genome wide association studies are based on the principle of linkage disequilibrium (LD) among genetic markers, such as SNPs (Visscher et al., 2012). Linkage disequilibrium is the non-random association of alleles at different loci and occurs due to genetic variation being transmitted from generation to generation in large chunks or haplotypes (Bush and Moore, 2012). Generally, markers that are closer together on a chromosome are in stronger LD than those further apart; therefore, the genomic distance at which LD decays determines how many genetic markers are needed to identify an associated haplotype.

A successful genome wide association study requires three essential elements (Cantor et al., 2010):

1) Sufficiently large study samples from a population that effectively provides genetic information on the trait in question
2) Polymorphic alleles that can be cheaply and efficiently genotyped and that cover the whole genome adequately
3) Statistically powerful analytical methods that can be used to detect the genetic associations

In order to achieve sufficient statistical power in a GWAS, a number of factors must be accounted for (Allen et al., 2010):

1) *Phenotypes must be well defined.* At the centre of many problems with association studies is the issue of phenotype definition (MacRae and Vasan, 2011). In GWAS where precise phenotypes are used, small population sizes have been sufficient to identify alleles of large effect size. In comparison, in studies where the phenotype under investigation is less precise, GWAS have often yielded limited success even with large population sizes (MacRae and Vasan, 2011).
2) *Population size and stratification.* Small study sizes and the substructure of a population, that is not properly accounted for, can often lead to false associations between genotype and phenotype (Allen et al., 2010). Larger populations can also increase the chances of detecting low frequency alleles with a small effect size on the phenotype.
3) *The extent of LD between the marker allele and the suspected causal allele.* The extent of LD between the marker allele and the suspected causal allele can affect statistical power. Weaker LD between these alleles would require more markers (i.e., a higher density genotyping panel) and a larger population size to effectively identify an association between phenotype and genotype.
4) *The effect size of the variant on the phenotype.* Complex traits are likely influenced by many SNPS across the genome, each with a small effect on the phenotype. High density genotypes are often required in these cases to increase the statistical power enough to detect all these associations.

1.4.4.1 Genotype Quality Control

Quality control is carried on all genotypes prior to analysis to ensure the integrity of the data. These quality control measures include:

1) *Animal and SNP call rate.* Prior to starting any genotype analysis, edits are carried out based on both animal and SNP call rates as a low call rate usually indicates poor quality DNA (Wiggans et al., 2010). Generally, animals that are missing >5% of their genotypes are typically removed from the data set (Su et al., 2012); however, research published by Purfield et al. (2016) recommends a minimum threshold animal genotype call rate of 0.85 in cattle. SNPs are usually only retained if they have a genotyping rate >90% (Wiggans et al., 2010; Calus et al., 2011) though the call rate of SNPs included in animal studies has varied from 0.80 to 0.95.
2) *Minor allele frequency.* The minor allele frequency (MAF) simply refers to the frequency at which the least common allele in a population occurs as a percentage of the total called genotypes for that SNP. In general, alleles with a MAF between 1-5% are removed prior to analysis (Hirschhorn and Daly, 2005; Wiggans et al., 2009) as the low frequency makes these alleles difficult to call and they can cause false-positives in association studies (Anderson et al., 2010). However, as population sizes increase, the lower threshold on MAF can be relaxed.
3) *Hardy-Weinberg equilibrium.* Hardy-Weinberg equilibrium (HWE) states that allele and genotype frequencies remain constant from generation to generation in large, randomly mating populations. Given a frequency of q for a biallelic SNP, the expected frequencies of the 3 three possible genotypes (aa, Aa, and AA) are $(1-q)^2$, $2q(1-q)$, and q^2, respectfully, assuming the locus is in HWE. Deviation

from HWE can be a sign of genotype calling error but can also be indicative that the population has undergone selection at that locus (Anderson et al., 2010). In animal breeding, generally all populations undergo selection or migration (first-cross or admixed populations) so deviations from HWE are expected but not necessarily at every locus (Garrick and Fernando, 2013). Thus, low HWE cut off points (at least $p < 10^{-4}$) are often implemented in animal association studies (Bolormaa et al., 2010; Sanchez et al., 2017).

4) *Mendelian inconsistencies.* A Mendelian inconsistency occurs when the genotype of an autosomal SNP in an animal and is in disagreement with what is expected based on the genotype of its parents. For example, the animal might be homozygous for one allele (e.g. AA) at a specific locus but the validated sire might be homozygous for the opposite allele (e.g. GG). Therefore that animal could not have inherited that allele from that sire. Mendelian inconsistencies may be due to simply mixing up DNA samples, due to an error in recorded pedigree or from genotyping errors, or sometimes but rarely, due to mutations (Calus et al., 2011). The genotyping of parent-offspring pairs in animal breeding allows these Mendelian inconsistencies to be identified and removed.

1.4.4.2 Statistical Approaches to Genome Wide Association Studies

There are two broad categories of statistical approaches to undertake a GWAS: a frequentist approach and a Bayesian approach. The main difference between these two approaches is that the frequentist approach focuses on the probability of the data given the hypothesis, whereas the Bayesian approach focuses on the probability of the hypothesis, given the data (Wagenmakers et al., 2008). This means that in the frequentist approach, data are treated as random and the hypothesis is fixed while in the

Bayesian approach the data are treated as fixed and the hypothesis as random (Wagenmakers et al., 2008).

1) Frequentist approaches

Regression approaches to GWAS are often frequentist in nature. There are two types of regression approaches in GWAS, linear and logistic. Linear regression is used when the phenotype being tested is continuous whereas logistic regression is used when the phenotype is a binary trait.

The simple linear regression model used to test for association is:

$$\boldsymbol{y} = \boldsymbol{\mu_m} + \boldsymbol{x_m}\boldsymbol{\beta_m} + \boldsymbol{e_m}$$

where **y** is a vector of an individual's phenotype, $\boldsymbol{\mu_m}$ is the intercept term, $\boldsymbol{x_m}$ is a vector of the individual's genotype at the *m*th SNP, $\boldsymbol{\beta_m}$ is the regression coefficient corresponding to the *m*th SNP, and $\boldsymbol{e_m}$ is the residual (Chen and Witte, 2007). Fitting this equation for each SNP separately gives the maximum-likelihood coefficient estimate for the association between the *m*th SNP and the phenotype. By dividing the regression coefficient by its standard error, the statistical significance of the association between the phenotype and that SNP can be obtained. Ranking these SNPs by their significance identifies regions of the genome to be further investigated.

The main advantage of linear regression models for GWAS is that only one marker is tested at a time. Therefore these models are capable of handling very large numbers of SNPs (Wang et al., 2016). The main disadvantage of using linear regression models in association analyses is the expected large number of false positives (Cantor et al., 2010). The most widely used approach to correct for these false positives is to use a Bonferroni correction (Weisstein, 2004) where each significance value is divided by the

number of (independent) tests performed. Another less conservative correction to reduce the number of false-positives is the Benjamini-Hochberg correction (Benjamini and Hochberg, 1995). This is undertaken by ordering the p-values by size from smallest to largest. The smallest p-value is given a rank of i=1, the second smallest i=2, all the way up to the largest p-value where i=n (the total number of tests). The critical value is then calculated as $(i/n)Q$, where Q is the chosen false discovery rate. The largest p-value that remains significant is the largest value of $p < (i/n)Q$; all the p-values smaller than this are also deemed to be significant. Either of these corrections are not always ideal however, as they assume each association test is independent of all other tests, which is simply not true due to the presence of LD between markers. Using overly conservative corrections can lead to an increase in the number of false negatives identified; therefore, some degree of caution should be exercised when declaring something is significant or even non-significant.

2) Bayesian hierarchical modelling

Bayesian models differ from linear regression analysis in that the Bayesian approach can fit all SNPs simultaneously, whereas the linear regression approach tests each SNP independently (Fernando and Garrick, 2013). Therefore, Bayesian models can handle high density data, where the number of SNPs is greater than the number of observations, to estimate all SNP effects simultaneously. Due to this simultaneous fitting of SNPs, Bayesian approaches are less sensitive to false-positive associations than regression analyses. Bayesian models can also combine prior information about the data and from prior distributions for inference (Fernando and Garrick, 2013). A Markov chain Monte Carlo is often used to sample and draw these inferences from the dataset (Fernando and Garrick, 2013). Prior information may include information from previous studies such as previously identified associations or the number of SNPs

(MacLeod et al., 2016) that are expected to be associated with the phenotype. The prior information from SNPs, such as previous associations or effect sizes, that have been previously identified as associated with a phenotype can be the same for all SNPs or can vary between SNPs depending on the Bayesian approach taken. Information about the SNP such as the MAF, the genomic location, and evidence for selection on the SNP should be taken into account when considering a prior value for each SNP.

1.4.4.3 Advantages & Disadvantages of Genome Wide Association Studies

Genome-wide association studies have revolutionized the field of complex trait genetics over the past decade (Tam et al., 2019). These studies have led to insights into the genetic architecture of complex traits but are not without their limitations and/or controversy (Figure 1.10).

The main advantage of modern-day GWAS is the ability to scan the entire genome for genomic variation associated with a phenotype to identify novel variant-trait associations and underlying biological mechanism(s) (Fan et al., 2010; Tam et al., 2019). Before the advent of GWAS, candidate gene and QTL mapping strategies were used extensively in bovine genetics. In these scenarios only pre-specified genes, where prior biological knowledge existed, were tested for associations (Fan et al., 2010). These techniques were somewhat successful; genetic evaluations including the QTL effect or candidate genes resulted in more accurate selection of the elite animals (Israel and Weller, 1998; Israel and Weller, 2002).

A major disadvantage of GWAS is the ability to miss rare alleles that may be associated with a phenotype. Minor allele frequency edits are normally applied to genotype data prior to analysis which may remove rare alleles that may be of significance to the population. In addition to this, variants identified as associated with a

trait during a GWAS tend to only account for a modest proportion of the estimated heritability of most complex traits (Tam et al., 2019). Several reasons have been proposed for this seemingly missing heritability; some SNPs of modest effect are missed because they do not meet the significance threshold or have been removed during data edits. However, with increasing population sizes being used in GWAS, the statistical power to discover the rarer variants should be increased and (some of) the missing heritability may soon be accounted for (Tam et al., 2019). Large genotyped populations are essential to achieve sufficient statistical power, especially for lowly heritability traits.

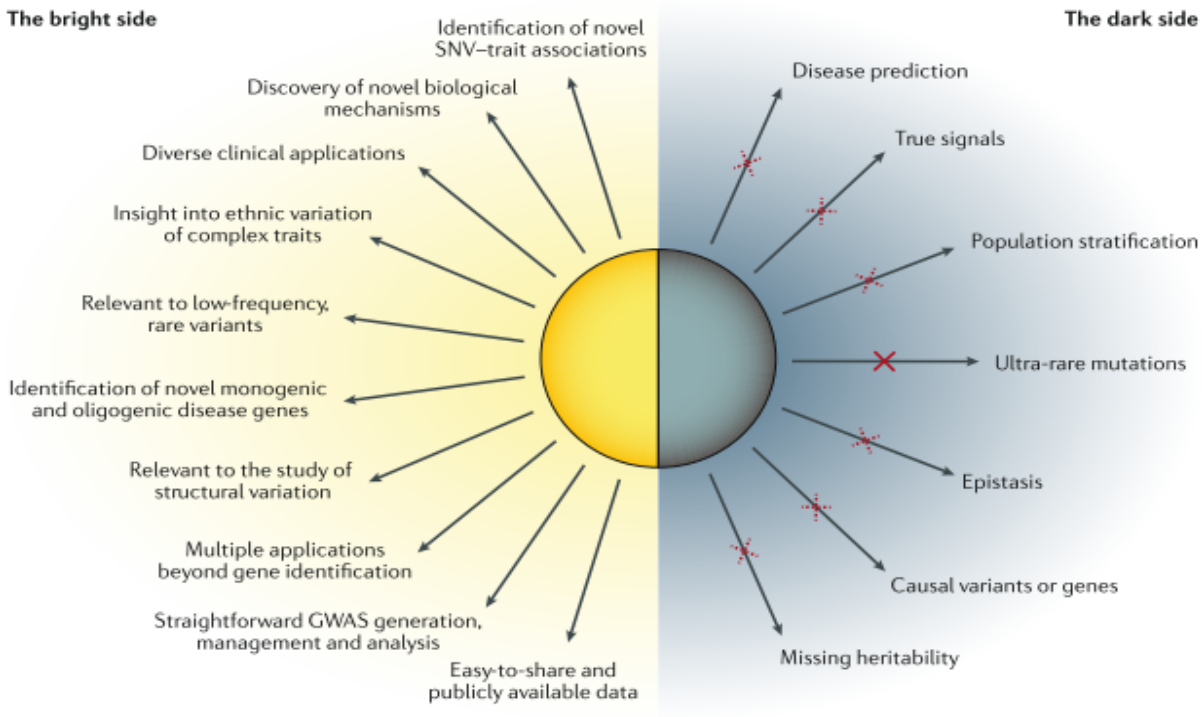


Figure 1.10 Benefits and limitations of genome-wide association studies as observed in human genomics (Tam et al., 2019). A visual depiction of the perceived benefits (the bright side) and limitations (the dark side) of genome-wide association studies. The solid X indicates a permanent limitation while the dotted Xs represent limitations that could potentially be overcome in the future. SNV = single-nucleotide variant.

1.4.5 Pathway Analysis

After the identification of possible candidate genes during a GWAS, pathway analysis can provide further insight into the biology of a complex trait. Pathway analysis of GWAS data sets is a potentially powerful approach to searching for causal genes for complex traits (Jia et al., 2011). However, this assumes a complex trait results from a number of genes which disrupt one or more biological pathways (Jia et al., 2011). Grouping long lists of genes from an association study into smaller sets of related genes or proteins involved in a biological pathway, which may relate to the expression of the phenotype, can reduce the complexity of the analysis (Khatri et al., 2012). There are two main pathway repositories, Kyoto Encyclopaedia of Genes and Genomes (KEGG) and Gene Ontology (GO).

The KEGG pathway database is a collection of pathway maps representing current knowledge on genes, RNA, proteins, chemical compounds, glycans, and chemical reactions, as well as disease genes and drug targets. This database divides the biological processes into 7 groups: metabolism, genetic information processing, environmental information processing, cellular processes, organismal systems, human diseases, and drug development. Over-represented KEGG pathways are identified by comparing lists of candidate genes to a background of all genes in the genome.

The GO database is a collection of ontologies that focus on the function of the gene and gene products by relating them to their biological properties. The GO database divides these biological properties into three groups, biological processes, cellular components, and molecular functions, with each containing a set of GO terms relating to a specific function. Thus, in pathway analysis, the hypothesis involves testing whether the number of genes relating to a GO term is greater than expected by chance (Szkiba et al., 2014).

1.5 Linear Type Traits

Linear type trait scoring in Irish beef cattle was first introduced in the 1990s by the Limousin Society, followed shortly by the Charolais Cattle Society and the Irish Simmental Society. The majority of other beef breeds followed suit in subsequent years (www.icbf.com, date accessed: 20/03/2020). At this time, each breed society employed their own classifiers. In 2002, the ICBF introduced the 'across-breed' linear scoring system which remains today, whereby classifiers score animals from a range of different breeds using a harmonised scale agreed by all beef herdbooks. The linear type traits scored describe the functional, muscular and skeletal characteristics (Table 1.1) of the animal. Historically, linear scores were measured by trained classifiers on weanling beef animals between 150 and 300 days old; since 2016, however, scoring has been undertaken on a whole herd basis to reflect the need to gather more data on the breeding females in the seedstock beef population (www.icbf.com; date accessed 20/03/2020).

In some countries, type traits have been recorded (non-linearly) on dairy cattle since the early 20th century, but, it was only in the 1980s that linear classification was introduced to type evaluations (Short and Lawlor, 1992). In Irish dairy cows, linear assessment is undertaken for 22 traits (Table 1.2) across 4 major body structures; udder, rump, feet & legs, and dairy strength. Scoring of these traits is undertaken by trained classifiers from the Irish Holstein-Friesian Association (IHFA).

As the breeding goals for dairy and beef cattle differ substantially, the majority of type traits are uniquely scored in either the beef or dairy breeds. Some traits such as locomotion, hind-leg rear view, hind-leg side view, and chest depth are scored similarly, but on a different scale, in both dairy and beef. Other traits are common across breeds; for example, stature in dairy cattle is comparable to wither height in beef cattle, while rump width in the dairy cattle is comparable to hip width in beef cattle. While these

traits are similar, the comparison of the genetic parameters and genomic regions associated with these traits is not readily carried out across beef and dairy breeds as research groups do not always have access to data from the different industries. Ireland is in a unique position as the ICBF database contains a wide range of information, including linear type scores, from both the beef and dairy sectors.

1.5.1 Heritability Estimates of Linear Type Traits

Heritability estimates for the functional traits are generally low in beef and dairy cattle, while the heritability estimates for the skeletal and muscular traits in both beef and dairy cattle are generally moderate to high.

1.5.1.1 Functional Traits

Functional traits scored include locomotion and those describing the feet and legs such as: foreleg front view, hind-leg side view, and hind leg rear view. Previously published estimates of heritability of locomotion have all been low ranging from 0.02 in Charolais beef cattle (Vallee et al., 2015) to 0.14 in Holstein-Friesian dairy cattle (Berry et al., 2004). A previous study undertaken by McHugh et al. (2012) in a multi-breed Irish beef population reported the heritability of locomotion to be 0.13.

McHugh et al. (2012) previously reported a heritability of 0.05 for foreleg front view in a multi-breed beef population. No other heritability estimates exist for this trait in the literature as it is only scored in Ireland. Other heritability estimates available for functional traits in beef breeds of cattle include an estimate for heritability of forelegs in Piedmontese cattle (0.09; Mantovani et al., 2010) and for front legs in Charolais (0.07; Vallee et al., 2015). Wiggins et al. (2004) reported estimates of 0.15, 0.18, 0.16, 0.07, and 0.11 for the heritability of hind-leg side view in Ayrshire, Brown Swiss, Guernseys,

Jerseys, and milking Shorthorns, respectively. Previous heritability estimates for hind-leg side view in Holstein-Friesians range from 0.18 (Veerkamp and Brotherstone, 1997) to 0.20 (Royal et al., 2002). McHugh et al. (2012) estimated the heritability of hind-leg rear view to be 0.11 in a multi-breed beef population.

1.5.1.2 Muscular Traits

There are six main muscular related type traits scored in the Irish beef herd: development of the hind quarter, the loin, and the inner thigh, and the width of the thighs, the withers, and behind the withers (Figure 1.11). Muscular traits are not generally measured in dairy cattle. Few heritability estimates of muscular traits have actually been published for these traits in beef cattle. The heritability of development of inner thigh has previously been estimated to be 0.22 in Asturiana de los Valles cattle (Gutierrez and Goyache, 2002). While heritability estimates for thigh width are non-existent in the literature, heritability estimates for two comparable traits have previously been documented; the heritability of thigh thickness in Piedmontese cattle has previously been estimated as 0.15 (Mantovani et al., 2010) and the heritability of thigh rear view, which also measures thigh thickness, in Rendena Dual-Purpose cattle has previously been estimated as 0.32 (Mazza et al., 2014). Overall muscularity at weaning and at 15months in the LM have previously been estimated as 0.35 and 0.51, respectively (Bouquet et al., 2010).

1.5.1.3 Skeletal Traits

Similar to the muscular traits, there are six main skeletal type traits scored in the Irish beef herd: chest width, chest depth, wither height, pelvic length, back length, and hip width (Figure 1.12). There are also six skeletal traits scored in the Irish dairy herd; stature, chest width, body depth, rump angle, rump width, and angularity. Heritability estimates for skeletal traits in beef cattle and their comparable traits in dairy cattle have

been well documented. Vesela et al. (2005) estimated the heritability for chest width to be 0.27 in Czech Republic beef cattle. In Holstein-Friesian dairy cattle, the heritability for this trait has been estimated to be between 0.25 and 0.29 (Veerkamp and Brotherstone, 1997; Royal et al., 2002; Berry et al., 2004). Chest depth is typically only measured in beef cattle; the heritability of this trait has been estimated to be 0.30 in Rendena Dual-Purpose cattle (Mazza et al., 2014) and 0.32 in beef cattle in the Czech Republic (Vesela et al., 2005).

Heritability estimates for wither height, or stature as it is called in dairy animals, is the most well documented heritability for a linear type trait in the literature. Wither height is generally thought to be the most heritable type trait with estimates of heritability ranging from 0.31 in Piedmontese cattle (Mantovani et al., 2010) to 0.54 in Ayrshire cows (Wiggans et al., 2004). Heritability estimates for stature in Holstein-Friesian dairy cows vary from 0.43 (Pérez-Cabal and Alenda, 2002; Berry et al., 2004) to 0.50 (Veerkamp and Brotherstone, 1997).

The heritability of pelvic length was previously estimated to be 0.31 based on an evaluation of Czech Republic beef cattle (Vesela et al., 2005). While back length is not widely assessed outside of Ireland, body length as assessed in other populations may be comparable. Both Mazza et al. (2014) and Vesela et al. (2005) reported heritability estimates for body length in Rendena Dual-Purpose cattle and Czech Republic beef cattle as 0.55 and 0.25, respectively.

Table 1.1 Scale of measurement and description of the linear type traits scored on beef cattle by the Irish Cattle Breeding Federation

Trait	Scale	Description
Functional	*1 - 10*	
Locomotion	low - high	The level of correctness when a weanling walks
Foreleg front view	toes out - toes in	The angle of the forelegs when viewed from the front
Hind-leg side view	straight - sickled	The angle of the hind legs when viewed from the side
Hind-leg rear view	toes out - toes in	The angle of the hind legs when viewed from behind
Muscular	*1 - 15*	
Development of hind quarter	low - high	The level of roundness and fill of the rear point above the rear legs when viewed from the side
Development of loin	low - high	The overall width, length, and fill of the loin
Thigh width	narrow - wide	The overall width of the hindquarters when viewed from behind
Development of inner thigh	low - high	The level of fill/development of muscle between the back legs
Width of withers	narrow - wide	The width of the animal at the highest point above the front legs
Width behind withers	narrow - wide	The width of the animal behind the highest point above the front legs
Skeletal	*1 - 10*	
Chest width	narrow - wide	The width of the animal between the front shoulders
Chest depth	shallow - deep	The vertical distance from the withers to the bottom of the chest, behind the front legs
Wither height	small- tall	The height of the animal at the highest point above the front legs
Pelvic length	short - long	The distance between the hip bone and the rear of the animal
Back length	short - long	The distance between the withers and the hip bone
Hip width	narrow - wide	The width between the hip bones
Other	*1 - 10*	
Body condition score	lean - fat	The overall fleshiness of an animal
Docility	aggressive - docile	How the cattle behaves when being handled by humans

Table 1.2 Scale of measurement and description of the linear type traits scored on dairy cattle by the Irish Holstein-Friesian Association

Trait	Scale (1 - 9)	Description
Stature	130cm - 154cm	The height of the animal at the rump
Chest width	narrow - wide	The width of the animal between the top of the front legs
Body depth	shallow - deep	The vertical distance from the top of the spine to the bottom of the barrel at the last rib
Rump angle	high -wide	The angle of the rump structure from hooks to pins
Rump width	narrow - wide	The width between the most posterior point of the pin bones
Angularity	thick - open ribbed & angular	The angle and openness of the ribs, combined with flatness of bone avoiding coarseness
Dairy strength	frail & heavy - strong &dairy	The overall appearance of the animal
Rear legs (side view)	straight - sickled	The angle measured at the front of the hock
Rear legs (rear view)	toes out - straight	The angle of the hind legs when viewed from behind
Foot angle	shallow - steep	The angle at the front of the right rear hoof measured from the floor to the hairline of the hoof
Locomotion	lame - excellent	The level of correctness when an animal walks
Bone Quality	thick - sharp	The quality of bones in the animals legs
Fore udder	loose - tight	The attachment of the fore udder to the barrel
Rear udder height	low - high	The distance between the bottom of the vulva and the milk secreting tissue in relation to the height of the animal
Rear udder width	narrow - wide	The distance between the the rear udders
Central ligament	convex - deep definition	The depth of cleft, measured at the base of the rear udder
Udder depth	below hocks - 22cm above hocks	The depth of the udder relative to the animals hocks
Front teat placement	on outside of udder - on inside of udder	The positions of the front teat from centre of quarter
Teat placement (side view)	close - wide	The positions of the front teat when viewed from the side of the animal
Teat length	short - long	The length of the teat scored on front left hand side
Rear teat placement	on outside - back teats crossed	The positions of the back teat from centre of quarter
Udder texture	fleshy - like silk	The fineness of the skin on the udder

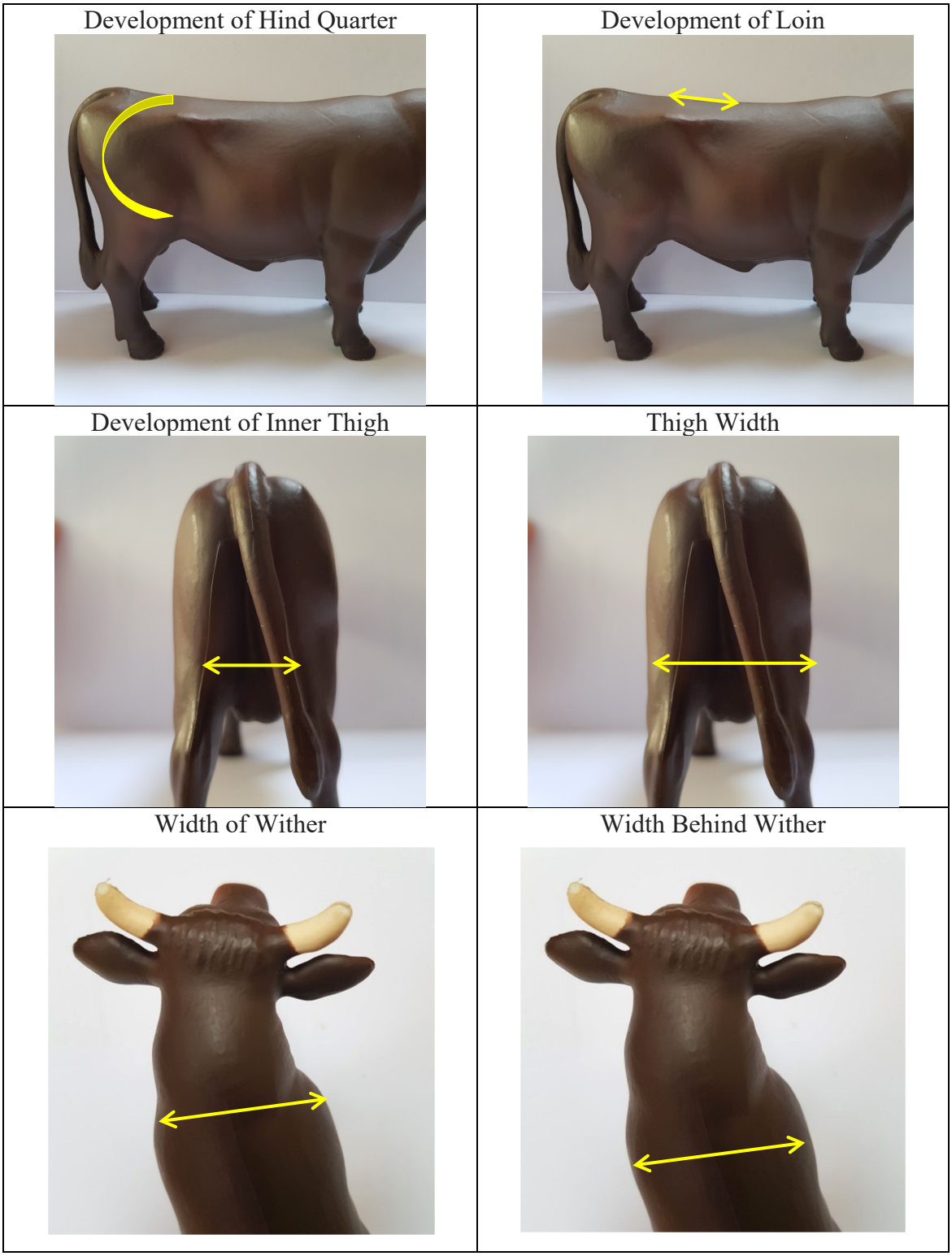


Figure 1.11 Visual descriptions of the muscular linear type traits in beef cattle

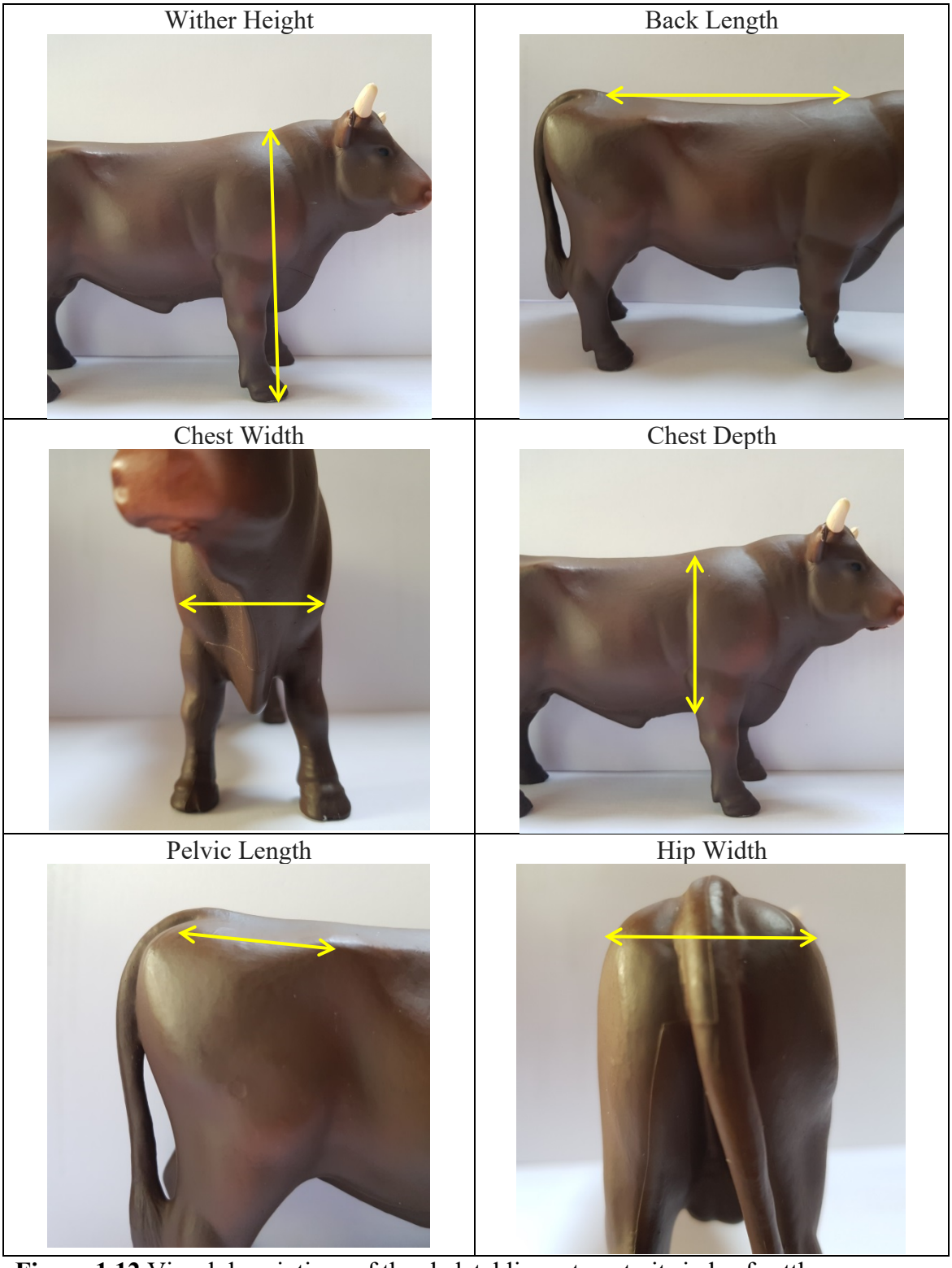


Figure 1.12 Visual descriptions of the skeletal linear type traits in beef cattle

1.5.2 Phenotypic & Genetic Correlations among Type Traits

The majority of studies that report phenotypic and genetic correlations among linear type traits have focused on dairy cattle. It is well documented that the genetic correlations are typically stronger than the respective phenotypic correlations and in the same direction (Vesela et al., 2005; Mc Hugh et al., 2012). It is also well documented that the strongest genetic correlations among type traits are generally among the muscular and skeletal traits, while the weakest genetic correlations are among the functional traits (Gutierrez and Goyache, 2002; Mc Hugh et al., 2012). Strong correlations among skeletal traits (0.56 – 0.96) and among muscular traits (0.92 – 0.98) have been reported by Vesela et al. (2005) in Czech Republic beef cattle. Strong correlations were also observed between muscular and skeletal traits (0.74 – 0.88) in Asturiana de Los Valles beef cattle (Gutierrez and Goyache, 2002).

1.5.3 Genome Wide Association Studies of Linear Type Traits

Genomic regions underlying the linear type traits in beef or dairy cattle are not well documented in the literature; many previous GWAS on this topic have solely focused on stature (Pryce et al., 2011; Nishimura et al., 2012; Bolormaa et al., 2014; Bouwman et al., 2018) or general muscling in cattle (Saatchi et al., 2014b; Vallée et al., 2016). It is plausible however, that due to the genetic correlations among the type traits, loci documented to be associated with stature or muscle would also be associated with at least some of the other linear type traits.

The majority of previous GWAS studies in cattle have typically focused on either one specific breed of cattle or just dairy cattle or just beef cattle. Few studies have examined the differences and similarities in the underlying genetic architecture of traits across breeds (Saatchi et al., 2014a; Saatchi et al., 2014b; Bouwman et al., 2018). Allele

effects and frequencies have previously been found to vary among breeds (Spangler and Van Eenennaam, 2010). Thus, the knowledge of these similarities and differences is important to the implementation of accurate across breed genomic selection (Spangler and Van Eenennaam, 2010).

1.5.3.1 Skeletal Traits

While wither height, or stature, is an easy to measure phenotypic trait, it is a genetically complex trait which is known to be affected by multiple genes (Gudbjartsson et al., 2008; Bouwman et al., 2018). Many genome wide association studies have been carried out on stature in countless different species including humans (Gudbjartsson et al., 2008), cattle (Pryce et al., 2011; Bouwman et al., 2018), horses (Tetens et al., 2013), and dogs (Hayward et al., 2016). In a GWAS of human stature, 27 regions of the genome were identified that were significantly associated with stature in Caucasians (Gudbjartsson et al., 2008). A GWAS of cattle stature identified 163 variants, 160 SNP and 3 indels, associated with stature across 17 populations of cattle (Bouwman et al., 2018). Many candidate regions identified in different species tend to overlap with each other including those identified in cattle (Pryce et al., 2011; Bouwman et al., 2018), humans (Gudbjartsson et al., 2008), and horses (Tetens et al., 2013). The most promising of these regions is located on chromosome 6 and includes LCORL (ligand dependent nuclear receptor corepressor like) and NCAPG (non-SMC condensin I complex subunit G). While the NCAPG-LCORL locus has been identified in multiple studies in multiple breeds as a locus associated with height, there is no information linking the function of either of these genes to the actual height of an animal or human. LCORL is a transcription factor that may function in the testes during spermatogenesis, while NCAPG is a regulatory subunit of the mammalian condensing I complex that is important during mitotic cell division (Takasuga, 2016).

Another locus on bovine chromosome 14 which has been associated with height in cattle (Nishimura et al., 2012) contains the PLAG1 (Pleomorphic adenoma gene 1) gene. This gene has also been associated with height in humans (Gudbjartsson et al., 2008; Lettre et al., 2008) although it is located on chromosome 8 in the human genome. PLAG1 encodes a zinc finger protein with 2 putative nuclear localization signals which is thought to be crucial for the formation of pleomorphic adenomas of the salivary glands (Takasuga, 2016). A previous study conducted on genetically modified PLAG1 null mice (Hensen et al., 2004) revealed that the full disruption of PLAG1 resulted in delayed development and ultimately, growth retardation, supporting the theory that PLAG1 may have a role to play in the height of other mammals.

1.5.3.2 Muscular Traits

It has been known for many years that the myostatin (MSTN) gene, also known as growth and differentiation factor 8 (GDF8), actively represses muscle growth (Grobet et al., 1997). Thus, it is no surprise that MSTN has previously been linked to muscularity in cattle (Esmailizadeh et al., 2008; Alexander et al., 2009; O'Rourke et al., 2012). There are several mutations in the MSTN gene that are associated with different levels of change in muscle morphology (Saatchi et al., 2014b). Most research into this gene has focused on the double muscling (muscle hypertrophy) phenotype found in certain breeds of cattle such as Belgian Blues and Piedmontese. Despite what the name suggests, double muscled animals do not have more muscles than a normal animal; they do however, have an increase in the number of muscle fibres and enlargement of these fibres (Bellinge et al., 2005). Double muscled animals also have less fat and less bone than that of a “normal phenotype” animal and often generate more revenue when slaughtered. MSTN is also known to be associated with growth traits, birth weight,

calving ease, marbling, rib eye area, and weaning weight in Limousin cattle (Saatchi et al., 2014b).

1.6 Sexual Dimorphism

Sexual dimorphism is the phenomenon whereby males and females of the same species are distinctive in size or appearance (Berns, 2013). In the strictest sense of the word, 'dimorphism' typically only refers to morphology; however, the term sexual dimorphism is used to include all aspects of differentiation of males and females (Fairbairn and Roff, 2006). This differentiation of males and females is widely observed throughout the animal kingdom for a plethora of traits, including coloration, vocalisation, ornamentation, foraging and mating behaviours, and of course, body size (McPherson and Chenoweth, 2012; Berns, 2013). In some species, males and females can be unrecognisable as the same species due to the phenotypic differences between them; however, the genomes of the different sexes are close to identical (Fairbairn and Roff, 2006).

Sexual dimorphism is not solely due to the differences in sex chromosomes between males and females as extreme sexual dimorphism occurs in many animals where sex is determined by environmental causes or in response to age or body size changes (Fairbairn and Roff, 2006). Thus, sexual dimorphism is thought to be attributable to a combination of sex-specific genes on sex chromosomes, sex-specific expression of genes present in both sexes, and other regulatory mechanisms that are not yet widely understood (Pointer et al., 2013). It is commonly thought that this phenomenon historically occurred in mammals due to evolution by natural selection, specifically sexual selection, which arose due to the competition among the same sex of

a species for mating rights and due to mating preferences of one sex to the other (Kirkpatrick, 1987; Katz, 2008; McPherson and Chenoweth, 2012).

In domesticated animals, the sexual selection that occurs in wild animals does not typically happen as those animals selected for breeding are generally selected on numerous economically important traits not just for their aggression and/or size as would happen in the wild. Nevertheless, sexual dimorphism is evident in beef cattle for traits such as growth rate (Koch and Clark, 1955; Marlowe and Gaines, 1958), and birth weight, weaning weight, and post-weaning gain (van der Heide et al., 2016).

1.6.1 Sexual Size Dimorphism

Sexual size dimorphism is a frequent phenomenon whereby the size of males and females of the same species differ (Berns, 2013). When this phenomenon occurs in closely related species, such as cattle, buffalo, and yaks, it can result in distinct patterns of among-species size dimorphism (Polák and Frynta, 2010; Berns, 2013). One of these patterns is known as Rensch's rule which claims that the slope of the allometric relationship between male and female body size is greater than one (Figure 1.13; (Rensch, 1959); i.e., the degree of sexual dimorphism increases with body size in species where males are the larger sex and decreases in species where females are the larger sex (Polák and Frynta, 2010; Berns, 2013).

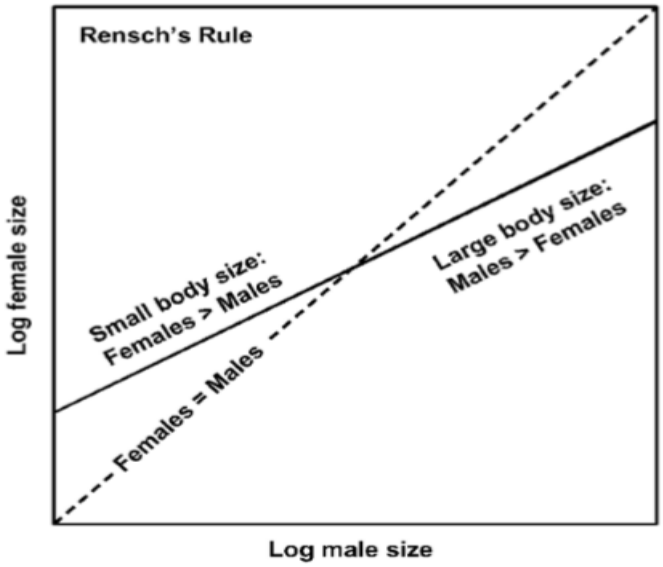


Figure 1.13 Rensch's Rule can be visualized by plotting the log of female body size against the log of male body size in different species. In species above the broken line the females are larger than the males and in species below the broken line the males are larger than the females (Berns, 2013)

There have been a number of hypotheses proposed to explain Rensch's rule (Berns, 2013):

1) The combination of genetic correlations between male and female size with directional sexual selection for male size leads to the evolution of larger males relative to females
2) Sexual size dimorphism evolved through intraspecific competition between the sexes when foraging is related to size
3) Sexual size dimorphism may have evolved due to larger females having a higher chance of reproducing effectively and having larger eggs/offspring

Examples of Rensch's rule and support for all three of these hypotheses are abundant in nature in organisms from hummingbirds (Colwell, 2000) to turtles (Berry and Shine, 1980) to salmon (Young, 2005) and shorebirds (Székely et al., 2004). Previous research into Rensch's rule into domestic cattle breeds has determined that,

despite the evolutionary changes in morphology and size associated with domestication of cattle, Rensch's rule is still adhered to for body mass ratio although no clear relationship was found when other size traits were analysed (Polák and Frynta, 2010).

1.7 Gaps in Knowledge

The gaps in knowledge that will be examined by this thesis include:

- Whether the genetic parameters for functional, skeletal, and muscular linear type traits differ across cattle breeds
- The size and location of SNP effects and the identification of possible candidate genes for muscular and skeletal traits in both dairy and beef cattle, and whether these are common across breeds and across traits
- Whether the genetic parameters for linear type traits in bulls and heifers differ from each other but also if the effects of the underlying SNPs differ by sex, a phenomenon known as sexual dimorphism

This knowledge will help inform breeding programmes of the importance, or lack thereof, of considering a trait in different breeds, or indeed different sexes, to be genetically different traits in order to improve the accuracy of evaluations.

Chapter 2

Genetic co-variance components within and among linear type traits differ among contrasting beef cattle breeds

2.1 Preface

At the time of thesis submission this chapter was published in the Journal of Animal Science (Accepted April 10, 2018; doi: 10.1093/jas/sky076). The full reference is Doyle JL, Berry DP, Walsh SW, Veerkamp RF, Evans RD, Carthy TR: Genetic co-variance components within and among linear type traits differ among contrasting beef cattle breeds. Journal of Animal Science 2018, 96(5):1628-1639.

Jennifer Doyle was primary author, performed the data edits and analysis and drafted the manuscript. Donagh Berry, Siobhan Walsh, Tara Carthy and Roel Veerkamp conceived the study, participated in the design and co-ordination of this study and helped draft the manuscript. Ross Evans supplied the data the analysis was performed on.

Formatting and referencing style has been edited for consistency throughout the thesis/ Figure and table captions have been assigned with a chapter prefix. Competing interests and acknowledgements have been removed. All other aspects are consistent with the published manuscript.

2.2 Abstract

Linear type traits describing the skeletal, muscular and functional characteristics of an animal are routinely scored on live animals in both the dairy and beef cattle industries. Previous studies have demonstrated that genetic parameters for certain performance traits may differ between breeds; no study, however, has attempted to determine if differences exist in genetic parameters of linear type traits among breeds or sexes. Therefore, the objective of the present study was to determine if genetic co-variance components for linear type traits differed among five contrasting cattle breeds, and to also investigate if these components differed by sex. A total of 18 linear type traits scored on 3,356 Angus (AA), 31,049 Charolais (CH), 3,004 Hereford (HE), 35,159 Limousin (LM), and 8,632 Simmental (SI) were used in the analysis. Data were analyzed using animal linear mixed models which included the fixed effects of sex of the animal (except in the investigation into the presence of sexual dimorphism), age at scoring, parity of the dam, and contemporary group of herd-date of scoring. Differences ($p < 0.05$) in heritability estimates, between at least two breeds, existed for 13 out of 18 linear type traits. Differences ($p < 0.05$) also existed between the pairwise within-breed genetic correlations among the linear type traits. Overall, the linear type traits in the continental breeds (i.e. CH, LM, SI) tended to have similar heritability estimates to each other as well as similar genetic correlations among the same pairwise traits, as did the traits in the British breeds (i.e. AA, HE). The correlation between a linear function of breeding values computed conditional on co-variance parameters estimated from the CH breed with a linear function of breeding values computed conditional on co-variance parameters estimated from the other breeds was estimated. Replacing the genetic (co)variance components estimated in the CH breed with those of the LM had least effect but the impact was considerable when the genetic (co)variance components of the

AA were used. Genetic correlations between the same linear type traits in the two sexes were all close to unity (≥ 0.90) suggesting little advantage in considering these as separate traits for males and females. Results for the present study indicate the potential increase in accuracy of estimated breeding value prediction from considering, at least, the British breed traits separate to continental breed traits.

2.3 Introduction

Linear type traits describing skeletal, muscular and functional characteristics of the animal are routinely scored globally in both dairy (Veerkamp and Brotherstone, 1997; Berry et al., 2004; Kern et al., 2015) and beef (Mc Hugh et al., 2012; Mazza et al., 2014) cattle. While genetic parameters of type traits have been extensively researched in Holstein-Friesian dairy cattle (VanRaden et al., 1990; Veerkamp and Brotherstone, 1997; Kern et al., 2015), fewer studies have been undertaken in beef cattle. Nonetheless, type traits are often included in multi-trait genetic evaluations as predictors of performance in both dairy (VanRaden et al., 1990; Berry et al., 2004) and beef (Gutierrez and Goyache, 2002; Mc Hugh et al., 2012) cattle.

The majority of previous studies have considered type traits in both males and females as being genetically the same trait. It is possible, however, that the genetic control of such traits may be sex-dependent (van der Heide et al., 2016). If sexual dimorphism exists for type traits, then these traits may need to be considered as genetically different traits in genetic evaluations. Genetic parameters for type traits may also differ by breed, similar to what has been previously demonstrated for other performance traits in cattle (Utrera and Van Vleck, 2004; Hickey et al., 2007). Knowledge of possible differences in genetic parameters among breeds is of increasing

importance as some populations move towards using a multi-breed, multi-trait statistical model in the pursuit of greater precision of genetic evaluations. The objective, therefore, of the present study was to determine if genetic co-variance components for linear type traits differed among five contrasting cattle breeds and also if these traits differed genetically by sex. The results from the present study will be useful in informing breeding programmes of the importance, or lack thereof, of considering a trait in different sexes or breeds to be genetically different traits.

2.4 Materials & Methods

2.4.1 Linear Type Trait Data

As part of the Irish national beef breeding program, routine scoring of linear type traits is carried out on both registered and commercial beef herds by trained classifiers (Mc Hugh et al., 2012; Berry and Evans, 2014); each classifier scores animals from a range of different breeds and crossbreeds. A total of 18 linear type traits assessed across all breeds were retained for analysis in the present study. Traits analysed represented muscular (n=6), skeletal (n=6) and functional (n=4) characteristics of the animal, as well as docility and body condition score.

Linear type trait data were available on 248,181 animals. Animals were discarded if the sire, herd, or classifier were unknown or the parity of the dam was not recorded; 230,109 records remained. Parity of the dam was stratified into 1, 2, 3, 4, and ≥5. Only animals scored between 6 and 16 months between the years 2000 and 2016 were retained; 179,921 records remained. Only animals that were deemed to be ≥87.5% Angus, Charolais, Hereford, Limousin or Simmental based on the available pedigree information were retained; 140,936 records remained. Only animals from sires with at least five progeny in the data set were retained. Furthermore, only data from classifiers

that scored ≥500 animals since the year 2000 were kept. Contemporary group was defined as herd-by-scoring date. Each contemporary group had to have at least five records and all records within contemporary group were from a single breed. Each trait was separately standardized to a common variance within classifier-by-year as described in detail by Brotherstone (1994). Following all edits, data were available on 81,200 animals in 1,811 herds all scored by 20 classifiers; 3,356 Angus (AA), 31,049 Charolais (CH), 3,004 Hereford (HE), 35,159 Limousin (LM) and 8,632 Simmental (SI).

2.4.2 Analysis

Co-variance components for each trait in each breed were estimated using linear animal mixed models in ASREML (Gilmour et al., 2009). Preliminary analyses were undertaken to detect any dam permanent environmental effect or genetic contribution of the dam to the linear scores, but neither improved the fit to the data and so were not considered further in the mixed model. The following model was used in all analyses:

$$y_{jklm} = HSD_i + Sex_j + AM_k + DP_l + Animal_m + e_{ijklm}$$

where y_{ijklm} is the linear type trait, HSD_i is the fixed effect of herd-by-scoring date (i=8,844 levels), Sex_j is the fixed effect of the sex of the animal (j=male or female), AM_k is the fixed effect of the age in months of the animal at scoring (k=11 classes from 6 to 16 months), DP_l is the fixed effect of the parity of the dam (l = 1, 2, 3, 4, or ≥5), $animal_m$ is the random additive genetic effect of animal m where $a \sim N(0, \mathbf{A}\sigma_a^2)$ with with **A** representing the additive genetic relationship matrix and σ_a^2 representing the additive genetic variance; and e_{ijklm} is the random residual effect, where $e \sim N(0, \mathbf{I}\sigma_e^2)$, **I**

is the identity matrix and σ_e^2 represents the residual variance. Box's M (Box, 1949) was then used to test the homogeneity of the co-variance matrices among the breeds.

In a separate series of analyses, the CH and LM datasets (i.e., the largest datasets) were separately stratified by sex. Further edits were carried out to ensure each sex-specific contemporary group still had >5 animals. Of the remaining 29,542 CH animals, there were 14,253 females and 15,288 males; of the remaining 34,071 LM animals, there were 16,634 females and 17,437 males. Univariate and bivariate analyses were conducted in ASREML using the previously described model without the fixed effect of the sex of the animal.

Likelihood ratio tests were used to evaluate whether sexual dimorphism existed. The log-likelihood value from the original unconstrained bivariate model was compared to that from a constrained model where either the genetic variance in both sexes were constrained to be identical or the genetic correlation between the sexes was constrained to be 0.99.

2.4.3 Eigenstructures

Eigenstructures were calculated to determine if the (co)variance structures among traits within a trait category (i.e., skeletal, muscular, or functional) differed by breed. (Co)variance components estimated from the bivariate analyses were arranged into a multi-trait (co)variance matrix within the skeletal, muscular and functional traits separately. Any non-positive definite (co)variance matrix were bended. Eigenvectors and eigenvalues were calculated using the (co)variance matrices in the individual breeds for the muscular traits, the skeletal traits, and the functional traits separately.

Differences in the (co)variance structures among traits were evaluated as:

$$\boldsymbol{E'_{CH} CO_i E_{CH} = D}$$

where $\mathbf{E_{CH}}$ is a matrix consisting of the eigenvectors in CH, $\mathbf{CO_i}$ is the estimated (co)variance matrix among traits in breed$_i$ and **D** is the resulting matrix. **D** was then rescaled to $\boldsymbol{\tilde{D}}$, a matrix with diagonal elements of 1. Whether the off-diagonals of the $\boldsymbol{\tilde{D}}$ matrix were different from zero was investigated when the $\mathbf{CO_i}$ was used; the closer to zero the off-diagonal elements were i.e. the lower the standard deviation, the more similar the co-variance matrices were to the CH. Further analysis was conducted using AA as the reference breed in place of CH to determine the differences in the (co)variance structures between AA and HE.

2.4.5 Impact of Incorrect (Co)variance Parameters on the Estimation of Breeding Value

Calculations were undertaken to quantify the impact of using the (co)variance components of a given breed to estimate the breeding values for an unmeasured trait in another breed. For illustrative purposes, wither height was assumed to represent the trait where estimated breeding values were desired but no estimated breeding values were assumed available for this trait; the CH was used as the reference breed for comparison purposes. Five linear type traits, namely chest width, hind-leg rear view, body condition score, development of loin and development of inner thigh were chosen as predictor traits. These traits were chosen based on a function of both the strength of their genetic correlation with height at withers (favouring the stronger correlation) and the variability in the correlation across breeds, taking cognisance of the genetic correlation between that trait and the index traits already included in the index.

The efficiency of the index (E_u) was calculated as outlined by Ochsner et al. (2017) as:

$$\boldsymbol{E_u = \frac{b'_u G_{12_t}}{\sqrt{b'_u G_{11_t} b_u}} \left(\sqrt{b'_t G_{12_t}}\right)^{-1}}$$

where $\boldsymbol{G_{12_t}}$ represents the true genetic co-variances between height at withers (i.e, goal traits) in CH and the 5 predictor traits, $\boldsymbol{G_{11_t}}$ is a *5* × *5* matrix representing the true genetic co-variances among the *5* predictor traits in CH, $\mathbf{b_t}$ is a *n* × 1 vector of the coefficients applied to the estimated breeding values derived as:

$$\boldsymbol{b_t = G_{11_t}^{-1} G_{12_t}}$$

and, $\mathbf{b_u}$ is a *n* × 1 vector of the coefficients estimated as above but by replacing genetic (co)variances from the CH breed (i.e., the "true" parameters) with those of the breed under investigation.

2.5 Results

2.5.1 Variance Components of the Linear Type Traits by Breed

The within breed heritability estimates for the linear type traits (Table 2.1 and Table 2.2) ranged from 0.00 (three of the four functional traits in HE) to 0.43 (height in CH). Heritability estimates for the functional traits were generally the lowest of all the traits, and were all ≤ 0.13 (standard error (SE) ≤ 0.04). Heritability for the muscular traits varied from 0.10 (SE=0.04) for development of loin in HE to 0.30 (SE = 0.02) for development of hind quarter in CH. Heritability for the skeletal traits ranged from 0.00 for both chest width and hip width in HE to 0.43 (SE = 0.02) for wither height in the CH.

The CH animals generally had the highest heritability estimates for the linear type traits describing the size of the animal; wither height (0.43; SE = 0.02), back length (0.30; SE = 0.02), development of hind quarter (0.30; SE = 0.02), development of inner thigh (0.28; SE = 0.02) and body condition score (0.13; SE = 0.02). For 13 of the 18 linear type traits, heritability estimates differed ($P < 0.05$) between at least two breeds. Heritability estimates for width of withers, width behind withers, chest depth, pelvic length, and hind-leg side view did not differ between breeds. The genetic standard deviation of the linear type traits differed greatly between the breeds with no genetic variation in six of the traits (i.e., locomotion, foreleg front view, hind-leg rear view, chest width, hip width, and body condition score) being detected in HE.

2.5.2 Within Breed Phenotypic & Genetic Correlations among the Linear Type Traits

Irrespective of breed, the strongest positive phenotypic correlation existed between width of withers and width behind withers, ranging from 0.81 (SE = 0.01) in SI to 0.87 (SE = 0.01) in CH (Table 2.4). The strongest negative phenotypic correlations generally existed among the functional traits or between the functional and muscular traits; hind-leg side view and locomotion in CH (-0.57; SE = 0.01), hind-leg rear view and locomotion in LM (-0.11; SE = 0.01; Table 2.6), hind-leg side view and development of loin in AA (-0.38; SE = 0.02; Table 2.3), hind-leg side view and development of inner thigh in HE (-0.16; SE = 0.02; Table 2.5).

In general, the pair-wise genetic correlations among traits were stronger than their respective phenotypic correlations but of the same sign. The genetic correlations among the muscular traits and among the skeletal traits were typically stronger in the continental breeds (CH, LM, SI) than in the British breeds (AA, HE). Within breed,

genetic correlations among the muscular traits were moderate to strong, varying from 0.58 (SE = 0.15) for development of loin and width of withers in HE (Table 2.5) to 0.99 (SE = 0.01) for development of hind quarter and development of inner thigh in CH (Table 2.4). Moderate to strong genetic correlations also existed between the skeletal traits in all five breeds, ranging from 0.33 (SE = 0.12) for pelvic length and chest width in SI (Table 2.7) to 0.98 (SE = 0.01) between wither height and both pelvic length and back length in CH. The genetic correlations among the functional traits varied considerably among the breeds ranging from -0.08 (SE = 0.29) between foreleg front view and locomotion in SI to 0.87 (SE = 0.14) between the same traits in AA (Table 2.3).

Box's M test for homogeneity of the co-variance matrices among the breeds revealed that all co-variance matrices estimated within breed differed from each other except for when the AA and HE were compared. The majority of the pair-wise estimated within-breed genetic correlations differed ($P < 0.05$) between at least two breeds. The fewest differences in correlations were between when the AA and HE were compared; the greatest number of within-breed estimated genetic correlations among traits was observed when the CH was compared to either the AA or the HE.

Table 2.1 Scale of measurement, number of records (n), mean (μ), genetic standard deviation (SD_g) and heritability estimates (h^2) of the functional and muscular linear type traits.

		Angus[1]			Charolais[1]			Hereford[1]			Limousin[1]			Simmental[1]		
		n = 3,220 - 3,356			n = 23,070 - 31,048			n = 2,390 - 3,004			n = 30,491 - 35,158			n = 6,638 - 8,632		
Trait	Scale	μ	SDg	h^2	μ	SDg	h^2	μ	SDg	h^2	μ	SDg	h^2	μ	SDg	h^2
Functional	1 - 10															
Locomotion	low - high	7.69	0.28	0.12	7.66	0.32	0.12	7.80	0.00	0.00	8.11	0.17	0.04	8.10	0.18	0.04
Foreleg front view	toes out - toes in	5.27	0.24	0.13	6.21	0.24	0.09	5.51	0.00	0.00	6.21	0.16	0.06	6.70	0.20	0.06
Hind-leg side view	straight - sickled	7.17	0.21	0.08	7.30	0.27	0.09	7.34	0.24	0.11	7.58	0.24	0.08	7.40	0.21	0.06
Hind-leg rear view	toes out - toes in	5.26	0.16	0.04	5.98	0.26	0.06	5.61	0.00	0.00	6.43	0.21	0.04	5.65	0.25	0.06
Muscular	1 - 15															
Development of hind quarter	low - high	8.03	0.43	0.22	9.71	0.60	0.30	8.06	0.35	0.14	11.47	0.52	0.25	10.91	0.51	0.24
Development of loin	low - high	8.21	0.37	0.13	9.45	0.52	0.21	8.66	0.31	0.10	10.58	0.45	0.17	9.88	0.47	0.18
Thigh width	narrow - wide	8.21	0.38	0.14	9.75	0.55	0.22	8.24	0.40	0.16	10.22	0.53	0.23	9.92	0.55	0.24
Development of inner thigh	low - high	8.47	0.37	0.14	10.39	0.62	0.28	8.28	0.43	0.20	11.14	0.54	0.24	10.44	0.51	0.23
Width of withers	narrow - wide	8.91	0.51	0.22	9.36	0.51	0.21	8.88	0.41	0.16	10.32	0.46	0.19	10.17	0.54	0.22
Width behind withers	narrow - wide	7.51	0.39	0.13	8.64	0.46	0.18	7.94	0.40	0.15	9.46	0.43	0.17	9.11	0.48	0.18

[1] Standard error of the heritability estimates in Angus ≤ 0.05. Standard error of the heritability estimates in Charolais ≤ 0.02. Standard error of the heritability estimates in Hereford ≤ 0.05. Standard error of the heritability estimates in Limousin ≤ 0.02. Standard error of the heritability estimates in Simmental ≤ 0.03.

Table 2.2 Scale of measurement, number of records (n), mean (μ), genetic standard deviation (SD_g) and heritability estimates (h^2) of the skeletal and other linear type traits

		Angus[1]			Charolais[1]			Hereford[1]			Limousin[1]			Simmental[1]		
		n = 3,124 - 3,356			n = 21,341 - 31,044			n = 2,993 - 3,004			n = 30,494 - 35,156			n = 6,637 - 8,631		
Trait	Scale	μ	SDg	h^2	μ	SDg	h^2	μ	SDg	h^2	μ	SDg	h^2	μ	SDg	h^2
Skeletal	1 - 10															
Width of chest	narrow - wide	6.56	0.20	0.07	6.94	0.24	0.10	6.53	0.00	0.00	6.19	0.24	0.10	6.80	0.30	0.15
Depth of chest	shallow - deep	7.40	0.29	0.15	7.17	0.24	0.13	7.26	0.36	0.25	6.96	0.27	0.15	7.64	0.26	0.14
Height of withers	small- tall	5.81	0.38	0.19	6.76	0.65	0.43	5.69	0.44	0.30	6.61	0.47	0.29	7.17	0.52	0.34
Length of pelvis	short - long	7.07	0.35	0.17	7.41	0.42	0.23	7.01	0.45	0.27	7.83	0.37	0.19	8.05	0.37	0.20
Length of back	short - long	6.83	0.36	0.17	7.71	0.49	0.30	6.78	0.47	0.29	7.68	0.42	0.23	7.97	0.37	0.20
Width at hips	narrow - wide	6.49	0.21	0.06	6.89	0.29	0.13	6.86	0.00	0.00	6.68	0.30	0.14	7.06	0.30	0.14
Other	1 - 10															
Body condition score	lean - fat	7.04	0.18	0.03	5.84	0.35	0.13	7.21	0.00	0.00	6.57	0.31	0.11	7.14	0.23	0.05
Docility	aggressive - docile	8.74	0.36	0.21	8.86	0.34	0.15	9.24	0.26	0.11	9.22	0.37	0.17	9.26	0.30	0.09

[1] Standard error of the heritability estimates in Angus ≤ 0.05. Standard error of the heritability estimates in Charolais ≤ 0.02. Standard error of the heritability estimates in Hereford ≤ 0.06. Standard error of the heritability estimates in Limousin ≤ 0.02. Standard error of the heritability estimates in Simmental ≤ 0.03.

2.5.3 Eigenstructures

The rescaled $\widetilde{\boldsymbol{D}}$ matrices calculated using the breed-specific co-variance matrices of the skeletal, functional and muscular traits in LM (compared to the CH as the reference breed) had off-diagonal elements close to zero; the mean (standard deviation) of the absolute values of the off-diagonals was 0.14 (0.17) for the skeletal traits, 0.17 (0.23) for the muscular traits and 0.05 (0.04) for the functional traits. The off-diagonal elements of $\widetilde{\boldsymbol{D}}$ calculated from the co-variance matrices of the linear type traits in AA were furthest from zero; the mean (standard deviation) of the absolute values of the off-diagonals was 0.21 (0.23) for the skeletal traits, 0.19 (0.19) for the muscular traits and 0.12 (0.04) for the functional traits.

2.5.4 Impact of Incorrect (co)variance Parameters on the Estimation of Breeding Value

The impact of using the genetic (co)variance components of the LM to predict genetic merit for height at withers in CH was least but still the efficiency of selection was just 0.62; the efficiency was 0.61 when the (co)variance components of the SI were used. When the genetic (co)variance components of the CH were replaced with those of the AA, the efficiency of the index was just 0.29.

Table 2.3 Phenotypic (below the diagonal) and genetic (above the diagonal) correlations between linear type traits in Angus[1,2]

	LOCO	FL-FV	HL-SV	HL-RV	DHQ	DL	TW	DIT	WOW	WBW	CW	CD	WH	PL	BL	HW	BCS	DOC
LOCO		0.87	-0.57	0.41	0.34	0.43	0.23	0.12	0.52	0.84	0.33	0.19	-0.10	0.21	0.15	0.15	0.66	0.52
FL-FV	0.28		-0.10	0.17	0.62	0.83	0.78	0.63	0.66	0.64	0.33	0.45	0.31	0.49	0.38	0.68	0.68	0.37
HL-SV	-0.38	-0.02		0.16	-0.50	-0.66	-0.13	-0.46	0.07	-0.13	0.47	0.01	0.01	-0.05	-0.02	0.09	0.15	-0.09
HL-RV	0.21	0.08	-0.04		0.33	0.07	0.05	0.30	0.12	0.33	0.43	-0.29	-0.46	-0.28	-0.61	0.09	-0.17	0.36
DHQ	0.20	0.13	-0.13	0.18		0.87	0.83	0.77	0.66	0.78	0.57	0.27	0.13	-0.003	0.35	0.41	0.69	0.43
DL	0.28	0.13	-0.14	0.13	0.64		0.74	0.93	0.77	0.89	0.33	0.26	0.22	0.18	0.37	0.43	0.53	0.76
TW	0.16	0.13	-0.07	0.15	0.64	0.64		0.71	0.77	0.75	0.75	0.45	0.18	0.29	0.40	0.85	0.75	0.63
DIT	0.14	0.11	-0.14	0.15	0.77	0.62	0.68		0.73	0.82	0.43	0.22	0.15	0.15	0.30	0.46	0.78	0.68
WOW	0.22	0.11	-0.08	0.07	0.58	0.68	0.66	0.59		0.95	0.65	0.18	0.25	0.20	0.30	0.73	0.63	0.67
WBW	0.25	0.09	-0.10	0.10	0.61	0.72	0.63	0.60	0.85		0.65	0.23	0.28	0.13	0.36	0.87	0.84	0.56
CW	0.14	0.14	-0.04	0.14	0.39	0.45	0.52	0.41	0.48	0.48		0.72	0.84	0.46	0.69	0.84	0.63	0.35
CD	0.09	0.09	0.01	0.05	0.30	0.39	0.47	0.33	0.41	0.38	0.70		0.84	0.87	0.68	0.86	0.60	0.55
WH	0.08	0.07	-0.01	0.05	0.23	0.32	0.40	0.27	0.39	0.34	0.45	0.59		0.8	0.95	0.86	0.42	0.53
PL	0.12	0.09	-0.01	0.08	0.25	0.33	0.40	0.29	0.37	0.33	0.42	0.54	0.73		0.71	0.84	-0.16	0.63
BL	0.12	0.08	-0.03	0.05	0.27	0.36	0.41	0.30	0.39	0.36	0.41	0.51	0.70	0.70		0.78	0.60	0.34
HW	0.15	0.13	-0.01	0.13	0.32	0.43	0.53	0.39	0.49	0.46	0.50	0.52	0.54	0.55	0.67		0.87	0.50
BCS	0.20	0.15	-0.07	0.12	0.39	0.48	0.47	0.42	0.43	0.48	0.41	0.46	0.28	0.29	0.27	0.44		0.81
DOC	0.18	0.07	-0.13	0.06	0.22	0.24	0.24	0.24	0.27	0.24	0.16	0.16	0.17	0.18	0.18	0.20	0.18	

[1] LOCO = locomotion, FL-FV = foreleg front view, HL-SV = hind-leg side view, HL-RV = hind-leg rear view, DHQ = development of hind quarter, DL = development of loin, TW = thigh width, DIT = development of inner thigh, WOW = width of withers, WBW = width behind withers, CW = chest width, CD = chest depth, WH = wither height, PL = pelvic length, BL = back length, HW = hip width, BCS = body condition score, DOC = docility.

[2] Standard errors for the phenotypic correlations ranged from 0.01 to 0.02. Standard errors for the genetic correlations varied from 0.02 to 0.49.

Table 2.4 Phenotypic (below the diagonal) and genetic (above the diagonal) correlations between linear type traits in Charolais[1,2]

	LOCO	FL-FV	HL-SV	HL-RV	DHQ	DL	TW	DIT	WOW	WBW	CW	CD	WH	PL	BL	HW	BCS	DOC
LOCO		0.12	-0.90	-0.35	0.05	0.16	0.10	0.14	0.14	0.23	0.15	-0.16	0.13	0.07	0.10	0.17	0.23	0.04
FL-FV	0.19		0.41	-0.05	0.01	0.03	0.06	-0.13	0.08	0.05	0.35	0.03	0.37	0.30	0.34	0.24	-0.11	0.26
HL-SV	-0.57	0.17		0.26	0.16	0.09	0.11	0.01	0.11	0.04	0.03	0.21	0.49	0.29	0.43	0.15	-0.08	0.06
HL-RV	-0.06	0.05	0.04		0.56	0.56	0.51	0.54	0.49	0.48	0.26	-0.20	-0.44	-0.22	-0.35	0.25	0.34	0.06
DHQ	0.10	0.06	0.06	0.16		0.95	0.93	0.99	0.88	0.87	0.60	0.11	0.04	0.05	0.08	0.60	0.76	0.37
DL	0.12	0.07	0.03	0.16	0.70		0.94	0.94	0.92	0.94	0.66	0.21	0.15	0.19	0.21	0.71	0.81	0.43
TW	0.10	0.08	0.06	0.16	0.70	0.72		0.93	0.92	0.92	0.85	0.40	0.22	0.25	0.27	0.83	0.84	0.40
DIT	0.12	-0.01	-0.04	0.16	0.78	0.68	0.71		0.87	0.86	0.66	0.18	0.02	-0.02	0.03	0.54	0.82	0.40
WOW	0.10	0.09	0.07	0.14	0.65	0.72	0.73	0.64		0.96	0.75	0.36	0.20	0.21	0.22	0.71	0.77	0.38
WBW	0.14	0.09	0.03	0.13	0.65	0.74	0.72	0.65	0.87		0.67	0.34	0.17	0.22	0.18	0.72	0.83	0.32
CW	0.07	0.08	0.00	0.09	0.38	0.43	0.52	0.41	0.48	0.47		0.78	0.70	0.61	0.65	0.88	0.74	0.45
CD	0.06	0.05	0.01	0.04	0.28	0.37	0.46	0.33	0.43	0.42	0.52		0.95	0.90	0.84	0.84	0.50	0.32
WH	0.12	0.18	0.13	-0.03	0.17	0.27	0.33	0.18	0.34	0.33	0.48	0.67		0.98	0.98	0.71	0.14	0.17
PL	0.09	0.12	0.08	0.01	0.18	0.28	0.34	0.19	0.34	0.33	0.41	0.57	0.67		0.90	0.76	0.03	0.15
BL	0.10	0.14	0.10	0.00	0.20	0.30	0.36	0.20	0.35	0.34	0.43	0.57	0.78	0.81		0.77	0.08	0.14
HW	0.09	0.12	0.09	0.08	0.41	0.49	0.58	0.41	0.54	0.53	0.49	0.52	0.48	0.48	0.48		0.71	0.44
BCS	0.12	0.05	-0.04	0.12	0.45	0.51	0.55	0.50	0.52	0.53	0.43	0.42	0.30	0.26	0.27	0.41		0.28
DOC	0.09	0.03	-0.09	0.05	0.16	0.16	0.17	0.19	0.18	0.16	0.10	0.12	0.08	0.09	0.08	0.12	0.12	

[1] LOCO = locomotion, FL-FV = foreleg front view, HL-SV = hind-leg side view, HL-RV = hind-leg rear view, DHQ = development of hind quarter, DL = development of loin, TW = thigh width, DIT = development of inner thigh, WOW = width of withers, WBW = width behind withers, CW = chest width, CD = chest depth, WH = wither height, PL = pelvic length, BL = back length, HW = hip width, BCS = body condition score, DOC = docility.

[2] Standard error for the phenotypic correlations was 0.01. Standard errors for the genetic correlations varied between 0.01 and 0.12.

Table 2.5 Phenotypic (below the diagonal) and genetic (above the diagonal) correlations between linear type traits in Hereford[1,2]

	HL-SV	DHQ	DL	TW	DIT	WOW	WBW	CD	WH	PL	BL	DOC
HL-SV		-0.34	-0.25	-0.41	-0.4	-0.24	-0.26	0.18	-0.27	-0.23	-0.2	-0.08
DHQ	-0.13		0.62	0.81	0.88	0.58	0.70	0.26	0.29	0.31	0.42	0.44
DL	-0.08	0.55		0.84	0.69	0.82	0.84	0.47	0.54	0.56	0.82	0.58
TW	-0.11	0.64	0.71		0.73	0.70	0.85	0.45	0.38	0.44	0.45	0.40
DIT	-0.16	0.80	0.55	0.70		0.74	0.86	0.49	0.41	0.49	0.46	0.34
WOW	-0.11	0.52	0.67	0.62	0.56		0.90	0.53	0.38	0.53	0.57	0.32
WBW	-0.12	0.63	0.76	0.72	0.58	0.82		0.21	0.43	0.50	0.56	0.48
CD	-0.02	0.30	0.45	0.46	0.36	0.46	0.40		0.84	0.82	0.70	0.60
WH	-0.02	0.25	0.37	0.40	0.29	0.41	0.35	0.56		0.76	0.94	0.29
PL	-0.02	0.26	0.34	0.39	0.29	0.38	0.32	0.54	0.77		0.80	0.16
BL	-0.03	0.27	0.37	0.38	0.29	0.38	0.33	0.48	0.70	0.77		0.43
DOC	-0.07	0.17	0.18	0.30	0.17	0.19	0.17	0.17	0.22	0.75	0.15	

[1] HL-SV = hind-leg side view, DHQ = development of hind quarter, DL = development of loin, TW = thigh width, DIT = development of inner thigh, WOW = width of withers, WBW = width behind withers, CD = chest depth, WH = wither height, PL = pelvic length, BL = back length, DOC = docility.

[2] Standard errors for the phenotypic correlations ranged from 0.01 to 0.02. Standard errors for the genetic correlations varied between 0.02 and 0.29.

Table 2.6 Phenotypic (below the diagonal) and genetic (above the diagonal) correlations between linear type traits in Limousin[1,2]

	LOCO	FL-FV	HL-SV	HL-RV	DHQ	DL	TW	DIT	WOW	WBW	CW	CD	WH	PL	BL	HW	BCS	DOC
LOCO		0.10	0.01	-0.51	0.06	0.09	0.09	-0.08	0.18	0.08	0.21	-0.08	0.11	0.05	0.06	0.11	0.09	0.10
FL-FV	0.21		0.19	0.06	-0.14	-0.03	-0.01	-0.08	-0.10	-0.08	0.16	0.13	0.03	0.06	0.04	0.01	0.13	-0.02
HL-SV	-0.08	-0.01		0.29	0.02	-0.02	0.00	0.10	-0.03	-0.03	0.13	0.20	0.10	0.09	0.12	0.05	0.08	-0.23
HL-RV	-0.11	0.03	0.06		0.39	0.34	0.35	0.47	0.24	0.24	0.22	0.04	-0.24	-0.12	-0.19	0.07	0.30	0.24
DHQ	0.12	0.03	-0.01	0.11		0.81	0.86	0.87	0.73	0.80	0.63	0.21	-0.07	-0.06	0.01	0.47	0.63	0.27
DL	0.14	0.05	-0.03	0.11	0.60		0.87	0.86	0.90	0.92	0.80	0.47	0.12	0.19	0.21	0.64	0.83	0.31
TW	0.14	0.06	-0.01	0.12	0.64	0.63		0.89	0.86	0.90	0.91	0.58	0.21	0.29	0.28	0.83	0.77	0.28
DIT	0.12	0.04	-0.01	0.12	0.71	0.61	0.67		0.74	0.80	0.70	0.38	-0.04	-0.02	0.04	0.56	0.78	0.33
WOW	0.12	0.04	-0.02	0.10	0.56	0.65	0.66	0.58		0.96	0.84	0.60	0.27	0.30	0.30	0.80	0.74	0.39
WBW	0.14	0.05	-0.03	0.10	0.58	0.69	0.67	0.59	0.82		0.83	0.55	0.25	0.27	0.27	0.75	0.77	0.35
CW	0.12	0.08	-0.02	0.08	0.39	0.46	0.54	0.43	0.51	0.50		0.86	0.56	0.57	0.52	0.92	0.81	0.31
CD	0.08	0.04	0.01	0.06	0.31	0.39	0.48	0.35	0.45	0.44	0.51		0.81	0.81	0.7	0.83	0.59	0.21
WH	0.10	0.03	-0.02	0.02	0.15	0.26	0.33	0.20	0.34	0.33	0.43	0.59		0.97	0.95	0.60	0.04	0.12
PL	0.08	0.03	-0.01	0.03	0.17	0.27	0.34	0.20	0.33	0.31	0.37	0.51	0.61		0.96	0.70	0.01	0.12
BL	0.09	0.03	-0.02	0.03	0.17	0.27	0.33	0.20	0.32	0.32	0.38	0.49	0.69	0.53		0.62	0.01	0.22
HW	0.12	0.05	-0.01	0.06	0.36	0.44	0.56	0.41	0.51	0.50	0.50	0.52	0.43	0.45	0.41		0.63	0.26
BCS	0.12	0.06	-0.03	0.10	0.43	0.49	0.53	0.48	0.48	0.51	0.45	0.39	0.26	0.22	0.22	0.40		0.25
DOC	0.09	0.02	-0.08	0.05	0.27	0.14	0.14	0.16	0.16	0.15	0.09	0.11	0.07	0.09	0.09	0.11	0.09	

[1] LOCO = locomotion, FL-FV = foreleg front view, HL-SV = hind-leg side view, HL-RV = hind-leg rear view, DHQ = development of hind quarter, DL = development of loin, TW = thigh width, DIT = development of inner thigh, WOW = width of withers, WBW = width behind withers, CW = chest width, CD = chest depth, WH = wither height, PL = pelvic length, BL = back length, HW = hip width, BCS = body condition score, DOC = docility.

[2] Standard error for the phenotypic correlations was 0.01. Standard errors for the genetic correlations varied between 0.01 and 0.13.

Table 2.7 Phenotypic (below the diagonal) and genetic (above the diagonal) correlations between the linear type traits in Simmental[1,2]

	LOCO	FL-FV	HL-SV	HL-RV	DHQ	DL	TW	DIT	WOW	WBW	CW	CD	WH	PL	BL	HW	BCS	DOC
LOCO		-0.08	0.27	0.61	0.34	0.70	0.64	0.27	0.56	0.14	-0.17	-0.31	-0.01	0.13	0.58	0.77	0.76	0.49
FL-FV	0.22		0.44	0.45	0.42	0.32	0.34	0.25	0.14	0.27	0.21	0.01	0.20	-0.10	0.12	0.08	0.14	-0.07
HL-SV	0.06	0.13		0.22	0.18	0.08	0.39	0.20	0.10	0.20	0.29	0.12	0.09	0.27	0.20	0.25	0.34	-0.70
HL-RV	0.12	0.09	0.05		0.62	0.85	0.66	0.78	0.67	0.65	0.89	0.37	0.15	0.20	0.32	0.57	0.70	0.49
DHQ	0.29	0.09	0.01	0.15		0.81	0.83	0.92	0.72	0.79	0.74	0.20	0.04	0.08	0.25	0.45	0.82	0.37
DL	0.24	0.09	-0.02	0.16	0.59		0.87	0.78	0.97	0.98	0.84	0.43	0.04	0.13	0.30	0.53	0.55	0.29
TW	0.24	0.12	0.04	0.13	0.62	0.62		0.81	0.80	0.76	0.94	0.51	0.22	0.32	0.37	0.81	0.88	0.22
DIT	0.19	0.06	-0.03	0.15	0.76	0.60	0.66		0.76	0.82	0.78	0.25	-0.02	0.06	0.15	0.51	0.76	0.35
WOW	0.23	0.09	-0.01	0.16	0.54	0.64	0.63	0.57		0.96	0.82	0.38	0.06	0.04	0.25	0.66	0.72	0.39
WBW	0.18	0.10	-0.02	0.16	0.57	0.73	0.62	0.58	0.81		0.76	0.26	0.03	0.04	0.21	0.66	0.59	0.39
CW	-0.07	0.14	0.01	0.14	0.46	0.50	0.57	0.46	0.52	0.51		0.76	0.42	0.33	0.53	0.91	0.73	0.18
CD	-0.14	0.07	0.02	0.08	0.30	0.36	0.44	0.31	0.39	0.37	0.47		0.90	0.80	0.81	0.60	0.64	0.10
WH	-0.07	0.07	0.01	0.05	0.17	0.25	0.30	0.18	0.36	0.26	0.36	0.59		0.95	0.95	0.59	0.02	0.13
PL	0.06	0.07	0.09	0.06	0.20	0.27	0.31	0.20	0.27	0.27	0.33	0.49	0.64		0.96	0.59	0.32	0.24
BL	0.12	0.04	0.10	0.06	0.19	0.26	0.30	0.20	0.27	0.28	0.34	0.45	0.66	0.62		0.71	0.10	0.35
HW	0.18	0.09	0.04	0.13	0.35	0.45	0.52	0.40	0.45	0.45	0.52	0.46	0.39	0.42	0.38		0.47	0.16
BCS	0.19	0.11	-0.03	0.12	0.46	0.48	0.53	0.46	0.50	0.47	0.43	0.38	0.23	0.23	0.21	0.43		0.14
DOC	0.13	0.04	-0.10	0.06	0.15	0.15	0.13	0.18	0.18	0.16	0.11	0.12	0.07	0.10	0.10	0.13	0.11	

[1] LOCO = locomotion, FL-FV = foreleg front view, HL-SV = hind-leg side view, HL-RV = hind-leg rear view, DHQ = development of hind quarter, DL = development of loin, TW = thigh width, DIT = development of inner thigh, WOW = width of withers, WBW = width behind withers, CW = chest width, CD = chest depth, WH = wither height, PL = pelvic length, BL = back length, HW = hip width, BCS = body condition score, DOC = docility.

[2] Standard error for the phenotypic correlations was 0.01. Standard errors for the genetic correlations varied between 0.02 and 0.29

2.5.5 Sexual Dimorphism

Although the genetic variance for the linear type traits was greater in male than female LM, no differences existed in the heritability estimates between the two sexes in LM (Table 2.8). The genetic variance of the type traits in CH was numerically greater in males than females for 14 of the 18 traits. The genetic variance of the type traits in CH was greater in females than males for development of hind quarter, hip width, body condition score and docility. Nevertheless, differences ($p < 0.05$) in the heritability estimates between sexes only existed for back length (males 0.36; females 0.12), wither height (males 0.68; females 0.29) and development of hind quarter (males 0.23; females 0.33) in CH. In both CH and LM, genetic correlations between the same linear type traits in both sexes were all greater than 0.90 (Table 2.9).

2.6 Discussion

Even though the homogeneity of co-variance matrices has long been a topic of interest in multivariate analysis (Box, 1949; Box 1953), previous studies that estimated the genetic parameters of linear type traits in beef cattle either did so on a single breed (Gutierrez and Goyache, 2002; Mantovani et al., 2010; Mazza et al., 2014; Vallee et al., 2015) or by collating multiple breeds and crosses into a single analysis (Mc Hugh et al., 2012); none have attempted to quantify if differences among breeds exist in genetic parameters of linear type traits. The absence of such information in the scientific literature may be due to classifiers often only performing linear type scoring on a single breed, thus contributing to confounding between a classifier effect and breed; 12 of the 20 classifiers included in the present study scored at least four of the five breeds. Linear type trait information from all breeds (and crossbreds) is collated into a centralized database in Ireland thus facilitating the analysis in the present study; such a centralized

system is not present in many countries with some breed societies responsible for the collection, collation and analysis of the data relating to their breed only. While differences in variance components of linear type traits among breeds have not been quantified previously, differences in genetic parameters among breeds have been reported previously for carcass traits (Marshall, 1994; Utrera and Van Vleck, 2004; Hickey et al., 2007; Pabiou et al., 2009; Kause et al., 2015) and birth and weaning weights (Phocas and Laloë, 2004). Studies are also lacking that investigated the possible existence of sexual dimorphism on variance components for linear type traits in beef cattle. Knowledge of the extent, if any, of breed differences in variance components, as well as the presence of sexual dimorphism for variance components is becoming more important as initiatives attempt to combine data from multiple sources in the pursuit of more accurate genomic evaluations.

2.6.1 Sexual Dimorphism

Sexual dimorphism in mammals and many other organisms is due to evolution by natural selection, specifically sexual selection. Sexual selection, a concept coined by Darwin, arises due to competition among the same sex of a species and due to mating preferences of one sex to the other (Kirkpatrick, 1987). Male and female mammals differ in many anatomical and physiological features concerning their role in the development and maintenance of their offspring, their body size, coloration, display characteristics and mating behaviour (McPherson and Chenoweth, 2012; van der Heide et al., 2016). Historically, sexual dimorphism tended to occur in mammals due to competition among males for access to females; males would fight one another and the winner, generally the biggest, strongest animal would mate with the females (Katz, 2008). This competition is, however, reduced in domestic animals where breeding

males are less likely to be selected for their size or aggressiveness but are selected on numerous other desirable traits. Sexual dimorphism has previously been researched in beef cattle for numerous important traits, such as growth rate (Koch and Clark, 1955; Marlowe and Gaines, 1958) and birth weight, weaning weight and post-weaning gain (van der Heide et al., 2016), but no study has been published that investigated the existence of sexual dimorphism on variance components for linear type traits in beef cattle.

While differences in the heritability estimates existed in three of the 18 linear type traits between the sexes in CH (development of hind quarter, wither height and back length), no differences existed in LM. All genetic correlations, in both CH and LM were ≥ 0.90. It has been proposed previously that traits with a correlation > 0.80 can be assumed to be genetically the same trait (Robertson, 1959), despite that fact that a correlation of 0.80 translates to only 64% of the variance in one trait being explained by the other. Combined, the results from the present study suggest little existence of appreciable sexual dimorphism on variance components in linear type traits and thus stratifying genetic evaluations into males and females is unlikely to be beneficial.

Table 2.8 The genetic standard deviation (SD_g), heritability estimate (h^2) and genetic correlation (r_g) of the linear type traits in male and female Limousin and Charolais animals.

	Limousin[1]					Charolais[2]				
	Male		Female			Male		Female		
Trait	SDg	h^2	SDg	h^2	r_g	SDg	h^2	SDg	h^2	r_g
Functional										
Locomotion	0.19	0.05	0.13	0.03	0.98	0.29	0.08	0.26	0.09	0.97
Foreleg front view	0.14	0.04	0.00	0.00	--	0.26	0.09	0.17	0.06	0.99
Hind-leg side view	0.24	0.08	0.15	0.04	0.99	0.23	0.06	0.16	0.04	0.98
Hind-leg rear view	0.19	0.03	0.16	0.02	0.99	0.24	0.05	0.17	0.03	0.98
Muscular										
Development of hind quarter	0.54	0.26	0.45	0.23	0.93	0.50	0.23*	0.57	0.33*	0.97
Development of loin	0.46	0.18	0.37	0.14	0.97	0.50	0.21	0.43	0.21	0.96
Thigh width	0.51	0.21	0.47	0.23	0.92	0.56	0.23	0.48	0.23	0.93
Development of inner thigh	0.53	0.21	0.49	0.23	0.94	0.53	0.23	0.53	0.27	0.96
Width of withers	0.45	0.19	0.38	0.17	0.97	0.53	0.23	0.44	0.21	0.96
Width behind withers	0.43	0.18	0.34	0.13	0.95	0.46	0.2	0.40	0.18	0.91
Skeletal										
Chest width	0.19	0.07	0.18	0.06	0.98	0.23	0.09	0.15	0.05	0.90
Chest depth	0.26	0.14	0.20	0.09	0.98	0.17	0.06	0.17	0.07	0.98
Wither height	0.47	0.27	0.40	0.24	0.94	0.89	0.68**	0.47	0.29**	0.99
Pelvic length	0.29	0.12	0.32	0.12	0.99	0.37	0.18	0.32	0.16	0.99
Back length	0.37	0.19	0.36	0.19	0.98	0.55	0.36**	0.36	0.12**	0.99
Hip width	0.23	0.09	0.27	0.12	0.95	0.18	0.06	0.25	0.11	0.98
Other										
Body condition score	0.32	0.13	0.28	0.10	0.99	0.24	0.07	0.30	0.12	0.98
Docility	0.30	0.11	0.32	0.13	0.98	0.29	0.11	0.31	0.14	0.99

[1] Standard errors for h^2 and r_g in Limousin were all ≤ 0.03. [2]Standard errors for h^2 and r_g in Charolais were all ≤ 0.04.

*p<0.05, **p<0.01

2.6.2 Muscularity Traits

The heritability estimates of the muscular linear type traits in the present study are in the range of what has been reported previously in Chianina beef cows (0.16 to 0.23; Forabosco et al., 2005), in the Rendena dual-purpose cattle breed (0.27 to 0.32; Mazza et al., 2014) and in beef cattle from the Czech Republic (0.26 to 0.35; Vesela et al., 2005). Irrespective of breed, the heritability estimates for the muscular traits in the present study were generally the greatest of all other traits assessed, which is consistent with previously documented heritability estimates in both beef (El-Saied et al., 2006) and dairy cattle (Brotherstone, 1994). Overall, CH tended to have the highest heritability estimates for muscular traits followed by LM and SI, while AA and HE had the lowest. While the heritability estimates and the genetic standard deviations for the muscular traits were greater in the continental breeds than in the British breeds, when rescaled to the mean, the extent of additive genetic variance was similar for all muscular traits.

The strong genetic correlations among the muscular traits are consistent with the correlations reported previously in Chianina beef cows (Forabosco et al., 2005), in the Rendena dual-purpose cattle breed (Mazza et al., 2014), and in Piemontese cows (Mantovani et al., 2010). Genetic correlations between width of withers and width behind withers were extremely strong across all five breeds, ranging from 0.90 to 0.96 suggesting redundancy; this is not unexpected since both traits are measures of animal width taken in close spatial proximity. Moreover, the redundancy is present across all breeds.

The mean and standard deviation of the off-diagonal elements of $\widetilde{\boldsymbol{D}}$ when calculated from the co-variance matrix of LM were close to zero, indicating that the co-variance matrix of LM and the co-variance matrix of CH were the most similar.

Similarly, the mean and standard deviation of $\widetilde{\boldsymbol{D}}$ calculated using the eigenvector matrix from AA with the co-variance matrix from HE suggests the co-variance matrices of these breeds were similar to one another.

2.6.3 Skeletal Traits

The heritability estimates of the skeletal linear type traits (0.00 – 0.43) are in the range of previous estimates reported in beef cattle; Gutierrez and Goyache (2002) reported heritability estimates of between 0.10 and 0.23 for the skeletal traits in Asturiana de los Valles beef cattle while Forabosco et al. (2005) reported heritability estimates in the range of 0.21 to 0.30 for Chianina beef cattle. The heritability estimates reported in the present study are also consistent with heritability estimates (0.23 to 0.38) relating to skeletal traits in dairy cattle (Veerkamp and Brotherstone, 1997; Berry et al., 2014).

The greatest differences in within-breed heritability estimates existed for wither height (0.19 in AA; 0.43 in CH) and back length (0.17 in AA; 0.30 in CH). The higher heritability estimate for wither height in CH is due to a larger genetic standard deviation (0.65) concurrent with a marginally smaller residual standard deviation (0.75) in CH than in AA. The differences in heritability estimates of back length in CH and AA are due to CH having a slightly lower residual standard deviation (0.75) than AA (0.79). The lower heritability and genetic variation for height at withers in AA may be related to AA, not only being generally smaller than CH, but also reaching mature height earlier than CH (Arango et al., 2002) and thus having less variability in height at withers at younger ages.

Excluding height at withers, and with the exception of the two skeletal traits in the HE with no genetic variation (chest width and hip width), the other skeletal traits

across all breeds expressed similar genetic variation when rescaled to the respective breed mean (0.03 to 0.07). This implies that, once scaled to the breed mean for that trait, the extent of additive genetic variance was similar within these traits and across the breeds.

Regardless of breed, genetic correlations among the skeletal traits were all generally moderate to strong, corroborating genetic correlation estimates among skeletal traits in beef cattle from other populations such as Asturiana de los Valles beef cattle (Gutierrez and Goyache, 2002) and beef cattle from the Czech Republic (Vesela et al., 2005). Overall, the strongest pairwise genetic correlations existed between the skeletal traits in CH while the weakest genetic correlations existed in AA. The skeletal traits were also moderately to strongly correlated with the muscular traits, signifying that the more muscular animals also have a tendency to score higher for skeletal type traits (Grona et al., 2002).

2.6.4 Functional Traits

While few previous studies have reported heritability estimates for functional traits in beef cattle, the heritability estimates reported in the present study are comparable to what has been reported previously in CH cattle (0.02 to 0.11; Vallee et al., 2015). The heritability estimates are, however, slightly lower than those reported in Brazilian Holstein cows (0.08 to 0.19; Kern et al., 2015) and Irish Holstein-Friesians (0.14 to 0.19; Berry et al., 2004). Excluding the three functional traits with no genetic variation in HE, the genetic standard deviation was similar (0.16 to 0.32) across the other functional traits in all breeds.

Unlike the skeletal and muscular traits, the genetic correlations among the functional traits did not follow any particular pattern with the correlations differing greatly among the breeds. The differences among the breeds, in both the variances and genetic correlations among the functional traits, may be due to the level of environmental influence. Environmental factors such as housing type, diet and hoof trimming schedules will affect the feet and legs of an animal potentially influencing the linear type classification (Fatehi et al., 2003). The differences in heritability estimates and genetic correlations observed among the linear type traits in the five breeds may be real differences in parameters among populations or may be due to the small sample size of HE and AA available for analysis in comparison to the continental breeds, or a combination of the two (Koots and Gibson, 1996).

The co-variance structures of the functional traits in LM and SI were similar to the co-variance structure of CH, as indicated by the means and standard deviations of the off-diagonal elements of $\widetilde{\boldsymbol{D}}$**.** Larger differences in the co-variance structures existed between CH and AA, signifying these breeds were more different to one another than CH was to LM or SI.

2.6.5 Efficiency of Index when using Incorrect Genetic Parameters

The efficiency associated with using the genetic parameters of LM in place of CH was poor despite the general similarity in estimated inter-trait genetic (co)variances among all 18 traits; this observed poor efficiency was due to the traits chosen to be included in the selection index to be contributors to variability in the goal trait but also variable between breeds. The efficiency of the index when CH was replaced by SI was similar to that observed for the LM which is not overly surprising given the relative similarities of

in the origin of these breeds (Kelleher et al., 2016). The large reduction in index efficiency associated with using the genetic parameters of AA suggests that AA should ideally not be included in multi-breed genetic evaluations with the continental breeds as this may lead to a marked reduction in accuracy. It should be noted nonetheless, that the index efficiency presented here was based on true breeding values which may not always be available and thus represents an upper threshold to this efficiency.

2.7 Conclusion

While the sex of an animal had little to no effect on the heritability estimates and genetic correlations among linear type traits, differences among the breeds in both the heritability estimates and in the genetic correlations among the linear type traits did exist. The greatest differences existed between the continental breeds and the British breeds suggesting that the accuracy of genetic evaluations may benefit from considering, at least, these breed groups separately in future evaluations.

Chapter 3

Genomic regions associated with muscularity in beef cattle differ in five contrasting cattle breeds

3.1 Preface

At the time of thesis submission this chapter was published in the Genetics Selection Evolution (Accepted January 17, 2020; doi: 10.1186/s12711-020-0523-1). The full reference is Doyle JL, Berry DP, Veerkamp RF, Carthy TR, Evans RD, Walsh SW, Purfield DC: Genomic regions associated with muscularity in beef cattle differ in five contrasting cattle breeds. Genetic Selection Evolution 2020, 52:2

Jennifer Doyle was primary author, performed the data edits and analysis and drafted the manuscript. Donagh Berry, Deirdre Purfield and Roel Veerkamp conceived the study, participated in the design and co-ordination of this study and helped draft the manuscript. Ross Evans supplied the data the analysis was performed on. All authors read and approved the final manuscript.

Formatting and referencing style has been edited for consistency throughout the thesis. Figure and table captions have been assigned with a chapter prefix. Competing interests and acknowledgements have been removed. All other aspects are consistent with the published manuscript.

3.2 Abstract

Linear type traits, which reflect the muscular characteristics of an animal, could provide insight into how, in some cases, morphologically very different animals can yield the same carcass weight. Such variability may contribute to differences in the overall value of the carcass since primal cuts vary greatly in price; such variability may also hinder successful genome-based association studies. Therefore, the objective of our study was to identify genomic regions that are associated with five muscularity linear type traits and to determine if these significant regions are common across five different breeds. Analyses were carried out using linear mixed models on imputed whole-genome sequence data in each of the five breeds, separately. Then, the results of the within-breed analyses were used to conduct an across-breed meta-analysis per trait. We identified many quantitative trait loci (QTL) that are located across the whole genome and associated with each trait in each breed. The only commonality among the breeds and traits was a large-effect pleiotropic QTL on BTA2 that contained the *MSTN* gene, which was associated with all traits in the Charolais and Limousin breeds. Other plausible candidate genes were identified for muscularity traits including *PDE1A*, *PPP1R1C* and multiple collagen and *HOXD* genes. In addition, associated (gene ontology) GO terms and KEGG pathways tended to differ between breeds and between traits especially in the numerically smaller populations of Angus, Hereford, and Simmental breeds. Most of the SNPs that were associated with any of the traits were intergenic or intronic SNPs located within regulatory regions of the genome. The commonality between the Charolais and Limousin breeds indicates that the genetic architecture of the muscularity traits may be similar in these breeds due to their similar origins. Conversely, there were vast differences in the QTL associated with muscularity in Angus, Hereford, and Simmental. Knowledge of these differences in genetic

architecture between breeds is useful to develop accurate genomic prediction equations that can operate effectively across breeds. Overall, the associated QTL differed according to trait, which suggests that breeding for a morphologically different (e.g, longer and wider versus shorter and smaller), more efficient animal may become possible in the future.

3.3 Introduction

Linear type traits have been used extensively to characterize conformation in both dairy (Veerkamp and Brotherstone, 1997; Berry et al., 2004; Kern et al., 2015) and beef cattle (Mc Hugh et al., 2010; Mazza et al., 2014). Muscularity linear type traits have previously been documented as moderate to highly heritable traits in beef cattle (Chapter 2; Forabosco et al., 2005; Mazza et al., 2014) and are known to be genetically associated with carcass merit (Mukai et al., 1995; Conroy et al., 2010) and with both animal live weight and price (Mc Hugh et al., 2010). Therefore, the genetic merit of a young animal for these traits may be a good representation of its merit for carcass traits. While both carcass value and conformation have been reported to be correlated with linear type traits (Conroy et al., 2010), the correlation with any one type trait is not equal to 1 which implies that the same carcass value can be achieved with morphologically different animals; by extension then, this implies that, for example, an animal with a better developed loin and a shallow chest may have the same yield as an animal with a lesser developed loin and a deep chest. Such morphological differences could contribute, in turn, to differences in individual carcass retail cut weights, and thus overall carcass value.

Many previous genomic studies in cattle have focused on live weight and

carcass traits as the phenotypes of interest (McClure et al., 2010; Bolormaa et al., 2011; Nishimura et al., 2012), but only a few have been published on the underlying features that contribute to differences in linear type traits in either beef cattle (Vallée et al., 2016) or dairy cattle (Wu et al., 2013). While previous studies have attempted to compare and contrast putative mutations, genes, and associated biological pathways across multiple breeds of beef cattle for carcass traits (Saatchi et al., 2014b), no study has attempted to do this using linear type traits. Knowledge of any kind of similarities or differences between breeds could enable the introduction of more accurate multi-breed genomic evaluations for both pure and crossbred animals. Therefore, the objective of the present study was to identify genomic regions associated with five muscularity linear type traits and to determine if these associated regions are common across multiple beef cattle breeds.

3.4 Materials & Methods

3.4.1 Phenotypic Data

As part of the Irish national beef breeding program, routine scoring of linear type traits is carried out on both registered and commercial beef herds by trained classifiers who are employed by the Irish Cattle Breeding Federation (Mc Hugh et al., 2010; Berry and Evans, 2014), with each classifier scoring animals from a range of different breeds. The muscularity type traits used in the present study describe the development of the hind quarter (DHQ), inner thigh (DIT), and loin (DL), and the width of the thigh (TW) and withers (WOW). Each trait was scored on a scale from 1 to 15 where 1 = low and 15 = high for DHQ, DIT and DL, and 1 = narrow and 15 = wide for TW and WOW (Appendix A1). Data on these five linear type traits were available for 147,704 purebred Angus (AA), Charolais (CH), Hereford (HE), Limousin (LM), or Simmental (SI) beef

cattle scored between the age of 6 and 16 months from 2000 to 2016 (Chapter 2).

Animals were discarded from the dataset if the sire, dam, herd, or classifier was unknown, or if the parity of the dam was not recorded. Parity of the dam was recoded as 1, 2, 3, 4, and ≥ 5. Contemporary group was defined as herd-by-scoring date generated separately per breed. Each contemporary group had to have at least five records. Following these edits, data were available on 81,200 animals: 3,356 AA, 31,049 CH, 3,004 HE, 35,159 LM and 8,632 SI.

3.4.2 Generation of Adjusted Phenotypes

Prior to inclusion in the analysis, all phenotypes were first adjusted within-breed in ASREML (Gilmour et al., 2009) using the model:

$$y_{jklm} = HSD_i + Sex_j + AM_k + DP_l + Animal_m + e_{ijklm}$$

where y_{ijklm} is the linear type trait, HSD_i is the fixed effect of herd by scoring date (11,130 levels), Sex_j is the fixed effect of the sex of the animal (male or female), AM_k is the fixed effect of the age in months of the animal (11 classes from 6 to 16 months), DP_l is the fixed effect of the parity of the dam (1, 2, 3, 4 and ≥ 5), $Animal_m$ is the random additive effect of the animal, and e_{ijklm} is the random residual effect. The adjusted phenotype was the raw phenotype minus the fixed effect solutions of HSD, Sex, AM and DP.

3.4.3 Genotype Data

Of the 81,200 animals with linear type trait information, 19,449 animals from five beef breeds (1,444 AA, 6,433 CH, 1,129 HE, 8,745 LM, and 1,698 SI) were imputed to whole-genome sequence as part of a larger dataset of 638,662 multi-breed genotyped animals. All 638,662 animals were genotyped using the Bovine Illumina SNP50 panel (n = 5,808; 54,001 single nucleotide polymorphisms, the Illumina High Density (HD) panel (n = 5,504; 777,972 SNPs), the Illumina 3k panel (n = 2,256; 2900 SNPs), the Illumina low-density (LD) genotyping panel (n = 15,107; 6909 SNPs) or a bespoke genotype panel (IDB) developed in Ireland (Mullen et al., 2013) with three versions, i.e. version 1 (n = 28,288; 17,137 SNPs), version 2 (n = 147,235; 18,004 SNPs) and version 3 (n = 434,464; 53,450 SNPs). Each animal had a call rate higher than 90% and only autosomal SNPs, SNPs with a known chromosome and position on UMD 3.1, and SNPs with a call rate higher than 90% within a panel were retained for imputation.

All genotyped animals were imputed to HD using a two-step approach in FImpute2 with pedigree information (Sargolzaei et al., 2014); this involved imputing the 3k, LD and IDB genotyped animals to the Bovine SNP50 density, and consequently imputing all resulting genotypes (including the Bovine SNP50 genotypes) to HD using a multi-breed reference population of 5,504 influential sires genotyped on the HD panel. Imputation to whole-genome sequence (WGS) was then undertaken using a reference population of 2,333 *Bos taurus* animals from multiple breeds from Run6.0 of the 1000 Bull Genomes Project (Daetwyler et al., 2014). All variants in the sequence reference population were called using SAMtools and genotype calls were improved using the Beagle software to provide a consensus SNP density across all animals. Details of the alignment to UMD 3.1 bovine reference genome, variant calling and quality controls completed within the multi-breed reference population are described in Daetwyler et al.

(Daetwyler et al., 2014). In total, 41.39 million SNPs were identified across the genome and the average coverage was 12.85X. Imputation of the HD genotypes to WGS was completed by first phasing all 638,662 imputed HD genotypes using Eagle (version 2.3.2; Loh et al., 2016), and subsequently imputing to WGS using minimac3 (Das et al., 2016). The average genotype concordance of imputation to WGS, defined as the proportion of correctly called SNPs versus all SNPs using a validation set of 175 Irish animals, was estimated to be 0.98 (Purfield et al., 2019).

Quality control edits were imposed on the imputed sequence genotypes within each breed, separately. Regions of poor WGS imputation accuracy, which could be due to local mis-assemblies or mis-orientated contigs, were removed. These regions were identified using an additional dataset of 147,309 verified parent progeny relations as described by (Purfield et al., 2019), which removed 687,352 SNPs from each breed. Then, all SNPs with a minor allele frequency (MAF) lower than 0.002 were removed. Following all SNP edits, 16,342,970, 17,733,147, 16,638,022, 17,803,135 and 17,762,681 autosomal SNPs remained for the analysis of the AA, CH, HE, LM, and SI populations, respectively.

3.4.4 Association Analyses

The association analyses were performed within each breed separately using a linear mixed model in the GCTA software (Yang et al., 2011). Autosomal SNPs from the original HD panel (i.e., 734,159 SNPs) were used to construct the genomic relationship matrix (GRM). The model used for the within-breed analysis was the following:

$$\mathbf{y} = \mu + \mathbf{xb} + \mathbf{u} + \mathbf{e},$$

where **y** is a vector of preadjusted phenotypes, μ is the overall mean, **x** is the vector of imputed genotypes, **b** is the additive fixed effect of the candidate SNP to be tested for association, $\mathbf{u} \sim N(0,\mathbf{G}\sigma_u^2)$ is the vector of additive genetic effects, where **G** is the genomic relationship matrix calculated from the imputed HD SNP genotypes, and σ_u^2 is the additive genetic variance, and $\mathbf{e} \sim N(0,\mathbf{I}\sigma_e^2)$ is the vector of random residual effects, with **I** representing the identity matrix and σ_e^2 the residual variance. Manhattan plots were created for each trait within each breed separately by using the QQman package (Turner, 2014) in R.

3.4.5 QTL Detection, Gene Annotation & Variance Explained

A genome-wide SNP significance threshold of $p \leq 1\times10^{-8}$ and a suggestive threshold of $p \leq 1\times10^{-5}$ were applied to each trait. SNPs in close proximity to each other (< 500 kb) were classified as being located within the same QTL. Genes within 500 kb of the most significant SNP in a peak above the genome-wide threshold were identified using Ensembl 94 (Zerbino et al., 2017) on the UMD 3.1 bovine genome assembly. Moreover, the functional consequence of all significantly associated SNPs was predicted using the Variant Effect Predictor tool (McLaren et al., 2016) from Ensembl. The Cattle QTLdb (https://www.animalgenome.org/cgi-bin/QTLdb/BT/index) was used to identify QTL that were known to be associated with other traits in cattle. To identify QTL regions that were suggestive in more than one breed, each chromosome was split into 1-kb genomic windows, and windows containing suggestive SNPs ($p \leq 1\times10^{-5}$) were compared across the breeds.

The proportion of genetic variance of a trait explained by a SNP was calculated as:

$$\frac{2p(1-p)a^2}{\sigma_g^2}$$

where p is the frequency of the minor allele, a is the allele substitution effect and σ_g^2 is the genetic variance of the trait in question.

3.4.6 Meta-analysis

Following the within-breed association analyses, meta-analyses were conducted for all traits across all five beef breeds using the weighted Z-score method in METAL (Willer et al., 2010); only SNPs that were included in the analyses of all of the individual breeds were considered here. METAL combines the p-values and the direction of SNP effects from individual analyses, and weights the individual studies based on the sample size to compute an overall Z-score:

$$Z = \frac{\Sigma_i z_i w_i}{\sqrt{\Sigma_i w_i^2}}$$

where w_i is the square root of the sample size of breed i, and z_i is the Z-score for breed i calculated as $z_i = \phi^{-1}\left(1 - \frac{P_i}{2}\right)\Delta_i$, where ϕ is the cumulative distribution function, and P_i and Δ_i are the P-value and direction of effect for breed i, respectively.

3.4.7 Conditional Analyses

The summary statistics from the individual analyses for the CH population were further used to conduct conditional analyses on BTA2 based on the Q204X mutation, which was previously reported to be associated with muscularity traits in cattle (Grobet et al., 1998). These analyses were undertaken for each trait in the CH population using the conditional and joint association analysis (COJO) method in GCTA (Yang et al., 2012). The Q204X mutation was included as a fixed effect in the association analysis model and the allele substitution effect of all remaining SNPs were re-estimated.

3.4.8 Pathway & Enrichment Analyses

Pathway analysis was conducted on all plausible candidate genes within a 500-kb region up- and downstream of SNPs that were discovered to be suggestively or significantly associated with each trait in each breed. For each gene list, DAVID 6.8 (Huang et al., 2008) was used to identify gene ontology (GO) terms and KEGG pathways which were significantly overrepresented ($p < 0.05$) by the set of genes. Enrichment analyses among the suggestive and significant SNPs were performed to estimate if the number of SNPs in each annotation class was greater than that expected by chance for each trait per breed (Bouwman et al., 2018); this was done separately per trait and per breed and was calculated as:

$$\text{Enrichment} = \frac{a}{b}\left[\frac{c}{d}\right]^{-1}$$

where a is the number of suggestive and/or significant SNPs in the annotation class of interest, b is the total number of suggestive and/or significant SNPs that were associated with the trait of interest, c is the total number of SNPs in the annotation class in the

association analysis, and d is the overall number of SNPs included in the association analysis.

3.5 Results

Summary statistics of the five linear type traits for each breed are in Appendix A1. Significant ($p \leq 1 \times 10^{-8}$) and/or suggestive ($p \leq 1 \times 10^{-5}$) SNPs were detected in all traits for the five breeds but the exact locations of these SNPs and the direction of the effects of these SNPs differed by breed. Manhattan plots for all the analyses are available in Appendix A1.

3.5.1 Within-breed Analyses

3.5.1.1 Angus

Whereas no significant SNPs were detected for any of the muscularity linear type traits in the AA population, suggestive SNPs ($p \leq 1 \times 10^{-5}$) were identified for all five traits. No genomic region was common to all five type traits (Appendix A2). However, there was some overlap in suggestive 1-kb windows between the traits DIT and TW; 11 windows contained SNPs of suggestive significance and the gene *EMILIN22* on BTA24 was identified within those windows for both traits. Nine genomic windows were associated with both the DL and WOW traits, i.e. on BTA6 (n = 2), BTA15 (n = 6), and BTA22 (n = 1). The windows on BTA15 contained suggestive SNPs that were located within the *UCP3* and *CHRDL2* genes.

Eighty-four SNPs within nine QTL were suggestively associated with the DHQ trait. Among these, the most strongly associated ($p = 3.34 \times 10^{-7}$) SNP was rs433492843

on BTA23 located in an intron of the *PTCHD4* gene (Table 3.1); it accounted for 0.002% of the genetic variance in this trait. A QTL on BTA1 was also strongly associated with DL with the most strongly associated SNP being rs465472414 ($p = 1.06 \times 10^{-6}$), which accounted for 0.08% of the genetic variance in this trait (Table 3.2). Other SNPs suggestively associated with DL were also identified within the *TMEM178A* gene on BTA11 and within the *UCP3* and *CHRDL2* genes on BTA15.

An intergenic SNP located on BTA29, rs109229230, was the most strongly associated ($p = 1.82 \times 10^{-7}$) with DIT (Table 3.3). Ninety-eight SNPs were suggestively associated with TW. The strongest QTL association with TW was on BTA13, on which 10 SNPs of suggestive significance were identified in a 1-Mb region (Table 3.4); rs137458299 displayed the strongest association ($p = 2.99 \times 10^{-7}$) and explained 0.9% of the genetic variation in TW. One hundred and seventy-three SNPs were associated with WOW in the AA population; among these 29.4% were located on BTA14 (Table 3.5) and the most strongly associated SNP, rs468048676, ($p= 2.34 \times 10^{-9}$), was an intergenic variant on BTA6.

Table 3.1 Location of the most significant QTL, limited to the top five per breed, which were associated with development of hind quarter and the genes located within these QTL within each breed

Breed	Chr	Start	End	Number of suggestive and significant SNPs	Most significant SNP	P-value	Allele frequency of + allele					Candidate genes within this QTL
							AA	CH	HE	LM	SI	
AA	1	72069526	130071811	8	120584401[a]	$3.69x10^{-6}$	0.940	0.940	0.061	0.949	0.951	*DLG1*, *FAM43A*, *APOD*, *OPA1*, *OSTN*, *GHSR*
	8	72017409	73103211	16	72569526[b]	$3.25x10^{-6}$	0.276	0.565	0.264	0.332	0.687	*ADAM28*
	10	5155837	6179062	11	5655837[d]	$7.60x10^{-6}$	0.032	0.133	0.000	0.073	0.036	*SFXN1*, *DRD1*[e]
	23	20541063	21541072	2	21041063[b]	$3.34x10^{-7}$	0.995	0.977	0.000	0.000	0.005	*PTCHD4*[e]
	24	35913368	37107434	16	36567715[a]	$8.07x10^{-7}$	0.918	0.050	0.958	0.941	0.903	*ENOSF1*, *ADCYAP1*
CH	2	35194	10711228	5128	6808074[a]	$9.07x10^{-49}$	0.000	0.079	0.000	0.028	0.004	*WDR75*[f], *ASNSD1*[f], *ARHGEF4*[f], *MYO7B*[f], *IWS1*[e], *NAB1*[f], *MFSD6*[f], *MSTN*[f], *PMS1*[f], *ORMDL1*[e], *COL3A1*[f], *COL5A2*[f], *ANKAR*[f], *SLC40A1*[f]
	4	89299122	90487119	12	89799122[a]	$4.67x10^{-6}$	0.000	0.007	0.002	0.000	0.992	*GPR37*[e], *POT1*
	14	33353270	34360874	4	33855595[a]	$2.20x10^{-7}$	0.013	0.026	0.054	0.054	0.982	*PREX2*
	21	34538609	35572213	47	35062974[b]	$9.70x10^{-7}$	0.617	0.521	0.594	0.000	0.451	*UBL7*, *SEMA7A*, *PML*
	28	15669176	18516719	77	16273851[a]	$2.95x10^{-8}$	0.574	0.827	0.580	0.416	0.520	*ANK3*, *CDK1*, *RHOBTB1*
HE	1	68813144	73614763	228	69348458[b]	$5.23x10^{-7}$	0.764	0.531	0.237	0.317	0.480	*KALRN*[e], *ITGB5*[e]
	7	83591608	84673367	26	84122153[a]	$3.16x10^{-7}$	0.177	0.851	0.913	0.000	0.910	*ACOT12*, *ATG10*,
	12	48965278	50129513	4	49465278[a]	$1.84x10^{-6}$	0.020	0.995	0.996	0.983	0.000	*KLF12*
	13	2761800	3761897	2	3261897[a]	$7.44x10^{-7}$	0.681	0.761	0.795	0.269	0.695	*MRPL33*
	23	12056019	13111217	5	12571991[b]	$1.09x10^{-6}$	0.220	0.000	0.094	0.885	0.100	*GLO1*, *GLP1R*
LM	2	4293223	12640428	2610	6622189[a]	$3.22x10^{-30}$	0.491	0.447	0.250	0.043	0.799	*WDR75*, *ASNSD1*[f], *MYO7B*, *IWS1*, *NAB1*[f], *MFSD6*[e], *MSTN*[f], *PMS1*[f], *ORMDL1*[e], *COL3A1*[f], *COL5A2*[f], *ANKAR*[f], *SLC40A1*[f], *ZNF804A*
	2	13916060	14987913	113	14446207[bc]	$4.83x10^{-8}$	0.567	0.562	0.374	0.394	0.523	*PDE1A*[e], *PPP1R1C*
	5	59612855	60696179	7	60112855[a]	$1.58x10^{-7}$	0.000	0.000	0.000	0.997	0.996	*AMDHD1*
	11	12163298	13181229	7	12676690[a]	$1.56x10^{-6}$	0.486	0.283	0.000	0.231	0.163	*CYP26B1*, *DYSF*, *ZNF638*
SI	6	76112104	77238886	16	76612104[a]	$1.28x10^{-6}$	0.997	0.993	0.000	0.000	0.993	*ENSBTAG00000043492*
	7	103195804	104247192	3	103695804[b]	$1.78x10^{-6}$	0.000	0.000	0.000	0.005	0.005	*SLCO4C1*[e], *SLC06A1*
	9	65068999	66303927	4	65712927[a]	$4.56x10^{-7}$	0.000	0.000	0.000	0.003	0.996	*TBX18*[e], *MRAP2*
	23	41204249	42287354	9	41747427[a]	$1.48x10^{-7}$	0.009	0.008	0.994	0.962	0.990	*JARID2*
	25	22649471	23937019	6	23400365[a]	$1.00x10^{-7}$	0.985	0.993	0.000	0.993	0.998	*AQP8*, *ZKSCAN2*

SNP classification: [a]intergenic, [b]intron, [c]upstream gene variant [d]downstream gene variant

Significance of SNPs within genes: [e]gene contained at least one suggestive SNP, [f]gene contained at least one significant SNP.

Table 3.2 Location of the most significant QTL, limited to the top 5 per breed which were associated with development of loin, and the genes located within these QTL within each breed

Breed	Chr	Start	End	Number of suggestive and significant SNPs	Most significant SNP	P-value	Allele frequency of + allele					Candidate genes within this QTL
							AA	CH	HE	LM	SI	
AA	1	39155170	40155196	3	39655188[a]	$1.06x10^{-6}$	0.003	0.015	0.000	0.967	0.976	
	10	5741208	6752759	54	6241208[a]	$3.45x10^{-6}$	0.181	0.105	0.090	0.773	0.872	*GCNT4*, *HMGCR*
	11	21539414	22560915	4	22049725[c]	$6.26x10^{-6}$	0.032	0.928	0.904	0.000	0.068	*CDKL8*, *MAP4K3*
	15	53716499	55635294	27	55096343[a]	$1.43x10^{-6}$	0.993	0.000	0.000	0.000	0.996	*RAB6A*, *MRPL48*, *UCP2*, *UCP3*[d], *PPME1*, *NEU3*
	19	12894306	13934964	9	13430257[a]	$2.23x10^{-6}$	0.045	0.088	0.969	0.102	0.044	*USP32*, *MYO19*, *ACACA*
CH	1	142705947	143712126	78	143205947[c]	$3.50x10^{-6}$	0.000	0.010	0.005	0.972	0.993	*BACE2*, *RIPK4*, *PRDM15*, *C2CD2*
	2	37387	8714844	1728	6808074[a]	$1.19x10^{-26}$	0.000	0.079	0.000	0.028	0.004	*WDR75*[e], *ASNSD1*[e], *MFSD6*[e], *MSTN*[e], *PMS1*[e], *ORMDL1*, *COL3A1*[e], *COL5A2*[e], *ANKAR*[e], *SLC40A1*[e]
	10	84923776	85960329	2	85423776[a]	$4.45x10^{-7}$	0.000	0.007	0.000	0.000	0.024	*PSEN1*, *ACOT2*, *ACOT4*, *DNAL1*, *ZNF410*, *FAM161B*, *COQ6*
	16	32025893	33120555	127	32558878[a]	$4.12x10^{-7}$	0.975	0.884	0.135	0.898	0.093	*SMYD3*, *KIF26B*, *EFCAB2*
	24	45933369	46937392	5	46437392[b]	$1.20x10^{-6}$	0.000	0.006	0.012	0.000	0.019	*PSTPIP2*, *ST8SIA5*
HE	2	79648265	81037622	15	80168803[b]	$1.42x10^{-6}$	0.000	0.994	0.996	0.987	0.997	*STAT1*, *MYO1B*[d], *NABP1*
	4	3339555	4351559	4	3851559[a]	$1.16x10^{-7}$	0.014	0.034	0.025	0.884	0.952	*ENSBTAG00000044810*
	11	17934758	18942324	3	18442324[a]	$1.11x10^{-6}$	0.876	0.011	0.011	0.964	0.988	*CRIM1*
	16	75812761	76823259	3	76312761[a]	$1.94x10^{-6}$	0.911	0.000	0.030	0.000	0.937	*ENSBTAG00000044497*
	29	9233633	10275049	26	9733633[a]	$1.80x10^{-6}$	0.650	0.417	0.304	0.378	0.801	*EED*, *SYTL2*, *CREBZF*, *TMEM126A*, *TMEM126B*
LM	2	5545383	8287013	748	6747317[a]	$6.69x10^{-10}$	0.198	0.409	0.617	0.073	0.511	WDR75[d], ASNSD1[d], MFSD6, MSTN[d], PMS1[d], ORMDL1, *COL3A1*, *COL5A2*[d], *ANKAR*[e], *SLC40A1*[d]
	3	99009887	100073842	5	99509887[b]	$5.78x10^{-7}$	0.041	0.000	0.047	0.934	0.883	*CYP4X1*, *CYP4A22*
	5	71738658	72751064	6	72238658[b]	$2.70x10^{-6}$	0.000	0.819	0.163	0.251	0.830	*SYN3*, *TIMP3*, *MGC137211*, *MGC137014*, *LARGE1*[d]
	6	110629080	112106706	11	111176155[a]	$3.30x10^{-7}$	0.126	0.868	0.898	0.830	0.211	*HS3ST1*
	12	32803688	33961156	7	33461156[b]	$5.12x10^{-7}$	0.000	0.000	0.000	0.997	0.975	*USP12*, *SHISA2*
SI	7	58665425	60243933	10	59590500[a]	$3.37x10^{-7}$	0.002	0.016	0.046	0.000	0.990	*SH3RF2*[d]
	8	90813081	91857894	14	91313081[a]	$5.24x10^{-7}$	0.957	0.627	0.064	0.346	0.291	*SPIN1*, *FBXW12*
	14	79527472	81070931	18	80091780[b]	$3.45x10^{-7}$	0.908	0.052	0.981	0.114	0.070	*E2F5*
	17	69082585	70268366	4	69646862[b]	$1.04x10^{-7}$	0.975	0.010	0.987	0.992	0.003	*PITPNB*[d]
	22	33231032	34644069	14	34044822[b]	$9.77x10^{-9}$	0.000	0.960	0.946	0.000	0.004	*FAM19A1*, *SUCLG2*[e], *KBTBD8*

SNP classification: [a]intergenic, [b]intron, [c]downstream gene variant

Significance of SNPs within genes: [e]gene contained at least one suggestive SNP, [f]gene contained at least one significant SNP.

Table 3.3 Location of the most significant QTL, limited to the top 5 per breed, which were associated with development of inner thigh, and the genes located within these QTL within each breed

Breed	Chr	Start	End	Number of suggestive and significant SNPs	Most significant SNP	P-value	Allele frequency of + allele					Candidate genes within this QTL
							AA	CH	HE	LM	SI	
AA	1	146687440	147685998	7	147187440[b]	$9.55x10^{-7}$	0.988	0.712	0.073	0.179	0.631	*TRAPPC10*, *COL18A1*, *SLC19A1*, *PCBP3*[e], *COL6A1*, *COL6A2*
	4	69229353	70999401	38	70373241[a]	$4.19x10^{-7}$	0.991	0.015	0.990	0.014	0.900	*HOXA1*, *HOXA2*, *HOXA3*, *HOXA4*, *HOXA5*, *HOXA6*, *HOXA7*, *HOXA9*, *HOXA10*, *HOXA11*, *HOXA13*
	24	36998437	38030390	4	37530390[bd]	$2.24x10^{-7}$	0.977	0.000	0.014	0.036	0.983	*NDC80*, EMILIN2[e], *MYOM1*, *MYL12A*, *MYL12B*
	25	35042698	36122096	6	35542698[a]	$3.50x10^{-7}$	0.998	0.000	0.000	0.000	0.002	*POLR2J*, *MYL10*, *ALKBH4*, *COL26A1*
	29	23787949	24826548	7	24290699[a]	$1.82x10^{-7}$	0.048	0.043	0.937	0.939	0.919	*SLC6A5*, *PRMT3*
CH	2	7850	10711228	5075	6808074[a]	$9.07x10^{-49}$	0.000	0.079	0.000	0.028	0.005	WDR75[f], *ASNSD1*[f], *NAB1*[f], *MFSD6*[f], *MSTN*[f], *PMS1*[f], *ORMDL*[e] *COL3A1*, *COL5A2*[f], *ANKAR*[f], *SLC40A1*[f]
	14	33353270	34360874	4	33855595[a]	$2.20x10^{-7}$	0.013	0.026	0.054	0.054	0.018	*ARFGEF1*, *CPA6*, *PREX2*
	14	67849241	68850085	4	6834924[a]	$2.69x10^{-6}$	0.000	0.062	0.000	0.000	0.000	*STK3*, *KCNS2*, *POP1*, *RPL30*, *MATN2*
	16	60443313	62320499	4	60943313[a]	$6.72x10^{-6}$	0.904	0.230	0.063	0.853	0.821	*RASAL2*, *ANGPTL1*, *TOR3A*, *ABL2*, *SOAT1*
	29	21313583	22460213	38	21917306[a]	$3.58x10^{-6}$	0.000	0.989	0.000	0.978	0.000	*GAS2*, *FANCF*
HE	4	3924012	5000928	3	4424012[a]	$4.05x10^{-7}$	0.755	0.032	0.082	0.000	0.939	*ENSBTAG00000023806*
	7	72100887	73679002	67	72608186[b]	$9.09x10^{-7}$	0.000	0.998	0.046	0.000	0.995	*EBF1*[e], *ADRA1B*
	13	6038290	8341939	9	7003978[a]	$1.84x10^{-7}$	0.000	0.000	0.002	0.004	0.000	*ESF1*, *NDUFAF5*, *FLRT3*
	14	57257740	58269158	3	57757740[a]	$7.98x10^{-7}$	0.262	0.122	0.758	0.104	0.000	*TRHR*
	25	37995664	39005834	3	38505834[c]	$2.68x10^{-7}$	0.160	0.928	0.097	0.056	0.071	*LMTK2*, *PMS2*, *EIF2AK1*, *USP42*
LM	2	313343	3758925	102	3226165[a]	$5.94x10^{-10}$	0.907	0.032	0.997	0.066	0.057	*NIPA1*, *NIPA2*[e], *ARHGEF4*
	2	4973733	11101064	2441	6747317[a]	$2.20x10^{-28}$	0.802	0.409	0.617	0.074	0.490	*WDR75*, *ASNSD1*[f], *NAB1*[f], *MFSD6*[e], *MSTN*[f], *PMS1*[f], *ORMDL1*[e], *COL3A1*[e], *COL5A2*[f], *ANKAR*[f], *SLC40A1*[f]
	2	11556240	12618550	36	12116324[a]	$1.36x10^{-8}$	0.308	0.402	0.231	0.195	0.556	*ZNF804A*
	2	13935604	14957932	55	14447892[bc]	$1.67x10^{-6}$	0.433	0.562	0.374	0.394	0.524	*PDE1A*[e], *PPP1R1C*, *NEUROD1*
	4	23124509	24137261	59	23630609[b]	$2.52x10^{-7}$	0.096	0.027	0.996	0.994	0.010	*AGMO*[e], *MEOX2*
SI	1	22558133	23622641	49	23117106[a]	$5.21x10^{-6}$	0.949	0.043	0.866	0.060	0.870	*ENSBTAG00000046369*
	14	79492849	80610485	17	80090294[b]	$2.33x10^{-7}$	0.104	0.052	0.980	0.886	0.072	*E2F5*
	17	21615210	22651417	8	22115210[a]	$1.47x10^{-6}$	0.474	0.619	0.639	0.359	0.297	*ENSBTAG00000044703*
	21	17057917	18336523	18	17836523[a]	$5.30x10^{-8}$	0.003	0.010	0.000	0.000	0.997	*ENSBTAG00000045960*
	22	33706576	34737653	10	34236342[b]	$2.08x10^{-6}$	0.000	0.995	0.000	0.987	0.002	*SUCLG2*[e], *KBTBD8*

SNP classification: [a]intergenic, [b]intron, [c]upstream gene variant, [d]downstream gene variant

Significance of SNPs within genes: [e]gene contains at least one suggestive SNP, [f]gene contains at least one significant SNP

Table 3.4 Location of the most significant QTL, limited to the top 5 per breed, which were associated with thigh width, and the genes located within these QTL within each breed

Breed	Chr	Start	End	Number of suggestive and significant SNPs	Most significant SNP	P-value	Allele frequency of + allele					Candidate genes within this QTL
							AA	CH	HE	LM	SI	
AA	11	53741952	54849531	2	54349531[a]	$2.65x10^{-6}$	0.017	0.047	0.000	0.997	0.994	*CTNNA2*
	13	78082912	79223584	10	78722523[b]	$2.99x10^{-7}$	0.191	0.232	0.866	0.157	0.108	*ZNFX1*, *B4GALT5*, *SLC9A8*[f], *UBE2V1*
	14	22612620	23619374	3	23116129[a]	$1.99x10^{-6}$	0.026	0.994	0.074	0.962	0.971	*OPRK1*, *ATP6V1H*, *RGS20*
	16	1654734	2669694	4	2169248[a]	$5.14x10^{-6}$	0.986	0.980	0.985	0.056	0.979	*ETNK2*, *GOLT1A*, *PPP1RI5B*, *PIK3C2B*, *NFASC*
	16	63066185	64073838	2	63066185[d]	$4.07x10^{-6}$	0.027	0.948	0.991	0.072	0.000	*ACBD6*, *STX6*, *MR1*
CH	2	7850	10186234	1860	6808074[a]	$4.09x10^{-25}$	0.000	0.079	0.000	0.028	0.004	*WDR75*[g], *ASNSD1*[g], *NAB1*[g], *MFSD6*[g], *MSTN*[g], *PMS1*[g], *ORMDL1*, *COL3A1*[g], *COL5A2*[g], *ANKAR*[g], *SLC40A1*[g]
	9	12470065	13731582	11	12970065[e]	$8.65x10^{-9}$	0.000	0.013	0.000	0.003	0.000	*KHDC3L*[g], *EEF1A1*
	20	61727533	63246384	5	62296495[b]	$7.78x10^{-9}$	0.997	0.010	0.000	0.000	0.000	*CTNND2*[f], *DAP*, *FAM173B*
	28	9234253	24087892	1025	16275379[a]	$4.17x10^{-9}$	0.581	0.829	0.476	0.430	0.523	*ANK3*[f], *CDK1*, *RHOBTB1*, *EGR2*
	28	24331178	40249741	1322	26134688[a]	$3.87x10^{-9}$	0.388	0.810	0.424	0.591	0.486	*SIRT1*, *MYPN*[f], *DNA2*[f], *SLC25A16*, *SRGN*[f], *COL13A1*[f], *AIFM2*, *ADAMTS14*[f], *SGLP1*, *PCBD1*, *SPOCK2*[f], *ANAPC16*[f], *DDIT4*, *MYOZ1*
HE	5	17326996	18595093	3	18095093[a]	$1.75x10^{-6}$	0.000	0.996	0.003	0.985	0.000	*C5H12orf50*, *C5H12orf29*, *CEP290*
	7	75039465	75039465	2	75539465[d]	$4.24x10^{-6}$	0.990	0.000	0.997	0.000	0.000	*GABRA6*[f]
	10	87480123	88487031	3	87980123[b]	$1.72x10^{-6}$	0.959	0.034	0.020	0.981	0.033	*TTLL5*[f],*TGFB3*
	20	35840468	37412173	5	36340468[a]	$1.30x10^{-6}$	0.342	0.589	0.421	0.230	0.328	*LIFR*, *EGFLAM*, *GDNF*
	20	39512041	41162914	10	40272197[b]	$1.01x10^{-6}$	0.015	0.985	0.985	0.994	0.003	C1QTNF3, ADAMTS12[f]
LM	2	4973607	10670961	1526	7772897[a]	$2.88x10^{-15}$	0.617	0.571	0.000	0.116	0.642	*WDR75*[f], *ASNSD1*[g], *NAB1*[f], *MFSD6*[f], *MSTN*[f], *PMS1*[f], *ORMDL1*, *COL3A1*, *COL5A2*, *ANKAR*[g], *SLC40A1*[g]
	2	11570479	12640428	43	12083258[a]	$1.01x10^{-7}$	0.561	0.205	0.057	0.097	0.809	*ZNF804A*
	2	13916060	14952210	55	14450953[b]	$1.92x10^{-6}$	0.434	0.565	0.374	0.394	0.477	*PDE1A*[f], *PPP1R1C*
	9	36716988	37719425	4	37216988[a]	$6.31x10^{-7}$	0.245	0.649	0.547	0.247	0.286	*HS3ST5*
	27	35767035	36785436	7	36267035[c]	$8.03x10^{-6}$	0.057	0.000	0.998	0.997	0.000	*SFRP1*, *GOLGA7*, *GPAT4*, *ANK1*
SI	9	32996000	34105290	5	33604527[b]	$4.82x10^{-7}$	0.969	0.980	0.981	0.985	0.019	*NEPN*, *GOPC*
	14	76746937	78193017	13	77271966[a]	$1.65x10^{-7}$	0.712	0.107	0.052	0.916	0.085	*MMP16*[f], *SLC2A5*
	24	17511696	18515510	3	18015356[a]	$8.99x10^{-7}$	0.461	0.449	0.299	0.506	0.378	*ENSBTAG00000045320*, *ENSBTAG00000011094*
	26	35041033	36071133	3	35541033[a]	$1.97x10^{-6}$	0.000	0.006	0.004	0.994	0.995	*AFAP1L2*, *ABLIM1*, *ATRNL1*
	29	47398869	48688066	30	47898869[a]	$5.42x10^{-7}$	0.790	0.197	0.886	0.850	0.108	*CCND1*, *FGF19*, *FGF4*

SNP classification: [a]intergenic, [b]intron, [c]upstream gene variant, [d]downstream gene variant, [e]synonymous gene variant
Significance of SNPs within genes: [f]gene contains at least one suggestive SNP, [g]gene contains at least one significant SNP.

Table 3.5 Location of the most significant QTL, limited to the top 5 per breed, which were associated with width of withers, and the genes located within these QTL within each breed

Breed	Chr	Start	End	Number of suggestive and significant SNPs	Most significant SNP	P-value	Allele frequency of + allele					Candidate genes within this QTL
							AA	CH	HE	LM	SI	
AA	1	40642399	41670931	3	41142399[b]	$2.49x10^{-6}$	0.997	0.998	0.009	0.991	0.000	*EPHA6*[d]
	6	90992982	92064889	10	91513217[a]	$2.34x10^{-9}$	0.008	0.019	0.009	0.995	0.005	*EREG*, *AREG*, *RCHY1*
	14	9414135	10414890	9	9914135[b]	$3.85x10^{-6}$	0.016	0.936	0.144	0.969	0.951	*TG*, *KCNQ3*[d]
	14	17744843	18889304	16	18244843[b]	$1.59x10^{-6}$	0.957	0.007	0.977	0.000	0.977	*ANXA13*, *KLHL38*, *FBXO32*
	15	66035436	67113172	4	66535436[b]	$3.42x10^{-7}$	0.004	0.996	0.930	0.003	0.992	*APIP*, *PDHX*
CH	2	218127	8714844	1227	6808074[a]	$2.02x10^{-21}$	0.000	0.079	0.000	0.028	0.004	*WDR75*[e], *ASNSD1*[d], *NAB1*[e], *MFSD6*[e], *MSTN*[e], *PMS1*[e], *ORMDL1*, *COL3A1*[e], *COL5A2*[e], *ANKAR*[e], *SLC40A1*[e]
	4	63740415	65180447	6	64680447[a]	$3.32x10^{-7}$	0.000	0.003	0.000	0.000	0.991	*NT5C3A*, *KBTBD2*
	14	67870173	68870239	7	68370173[a]	$2.98x10^{-7}$	0.028	0.060	0.965	0.053	0.024	*STK3*, *KCNS2*, *POP1*, *RPL30*, *MATN2*
	28	15669176	17454059	6	16943776[a]	$1.52x10^{-8}$	0.640	0.868	0.525	0.690	0.729	*ANK3*, *CDK1*, *RHOBTB1*
	28	31765108	33148059	7	32634467[a]	$1.61x10^{-7}$	0.367	0.889	0.664	0.511	0.468	*KCNMA1*
HE	7	63750754	64814905	15	64309256[b]	$1.46x10^{-7}$	0.022	0.949	0.008	0.011	0.009	*RPS14*, *MYOZ3*, *ZNF300*, *GPX3*[d], *ANXA6*
	11	89579990	90599173	5	90079990[a]	$5.13x10^{-6}$	0.664	0.280	0.163	0.800	0.217	*RNF144A*, *RSAD2*
	18	11547513	12549350	5	12047513[a]	$4.85x10^{-6}$	0.479	0.565	0.507	0.532	0.446	*GSE1*, *IRF8*, *FOXC2*
	20	25429449	27124758	9	25929649[a]	$3.38x10^{-6}$	0.532	0.000	0.559	0.000	0.480	*FST*
	26	32810942	33810987	3	33310942[b]	$2.78x10^{-6}$	0.000	0.000	0.998	0.000	0.000	*GPAM*, *ACSL5*
LM	2	5547713	8495179	725	6622189[a]	$8.77x10^{-12}$	0.490	0.447	0.250	0.043	0.800	*WDR75*, *ASNSD1*[d], *NAB1*, *MFSD6*, *MSTN*, *PMS1*[d], *ORMDL1*, *COL3A1*, *COL5A2*[e], *ANKAR*[d], *SLC40A1*[e]
	2	9053737	10527711	26	9559686[a]	$1.53x10^{-6}$	0.464	0.504	0.213	0.127	0.456	*ITGAV*, *ZC3H15*
	2	13916060	14957655	68	14450953[b]	$6.10x10^{-8}$	0.433	0.564	0.626	0.394	0.479	*PDE1A*[d], *PPP1R1C*, *NEUROD1*
	2	20539841	21539862	3	21039862[a]	$6.85x10^{-10}$	0.000	0.000	0.000	0.998	0.000	*HOXD1*, *HOXD3*, *HOXD4*, *HOXD9*, *HOXD10*, *HOXD11*, *HOXD12*, *HOXD13*
	6	19014612	21121181	13	19817910[a]	$8.15x10^{-9}$	0.000	0.000	0.008	0.995	0.000	*NPNT*, *GSTCD*[d]
SI	1	79028842	80104503	3	79604503[b]	$3.81x10^{-7}$	0.022	0.040	0.000	0.964	0.004	*LPP*[d]
	4	57566849	58584434	5	58084434[d]	$6.35x10^{-8}$	0.902	0.261	0.457	0.248	0.229	*IMMP2L*[d], *LRRN3*
	9	32996000	34471196	16	33604527[b]	$5.67x10^{-8}$	0.969	0.979	0.019	0.985	0.019	*NEPN*, *GOPC*[d]
	12	28061051	29073791	3	28561051[d]	$7.49x10^{-7}$	0.008	0.020	0.084	0.016	0.998	*PDS5B*
	20	65137216	66168943	4	65668943[a]	$1.76x10^{-7}$	0.000	0.009	0.000	0.977	0.007	*FASTKD3*, *ADCY2*[d]

SNP classification: [a]intergenic, [b]intron, [c]downstream gene variant

Significance of SNPs within genes: [d]gene contains at least one suggestive SNP, [e]gene contains at least one significant SNP.

3.5.1.2 Hereford

No significant SNPs were detected for any of the muscularity linear type traits in the HE population, although suggestive SNPs were identified for all five traits. However, no genomic window was common to all five type traits (Appendix A2); six 1-kb windows i.e. on BTA5 (n = 1), BTA7 (n = 4), and BTA25 (n = 1) were shared between DHQ and DIT with three 1-kb regions on BTA20 shared between DIT and TW.

Three hundred and eleven SNPs were suggestively associated with DHQ. The strongest association with DHQ was located within a 1-Mb QTL on BTA7 where 26 SNPs of suggestive significance were identified (Table 3.1). The intergenic SNP, rs446625612 ($p = 1.16 \times 10^{-7}$) was the most strongly associated with DL and located within a QTL on BTA4 encompassing the *ENSBTAG00000044810* gene. Most interestingly, the strongest association within the QTL on BTA2 with DL was an intronic variant, which explained 0.7% of the genetic variance and was located within the muscle related gene *MYO1B*.

In total, 155 SNPs were suggestively or significantly associated with DIT, and 43% of these were located within a 1-Mb QTL on BTA7 (Table 3.3) where a number of significant SNPs were located within the *EBF1* gene. For TW, four putative candidate genes were identified (Table 3.4): *GABRA6* on BTA7, *TTLL5* on BTA10, and both *ADAMTS12* and *GDNF* on BTA20. The SNP, rs380761563, which displayed the strongest association with WOW, explained 1% of the genetic variance and was located in an intron of the gene *TNIP1* on BTA7 (Table 3.5).

3.5.1.3 Charolais

There were 483 1-kb suggestive genomic windows common to all five type traits in the CH population (Appendix A2), among which the vast majority (n = 482) were located on BTA2 in a region encompassing the *MSTN* gene. The final region that was shared between all five traits was on BTA11. More overlaps were found for DHQ and DIT with 904 windows being common to just these two traits, 146 windows common to DHQ, DIT, and DL, 304 windows common to DHQ, DIT, DL, and TW, and 178 windows common to DHQ, DIT, and TW. The majority of all these windows were also located on BTA2.

For each of the muscularity linear traits, we identified a QTL on BTA2 in the CH population. DHQ had the largest number of associated SNPs, i.e. 3707 suggestive and 1851 significant SNPs (Table 3.1), all of which were located on BTA2 within a single QTL between positions 0.35 and 9.79 Mb. In total, 41 genes including *MFSD6*, *MSTN*, and *MYO7B* were located in this QTL. For DIT, a 10-Mb QTL on BTA2 was identified that contained 5075 SNPs, of which 1796 had a p-value that met the significance threshold (Table 3.3), whereas 178 SNPs on BTA2 in the region between 54.1 and 86.1 Mb were significantly associated with TW (Table 3.4). The same SNP, an intergenic variant rs799943285, showed the strongest association with all traits. The well-known Q204X mutation within the *MSTN* gene was significantly associated with DHQ, DIT and TW, and this SNP explained 4.9, 0.05, and 0.01% of the genetic variation of each trait, respectively.

In the conditional analyses within the CH population, where the Q204X mutation was included as a fixed effect in the model, the most significant SNPs from the original analyses of each trait generally reduced in significance. The most

significant SNP for all traits in the original analyses was rs799943285 (p-value ranging from 9.07 x 10^{-49} for DIT and DHQ to 2.02 x 10^{-21} for WOW). In the conditional analyses, this SNP was non-significant for DL, TW, and WOW but remained suggestive for both DIT (p = 4.02 x 10^{-6}) and DHQ (p = 4.62 x 10^{-6}). The most significant SNP in the conditional analyses of DHQ, DL, DIT, and TW was rs41638272, which is an intergenic SNP located 10 kb from the *SLC40A1* gene; this SNP was significant in the original analyses but its significance actually increased when the Q204X mutation was included as a fixed effect. The most significant SNP in the conditional analysis of WOW was an intergenic variant, rs457456302 (p = 4.78 x 10^{-10}) that was located 0.1 Mb from the *MSTN* gene.

3.5.1.4 Limousin

There were 164 1-kb suggestive genomic regions that were common across all muscularity traits in the LM population (Appendix A2); another 232 regions were common to the three traits DHQ, DIT, and TW, while 326 were common to just DHQ and DIT. All five traits had significant QTL located on BTA2, with four genes common to all traits located within these QTL, namely *ASNSD1*, *GULP1*, *SLC40A1*, and *ANKAR*.

For DHQ, there were 2983 SNPs above the suggestive threshold and most of these (n = 2610) were located in a single QTL on BTA2. The most significant SNP, rs211140207 (p = 3.22 x 10^{-30}), was located within an 8-Mb QTL on BTA2 that contains 20 genes (Table 3.1). The Q204X stop-gain mutation (rs110344317) located within this QTL was significantly associated with DHQ and accounted for 2.4% of the genetic variation in this trait, although the allele frequency of the favourable mutation was only 0.02% in the LM population. The well-known *MSTN* mutation in the

Limousin breed, F94L (MAF = 0.3798), did not meet the suggestive threshold for association with any of the traits. Similar to DHQ, a QTL located between 4.9 and 11 Mb on BTA2 was associated with both DIT (Table 3.3) and TW (Table 3.4). In total, 2441 and 1526 SNPs were above the suggestive threshold within this QTL on BTA2, and the variant rs110344317, which was significantly associated with DHQ, was also significantly associated with both DIT and TW. For the DL trait, 748 SNPs were suggestively associated and located between 55.4 and 82.8 Mb on BTA2. The most significant SNP associated with DL (rs379791493; $p = 6.69 \times 10^{-10}$) was also the most significantly associated SNP with DIT ($p = 2.20 \times 10^{-28}$). The most significant SNP associated with WOW, rs211140207, ($p = 8.77 \times 10^{-12}$), was an intergenic SNP that accounted for 0.4% of the genetic variance in this trait and was located in a QTL (between 5.9 and 8.4 Mb) that included 724 other significantly-associated SNPs (Table 3.5).

Suggestive QTL were also detected on autosomes other than BTA2 for all traits in the LM population except for DIT. A small QTL on BTA11 containing seven suggestive SNPs was associated with DHQ. The SNP with the strongest association, rs43666945 ($p = 1.56 \times 10^{-6}$), was an intergenic SNP located 2.2 Mb from the *DYSF* gene. Both DHQ and DL had suggestively associated QTL on BTA5. The most strongly associated SNP for DHQ ($p = 1.58 \times 10^{-7}$) was an intergenic SNP, rs718375830, located within a QTL between positions 59.6 and 60.6 Mb, whereas the most strongly associated SNP with DL ($p = 2.70 \times 10^{-6}$) was also an intergenic SNP, rs109909829, but was located within a QTL between 71.7 to 72.8 Mb.

3.5.1.5 Simmental

For the SI breed, only a few suggestive 1-kb genomic regions overlapped for more than two traits. Sixteen 1-kb windows were suggestively associated with both DHQ and DL, eight of which were located on BTA6, seven on BTA22, and one on BTA18 (Appendix A2). Five 1-kb windows on BTA23 and one on BTA4 were common to both DHQ and DIT, while another 15 suggestive windows were associated with DHQ and WOW, 12 of which were located on BTA22.

The intergenic SNP, rs437686690 on BTA25, was the most strongly associated ($p = 1.00 \times 10^{-7}$) with DHQ in the SI population and accounted for 0.6% of the genetic variance in DHQ (Table 3.1). In total, 199 SNPs were associated with DL in the SI population, among which four met the significance threshold. The most significant SNP, rs482545354 ($p = 9.77 \times 10^{-9}$), was located in an intronic region of the *SUCGL2* gene (Table 3.2) on BTA22. Although 194 SNPs were suggestively associated with DIT, only one, i.e., rs798946118 ($p = 5.30 \times 10^{-8}$), achieved the significance threshold which was located on BTA21 within a 1-Mb block containing 17 other suggestive SNPs (Table 3.3) and accounted for 0.6% of the genetic variance of DIT. The largest 1-Mb QTL associated with TW was located on BTA29 and contained 30 suggestive SNPs (Table 3.4). QTL putatively associated with WOW were located on BTA1, 4, 9, 12, and 20 (Table 3.14) where the most significant SNP, rs801295753 ($p = 5.67 \times 10^{-8}$), was an intronic SNP on BTA9 located within both the *ROS1* and *ENSBTAG000000039574* genes.

3.5.2 Meta-analyses

Within each of the five meta-analyses (Appendix A3), a strong association peak on BTA2 around the *MSTN* gene was detected, which is consistent with the individual association results identified in the CH and LM populations. For DIT, TW, and WOW, the most significantly associated SNP was the intergenic SNP, rs799943285 ($p = 5.51\times10^{-24}$), which was previously identified as the most strongly associated SNP in the CH population for each of these traits. This variant, rs799943285, was also the most significantly associated with DL in the meta-analysis, whereas the most significantly associated SNP with DHQ, rs482419628 ($p = 2.06 \times 10^{-47}$), was located further downstream on BTA2 within 5 kb of the *ASNSD1* gene.

Although the QTL on BTA2 was the most strongly associated with each of the traits analysed, we also identified several other QTL associated with muscularity. In the meta-analysis of DHQ, the most strongly associated SNP on BTA11, rs43666945 ($p = 1.93 \times 10^{-7}$), was previously identified as being associated with DHQ in the LM population, but the level of significance increased in the meta-analysis and the QTL contained three times the number of suggestive SNPs compared to that found for the LM breed only. A 1-Mb QTL on BTA7 containing the *SPRY4* and *FGF1* genes was associated with both DL and WOW in the meta-analysis; the most significant SNPs in this QTL, however, differed according to trait (Appendix A3).

3.5.3 Enrichment of SNPs

With the exception of WOW in the AA population, intergenic SNPs were the most common annotation class of SNPs that were significantly associated with all traits in all breeds. The 3' UTR class was enriched for all traits in the CH and LM populations,

whereas there were more downstream gene variants significantly associated with DHQ and DL in the AA, CH and HE populations, and with TW in the CH, HE, and SI populations than expected by chance (Table 3.6). The intronic class of SNPs was enriched for all five traits in HE, for four traits (DHQ, DL, TW, and DIT) in SI, three traits in both AA (DHQ, DL, and WOW) and CH (DL, TW, and WOW) and two traits in LM (DHQ and DIT).

3.5.4 Gene Ontology and KEGG Pathways

Several GO terms and KEGG pathways were over-represented by the genes identified in each analysis, although this tended to differ per breed and per trait especially in the smaller AA, HE, and SI populations. In CH and LM, five GO terms were associated with each trait: skin development (GO:0043588), collagen fibril organisation (GO:0030199), extracellular matrix structural constituent (GO:0005201), cellular response to amino acid stimulus (GO:0071230), transforming growth factor beta receptor signalling pathway (GO:0007179). One KEGG pathway, i.e. protein digestion and absorption (KEGG:map04974), was also significantly associated with all traits in CH and LM. Apart from this overlap, only a limited number of terms and pathways were over-represented across breeds. The GO term mitochondrial inner membrane (GO:0005743) was significantly over-represented for the DL trait in AA and the WOW trait in HE, although none of the same genes were significantly associated with both traits. Another GO term collagen trimer (GO:0005581) was over-represented for DIT in AA and DL in LM.

Table 3.6 Fold enrichment/depletion of SNPs in each annotation class for each trait in each breed

		3' UTR variant	5' UTR variant	Coding sequence variant	Downstream gene variant	Intergenic variant	Intron variant	Missense variant	Non-coding transcript	Splice acceptor variant	Splice region variant	Stop gained	Synonymous variant	Upstream gene variant
DHQ	AA	-	-	-	3.76	0.89	1.02	-	-	-	-	-	-	0.73
	CH	3.40	0.46	-	1.08	1.09	0.84	0.62	-	-	-	7.12	0.97	0.39
	HE	1.80	-	-	4.43	0.36	2.32	-	-	-	-	-	0.91	0.49
	LM	1.30	0.85	-	0.50	1.02	1.09	1.15	-	-	0.67	12.89	0.47	0.37
	SI	-	-	-	0.40	0.86	1.54	-	-	-	-	-	-	0.39
DL	AA	-	-	-	1.11	1.15	0.63	-	-	-	-	-	3.98	0.64
	CH	7.36	-	-	2.10	0.96	1.00	0.72	1.60	-	-	18.71	1.07	0.50
	HE	-	-	-	1.21	0.77	1.64	-	-	-	-	-	3.63	0.39
	LM	2.33	-	-	0.20	1.25	0.58	-	-	79.49	-	-	-	0.19
	SI	-	-	-	0.94	0.93	1.13	3.90	-	-	-	-	-	1.52
DIT	AA	-	-	-	0.14	1.03	1.12	-	-	-	-	-	-	0.66
	CH	3.47	0.48	-	0.94	1.11	0.79	0.58	-	-	-	7.47	0.96	0.40
	HE	-	-	-	0.40	0.48	2.43	2.49	-	-	-	-	-	0.98
	LM	1.16	1.77	99.15	0.86	0.99	1.09	2.28	-	-	0.70	13.49	0.59	0.55
	SI	-	-	-	0.86	1.04	1.05	-	-	-	-	-	-	0.62
TW	AA	-	-	-	0.64	1.19	0.64	-	-	-	-	-	2.91	0.31
	CH	3.07	0.54	-	1.22	0.99	1.01	0.24	1.43	-	-	8.40	1.44	0.83
	HE	-	-	-	1.10	0.87	1.50	-	-	-	-	-	-	-
	LM	3.17	1.45	-	0.45	1.21	0.61	0.22	-	-	-	22.05	0.32	0.28
	SI	2.10	9.65	-	3.62	0.55	1.26	0.74	-	-	23.85	-	-	5.36
WOW	AA	3.28	-	-	0.18	0.84	1.44	-	-	-	-	-	1.75	1.67
	CH	6.90	-	-	1.16	1.00	1.00	-	-	-	-	27.28	1.09	0.60
	HE	-	40.89	-	1.00	0.93	1.17	-	-	-	-	-	4.95	0.52
	LM	2.15	2.46	-	0.58	1.12	0.84	0.37	-	-	-	-	0.32	0.41
	SI	1.99	-	-	1.35	1.12	0.65	2.79	-	-	-	-	-	1.03

3.6 Discussion

Whereas a number of across-breed and breed-specific pleiotropic QTL have been documented for carcass traits, birth weight, weaning weight, and mature weight in beef cattle (Saatchi et al., 2014b), as well as for dry matter intake and growth and feed efficiency (Saatchi et al., 2014a), no study has attempted to detect across-breed or breed-specific pleiotropic QTL for muscularity linear type traits. Previous studies have been conducted on the genetic correlations between the linear type traits themselves (Chapter 2) and between both meat yield and carcass cuts with the muscularity linear type traits (Pabiou et al., 2012). While these genetic correlations are moderate to strong, none is equal to 1, which implies that two animals that yield a carcass of similar merit could be morphologically different. In fact, a shorter and more muscular animal or a taller and less muscular animal could have the same total carcass weight. In turn, these animals could yield very different carcass values owing to their distribution of primal cuts. For example, the loin of an animal harbours generally the most valuable cuts (Unnevehr and Bard, 1993; Connolly et al., 2018). Therefore, selection for a better-developed loin could lead to a more valuable carcass in comparison to a carcass with a lesser-developed loin if that carcass was still within the factory specification for weight and conformation.

Here, we have detected several genomic regions that are strongly associated with each of the muscularity traits analysed. However, most of these regions were unique to each trait or each breed, which indicates the existence of trait-specific and breed-specific QTL for muscularity traits. Thus, it is plausible to hypothesise that through more precise (i.e., targeting individual QTL) genome-based evaluations and selection, the morphology of an animal could be targeted to increase the output of high-quality

carcass cuts and consequently improve the profitability of the farm system and the value to the meat processor (Connolly et al., 2018). While a similar conclusion could be achieved through traditional breeding means, exploiting the breed- and trait-specific QTL could be more efficient.

This is the first published genome study on muscularity linear type traits in beef cattle using sequence data and is one of the few genome-based studies that compare multiple breeds of beef cattle. The number of animals used in our study is comparable to the number of animals used in a previous across-breed comparison that focused on carcass and birth traits in 10 cattle breeds (Saatchi et al., 2014b) and was thought to be the largest genome based-study ever performed in beef cattle at that time. This previous across-breed study was undertaken on 12 traits including birth weight, calving ease, carcass weight, and mature weight across 10 breeds and the results were similar to what we observed here for the muscularity traits. Saatchi et al. (2014b) identified 159 unique QTL associated with 12 traits, but only four QTL had pleiotropic effects and segregated in more than one breed. Similar results were observed in an across-breed study on dry matter intake, growth and feed efficiency in four beef cattle breeds (Saatchi et al., 2014a). The QTL identified for these traits were also breed-specific with little overlap among the breeds. This is comparable to our findings that show that the majority of the QTL were also trait-specific and breed-specific.

In total, approximately 83% of all QTL that are suggestively or significantly associated with a trait in our study overlapped with previously reported QTL associated with other production traits in dairy or beef cattle in the Cattle QTLdb (accessed 08 January 2019). Approximately 36% of all QTL overlapped with other traits that were specifically related to muscle in beef cattle such as body weight, carcass weight and

marbling score (Saatchi et al., 2014b), calving traits (Sahana et al., 2011), Warner-Bratzler shear force (McClure et al., 2012), and longissimus muscle area (Peters et al., 2012). One QTL on BTA17 that was associated with DIT in the SI breed was previously associated with ribeye area in a composite beef cattle breed composed of 50% Red Angus, 25% Charolais, and 25% Tarentaise (Hamidi Hay and Roberts, 2018). Our study is further validated by the presence of significantly associated QTL regions on BTA2, which harbours the *MSTN* gene, with the five muscularity traits in the CH and LM breeds, and within the meta-analysis. In a previous study on five muscularity type traits, which were combined into one singular muscular development trait in CH, a QTL on BTA2, which contained *MSTN*, was the only region significantly associated with these traits (Vallée et al., 2016).

In general, the suggestive and significant QTL, and thus genes, associated with each trait and each breed were both trait-specific and breed-specific. The low commonality of QTL among the breeds may be due to different genetic architectures underlying the traits in these breeds, or to gene-by-environment or epistatic interactions (Saatchi et al., 2014a), or to differences in the power to detect QTL due to the large differences in population sizes between the breeds. In many cases, the significant alleles were simply not segregating in all five breeds. The differences between breeds may also be due to limitations in the imputation process with the imputation accuracy being too low to determine strong associations between a SNP and a trait; consequently, the minor suggestive associations were interpreted with caution because of the possibility of poor imputation. Overall, the largest number of overlaps among significant genes were found between the CH and LM breeds for all traits, which is not surprising considering the relative similarities in the origins of these breeds (Kelleher et al., 2016) and of the selection pressures they have experienced (Zhao et al., 2015).

3.6.1 Myostatin

MSTN was first observed as a negative regulator of skeletal muscle mass in mice (McPherron et al., 1997) and since then has been identified as responsible for muscular hypertrophy in cattle (McPherron and Lee, 1997; Grobet et al., 1997) and is widely known as the causal variant for many muscularity and carcass traits in cattle (Casas et al., 2000; Allais et al., 2010). The stop-gain mutation Q204X in *MSTN* was significantly associated with the muscularity traits in both the CH and LM populations in the present study. Previously published research showed that CH and LM calves carrying one copy of this mutated allele scored better for carcass traits than non-carrier animals and that young CH bulls carrying this mutation presented a carcass with less fat and more tender meat than non-carriers (Allais et al., 2010). In the present study, the CH and LM animals carrying one copy of the minor allele scored significantly ($p < 0.01$) higher for muscularity type traits. The Q204X mutation was not significant in the AA population and it was removed during the data-editing step in both HE and SI as it was non-segregating. When Q204X was included as a fixed effect in the model for the CH animals, no SNPs located within the *MSTN* gene itself remained significant. This indicates that the significant SNPs within this gene were in tight linkage disequilibrium with Q204X, which provides evidence that this mutation may be causative for the muscularity linear type traits in the CH breed. Other genes on BTA2 that were significantly associated with some or all of the traits in CH and LM were *ORMDL1*, *PMS1*, *MFSD6*, and *NAB1*, all of which are in strong linkage disequilibrium with *MSTN* in mammals (Grade et al., 2009).

3.6.2 Other Candidate Genes

While the major peaks on BTA2 in the analyses on CH and LM, and all the meta-analyses contain *MSTN*, a known contributor to muscle development, it is also plausible that other candidate genes within the QTL on BTA2 could also contribute to muscle development. Two such genes are *COL3A1* and *COL5A2*. Intronic variants in *COL3A1* and upstream and downstream gene variants in *COL5A2* were significantly associated with DHQ in both CH and LM; however, no SNPs within coding or non-coding regions of this gene were associated with any traits in AA, HE, or SI although the SNPs were indeed segregating. Collagen is abundant in muscle and the quantity and stability of these intramuscular fibres have previously been linked to eating palatability of beef (Miller et al., 1987). The quantity and stability of muscle collagen are known to differ by breed (Andersen et al., 1977), sex (Boccard et al., 1979), and age (Cross et al., 1973) of cattle. Other collagen genes, *COL6A1*, *COL6A2*, and *COL18A1*, on BTA1 were also identified as candidate genes for DIT in the AA breed. Both type VI collagen genes have previously been linked to various muscle disorders in humans since they are known to affect muscular regeneration (Urciuolo et al., 2013). Type XVIII collagen has previously been proposed as a useful marker for beef marbling because it is involved in fat deposition in ruminants (Inoue-Murayama et al., 2000).

Another QTL on BTA2 located in the region between 13.9 and 14.9 Mb and significantly associated with four of the traits (DHQ, DIT, TW, and WOW) in the LM breed contained the *PDE1A* and *PPP1R1C* genes. The most significant SNP in this region was an intronic SNP within *PDE1A*. The *PDE1A* gene is involved in a pathway related to myofibroblast formation in smooth muscle in humans (Zhou et al., 2010) while previous genome-wide studies in mice have identified the *PPP1R1C* gene as a possible candidate gene for muscle mass (Kärst et al., 2011). Overall, the allele

frequencies of the favorable alleles in this 1-Mb region were similar in all five breeds, which support a breed-specific association with DHQ, DIT, TW, and WOW in LM rather than an imputation error.

An additional breed-specific QTL on BTA2 that contains numerous *HOXD* genes was associated with WOW in the LM population. The *HOXD* genes are documented as having a role in limb (Zakany and Duboule, 2007) and digit (Delpretti et al., 2012) formation, thus they probably also play a role in skeletal muscle development. The most significantly associated SNPs with WOW in this region were only segregating in the LM breed and had a very high favorable allele frequency (0.998) in this breed. These SNPs were fixed or very close to fixation in the four other breeds.

In the meta-analyses of DHQ, associated variants in all the breeds analysed were identified, which may be beneficial for across-breed genomic prediction (Purfield et al., 2015). Although the associations detected in the meta-analysis corresponded to associations identified in the CH and LM breeds, three of these QTL on BTA5, 11, and 12 increased in significance when compared to the within-breed analysis. The QTL on BTA5 which contained the *AMDHD1* gene, was located close to a QTL previously associated with carcass composition [43], whereas the QTL on BTA11 contains *DYSF*, a gene known to be linked with muscular dystrophy in humans (Al-Zaidy et al., 2014). The QTL on BTA14 contained the *PREX2* gene which was previously linked to carcass weight in Hanwoo cattle (Edea et al., 2018).

Interestingly, in the meta-analyses of DL and WOW, a 1-Mb QTL on BTA7 containing the *SPRY4* and *FGF1* genes became suggestively associated, although it was not associated in any breed individually. The *SPRY4* gene was reported to be associated with feed intake in cattle (Chen et al., 2011), whereas *FGF1*, a member of the fibroblast

growth factor family, is thought to be involved in embryonic muscle formation (Hudson et al., 2013).

Similarly, in the meta-analysis of TW, a 3-Mb QTL on BTA6 containing the NCAPG/LCORL genes became suggestively associated, although it was not associated in any breed individually. These genes are associated with variation in body size and height in cattle (Bouwman et al., 2018), humans (Wood et al., 2014), and horses (Tetens et al., 2013), thus they are likely plausible candidate genes associated with muscularity linear type traits describing the size of the body.

3.6.3 Gene Ontology and KEGG Pathways

Linear type traits are complex traits that are governed by many genes each with a small effect, and hence, are likely involved in many biological systems. Several GO terms were only associated with a single trait or a single breed; hence there was limited commonality among traits or breeds suggesting the absence of a central biological process that links these traits together. Over-represented GO terms in multiple traits and breeds include those related to skin development, collagen fibril organisation, and the transforming growth factor beta receptor signalling pathway. Each of these GO terms was associated with genes located in the large QTL on BTA2 that contained MSTN. Excluding the major *MSTN* QTL in these breeds, which is known to have a large effect on muscularity, the various GO terms and KEGG pathways represented by the genes associated with the muscularity traits suggest that the majority of genes identified as significantly associated with a trait are not only breed-specific but also trait-specific in many cases.

3.6.4 Regulatory Regions Involved in the Development of Muscle

Although millions of SNPs were tested for association with each trait, only 79 of the SNPs suggestively or significantly associated with a trait were located in the coding region of a gene; the vast majority of the SNPs associated with the muscularity traits in any of the breeds were located outside of the coding regions. This is consistent with previous genomic studies for complex quantitative traits in cattle using HD SNP data (Koufariotis et al., 2014) or sequence data (Bouwman et al., 2018). While the coverage of the HD study (Koufariotis et al., 2014) may not have included the coding regions required to identify significant associations within these regions, our study and a previous study on cattle stature (Bouwman et al., 2018) used imputed sequence data, and thus, covered the entire genome.

Whereas many studies have previously acknowledged the importance of non-coding SNPs to genetic variability, little is actually known about the mechanisms by which these SNPs contribute to variation in complex traits (Visel et al., 2009), (Schierding et al., 2016). One possibility to explain the significance of these non-coding SNPs is that the non-coding regions contain gene regulatory sequences, called enhancers, that act over long distances possibly altering the expression of a gene nearby (Visel et al., 2009). Another possibility is that the folding of DNA into the 3-dimensional nucleus may cause distant loci, such as those in non-coding and coding regions, to become spatially close together thus enabling these regulatory regions to come into contact with genes far away or even on different chromosomes (Schierding et al., 2014).

Non-coding variants such as 3’ UTR, 5’ UTR and intergenic variants were enriched for most of the traits in each breed. Downstream and upstream gene variants

were also enriched in some traits. In general, the SNPs located close to and within the genes identified as candidate genes were located within non-coding or regulatory regions. For example, for DHQ in the CH breed, 60 suggestively and significantly associated SNPs were located within the *MSTN* gene; 10 of these were 3'UTR variants, 31 were downstream gene variants and 19 were intronic. Whereas regulatory regions may not have an effect on the coding sequence of any gene, they are thought to be particularly important for growth and development in humans (Schierding et al., 2014; Schierding et al., 2016) and cattle (Karim et al., 2011; Bouwman et al., 2018). Thus, similar to previous observations in humans and cattle, enrichment of the non-coding classes of SNPs in our study may indicate the importance of regulatory regions for cattle muscle development.

3.7 Conclusion

Although we identified many QTL associated with muscularity in beef cattle, our results suggest that these QTL tend to be not only trait-specific but also breed-specific. Overall, the significant SNPs contained in these QTL were more likely located in regulatory regions of genes, which suggest the importance of non-coding regions that may affect gene expression for muscle development in cattle. Some shared regions associated with muscularity were found between CH and LM, with a large-effect QTL on BTA2 containing *MSTN* being associated with the five traits analysed. This overlap between these breeds was somewhat expected, because they are subjected to similar selection pressures. Apart from this single QTL, extensive differences were observed between the breeds, which may be due to the much smaller sample sizes for AA, HE, and SI compared to the CH and LM populations that result in reduced power to detect QTL or they may be due to differences in genetic architecture of these traits among the

populations. In many cases, the strongly associated SNPs in one breed were not segregating in the other breeds, and thus, were missing from the analyses. Knowledge of any potential differences in genetic architecture among breeds is important to develop accurate genomic prediction equations in across-breed analyses.

Chapter 4

Genomic regions associated with skeletal type traits in beef and dairy cattle are common to regions associated with carcass traits, feed intake and calving difficulty

4.1 Preface

At the time of thesis submission this chapter was published in the Frontiers in Genetics (Accepted January 7, 2020; doi: 10.3389/fgene2020.00020). The full reference is Doyle JL, Berry DP, Veerkamp RF, Carthy TR, Evans RD, Walsh SW, Purfield DC: Genomic regions associated with skeletal type traits in beef and dairy cattle are common to regions associated with carcass traits, feed intake and calving difficulty. Frontiers in Genetics 2020, 11:20.

Jennifer Doyle was primary author, performed the data edits and analysis and drafted the manuscript. Donagh Berry, Deirdre Purfield and Roel Veerkamp conceived the study, participated in the design and co-ordination of this study and helped draft the manuscript. Ross Evans supplied the data the analysis was performed on. All authors read and approved the final manuscript.

Formatting and referencing style has been edited for consistency throughout the thesis. Figure and table captions have been assigned with a chapter prefix. Competing interests and acknowledgements have been removed. All other aspects are consistent with the published manuscript.

4.2 Abstract

Linear type traits describing the skeletal characteristics of an animal are moderately to strongly genetically correlated with a range of other performance traits in cattle including feed intake, reproduction traits and carcass merit; thus, type traits could also provide useful insights into the morphological differences among animals underpinning phenotypic differences in these complex traits. The objective of the present study was to identify genomic regions associated with 5 subjectively scored skeletal linear traits, to determine if these associated regions are common in multiple beef and dairy breeds, and also to determine if these regions overlap with those proposed elsewhere to be associated with correlated performance traits. Analyses were carried out using linear mixed models on imputed whole genome sequence data separately in 1,444 Angus, 1,129 Hereford, 6,433 Charolais, 8,745 Limousin, 1,698 Simmental, and 4,494 Holstein-Friesian cattle all scored for the linear type traits; there was, on average, 24 months difference for the age of assessment of the beef versus the dairy animals. While the majority of the identified quantitative trait loci (QTL), and thus genes, were both trait-specific and breed-specific, a large-effect pleiotropic QTL on BTA6 containing the *NCAPG* and *LCORL* genes was associated with all skeletal traits in the Limousin population and with wither height in the Angus. Other than that, little overlap existed in detected QTLs for the skeletal type traits in the other breeds. Only 2 QTLs overlapped the beef and dairy breeds; both QTLs were located on BTA5 and were associated with height in both the Angus and the Holstein-Friesian, despite the difference in age at assessment. Several detected QTL in the present study overlapped with QTL documented elsewhere to associate with carcass traits, feed intake, and calving difficulty. While most breeding programs select for the macro-traits like carcass weight, carcass conformation, and feed intake, the higher degree of granularity with selection on

the individual linear type traits in a multi-trait index underpinning the macro-level goal traits, presents an opportunity to help resolve genetic antagonisms among morphological traits in the pursuit of the animal with optimum performance metrics.

4.3 Introduction

Linear type traits have been used in both beef and dairy cattle since the early 20th century to characterize the skeletal characteristics of an animal (Berry et al., 2019). These type traits have previously been identified as being moderately to strongly genetically correlated with a range of performance traits in cattle including feed intake (Veerkamp and Brotherstone, 1997; Crowley et al., 2011), reproductive traits (Berry et al., 2004; Wall et al., 2005; Carthy et al., 2016), carcass merit (Mukai et al., 1995; Berry et al., 2019), animal value (Mc Hugh et al., 2010), and health (Ring et al., 2018). As type trait measurements are typically taken when an animal is young (Chapter 2), they may be useful as early predictors of the correlated traits which are often measured later in life or after the animal is slaughtered. While type traits are also moderately to strongly correlated with live-weight (Mc Hugh et al., 2010; Berry et al., 2019) and carcass weight (Conroy et al., 2010), none of these correlations are unity implying that two animals with the same weight may be morphologically very different; for example, a tall animal with a short back may have the same (carcass) weight as a short animal with a long back. Therefore, including linear type traits in future genetic and genomic evaluations as part of a multi-trait evaluation including also the goal trait of interest may provide additional information on what could be gleaned from the goal traits alone.

While many genomic studies have been carried out on stature in both beef and dairy cattle (Pryce et al., 2011; Bolormaa et al., 2014), few studies have been published on the underlying genomic features contributing to differences in other skeletal linear

type traits in either beef (Vallée et al., 2016) or dairy (Cole et al., 2011; Wu et al., 2013; Sahana et al., 2015) cattle. No previous study has attempted to identify quantitative trait loci (QTL) associated with the skeletal traits in multiple breeds or to compare and contrast detected QTLs to previously identified QTLs associated with correlated complex phenotypes such as carcass merit, feed intake and efficiency, and calving performance. Therefore, the objective of the present study was to identify genomic regions associated with 5 subjectively-scored skeletal linear traits to determine if these associated regions are common in multiple beef and dairy breeds and also to determine if these regions overlapped with previously identified QTLs associated with other correlated performance traits.

4.4 Materials & Methods

4.4.1 Beef Phenotypes

Routine scoring of linear type traits is carried out on both registered and commercial beef herds by trained classifiers from the Irish Cattle Breeding Federation as part of the Irish national beef breeding programme (Mc Hugh et al., 2010; Berry and Evans, 2014). Five skeletal type traits scored on a scale of 1 to 10 on beef cattle describing the wither height (WH), back length (BL), chest depth (CD), chest width (CW) and hip width (HW) were included for analysis in the present study (Appendix B1). Data on these linear type traits were available on 147,704 purebred Angus (AA), Charolais (CH), Hereford (HE), Limousin (LM), or Simmental (SI) beef cattle, all scored between the ages of 6 and 16 months between the years 2000 and 2016, with only one (i.e., the first) record per animal retained.

Animals were discarded from the dataset if the sire, dam, herd, or classifier was unknown. Only data from classifiers that scored ≥100 animals since the year 2000 were

kept. Animals were also discarded from the dataset if the parity of the dam was unknown; parity of the dam was subsequently recoded into 1, 2, 3, 4, and ≥5. Contemporary group was defined as herd-by-scoring date generated separately per breed. Each contemporary group had to have at least five records. Following edits, data were available on 81,200 animals, aged between 6 and 16 months, consisting of 3,356 AA, 31,049 CH, 3,004 HE, 35,159 LM, and 8,632 SI.

4.4.2 Dairy phenotypes

Scoring of linear type traits in the Irish dairy herd is undertaken by trained classifiers from the Irish Holstein-Friesian Association (Berry et al., 2004). For the purpose of the present study, 3 skeletal linear type traits that closely align to one of the 5 beef skeletal traits were selected for analysis. These traits were stature (STA which is comparable to WH in beef), rump width (RW which is comparable to HW in beef), and chest width (CWD which is comparable to CW in beef). In dairy cattle, these traits were scored on a scale of 1 to 9 (Appendix B2) with the direction of scale the same as the comparable traits in the beef herd. Linear type trait information on 239,776 first parity cows was available between the years 2000 and 2016; only the first record per cow was retained.

Animals were discarded from the dataset if the sire, dam, herd, or classifier was unknown. Records were also discarded from the data set if scored after 10 months of lactation. Only data from classifiers that scored >100 animals since the year 2000 were retained. Contemporary group was defined as herd-by-scoring date and each contemporary group had to have at least five records. Following edits, data were available on 117,151 primiparous Holstein-Friesian cows (HF) aged between 23 and 42 months at scoring.

4.4.3 Generation of Adjusted Phenotypes

Prior to inclusion in the analysis, all beef cattle phenotypes were adjusted, within breed, in ASREML (Gilmour et al., 2009) using the model:

$$Y_{ijklm} = HSD_m + Sex_j + AM_k + DP_l + Animal_i + \mathbf{e_{ijklm}}$$

where Y_{ijklm} is the linear type trait, HSD_m is the fixed effect of herd-by-scoring date (m=11,130 levels), Sex_j is jth sex of the animal (male or female), AM_k is the fixed effect of the age in months of the animal (k=11 classes from 6 to 16 months), DP_l is the fixed effect of the parity of the dam (l=1, 2, 3, 4 and ≥5), $animal_i$ is the random additive effect of animal i, and e_{ijklm} is the random residual effect. The adjusted phenotype was the raw phenotype less the fixed effect solutions of HSD, Sex, AM, and DP.

The dairy phenotypes were also adjusted in ASREML (Gilmour et al., 2009) using the model:

$$Y_{ijklm} = HSD_m + AM_j + CM_k + LS_l + Animal_i + \mathbf{e_{ijklm}}$$

where Y_{ijklm} is the linear type trait, HSD_m is the fixed effect of herd-by-scoring date (m=9,591 levels), AM_j is the fixed effect of the age in months of the animal at scoring (j=20 levels from 23 to 42 months), CM_k is the fixed effect of the month of calving (k=12 levels from 1 to 12), LS_l if the fixed effect of the stage of lactation of the animals (l=10 levels from 1 to 10 reflecting number of months of lactation), $animal_i$ is the random additive effect of animal i, and e_{ijklm} is the random residual effect. The adjusted phenotype was the raw phenotype less the fixed effect solutions of HSD, AM, CM and LS.

4.4.4 Genotype Data

Of the edited dataset of 81,200 beef animals and 117,151 dairy animals with linear type trait information, 23,943 animals from 6 breeds (1,444 AA, 6,433 CH, 1,129 HE, 8,745 LM, 1,698 SI, and 4,494 HF) also had genotype information available. These genotypes were imputed to whole genome sequence (WGS) as part of a larger dataset of 638,662 genotyped animals from multiple breeds as detailed by Purfield et al. (2019). All 638,662 genotyped animals were genotyped using either the Bovine Illumina SNP50 (n=5,808; 54,001 SNPs), the Illumina High Density (HD; n=5,504; 777,972 SNPs), the Illumina 3k panel (n=2,256, 2,900 SNPs), the Illumina LD genotyping panel (n=15,107, 6,909 SNPs) or a bespoke genotype panel (IDB) developed in Ireland (Mullen et al., 2013) which was either on version 1 (n=28,288; 17,137 SNPs), version 2 (n=147,235; 18,004 SNPs) or version 3 (n=434,464; 53,450 SNPs). Each animal had a call rate ≥ 90%. Only autosomal SNPs, SNPs with a call rate ≥ 90% and those with a known chromosome and position on UMD 3.1 were retained for imputation.

Imputation to HD was carried out on all genotyped animals using a two-step approach in FImpute2 with pedigree information (Sargolzaei et al., 2014); this involved imputing animals genotyped on the 3k, LD, or IDB panels to the Bovine SNP50 density and subsequently imputing all resulting genotypes (including the Bovine SNP50 genotypes) to HD using a multi-breed reference population of 5,504 influential sires genotyped on the HD panel. Imputation to WGS was then undertaken using a reference population of 2,333 *Bos Taurus* animals of multiple breeds from Run6.0 of the 1000 Bulls Genomes Project by first phasing all 638,662 imputed HD genotypes using Eagle (version 2.3.2; (Loh et al., 2016) and subsequently imputing to WGS using minimac3 (Das et al., 2016).

Quality control edits were imposed on the imputed sequence genotypes within each of the 6 breeds separately; all SNPs with a minor allele frequency (MAF) ≤ 0.002 were removed and regions of poor WGS imputation accuracy, identified using 147,309 verified parent-progeny relationships as previously described by Purfield et al. (2019), were then removed. Following all SNP edits, 16,342,970, 17,733,147, 16,638,022, 17,803,135, 17,762,681, and 15,542,919 autosomal SNPs remained for analysis in the AA, CH, HE, LM, SI and HF populations, respectively.

4.4.5 Association Analyses

The association analyses were performed, within each breed separately, using a mixed linear model in Genome-wide Complex Trait Analysis (GCTA; Yang et al., 2011). Autosomal SNPs from the original HD density panel (i.e., 734,159 SNPs) were used to construct the genomic relationship matrix (Yang et al., 2010). The model used for the within-breed analysis was:

$$\mathbf{y} = \mu + \mathbf{xb} + \mathbf{u} + \mathbf{e}$$

where **y** is a vector of preadjusted phenotypes, μ is the overall mean, **x** is the vector of imputed genotypes, **b** is the additive fixed effect of the candidate SNP to be tested for association, $\mathbf{u} \sim N(0,\mathbf{G}\sigma_u^2)$ is the vector of additive genetic effects, where **G** is the genomic relationship matrix calculated from the imputed HD SNP genotypes, and σ_u^2 is the additive genetic variance, and $\mathbf{e} \sim N(0,\mathbf{I}\sigma_e^2)$ is the vector of random residual effects, with **I** representing the identity matrix and σ_e^2 the residual variance

4.4.6 QTL Detection, Gene Annotation & Variance Explained

A significance threshold of $p \leq 1 \times 10^{-8}$ and a suggestive threshold of $p \leq 1 \times 10^{-5}$ were applied genome-wide for each SNP in each trait as per (Wang et al., 2016). Significant and/or suggestive SNPs that were within 500kb of each other were classed as being within the same QTL. Genes within these QTLs were then identified using Ensembl 94 (Zerbino et al., 2017) on the UMD 3.1 bovine genome assembly. Cattle QTLdb (https://www.animalgenome.org/cgi-bin/QTLdb/BT/index) was used to identify if any of the QTLs identified within the present study had previously been associated with any other traits in beef or dairy cattle. To identify QTL regions that were suggestive in more than 1 breed, each chromosome was split into 1kb genomic windows and windows containing suggestive SNPs ($p \leq 1 \times 10^{-5}$) were compared across the breeds.

The proportion of genetic variance of a trait explained by a SNP was calculated as:

$$\frac{2p(1-p)a^2}{\sigma_g^2}$$

where p is the frequency of the minor allele, a is the allele substitution effect and σ_g^2 is the genetic variance of the trait in question as calculated from the association analyses.

4.4.7 Meta-analyses

Following the within breed analyses, meta-analyses were conducted for CD and BL across the 5 beef breeds and for WH, CW and HW across all six breeds using the weighted Z-score method in METAL (Willer et al., 2010). METAL uses the p-values and the direction of SNP effects from the individual analysis and weights the individual studies based on the sample size to calculate an overall Z-score:

$$Z = \frac{\Sigma_i z_i w_i}{\sqrt{\Sigma_i w_i^2}}$$

where w is the square root of the sample size of the ith breed, and z is the z-score for the ith breed calculated as $z_i = \phi^{-1}\left(1 - \frac{P_i}{2}\right)\Delta_i$, where Φ is the cumulative distribution function, and P_i and Δ_i are the P-value and direction of effect for breed *i*, respectively.

4.4.8 Enrichment Analyses

Enrichment analysis was carried out among all suggestive and significant SNPs within each trait and each breed separately to estimate if the number of SNPs in each annotation class was greater than what would be expected by chance (Bouwman et al., 2018):

$$\text{enrichment} = \frac{a}{b}\left[\frac{c}{d}\right]^{-1}$$

where a is the number of suggestive and/or significant SNPs in the annotation class of interest, b is the total number of suggestive and/or significant SNPs that were associated, c is the total number of SNPs in the annotation class in the association analysis, and d is the overall number of SNPs included in the association analysis.

4.5 Results

The scale of measurement, number of records, mean, and standard deviation of the linear type traits in each breed is in Appendix B1 and B2. The average age of the beef cattle at measurement was 10 months while the average age of the dairy cows was 28 months; hence, there was, on average, a 2 year difference in age at classification

between the dairy and beef populations. Significant ($p \leq 1 \times 10^{-8}$) and/or suggestive ($p \leq 1 \times 10^{-5}$) SNPs were detected for all of the traits in all 6 breeds; however, the exact locations of these SNPS, and the direction of the effects of these SNPs, differed by breed.

4.5.1 Wither Height/Stature

No 1kb genomic window associated with height was common to all 6 breeds. There was, however, some overlap in suggestive 1kb windows between AA and LM where 79 suggestive windows located on BTA6 were common to both breeds (Appendix B3). Six genes were identified within these windows on BTA6 including *NCAPG* and *LCORL*. There were also 2 suggestive 1kb windows located at approximately 94.9 Mb on BTA5 common to both the AA and HF.

The strongest association in both the AA and LM were intergenic variants located in QTLs surrounding the *NCAPG* and *LCORL* genes on BTA6 (Table 4.1) and accounted for 0.6% and 0.04% of the genetic variation in WH in the AA and LM, respectively. Five intronic variants and three downstream gene variants located within the *LCORL* gene, and 12 intronic variants located within the *NCAPG* gene, were suggestively associated in the AA ($p < 9.18 \times 10^{-6}$) and significantly associated in the LM ($p < 1.29 \times 10^{-12}$). Interestingly, the positive (i.e., taller) allele of these SNPs occurred at similar frequencies (0.08 to 0.09) in both the AA and LM and had a similar effect size in both breeds. In comparison, while these SNPs were segregating in both the HE and HF, and had similar allele frequencies in the HE as in the AA and LM, none of these SNPs were near significance in either the HE ($p > 0.11$) or HF ($p > 0.88$). However, a suggestive association was detected 21 Mb further upstream of *LCORL* on

BTA6 in the HF where the strongest association within this QTL, rs209851496 ($p = 1.94 \times 10^{-6}$), was located 1kb upstream of the *CHRNA9* gene.

Of the 514 SNPs that were suggestively associated with stature in HF, 281 were located on BTA5. Both AA and HF had suggestive associations on this autosome; two intergenic SNPs, rs798298008 (AA) and rs475950607 (HF), located just 17 bp apart and 63 Kb from the *PTPRO* gene, were associated with WH in these breeds. The strongest associations in the remaining breeds were all intergenic SNPs, although their location differed by chromosome; the strongest association in CH was on BTA2 in a 1 Mb QTL containing *MSTN*; the strongest association in HE was in BTA7, with the strongest association for SI located on BTA12.

There were 1,055 suggestive and 36 significant SNPs associated with WH in the meta-analysis (Appendix B5). A single QTL on BTA15 containing multiple plausible candidate genes, such as *ALKBH8* and *RAB39A*, was the only QTL identified that had not previously been associated with WH in any of the within-breed analyses.

4.5.2 Chest Width

The window-based analyses revealed no 1kb genomic region suggestively associated with CW in more than one breed (Appendix B3). Similar to WH, BTA6 harbored the strongest QTL association for CW in LM. This QTL, which also encompassed the *NCAPG*/*LCORL* complex, contained 34 suggestively associated SNPs, of which the strongest (rs110194711) was in the *MEPE* gene. A similar genomic region on BTA6 was also associated with CW in HE, suggesting that the QTL region on BTA6 may harbor an across-breed pleiotropic association since it was also associated with WH in AA and LM. Although four of the 6 breeds (AA, CH, HE and HF) had QTLs on BTA10 suggestively associated with CW (Table 4.2), these all differed in their location across

the chromosome which may suggest that BTA10 contains multiple genomic regions influencing CW.

The meta-analysis of all 23,943 animals failed to identify a genomic region significantly associated with CW, but 170 SNPs were suggestively associated (Appendix B5). The majority of these associations were singular SNPs, although peaks of suggestive association were detected on BTA1, BTA2, BTA8, BTA16, and BTA19.

4.5.3 Hip Width/Rump Width

There were no 1kb suggestive windows common to any of the breeds associated with width of hips. The QTL on BTA6 surrounding the *NCAPG* and *LCORL* genes was again significant in the LM although it failed to reach significance in the remaining 5 breeds (Table 4.3). Of the 222 SNPs suggestively associated with HW in the HE population, 52% were located in a QTL on BTA4 surrounding the *CLEC5A* gene. Although *MSTN* may have been expected to influence HW in the CH, the QTL on BTA2 associated with HW was located much further down-stream, between 30.21 and 31.26 Mb (Table 4.3). Several plausible candidate genes were located within this QTL on BTA2 including multiple voltage-gated sodium-channel genes, *TTC21B*, and *CSRNP3*; nonetheless only 0.07% of the genetic variation in HW was explained by the strongest association within this QTL. In HF, the most significant SNP associated with RW was an intergenic SNP, rs382714953 (2.03×10^{-7}), located on BTA20.

In comparison to WH and CW, the lead variant within the top 5 QTLs associated with HW in the AA, CH, HE and SI breeds was near fixation (Table 4.3). All of the lead variants in the top 5 QTLs in the SI breed were close to the fixation for the positive (i.e., wider) allele in the SI and fixed for the negative (i.e., narrower) allele in the HE. In

contrast, the frequency of the positive alleles for each of the lead variants identified in the LM population ranged from low to moderate.

In the meta-analysis of HW and RW, suggestively associated QTL were located on BTA11, BTA15, BTA18, and BTA23 (Appendix B5); none of these QTL had been previously identified in the individual breed analyses but they contained multiple possible candidate genes

4.5.4 Back Length

The window-based analyses revealed that no 1kb genomic region was suggestively associated with BL in all breeds, but 40 1kb windows on BTA6 surrounding the *NCAPG* and *LCORL* genes were suggestively associated with BL in both the AA and LM (Appendix B3). In total, 96 SNPs within a QTL spanning from 37.9 to 40.4 Mb on BTA6 were suggestively associated with BL in AA, of which 12 SNPs were either intronic SNPs, or downstream or upstream variants of the *NCAPG* and *LCORL* genes (Table 4.4). In LM, the most strongly associated SNP, rs110343895 ($p = 4.24 \times 10^{-13}$), was an intronic SNP located within *NCAPG*. In total, 7 SNPs located within the *NCAPG* gene and 15 SNPs within the *LCORL* gene were suggestively associated with BL in LM. Of the 33 potentially disruptive variants within the *NCAPG* and *LCORL* complex that were tested for association, 6 were segregating in the LM population but none were significant. LM animals that had at least one copy of the minor allele for the top 3 associated SNPs, rs465117501, rs378370406 or rs110343895, within the *NCAPG* and *LCORL* complex had a longer back, 0.37 (SE = 0.18) units longer on average, than those with two copies of the major allele.

A QTL on BTA2 was significantly associated with BL in CH; this QTL stretched 10 Mb and contained 1,765 significant and 3,760 suggestive SNPs. Fifty

significant and 12 suggestive SNPs within this QTL were located within the *MSTN* gene; these SNPs included the well-known Q204X stop-gain mutation, rs110344317 ($p = 2.01 \times 10^{-35}$). Where Q204X was forced into the model as a fixed effect, the most significant of the remaining SNPs on BTA2 generally reduced in significance relative to when Q204X was not included in the model. The most significant SNP on BTA2 after accounted for the variability in the Q204X genotype was rs41638272, an intergenic SNP located 15kb from the *SLC40A1* gene. The QTL associated with BL also overlapped the QTL on BTA2 associated with WH suggesting this QTL may play a major role in affecting the morphology of an animal. No other significant associations with BL were identified in any of the remaining beef breeds.

In the meta-analysis of BL, significantly associated QTLs were identified on BTA2 and BTA6, similar to what was identified in the CH and LM breeds, respectively (Appendix B5). Other QTLs on BTA12 and BTA13 were also associated with BL; the QTL on BTA13 contained numerous possible candidate genes including *DNTTIP1*, *TNNC2*, *PLTP*, and *CDH22* while no obvious candidate genes were identified on BTA12.

4.5.5 Chest Depth

No suggestive or significant 1kb window associated with CD was common to more than one breed (Appendix B3). Only a single QTL on BTA6 containing the *NCAPG* and *LCORL* genes in LM was significantly associated with any of the breeds for CD (Table 4.5), suggesting that CD has a highly polygenic architecture in the beef breeds. Four of the 5 lead variants identified within the top 5 QTLs associated with CD in the AA were near fixation for the negative (i.e., narrower) allele while 4 of the 5 lead variants associated in SI were close to fixation for the positive (i.e., deeper) allele. Only 90 SNPs

were suggestively associated with CD in CH, of which 19 were located on BTA10, but the proportion of genetic variance accounted for by the strongest association on this autosome was minimal (0.001%).

In the meta-analysis, 3 SNPs were identified to be significantly associated with CD while 249 SNPs were suggestively associated. Three QTLs associated with CD in the meta-analysis were not significant in any of the single breed analyses and were located on BTA1, BTA5 and BTA13 (Appendix B5).

Table 4.1 The location of the most significant QTLs, limited to the top 5, which were associated with wither height or stature, and the genes located within these QTLs within each breed (AA=Angus; CH=Charolais;HE=Hereford; LM=Limousin; SE=Simmental; HF=Holstein-Friesian).

Breed	Chr	Start	End	No of suggestive & significant SNPs	Most significant SNP	P-Value	Allele frequency of positive allele						Candidate genes within this QTL
							AA	CH	HE	LM	SI	HF	
Angus	6	37859028	40529961	96	39955422[a]	$7.31x10^{-9}$	0.114	0.000	0.000	0.064	0.000	0.042	ABCG2, PKD2, SPP1, MEPE, LAP3, NCAPG*, LCORL*
	6	40760106	41784760	14	41276346[b]	$2.74x10^{-7}$	0.445	0.000	0.372	0.522	0.000	0.784	SLIT2*, PACRGL, KCNIP4
	16	72342264	73978632	25	72877647[a]	$1.46x10^{-7}$	0.995	0.002	0.003	0.000	0.996	0.996	RPS6KC1, BATF3, PPP2R5A*
	20	46866355	47884741	51	47372538[a]	$7.48x10^{-7}$	0.161	0.310	0.523	0.768	0.822	0.834	ENSBTAG00000048105
	26	40278450	41826296	23	41323903[c]	$2.21x10^{-7}$	0.993	0.980	0.982	0.017	0.983	0.000	WDR11*, PTPRG, FHIT
Charolais	2	5346602	6349651	2	5846602[a]	$6.02x10^{-8}$	0.690	0.585	0.703	0.000	0.586	0.390	NAB1, MSTN, MFSD6
	5	40455760	41765149	12	40955760[a]	$5.68x10^{-8}$	0.000	0.038	0.987	0.000	0.016	0.010	SLC2A13*, ABCD2
	6	33942529	35471763	9	34442529[a]	$7.78x10^{-6}$	0.998	0.011	0.000	0.003	0.000	0.000	CCSER1
	27	11896148	12929004	15	12428578[a]	$7.97x10^{-7}$	0.295	0.464	0.000	0.513	0.501	0.344	TENM3, DCTD
	28	11615130	12630615	8	12127037[a]	$6.97x10^{-7}$	0.758	0.975	0.273	0.138	0.153	0.146	
Hereford	3	74681893	76225687	5	75725687[b]	$5.73x10^{-7}$	0.005	0.003	0.976	0.000	0.995	0.390	CTH, LRRC7*, LRRC40
	5	79055337	80113473	8	79564409[a]	$3.32x10^{-7}$	0.536	0.606	0.975	0.487	0.403	0.900	SINHCAF
	7	81624551	82816882	4	82124551[a]	$1.88x10^{-7}$	0.994	0.003	0.995	0.000	0.003	0.991	TENM2, WWC1
	20	19842459	20942794	51	20401686[a]	$2.44x10^{-7}$	0.043	0.098	0.266	0.271	0.198	0.146	PDE4D, RAB3C
	23	50140690	51876442	10	51357892[b]	$8.96x10^{-7}$	0.277	0.755	0.229	0.786	0.184	0.582	SLC22A23, RIPK1, NQO2, GMDS*
Limousin	4	57644495	58664115	9	58148365[a]	$5.52x10^{-7}$	0.974	0.053	0.932	0.092	0.953	0.335	IMMPL2
	6	31747431	35203508	1588	33609037[a]	$1.17x10^{-18}$	0.249	0.879	0.415	0.151	0.260	0.812	SMARCAD1, ATOH1, CCSER1
	6	36934944	41871562	663	38035891[d]	$1.45x10^{-16}$	0.086	0.000	0.000	0.128	0.636	0.007	PPM1K^, ABCG2^, PKD2^, SPP1^, MEPE*, LAP3, NCAPG^, LCORL^
	6	42312608	43680601	17	42990479[b]	$1.48x10^{-7}$	0.000	0.006	0.000	0.029	0.000	0.000	ADGRA3, KCNIP4*
	11	104805923	105866536	3	105366536[b]	$1.04x10^{-7}$	0.032	0.979	0.008	0.010	0.983	0.035	BRD3, WDR5, CACNA1B*
Simmental	8	82805400	83805881	3	83305881[a]	$1.67x10^{-6}$	0.367	0.688	0.693	0.540	0.279	0.675	FANCC
	8	106857510	107869952	3	107357510[b]	$8.48x10^{-7}$	0.990	0.073	0.928	0.093	0.859	0.878	PAPPA*, TRIM32
	12	55018060	56018149	3	55518060[a]	$2.66x10^{-7}$	0.000	0.955	0.004	0.967	0.005	0.000	SPRY2
	12	89258864	90269817	3	89758864[a]	$2.78x10^{-6}$	0.015	0.988	0.992	0.028	0.950	0.982	ANKRD10, ING1, SOX1, TUBGCP3
	22	1921471	3018467	32	2517667[a]	$4.87x10^{-7}$	0.000	0.000	0.000	0.000	0.003	0.000	CMC1, AZI2
Holstein Friesian	4	108676456	109728131	8	109185322[a]	$1.49x10^{-6}$	0.096	0.203	0.081	0.775	0.365	0.794	TPK1
	5	59814571	62558882	76	60701477[a]	$4.28x10^{-8}$	0.257	0.900	0.953	0.874	0.949	0.894	NEUROD4, TSPA1, NTN4*, SNRPF*, AMDHD1*, LTA4H*, CDK17*, NEDD1
	5	104934097	106783101	135	106283101[a]	$3.77x10^{-8}$	0.096	0.679	0.437	0.000	0.802	0.475	ANO2, NTF3, KCNA1, NDUFA9, FGF6*, FGF23*, TIGAR*
	6	60485248	61489096	26	60985248[d]	$1.94x10^{-6}$	0.965	0.903	0.148	0.000	0.904	0.973	UBE2K, N4BP2, RHOH, CHRNA9*, RBM47
	7	23221527	24809431	46	23789810[b]	$1.24x10^{-7}$	0.110	0.834	0.952	0.000	0.031	0.903	IRF1, PDLIM4, P4HA2, IL3, ACSL6, FNIP1*, HINT1

Superscript denotes SNP classification: [a]intergenic, [b]intron, [c]upstream gene variant, [d]downstream gene variant. Symbols denote the significance of SNPs within genes: *gene contained at least one suggestive ($p \leq 1 \times 10^{-5}$) SNP ^ gene contained at least one significant ($p \leq 1 \times 10^{-8}$) SNP.

Table 4.2 The location of the most significant QTLs, limited to the top 5, which were associated with chest width, and the genes located within these QTLs within each breed (AA=Angus; CH=Charolais;HE=Hereford; LM=Limousin; SE=Simmental; HF=Holstein-Friesian).

Breed	Chr	Start	End	No of suggestive and significant SNPs	Most significant SNP	P-Value	Allele frequency of positive allele						Candidate genes within this QTL
							AA	CH	HE	LM	SI	HF	
Angus	8	19919026	20930648	3	20426751[b]	$4.18x10^{-6}$	0.057	0.980	0.014	0.896	0.112	0.989	ELAVL2*
	10	101530896	102548539	4	102040999[b]	$5.32x10^{-7}$	0.232	0.678	0.305	0.306	0.216	0.223	TTC8, FOXN3*
	11	52112729	53133828	13	52632756[a]	$1.85x10^{-6}$	0.639	0.241	0.895	0.908	0.898	0.105	
	12	12006341	13006349	2	12506341[a]	$2.69x10^{-8}$	0.073	0.047	0.943	0.000	0.918	0.963	VWA8, DGKH, TNFSF11, AKIP11
	28	2715813	4028680	11	3522966[d]	$1.14x10^{-6}$	0.832	0.806	0.201	0.000	0.808	0.220	SPRTN, TRIM67*
Charolais	3	75099566	76200376	43	75636445[b]	$2.40x10^{-7}$	0.111	0.071	0.000	0.000	0.903	0.168	CTH, LRRC7*, LRRC40
	9	12560255	13754168	9	13060255[a]	$3.52x10^{-7}$	0.990	0.016	0.991	0.985	0.987	0.000	MTO1, EEF1A1
	10	42104985	43116388	3	42604985[a]	$3.68x10^{-7}$	0.003	0.039	0.992	0.948	0.030	0.991	RPL36AL, MGAT2, ARF6, SOS2
	11	10962219	12944023	11	11462219[b]	$3.04x10^{-7}$	0.000	0.007	0.000	0.000	0.000	0.000	ALMS1, EGR4, SMYD5, CYP26B1, SFXN5*
	18	57584619	58600780	3	58084619[d]	$1.61x10^{-7}$	0.002	0.012	0.991	0.004	0.031	0.000	ENSBTAG00000014593*
Hereford	4	82061233	83396596	4	82561233[a]	$2.12x10^{-7}$	0.000	0.989	0.023	0.012	0.994	0.993	POU6F2
	6	38955125	39995325	14	39461621[a]	$6.63x10^{-7}$	0.308	0.000	0.852	0.000	0.000	0.604	LCORL
	7	79663134	80729587	3	80197062[a]	$8.60x10^{-9}$	0.013	0.014	0.996	0.000	0.012	0.978	
	10	56443792	57546809	7	57025496[a]	$9.84x10^{-8}$	0.961	0.000	0.967	0.050	0.893	0.062	WDR72
	23	8222426	9377363	9	8722426[d]	$1.37x10^{-7}$	0.010	0.009	0.005	0.005	0.010	0.979	UHRF1BP1*, HMGA1, NUDT3, SCUBE3
Limousin	1	61741512	63549298	3	63048403[a]	$1.08x10^{-6}$	0.439	0.000	0.650	0.730	0.449	0.225	
	6	37530341	38792617	34	38284104[b]	$1.09x10^{-7}$	0.298	0.000	0.685	0.565	0.722	0.460	PPM1K, ABCG2*, PKD2*, SPP1, MEPE*, LAP3
	18	9391406	10382598	13	9891406[b]	$1.69x10^{-6}$	0.211	0.249	0.109	0.137	0.447	0.136	CDH13*, HSBP1, MLYCD
	18	55221720	56247875	3	55721720[d]	$1.39x10^{-6}$	0.995	0.000	0.002	0.003	0.000	0.000	LIG1, KCNJ14, CYTH2, RPL18, PPP1R15A
	20	7457546	8466248	16	7959103[a]	$9.82x10^{-7}$	0.057	0.008	0.990	0.987	0.979	0.000	UTP15, ANKRA2
Simmental	13	70061402	71118226	7	70573855[a]	$2.82x10^{-6}$	0.010	0.050	0.987	0.897	0.055	0.998	TOP1, PLCG1, LPIN3
	23	10151181	11174475	3	10651181[b]	$1.20x10^{-6}$	0.071	0.913	0.054	0.128	0.068	0.959	CPNE5*, PIM1, TMEM217, TBC1D22B
	23	30033517	31047355	4	30533517[c]	$1.67x10^{-6}$	0.099	0.072	0.219	0.187	0.902	0.146	ZSCAN31, ZKSCAN4, HIST1H2BB
	25	8699062	9699134	5	9199062[a]	$6.83x10^{-7}$	0.000	0.000	0.996	0.000	0.005	0.996	EMP2, NUBP1, CLEC16A
	26	48445354	49451650	6	48945354[a]	$2.03x10^{-6}$	0.989	0.008	0.996	0.041	0.996	0.000	
Holstein Friesian	1	57000435	58139976	15	57582901[b]	$7.42x10^{-7}$	0.034	0.225	0.209	0.688	0.000	0.195	ABHD10, CD200*, ATG3, CCDC80
	2	30344158	31344250	3	30844250[a]	$7.56x10^{-7}$	0.003	0.000	0.994	0.987	0.998	0.990	TTC21B, GALNT3, CSRNP3
	10	39919494	41220895	5	40476976[a]	$1.04x10^{-6}$	0.004	0.004	0.975	0.991	0.982	0.861	MDGA3*
	13	78475631	79544490	9	79027846[a]	$9.15x10^{-7}$	0.059	0.098	0.045	0.765	0.840	0.073	SNAI1, UBE2V1, PTPN1
	24	827290	2268995	9	1331600[a]	$4.86x10^{-7}$	0.064	0.067	0.986	0.000	0.924	0.013	PQLC1, KCNG2, NFATC1, ATP9B

Superscript denotes SNP classification: [a]intergenic, [b]intron, [c]upstream gene variant, [d]downstream gene variant. Symbols denote the significance of SNPs within genes: *gene contained at least one suggestive ($p \leq 1 \times 10^{-5}$) SNP ^ gene contained at least one significant ($p \leq 1 \times 10^{-8}$) SNP.

Table 4.3 The location of the most significant QTLs, limited to the top 5, which were associated with hip width or rump width, and the genes located within these QTLs within each breed (AA=Angus; CH=Charolais;HE=Hereford; LM=Limousin; SE=Simmental; HF=Holstein-Friesian).

Breed	Chr	Start	End	No of suggestive and significant SNPs	Most significant SNP	P-Value	Allele frequency of positive allele						Candidate genes within this QTL
							AA	CH	HE	LM	SI	HF	
Angus	4	115417450	116432669	15	115922671[b]	$6.53x10^{-7}$	0.031	0.925	0.109	0.840	0.788	0.268	KMT2C, ACTR3B*,XRCC2, CCT8L2
	5	30902961	31924821	5	31402961[a]	$2.79x10^{-7}$	0.002	0.992	0.978	0.006	0.002	0.990	RHEBL1, PRKAG1, WNT1, WNT10B, CCDC65
	11	81485390	82623280	5	81985390[b]	$1.09x10^{-6}$	0.004	0.000	0.000	0.996	0.994	0.000	FAM49A*
	20	13855925	14889348	17	14374205[a]	$1.16x10^{-6}$	0.004	0.021	0.013	0.000	0.016	0.002	TRIM23, ADAMTS6
	25	15156974	16246007	4	15656974[a]	$5.63x10^{-7}$	0.003	0.983	0.011	0.973	0.974	0.000	XYLT1
Charolais	2	30205997	31264765	30	30705997[a]	$2.83x10^{-8}$	0.978	0.993	0.995	0.008	0.007	0.267	GALNT3*, SCN1A, SCN2A, SCN3A, TTC21B, CSRNP3
	8	4328030	5328051	4	4828030[b]	$1.09x10^{-6}$	0.000	0.005	0.004	0.028	0.026	0.997	GALNTL6*
	9	12598999	13731582	8	13113448[a]	$6.49x10^{-7}$	0.990	0.024	0.010	0.009	0.011	0.027	MTO1, EEF1A1
	15	7774063	8881109	3	8274063[b]	$2.59x10^{-7}$	0.000	0.004	0.000	0.005	0.998	0.000	ARHGAP42*
	28	5674318	6741712	5	6241712[c]	$1.09x10^{-6}$	0.002	0.004	0.996	0.000	0.006	0.000	PCNX2*
Hereford	4	105760789	106772084	113	106265147[a]	$2.78x10^{-7}$	0.596	0.432	0.695	0.000	0.572	0.521	TAS2R3, TAS2R4, TAS2R38
	8	4170402	5731161	6	4670402[b]	$3.39x10^{-6}$	0.000	0.989	0.997	0.000	0.000	0.000	GALNTL6*, GALNT7
	13	53374292	54375561	4	53874292[a]	$2.94x10^{-6}$	0.784	0.292	0.690	0.727	0.309	0.880	STK35, PDYN, SIRPA
	14	5352193	6396755	6	5852193[a]	$4.29x10^{-6}$	0.000	0.000	0.986	0.000	0.000	0.000	COL22A1, FAM135B
	18	21513927	22756651	3	22256651[b]	$3.63x10^{-6}$	0.983	0.006	0.008	0.993	0.991	0.040	CHD9, RBL2, RPGRIP1L*, FTO*, IRX3
Limousin	5	16612583	17626967	5	17112583[a]	$4.66x10^{-7}$	0.030	0.000	0.008	0.983	0.994	0.066	
	6	32350666	34490506	812	33611754[a]	$1.95x10^{-9}$	0.246	0.880	0.366	0.150	0.232	0.819	
	6	37341111	40835172	153	38030341[b]	$1.55x10^{-9}$	0.084	0.000	0.000	0.126	0.366	0.006	ABCG2^, PKD2^, SPP1*, MEPE, LAP3, NCAPG*, LCORL*
	13	76534127	77546426	23	77045666[d]	$4.01x10^{-6}$	0.962	0.000	0.988	0.041	0.023	0.101	NCOA3, SULF2
	21	38149733	39222453	23	38702258[a]	$3.41x10^{-7}$	0.000	0.940	0.003	0.002	0.997	0.000	
Simmental	1	79028842	80104503	3	79590057[b]	$1.77x10^{-7}$	0.022	0.040	0.000	0.040	0.005	0.027	LPP*
	10	86379935	87382277	3	86879935[c]	$1.13x10^{-6}$	0.009	0.988	0.000	0.000	0.995	0.000	YLPM1, PGF, EIF2B2, MLH3, ACYP1, ZC2HC1C, NEK9, TMED10
	11	24184879	25302455	4	24684879[a]	$1.36x10^{-6}$	0.000	0.000	0.000	0.000	0.998	0.006	PKDCC
	18	9064056	10795231	11	10281382[a]	$3.42x10^{-7}$	0.000	0.000	0.000	0.000	0.986	0.040	CDH13*, OSGIN1, MBTPS1, DNAAF1, TAF1C
	22	25717794	30456249	16	29136317[a]	$1.11x10^{-6}$	0.000	0.030	0.000	0.987	0.997	0.000	CHL1*, CNTN3, PDZRN3, GXYLT2
Holstein Friesian	1	8144528	9875908	27	9335614[a]	$1.37x10^{-6}$	0.097	0.209	0.206	0.000	0.783	0.226	ADAMTS1, ADAMTS5, APP
	9	31692809	33191394	7	32273403[a]	$3.55x10^{-6}$	0.005	0.021	0.006	0.973	0.030	0.995	MAN1A1*, ASF1A, CEP85L, PLN, SLC35F1
	13	78476376	79544490	4	78976376[a]	$1.23x10^{-6}$	0.862	0.640	0.923	0.629	0.701	0.230	SNAI1, UBE2V1, PTPN1
	20	63192522	64260191	3	63722163[a]	$2.03x10^{-7}$	0.003	0.025	0.000	0.023	0.995	0.995	TAS2R1, SEMA5A
	24	49503031	50528738	3	50024697[b]	$5.55x10^{-7}$	0.081	0.901	0.938	0.066	0.023	0.936	ACAA2, MYO5B*, MBD1, CXXC1

Superscript denotes SNP classification: [a]intergenic, [b]intron, [c]upstream gene variant, [d]downstream gene variant. Symbols denote the significance of SNPs within genes: *gene contained at least one suggestive ($p \leq 1 \times 10^{-5}$) SNP ^ gene contained at least one significant ($p \leq 1 \times 10^{-8}$) SNP.

Table 4.4 The location of the most significant QTLs, limited to the top 5, which were associated with back length, and the genes located within these QTLs within each breed (AA=Angus; CH=Charolais;HE=Hereford; LM=Limousin; SE=Simmental).

Breed	Chr	Start	End	No of suggestive and significant SNPs	Most significant SNP	P-Value	Allele frequency of positive allele					Candidate genes within this QTL
							AA	CH	HE	LM	SI	
Angus	6	37939769	40455422	70	38443019[a]	$5.79x10^{-7}$	0.139	0.000	0.847	0.207	0.311	PKD2, SPP1, MEPE, LAP3, NCAPG*, LCORL*
	6	40762050	42494936	24	41262050[b]	$8.44x10^{-7}$	0.032	0.003	0.000	0.000	0.000	SLIT2*, PACRGL, KCNIP4*
	9	11789073	12803143	4	12298383a	$1.17x10^{-6}$	0.008	0.032	0.979	0.987	0.969	RIMS1, KCNQ5
	12	84208854	85283107	29	84720853[a]	$6.13x10^{-8}$	0.949	0.000	0.013	0.035	0.981	
	13	68993173	70000878	3	69495192[a]	$6.75x10^{-7}$	0.026	0.000	0.032	0.073	0.060	
Charolais	2	1	10036842	5525	6808074[a]	$3.96x10^{-48}$	0.000	0.079	0.000	0.972	0.996	WDR75, ASNSD1^, ARHGEF4^, MYO7B^, NAB1^, MFSD6^, MSTN^, PMS1^, ORMDL1^, COL3A1^, COL5A2^, ANKAR^, SLC40A1^
	14	33353270	34356964	4	33853270[a]	$1.19x10^{-7}$	0.000	0.013	0.000	0.000	0.992	ARFGEF1, CPA6, PREX2
	14	44425358	45430890	3	44928243[a]	$7.51x10^{-7}$	0.209	0.273	0.605	0.423	0.423	STMN2, HEY1, MRPS28
	28	19217733	21371343	36	19836248[a]	$1.01x10^{-7}$	0.418	0.784	0.575	0.583	0.626	NRBF2, REEP3*
	28	30350477	31864396	38	31332353[a]	$6.88x10^{-9}$	0.450	0.859	0.629	0.426	0.629	KAT6B*, DUPD1, DUSP13, VDAC2
Hereford	4	1	910718	5	223774[a]	$1.15x10^{-6}$	0.975	0.984	0.981	0.981	0.000	VSTM2A*
	4	37522586	38567213	13	38055263[a]	$2.31x10^{-6}$	0.959	0.283	0.130	0.851	0.201	PCLO*
	8	85462715	87578203	16	86646431[a]	$1.63x10^{-6}$	0.000	0.000	0.998	0.000	0.000	OGN, ASPN, ECM2, IPPK, BICD2, FGD3, NINJ1, BARX1*,PTPDC1*
	14	30747311	31758061	7	31247311[a]	$3.41x10^{-6}$	0.429	0.333	0.485	0.636	0.648	BHLHE22, MTFR1
	18	29621954	30630622	5	30130622[a]	$8.58x10^{-7}$	0.996	0.995	0.996	0.985	0.010	CDH8
Limousin	1	66063243	67175049	15	66587440[b]	$2.16x10^{-7}$	0.002	0.030	0.983	0.997	0.018	GTF2E1, STXBP5L, POLQ*, FBXO40, HCLS1, GOLGB1
	3	24752329	26688150	3	26188150[d]	$9.48x10^{-7}$	0.000	0.897	0.908	0.917	0.888	SPAG17*, WDR3, MAN1A2, VTCN1*, TRIM45, TTF2, CD101, PTGFRN
	6	32025422	34384319	1058	33661101[a]	$5.14x10^{-13}$	0.753	0.904	0.407	0.142	0.259	ATOH1
	6	36996616	41253691	469	38792702[b]	$4.24x10^{-13}$	0.097	0.000	0.105	0.091	0.000	ABCG2^, PKD2^, SPP1*, MEPE, LAP3, NCAPG^, LCORL^, SLIT2
	21	33476048	34502357	6	33999605[a]	$1.55x10^{-6}$	0.006	0.017	0.005	0.017	0.005	CSPG4, SNX33, IMP3, PTPN9
Simmental	15	77047714	78087312	9	77558153[b]	$5.09x10^{-7}$	0.811	0.000	0.270	0.000	0.264	DGKZ, ATG13, ARHGAP1, ZNF408, CKAP5*
	16	10050545	11308116	5	10550545[a]	$6.88x10^{-7}$	0.000	0.000	0.000	0.000	0.981	
	17	62751558	63784022	12	63254862[b]	$1.24x10^{-6}$	0.047	0.940	0.977	0.930	0.969	LHX5*, PLDB2, OAS2, OAS1Y, OAS1X
	20	43798108	44854685	5	44298108[a]	$2.56x10^{-6}$	0.042	0.069	0.240	0.074	0.109	
	21	10803227	11841095	7	11303227[a]	$2.88x10^{-6}$	0.998	0.012	0.980	0.006	0.994	NR2F2

Superscript denotes SNP classification: [a]intergenic, [b]intron, [c]upstream gene variant, [d]downstream gene variant. Symbols denote the significance of SNPs within genes: *gene contained at least one suggestive ($p \leq 1 \times 10^{-5}$) SNP ^ gene contained at least one significant ($p \leq 1 \times 10^{-8}$) SNP.

Table 4.5 The location of the most significant QTLs, limited to the top 5, which were associated with chest depth, and the genes located within these QTLs within each breed (AA=Angus; CH=Charolais;HE=Hereford; LM=Limousin; SE=Simmental).

Breed	Chr	Start	End	No of suggestive and significant SNPs	Most significant SNP	P-Value	Allele frequency of positive allele					Candidate genes within this QTL
							AA	CH	HE	LM	SI	
Angus	4	109535218	110566320	118	110035226[a]	$2.08x10^{-7}$	0.003	0.000	0.000	0.000	0.998	CNOT4*
	8	51491571	52874502	6	52374502[b]	$1.55x10^{-7}$	0.004	0.011	0.998	0.990	0.000	OSTF1, PCSK5*
	18	42431986	42811277	6	41931986[a]	$8.41x10^{-8}$	0.004	0.003	0.000	0.014	0.996	
	19	25487490	26528596	129	25988404[b]	$2.68x10^{-7}$	0.003	0.953	0.003	0.977	0.076	PITPNM3*, UBE2G1, MYBBP1A, GGT6, PIMREG
	23	27713725	28798254	34	28273994[c]	$6.31x10^{-7}$	0.043	0.104	0.023	0.023	0.904	MIC1, TCF19, CCHCR1, VARS2, PPP1R18, TRIM26, TRIM15, TRIM10, TRIM40, TRIM31, TRIM39*, PPP1R11
Charolais	4	103847357	105940963	3	104347357[b]	$2.48x10^{-6}$	0.000	0.006	0.000	0.993	0.005	HIPK, SLC37A3, WEE2, SSBP1, PARP12*
	10	29295461	30295461	12	29796031[b]	$4.91x10^{-6}$	0.808	0.756	0.252	0.672	0.000	TMCO5B, SCG5
	10	75515119	76535772	7	76015119[b]	$1.17x10^{-6}$	0.006	0.995	0.995	0.980	0.987	KCNH5, PPP2R5E*, SYNE2
	12	81616525	82648669	15	82139001[a]	$4.14x10^{-6}$	0.053	0.100	0.000	0.027	0.921	NALCN, ITGB1
	14	49295193	50325837	6	49825837[b]	$1.24x10^{-6}$	0.916	0.795	0.285	0.185	0.860	UTP23, EIF3H*
Hereford	3	63308338	64320629	4	63808996[a]	$1.19x10^{-6}$	0.990	0.000	0.038	0.063	0.919	
	5	99016506	100071368	31	99516506[a]	$6.26x10^{-7}$	0.100	0.056	0.070	0.966	0.046	
	17	61625220	62663494	3	62157617[a]	$1.37x10^{-6}$	0.000	0.003	0.969	0.000	0.000	TBX3, TBX5
	18	41115715	42140232	4	41635699[a]	$3.00x10^{-6}$	0.997	0.014	0.002	0.997	0.000	ZNF536, TSHZ3
	20	9677922	10679487	5	10177922[b]	$2.93x10^{-6}$	0.863	0.257	0.741	0.666	0.219	MCCC2, BDP1, SERF1A, SMN2, SLC30A5
Limousin	5	26076148	27084460	3	26576148[c]	$8.02x10^{-7}$	0.000	0.004	0.007	0.009	0.000	HOXC4, HOXC5, HOXC6, HOXC8, HOXC8, HOXC9, HOXC10, HOXC11, HOXC12, HOXC13
	6	32350666	34308736	456	33560360[a]	$2.14x10^{-7}$	0.060	0.049	0.053	0.097	0.968	
	6	37037069	40568831	211	38075438[b]	$2.92x10^{-9}$	0.087	0.000	0.000	0.131	0.368	PPM1K, ABCG2^, PKD2^, SPP1, MEPE, LAP3, NCAPG*, LCORL*
	7	16966648	17927749	15	17466648[a]	$5.13x10^{-7}$	0.991	0.956	0.978	0.052	0.941	EBF1*
	11	77828096	78855720	3	78355720[a]	$5.74x10^{-7}$	0.000	0.000	0.000	0.003	0.000	GDF7, RHOB, SDC1
Simmental	2	97634951	98536954	3	98035848[b]	$2.77x10^{-7}$	0.000	0.002	0.000	0.000	0.004	KANSL1L, ACADL, MYL1
	11	42337336	43357452	3	42837336[a]	$4.45x10^{-7}$	0.865	0.815	0.000	0.975	0.991	BCL11A, GTF2A1L*
	21	50755259	51864196	11	51364196[a]	$4.44x10^{-8}$	0.000	0.002	0.000	0.002	0.998	LRFN5
	24	49238747	50334349	12	49739134[d]	$4.03x10^{-7}$	0.997	0.002	0.005	0.005	0.995	CDH2*, DYM, ACAA2, MYO5B
	27	9276392	10276408	3	9776396[a]	$3.29x10^{-7}$	0.000	0.007	0.000	0.975	0.998	

Superscript denotes SNP classification: [a]intergenic, [b]intron, [c]upstream gene variant, [d]downstream gene variant. Symbols denote the significance of SNPs within genes: *gene contained at least one suggestive ($p \leq 1 \times 10^{-5}$) SNP ^ gene contained at least one significant ($p \leq 1 \times 10^{-8}$) SNP.

Table 4.6 Fold enrichment/depletion of SNPs in each annotation class in each trait in each breed (AA=Angus; CH=Charolais; HE=Hereford; LM=Limousin; SE=Simmental)

		3' UTR variant	5' UTR variant	Downstream gene variant	Intergenic variant	Intron variant	Missense variant	Missense Variant & Splice	Non Coding Transcript	Splice region variant	Stop gained	Synonymous variant	Upstream gene variant
WH	AA	5.70	-	0.83	1.04	0.95	0.79	-	-	-	-	0.86	0.61
	CH	-	-	0.54	1.24	0.50	-	-	-	-	-	0.69	0.67
	HE	-	-	0.23	1.25	0.54	0.71	-	-	-	-	0.52	0.45
	LM	0.21	-	0.53	1.28	0.39	-	-	1.29	0.66	-	0.54	0.79
	SI	1.11	-	0.31	1.10	0.92	0.78	-	-	-	-	1.14	0.37
	HF	4.41	-	3.20	0.93	0.82	2.32	-	-	-	-	1.68	1.37
BL	AA	-	-	0.53	1.09	0.87	1.65	-	-	-	-	0.80	0.77
	CH	2.94	0.41	1.21	1.00	1.04	0.43	-	-	-	6.35	1.18	0.41
	HE	3.80	4.37	1.22	1.02	0.91	1.97	-	-	-	68.90	0.48	0.82
	LM	-	1.26	0.58	1.26	0.44	0.95	7.85	1.67	0.85	-	0.42	0.72
	SI	-	-	1.09	0.76	1.56	2.70	-	-	-	135.73	0.99	1.16
CW	AA	-	-	1.11	0.99	1.12	-	-	-	-	-	1.11	0.36
	CH	-	5.66	1.33	0.87	1.36	1.71	-	-	-	-	1.90	0.41
	HE	1.20	-	2.22	1.16	0.46	-	-	-	-	-	0.61	0.98
	LM	-	-	2.25	0.83	1.36	-	-	-	-	-	2.26	0.36
	SI	2.84	-	2.08	0.96	0.99	-	-	-	-	-	-	0.93
	HF	-	-	3.85	1.09	0.44	-	-	-	-	-	-	1.06
CD	AA	-	-	1.15	0.82	1.41	1.43	-	12.15	9.56	-	6.21	0.56
	CH	-	-	1.23	0.96	0.96	-	-	-	-	-	1.39	2.08
	HE	-	-	1.11	1.21	0.46	-	-	9.67	-	-	0.83	0.98
	LM	-	-	0.39	1.31	0.39	-	-	-	-	-	0.54	0.29
	SI	-	-	0.77	1.20	0.51	1.07	-	-	-	-	2.34	1.00
HW	AA	3.31	-	0.28	1.01	1.10	-	-	-	-	-	3.32	0.36
	CH	2.96	-	1.00	0.92	1.20	-	-	-	-	-	1.50	1.13
	HE	1.78	-	0.30	1.18	0.53	1.23	-	-	-	-	1.81	1.64
	LM	1.54	-	0.50	1.30	0.38	-	-	-	1.19	-	0.20	0.40
	SI	3.01	-	1.19	0.79	1.53	-	-	-	-	-	-	0.99
	HF	-	-	1.68	1.05	0.59	-	-	29.22	-	-	-	2.44

4.5.6 Across Trait Overlap

Quantitative trait loci associated with two or more skeletal traits were identified within each breed (Appendix B4). The *NCAPG* and *LCORL* genes were identified as pleiotropic genes associated with all 5 traits in the LM breed and with both WH and BL in the AA breed. There were also suggestive genomic windows in common between CW and HW in AA with 5 windows on BTA4 and a single window on BTA8 being common to both of these traits. These 5 windows on BTA4 contained six SNPs that were suggestively associated with both CW and HW; all 6 of these SNPs were intronic SNPs located within the *ENSBTAG00000008032* gene. No gene was located within the 1kb window on BTA8.

A greater overlap in QTLs associated with both WH and BL was identified in the CH and HE. Ten 1kb windows were associated with both WH and BL in the CH, nine of which were located on BTA28. Eight 1kb windows overlapped between WH and BL in the HE with 6 windows located on BTA23 encompassing the *GMDS* gene. Further overlap among traits was identified in the CH breed where 3 windows on BTA9 and 3 windows on BTA19 were associated with both WH and CW. The SI breed had the fewest number of pleiotropic associations of all beef breeds, as only one window on BTA12 near the *SPRY2* gene was suggestively associated with both WH and BL. The only overlap in associated QTLs between the beef and dairy breeds was in WH/Stature between AA and HF. These breeds had 2 overlapping 1kb windows on BTA5 but no obvious candidate genes were identified in this region.

4.5.7 Enrichment of SNPs

Intergenic SNPs were the most common annotation class of SNPs associated with each trait in each breed. This annotation class was enriched for all traits in HE, 4 traits in LM

(WH, BL, CD and HW), 3 traits in SI (WH, CW, and CD) and AA (WH, BL, and HW), and 2 in both CH (WH and BL) and HF (CWD and RW; Table 4.6). The second most common annotation was the intronic SNPs; this class was enriched for 3 traits in AA (CW, CD, and HW) and CH (BL, CW, and HW) and 2 traits in SI (BL and HW). Downstream gene variants were enriched in all breeds for CW and at least one breed for all the remaining traits (Table 4.6). Stop-gain SNPs that were significantly associated with BL were enriched in all breeds in which they were associated.

4.6 Discussion

Several QTLs were discovered in the present study to be associated with each of the skeletal type traits although the majority of these regions, excluding the *NCAPG*/*LCORL* locus in the LM population, were unique to a single trait or a single breed. This indicates the existence of breed-specific and trait-specific QTL for skeletal traits which has implications for the usefulness of such QTL in across breed genomic evaluations where only purebreds are used. Previous studies have documented both across-breed and breed-specific QTL associated with carcass traits, birth weight, weaning weight, and mature weight (Saatchi et al., 2014b), as well as dry matter intake, growth and feed efficiency (Saatchi et al., 2014a), carcass traits (Purfield et al., 2019), and muscular type traits (Chapter 3) in beef cattle. Excluding stature (Bouwman et al., 2018), the present study is the first published genome study on the skeletal linear type traits in beef cattle using imputed sequence data and is one of few genome-based studies comparing QTLs across multiple breeds of cattle. The present study, however, also incorporated imputed genome sequence information on 4,494 dairy cattle to compare to the beef animals. This comparison is rarely carried out (Purfield et al., 2015) as such multi-breed data are not always readily available for incorporation into the same study.

Nonetheless, the difference in age at classification between the beef and dairy animals varied substantially with the beef animals all being < 16 months and the dairy animals > 23 months when assessed. Previous heritability estimates of the linear type traits assessed in the dairy cows were all ≥ 0.26 (Berry et al., 2004) indicating these traits are, however, expected to be moderately to highly repeatable over time. This was substantiated by the fact that some common QTL were detected for Angus and Holstein-Friesian.

An earlier study on the beef cattle population from the dataset used in the present study (Chapter 2) summarized the heritability estimates of, and genetic correlations among, the skeletal type traits in each breed. In general, the genetic variance within each trait and the correlations between each trait differed by breed indicating that breed-specific and trait-specific QTL may be underlying these traits. Similarities were observed between CH and LM in terms of heritability estimates and genetic correlations (Chapter 2) from this it was theorized that the genetic architecture of these breeds may be quite similar. The present study is an advance of this study (Chapter 2) where the contributors to the genetic variation within and across breeds have been identified.

Type traits have previously been proposed as potential early predictors of carcass weight and conformation (Conroy et al., 2010) and of overall carcass merit (Berry et al., 2019) given the genetic correlations between these traits and linear type traits are generally moderate to strong. However, as these correlations are not unity, two animals with the same live-weight may be morphologically very different which may lead to very different carcass value owing to the distribution of primal cuts (Berry et al., 2019). Therefore, type traits may be useful in future multi-trait genetic and genomic evaluations as they provide more information than live-weight alone. Consequently,

knowledge of the QTLs associated with the skeletal traits could be used in these genome-based evaluations as part of a multi-trait evaluation targeting the altering of the morphology of an animal to increase the output of goal trait high quality primal cuts thus improving the profitability of the farm system.

In total, over 90% of the QTLs identified in the present study have been previously documented to be associated with other production traits in beef or dairy cattle when compared to those within the Cattle QTLdb database (Accessed 08 January 2019). Of the top 140 QTLs associated with the skeletal type traits (Tables 4.2 to 4.6), 80 of these had previously been identified as being associated with body weight at either birth (Lu et al., 2013), as a yearling (Snelling et al., 2010), as a weanling (Saatchi et al., 2014b), at slaughter (Sherman et al., 2008), or at maturity (Saatchi et al., 2014b). Furthermore, some of the top 140 QTLs were also previously associated with carcass weight (McClure et al., 2010; Saatchi et al., 2014a) and residual feed intake (Nkrumah et al., 2007; Lu et al., 2013; Saatchi et al., 2014a) in cattle. Nineteen QTLs identified in the present study have also been identified previously as being associated with linear type traits describing the muscular characteristics of cattle (Chapter 3).

4.6.1 Across-breed Comparison

With the exception of the *NCAPG* and *LCORL* genes, the majority of QTLs associated with the skeletal type traits were breed-specific and in many cases, also trait specific. The differences observed in associated QTLs among the breeds may be due to epistatic or gene-by-environment interactions, or simply due to differences in the power to detect significance due to the large differences in population sizes among the breeds (Saatchi et al., 2014b). The age difference between the dairy and beef animals when classified may also have contributed to some of the inconsistencies in discovered QTL between

the dairy and beef cattle. In many cases, the SNPs detected to associate with a trait in one breed were not segregating in all 5 breeds. Observed differences in detected QTL among the breeds may also be due to limitations in imputation where the imputed genotypes may not be perfect; this may result in the causal SNP not being identified as the most significant association especially if that SNP is rare among the populations (Bouwman et al., 2018).

Both *NCAPG* and *LCORL* are widely accepted as being associated with stature in many mammals including cattle (Bouwman et al., 2018), humans (Gudbjartsson et al., 2008), and horses (Tetens et al., 2013); therefore it was not unexpected that these genes were associated with all the skeletal traits in the LM population and with BL and WH in AA. The *NCAPG* and *LCORL* genes have also been previously linked to growth and carcass traits in SI (Zhang et al., 2018), carcass weight in AA, CH, and LM (Purfield et al., 2019), and with both feed intake and body weight gain in a population containing 14 different breeds of cattle (Lindholm-Perry et al., 2011). Interestingly, the QTL containing *NCAPG* and *LCORL* were not associated with any of the skeletal traits evaluated in SI or HF even though SNPs within these regions were segregating in both breeds. Although imputed sequence variants were used, we were unable to identify which of the two genes is causal; indeed none of the segregating missense variants within either gene were suggestively associated with any trait. However, a previous study that associated *LCORL* with growth and carcass traits in cattle, proposed that it is the non-coding and regulatory expression of *LCORL* that influences a trait (Han et al., 2017). This theory is further substantiated by the significant over-representation of the intergenic variant SNP class within the present study which suggests that it is the regulatory expression of many genes that influence animal morphology rather than the causative disruption of gene functionality.

4.6.2 Carcass Traits

Some skeletal linear type traits in beef cattle are moderately genetically correlated with carcass traits including carcass cut weights (Pabiou et al., 2012), primal cut yields (Berry et al., 2019), and rib and subcutaneous fat thickness (Mukai et al., 1995). Thus, it is not surprising that there was overlap among some of the QTLs associated with linear type traits in the present study with those previously reported for carcass traits. Across all breeds and traits, there were 22 QTLs associated with the skeletal type traits in the present study that have been previously associated with carcass weight (McClure et al., 2010; Nishimura et al., 2012; Sharma et al., 2014). Twelve of these QTL were located on BTA6 and incorporated the *NCAPG* and *LCORL* genes. Interestingly, the *NCAPG* and *LCORL* genes, while being associated with size have also been associated with subcutaneous fat thickness in beef cattle (Lindholm-Perry et al., 2011). More overlap among the QTLs associated with the skeletal type traits and fat thickness was on BTA2, where a QTL containing *MSTN* which was associated with BL and WH in CH has also been documented to be associated with fat thickness at the 12th rib (Casas et al., 1998).

In general, if an allele was associated with a wider or longer skeletal type trait, it also had the same effect direction on the other traits, i.e. if an allele was associated with wider CW it tended to be associated with deeper CD and vice versa. Interestingly, this was not always the case for the alleles associated with WH and BL indicating that some alleles associated with taller WH were associated with shorter BL; thus, the correlation between these two traits (Chapter 2) could be broken leading to a morphologically different animal. The knowledge of SNPs and QTLs that influence one or more traits of interest (e.g., a longer back but with better muscling) would enable the selection for the desired trait combinations despite any genetic antagonisms. Furthermore, including traits such as WH and BL in a multi-trait genetic evaluation for terminal beef cattle,

along with the other trait of interests (e.g., carcass weight, carcass conformation, and carcass fat) would provide more information on an animal's carcass and conformation than what is possible from the carcass traits alone.

4.6.3 Feed Intake & Efficiency

Feed intake is both genetically and phenotypically correlated with body weight and average daily gain (Arthur et al., 2001; Crowley et al., 2010); on average, bigger, heavier cattle tend to eat more. Feed is generally the greatest cost associated with beef production (Montano-Bermudez et al., 1990); thus, improvements in the efficiency of which feed is utilized should contribute to greater economic returns in the whole beef production system (Archer et al., 1999). Difficulty in selection for feed efficiency is mainly due to a lack of genetic evaluations for feed intake; data are generally readily available for the energy sink component of feed efficiency thus being hindered by data on feed intake. Feed intake is linked to the morphology of an animal (Crowley et al., 2011). While genomic evaluations for feed intake could be useful, the reference population required to generate accurate genomic evaluations are few. Having knowledge of potential QTL associated with feed intake, discovered using much larger dataset on correlated traits (i.e., the present study) could be used as prior information in such genomic evaluations (MacLeod et al., 2016); the correlated traits could also be considered in a multi-trait genomic evaluation.

Among the QTLs associated with at least one of the skeletal type traits, 51 QTLs were previously identified as being associated with feed intake (Nkrumah et al., 2007; Sherman et al., 2010; Lindholm-Perry et al., 2011; Lu et al., 2013; Saatchi et al., 2014a) while 80 were previously identified as being associated with body weight at various stages of the animal's life (Sherman et al., 2008; Snelling et al., 2010; Saatchi et al.,

2014a) and body weight gain (Snelling et al., 2010). Given the generally small dataset sizes used in genomic analyses of feed intake traits, the QTL detected from the present study could actually be used as prior information in Bayesian-type analyses for genomic analyses (including genomic predictions) for traits like feed intake where the dataset size is limiting; such an approach could be deployed using models similar to those proposed by (MacLeod et al., 2016).

4.6.4 Calving Difficulty

The difficulty or ease of calving has long been thought to be related to the conformation of the dam (Ali et al., 1984) and the size of the calf (Sieber et al., 1989). Cows with wider hips and long rumps generally have larger internal pelvic openings which in turn lead to an easier calving; cows with smaller pelvic areas have more difficulty calving (Ali et al., 1984). Moreover, bigger, heavier calves are often more difficult to calve than their smaller, lighter counterparts (Sieber et al., 1989). It is, therefore, no surprise that 58 QTLs associated with the skeletal (i.e., size) type traits have previously been documented to be associated with calving difficulty in cattle (Purfield et al., 2015; Sahana et al., 2015). Seven of these 58 QTLs were associated with HW or RW in the present study; these QTLs were located on BTA1 in AA, BTA14 in HE, BTA6, BTA13, and BTA21 in LM, BTA10 in SI, and BTA1 in HF. None of the lead SNPs in these QTLs were segregating in all 6 breeds and a number of the lead SNPs were close to fixation for either the positive (i.e., wider hips) or negative (i.e., narrower hips) allele depending on the breed. Knowledge of the underlying quantitative trait variant associated with different morphological characteristics facilitates the development of more precise mating advice systems over and above consideration of the holistic calving difficulty estimate breeding values based on genome-wide quantitative trait variants.

For example, the choice of mate for a female with a genetic predisposition for a wide pelvic is likely to differ from that of a female with a narrower pelvic area; knowledge of genetic merit of the mate for different skeletal characteristics, even with the same estimated breeding value for calving difficulty should be exploited in the decision.

4.6.5 Omnigenic Model of Complex Traits

It has long been hypothesized that many genes, each with a small effect size, underlie complex traits that do not exhibit simple Mendelian inheritance (Fisher, 1918). In recent years, and with the advancement of genomic technology, many studies have reported that even the most significant loci across the genome associated with a trait have small effect sizes and only explain a small percentage of the predicted genetic variance (Wood et al., 2014; Boyle et al., 2017). The term omnigenic has been used to describe the phenomenon whereby a very large number of genes with seemingly no relevance to the trait of interest are associated with that trait due to being in the same regulatory networks as the relevant genes (Boyle et al., 2017). The results of the individual genome-based analyses in the present study, where many SNPs of small effect, often located within regulatory regions were associated with each trait within each breed, confirms that a complex omnigenic genetic architecture underlies the skeletal type traits in the 6 cattle breeds.

Despite millions of SNPs being tested for associations with each of the skeletal traits investigated, only 140 of the SNPs suggestively or significantly associated with a trait were located within the coding regions of the genome. The majority (i.e., 57.2%) of SNPs associated with any trait were intergenic SNPs; the number of intergenic SNPs and also 3’ UTR and 5’ UTR variants were enriched for the majority of traits they were associated with in each breed, demonstrating the importance of regulatory networks

within the genome to the cattle skeletal traits. Inference could also be drawn, therefore, on the contribution of regulatory regions to the correlated traits like carcass merit and feed intake. Downstream and upstream gene variants were also enriched in many of the traits. In general, the SNPs located within, or close to, the genes identified as candidate genes were located within these non-coding or regulatory regions. For example, 22 SNPs that were suggestively or significantly associated with WH in LM were located within the *LCORL/NCAPG* gene; 19 of these were intronic variants and 3 were downstream gene variants. Thus regulatory non-coding regions, while not having an effect on the coding sequence of a gene, may be of particular importance for cattle skeletal development via the proposed omnigenic model (Boyle et al., 2017).

4.8 Conclusion

While many QTL were identified as being associated with each trait in each breed, a large-effect QTL on BTA6 containing the *NCAPG* and *LCORL* genes was the only QTL associated with more than two traits and in more than one breed. This indicates that while the *NCAPG* and *LCORL* genes may affect multiple traits in multiple breeds, the majority of QTL underlying the skeletal type traits are both trait-specific and breed-specific. This has implications on the perceived usefulness of across-breed genomic evaluations for the component traits as well as possibly their correlated economically-important traits (e.g., carcass merit, feed intake) based solely on purebreds. Many of the QTLs identified in the present study have previously been documented to be associated with a number of other performance traits in cattle, including carcass traits, feed intake and calving difficulty.

Chapter 5

Identification of genomic regions that exhibit sexual dimorphism for size and muscularity in cattle

5.1 Preface

At the time of thesis submission this chapter was submitted to Journal of Animal Science.

Jennifer Doyle was primary author, performed the data edits and analysis and drafted the manuscript. Donagh Berry and Tom Moore conceived the study. Deirdre Purfield and Tara Carthy helped with imputation of the X chromosome and helped draft the manuscript. Siobhan Walsh and Roel Veerkamp helped draft the manuscript. Ross Evans supplied the data the analysis was performed on.

5.2 Abstract

Sexual dimorphism, the phenomenon whereby males and females of the same species are distinctive in some aspect of appearance or size, has previously been documented in cattle for traits such as growth rate and carcass merit using a quantitative genetics approach. No previous study in cattle has attempted to document sexual dimorphism at a genome level; therefore, the objective of the present study was to determine if genomic regions associated with size and muscularity in cattle exhibited signs of sexual dimorphism. Analyses were undertaken on 10 linear type traits that describe the muscular and skeletal characteristics of both males and females of 5 beef cattle breeds; 1,444 Angus (AA), 6,433 Charolais (CH), 1,129 Hereford (HE), 8,745 Limousin (LM), and 1,698 Simmental (SI). Genome wide association analyses were undertaken using imputed whole-genome sequence data for each sex separately by breed. For each SNP that was segregating in both sexes, the difference between the allele substitution effect sizes for each sex, in each breed separately, was calculated. Suggestively ($p \leq 1 \times 10^{-5}$) sexually dimorphic SNPs that were segregating in both males and females were detected for all traits in all breeds, although the location of these SNPs differed by both trait and breed. Significantly ($p \leq 1 \times 10^{-8}$) dimorphic SNPs were detected for traits in just three traits in the AA, seven traits in the CH and three traits in the LM. The vast majority of all segregating autosomal SNPs (86% in AA to 94% in LM) had the same minor allele in both males and females. Differences ($p \leq 0.05$) in allele frequencies between the sexes were observed for between 36% (LM) and 66% (AA) of the total autosomal SNPs that were segregating in both sexes. Dimorphic SNPs were located within a number of genes related to muscularity and/or size including the NAB1, COL5A2, and IWS1 genes on BTA2 that are located close to, and thought to be co-inherited with, the MSTN gene. Overall, sexual dimorphism exists in cattle at the

genome level, but it is not consistent by either trait or breed. It is unlikely that consideration of sexual dimorphism in beef cattle will improve the accuracy of genomic predictions for the traits and breeds investigated in the present study at least.

5.3 Introduction

Sexual dimorphism is the phenomenon whereby males and females of the same species are distinctive in behaviour, size, or appearance (Berns, 2013). This is attributable to the combination of sex-specific genes on sex chromosomes, sex-specific expression of genes, and other regulatory mechanisms that are not yet widely understood (Pointer et al., 2013). Sex-dependent differences have been documented for a whole range of traits in different species ranging from colour, ornamentation, mating behaviour, and size (McPherson and Chenoweth, 2012; Berns, 2013; van der Heide et al., 2016). Sex is also known to have an influence on growth of body tissues and could, therefore, affect carcass composition and weight distribution within the body tissue (Berg and Butterfield, 1976). Sexual size dimorphism is likely to have originated in mammals during evolution due to competition among males for access to females; males would fight one another to gain access to females and the winner, generally the bigger, stronger animal would mate with more females (Kirkpatrick, 1987; Katz, 2008). In selective breeding systems, breeding males are selected on numerous desirable traits and consequently competition for mates has been diminished in domesticated animals. Nonetheless, evidence of sexual dimorphism based on quantitative genetics approaches have been reported for several economically important traits in cattle, including growth rate (Koch and Clark, 1955; Marlowe and Gaines, 1958; van der Heide et al., 2016) and carcass traits (Crews Jr and Kemp, 2001; Bittante et al., 2018).

Linear type traits describing the muscular and skeletal characteristics of an animal are scored globally in both dairy (Veerkamp and Brotherstone, 1997; Berry et al., 2004) and beef (Mc Hugh et al., 2012; Mazza et al., 2014) cattle. These traits are typically considered as being genetically the same in both males and females; estimated genetic correlations of near unity between the same linear type trait in different sexes of cattle substantiate this assumption (Doyle et al., 2018). Genetic correlations, however, are a manifestation of the cumulative effect of both linkage and pleiotropy across the entire genome and it is possible that the control of such traits by sex may differ in specific genomic locations. The objective, therefore, of the present study was to determine if genomic regions associated with size and muscularity in cattle exhibited signs of sexual dimorphism. This knowledge will be useful in informing breeding programmes of the potential improvement in accuracy achievable by evaluating males and females separately.

5.4 Materials & Methods

5.4.1 Phenotypic Data

Linear type traits are routinely scored in both registered and commercial beef herds by trained classifiers from the ICBF as part of the Irish national beef breeding programme (Mc Hugh et al., 2012; Berry and Evans, 2014). The type traits used in the present study describe the muscular and skeletal development of the animal and include development of the hind quarter (DHQ), inner thigh (DIT), and loin (DL), thigh width (TW), wither width (WOW), wither height (WH), back length (BL), hip width (HW), and chest width (CW) and depth (CD). The five muscular traits were scored (Appendix C.1) on a scale of 1 (narrow) to 15 (wide) while the five skeletal traits (Appendix C.1) were scored on a scale of 1 (short or narrow) to 10 (long/tall or wide). Data on these 10 traits were

available on 147,704 purebred Angus, Charolais, Hereford, Limousin, and Simmental beef cattle scored between 6 and 16 months of age between the years 2000 and 2016.

Data editing procedures and the justification for such edits are outlined in detail b in Chapters 2, 3, and 4. Animals were discarded from the dataset if the sire, dam, herd, or classifier was unknown, or the parity of the dam was not recorded. Parity of the dam was subsequently recoded into 1, 2, 3, 4, and ≥5. Contemporary group was defined as herd-by-scoring date generated separately within each breed; each contemporary group had to have at least five records. Each of the 10 traits were separately standardized to a common variance within classifier-by-year as described in detail by (Brotherstone, 1994). Following edits, data were available on 81,200 animals (Appendix C.2) consisting of 3,356 Angus (AA) , 31,049 Charolais (CH), 3,004 Hereford (HE), 35,159 Limousin (LM) and 8,632 Simmental (SI).

5.4.2 Generation of Adjusted Phenotypes

Prior to inclusion in the genome wide association analysis, all phenotypes were adjusted within breed in ASREML (Gilmour et al., 2009) using the model:

$$y_{ijkl} = \mu + HSD_i + AM_j + DP_k + Animal_l + e_{ijkl}$$

where y_{ijkl} is the linear type trait, μ is the overall mean, HSD_i is the fixed effect of herd-by-scoring date (11,130 levels), AM_j is the fixed effect of the age in months of the animal (11 classes from 6 to 16 months), DP_k is the fixed effect of the parity of the dam (1, 2, 3, 4 and ≥5), $Animal_l$ is the random additive genetic effect of the animal where $N(0,\mathbf{A}\sigma_a^2)$, and e is the random residual effect where $N(0,\mathbf{I}\sigma_e^2)$; σ_a^2 is the additive genetic variance, σ_e^2 is the residual variance, **A** is the numerator relationship matrix and **I** is an

identity matrix. The adjusted phenotype used in the subsequent analysis was the raw phenotype less the fixed effect solutions of HSD, AM and DP.

5.4.3 Genotype Data

Of the phenotypic dataset of 81,200 animals, 19,449 animals from the five beef breeds (Appendix C2) were imputed to whole genome sequence as part of a larger dataset of 638,662 multi-breed genotyped animals (Purfield et al., 2019). These 638,662 animals were genotyped using one of 7 different genotype panels as described previously in Chapters 3 and 4). The reference population used for imputation contained 90% male animals and 8% female animals; 2% of the reference population were of unknown sex. Each animal had to have a call rate ≥ 90% and only SNPs with a known chromosome and position on UMD 3.1, and SNPs with a call rate ≥ 90% within the panel were retained for imputation.

All autosomes of genotyped animals were imputed to whole genome sequence (WGS) following the steps outlined in Chapters 3 and 4. Imputation of the pseudo-autosomal region (PAR) and non-PAR regions of the X-chromosome was undertaken separately. The non-PAR region was imputed for males and females separately. The PAR region of the X chromosome was defined from 143,861,798 to 148,823,899 bp (Mao et al., 2016).

Regions of poor WGS imputation accuracy were discarded as described by Purfield et al. (2019). Furthermore, within each breed and each sex, all SNPs with a minor allele frequency (MAF) ≤ 0.002 were not considered further (Appendix C.2). The number of SNPs remaining for each sex in each breed is outlined in Appendix C.2.

5.4.4 Genome Wide Association Study

Whole genome association analyses were performed within each sex in each breed separately using a mixed linear model association analysis in GCTA (Yang et al., 2011). Autosomal SNPs from the original high density (HD) panel (i.e., 734,159 SNPs) were used to construct the genomic relationship matrix (GRM) for each sex within each breed as per Chapters 3 & 4 that used the data from the present study but in a combined analysis of both sexes. In the association analyses of the X chromosome, all males were coded as homozygous for one of the alleles for SNPs in the non-PAR region and heterozygous SNPs were accepted in the PAR region. The model used for the within-sex and within-breed analysis was

$$\mathbf{y} = \mu + \mathbf{xb} + \mathbf{u} + \mathbf{e}$$

where **y** is a vector of preadjusted phenotypes, μ is the overall mean, **x** is the vector of imputed genotypes, **b** is the vector of additive fixed effects of the candidate SNP to be tested for association, $\mathbf{u} \sim N(0,\mathbf{G}\sigma_u^2)$ is the vector of additive genetic effects, where **G** is the genomic relationship matrix calculated from the HD SNP genotypes, and σ_u^2 is the additive genetic variance, and $\mathbf{e} \sim N(0,\mathbf{I}\sigma_e^2)$ is the vector of random residual effects, where **I** is the identity matrix and σ_e^2 is the residual variance.

5.4.5 Dimorphism

For each SNP that was analysed in both sexes (i.e., segregating in both sexes), the difference between the allele substitution effect sizes for each sex, in each breed separately, was calculated using a t-test:

$$t = \left| \frac{b_m - b_f}{\sqrt{\frac{SE_m^2 + SE_f^2}{n_m + n_f}}} \right|$$

where bx is the allele substitution effect in males (m) and females (f), SE is the estimated standard error of the allele effect, and n is the respective sample size. The presence of dimorphism was determined at each SNP based on the calculated p-value from the t-test statistic. A SNP with a p-value $\leq 1 \times 10^{-5}$ was assumed to have a suggestively different allele effect in the two sexes while a SNP with a p-value $\leq 1 \times 10^{-8}$ was assumed to have a significantly different allele substitution effect in the two sexes.

5.4.6 QTL Detection

To identify QTL regions that were dimorphic in more than one trait or more than one breed, each chromosome was split into 1kb genomic windows and windows containing at least one suggestive ($p \leq 1 \times 10^{-5}$) or significant ($p \leq 1 \times 10^{-8}$) SNP were compared across the traits and breeds.

5.5 Results

The scale of measurement, mean, and standard deviation of the linear type traits in each sex in each breed is in Appendix C1. Single nucleotide polymorphisms with evidence of significant ($p \leq 1 \times 10^{-8}$) dimorphism were detected for some traits, while suggestively ($p \leq 1 \times 10^{-5}$) dimorphic SNP were detected for all of the traits in all 5 breeds; however, these SNPs differed both by trait and by breed.

5.5.1 Angus

A total of 16,541,913 SNPs were segregating in the 1,044 males and 15,402,160 SNPs were segregating in the 400 females. Of these, 15,008,408 SNPs were segregating in both the male and female populations (Appendix C.2). Significant dimorphism ($p \leq 1 \times 10^{-8}$) was evident for a total of 7 SNPs across just three traits (HW, TW, and DIT; Table 5.1) while suggestive dimorphism ($p \leq 1 \times 10^{-5}$) was evident for between 31 (DHQ) and 1,254 (HW) SNPs depending on the trait (Table 5.1). In general, the allele substitution effects of the dimorphic SNPs tended to be in opposite directions in each sex (i.e., if the allele effect in the male population was negative, then the allele effect of the same allele in the female population was positive or vice versa; Table 5.2 and 5.4). The allele effects in the male population also tended to be closer to zero than those in the female population (Table 5.2 and 5.4) and the most significantly dimorphic traits tended to have a very low MAF in the female population. Of the muscular traits investigated, the most significantly dimorphic SNP was an intronic SNP ($p = 3.45 \times 10^{-9}$) located within the ADGRA3 gene on BTA6 and was associated with DIT (Table 5.2); this SNP had an allele effect of +0.12 (SE = 0.13) and a MAF of 0.026 in males but an allele effect of -3.68 (SE 0.63) and a MAF of 0.003 in females. Of the skeletal traits, the most significantly dimorphic SNP was an intergenic SNP, rs208222963 ($p = 9.86 \times 10^{-10}$), located on BTA8 that was associated with HW (Table 5.3; this SNP had an allele substitution effect of -0.23 (SE = 0.10) in males but +2.29 (SE = 0.40) in females with a MAF of 0.025 in the males and 0.005 in the females. Suggestively associated dimorphic SNPs with a higher MAF tended to have a smaller allele effect size than those with a low MAF; one such SNP, rs109325958 ($p = 9.43 \times 10^{-7}$) with a MAF of 0.497 in males and 0.424 in females was an intronic variant located within the KCNIP4 gene on BTA6

that had a dimorphic association with CW (Table 5.3) and had an allele effect -0.05 in males but +0.27 in females.

Of the 1kb windows containing at least one suggestively associated dimorphic SNP, there was little overlap between the muscular and skeletal groups of traits. The only overlap in windows between the muscular and skeletal traits was between HW, DL and WOW (two windows on each of BTA23 located at 14.555Mb and 51.713Mb), between HW and TW (one window on BTA8 located at 66.264Mb), and between HW and DL (two windows on BTA27 at 18.141Mb and 18.403Mb). Only one 1kb genomic window was suggestively associated with three skeletal traits (WH, BL, and HW; Appendix C.3a) and this was located between 66.386Mb and 66.387Mb on BTA8, within the ENSBTAG00000006446 gene. The greatest overlap across traits in AA was between WH and HW where a total of 12 1kb windows across 5 chromosomes suggestively exhibited sexual dimorphism (Appendix C.3a). Similar to the skeletal traits, only one 1kb window was common to more than two muscular traits (DL, DIT, and TW) and this was located between 59.524Mb and 59.525Mb on BTA24 (Appendix C.4a). The largest overlap across all skeletal traits was between DL and TW where 7 1kb windows exhibited suggestive dimorphism. Minimal overlap was detected among the remaining skeletal traits

5.5.2 Charolais

A total of 18,054,274 SNPs were segregating in the 4,641 CH males and 17,448,948 SNPs were segregating in the 1,792 CH females. Of these, 17,227,625 SNPs were segregating in both the male and female animals. Evidence of suggestive dimorphism ($p \leq 1 \times 10^{-5}$) existed for between 51 (DIT) and 3,051 (CW) SNPs depending on the trait (Appendix C.2), while evidence of significant dimorphism ($p \leq 1 \times 10^{-8}$) was evident in all but three traits (i.e., DIT, HW and WH). Of the muscular traits, the most

significantly dimorphic SNP was rs110487743 ($p = 1.36 \times 10^{-10}$), an intronic SNP located within the NAB1 gene on BTA2 which exhibited dimorphic associations for DL (Table 5.4). Of the skeletal traits, the most significantly dimorphic SNP was for CW and was an intergenic SNP, rs446294174 ($p = 4.44 \times 10^{-16}$; Table 5.5) that had an allele effect of -0.03 (SE = 0.18) in males but -3.34 (SE = 0.36) in females with a MAF of 0.015 in males and 0.008 in females. Similar to the AA, SNPs with a higher MAF tended to have a smaller allele effect size; one such SNP, rs133078486 ($p = 1.50 \times 10^{-6}$) an intronic SNP located within the SPATA9 gene on BTA7 that had dimorphic associations with WH, had a MAF of 0.213 and an allele effect of -0.08 in males and a MAF of 0.254 and an allele effect of +0.09 in females.

Of the 1kb windows containing at least one dimorphic SNP, no windows were shared between the skeletal and muscular trait groups. Despite the lack of overlap between the skeletal and muscular traits in the CH, considerable dimorphism was detected across the muscular traits. Across trait dimorphism was detected in four of the five muscular traits (i.e., DHQ, DL, TW, and WOW; Appendix C.4b) where 8 1kb windows in common between 5.54Mb and 5.60Mb on BTA2 contained a suggestively associated dimorphic SNP; only one gene, NAB1, was located within this region. An additional 22 1kb windows on BTA2 were also deemed to exhibit across trait dimorphism for the muscular traits (Appendix C.4b). Compared to the muscular traits, fewer windows containing a suggestive SNP were common among the skeletal traits. Seven 1kb windows were common to CW and CD (Appendix C.3b), one window on each of BTA4, BTA5, BTA12, and BTAX and three windows on BTA13 that contained the BTBD3 gene. A single window on BTA15 at 72.31Mb was common to both WH and BL.

Table 5.1 The number of suggestively dimorphic ($p \leq 1 \times 10^{-5}$) and significantly dimorphic ($p \leq 1 \times 10^{-8}$; in parenthesis) SNPs for each trait in each breed. Where there is no parenthesis, no significantly dimorphic SNP was detected.

	Angus	Charolais	Hereford	Limousin	Simmental
Chest depth	259	1,439 (84)	148	699 (5)	679
Chest width	259	3,051 (172)	256	1,105 (27)	136
Back length	34	176 (1)	241	229	87
Hip width	1,254 (1)	272	264	107	178
Wither height	155	62	122	95	115
Development of hind quarter	31	341 (28)	82	433 (3)	115
Development of inner thigh	329 (5)	51	39	241	91
Development of loin	274	101 (9)	97	74	233
Thigh width	128 (1)	216 (9)	62	46	160
Wither width	42	125 (9)	67	268	92

Table 5.2 The location of the most significantly dimorphic SNPs, limited to the top 5, which were associated with the muscular traits in the Angus

Trait	Chr	Start	End	No. of dimorphic SNPS	Location of most significantly dimorphic SNP	Updated position of most significantly dimorphic SNP (ARS-UCD 1.2)	Male		Female		Significance of dimorphism
							Allele Frequency	Allele Effect (SE)	Allele Frequency	Allele Effect (SE)	
Development of hind quarter	10	101970186	102988802	4	102482010[a]	101424768	0.017	-0.25 (0.14)	0.008	1.53 (0.37)	$6.64x10^{-6}$
	12	61322071	62452555	6	61890529[a]	61531778	0.118	-0.12 (0.06)	0.118	0.41 (0.10)	$3.60x10^{-6}$
	14	33610209	34610573	2	34110209[b]	32047165	0.122	0.10 (0.06)	0.140	-0.37 (0.09)	$5.32x10^{-6}$
	20	15474737	16474737	1	15974737[a]	15988960	0.034	-0.41 (0.10)	0.056	0.10 (0.14)	$1.30x10^{-6}$
	21	53862500	54867369	5	54367218[a]	53876946	0.020	0.39 (0.13)	0.015	-0.90 (0.25)	$4.16x10^{-6}$
Development of inner thigh	5	10273554	11452988	5	10773554[b]	10714008	0.004	0.50 (0.31)	0.003	-3.28 (0.63)	$7.04x10^{-8}$
	6	42909192	44746796	7	43497285[b]	42045392	0.026	0.12 (0.13)	0.003	-3.68 (0.63)	$3.45x10^{-9}$
	14	33610209	35958833	69	34110209[b]	32047165	0.122	0.12 (0.06	0.140	-0.49 (0.10)	$8.22x10^{-8}$
	20	15474737	17005863	10	15974737[a]	15988960	0.034	-0.39 (0.12)	0.056	0.58 (0.14)	$1.50x10^{-7}$
	26	17899058	19802025	5	18537518[bc]	18672598	0.005	0.88 (0.30)	0.003	-4.42 (0.89)	$1.44x10^{-8}$
Development of loin	2	101653384	102770408	4	102270408[a]	101752093	0.005	-0.45 (0.27)	0.004	2.60 (0.56)	$9.12x10^{-7}$
	5	99020985	100028321	3	99527732[a]	99094471	0.025	0.30 (0.12)	0.039	-0.70 (0.15)	$3.75x10^{-7}$
	6	117849749	118880939	4	118349749[d]	113543086	0.007	-0.59 (0.24)	0.004	2.46 (0.55)	$4.29x10^{-7}$
	11	76054540	77056745	11	76556745[a]	76492699	0.207	0.07 (0.05)	0.133	-0.46 (0.09)	$2.36x10^{-7}$
	21	53940632	55017766	6	54450845[a]	53960572	0.011	0.40 (0.18)	0.014	-1.28 (0.27)	$4.06x10^{-7}$
Thigh width	8	65764830	67442321	23	66797263[a]	66310487	0.004	-1.25 (0.31)	0.003	2.78 (0.70)	$1.53x10^{-7}$
	9	29472999	30503566	6	29977509[a]	29598390	0.039	-0.36 (0.11)	0.086	0.39 (0.12)	$5.43x10^{-6}$
	10	54322355	55487705	8	54835748[b]	54777779	0.008	-0.66 (0.24)	0.004	2.35 (0.57)	$1.21x10^{-6}$
	12	49908375	50963553	5	50408375[a]	50050440	0.010	-0.71 (0.22)	0.003	2.78 (0.70)	$1.95x10^{-6}$
	24	59024887	60257758	24	59524887[a]	59018608	0.008	0.77 (0.23)	0.011	-1.49 (0.30)	$3.90x10^{-9}$
Width of withers	2	34064457	35064503	2	34564503[b]	34459316	0.329	0.12 (0.05)	0.386	-0.32 (0.07)	$6.09x10^{-7}$
	3	2519849	3522479	2	3019849[a]	2958782	0.009	0.69 (0.23)	0.008	-1.71 (0.43)	$1.03x10^{-6}$
	9	2753165	3753801	6	3253165[a]	3188640	0.015	-0.20 (0.17)	0.011	1.44 (0.31)	$3.22x10^{-6}$
	11	76054540	77056745	11	76556745[a]	76492699	0.207	0.13 (0.06)	0.133	-0.45 (0.11)	$1.27x10^{-6}$
	23	14054921	15055421	3	14554921[a]	14546664	0.018	-0.55 (0.17)	0.008	1.58 (0.43)	$5.12x10^{-6}$

. [a]intergenic variant, [b]intron variant, [c]downstream gene variant, [d]upstream gene variant

Table 5.3 The location of the most significantly dimorphic SNPs, limited to the top 5, which were associated with the skeletal traits in the Angus.

Trait	Chr	Start	End	No. of dimorphic SNPS	Location of most significantly dimorphic SNP (UMD 3.1)	Updated position of most significantly dimorphic SNP (ARS-UCD 1.2)	Male		Female		Significance of dimorphism
							Allele Frequency	Allele Effect (SE)	Allele Frequency	Allele Effect (SE)	
Wither height	5	110840144	111861436	4	111346952[b]	110829906	0.010	-0.25 (0.17)	0.006	1.63 (0.33)	$4.3x10^{-7}$
	8	60455638	62004178	27	61312774[a]	60912522	0.084	-0.20 (0.06)	0.014	1.07 (0.24)	$2.45x10^{-7}$
	10	31872810	32902266	7	32402196[b]	32316973	0.091	-0.12 (0.06)	0.036	0.61 (0.14)	$7.23x10^{-7}$
	18	6104766	7113421	4	6613421[a]	6584541	0.276	0.09 (0.04)	0.01	-0.32 (0.07)	$8.69x10^{-7}$
	29	48236260	49251635	2	48751635[a]	48087103	0.011	-0.21 (0.16)	0.006	1.63 (0.33)	$4.25x10^{-7}$
Back length	1	70204850	71422791	16	70835180[b]	70223171	0.259	0.11 (0.04)	0.224	-0.30 (0.07)	$2.45x10^{-7}$
	2	28943486	30056216	3	29556216[a]	29476541	0.009	0.68 (0.18)	0.026	-0.46 (0.17)	$3.28x10^{-6}$
	6	26975924	27975924	1	27475924[b]	26065467	0.067	-0.11 (0.07)	0.033	0.70 (0.15)	$1.19x10^{-6}$
	8	65886591	67412672	3	66912672[a]	66423762	0.019	-0.35 (0.13)	0.006	1.46 (0.36)	$2.12x10^{-6}$
	23	23943976	24943976	1	24443976[b]	24699576	0.008	-0.30 (0.18)	0.006	1.60 (0.35)	$1.30x10^{-6}$
Hip width	1	62250123	63302205	6	62787935[a]	62183938	0.023	-0.10 (0.10)	0.009	1.58 (0.30)	$9.27x10^{-8}$
	8	57066889	67412672	77	66296263[a]	65817557	0.025	-0.23 (0.10)	0.005	2.29 (0.40)	$9.86x10^{-10}$
	10	81021198	83021473	3	81521198[c]	81172739	0.011	-0.25 (0.14)	0.003	2.84 (0.55)	$5.48x10^{-8}$
	11	30092116	31173851	16	30670702[d]	30824114	0.022	-0.20 (0.12)	0.003	2.84 (0.55)	$6.34x10^{-8}$
	13	6530211	7544606	3	7030211[b]	6887666	0.007	-0.46 (0.19)	0.003	2.87 (0.56)	$1.72x10^{-8}$
Chest width	2	12527638	14665457	29	16317185[a]	16287817	0.008	0.48 (0.18)	0.003	-2.24 (0.52)	$6.46x10^{-7}$
	6	42435811	43438281	2	42938281[b]	41487714	0.497	-0.05 (0.03)	0.424	0.27 (0.06)	$9.43x10^{-7}$
	6	46088094	47172011	47	46597538[a]	45051418	0.200	-0.10 (0.04)	0.140	0.32 (0.08)	$1.08x10^{-6}$
	15	74696711	75749423	10	75205777[b]	74292869	0.085	-0.07 (0.06)	0.034	0.75 (0.15)	$2.27x10^{-7}$
	28	23985323	24991309	3	24485323[a]	24334640	0.080	0.11 (0.06)	0.068	-0.47 (0.10)	$4.54x10^{-7}$
Chest depth	3	87713085	88819263	2	88213085[a]	87638290	0.018	-0.29 (0.12)	0.015	0.82 (0.19)	$8.77x10^{-7}$
	6	76026228	77061527	89	76542923[a]	74889187	0.287	-0.10 (0.03)	0.228	0.22 (0.06)	$7.15x10^{-7}$
	11	48278616	49764300	31	48778616[a]	48908041	0.165	0.11 (0.04)	0.196	-0.26 (0.06)	$2.11x10^{-7}$
	23	3078545	4088215	44	3581168[b]	3660670	0.044	-0.13 (0.07)	0.008	1.28 (0.28)	$9.81x10^{-7}$
	27	16617773	17634805	3	17128718[a]	18055239	0.499	0.09 (0.03)	0.433	-0.17 (0.05)	$1.21x10^{-6}$

[a]intergenic variant, [b]intron variant, [c]missense variant, [d]3' UTR variant

Table 5.4 The location of the most significantly dimorphic SNPs, limited to the top 5, which were associated with the muscular traits in the Charolais.

Trait	Chr	Start	End	No. of dimorphic SNPS	Most significantly dimorphic SNP	Updated position of most significantly dimorphic SNP (ARS-UCD 1.2)	Male		Female		Significance of dimorphism
							Allele Frequency	Allele Effect (SE)	Allele Frequency	Allele Effect (SE)	
Development of hind quarter	1	103824917	105170930	25	104363453[a]	103548640	0.215	-0.04 (0.03)	0.208	0.21 (0.04)	$3.86x10^{-7}$
	2	1255471	9445764	113	5302042[a]	5369735	0.497	-0.21 (0.02)	0.496	0.12 (0.03)	$1.55x10^{-15}$
	12	45346059	48106494	85	47588546[a]	47283052	0.107	-0.08 (0.03)	0.085	0.25 (0.06)	$6.02x10^{-7}$
	15	60710700	61831852	4	61210700[a]	60427864	0.005	0.60 (0.16)	0.003	-0.89 (0.26)	$1.29x10^{-6}$
	16	70153655	71221659	3	70653655[b]	69141596	0.004	-0.45 (0.17)	0.004	1.05 (0.23)	$1.45x10^{-7}$
Development of inner thigh	1	117119798	118119961	2	117619798[b]	116726201	0.204	0.47 (0.14)	0.219	-0.86 (0.23)	$7.34x10^{-7}$
	2	37440689	38440693	2	37940693[a]	37839039	0.033	0.69 (0.32)	0.038	-2.01 (0.46)	$1.46x10^{-6}$
	19	8575554	9576597	2	9076597[a]	8844572	0.482	0.35 (0.13)	0.492	-0.80 (0.20)	$1.46x10^{-6}$
	26	45221800	46222229	4	45721883[b]	45385940	0.486	0.18 (0.12)	0.462	-0.94 (0.19)	$8.10x10^{-7}$
	X	107228339	108242126	8	107728339[a]	102268740	0.499	0.18 (0.08)	0.498	-0.80 (0.20)	$6.48x10^{-6}$
Development of loin	1	84531371	85539275	3	85034032[a]	84419934	0.003	-0.12 (0.19)	0.003	1.75 (0.31)	$2.74x10^{-7}$
	2	0	648674	11	148674[a]	225688	0.482	-0.08 (0.03)	0.470	0.14 (0.04)	$4.37x10^{-7}$
	2	4801997	6104335	12	5587046[b]	5654486	0.487	0.15 (0.03)	0.467	-0.13 (0.04)	$1.36x10^{-10}$
	3	51686577	52703486	2	52186577[a]	52028651	0.020	-0.09 (0.08)	0.022	0.56 (0.11)	$1.33x10^{-6}$
	15	26121639	27171054	3	26671054[b]	26249705	0.034	-0.16 (0.07)	0.017	0.60 (0.13)	$3.89x10^{-7}$
Thigh width	2	4801997	6104335	13	5587046[b]	5654486	0.487	0.17 (0.03)	0.467	-0.16 (0.04)	$8.62x10^{-13}$
	5	51522594	52583478	5	52022594[b]	51784070	0.016	0.15 (0.10)	0.012	-0.78 (0.16)	$8.37x10^{-7}$
	6	100685822	102462671	3	101962671[a]	100182435	0.236	-0.12 (0.03)	0.250	0.12 (0.04)	$1.64x10^{-6}$
	21	37454121	39897176	100	37969433[a]	37571983	0.420	0.05 (0.03)	0.425	-0.17 (0.04)	$8.68x10^{-7}$
	27	39790937	40817098	7	40317098[a]	40462706	0.295	-0.06 (0.03)	0.320	0.17 (0.04)	$9.75x10^{-7}$
Width of withers	1	25066241	26077136	3	25577136[a]	26076126	0.095	0.09 (0.04)	0.126	-0.24 (0.05)	$4.07x10^{-7}$
	2	5064592	6104335	9	5587046[b]	5654486	0.487	0.15 (0.03)	0.467	-0.12 (0.04)	$1.66x10^{-9}$
	4	26737341	27961476	2	27461476[b]	27487938	0.002	-1.17 (0.25)	0.003	0.73 (0.29)	$5.38x10^{-7}$
	9	60485268	61503827	4	60993475[a]	60117188	0.005	-0.73 (0.16)	0.003	1.04 (0.32)	$9.89x10^{-7}$
	21	39307646	40307667	4	39807694[a]	39393236	0.006	0.37 (0.15)	0.008	-0.80 (0.18)	$7.25x10^{-7}$

[a]intergenic variant, [b]intron variant, [c]downstream gene variant

Table 5.5 The location of the most significantly dimorphic SNPs, limited to the top 5, which were associated with the skeletal traits in the Charolais.

Trait	Chr	Start	End	No. of dimorphic SNPS	Most significantly dimorphic SNP	Updated position of most significantly dimorphic SNP (ARS-UCD 1.2)	Male		Female		Significance of dimorphism
							Allele Frequency	Allele Effect (SE)	Allele Frequency	Allele Effect (SE)	
Wither height	7	96716729	97718351	3	97483338[b]	94984764	0.213	-0.08 (0.02)	0.254	0.09 (0.02)	$1.50x10^{-6}$
	10	75208550	77306612	5	75708579[a]	75395884	0.029	-0.13 (0.05)	0.028	0.28 (0.07)	$5.89x10^{-6}$
	17	34267907	35272234	15	34770503[a]	34374526	0.120	-0.09 (0.03)	0.113	0.14 (0.04)	$1.64x10^{-6}$
	19	33223631	34225743	6	33723631[b]	33122406	0.451	0.03 (0.02)	0.499	-0.11 (0.03)	$5.95x10^{-6}$
	29	29284867	30316330	11	29807810[a]	29436465	0.246	0.03 (0.02)	0.170	-0.16 (0.03)	$1.46x10^{-6}$
Back length	1	60942796	62141649	4	61529004[a]	60977668	0.020	0.08 (0.06)	0.008	-0.76 (0.14)	$5.31x10^{-8}$
	8	65823397	67021631	17	66520586[bc]	66036297	0.014	0.24 (0.07)	0.014	-0.39 (0.11)	$7.09x10^{-7}$
	11	67394139	68443814	13	67936368[b]	67962565	0.039	0.11 (0.04)	0.033	-0.29 (0.07)	$1.18x10^{-6}$
	12	66787629	68105969	3	67354072[a]	66817243	0.005	0.19 (0.11)	0.003	-1.24 (0.22)	$6.64x10^{-9}$
	26	16741492	17746984	6	17242026[b]	17376909	0.198	-0.06 (0.02)	0.196	0.14 (0.03)	$2.43x10^{-7}$
Hip width	1	71113538	72230151	34	71680557[a]	71071338	0.007	-0.18 (0.09)	0.003	1.09 (0.24)	$7.09x10^{-7}$
	1	150460935	151629392	5	150960935[c]	*	0.031	0.11 (0.05)	0.046	-0.29 (0.06)	$1.96x10^{-7}$
	7	78851675	79854739	3	79352226[a]	77071480	0.027	-0.09 (0.05)	0.013	0.49 (0.11)	$9.91x10^{-7}$
	19	41254273	42757991	23	42257563[c]	41624383	0.127	-0.09 (0.02)	0.116	0.14 (0.04)	$1.32x10^{-6}$
	23	40292666	41337462	14	40792666[d]	41014434	0.044	0.13 (0.04)	0.052	-0.21 (0.06)	$1.49x10^{-6}$
Chest width	2	39826441	43069032	11	42375687[d]	*	0.003	0.46 (0.38)	0.002	-5.30 (0.71)	$1.01x10^{-12}$
	8	24830532	26713234	9	29877151[a]	29907983	0.015	-0.03 (0.18)	0.008	-3.34 (0.36)	$4.44x10^{-16}$
	9	21416858	24598647	19	21916858[a]	21654494	0.007	0.21 (0.26)	0.003	-4.10 (0.52)	$1.66x10^{-13}$
	14	5222811	6612129	22	6528804[a]	5500130	0.010	0.50 (0.23)	0.006	-3.14 (0.43)	$4.49x10^{-14}$
	28	0	865080	17	365080[a]	1347377	0.006	0.62 (0.28)	0.003	-4.79 (0.63)	$6.66x10^{-15}$
Chest depth	2	78501288	79652702	7	79001288[a]	78633046	0.005	0.29 (0.27)	0.002	-3.88 (0.56)	$2.50x10^{-11}$
	6	22501427	30700058	17	29804490[a]	28380999	0.006	0.17 (0.28)	0.003	-4.13 (0.57)	$1.55x10^{-11}$
	17	20292176	22351653	28	20871130[a]	20556557	0.016	0.01 (0.18)	0.008	-2.36 (0.30)	$1.22x10^{-11}$
	19	5471820	6722162	11	6222162[a]	6015594	0.013	0.15 (0.19)	0.008	-2.41 (0.33)	$1.32x10^{-11}$
	20	68770623	70729278	50	70127722[a]	*	0.002	0.23 (0.46)	0.002	-5.16 (0.57)	$2.42x10^{-13}$

[a]intergenic variant, [b]intron variant, [c]downstream gene variant, [d]upstream gene variant

5.5.3 Hereford

A total of 17,241,152 SNPs were segregating in the 727 HE males and 16,494,904 SNPs were segregating in the 402 HE females. Of these, 15,991,751 SNPs were segregating in both the male and female animals (Appendix C.2). In comparison to the AA and CH, evidence of suggestive dimorphism ($p \leq 1 \times 10^{-5}$) was evident for fewer SNPs in all of the traits, ranging from 39 (DIT) to 256 (CW) SNPs depending on the trait; there was no evidence of significant dimorphism ($p \leq 1 \times 10^{-8}$; Table 5.1). Similar to the AA, the allele substitution effect of the dimorphic SNPs in the males tended to be in the opposite direction to the allele effect of the same SNP in the females (Table 5.6 and 5.8). The most significantly dimorphic SNP in the muscular traits was rs798960299 ($p = 6.30 \times 10^{-8}$), an intronic variant located within the NR5A2 gene on BTA16 with an allele effect of +0.42 (SE = 0.13) in males but -0.61 (SE = 0.14) in females (Table 5.6). Of the skeletal traits, the most significantly dimorphic SNP was rs381085044, an intergenic SNP on BTA1 that had a dimorphic association with CW represented by an allele effect of -0.22 (SE = 0.06) in males but +0.32 (SE = 0.08) in females (Table 5.7).

No 1kb window that contained at least one dimorphic SNP was common to more than two traits. Limited dimorphism was found between CW and HW, where three windows between 40.75Mb and 40.78Mb on BTA24 containing the PTPRM gene were suggestively associated with both traits (Appendix C.3c). Chest width also had two separate windows exhibiting dimorphism on BTA13 between 27.46Mb and 27.47Mb in common with CD, and one window on BTA14 at 45.485Mb in common with WH. Two adjacent 1kb windows on BTA8 at 25.965Mb were common to both WH and BL. For the muscular traits, WOW had one window in common with each of DL (BTA10 at 48.327Mb), TW (BTA10 at 50.420Mb) and DHQ (BTA9 at 83.332Mb). One 1kb

window was also common to both TW and DHQ (BTA16 at 68.580Mb; Appendix C.4c).

5.5.4 Limousin

A total of 18,056,913 SNPs were segregating in the LM males and 17,767,237 SNPs were segregating in the LM females. Of these, 17,482,131 SNPs were segregating in both the male and female animals. Between 46 (TW) and 1,105 (CW) SNPs were suggestively dimorphic ($p \leq 1 \times 10^{-5}$) while 3 traits (i.e., CW, CD, and DHQ) had evidence of significant dimorphism ($p \leq 1 \times 10^{-8}$; Table 5.1). Of the muscular traits, the most significantly dimorphic SNP was rs42425148 ($p = 1.85 \times 10^{-9}$), an intergenic SNP located on BTA1 that had a dimorphic association with DHQ (Table 5.8) represented by an allele effect of -0.01 (SE = 0.15) in males but -1.64 (SE = 0.23) in females. The most significantly dimorphic SNP for the skeletal traits was also an intergenic SNP, rs478688690 ($p = 3.63 \times 10^{-11}$), located on BTA11, that had a dimorphic association with CW (Table 5.9) with an allele effect of +0.03 (SE = 0.29) in males and -3.86 (SE = 0.51) in females; this SNP had a MAF of 0.005 in males and 0.003 in females. Similar to the AA and CH, SNPs with a higher MAF tended to have an allele effect that was closer to zero. An intergenic SNPs with dimorphic associations in HW ($p = 5.71 \times 10^{-7}$) had a MAF of 0.493 in males and 0.486 in females but the allele substitution effects were just -0.07 (SE = 0.01) in males and +0.06 (SE = 0.02) in females.

Genomic regions that exhibited sexual dimorphism across traits in the LM were limited; 8 1kb windows were found to be suggestively associated with three of the skeletal traits (WH, BL, HW; Appendix C.3d) whereas only 3 windows were common between TW and WOW of the muscular traits (Appendix C.4d). All 8 windows associated with WH, BL and HW were on BTA6 but no obvious candidate gene was located in the vicinity. The 40 windows suggestively associated with both CW and CD

(Appendix C.3d) were located on 15 different autosomes, BTA5, BTA6, BTA8, BTA9, BTA10, BTA11, BTA13, BTA16, BTA17, BTA18, BTA20, BTA22, BTA24, BTA26, BTA29, BTAX.

5.5.5 Simmental

A total of 18,257,175 SNPs were segregating in the SI males and 17,814,297 SNPs were segregating in the SI females, while 17,319,250 of these SNPs were segregating in both the males and females. Between 87 (BL) to 679 (CD) SNPs were suggestively dimorphic ($p \leq 1 \times 10^{-5}$) while no SNP was significantly dimorphic ($p \leq 1 \times 10^{-8}$; Table 5.1). Once again, the most significantly dimorphic SNPs tended to have a low MAF and a large allele effect size. The most significantly dimorphic SNP associated with any of the muscular traits was rs110995439 ($p = 1.40 \times 10^{-8}$), an intron variant located within the GPC5 gene on BTA12 that had a dimorphic association with DIT; the allele substitution effect in the males was -0.85 (SE = 0.34) while the allele substitution effect in the females was +1.76 (SE = 0.30; Table 5.10). Of the skeletal traits, the most significantly dimorphic SNP was an intergenic SNP, rs437227524 ($p = 1.98 \times 10^{-8}$), that had a dimorphic association with CD (Table 5.11) and had an allele substitution effect of -0.08 (SE = 0.24) in the male population but -2.64 (SE = 0.42) in the female population with a MAF of 0.004 in the males and 0.002 in the females. An intronic SNP, rs133629874 ($p = 2.20 \times 10^{-7}$), located within the MMRN1 that had a dimorphic association with CW had a MAF of 0.439 in males and 0.472 in females with an effect size of +0.14 (SE = 0.03) in males and -0.13 (SE = 0.04) in females (Table 5.11).

Few 1kb windows containing a suggestive SNP overlapped among the muscular and skeletal traits. A single 1kb window on BTA2, approximately 0.1Mb from the IWS1 gene, contained suggestively dimorphic SNPs for all of DIT, DL, and CW. Of the skeletal traits, no genomic windows exhibited suggestive associated dimorphism

(Appendix C.3e) and no window was suggestively associated with three or more muscular traits (Appendix C.4e). Three 1kb windows, all located on BTA18 between 3.78Mb and 3.80Mb, contained dimorphic SNPs for both TW and WOW (Appendix C.4e); these windows were located approximately 0.3Mb from the CNTNAP4 gene. One window located on BTA28 at 41.045Mb and 6 windows on the X chromosome between 114.572Mb and 114.578Mb were dimorphic for both TW and DL. A single 1kb window was common to each of DIT and TW (13.541Mb on BTA8), DHQ and WOW (102.906Mb on BTA6), and WH and BL (145.410Mb on BTAX).

Table 5.6 The location of the most significantly dimorphic SNPs, limited to the top 5, which were associated with the muscular traits in the Hereford.

Trait	Chr	Start	End	No. of dimorphic SNPS	Most significantly dimorphic SNP	Updated position of most significantly dimorphic SNP (ARS-UCD 1.2)	Male		Female		Significance of dimorphism
							Allele Frequency	Allele Effect (SE)	Allele Frequency	Allele Effect (SE)	
Development of hind quarter	1	47475388	48477830	9	47975875[a]	47609095	0.440	-0.14 (0.05)	0.460	0.20 (0.06)	$5.34x10^{-6}$
	7	38157809	39157899	3	38657899[a]	37308156	0.307	-0.20 (0.05)	0.291	0.17 (0.06)	$1.58x10^{-6}$
	12	33106286	34154759	5	33654759b	33414723	0.008	-0.84 (0.24)	0.002	1.90 (0.56)	$6.14x10^{-6}$
	16	79654320	81344038	20	80834643[b]	78911900	0.032	0.42 (0.13)	0.041	-0.61 (0.14)	$6.30x10^{-8}$
	29	48954403	49974227	13	49471817[a]	*	0.080	0.29 (0.09)	0.092	-0.31 (0.09)	$4.83x10^{-6}$
Development of inner thigh	1	139386569	140386569	1	139886569[b]	138390445	0.131	-0.24 (0.07)	0.098	0.31 (0.09)	$3.36x10^{-6}$
	6	104110238	105116799	3	104610238[b]	102834429	0.030	0.46 (0.14)	0.021	-0.59 (0.18)	$4.31x10^{-6}$
	13	42113572	43125008	4	42613572[a]	42236685	0.003	-1.57 (0.41)	0.005	1.22 (0.4)	$1.09x10^{-6}$
	19	48026847	49047485	15	48526847[b]	47877987	0.078	-0.23 (0.09)	0.095	0.33 (0.09)	$6.22x10^{-6}$
	20	45699238	46731354	6	46213822[a]	46189379	0.169	-0.18 (0.06)	0.193	0.32 (0.07)	$1.21x10^{-7}$
Development of loin	3	89108056	90628279	9	90128279[d]	89550331	0.072	0.14 (0.09)	0.040	-0.82 (0.16)	$1.76x10^{-7}$
	16	16872687	17885604	2	17372687[a]	16732806	0.211	0.19 (0.06)	0.183	-0.29 (0.08)	$9.83x10^{-7}$
	21	57418522	58508022	5	57972000[b]	57385760	0.018	0.44 (0.18)	0.019	-1.06 (0.23)	$2.06x10^{-7}$
	23	44376366	45381654	9	44876460[b]	45012825	0.433	-0.18 (0.05)	0.445	0.19 (0.06)	$2.30x10^{-6}$
	26	24760894	25969246	11	25358414[d]	25093228	0.052	0.06 (0.10)	0.016	-1.32 (0.25)	$2.54x10^{-7}$
Thigh width	4	30981541	31996449	3	31481541[a]	31358361	0.052	-0.24 (0.11)	0.050	0.61 (0.15)	$2.84x10^{-6}$
	11	34683559	35683570	2	35183559[a]	35342330	0.210	-0.17 (0.06)	0.249	0.28 (0.07)	$1.98x10^{-6}$
	15	48172990	49186437	10	48675268[c]	48028112	0.010	-0.59 (0.25)	0.016	1.07 (0.25)	$3.00x10^{-6}$
	27	37804950	38804962	3	38304950[a]	*	0.004	-1.24 (0.38)	0.010	1.25 (0.32)	$3.96x10^{-7}$
	29	32341813	33341891	5	32841813[a]	32298517	0.010	-0.24 (0.26)	0.005	2.33 (0.45)	$6.10x10^{-7}$
Width of withers	1	23635565	24635601	2	24135565[a]	24617417	0.044	0.19 (0.12)	0.027	-0.87 (0.20)	$6.32x10^{-6}$
	11	88551389	89554029	5	89052894[a]	89077280	0.263	-0.11 (0.06)	0.195	0.34 (0.08)	$4.63x10^{-6}$
	16	16859305	17861510	3	17361510[a]	16721630	0.297	0.21 (0.06)	0.320	-0.22 (0.07)	$9.32x10^{-7}$
	23	28067269	29080639	4	28580639[c]	28787480	0.023	-0.66 (0.18)	0.035	0.58 (0.17)	$7.84x10^{-7}$
	29	3021804	4053998	21	3553943[a]	3466707	0.483	0.14 (0.05)	0.469	-0.28 (0.06)	$6.82x10^{-7}$

[a]intergenic variant, [b]intron variant, [c]downstream gene variant, [d]upstream gene variant

Table 5.7 The location of the most significantly dimorphic SNPs, limited to the top 5, which were associated with the skeletal traits in the Hereford.

Trait	Chr	Start	End	No. of dimorphic SNPS	Most significantly dimorphic SNP	Updated position of most significantly dimorphic SNP (ARS-UCD 1.2)	Male		Female		Significance of dimorphism
							Allele Frequency	Allele Effect (SE)	Allele Frequency	Allele Effect (SE)	
Wither height	1	1925823	2922535	2	2422535[b]	3142926	0.010	0.63 (0.20)	0.009	-0.99 (0.27)	$1.22x10^{-6}$
	10	49372889	50457497	10	49872889[b]	49812963	0.004	-1.27 (0.30)	0.006	0.83 (0.31)	$1.18x10^{-6}$
	10	68414747	69451655	11	68933579[b]	68686530	0.107	0.27 (0.06)	0.137	-0.20 (0.07)	$5.97x10^{-7}$
	16	916164	1945553	12	1441711[c]	1624685	0.025	-0.36 (0.13)	0.016	0.78 (0.20)	$1.05x10^{-6}$
	24	10954731	11969624	6	11461076[a]	11159461	0.096	-0.27 (0.06)	0.067	0.29 (0.09)	$2.65x10^{-7}$
Back length	4	12086873	13087742	3	12587742[a]	12740501	0.008	-0.90 (0.24)	0.005	1.22 (0.39)	$3.37x10^{-6}$
	8	25464831	26824440	23	25966773[a]	25943096	0.110	0.32 (0.07)	0.132	-0.23 (0.08)	$3.56x10^{-7}$
	14	30747311	31768991	12	31247311[a]	29553248	0.461	0.14 (0.04)	0.471	-0.20 (0.06)	$1.38x10^{-6}$
	25	13061095	14195964	4	13561095[b]	*	0.003	-1.27 (0.36)	0.012	0.75 (0.24)	$3.63x10^{-6}$
	28	32468295	34072536	3	33572536[b]	33371736	0.146	-0.10 (0.06)	0.148	0.37 (0.09)	$3.18x10^{-6}$
Hip width	5	75173455	76244035	40	75719616[a]	75344242	0.006	-0.73 (0.28)	0.004	1.85 (0.44)	$8.46x10^{-7}$
	11	88093610	89599089	7	89005217[a]	89030779	0.263	-0.13 (0.06)	0.195	0.40 (0.09)	$8.95x10^{-7}$
	12	82288642	83293488	3	82793488[a]	78806868	0.155	-0.17 (0.06)	0.208	0.29 (0.07)	$3.07x10^{-7}$
	23	51480640	52527183	5	52013801[b]	52167774	0.089	-0.16 (0.08)	0.091	0.46 (0.10)	$3.27x10^{-7}$
	27	8148686	10276462	17	8648686[a]	9657276	0.107	-0.24 (0.06)	0.106	0.30 (0.09)	$5.65x10^{-7}$
Chest width	1	62368829	63399292	5	62868829[a]	62264629	0.082	-0.22 (0.06)	0.108	0.32 (0.08)	$5.58x10^{-8}$
	5	49311974	50341913	23	49822014[b]	49592445	0.428	-0.17 (0.04)	0.384	0.13 (0.05)	$5.20x10^{-7}$
	6	41424236	42580067	9	42029809[b]	40571631	0.003	1.31 (0.31)	0.005	-1.13 (0.33)	$7.10x10^{-8}$
	9	31485678	32498136	4	31985678[a]	31570153	0.010	0.42 (0.19)	0.004	-1.78 (0.38)	$2.34x10^{-7}$
	13	2019339	3201947	16	2522257[d]	2613918	0.066	0.10 (0.07)	0.030	-0.67 (0.13)	$1.94x10^{-7}$
Chest depth	10	2981982	4121382	3	3621377[a]	3672531	0.008	0.11 (0.19)	0.002	2.49 (0.48)	$3.42x10^{-6}$
	11	67322056	68784534	4	68278582[d]	68305651	0.003	-0.87 (0.32)	0.002	2.05 (0.48)	$4.15x10^{-7}$
	16	9638392	10667968	6	10150979[a]	9550179	0.044	-0.24 (0.08)	0.040	0.46 (0.13)	$3.62x10^{-6}$
	19	42220630	43244434	32	42738845[d]	42096896	0.349	0.11 (0.03)	0.322	-0.20 (0.06)	$2.67x10^{-6}$
	22	25263617	26332810	15	25767332[a]	25654951	0.054	-0.15 (0.07)	0.050	0.49 (0.11)	$7.27x10^{-7}$

[a]intergenic variant, [b]intron variant, [c]downstream gene variant, [d]upstream gene variant

5.5.6 Across-breed

The numerically smaller breeds of AA, HE, and SI, had approximately 1.5 times more SNPs segregating in only one sex than the numerically larger breeds of CH and LM. Of the total autosomal SNPs that were segregating in both sexes, between 86% (AA) and 94% (LM) had the same minor alleles in both sexes. However, differences ($p \leq 0.05$) in allele frequencies between the sexes were observed for between 36% (LM) and 66% (AA) of the total autosomal SNPs that were segregating in both sexes.

The vast majority of SNPs and 1kb windows associated with any one trait were breed-specific. The most windows displaying dimorphic characteristics in common for more than one breed was for CW in both the CH and LM with 9 1kb windows common to these breeds occurring on BTA2 at 89.527Mb (n=1), on BTA12 between 87.128Mb and 87.303Mb and containing the ENSBTAG00000032038 gene (n=6), and on BTA24 between 31.661Mb and 31.671Mb (n=2). Also for CW, a single 1kb window on BTA6 approximately 0.1Mb from the SLC34A2 gene displayed dimorphic associations in both the AA and the CH. A single 1kb window containing the ENSBTAG00000046311 gene on BTA10 had dimorphic associations with HW in both the AA and SI. In the CH and the SI, two common windows, one at 91.126Mb on BTA7 and one at 70.827Mb on BTA9, exhibited dimorphism with CD in both breeds.

5.6 Discussion

While several studies in cattle have investigated the presence of sexual dimorphism using a quantitative genetics approach (Chapter 2; Crews Jr and Kemp, 2001; van der Heide et al., 2016; Bittante et al., 2018) no previous study in cattle has attempted to detect evidence of sexual dimorphism at the genome level or to compare these effects

across multiple breeds of cattle. From previous quantitative genetics studies, differences in genetic parameters by cattle sex have been observed for growth rate (Koch and Clark, 1955; Marlowe and Gaines, 1958), post-weaning gain (van der Heide et al., 2016), feed intake and efficiency, fleshiness scores, carcass weight and yield (Bittante et al., 2018), as well as longissimus muscle area and backfat (Crews Jr and Kemp, 2001). In contrast, negligible differences in genetic parameters between the sexes were detected for early growth traits such as weaning weight (Koch and Clark, 1955; van der Heide et al., 2016).

A single trait measured in different environments (or different sexes) can be regarded as separate traits which are genetically correlated (Falconer, 1952). In many situations, the genetic correlations of the same trait taken in different environments are less than unity, indicating that selection occurring in one environment may not be optimal for performance in the other environment (Mulder and Bijma, 2005). This represents a genotype-by-environment interaction although Robertson (1959) postulated that the genetic correlation between environments would need to be weaker than 0.80 to be considered of importance for breeding purposes. In quantitative genetics studies on dimorphism, it is not necessarily the environment that is causing the genetic correlations to differ from unity, but possibly the effect of dimorphism (which is often cofounded with environment). Weaker than unity genetic correlations between the sexes may be indicative of many factors including the alleles having a different substitution effect in each sex; these differences may be due to inter-sex differences in effect sizes or the sign of the allele substitution effect differing by sex.

Bittante et al. (2018), while investigating the effects of sexual dimorphism on the fattening performance and muscling of young Belgian Blue and Piedmontese dairy cross bulls and heifers, noticed that the effects of dimorphism were greater in the

Belgian Blues than the Piedmontese suggesting that the effects of sexual dimorphism may actually differ by breed. Because linear type traits which describe the skeletal and muscular conformation of an animal are related to many performance traits such as animal live weight (Mc Hugh et al., 2012), carcass merit (Mukai et al., 1995; Conroy et al., 2010), primal cut yields (Berry et al., 2019), and feed intake (Veerkamp and Brotherstone, 1997; Crowley et al., 2011), it is plausible that the underlying variome of animals contributing to differences in their skeletal or muscular characteristics may also exhibit sexual dimorphism; it is also plausible that these regions exhibiting dimorphism may differ by breed. The data set used in the present study was particularly useful to test this hypothesis in that all linear type traits were assessed in all breeds and sexes using the same scale by the same classifiers and, therefore, a direct comparison of sex effects as well as commonalities of detected regions across breeds was possible.

Using a dataset of 32,725 males and 30,887 females of the CH and LM breeds, Chapter 2 estimated variance components for 18 type traits in both sexes separately; the type traits included in the present study were those represented in that study and included functional, skeletal and muscular subjective measures. Numerical differences in variance estimates were detected between both sexes in each breed while inter-sex differences in heritability estimates were only significant ($p < 0.05$) for BL, WH, and DHQ in the CH, with no differences observed in the LM (Chapter 2). Within trait genetic correlations between each of the 18 type traits in each sex were all stronger than 0.90 (Chapter 2); because Robertson (Robertson, 1959) concluded that genetic correlations had to be weaker than 0.80 to be impactful, Chapter 2 concluded a lack of dimorphism in that study. Nonetheless, genetic correlations are derived from the entire genome and therefore may not capture the granularity achieved by investigation of specific regions of the genome, as undertaken in the present study. This is especially

true when only a few regions exhibit dimorphism and the extent of dimorphism in these regions may be small. Indeed the results from the present study indicate that, in fact, only a few regions exhibit dimorphism and the effects are small.

5.6.1 The X Chromosome

The X chromosome is the second largest chromosome in the bovine genome and accounts for over 6% of the total physical genome (148,823,899 bp; Zimin et al., 2009); it is, however, regularly discarded from genome-based studies in cattle due the inheritance of the X chromosome being different to the autosomes. Males are heterogametic (XY) and the females are homogametic (XX; Fernando and Grossman, 1990); therefore, male offspring inherit their X chromosome from their dam only, while female offspring inherit one copy of the X chromosome from their dam and the other from their sire. Furthermore, a small region of the X chromosome, known as the pseudo-autosomal region (PAR), is homologous to the Y chromosome and is inherited like an autosome with some recombination occurring during meiosis (Van Leare, et al., 2008; Su et al., 2014).

Ignoring the X chromosome could lead to important biological functions being missed and could also impact the accuracy of genomic evaluations (Lyons et al., 2014; Su et al., 2014; Mao et al., 2016). A previous study on the role of the sex chromosomes in dimorphism (Rice, 1984) revealed that while sex chromosomes were not required for the evolution of sexual dimorphism, they facilitated the evolution of sexual dimorphism for a wider range of traits than would have occurred without them and that X-linked genes in particular had a large role in the evolution of sexually dimorphic traits. In the present study, only 406 SNPs located on the X chromosome expressed sexual

dimorphism in at least one breed or trait. The low number of dimorphic SNPs on the X chromosome may be a function of possible allelic content variation where females have two copies of an allele and males only have one, thus interfering with the detection of dimorphic SNPs on the X chromosome. However, previous studies have discovered that in most cases, sex chromosomes are only required to initiate sexual dimorphism and the corresponding genes are mostly located on the autosomes (Saifl and Chandra, 1999; Fairbairn and Roff, 2006).

Previous studies have linked mutations on the X chromosome to andrological and growth traits in beef cattle (Lyons et al., 2014) as well as the length of productive life in dairy cattle (Saowaphak et al., 2017). In the present study, all sexually dimorphic SNPs located on the X chromosome for any trait in any of the breeds had a greater allele effect size in females than in males; this is in agreement with a previous study on sexual dimorphic gene expression in cattle using RNA-seq which stated that all X-chromosomal sexually dimorphic genes had a greater effect in females than males (Seo et al., 2016). In other species, such as Drosophila melanogaster, it has been proven that an X-linked recessive mutation that benefits males will accumulate faster as expression in males is hemizygous and there will be no masking by dominance (Gibson et al., 2002); however, the male-biased expression of these alleles reduces as the allele becomes more frequent in the population which enables counter-selection in the females to halt the spread of this male-biased allele.

Table 5.8 The location of the most significantly dimorphic SNPs, limited to the top 5, which were associated with the muscular traits in the Limousin.

Trait	Chr	Start	End	No. of dimorphic SNPS	Most significantly dimorphic SNP	Updated position of most significantly dimorphic SNP (ARS-UCD 1.2)	Male		Female		Significance of dimorphism
							Allele Frequency	Allele Effect (SE)	Allele Frequency	Allele Effect (SE)	
Development of hind quarter	10	87342478	88342595	3	87842478[a]	86838930	0.004	-0.01 (0.15)	0.003	-1.64 (0.23)	$1.85x10^{-9}$
	12	88313474	89356176	25	88814730[a]	84800697	0.011	0.16 (0.08)	0.005	-0.84 (0.17)	$1.96x10^{-7}$
	14	80691409	81691453	2	81191409[a]	78818851	0.002	0.10 (0.19)	0.002	-1.46 (0.24)	$4.02x10^{-7}$
	16	7877931	9577924	42	8703762[a]	8116291	0.003	0.12 (0.16)	0.002	-1.50 (0.24)	$2.53x10^{-8}$
	18	53678352	54842543	90	54323356[a]	53885291	0.003	0.36 (0.17)	0.002	-1.28 (0.25)	$1.03x10^{-7}$
Development of inner thigh	1	134576003	135674198	26	135111581[a]	*	0.024	-1.21 (0.31)	0.035	1.27 (0.38)	$4.64x10^{-7}$
	7	94624535	94624535	9	95351376[a]	92825575	0.021	1.27 (0.31)	0.082	-0.92 (0.26)	$9.29x10^{-8}$
	8	83362119	84373117	16	83862344[a]	82440029	0.026	-0.71 (0.29)	0.029	1.77 (0.43)	$2.06x10^{-6}$
	14	12997796	14065940	91	13497889[a]	12381763	0.351	-0.39 (0.11)	0.373	0.60 (0.16)	$3.15x10^{-7}$
	20	15771955	16909647	12	16366926[a]	16380235	0.003	3.16 (0.79)	0.005	-2.93 (0.93)	$6.26x10^{-7}$
Development of loin	2	90388815	91414365	6	90901181[a]	90485467	0.170	-0.09 (0.03)	0.168	0.12 (0.03)	$1.37x10^{-6}$
	6	95583095	96694369	5	96194369[a]	94427267	0.004	-0.43 (0.15)	0.002	1.00 (0.27)	$4.45x10^{-6}$
	15	81201476	82201561	2	81701476[a]	80411671	0.149	0.07 (0.03)	0.171	-0.14 (0.03)	$1.04x10^{-6}$
	25	3582865	4590247	3	4090247[bc]	4072335	0.492	0.09 (0.02)	0.460	-0.07 (0.03)	$5.72x10^{-7}$
	X	33834595	34867127	35	34354746[a]	34130068	0.376	-0.03 (0.01)	0.404	0.11 (0.03)	$5.35x10^{-6}$
Thigh width	2	19380958	20380975	3	19880958[a]	19838410	0.014	0.27 (0.09)	0.011	-0.38 (0.11)	$6.57x10^{-6}$
	2	20700909	21720973	6	21217950[a]	21181655	0.100	0.06 (0.03)	0.090	-0.19 (0.05)	$7.05x10^{-6}$
	14	6060940	7069392	3	6569392[a]	5540717	0.004	-0.36 (0.16)	0.003	0.87 (0.21)	$3.45x10^{-6}$
	23	40690020	41776379	6	41276379[b]	41595913	0.009	-0.15 (0.10)	0.003	0.94 (0.20)	$1.11x10^{-6}$
	X	16962347	17966732	2	17466732[b]	17527360	0.415	-0.02 (0.01)	0.445	0.12 (0.03)	$3.38x10^{-6}$
Width of withers	1	51070933	52117339	12	51577476[a]	51177783	0.289	0.07 (0.02)	0.294	-0.10 (0.03)	$1.96x10^{-6}$
	2	85735779	86812555	7	86265284[a]	85863432	0.019	-0.18 (0.08)	0.032	0.31 (0.07)	$2.80x10^{-6}$
	10	28209680	29266690	175	28742392[b]	28681684	0.378	-0.07 (0.02)	0.353	0.10 (0.03)	$2.53x10^{-6}$
	12	66608973	67613368	3	67108973[b]	66572583	0.011	-0.20 (0.10)	0.009	0.60 (0.13)	$9.62x10^{-7}$
	19	7961956	8967691	5	8466831[b]	8237090	0.003	-0.95 (0.19)	0.002	0.61 (0.27)	$1.69x10^{-6}$

[a]intergenic variant, [b]intron variant, [c]upstream gene variant

Table 5.9 The location of the most significantly dimorphic SNPs, limited to the top 5, which were associated with the skeletal traits in the Limousin.

Trait	Chr	Start	End	No. of dimorphic SNPS	Most significantly dimorphic SNP	Updated position of most significantly dimorphic SNP (ARS-UCD 1.2)	Male		Female		Significance of dimorphism
							MAF	Allele Effect (SE)	MAF	Allele Effect (SE)	
Wither height	3	9818384	10820240	2	10318384[a]	*	0.022	-0.20 (0.05)	0.013	0.26 (0.08)	$1.08x10^{-6}$
	6	36963566	40499953	40	39760256[a]	38319979	0.484	-0.07 (0.02)	0.487	0.09 (0.02)	$1.07x10^{-8}$
	11	48105352	49267580	9	48766334[a]	48896156	0.006	0.26 (0.09)	0.006	-0.49 (0.13)	$3.40x10^{-6}$
	19	30754934	31762424	2	31254934[a]	30619941	0.257	-0.06 (0.02)	0.248	0.09 (0.02)	$1.94x10^{-7}$
	X	16956607	17966732	6	17462347[b]	17522975	0.428	-0.01 (0.01)	0.454	0.11 (0.02)	$1.19x10^{-6}$
Back length	2	19533709	20573649	21	20065919[a]	20029900	0.096	-0.06 (0.02)	0.089	0.16 (0.04)	$1.02x10^{-6}$
	15	7167859	8174094	3	7674094[b]	7433522	0.143	-0.05 (0.02)	0.118	0.13 (0.03)	$3.14x10^{-6}$
	16	64732311	65827907	8	65324669[a]	63842035	0.003	0.39 (0.14)	0.003	-0.74 (0.19)	$1.47x10^{-6}$
	21	10874079	11978348	2	11374079[a]	11140251	0.009	-0.12 (0.07)	0.006	0.60 (0.14)	$4.03x10^{-6}$
	23	19728831	20739036	5	20228831[b]	20236461	0.005	-0.34 (0.10)	0.004	0.55 (0.16)	$3.04x10^{-6}$
Hip width	4	1742345	2747816	5	2245736[a]	2343834	0.018	0.15 (0.05)	0.021	-0.27 (0.07)	$1.83x10^{-6}$
	6	39032482	40040409	8	39539558[a]	38099446	0.493	-0.07 (0.01)	0.486	0.06 (0.02)	$5.71x10^{-7}$
	12	5047438	6074330	3	5547438[a]	5568301	0.020	-0.14 (0.05)	0.009	0.39 (0.10)	$1.23x10^{-6}$
	21	23430143	24464331	28	23931207[a]	23468504	0.237	-0.01 (0.02)	0.202	0.11 (0.03)	$1.74x10^{-7}$
	29	13364114	14373070	4	13872420[a]	13797973	0.011	0.21 (0.07)	0.011	-0.34 (0.09)	$1.11x10^{-6}$
Chest width	11	15939177	16959054	3	16459054[a]	16439125	0.005	0.03 (0.29)	0.003	-3.86 (0.51)	$3.63x10^{-11}$
	20	60692205	68001633	7	62047530[a]	61941499	0.008	0.03 (0.22)	0.002	-3.67 (0.58)	$2.64x10^{-9}$
	22	55794366	59621588	10	56329073[a]	55689684	0.031	0.31 (0.12)	0.017	-1.26 (0.22)	$2.01x10^{-10}$
	26	38382849	39632022	5	38919539[a]	*	0.006	0.40 (0.26)	0.002	-3.51 (0.58)	$8.26x10^{-10}$
	X	15184800	16185025	2	15684800[a]	15786355	0.004	0.36 (0.21)	0.003	-3.37 (0.53)	$6.15x10^{-11}$
Chest depth	3	26560340	27936092	8	27203740[a]	*	0.003	0.59 (0.39)	0.002	-3.35 (0.58)	$1.52x10^{-8}$
	6	65184840	66194780	2	65684840[a]	64045311	0.003	0.54 (0.39)	0.002	-4.10 (0.65)	$8.70x10^{-10}$
	10	66552011	70565607	71	70065607[c]	69820365	0.005	0.55 (0.30)	0.002	-3.16 (0.58)	$1.69x10^{-8}$
	17	24086084	25381796	56	24844625[a]	*	0.003	0.42 (0.26)	0.002	-2.21 (0.39)	$1.95x10^{-8}$
	20	60681167	62562427	6	61192205[a]	61100715	0.004	0.21 (0.35)	0.002	-3.81 (0.60)	$9.40x10^{-9}$

[a]intergenic variant, [b]intron variant, [c]downstream gene variant

Table 5.10 The location of the most significantly dimorphic SNPs, limited to the top 5, which were associated with the muscular traits in the Simmental.

Trait	Chr	Start	End	No. of dimorphic SNPS	Most significantly dimorphic SNP	Updated position of most significantly dimorphic SNP (ARS-UCD 1.2)	Male		Female		Significance of dimorphism
							Allele Frequency	Allele Effect (SE)	Allele Frequency	Allele Effect (SE)	
Development of hind quarter	1	61306996	70287318	7	61948489[a]	61358095	0.014	0.63 (0.20)	0.013	-0.75 (0.21)	$1.46x10^{-6}$
	16	25091200	26156103	7	25645958[a]	*	0.007	-0.81 (0.28)	0.008	1.12 (0.25)	$2.47x10^{-7}$
	18	38816641	40478326	7	39974706[a]	39838731	0.003	1.27 (0.40)	0.005	-1.25 (0.32)	$1.05x10^{-6}$
	19	48915944	50012600	4	49415944[a]	48765663	0.011	-0.40 (0.23)	0.003	2.08 (0.47)	$1.86x10^{-6}$
	20	12847020	13908908	29	13347020[b]	13394150	0.006	0.99 (0.30)	0.005	-1.10 (0.32)	$1.50x10^{-6}$
Development of inner thigh	6	106083805	107083825	2	106583805[a]	114749586	0.002	2.31 (0.55)	0.004	-0.98 (0.38)	$8.40x10^{-7}$
	8	13016740	14043577	6	13541906[b]	13620555	0.051	0.51 (0.12)	0.082	-0.22 (0.09)	$7.91x10^{-7}$
	12	65403646	67717725	14	67209065[a]	66672413	0.007	-0.85 (0.34)	0.006	1.76 (0.30)	$1.40x10^{-8}$
	17	29020755	30121482	5	29520755[b]	29099068	0.008	-1.06 (0.30)	0.003	1.73 (0.48)	$7.74x10^{-7}$
	22	2823206	3890106	3	3390106[c]	3346386	0.006	-1.65 (0.34)	0.005	0.72 (0.31)	$3.22x10^{-7}$
Development of loin	2	127861842	128864558	2	128364558[a]	127768904	0.025	-0.23 (0.16)	0.013	1.14 (0.22)	$7.57x10^{-7}$
	10	89463351	90472920	2	89963351[c]	88893256	0.006	1.09 (0.30)	0.005	-1.23 (0.36)	$7.37x10^{-7}$
	12	84618243	85621166	2	85118243[a]	81125408	0.087	-0.45 (0.10)	0.061	0.28 (0.11)	$4.18x10^{-7}$
	25	35252984	36256151	4	35752984[a]	35196894	0.078	-0.27 (0.10)	0.068	0.43 (0.10)	$9.22x10^{-7}$
	28	3386654	4391033	2	3886654[a]	3541484	0.003	-2.29 (0.50)	0.005	1.06 (0.36)	$6.03x10^{-8}$
Thigh width	1	131926057	133026755	19	132524816[a]	131446846	0.186	0.23 (0.07)	0.173	-0.24 (0.07)	$7.78x10^{-7}$
	2	55252366	56331341	6	55802026[a]	55583385	0.022	0.40 (0.19)	0.013	-0.96 (0.23)	$3.25x10^{-6}$
	2	99511847	100585115	3	100011847[a]	99572436	0.014	1.00 (0.23)	0.013	-0.52 (0.21)	$1.23x10^{-6}$
	14	68627562	70444960	30	69137624[a]	66845331	0.059	-0.46 (0.11)	0.037	0.32 (0.13)	$2.92x10^{-6}$
	18	3288714	4902994	7	3802911[a]	3762701	0.015	0.93 (0.22)	0.026	-0.35 (0.16)	$2.56x10^{-6}$
Width of withers	3	97729943	99024542	2	98229943[a]	*	0.015	1.04 (0.24)	0.017	-0.50 (0.22)	$2.13x10^{-6}$
	16	13116847	14137101	55	13629940[a]	13011210	0.446	-0.19 (0.06)	0.487	0.18 (0.05)	$3.41x10^{-6}$
	18	3288714	4302911	3	3802911[a]	3762701	0.015	1.15 (0.24)	0.026	-0.27 (0.17)	$1.64x10^{-6}$
	22	5987150	6987507	3	6487150[a]	6411683	0.062	0.23 (0.11)	0.040	-0.63 (0.14)	$1.57x10^{-6}$
	29	46651577	47656859	2	47156859[a]	46499562	0.027	-0.69 (0.17)	0.032	0.37 (0.15)	$2.27x10^{-6}$

[a]intergenic variant, [b]intron variant, [c]upstream gene variant

Table 5.11 The location of the most significantly dimorphic SNPs, limited to the top 5, which were associated with the skeletal traits in the Simmental.

Trait	Chr	Start	End	No. of dimorphic SNPS	Most significantly dimorphic SNP	Updated position of most significantly dimorphic SNP (ARS-UCD 1.2)	Male		Female		Significance of dimorphism
							Allele Frequency	Allele Effect (SE)	Allele Frequency	Allele Effect (SE)	
Wither height	10	39416014	40523072	17	39916014[b]	39827207	0.003	-0.86 (0.32)	0.004	1.05 (0.29)	$8.11x10^{-6}$
	15	79224399	80224427	2	79724399[d]	78498073	0.031	-0.23 (0.09)	0.016	0.55 (0.15)	$9.47x10^{-6}$
	21	43816205	44823203	2	44316205[a]	43898184	0.005	-0.82 (0.25)	0.005	0.94 (0.27)	$1.70x10^{-6}$
	23	29940021	31111779	72	30451606[a]	30700499	0.314	-0.14 (0.04)	0.319	0.13 (0.04)	$9.09x10^{-7}$
	X	144910600	145867274	8	145410600[b]	134272382	0.007	-0.26 (0.19)	0.011	0.30 (0.22)	$3.54x10^{-8}$
Back length	1	68029603	69086059	16	68580516[bd]	67982646	0.476	0.11 (0.04)	0.483	-0.15 (0.04)	$7.57x10^{-7}$
	6	12731620	13787338	4	13287338[a]	*	0.004	1.11 (0.26)	0.009	-0.44 (0.18)	$1.01x10^{-6}$
	9	84630969	85633807	2	85130969[b]	84014933	0.025	-0.52 (0.12)	0.018	0.40 (0.14)	$6.71x10^{-7}$
	26	23562542	24623399	7	24110489[a]	23866480	0.086	-0.20 (0.06)	0.059	0.30 (0.08)	$6.31x10^{-7}$
	28	39092226	40119176	5	39592226[a]	39277437	0.348	-0.20 (0.04)	0.347	0.08 (0.04)	$5.94x10^{-7}$
Hip width	7	2278460	3359162	15	2785141[a]	2871594	0.109	0.16 (0.05)	0.071	-0.35 (0.08)	$2.81x10^{-7}$
	7	92446667	93472476	2	92972476[a]	90587933	0.029	-0.25 (0.10)	0.020	0.67 (0.15)	$1.94x10^{-7}$
	8	23629760	24686981	6	24129760[c]	24165949	0.004	-0.84 (0.27)	0.003	1.66 (0.39)	$1.51x10^{-7}$
	21	3783435	4810305	6	4308308[a]	4167947	0.129	-0.23 (0.05)	0.125	0.17 (0.06)	$5.70x10^{-7}$
	X	138125669	139135827	3	138625669[a]	138267978	0.067	0.02 (0.06)	0.060	0.29 (0.10)	$5.57x10^{-7}$
Chest width	2	116794370	119288912	3	117931744	*	0.009	0.31 (0.18)	0.005	-1.24 (0.26)	$5.88x10^{-7}$
	6	35397549	36711379	11	36185495[b]	34752887	0.439	0.14 (0.03)	0.472	-0.13 (0.04)	$2.20x10^{-7}$
	12	53187285	54432359	5	53687285[a]	53340630	0.208	0.07 (0.04)	0.228	-0.21 (0.05)	$3.91x10^{-6}$
	16	57116725	58168280	26	57657852[b]	56195234	0.007	-0.82 (0.19)	0.011	0.43 (0.18)	$1.85x10^{-6}$
	22	23885582	25123955	9	24451172[a]	24349773	0.086	0.18 (0.06)	0.121	-0.21 (0.06)	$2.41x10^{-6}$
Chest depth	1	8065635	9123471	237	88582433[a]	87956643	0.004	0.08 (0.24)	0.002	-2.64 (0.42)	$1.98x10^{-8}$
	6	117081578	119074538	2	117581578[a]	112778288	0.008	0.54 (0.17)	0.003	-1.47 (0.34)	$1.95x10^{-7}$
	7	90595916	91626272	3	91121080[a]	88762617	0.004	0.54 (0.25)	0.002	-2.06 (0.42)	$1.27x10^{-7}$
	11	56218949	59206326	7	58002539[a]	58081253	0.002	0.36 (0.38)	0.002	-3.48 (0.60)	$6.15x10^{-8}$
	12	5835154	6893124	2	6335154[a]	6351298	0.004	-0.07 (0.24)	0.002	-3.48 (0.60)	$1.12x10^{-7}$

[a]intergenic variant, [b]intron variant, [c]downstream gene variant, [d]upstream gene variant

5.6.2 Allele Frequencies

The number of SNPs included in each of the analyses in the present study differed by both sex and breed due solely to the SNP not always segregating in each sub-population. While the vast majority of SNPs in the present study had the same minor allele in both sexes, differences between sexes in allele frequencies still exist. In humans, it has been hypothesised that inter-sex differences in allele frequencies occur during the initial evolution of sexual dimorphism due to males and females having different fitness optima for a phenotype (Rice, 1984; Lucotte et al., 2016). In cattle, differences in allele frequencies between the sexes may also have arisen due to these differential fitness optima for males and females where selection occurred at a given locus in one sex but not the other, or due to different selection pressures at that locus in each sex (Lucotte et al., 2016); for example, selection for a female trait could have very different effects on genetically correlated traits in females compared to those traits in males (Bittante et al., 2018).

Inter-sex differences in the frequency of a given allele may also be due to inter-sex differences in recombination rates; up to 75% of species that undergo recombination in their genome have different recombination rates per sex (Burt et al., 1991; Wyman and Wyman, 2013). In the majority of species, the male recombination rate is generally lower than in females (Poissant et al., 2010; Wyman and Wyman, 2013). It is thought that this lower recombination rate is advantageous to males as it maintains combinations of beneficial genes that have undergone sexual selection (Trivers, 1988); however, studies in both cattle (Ma et al., 2015) and sheep (Maddox and Cockett, 2007) reported that the male recombination rate in these species is actually higher than the females.

5.6.3 Across Breed Genetics

Based on a series of within-breed genome-wide associations undertaken across both sexes combined using the data from the present study, Doyle et al. (2020a; 2020b) detected little to no commonality in the genomic regions associated with each type trait across breeds. This indicates that the underlying genetic basis of the same trait in each breed is quite different; therefore it was somewhat expected that the regions exhibiting dimorphism may also differ by breed. In general, the British breeds (AA and HE) had fewer suggestively or significantly dimorphic SNPs than the continental breeds but the British breeds had a greater percentage of SNPs exhibiting significant differences in allele frequencies between the sexes than the continental breeds. The location of the most significantly dimorphic SNPs also differed across the breeds. The differences observed between the breeds may be due to actual differences in the genetic basis of sexual dimorphism among the breeds, as previously observed between the Belgian Blue and Piemontese cattle breeds (Bittante et al., 2018), or may be simply due to differences in the statistical power to detect QTL due to the differences in breed-specific population sizes (Chapter 3). Actual differences in the genetics underlying each trait may be attributable to different mutations affecting specific genes in each breed, such as the breed-specific MSTN mutations, or may, more likely be attributable to different QTL being affected by different selection pressures within each breed (Bittante et al., 2018).

5.6.4 Genes Exhibiting Dimorphism

Single nucleotide polymorphisms exhibiting sexually dimorphism were located within, or close to a number of different genes that have previously been associated with muscularity and/or size in cattle in Chapters 3 and 4. Three genes on BTA2

(NAB1, COL5A2 and IWS1) containing dimorphic SNPs associated with multiple traits are thought to either be in strong linkage disequilibrium with MSTN (Grade et al., 2009) or have been previously identified as being located within a QTL also containing MSTN that was associated with muscularity in beef cattle (Chapter 3). The MSTN gene has already been documented as being responsible for muscular hypertrophy in cattle (Grobet et al., 1997; McPherron and Lee, 1997) and is widely known as the causal variant for many muscularity and carcass traits in cattle (Casas et al., 2000; Allais et al., 2010; Purfield et al., 2019). Another candidate gene for muscularity that exhibited evidence of sexual dimorphism was the PDHX gene on BTA15 that contained dimorphic SNPs for 3 of the muscularity traits in CH and has previously been associated with carcass quality traits in beef cattle (Karisa et al., 2013).

5.7 Conclusion

While many significantly and suggestively sexually dimorphic SNPs associated with the muscular and skeletal type traits were identified in the present study, the location and effect sizes of these tended to be both trait-specific and breed-specific. Both the allele substitution effect sizes and the allele frequencies of the dimorphic SNPs also differed by sex. This indicates that while sexual dimorphism exists in cattle at a genome level, it occurs at a low frequency but also differs both by trait and by breed.

Chapter 6

Thesis Summary, Conclusions & Implications

6.1 Summary of Thesis

The overall objective of this thesis was to determine if the genetic architecture of muscular and skeletal traits differed either by breed or by sex. This was explored using both traditional quantitative genetics approaches based on the estimation of the genetic parameters for these type traits by both breed and sex (Chapter 2) but also by studying genomic regions associated with each trait in each breed; the latter was undertaken for both muscular (Chapter 3) and skeletal (Chapter 4) traits. A final objective was to determine if genomic regions exhibited sexual dimorphism for any linear type trait (Chapter 5). Data originated from the national database of the Irish Cattle Breeding Federation with the imputation of whole genome sequence using data from the 1000 Bulls' genome project. Data on a total of 18 linear type traits were available on 81,200 beef cattle and 117,151 Holstein-Friesian (HF) dairy cows while genotypes were available for 19,449 beef cattle and 4,494 dairy cows.

Genetic evaluations for most performance traits of beef cattle in Ireland are based on data from crossbred animals; in such evaluations, all data are analysed together with a single variance component used per trait for the entire population. Whether these (co)variance components vary by breed has never been explored in Ireland; while data for purebred animals exists for a selection of traits, many of these phenotypes could exhibit bias. For example, only the poorer quality or injured purebred animals are slaughtered at a young age and the performance of many purebred cows would not reflect normal commercial practices. Purebred data does, however, exist for linear scores which are scored using the same scale across breeds by the same classifiers thus providing a rich source of data to explore if (co)variance components and the genomic architecture of animal characteristics in beef cattle differ by breed.

6.1.1 Chapter 1: Literature Review

Objective: To review the existing literature on the genetics and genomics of linear type traits and sexual dimorphism

- The ICBF publishes two genetic indexes for beef cattle; the Terminal Index and the Replacement Index, both of which are multi-breed indexes with a strong emphasis on carcass traits
- Linear type traits are useful early indicators of carcass merit and can often provide more granular information of the carcass than just carcass weight, fat, or conformation alone
- Heritability estimates for the functional traits are generally low (0.04 – 0.13), while the heritability estimates for the skeletal and muscular traits in both beef and dairy cattle are generally moderate to high (0.06 – 0.43)
- Many previous genome wide association studies have generally focused on stature, overall muscling or carcass traits but in only a single breed of cattle
- The allele frequency per SNP as well as the associated allele substitution effects have previously been documented to vary among breeds of cattle
- Gaps in knowledge:
    - Do the genetic parameters for functional, skeletal, and muscular linear type traits differ by breed of cattle?
    - Do SNPs and possible candidate genes associated with muscular and skeletal traits in both dairy and beef cattle vary across breeds and across traits?
    - Evidence of sexual dimorphism at the genome level for muscular and skeletal traits in beef cattle

6.1.2 Chapter 2: Genetic (co)variance components within and among linear type traits differ among contrasting cattle breeds

Objective: To determine if genetic (co)variance components for linear type traits differed among five contrasting cattle breeds and, also, if these traits differed genetically by sex.

- Linear type trait data on 18 traits from 81,200 beef animals of 5 breeds
- Genetic (co)variance components estimated within each breed and each sex separately using animal linear mixed models
- Heritability estimates ranged from 0.00 to 0.13 for the functional traits, from 0.10 to 0.28 for the muscular traits and from 0.06 to 0.43 for the skeletal traits
- Breed differences existed in the heritability estimates for 13 out of the 18 type traits analysed and between the pairwise within-breed genetic correlations among the linear type traits.
- Genetic correlations between the same linear type traits in both sexes were all close to unity (>0.90).
- Overall, the linear type traits in the continental breeds (i.e. Charolais (CH), Limousin (LM), Simmental (SI)) tended to have similar heritability estimates to each other as well as similar genetic correlations among the same pairwise traits, as did the traits in the British breeds (i.e. Angus (AA), Hereford (HE).
- In a selection index, the impact of using the genetic (co)variance components of the LM to predict genetic merit for height of withers in CH was less than if the genetic (co)variance components of the AA were used
- There is little advantage in considering the sexes as separate traits but improved accuracy of estimated breeding value could be achieved by considering, at least, the British breeds separate to the continental breeds

6.1.3 Chapter 3: Genomic regions associated with muscularity in beef cattle differ in five contrasting cattle breeds

Objective: To identify genomic regions associated with five muscularity linear type traits and to determine if these associated regions are common across multiple beef cattle breeds

- Phenotypic data on 5 muscular type traits coupled with imputed whole genome sequence data on 19,449 animals
- Association analyses undertaken using a linear mixed model fitting each SNP separately
- Several quantitative trait loci located across the entire genome
- The only region common to more than one breed and one trait was a large-effect pleiotropic QTL on BTA2 containing the *MSTN* gene
- Other plausible candidate genes for muscularity traits included *PDE1A*, *PPP1R1C* and multiple collagen and *HOXD* genes
- Associated GO terms and KEGG pathways tended to differ by breed and trait
- Most of the SNPs associated with any of the traits were intergenic or intronic SNPs located within regulatory regions of the genome
- The extensive differences observed between the breeds may be due to the much smaller sample sizes for AA, HE, and SI compared to the CH and LM populations that result in reduced power to detect QTL or they may be due to differences in genetic architecture of these traits among the populations
- Knowledge of these differences in genetic architecture among breeds will be useful for developing accurate genomic prediction equations that can operate robustly across breeds

6.1.4 Chapter 4: Genomic regions associated with skeletal type traits in beef and dairy cattle are common to regions associated with carcass traits, feed intake and calving difficulty

Objective: To identify genomic regions associated with five skeletal linear type traits and to determine if these associated regions are common across multiple beef cattle breeds

- Phenotypic data from 5 skeletal type traits in beef cattle with 3 comparable skeletal type traits in dairy cattle along with imputed whole genome sequence data on 19,449 beef cattle and 4,494 Holstein-Friesian dairy cows
- Association analyses undertaken using a linear mixed model fitting each SNP separately
- Majority of QTLs identified were both trait-specific and breed-specific
- A large-effect pleiotropic QTL on BTA6 containing the *NCAPG* and *LCORL* genes that was associated with all 5 traits in the Limousin and wither height in the Angus was the only overlap among the beef breeds
- Only two QTLs, both of which were on BTA5 and were associated with height in the Angus and the Holstein-Friesian, overlapped the beef and dairy breeds
- Several detected QTLs overlapped with QTLs documented elsewhere as being associated with carcass traits, feed intake, and calving difficulty
- The associated QTL differed by trait, which suggests that breeding for a morphologically different (i.e., longer and taller versus shorter and smaller) more efficient animal may become possible in the future

6.1.5 Chapter 5: Genomic regions exhibit sexual dimorphism for size and muscularity in cattle

Objective: To determine if genomic regions associated with size and muscularity in cattle exhibited signs of sexual dimorphism

- Phenotypic data on 5 skeletal and 5 muscular type traits from 5 beef cattle breeds along with imputed whole genome sequence data from 19,449 beef cattle including the X chromosome
- Association analyses carried out using linear mixed models in each sex and each breed separately
- Significantly ($p \leq 1 \times 10^{-8}$) sexually dimorphic SNPs were detected for three traits in the Angus, seven traits in the Charolais, and three in the Limousin
- Suggestively ($p \leq 1 \times 10^{-5}$) sexually dimorphic SNPs were identified for all traits in all breeds
- Between 86% (Angus) and 94% (Limousin) of SNPs that were segregating in both sexes had the same minor allele
- Differences ($p \leq 0.05$) in allele frequencies between the sexes were observed for between 36% (Limousin) and 66% (Angus) of the total SNPs
- Dimorphic SNPs were located within a number of genes including the *NAB1*, *COL5A2*, and *IWS1* genes on BTA2 that are located close to the *MSTN* gene
- Sexual dimorphism does exist in cattle at a genome level, but it is not consistent by either trait or breed
- It is unlikely that consideration of sexual dimorphism in beef cattle will improve the accuracy of genomic predictions, at least for the traits and breeds investigated in the present study

6.2 Overall Conclusions & Implications

Both the replacement and terminal selection indexes published for beef cattle in Ireland are calculated using multi-breed genetic evaluations with both having some emphasis on carcass traits. Results from this thesis described the differences in genetic parameters for linear type traits between breeds and sexes (Chapter 2), the interbreed differences in QTL associated with muscular (Chapter 3) and skeletal traits (Chapter 4), while also investigating the presence of sexual dimorphism at a genome level (Chapter 5). Furthermore, this thesis attempted to identify QTL that overlapped between the scored linear type traits and economically important carcass merit (Chapter 3 & 4). Therefore, the implications of this thesis are: 1) identified breed differences in (co)variance components and QTL that may be crucial to improving the accuracy of future across-breed genetic and genomic evaluations, 2) identified QTL associated with linear type traits that co-located with previously documented QTLs for carcass traits, which may therefore be useful in improving carcass merit, 3) confirmed the presence of sexual dimorphism at a genome level in beef cattle but not at a frequency that consideration of this phenomenon would noticeably improve genomic predictions.

6.2.1 Across-Breed Genetics & Genomics

6.2.1.1 Current multi-breed genetic evaluations

The multi-breed genetic evaluation models used by the ICBF vary based on the trait being evaluated with the models used to calculate the six beef performance traits (i.e., carcass weight, carcass conformation, carcass fat, feed intake, cow live weight and cull cow weight) included in the selection indexes split into three multi-trait models (ICBF, 2020): 1. carcass weight, 2. carcass conformation, and 3. feed intake

models. There are 12 traits included in the carcass weight model: carcass weight, 150-250 day weight, 250-350 day weight, 350-450 day weight, 450-550 day weight, 550-700 day weight, cow live weight, cull cow carcass weight, skeletal type trait composite score, foreign weaning weight EBV, foreign skeletal EBV and foreign carcass weight EBV. There are 9 traits included in the carcass conformation model: carcass conformation, cow carcass conformation, muscle type trait composite score, calf quality, calf price, weanling price, post weanling price, foreign muscle EBV and foreign skeletal EBV. The 11 traits included in the feed intake model are: feed intake, carcass weight, carcass conformation, carcass fat, 350-450 day weight, 450-550 day weight, 550-700 day weight, skeletal type trait composite score, foreign weaning weight EBV, foreign carcass weight EBV and foreign carcass conformation EBV. The beef performance traits are all expressed relative to the same base population and provide one variance component for all breeds and sexes.

Since 2008, foreign EBVs, supplied by France and the UK, for traits such as calving difficulty, maternal calving difficulty, direct and maternal weaning weight, linear type traits, and carcass weight have been included in the Irish national genetic evaluations. These EBVs are provided for 10 major breeds including the AA, CH, HE, LM, SI, Blonde d'Aquitaine, Partenaise, Saler, Aubrac and Belgian Blue. The foreign EBVs are modelled using approximated daughter yield deviations (Bonaiti and Boichard, 1995) and are then weighted in the Irish genetic evaluation depending on the accuracy/reliability from the country of origin.

The ICBF currently produce EBVs for 13 of the linear type traits described in this thesis; width at withers, width behind withers, development of the loin, development of the hindquarter, thigh width, wither height, back length, pelvic length, hip width, locomotion, hind leg side view, hind leg rear view, and front leg

front view. The genetic evaluations take into account the level of heterosis, the fraction of dairy bloodline in the animals, the age of the animal, the animal's birth year, the dam age nested within the dam parity, the dam's permanent environment, and the scorer that scored the animal as effects in the model. As the genetic evaluation of the linear type traits is part of a multi-breed evaluation, only one variance component is used for each trait in the evaluation, regardless of breed. The genetic correlations estimated between these traits in the Irish national evaluation are all stronger than those estimated between the same traits, within breed, in Chapter 2. The efficiency of selection using the (co)variance components adopted in the current Irish national genetic evaluation versus the assumed correct (co)variance components estimated in Chapter 2 for each breed was quantified using the method for estimating index sensitivity as outlined in Chapter 2 and adapted from Ochsner et al. (2017). The efficiency of selection was as low as 0.62 when the parameters from the current Irish national genetic evaluation were used in place of the parameters from the CH breed. This efficiency reduced to 0.59 in the AA, 0.58 in the HE, 0.55 in the LM, and 0.53 in the SI when the parameters of these breeds were replaced by those used in the current Irish national genetic evaluation.

6.2.1.2 Why Linear Type Traits?

Carcass traits, including carcass fat, carcass conformation, and carcass weight account for 14% of the emphasis within the Irish beef replacement index and 57% of the emphasis within the Irish beef terminal index. Carcass weight is defined as the weight of both half carcasses after being bled, eviscerated and after removal of skin, genitalia, limbs, head, tail, kidneys and the udder (ICBF, 2020). Carcass conformation is the shape and development of the carcass and is denoted by the

letters E, U, R, O, P with E being the best and P the poorest and subsequently divided into a 15-point scale with the use of +, =, and – for each letter grade. Carcass fat is also scored on a 15 point scale denoted by the number 1 to 5, where 1 is lean and 5 is fattest, and subsequently divided into 15 points using +,=, and – for each number. Carcass traits cannot be measured on live animals so EBVs are based on the parental average, the genotype of the animal, correlated traits, and progeny records.

Linear type traits measured on young, live beef animals are known to be strongly genetically correlated with carcass merit (Mukai et al., 1995; Conroy et al., 2010), meaning linear type traits may be useful as early phenotypic predictors of the carcass traits. In fact, linear type traits are currently included in the carcass trait genetic evaluations as a composite score based on overall muscularity and skeletal conformation. Selection based on these composite traits would increase overall muscularity or conformation; however, the correlation between carcass merit with any one type trait is not equal to 1 which implies that the same carcass value can be achieved with morphologically different animals (Figure 6.1). By extension then, this implies that, for example, an animal with a longer better developed back and a shallow chest may have the same meat yield as an animal with a lesser developed shorter back and a deep chest. Such morphological differences could contribute, in turn, to differences in individual carcass retail cut weights, and thus overall carcass value.

Overall, 35% of the QTL associated with the linear type traits in Chapters 3 and 4 had previously been associated with carcass merit, muscle or body weight (McClure et al., 2010; Snelling et al., 2010; Lu et al., 2013; Saatchi et al., 2014). The QTL associated with body weight were associated with each of the five skeletal type traits; this suggests that breeding for a morphologically different animal, with a

higher value carcass, while maintaining a similar body weight may become possible in the future. Traditionally, adding value to a carcass was accomplished by selecting for heavier animals; however, the impact of animal weight on production efficiency has questioned this practice (Ferrell and Jenkins, 1984). Furthermore, previous research into the genetic variability of individual primal cut weights in cattle demonstrated that there was significant exploitable genetic variability in cut weight even if the goal is to increase primal cut weight without an increase in carcass weight (Judge et al., 2019). Therefore, the ability to morphologically change a carcass and add value to it without increasing carcass weight may be of benefit to beef producers.

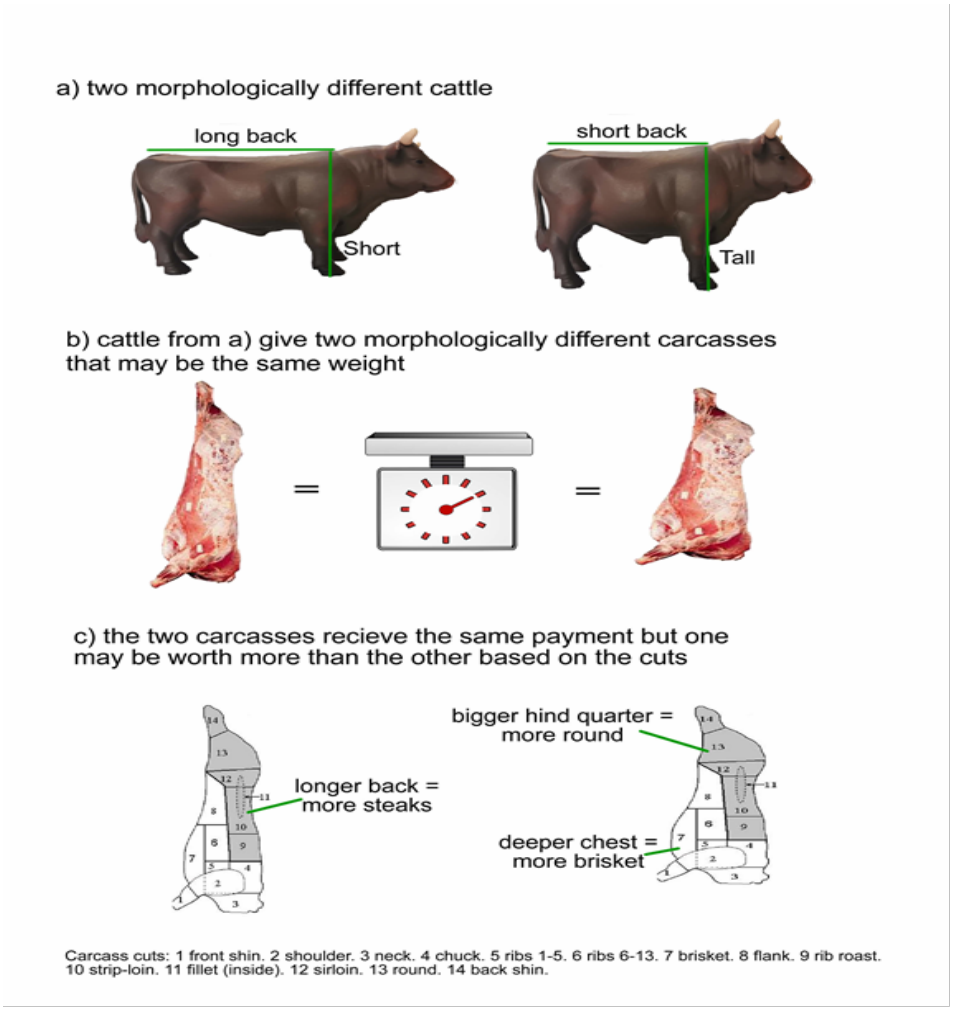


Figure 6.1 Graphical depiction of morphologically different animals and the differences in yield that may occur.

6.2.1.3 Genetic Architecture

Genetic correlations between two traits may be due to pleiotropy or genetic linkage; understanding the contribution of each to the genetic correlation provides information on the capacity to improve both traits simultaneously through a selection index framework. The traits with stronger genetic correlations between them would be expected to either have a greater number of QTL in common due to the presence of pleiotropy, or have a smaller number of QTL with large effect sizes common to both. The correlations among the type traits were documented in Chapter 2 and the QTL associated with each trait were documented in Chapter 3 and Chapter 4. In general, within each breed, the number of QTL in common between traits increased when the genetic correlations between the traits strengthened (Figure 6.2).

The number of QTL in common between traits, regardless of the strength of the genetic correlation, was greater in the CH and LM breeds when compared to the other three breeds; this may be due to the common geographical origin of these breeds and the interbreeding of these populations over time. Each of the muscular traits in the CH and the LM were associated with *MSTN* (Chapter 3) while each of the skeletal traits in the LM were associated with *NCAPG/LCORL* (Chapter 4). However, the other QTL were all trait and breed-specific.

The different QTL, and genes, associated with the traits in each breed are due to allele effects differing within each breed but may also be due to the linkage phase between the genotyped allele and the causal mutations varying by breed. These substantial differences in associated QTL may have implications on the accuracy of the multi-breed genetic evaluations and also on any across breed genomic evaluations.

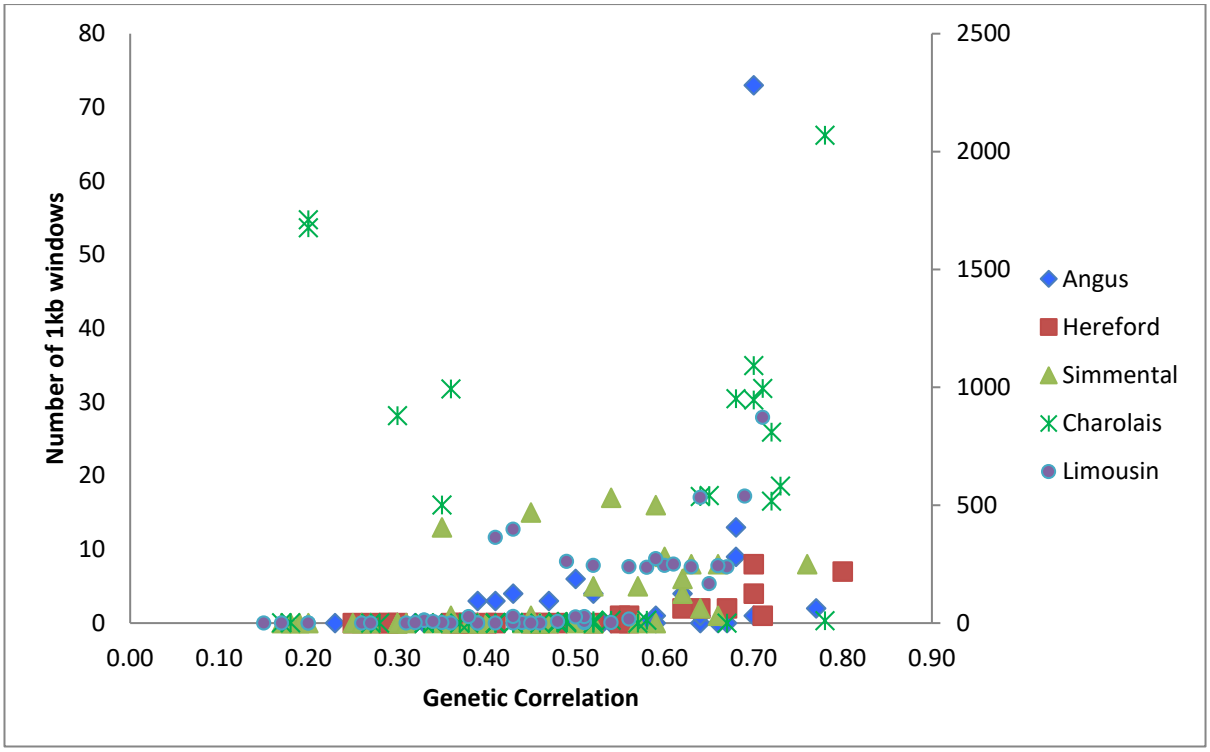


Figure 6.2 The genetic correlations between the type traits and the number of 1kb windows in common between those traits. Angus, Hereford, and Simmental on the left axis. Charolais and Limousin on the right axis.

6.2.2 Considerations for Future Evaluations

Chapter 2 of this thesis demonstrates that although genetic variation exists for the vast majority of the linear type traits in each breed, the variance components (and their ratio) of each linear type trait and the genetic correlations between each linear type trait differs by breed (Table 6.1) which may have implications for the current national multi-breed Irish genetic evaluations. Overall, the (co)variance components estimated in the present study for the British breeds of AA and HE were very similar to one another, as were the continental breeds of CH, LM and SI. The impact of using incorrect (co)variance components (i.e. those of a different breed to the target breed) in a selection index demonstrated that the loss of accuracy was less when the CH parameters were substituted with the LM or SI parameters over the AA

parameters (Chapter 2). This indicates that future genetic evaluations may benefit from considering, at least, the continental breeds separately to the British breeds. Nonetheless, the added complexity of the models should be considered should separate variance components be fitted for each breed; moreover, it is not clear how to consider crossbred data which predominate in Ireland for commercially relevant traits.

Furthermore, one of the main advantages of the current multi-breed genetic evaluation system is the ability to objectively compare all animals of different breeds (and crossbreds) against each other. This is hugely important for both beef and dairy producers since producers are not always wedded to a single breed and are appreciative of the capability to compare candidate sires of all breeds against each other. The end result is that one may have to accept some level of inefficiency in the genetic evaluation in the pursuit of across-breed genetic evaluations where selection bias is also properly accounted for through the inclusion of correlated traits in the multi-trait evaluation. A possible source of inefficiency in a genetic evaluation stems from the use of incorrect variance components in the calculation which can lead to either over- or under-prediction of EBVs. As selection decisions in livestock genetic improvement programmes are typically based on EBVs, the accuracy and credibility of the EBV is of particular importance (Bijma, 2012). An overestimated EBV would place an animal higher up a ranking list however, the progeny of this animal will generally not perform as well as expected. This can lead to a reduced selection response and in turn lead to a decrease in the rate of genetic gain. In more practical terms, any long-term or marked inefficiency in genetic evaluations can lead to producers losing faith in the evaluations and moving away from utilizing EBVs in their selection decisions; again, leading to a decrease in the rate of genetic gain.

Thus, in the pursuit of accurate across-breed genetic evaluations, bias that may occur among the different breeds should be properly accounted for by using the most accurate set of genetic variance components available in order to limit any inefficiency that may occur.

Table 6.1 Heritability estimates (h^2) and standard errors (SE) of the linear type traits when all breeds were considered as one population and each breed being was evaluated independently.

	All Breeds	**Angus**	**Charolais**	**Hereford**	**Limousin**	**Simmental**
Trait	**h^2 (SE)**	**h^2 (SE)**	**h^2 (SE)**	**h^2 (SE)**	**h^2 (SE)**	**h^2 (SE)**
Functional						
Locomotion	0.08 (0.01)	0.12 (0.04)	0.12 (0.02)	0.00	0.04 (0.01)	0.04 (0.02)
Foreleg front view	0.08 (0.01)	0.13 (0.04)	0.09 (0.01)	0.00	0.06 (0.01)	0.06 (0.02)
Hind-leg side view	0.11 (0.01)	0.08 (0.03)	0.09 (0.01)	0.11 (0.04)	0.08 (0.01)	0.06 (0.02)
Hind-leg rear view	0.09 (0.01)	0.04 (0.03)	0.06 (0.01)	0.00	0.04 (0.01)	0.06 (0.02)
Muscular						
Development of hind quarter	0.39 (0.01)	0.22 (0.05)	0.30 (0.02)	0.14 (0.04)	0.25 (0.02)	0.24 (0.03)
Development of loin	0.24 (0.01)	0.13 (0.04)	0.21 (0.02)	0.10 (0.04)	0.17 (0.01)	0.18 (0.03)
Thigh width	0.29 (0.01)	0.14 (0.04)	0.22 (0.22)	0.16 (0.05)	0.23 (0.01)	0.24 (0.03)
Development of inner thigh	0.36 (0.01)	0.14 (0.04)	0.28 (0.02)	0.20 (0.05)	0.24 (0.02)	0.23 (0.03)
Width of withers	0.26 (0.01)	0.22 (0.05)	0.21 (0.02)	0.16 (0.05)	0.19 (0.01)	0.22 (0.03)
Width behind withers	0.25 (0.01)	0.13 (0.04)	0.18 (0.02)	0.15 (0.05)	0.17 (0.01)	0.18 (0.03)
Skeletal						
Width of chest	0.14 (0.01)	0.07 (0.03)	0.10 (0.02)	0.00	0.10 (0.01)	0.15 (0.03)
Depth of chest	0.17 (0.01)	0.15 (0.04)	0.13 (0.02)	0.25 (0.06)	0.15 (0.01)	0.14 (0.03)
Height of withers	0.36 (0.01)	0.19 (0.05)	0.43 (0.02)	0.30 (0.06)	0.29 (0.02)	0.34 (0.03)
Length of pelvis	0.21 (0.01)	0.17 (0.04)	0.23 (0.02)	0.27 (0.06)	0.19 (0.01)	0.20 (0.03)
Length of back	0.27 (0.01)	0.17 (0.04)	0.30 (0.02)	0.29 (0.06)	0.23 (0.02)	0.20 (0.03)
Width at hips	0.14 (0.01)	0.06 (0.04)	0.13 (0.01)	0.00	0.14 (0.01)	0.14 (0.03)
Other						
Body condition score	0.12 (0.01)	0.03 (0.03)	0.13 (0.02)	0.00	0.11 (0.01)	0.05 (0.02)
Docility	0.16 (0.01)	0.21 (0.05)	0.15 (0.02)	0.11 (0.04)	0.17 (0.01)	0.09 (0.02)

6.2.3 Sexual Dimorphism

The Irish national beef genetic evaluations consider the sexes as genetically similar; therefore traits in each sex are evaluated together. In Chapter 2 of this thesis, the genetic correlations between the pairwise linear type traits in the different sexes were documented to all be greater than >0.90. Furthermore, the percentage of SNPs documented as being sexually dimorphic was less than 0.01% in each trait and each breed (Table 6.2). Sexual dimorphism does exist at the genome level in cattle; however, it is not common among breeds or traits. Therefore, evaluating the sexes as one in genetic evaluations is not thought to affect the accuracy of the evaluations.

Table 6.2 The percentage of SNPs in each trait in each breed that were sexually dimorphic ($p \leq 1 \times 10^{-5}$)

	Angus	Charolais	Hereford	Limousin	Simmental
Chest depth	0.0017	0.0088	0.0009	0.0040	0.0039
Chest width	0.0017	0.0187	0.0016	0.0065	0.0008
Back length	0.0002	0.0010	0.0015	0.0013	0.0005
Hip width	0.0084	0.0016	0.0017	0.0006	0.0010
Wither height	0.0010	0.0004	0.0008	0.0005	0.0007
Development of hind quarter	0.0002	0.0021	0.0005	0.0025	0.0007
Development of inner thigh	0.0022	0.0003	0.0002	0.0014	0.0005
Development of loin	0.0018	0.0006	0.0006	0.0004	0.0013
Thigh width	0.0009	0.0013	0.0004	0.0003	0.0009
Wither width	0.0003	0.0008	0.0004	0.0015	0.0005

6.3 Future Research

While this thesis focused on the genetic architecture of purebred cattle, the majority of beef cattle in Ireland are crossbred. Crossbred cattle are thought to have many advantages over a purebred animal due to the benefit of both breed complementarity and heterosis. Previous research on crossbred beef animals has determined that they matured earlier and were heavier at maturity than purebred animals (Mendonça et al., 2019). Further research in the beef industry could examine the effect of crossbreeding on the genetics and genomics of the linear type traits and how crossbred animals would be evaluated if the continental breeds were to be considered separately to the British breeds in future evaluations.

The sheep genetic evaluations in Ireland are also multi-breed so the same issues identified in the cattle evaluations apply to the sheep; all breeds are evaluated as one breed and one (co)variance component is calculated for each trait regardless of breed. Future research could therefore be carried out on the genetic architecture of traits in the different sheep breeds, or indeed any other species, where multi-breed evaluations are used.

Further research on carcass merit could potentially focus on the morphology of the carcass and its potential for increasing carcass value. Previous research has discovered that the muscular linear type traits have some predictive ability for genetic merit of certain carcass cuts (Berry et al., 2019) and that moderate genetic correlations exist between the muscular linear type traits and kill out percentage (Berry et al., 2020), suggesting that type traits may be useful in a breeding program to help achieve accurate genetic evaluations for carcass merit. In future, a controlled study could be carried out whereby, after slaughter, animals with a similar carcass

weight but a different morphology are identified and the quantity of each cut is measured and compared between carcasses. This would assist in the definitive conclusion as to whether breeding for a different morphology is indeed a useful strategy to increase overall carcass value by increasing the amount of higher value cuts of meat.

Chapter 7
References

Alexander, L., L. Kuehn, T. Smith, L. Matukumalli, B. Mote, J. Koltes, J. Reecy, T. Geary, D. Rule, and M. MacNeil. 2009. A Limousin specific myostatin allele affects longissimus muscle area and fatty acid profiles in a Wagyu-Limousin F 2 population 1 2. Journal of Animal Science. 87(5):1576-1581.

Ali, T. E., E. Burnside, and L. Schaeffer. 1984. Relationship between external body measurements and calving difficulties in Canadian Holstein-Friesian cattle. Journal of Dairy Science. 67(12):3034-3044.

Allais, S., H. Levéziel, N. Payet-Duprat, J.F. Hocquette, J. Lepetit, S. Rousset, C. Denoyelle, C. Bernard-Capel, L. Journaux, and A. Bonnot. 2010. The two mutations, Q204X and nt821, of the myostatin gene affect carcass and meat quality in young heterozygous bulls of French beef breeds. Journal of Animal Science. 88(2):446-454.

Allen, A., G. Minozzi, E. Glass, R. Skuce, S. McDowell, J. Woolliams, and S. Bishop. 2010. Bovine tuberculosis: the genetic basis of host susceptibility. Proceedings of the Royal Society B: Biological Sciences 277(1695):2737-2745.

Al-Zaidy, S., L. Rodino-Klapac, and J. R. Mendell. 2014. Gene therapy for muscular dystrophy: moving the field forward. Pediatric neurology 51(5):607-618.

Amer, P. R., G. Simm, M. G. Keane, M. G. Diskin, and B. W. Wickham. 2001. Breeding objectives for beef cattle in Ireland. Livestock Production Science 67(3):223-239.

Andersen, B. B., T. Lykke, K. Kousgaard, L. Buchter, and J. Pedersen. 1977. Growth, feed utilization, carcass quality and meat quality in Danish dual-purpose cattle: Genetic analysis concerning 5 years data from the progeny tests for beef production at Egtved. Beretning fra Statens Husdyrbrugsforsoeg .

Anderson, C. A., F. H. Pettersson, G. M. Clarke, L. R. Cardon, A. P. Morris, and K. T. Zondervan. 2010. Data quality control in genetic case-control association studies. Nature protocols 5(9):1564.

Arango, J. A., L. V. Cundiff, and L. D. Van Vleck. 2002. Breed comparisons of Angus, Charolais, Hereford, Jersey, Limousin, Simmental, and South Devon for weight,

weight adjusted for body condition score, height, and body condition score of cows. Journal of Animal Science. 80: 3123-3132.

Archer, J., E. Richardson, R. Herd, and P. Arthur. 1999. Potential for selection to improve efficiency of feed use in beef cattle: a review. Australian Journal of Agricultural Research. 50(2):147-162.

Arthur, P., J. Archer, D. Johnston, R. Herd, E. Richardson, and P. Parnell. 2001. Genetic and phenotypic variance and (co)variance components for feed intake, feed efficiency, and other postweaning traits in Angus cattle. Journal of Animal Science. 79(11):2805-2811.

Baum, L. E., and J. A. Eagon. 1967. An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. Bulletin of the American Mathematical Society 73(3):360-363.

Baum, L. E., and T. Petrie. 1966. Statistical inference for probabilistic functions of finite state Markov chains. The Annals of Mathematical Statistics 37(6):1554-1563.

Baum, L. E., T. Petrie, G. Soules, and N. Weiss. 1970. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. The Annals of Mathematical Statistics 41(1):164-171.

Bellinge, R., D. Liberles, S. Iaschi, P. O'brien, and G. Tay. 2005. Myostatin and its implications on animal breeding: a review. Animal Genetics. 36(1):1-6.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society. Series B (Methodological):289-300.

Berns, C. M. 2013. The evolution of sexual dimorphism: understanding mechanisms of sexual shape differences. Sexual dimorphism. Rijeka: InTech:1-16.

Berry, D. P., R. Buckley, P. Dillon, R. D. Evans, and R. F. Veerkamp. 2004. Genetic relationships among linear type traits, milk yield, body weight, fertility and somatic

cell count in primiparous dairy cows. Irish Journal of Agricultural and Food Research 43(2):161-176.

Berry, D., L. Shalloo, A. Cromie, R. Veerkamp, P. Dillon, P. Amer, J. Kearney, R. Evans, and B. Wickham. 2007. The economic breeding index: a generation on. Irish Cattle Breeding Federation:29-34.

Berry, D. P., M. L. Bermingham, M. Good, and S. J. More. 2011. Genetics of animal health and disease in cattle. Irish Veterinary Journal 64(1):5.

Berry, D. P., and R. D. Evans. 2014. Genetics of reproductive performance in seasonal calving beef cows and its association with performance traits. Journal of Animal Science. 92(4):1412-1422.

Berry, D.P., E. Wall, and J. Pryce. 2014. Genetics and genomics of reproductive performance in dairy and beef cattle. Animal. 8: 105-121.

Berry, D. P. 2015. Breeding the dairy cow of the future: what do we need? Animal Production Science 55(7):823-837.

Berry, D.P., A. Twomey, R. Evans, A. Cromie, and S. Ring. 2019. Heritability–what is it, and what is it not; implications for improving cattle health. Cattle Practice 27(1)

Berry, D.P., Amer, R. Evans, T. Byrne, A. Cromie, and F. Hely. 2019. A breeding index to rank beef bulls for use on dairy females to maximize profit. Journal of Dairy Science. 102(11):10056-10072.

Berry, D. P., T. Pabiou, R. Fanning, R. D. Evans, and M. M. Judge. 2019. Linear classification scores in beef cattle as predictors of genetic merit for individual carcass primal cut yields. Journal of Animal Science. 97(6):2329-2341.

Berry, D. P., J.M. Coyne, J.L. Doyle, and R.D. Evans. 2020. The use of subjectively assessed muscular and skeletal traits on live cattle to aid in differentiation between animal genetically divergent in carcass kill out metrics. Livestock Science, 103984.

Berry, J. F., and R. Shine. 1980. Sexual size dimorphism and sexual selection in turtles (Order Testudines). Oecologia 44(2):185-191.

Bern, E. L., J. A. Cobuci, C. N. Costa, C. M. McManus, and J. Braccini. 2015. Genetic association between longevity and linear type traits of Holstein cows. Scientia Agricola. 72: 203-209.

Bijma, P. 2012. Accuracies of estimated breeding values from ordinary genetic evaluations do not reflect the correlation between true and estimated breeding values in selected populations. Journal of Animal Breeding and Genetics 129(5):345-358.

Bittante, G., A. Cecchinato, F. Tagliapietra, R. Verdiglione, A. Simonetto, and S. Schiavon. 2018. Crossbred young bulls and heifers sired by double-muscled Piemontese or Belgian Blue bulls exhibit different effects of sexual dimorphism on fattening performance and muscularity but not on meat quality traits. Meat science 137:24-33.

Boccard, R., R. Naude, D. Cronje, M. C. Smit, H. Venter, and E. J. Rossouw. 1979. The influence of age, sex and breed of cattle on their muscle characteristics. Meat Science 3(4):261-280.

Boichard, D., H. Chung, R. Dassonneville, X. David, A. Eggen, S. Fritz, K. J. Gietzen, B. J. Hayes, C. T. Lawley, and T. S. Sonstegard. 2012. Design of a bovine low-density SNP array optimized for imputation. PloS one 7(3)

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

Bolormaa, S., J. Pryce, B. Hayes, and M. Goddard. 2010. Multivariate analysis of a genome-wide association study in dairy cattle. Journal of Dairy Science. 93(8):3818-3833.

Bolormaa, S., L. Neto, Y. Zhang, R. Bunch, B. Harrison, M. Goddard, and W. Barendse. 2011. A genome-wide association study of meat and carcass traits in Australian cattle. Journal of Animal Science. 89(8):2297-2309.

Bonaiti, B. and D. Boichard. 1995. Accounting for foreign information in genetic evaluation. Interbull Bulletin. 11.

Bouquet, A., M.N. Fouilloux, G. Renand, and F. Phocas. 2010. Genetic parameters for growth, muscularity, feed efficiency and carcass traits of young beef bulls. Livestock Science. 129(1-3):38-48.

Bouwman, A. C., H.D. Daetwyler, A.J. Chamberlain, C.H. Ponce, M. Sargolzaei, F.S. Schenkel, G. Sahana, A. Govignon-Gion, S, Boitard, M. Dolezal, H. Pausch, R.F. Brøndum, P.J. Bowman, B. Thomsen, B. Guldbrandtsen. M.S. Lund, B. Servin, D.J. Garrick, J. Reecy, J. Vilkki, A. Bagnato, M. Wang, J.L Hoff, R.D Schnabel, J.F. Taylor, A.A.E. Vinkhuyzen, F. Panitz, C. Bendixen, L.E. Holm, B. Gredler, C. Hozé, M. Boussaha, M.P. Sanchez, D. Rocha, A. Capitan, T. Tribout, A. Barbat, P. Croiseau, C. Drögemüller, V. Jagannathan, C. Vander Jagt, J.J. Crowley, A. Bieber, D.C. Purfield, DP. Berry, R. Emmerling, K.U. Götz, M. Frischknecht, I. Russ, J. Sölkner, C.P. Van Tassell, R. Fries, P. Stothard, R.F. Veerkamp, D. Boichard, M.E. Goddard and B.J. Hayes. (2018). Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. Nature Genetics. 50(3): 362-367.

Box, G. E. 1949. A general distribution theory for a class of likelihood criteria. Biometrika. 36: 317-346.

Box, G. E. 1953. Non-normality and tests on variances. Biometrika 40: 318-335.

Boyle, E. A., Y. I. Li, and J. K. Pritchard. 2017. An expanded view of complex traits: from polygenic to omnigenic. Cell 169(7):1177-1186.

Brotherstone, S. 1994. Genetic and phenotypic correlations between linear type traits and production traits in Holstein-Friesian dairy cattle. Animal Science. 59(2):183-187.

Burt, A., G. Bell, and P. H. Harvey. 1991. Sex differences in recombination. Journal of Evolutionary Biology. 4(2):259-277.

Bush, W. S., and J. H. Moore. 2012. Chapter 11: Genome-wide association atudies. PLoS Computational Biology. 8(12):e1002822.

Calus, M. P., H. A. Mulder, and J. W. Bastiaansen. 2011. Identification of Mendelian inconsistencies between SNP and pedigree information of sibs. Genet. Sel. Evol. 43(1):34.

Cantor, R. M., K. Lange, and J. S. Sinsheimer. 2010. Prioritizing GWAS results: a review of statistical methods and recommendations for their application. The American Journal of Human Genetics 86(1):6-22.

Carthy, T., D. Ryan, A. Fitzgerald, R. Evans, and D. P. Berry. 2016. Genetic relationships between detailed reproductive traits and performance traits in Holstein-Friesian dairy cattle. Journal of Dairy Science. 99(2):1286-1297.

Casas, E., J. Keele, S. Shackelford, M. Koohmaraie, T. Sonstegard, T. Smith, S. Kappes, and R. Stone. 1998. Association of the muscle hypertrophy locus with carcass traits in beef cattle. Journal of Animal Science. 76(2):468-473.

Casas, E., S. Shackelford, J. Keele, R. Stone, S. Kappes, and M. Koohmaraie. 2000. Quantitative trait loci affecting growth and carcass composition of cattle segregating alternate forms of myostatin. Journal of Animal Science. 78(3):560-569.

Casas, E., S. Shackelford, J. Keele, R. Stone, S. Kappes, and M. Koohmaraie. 2000. Quantitative trait loci affecting growth and carcass composition of cattle segregating alternate forms of myostatin. Journal of Animal Science. 78(3):560-569.

Chen, G. K., and J. S. Witte. 2007. Enriching the analysis of genomewide association studies with hierarchical modeling. The American Journal of Human Genetics 81(2):397-404.

Chen, Y., C. Gondro, K. Quinn, R. Herd, P. Parnell, and B. Vanselow. 2011. Global gene expression profiling reveals genes expressed differentially in cattle with high and low residual feed intake. Animal Genetics. 42(5):475-490.

Cole, J. B., G. R. Wiggans, L. Ma, T. S. Sonstegard, T. J. Lawlor, B. A. Crooker, C. P. Van Tassell, J. Yang, S. Wang, L. K. Matukumalli, and Y. Da. 2011. Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. BMC Genomics 12(1):408.

Colwell, R. K. 2000. Rensch's rule crosses the line: convergent allometry of sexual size dimorphism in hummingbirds and flower mites. The American Naturalist 156(5):495-510.

Connolly, S. M., A. R. Cromie, R. D. Sleator, and D. P. Berry. 2019. Predicted carcass meat yield and primal cut yields in cattle divergent in genetic merit for a terminal index. Translational Animal Science 3(1):1-3.

Conroy, S., M. Drennan, D. Kenny, and M. McGee. 2010. The relationship of various muscular and skeletal scores and ultrasound measurements in the live animal, and carcass classification scores with carcass composition and value of bulls. Livestock Science 127(1):11-21.

Crews Jr, D., and R. Kemp. 2001. Genetic parameters for ultrasound and carcass measures of yield and quality among replacement and slaughter beef cattle. Journal of Animal Science. 79(12):3008-3020.

Cromie, A. R., R. D. Evans, J. F. Kearney, M. McClure, J. McCarthy, and D. P. Berry. 2016. Multi-breed Genomic Evaluations for 1 million Beef Cattle in Ireland. In: American Society of Animal Science 2016 Joint Annual Meeting, Salt Lake City, UT.

Cross, H., Z. Carpenter, and G. Smith. 1973. Effects of intramuscular collagen and elastin on bovine muscle tenderness. Journal of Food Science. 38(6):998-1003.

Crowley, J., M. McGee, D. Kenny, D. Crews Jr, R. Evans, and D. Berry. 2010. Phenotypic and genetic parameters for different measures of feed efficiency in different breeds of Irish performance-tested beef bulls. Journal of Animal Science. 88(3):885-894.

Crowley, J., R. Evans, N. Mc Hugh, T. Pabiou, D. Kenny, M. McGee, D. Crews Jr, and D. Berry. 2011. Genetic associations between feed efficiency measured in a performance test station and performance of growing cattle in commercial beef herds. Journal of Animal Science. 89(11):3382-3393.

CSO. 2019. Livestock Survey December 2018. In: C. S. Office (ed.) https://www.cso.ie/en/releasesandpublications/er/lsd/livestocksurveydecember2018/.

Daetwyler, H. D., A. Capitan, H. Pausch, P. Stothard, R. Van Binsbergen, R. F. Brøndum, X. Liao, A. Djari, S. C. Rodriguez, and C. Grohs. 2014. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nature Genetics. 46(8):858.

DAFM. 2019. Annual review and outlook for agriculture, food and the marine. Department of Agriculture, Food, and the Marine (ed.).

Das, S., L. Forer, S. Schönherr, C. Sidore, A. E. Locke, A. Kwong, S. I. Vrieze, E. Y. Chew, S. Levy, and M. McGue. 2016. Next-generation genotype imputation service and methods. Nature Genetics. 48(10):1284.

Delpretti, S., J. Zakany, and D. Duboule. 2012. A function for all posterior Hoxd genes during digit development? Developmental Dynamics. 241(4):792-802.

Edea, Z., Y. H. Jeoung, S.-S. Shin, J. Ku, S. Seo, I.-H. Kim, S.-W. Kim, and K.-S. Kim. 2018. Genome–wide association study of carcass weight in commercial Hanwoo cattle. Asian-Australasian Journal of Animal Sciences. 31(3):327.

El-Saied, U., L. De la Fuente, R. Rodríguez, and F. San Primitivo. 2006. Genetic parameter estimates for birth and weaning weights, pre-weaning daily weight gain and three type traits for Charolais beef cattle in Spain. Spanish Journal of Agricultural Research. 4: 146-155.

Esmailizadeh, A., C. Bottema, G. Sellick, A. Verbyla, C. Morris, N. Cullen, and W. Pitchford. 2008. Effects of the myostatin F94L substitution on beef traits. Journal of Animal Science. 86(5):1038-1046.

Evans, R. D., T. Pabiou, A. R. Cromie, F. Kearney, and B. W. Wickham. 2007. Genetic improvement in the Irish suckler beef herd: industry expectation and experience so far. Interbull Bulletin 36.

Fairbairn, D. J., and D. A. Roff. 2006. The quantitative genetics of sexual dimorphism: assessing the importance of sex-linkage. Heredity 97(5):319-328.

Falconer, D. S. 1952. The problem of environment and selection. The American Naturalist 86(830):293-298.

Fan, B., Z.Q. Du, D. M. Gorbach, and M. F. Rothschild. 2010. Development and application of high-density SNP arrays in genomic studies of domestic animals. Asian-Australasian Journal of Animal Sciences 23(7):833-847.

Fatehi, J., A. Stella, J. J. Shannon, and P. J. Boettcher. 2003. Genetic parameters for feet and leg traits evaluated in different environments. Journal of Dairy Science. 86: 661-666.

Fernando, R.L. and M. Grossman. 1990. Genetic evaluation with autosomal and X-chromosomal inheritance. Theoretical and Applied Genetics. 80(1):75-80.

Fernando, R. L., and D. Garrick. 2013. Bayesian methods applied to GWAS, Genome-wide association studies and genomic prediction. Springer. p. 237-274.

Ferrell, C. L., and T. G. Jenkins. 1984. Cow type and the nutritional environment: Nutritional aspects. Journal of Animal Science. 61:725–741.

Fisher, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. Earth and Environmental Science Transactions of the Royal Society of Edinburgh 52(2):399-433.

Forabosco, F., R. Bozzi, P. Boettcher, F. Filippini, P. Bijma, and J. Van Arendonk. 2005. Relationship between profitability and type traits and derivation of economic values for reproduction and survival traits in Chianina beef cows. Journal of Animal Science. 83(9):2043-2051.

Garrick, D. J., and R. L. Fernando. 2013. Implementing a QTL detection study (GWAS) using genomic prediction methodology, Genome-wide association studies and genomic prediction. Springer. p. 275-298.

Gibson, J. R., A. K. Chippindale, and W. R. Rice. 2002. The X chromosome is a hot spot for sexually antagonistic fitness variation. Proceedings of the Royal Society of London. Series B: Biological Sciences 269: 499-505.

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. 2009. ASReml User Guide Release 3.0. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Goddard, M. E., and B. J. Hayes. 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nature Reviews Genetics 10(6):381-391.

Grade, C. V. C., M. S. Salerno, F. R. Schubert, S. Dietrich, and L. E. Alvares. 2009. An evolutionarily conserved Myostatin proximal promoter/enhancer confers basal levels of transcription and spatial specificity in vivo. Development Genes and Evolution volume.. 219(9-10):497-508.

Grobet, L., D. Poncelet, L. J. Royo, B. Brouwers, D. Pirottin, C. Michaux, F. Ménissier, M. Zanotti, S. Dunner, and M. Georges. 1998. Molecular definition of an allelic series of mutations disrupting the myostatin function and causing double-muscling in cattle. Mammalian Genome 9(3):210-213.

Grobet, L., L. J. R. Martin, D. Poncelet, D. Pirottin, B. Brouwers, J. Riquet, A. Schoeberlein, S. Dunner, F. Ménissier, and J. Massabanda. 1997. A deletion in the bovine myostatin gene causes the double–muscled phenotype in cattle. Nature Genetics. 17(1):71.

Grona, A., J. Tatum, K. Belk, G. Smith, and F. Williams. 2002. An evaluation of the USDA standards for feeder cattle frame size and muscle thickness. Journal of Animal Science. 80: 560-567.

Gudbjartsson, D. F., G. B. Walters, G. Thorleifsson, H. Stefansson, B. V. Halldorsson, P. Zusmanovich, P. Sulem, S. Thorlacius, A. Gylfason, and S. Steinberg. 2008. Many sequence variants affecting diversity of adult human height. Nature Genetics. 40(5):609.

Gutierrez, J. P., and F. Goyache. 2002. Estimation of genetic parameters of type traits in Asturiana de los Valles beef cattle breed. Journal of Animal Breeding and Genetics. 119(2):93-100.

Hamidi Hay, E., and A. Roberts. 2018. Genome-wide association study for carcass traits in a composite beef cattle breed. Livestock Science. 213, 35-43.

Han, Y., Y. Chen, Y. Liu, and X. Liu. 2017. Sequence variants of the LCORL gene and its association with growth and carcass traits in Qinchuan cattle in China. Journal of Genetics. 96(1):9-17.

Hayes, B., R. Fries, M. S. Lund, D. Boichard, P. Stothard, R. F. Veerkamp, C. Tassell, C. Anderson, B. Hulsegge, and B. Guldbrandtsen. 2012. 1000 bull genomes consortium project. In: Plant and Animal Genome Conference, San Diego, CA, USA. p W139.

Hayward, J. J., M. G. Castelhano, K. C. Oliveira, E. Corey, C. Balkman, T. L. Baxter, M. L. Casal, S. A. Center, M. Fang, and S. J. Garrison. 2016. Complex disease and phenotype mapping in the domestic dog. Nature communications 7(1):1-11.

Hazel, L. N. 1943. The genetic basis for constructing selection indexes. Genetics 28(6):476-490.

Henderson, C. R. 1949. Estimation of changes in herd environment. Journal of Dairy Science. 32(1):706715.

Henderson, C. R. 1950. Estimation of genetic parameters. In: Biometrics. p 186-187.

Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics:423-447.

Hennessy, T. 2018. The Economic and Societal Importance of the Irish Suckler Beef Sector, Department of Economics, University College Cork.

Hensen, K., C. Braem, J. Declercq, F. Van Dyck, M. Dewerchin, L. Fiette, C. Denef, and W. J. Van de Ven. 2004. Targeted disruption of the murine Plag1 proto-oncogene causes growth retardation and reduced fertility. Development, growth & differentiation 46(5):459-470.

Hickey, J. M., G. Gorjanc, R. K. Varshney, and C. Nettelblad. 2015. Imputation of single nucleotide polymorphism genotypes in biparental, backcross, and topcross populations with a hidden Markov model. Crop Science. 55(5):1934-1946.

Hickey, J. M., M. G. Keane, D. A. Kenny, A. R. Cromie, and R. F. Veerkamp. 2007. Genetic parameters for EUROP carcass traits within different groups of cattle in Ireland. Journal of Animal Science. 85: 314-321.

Hirschhorn, J. N., and M. J. Daly. 2005. Genome-wide association studies for common diseases and complex traits. Nature Reviews Genetics 6(2):95-108.

Hozé, C., M.-N. Fouilloux, E. Venot, F. Guillaume, R. Dassonneville, S. Fritz, V. Ducrocq, F. Phocas, D. Boichard, and P. Croiseau. 2013. High-density marker imputation accuracy in sixteen French cattle breeds. Genetics Selection Evolution. 45(1):33.

Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2008. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Research. 37(1):1-13.

Hudson, N. J., R. E. Lyons, A. Reverter, P. L. Greenwood, and B. P. Dalrymple. 2013. Inferring the in vivo cellular program of developing bovine skeletal muscle from expression data. Gene Expression Patterns 13(3-4):109-125.

ICBF. 2013. Euro-Star Index & latest breeding developments. In: ICBF/Teagasc/AHI Suckler Event

ICBF. 2019. Genetic Indexes in Beef Cattle. In: World Charolais Technical Congress

ICBF. 2020. Beef Evaluations Supporting Document.

Illumina. 2011. GoldenGate® Bovine3K Genotyping BeadChip.

Inoue-Murayama, M., Y. Sugimoto, Y. Niimi, and H. Aso. 2000. Type XVIII collagen is newly transcribed during bovine adipogenesis. Differentiation 65(5):281-285.

International Hapmap Consortium. 2003. The international HapMap project. Nature 426(6968):789.

Israel, C., and J. Weller. 1998. Estimation of candidate gene effects in dairy cattle populations. Journal of Dairy Science. 81(6):1653-1662.

Israel, C., and J. Weller. 2002. Estimation of quantitative trait loci effects in dairy cattle populations. Journal of Dairy Science. 85(5):1285-1297.

Jia, P., L. Wang, H. Y. Meltzer, and Z. Zhao. 2011. Pathway-based analysis of GWAS datasets: effective but caution required. International Journal of Neuropsychopharmacology 14(4):567-572.

Judge, M., D. Purfield, R. Sleator, and D. Berry. 2017. The impact of multi-generational genotype imputation strategies on imputation accuracy and subsequent genomic predictions. Journal of Animal Science. 95(4):1489-1501.

Judge, M., J. Kearney, M. McClure, R. Sleator, and D. Berry. 2016. Evaluation of developed low-density genotype panels for imputation to higher density in independent dairy and beef cattle populations. Journal of Animal Science. 94(3):949-962.

Judge, M., T. Pabiou, J. Murphy, S. Conroy, P. Hegarty, and D. Berry. 2019. Potential exists to change, through breeding, the yield of individual primal carcass cuts in cattle without increasing overall carcass weight. Journal of Animal Science. 97(7), 2769-2779.

Kachman, S. D., M. L. Spangler, G. L. Bennett, K. J. Hanford, L. A. Kuehn, W. M. Snelling, R. M. Thallman, M. Saatchi, D. J. Garrick, R. D. Schnabel, J. F. Taylor, and E. J. Pollak. 2013. Comparison of molecular breeding values based on within- and across-breed training in beef cattle. Genetics Selection Evolution. 45:9.

Karim, L., H. Takeda, L. Lin, T. Druet, J. A. Arias, D. Baurain, N. Cambisano, S. R. Davis, F. Farnir, and B. Grisart. 2011. Variants modulating the expression of a chromosome domain encompassing PLAG1 influence bovine stature. Nature Genetics. 43(5):405.

Karisa, B., J. Thomson, Z. Wang, H. Bruce, G. Plastow, Moore, and SS. 2013. Candidate genes and biological pathways associated with carcass quality traits in beef cattle. Canadian Journal of Animal Science 93(3):295-306.

Kärst, S., R. Cheng, A. O. Schmitt, H. Yang, F. P. M. De Villena, A. A. Palmer, and G. A. Brockmann. 2011. Genetic determinants for intramuscular fat content and water-

holding capacity in mice selected for high muscle mass. Mammalian Genome 22(9-10):530.

Katz, L. S. 2008. Variation in male sexual behavior. Animal Reproduction Science. 105(1):64-71.

Kause, A., L. Mikkola, I. Stranden, and K. Sirkko. 2015. Genetic parameters for carcass weight, conformation and fat in five beef cattle breeds. Animal. 9: 35-42.

Kelleher, M. M., D. P. Berry, J. F. Kearney, S. McParland, F. Buckley, and D. C. Purfield. 2016. Inference of population structure of purebred dairy and beef cattle using high-density genotype data. Animal 11(1):15-23.

Kern, E. L., J. A. Cobuci, C. N. Costa, C. M. McManus, and J. Braccini. 2015. Genetic association between longevity and linear type traits of Holstein cows. Scientia Agricola 72(3):203-209.

Khatri, P., M. Sirota, and A. J. Butte. 2012. Ten years of pathway analysis: current approaches and outstanding challenges. PLoS Computational Biology. 8(2)

Kirkpatrick, M., 1987. Sexual selection by female choice in polygynous animals. Annual Review of Ecology and Systematics. 18: 43-70.

Klein, R. J., C. Zeiss, E. Y. Chew, J.-Y. Tsai, R. S. Sackler, C. Haynes, A. K. Henning, J. P. SanGiovanni, S. M. Mane, and S. T. Mayne. 2005. Complement factor H polymorphism in age-related macular degeneration. Science 308(5720):385-389.

Koch, R. M., and R. Clark. 1955. Influence of sex, season of birth and age of dam on economic traits in range beef cattle. Journal of Animal Science.14(2):386-397.

Koots, K. R., and J. P. Gibson. 1996. Realized sampling variances of estimates of genetic parameters and the difference between genetic and phenotypic correlations. Genetics. 143: 1409-1416.

Koufariotis, L., Y.-P. P. Chen, S. Bolormaa, and B. J. Hayes. 2014. Regulatory and coding genome regions are enriched for trait associated variants in dairy and beef cattle. BMC Genomics 15(1):436.

Li, Y., C. Willer, S. Sanna, and G. Abecasis. 2009. Genotype imputation. Annual Review of Genomics and Human Genetics 10:387-406.

Lindholm-Perry, A. K., A. K. Sexten, L. A. Kuehn, T. P. Smith, D. A. King, S. D. Shackelford, T. L. Wheeler, C. L. Ferrell, T. G. Jenkins, and W. M. Snelling. 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 Genetics. 12(1):103.

Loh, P.-R., P. Danecek, P. F. Palamara, C. Fuchsberger, Y. A. Reshef, H. K. Finucane, S. Schoenherr, L. Forer, S. McCarthy, and G. R. Abecasis. 2016. Reference-based phasing using the Haplotype Reference Consortium panel. Nature Genetics. 48(11):1443.

Long, C. R. 1980. Crossbreeding for beef production: experimental results. Journal of Animal Science. 51(5):1197-1223.

Lu, D., S. Miller, M. Sargolzaei, M. Kelly, G. Vander Voort, T. Caldwell, Z. Wang, G. Plastow, and S. Moore. 2013. Genome-wide association analyses for growth and feed efficiency traits in beef cattle. Journal of Animal Science. 91(8):3612-3633.

Lucotte, E. A., R. Laurent, E. Heyer, L. Ségurel, and B. Toupance. 2016. Detection of allelic frequency differences between the sexes in humans: A signature of sexually antagonistic selection. Genome Biology and Evolution 8(5):1489-1500.

Lush, J. L. 1949. Heritability of quantitative characters in farm animals. Heritability of quantitative characters in farm animals.

Lyons, R. E., N. T. Loan, L. Dierens, M. R. Fortes, M. Kelly, S. S. McWilliam, Y. Li, R. J. Bunch, B. E. Harrison, and W. Barendse. 2014. Evidence for positive selection of taurine genes within a QTL region on chromosome X associated with testicular size in Australian Brahman cattle. BMC Genetics. 15(1):6.

Ma, L., J. R. O'Connell, P. M. VanRaden, B. Shen, A. Padhi, C. Sun, D. M. Bickhart, J. B. Cole, D. J. Null, and G. E. Liu. 2015. Cattle sex-specific recombination and genetic control from a large pedigree analysis. PLoS Genetics. 11(11).

MacLeod, I., P. Bowman, C. Vander Jagt, M. Haile-Mariam, K. Kemper, A. Chamberlain, C. Schrooten, B. Hayes, and M. Goddard. 2016. Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. BMC Genomics 17(1):144.

MacRae, C. A., and R. S. Vasan. 2011. Next-generation genome-wide association studies: time to focus on phenotype? American Heart Association. 334-336.

Maddox, J., and N. Cockett. 2007. An update on sheep and goat linkage maps and other genomic resources. Small ruminant research 70(1):4-20.

Mantovani, R., M. Cassandro, B. Contiero, A. Albera, and G. Bittante. 2010. Genetic evaluation of type traits in hypertrophic Piemontese cows. Journal of Animal Science. 88(11):3504-3512.

Mao, X., A. M. Johansson, G. Sahana, B. Guldbrandtsen, and D.-J. De Koning. 2016. Imputation of markers on the bovine X chromosome. Journal of Dairy Science. 99(9):7313-7318.

Marchini, J., and B. Howie. 2010. Genotype imputation for genome-wide association studies. Nature Reviews Genetics 11(7):499.

Marlowe, T. J., and J. A. Gaines. 1958. The influence of age, sex, and season of birth of calf, and age of dam on preweaning growth rate and type score of beef calves. Journal of Animal Science. 17(3):706-713.

Marshall, D. 1994. Breed differences and genetic parameters for body composition traits in beef cattle. Journal of Animal Science. 72: 2745-2755.

Matukumalli, L. K., C. T. Lawley, R. D. Schnabel, J. F. Taylor, M. F. Allan, M. P. Heaton, J. O'connell, S. S. Moore, T. P. Smith, and T. S. Sonstegard. 2009. Development and characterization of a high density SNP genotyping assay for cattle. PloS one 4(4):e5350.

Mazza, S., N. Guzzo, C. Sartori, D. P. Berry, and R. Mantovani. 2014. Genetic parameters for linear type traits in the Rendena dual-purpose breed. Journal of Animal Breeding and Genetics. 131(1):27-35.

Mc Hugh, N., A. Fahey, R. Evans, and D. P. Berry. 2010. Factors associated with selling price of cattle at livestock marts. Animal 4(8):1378-1389.

Mc Hugh, N., R. D. Evans, A. G. Fahey, and D. P. Berry. 2012. Animal muscularity and size are genetically correlated with animal live-weight and price. Livestock Science 144(1-2):11-19.

McClure, M., H. Ramey, M. Rolf, S. McKay, J. Decker, R. Chapple, J. Kim, T. Taxis, R. Weaber, and R. Schnabel. 2012. Genome-wide association analysis for quantitative trait loci influencing Warner–Bratzler shear force in five taurine cattle breeds. Animal. Genetics. 43(6):662-673.

McClure, M., N. Morsci, R. Schnabel, J. Kim, P. Yao, M. Rolf, S. McKay, S. Gregg, R. Chapple, and S. Northcutt. 2010. A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle. Anim. Genet. 41(6):597-607.

McLaren, W., L. Gil, S. E. Hunt, H. S. Riat, G. R. Ritchie, A. Thormann, P. Flicek, and F. Cunningham. 2016. The ensembl variant effect predictor. Genome biology 17(1):122.

McPherron, A. C., A. M. Lawler, and S.-J. Lee. 1997. Regulation of skeletal muscle mass in mice by a new TGF-p superfamily member. Nature 387(6628):83.

McPherron, A. C., and S.J. Lee. 1997. Double muscling in cattle due to mutations in the myostatin gene. Proceedings of the National Academy of Sciences 94(23):12457-12461.

McPherson, F. J., and P. J. Chenoweth. 2012. Mammalian sexual dimorphism. Animal reproduction science. 131(3):109-122.

Mendonça, F. S., M.D. MacNeil, W.S. Leal, R.C. Azambuja, P.F. Rodrigues, and F.F. Cardoso. 2019. Crossbreeding effects on growth and efficiency in beef cow–calf systems: evaluation of Angus, Caracu, Hereford and Nelore breed direct, maternal and heterosis effects. Translational Animal Science, 3(4), 1286-1295.

Miller, M., H. Cross, J. Crouse, and T. Jenkins. 1987. Effect of feed energy intake on collagen characteristics and muscle quality of mature cows. Meat Science 21(4):287-294.

Montano-Bermudez, M., M. Nielsen, and G. Deutscher. 1990. Energy requirements for maintenance of crossbred beef cattle with different genetic potential for milk. Journal of Animal Science. 68(8):2279-2288.

Mrode, R. A. 2014. Linear models for the prediction of animal breeding values. Cabi.

Mukai, F., K. Oyama, and S. Kohno. 1995. Genetic relationships between performance test traits and field carcass traits in Japanese Black cattle. Livestock Production Science 44(3):199-205.

Mulder, H., and P. Bijma. 2005. Effects of genotype× environment interaction on genetic gain in breeding programs. Journal of Animal Science. 83(1):49-61.

Mullen, M. P., M. C. McClure, J. F. Kearney, S. M. Waters, R. Weld, P. Flynn, C. J. Creevey, A. R. Cromie, and D. P. Berry. 2013. Development of a custom SNP chip for dairy and beef cattle breeding, parentage and research. Proceedings of the Interbull Meeting; August 23–25, 2013; Nantes, France

Nettelblad, C. 2012. Inferring haplotypes and parental genotypes in larger full sib-ships and other pedigrees with missing or erroneous genotype data. BMC Genetics. 13(1):85.

Nishimura, S., T. Watanabe, K. Mizoshita, K. Tatsuda, T. Fujita, N. Watanabe, Y. Sugimoto, and A. Takasuga. 2012. Genome-wide association study identified three major QTL for carcass weight including the PLAG1-CHCHD7 QTN for stature in Japanese Black cattle. BMC Genetics. 13(1):40.

Nkrumah, J., E. Sherman, C. Li, E. Marques, D. Crews Jr, R. Bartusiak, B. Murdoch, Z. Wang, J. Basarab, and S. Moore. 2007. Primary genome scan to identify putative quantitative trait loci for feedlot growth rate, feed intake, and feed efficiency of beef cattle. Journal of Animal Science. 85(12):3170-3181.

Ochsner, K., M. MacNeil, R. Lewis, and M. Spangler. 2017. Economic selection index development for Beefmaster cattle I: Terminal breeding objective. Journal of Animal Science. 95(3):1063-1070.

Ochsner, K.P, M. D. MacNeil, R.M. Lewis, and M.L. Spangler. 2017. Economic selection index development for Beefmaster cattle II: General-purpose breeding objective. Journal of Animal Science. 95:1913–1920

O'Rourke, B., P. Greenwood, P. Arthur, and M. Goddard. 2012. Inferring the recent ancestry of myostatin alleles affecting muscle mass in cattle. Animal Genetics. 44(1):86-90.

Pabiou, T., W. Fikse, A. Näsholm, A. Cromie, M. Drennan, M. Keane, and D. Berry. 2009. Genetic parameters for carcass cut weight in Irish beef cattle. Journal of Animal Science. 87: 3865.

Pabiou, T., W. Fikse, P. Amer, A. Cromie, A. Näsholm, and D. P. Berry. 2012. Genetic relationships between carcass cut weights predicted from video image analysis and other performance traits in cattle. Animal 6(9):1389-1397.

Pearson, T. A., and T. A. Manolio. 2008. How to interpret a genome-wide association study. Jama 299(11):1335-1344.

Pérez-Cabal, M. A., and R. Alenda. 2002. Genetic relationships between lifetime profit and type traits in Spanish Holstein cows. Journal of Dairy Science. 85(12):3480-3491.

Peters, S., K. Kizilkaya, D. Garrick, R. Fernando, J. Reecy, R. Weaber, G. Silver, and M. Thomas. 2012. Bayesian genome-wide association analysis of growth and yearling ultrasound measures of carcass traits in Brangus heifers. Journal of Animal Science. 90(10):3398-3409.

Phocas, F., and D. Laloë. 2004. Genetic parameters for birth and weaning traits in French specialized beef cattle breeds. Livestock Production Science. 89: 121-128.

Pointer, M. A., P. W. Harrison, A. E. Wright, and J. E. Mank. 2013. Masculinization of gene expression is associated with exaggeration of male sexual dimorphism. PLoS Genetics. 9(8).

Poissant, J., J. T. Hogg, C. S. Davis, J. M. Miller, J. F. Maddox, and D. W. Coltman. 2010. Genetic linkage map of a wild genome: genomic structure, recombination and sexual dimorphism in bighorn sheep. BMC Genomics 11(1):524.

Polák, J., and D. Frynta. 2010. Patterns of sexual size dimorphism in cattle breeds support Rensch's rule. Evolutionary Ecology . 24(5):1255-1266.

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(3):981-984.

Purfield, D. C., D. G. Bradley, R. D. Evans, F. J. Kearney, and D. P. Berry. 2015. Genome-wide association study for calving performance using high-density genotypes in dairy and beef cattle. Genetics Selection Evolution. 47(1):47.

Purfield, D. C., R. D. Evans, and D. P. Berry. 2019. Reaffirmation of known major genes and the identification of novel candidate genes associated with carcass-related metrics based on whole genome sequence within a large multi-breed cattle population. BMC Genomics 20(1), 720.

Rabiner, L. R. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE 77(2):257-286.

Randall, J. C., T. W. Winkler, Z. Kutalik, S. I. Berndt, A. U. Jackson, K. L. Monda, T. O. Kilpeläinen, T. Esko, R. Mägi, and S. Li. 2013. Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. PLoS Genetics. 9(6):e1003500.

Rensch, B. 1959. Evolution above the species level. Columbia University Press.

Rice, W. R. 1984. Sex chromosomes and the evolution of sexual dimorphism. Evolution:735-742.

Ring, S., A. Twomey, N. Byrne, M. Kelleher, T. Pabiou, M. Doherty, and D. Berry. 2018. Genetic selection for hoof health traits and cow mobility scores can accelerate the rate of genetic gain in producer-scored lameness in dairy cows. Journal of Dairy Science. 101(11):10034-10047.

Robertson, A. 1959. The sampling variance of the genetic correlation coefficient. Biometrics 15(3):469-485.

Royal, M. D., J. E. Pryce, J. A. Woolliams, and A. P. F. Flint. 2002. The genetic relationship between commencement of luteal activity and calving interval, body condition score, production, and linear type traits in Holstein-Friesian dairy cattle. Journal of Dairy Science. 85(11):3071-3080.

Saatchi, M., J. E. Beever, J. E. Decker, D. B. Faulkner, H. C. Freetly, S. L. Hansen, H. Yampara-Iquise, K. A. Johnson, S. D. Kachman, M. S. Kerley, J. Kim, D. D. Loy, E. Marques, H. L. Neibergs, E. J. Pollak, R. D. Schnabel, C. M. Seabury, D. W. Shike, W. M. Snelling, M. L. Spangler, R. L. Weaber, D. J. Garrick, and J. F. Taylor. 2014a. QTLs associated with dry matter intake, metabolic mid-test weight, growth and feed efficiency have little overlap across 4 beef cattle studies. BMC Genomics 15(1):1004.

Saatchi, M., R. D. Schnabel, J. F. Taylor, and D. J. Garrick. 2014b. Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds. BMC Genomics 15(1):442.

Sahana, G., B. Guldbrandtsen, and M. S. Lund. 2011. Genome-wide association study for calving traits in Danish and Swedish Holstein cattle. Journal of Dairy Science. 94(1):479-486.

Sahana, G., J. K. Höglund, B. Guldbrandtsen, and M. S. Lund. 2015. Loci associated with adult stature also affect calf birth survival in cattle. BMC Genetics. 16(1):47.

Saifl, G. M. and H. S. Chandra (1999). An apparent excess of sex–and reproduction–related genes on the human X chromosome. Proceedings of the Royal Society of London. Series B: Biological Sciences 266(1415): 203-209.

Sanchez, M.-P., A. Govignon-Gion, P. Croiseau, S. Fritz, C. Hozé, G. Miranda, P. Martin, A. Barbat-Leterrier, R. Letaïef, D. Rocha, M. Brochard, M. Boussaha, and D. Boichard. 2017. Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. Genetics Selection. Evolution. 49(1):68.

Saowaphak, P., M. Duangjinda, S. Plaengkaeo, R. Suwannasing, and W. Boonkum. 2017. Genetic correlation and genome-wide association study (GWAS) of the length of productive life, days open, and 305-days milk yield in crossbred Holstein dairy cattle. Genetics and Molecular Research. 16(2).

Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15(1):478.

Schierding, W. S., W. S. Cutfield, and J. M. O'Sullivan. 2014. The missing story behind genome wide association studies: single nucleotide polymorphisms in gene deserts have a story to tell. Frontiers in genetics 5:39.

Schierding, W., J. Antony, W. S. Cutfield, J. A. Horsfield, and J. M. O'sullivan. 2016. Intergenic GWAS SNPs are key components of the spatial and regulatory network for human growth. Human Molecular Genetics 25(15):3372-3382.

Schneeberger, M., S. A. Barwick, G. H. Crow, and K. Hammond. 1992. Economic indices using breeding values predicted by BLUP. Journal of Animal Breeding and Genetics, 109(1-6):180-187.

Seo, M., K. Caetano-Anolles, S. Rodriguez-Zas, S. Ka, J. Y. Jeong, S. Park, M. J. Kim, W.-G. Nho, S. Cho, and H. Kim. 2016. Comprehensive identification of sexually dimorphic genes in diverse cattle tissues using RNA-seq. BMC Genomics 17(1):81.

Sharma, A., C. Dang, K. Kim, J. Kim, H. Lee, H. Kim, S. Yeon, H. Kang, and S. Lee. 2014. Validation of genetic polymorphisms on BTA 14 associated with carcass trait in a commercial H anwoo population. Animal Genetics. 45(6):863-867.

Sherman, E., J. Nkrumah, and S. Moore. 2010. Whole genome single nucleotide polymorphism associations with feed intake and feed efficiency in beef cattle. Journal of Animal Science. 88(1):16-22.

Sherman, E., J. Nkrumah, B. Murdoch, C. Li, Z. Wang, A. Fu, and S. Moore. 2008. Polymorphisms and haplotypes in the bovine neuropeptide Y, growth hormone receptor, ghrelin, insulin-like growth factor 2, and uncoupling proteins 2 and 3 genes and their associations with measures of growth, performance, feed efficiency, and carcass merit in beef cattle. Journal of Animal Science. 86(1):1-16.

Short, T., and T. Lawlor. 1992. Genetic parameters of conformation traits, milk yield, and herd life in Holsteins. Journal of Dairy Science. 75(7):1987-1998.

Sieber, M., A. Freeman, and D. Kelley. 1989. Effects of body measurements and weight on calf size and calving difficulty of Holsteins. Journal of Dairy Science. 72(9):2402-2410.

Snelling, W., M. Allan, J. Keele, L. Kuehn, T. Mcdaneld, T. Smith, T. Sonstegard, R. Thallman, and G. Bennett. 2010. Genome-wide association study of growth in crossbred beef cattle. Journal of Animal Science. 88(3):837-848.

Sollero, B. P., J. T. Howard, and M. L. Spangler. 2019. The impact of reducing the frequency of animals genotyped at higher density on imputation and prediction accuracies using ssGBLUP. Journal of Animal Science. 97(7):2780-2792.

Spangler, M., and A. Van Eenennaam. 2010. Utilizing molecular information in beef cattle selection. Beef Sire Selection Manual:79.

Stranger, B. E., E. A. Stahl, and T. Raj. 2011. Progress and promise of genome-wide association studies for human complex trait genetics. Genetics 187(2):367-383.

Su, G., B. Guldbrandtsen, G. P. Aamand, I. Strandén, and M. S. Lund. 2014. Genomic relationships based on X chromosome markers and accuracy of genomic predictions with and without X chromosome markers. Genetics Selection Evolution. 46(1):47.

Su, G., O. F. Christensen, T. Ostersen, M. Henryon, and M. S. Lund. 2012. Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers. PloS one 7(9)

Székely, T., R. P. Freckleton, and J. D. Reynolds. 2004. Sexual selection explains Rensch's rule of size dimorphism in shorebirds. Proceedings of the National Academy of Sciences 101(33):12224-12227.

Szkiba, D., M. Kapun, A. von Haeseler, and M. Gallach. 2014. SNP2GO: functional analysis of genome-wide association studies. Genetics 197(1):285-289.

Takasuga, A. 2016. PLAG1 and NCAPG-LCORL in livestock. Animal Science Journal 87(2):159-167.

Tam, V., N. Patel, M. Turcotte, Y. Bossé, G. Paré, and D. Meyre. 2019. Benefits and limitations of genome-wide association studies. Nature Reviews Genetics 20(8):467-484.

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. Animal Genetics. 44(4):467-471.

Trivers, R. 1988. Sex differences in rates of recombination and sexual selection. The evolution of sex 270:286.

Turner, S. D. 2014. qqman: an R package for visualizing GWAS results using QQ and manhattan plots. BioRxiv:005165.

Unnevehr, L. J., and S. Bard. 1993. Beef quality: will consumers pay for less fat? Journal of Agricultural and Resource Economics.:288-295.

Urciuolo, A., M. Quarta, V. Morbidoni, F. Gattazzo, S. Molon, P. Grumati, F. Montemurro, F. S. Tedesco, B. Blaauw, and G. Cossu. 2013. Collagen VI regulates satellite cell self-renewal and muscle regeneration. Nature Communications 4:1964.

Utrera, A. R., and L. D. Van Vleck. 2004. Heritability estimates for carcass traits of cattle: a review. Genetics Molecular Research. 3: 380-394.

Vallee, A., I. Breider, J. A. M. van Arendonk, and H. Bovenhuis. 2015. Genetic parameters for large-scale behavior traits and type traits in Charolais beef cows. Journal of Animal Science. 93(9):4277-4284.

Vallée, A., J. Daures, J. van Arendonk, and H. Bovenhuis. 2016. Genome-wide association study for behavior, type traits, and muscular development in Charolais beef cattle. Journal of Animal Science. 94(6):2307-2316.

van der Heide, E. M. M., D. A. L. Lourenco, C. Y. Chen, W. O. Herring, R. L. Sapp, D. W. Moser, S. Tsuruta, Y. Masuda, B. J. Ducro, and I. Misztal. 2016. Sexual dimorphism in livestock species selected for economically important traits. Journal of Animal Science. 94(9):3684-3692.

Van Laere, A.S., W. Coppieters, and M. Georges. 2008. Characterization of the bovine pseudoautosomal boundary: Documenting the evolutionary history of mammalian sex chromosomes. Genome Research. 18(12):1884-1895.

VanRaden, P. M., C. Sun, and J. R. O'Connell. 2015. Fast imputation using medium or low-coverage sequence data. BMC Genetics. 16(1):82.

VanRaden, P. M., E. L. Jensen, T. J. Lawlor, and D. A. Funk. 1990. Prediction of transmitting abilities for Holstein type traits. Journal of Dairy Science 73: 191-197.

Veerkamp, R. F., and S. Brotherstone. 1997. Genetic correlations between linear type traits, food intake, live weight and condition score in Holstein Friesian dairy cattle. Animal Science. 64(3):385-392.

Veneroni-Gouveia, G., S. Meirelles, D. Grossi, A. Santiago, T. Sonstegard, M. Yamagishi, L. Matukumalli, L. Coutinho, M. Alencar, and H. N. d. Oliveira. 2012. Whole-

genome analysis for backfat thickness in a tropically adapted, composite cattle breed from Brazil. Animal Genetics. 43(5):518-524.

Ventura, R. V., S. P. Miller, K. G. Dodds, B. Auvray, M. Lee, M. Bixley, S. M. Clarke, and J. C. McEwan. 2016. Assessing accuracy of imputation using different SNP panel densities in a multi-breed sheep population. Genetics Selection Evolution. 48(1):71.

Vesela, Z., J. Pribyl, P. Safus, L. Vostry, K. Seba, and L. Stolc. 2005. Breeding value for type traits in beef cattle in the Czech Republic. Czech Journal of Animal Science 50(9):385-393. (

Vignal, A., D. Milan, M. SanCristobal, and A. Eggen. 2002. A review on SNP and other types of molecular markers and their use in animal genetics. Genetics Selection Evolution. 34(3):275.

Visel, A., E. M. Rubin, and L. A. Pennacchio. 2009. Genomic views of distant-acting enhancers. Nature 461(7261):199.

Visscher, P. M., M. A. Brown, M. I. McCarthy, and J. Yang. 2012. Five years of GWAS discovery. The American Journal of Human Genetics 90(1):7-24.

Visscher, P. M., N. R. Wray, Q. Zhang, P. Sklar, M. I. McCarthy, M. A. Brown, and J. Yang. 2017. 10 years of GWAS discovery: biology, function, and translation. The American Journal of Human Genetics 101(1):5-22.

Wagenmakers, E.-J., M. Lee, T. Lodewyckx, and G. J. Iverson. 2008. Bayesian versus frequentist inference, Bayesian evaluation of informative hypotheses. Springer. p. 181-207.

Wall, E., I. White, M. Coffey, and S. Brotherstone. 2005. The relationship between fertility, rump angle, and selected type information in Holstein-Friesian cows. Journal of Dairy Science. 88(4):1521-1528.

Wang, S.-B., J.-Y. Feng, W.-L. Ren, B. Huang, L. Zhou, Y.-J. Wen, J. Zhang, J. M. Dunwell, S. Xu, and Y.-M. Zhang. 2016. Improving power and accuracy of genome-

wide association studies via a multi-locus mixed linear model methodology. Scientific reports 6:19444.

Wang, X., N. R. Tucker, G. Rizki, R. Mills, P. H. Krijger, E. de Wit, V. Subramanian, E. Bartell, X.-X. Nguyen, and J. Ye. 2016. Discovery and validation of sub-threshold genome-wide association study loci using epigenomic signatures. Elife 5:e10557.

Weisstein, E. W. 2004. Bonferroni correction. https://mathworld.wolfram com/.

Wiggans, G. R., N. Gengler, and J. R. Wright. 2004. Type trait (Co)variance components for five dairy breeds. J Journal of Dairy Science. 87(7):2324-2330.

Wiggans, G., P. VanRaden, L. Bacheller, M. Tooker, J. Hutchison, T. Cooper, and T. Sonstegard. 2010. Selection and management of DNA markers for use in genomic evaluation. Journal of Dairy Science. 93(5):2287-2292.

Wiggans, G., T. Sonstegard, P. VanRaden, L. Matukumalli, R. Schnabel, J. Taylor, F. Schenkel, and C. Van Tassell. 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. Journal of Dairy Science. 92(7):3431-3436.

Willer, C. J., Y. Li, and G. R. Abecasis. 2010. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26(17):2190-2191.

Wood, A. R., T. Esko, J. Yang, S. Vedantam, T. H. Pers, S. Gustafsson, A. Y. Chu, K. Estrada, J. a. Luan, and Z. Kutalik. 2014. Defining the role of common variation in the genomic and biological architecture of adult human height. Nature Genetics. 46(11):1173.

Wu, X., M. Fang, L. Liu, S. Wang, J. Liu, X. Ding, S. Zhang, Q. Zhang, Y. Zhang, L. Qiao, M. S. Lund, G. Su, and D. Sun. 2013. Genome wide association studies for body conformation traits in the Chinese Holstein cattle population. BMC Genomics 14(1):897.

Wyman, M. J., and M. C. Wyman. 2013. Sex-specific recombination rates and allele frequencies affect the invasion of sexually antagonistic variation on autosomes. Journal of Evolutionary Biology. 26(11):2428-2437.

Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A. Madden, A. C. Heath, N. G. Martin, and G. W. Montgomery. 2010. Common SNPs explain a large proportion of the heritability for human height. Nature Genetics. 42(7):565.

Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher. 2011. GCTA: A Tool for genome-wide complex trait analysis. The American Journal of Human Genetics. 88(1):76-82.

Yang, J., T. Ferreira, A. P. Morris, S. E. Medland, P. A. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, M. N. Weedon, and R. J. Loos. 2012. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. Nature Genetics. 44(4):369.

Young, K. A. 2005. Life–history variation and allometry for sexual size dimorphism in Pacific salmon and trout. Proceedings of the Royal Society B: Biological Sciences 272(1559):167-172.

Zakany, J., and D. Duboule. 2007. The role of Hox genes during vertebrate limb development. Current opinion in genetics & development,. 17(4):359-366.

Zerbino, D. R., P. Achuthan, W. Akanni, M. R. Amode, D. Barrell, J. Bhai, K. Billis, C. Cummins, A. Gall, and C. G. Girón. 2017. Ensembl 2018. Nucleic Acids Research. 46(D1):D754-D761.

Zhang, W., L. Xu, H. Gao, Y. Wu, X. Gao, L. Zhang, B. Zhu, Y. Song, J. Bao, J. Li, and Y. Chen. 2018. Detection of candidate genes for growth and carcass traits using genome-wide association strategy in Chinese Simmental beef cattle. Animal Production Science. 58(2):224-233.

Zhao, F., S. McParland, F. Kearney, L. Du, and D. P. Berry. 2015. Detection of selection signatures in dairy and beef cattle using high-density genomic information. Genetics Selection Evolution. 47(1):49.

Zhou, H.-Y., W.-D. Chen, D.-L. Zhu, L.-Y. Wu, J. Zhang, W.-Q. Han, J.-D. Li, C. Yan, and P.-J. Gao. 2010. The PDE1A-PKCα signaling pathway is involved in the upregulation of α-smooth muscle actin by TGF-β1 in adventitial fibroblasts. Journal of Vascular Research. 47(1):9-15.

Zimin, A. V., A. L. Delcher, L. Florea, D. R. Kelley, M. C. Schatz, D. Puiu, F. Hanrahan, G. Pertea, C. P. Van Tassell, and T. S. Sonstegard. 2009. A whole-genome assembly of the domestic cow, Bos taurus. Genome biology. 10(4):R42.

Chapter 8

Publications and Contributions from this Thesis

8.1 Peer Reviewed Publications

Doyle J.L., D.P. Berry, S.W. Walsh, R.F. Veerkamp, R.D. Evans, and T.R. Carthy, 2018. Genetic co-variance components within and among linear type traits differ among contrasting beef cattle breeds. Journal of Animal Science, 96(5):1628-1639.

Doyle J.L., D.P. Berry, R.F. Veerkamp, T.R. Carthy, R.D. Evans, S.W. Walsh, and D.C. Purfield, 2020. Genomic regions associated with muscularity in beef cattle differ in five contrasting cattle breeds. Genetic Selection Evolution, 52:2

Doyle J.L., D.P. Berry, R.F. Veerkamp, T.R. Carthy, R.D. Evans, S.W. Walsh, and D.C. Purfield, 2020. Genomic regions associated with skeletal type traits in beef and dairy cattle are common to regions associated with carcass traits, feed intake and calving difficulty. Frontiers in Genetics, 11:20.

Doyle J.L., D.C. Purfield, T. Moore, T.R. Carthy, S.W. Walsh, R.F. Veerkamp, R.D. Evans, and D.P. Berry. Identification of genomic regions that exhibit sexual dimorphism for size and muscularity in cattle. Journal of Animal Science. Submitted

Berry D.P., J.M. Coyne, **J.L. Doyle**, and R.D. Evans, 2020. The use of subjectively assessed muscular and skeletal traits on live cattle to aid in differentiation between animal genetically divergent in carcass kill out metrics. Livestock Science, 103984.

Costilla R., K.E. Kemper, E.M. Byrne, L.R. Porto-Neto, R. Carvalheiro, D.C. Purfield, **J.L. Doyle**, D.P. Berry, S.S. Moore, N.R. Wray, and B.J. Hayes, 2020. Genetic control of temperament traits across species: association of autism spectrum disorder risk genes with cattle temperament. Genetics Selection Evolution. 2020 Dec;52(1):1-4.

8.2 Conferences

Doyle J.L., T.R. Carthy, S.W. Walsh, R.D. Evans, and D.P. Berry, 2017. Heritability estimates of linear type traits in the Irish beef herd. WIT Research Day, Co. Waterford, Ireland

Doyle J.L., D.P. Berry, S.W. Walsh, R.F. Veerkamp, R.D. Evans, and T.R. Carthy, 2018. Genetic co-variance components within and among muscular, skeletal and functional traits differ among contrasting beef breeds. World Congress on Genetics Applied to Livestock Production, Auckland, New Zealand.

Doyle J.L., D.P. Berry, R.F. Veerkamp, and D.C. Purfield, 2018. Genes associated with development of loin in 5 beef cattle breeds. Irish Cattle Breeding Federation and Sheep Ireland Genetics conference, Athlone, Ireland.

Doyle J.L., D.P. Berry, R.F. Veerkamp, T.R. Carthy, S.W. Walsh and D.C. Purfield, 2019. Whole genome sequence GWAS reveals muscularity in beef cattle differs across five cattle breed. European Federation of Animal Science, Ghent, Belgium.

8.3 Industry Dissemination

Hurley A.M., D.T. Byrne, N. McHugh, **J.L Doyle,** J.M Coyne and D.P. Berry, 2017. Futuristic traits for inclusion in the EBI. Teagasc Moorepark Open Day.

Purfield, D.C., **J.L. Doyle**, J. Newtown and A. Blom, 2019. Advancements in genomic evaluations. Teagasc Moorepark Open Day.

8.4 Courses Attended

8.3.1 Generic Skills for Postgraduate Researchers, Waterford IT

- Good writing in Science and Engineering – 01 March 2017
- Academic Writing: Writing for Publications – 08 March 2017
- How to do the Lit Review – 29 November 2017
- Tranferring from Masters to PhD – 22 March 2018

8.3.2 Other Courses

Analysis and Interpretation of Experimental Data with Mathematical and Statistical Tools
University College Dublin (AFGDP course); 28th – 30th August 2017
Topics covered in this course included creating and managing data files, hypothesis testing, multi-variate data analysis using suitable parametric and non-parametric tests, linear modelling, and correlation and regression analysis. This course was conducted through lectures, exercises and self-study sessions. ***(5ects)***

Leadership skills for Agri-Food Researchers
Teagasc Oakpark (AFGDP course); 21st – 23rd November 2017
Topics covered in this course included developing the tools to improve leadership skills and team working skills, networking, effective communication, presenting, managing conflict, providing and receiving feedback as well as carrying out DiSC and Belbin analysis. This course was conducted through lectures, exercises and self-study sessions. ***(5ects)***

Science writing and Presenting skills for the agri-food researcher
University College Dublin (AFGDP course); 28th – 30th May 2018
This module covered topics such as: the scientific publication process; good scientific writing principles; getting published and making an impact; when to publish and when to patent; writing a research proposal; writing your thesis; ethics and publication; communicating with non-scientific audiences; writing a press release; communicating with Stakeholders, an oral presentation masterclass, using digital media platforms to promote your message. This course was conducted through lectures, exercises and self-study sessions. ***(5ects)***

Quantitative Genetics Applied in Animal Breeding
Mikkeli, Finland (NOVA course); 06th – 10th August 2018
This course covered the following topics: basics of quantitative traits and quantitative variation, genotypic variance and its components, the normal distribution and the infinitesimal model, variance components estimation and heritability, epistasis, relatedness among relatives, the relationship matrix, inbreeding and heterosis, The multivariate normal distribution, two-trait models, genetic and environmental correlation, genotype by environment interactions, norm of reaction, maternal effects, genetic drift, wright's F statistics, FST, FIS, effective population size, factors affecting Ne, linkage disequilibrium, haplotypes, the Bulmer effect, QTL mapping, genetic markers, and association studies. This course was conducted through lectures, exercises and self-study sessions. **(3 ects)**

Career Management
University College Cork (AFGDP course); 10th – 12th December 2018
Topics covered in this course included an introduction to career planning, SWOT analyses, networking, CV preparation, cover letter writing, interview skills, using LinkedIn effectively, mentorship, and personal development. This course was conducted through lectures, exercises and self-study sessions. ***(5ects)***

Laboratory Animal Science &Training (Rodent)
University College Dublin (LAST Ireland course); 12th – 13th November 2018
This course allowed me to become certified to work on an in-vivo mice trial. Topics studied included: the national legislation, ethics and animal welfare, animal care, health and management, recognition of pain, suffering and distress, humane methods of killing, design of procedures and projects, rodent biology and management. This course was conducted through lectures, exercises and an end of module assessment. ***(5ects)***

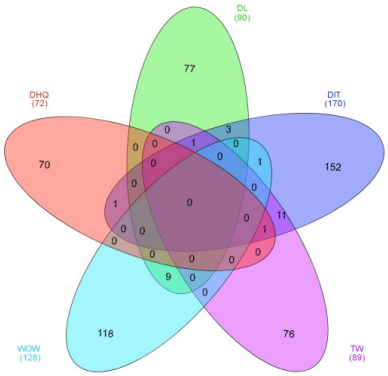
Appendix A

Genomic regions associated with muscularity in beef cattle differ in five contrasting cattle breeds

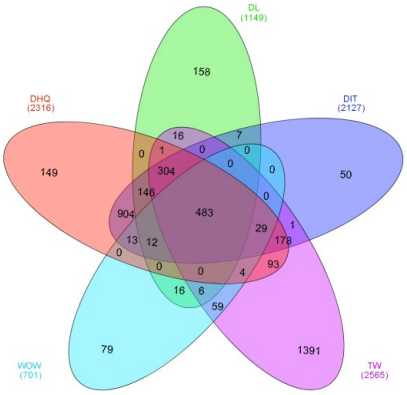
A1 The number of records, the mean, and the standard deviation of each linear type trait within each breed.

Trait	Scale 1-15	Angus			Charolais			Hereford			Limousin			Simmental		
		n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
Development of hindquarter	low - high	1444	7.71	1.28	6433	9.48	1.38	1129	7.35	1.34	8745	10.02	1.34	1698	9.31	1.39
Development of inner thigh	low - high	1434	7.47	1.43	6253	9.23	1.54	1128	7.07	1.51	8537	9.70	1.47	1619	9.02	1.46
Development of loin	low - high	1444	7.96	1.31	6433	9.61	1.52	1129	7.92	1.26	8745	9.62	1.47	1698	9.29	1.54
Thigh width	narrow - wide	1444	7.62	1.37	6433	9.31	1.57	1129	7.38	1.38	8745	9.27	1.62	1698	9.10	1.57
Width of withers	narrow - wide	1440	8.04	1.56	6412	9.54	1.63	1129	7.80	1.44	8710	9.61	1.54	1682	9.16	1.74

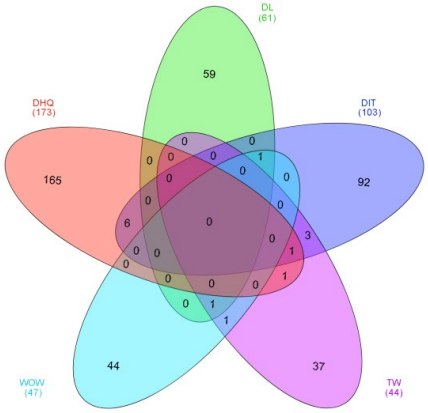
a)



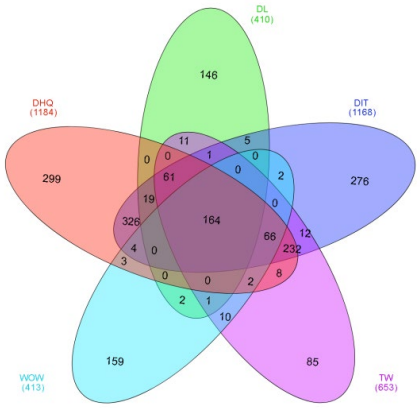
b)



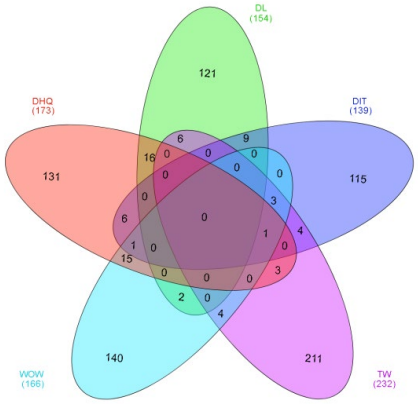
c)



d)



e)



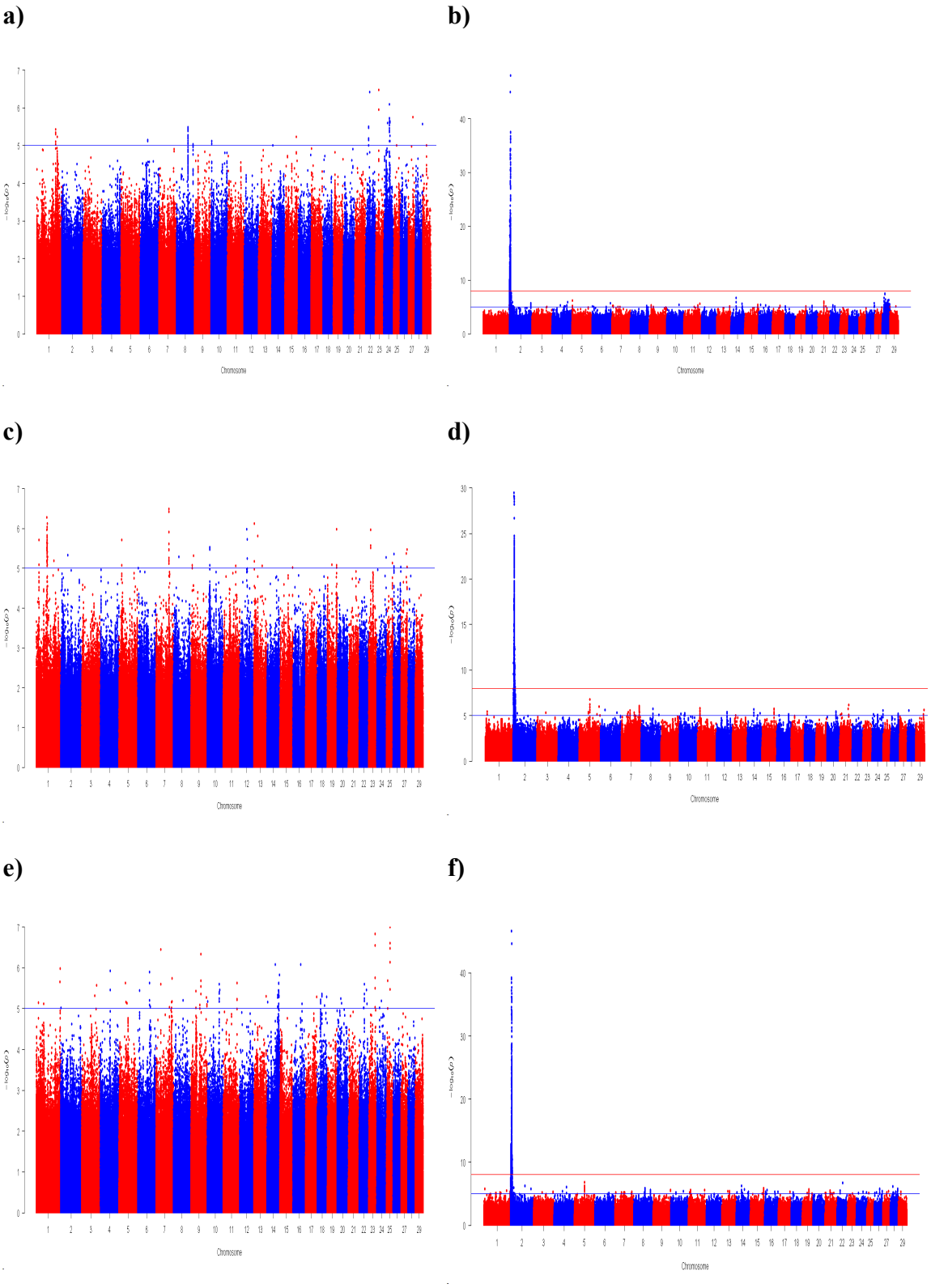
A2 Venn diagrams of overlapping 1kb regions that contain at least one suggestive or significant SNP for the 5 muscular traits in a) Angus, b) Charolais, c) Hereford, d) Limousin and e) Simmental

A3 The location of the top 5 most significant QTLs associated with each of the traits in the meta-analysis containing all 5 breeds

Trait	Chr	Start	End	No of suggestive and significant SNPs	Most significant SNP	P-Value	Candidate genes within this QTL
DHQ	2	35194	12618550	3545	6630781[a]	$2.06x10^{-47}$	WDR75^, ASNSD1^, ARHGEF4*, MYO7B^, NAB1^, MFSD6^, MSTN^, PMS1^, ORMDL1*, COL3A1^, COL5A2^, ANKAR^, SLC40A1
	4	73026098	74288232	3	73526098[b]	$8.36x10^{-7}$	ZNF804B*, TEX47
	5	59612855	60677712	6	60112855[a]	$1.46x10^{-7}$	AMDHD1, TESPA1, NTN4, SNRPF
	8	65991701	67324054	5	66811769[a]	$1.45x10^{-6}$	
	14	33353270	34360874	6	33853270[a]	$5.37x10^{-7}$	ARFGEF1, PREX2
DIT	2	156108	11096370	3358	6808074[a]	$2.22x10^{-44}$	WDR75^, ASNSD1^, ARHGEF4*, MYO7B^, NAB1^, MFSD6^, MSTN^, PMS1^, ORMDL1*, COL3A1^, COL5A2^, ANKAR^, SLC40A1^
	9	7672322	8675817	6	8173505[b]	$3.13x10^{-7}$	ADGRB3*
	13	7208015	7341939	5	7708015[a]	$1.85x10^{-7}$	TASP1
	13	46254146	47297509	37	46754146[a]	$4.36x10^{-7}$	IDI1, GTPBP4
	14	33063362	34360874	3	33853270[a]	$5.61x10^{-8}$	ARFGEF1, PREX2
TW	2	35194	10589284	1806	6808074[a]	$1.85x10^{-24}$	WDR75^, ASNSD1*, ARHGEF4, MYO7B*, NAB1*, MFSD6*, MSTN^, PMS1^, ORMDL1, COL3A1*, COL5A2*, ANKAR*, SLC40A1^
	11	94484006	96696774	8	96162620[b]	$1.09x10^{-8}$	LHX2, PSMB7, WDR38, GOLGA1
	13	74868306	75902585	94	75384004[b]	$1.20x10^{-7}$	DNTTIP1, TNNC2, ACOT8
	24	56615639	57674865	21	57144827[a]	$1.37x10^{-7}$	WDR7, FECH
	28	24372370	25568836	5	25068836[c]	$1.21x10^{-7}$	SIRT1, MYPN, DNA2, SLC25A16
DL	2	167014	9934770	845	6808074[a]	$5.51x10^{-24}$	WDR75^, ASNSD1*, ARHGEF4, MYO7B*, NAB1*, MFSD6*, MSTN^, PMS1^, ORMDL1, COL3A1^, COL5A2^, ANKAR^, SLC40A1^
	7	55126795	56162374	82	55642210[a]	$3.49x10^{-7}$	SPRY4, FGF1
	12	32803688	33961156	5	33461156[b]	$6.02x10^{-7}$	USP12, SHISA2
	16	79153910	80205278	3	79653910[a]	$7.65x10^{-7}$	ATP6V1G3, PTPRC
	19	36066854	37067996	6	36567996[a]	$2.22x10^{-8}$	TOB1, WFIKKN2, MYCBPAP
WOW	2	1889616	8714844	735	6727404[b]	$1.86x10^{-19}$	WDR75^, ASNSD1*, ARHGEF4, MYO7B*, NAB1, MFSD6, MSTN^, PMS1^, ORMDL1, COL3A1^, COL5A2^, ANKAR^, SLC40A1^
	6	19014612	20892119	10	19708641[a]	$1.63x10^{-8}$	NPNT, GSTCD
	7	55104801	56162374	90	55628366[a]	$3.40x10^{-7}$	SPRY4, FGF1*
	9	69711	2115412	3	618653[a]	$1.25x10^{-7}$	PTP4A1, PHF3
	24	49067824	50114190	28	49567824[b]	$3.78x10^{-7}$	SMAD7, ACAA2, MYO5B

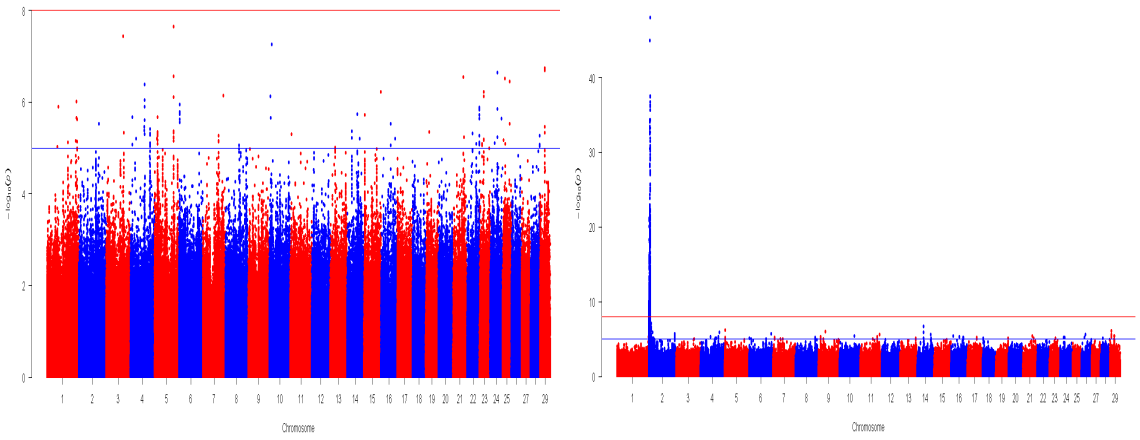
Superscript denotes SNP classification: [a]intergenic, [b]intron, [c]upstream gene variant. Symbols denote the significance of SNPs within genes: *gene contained at least one suggestive SNP, ^ gene contained at least one significant SNP.

A4 Manhattan plots for the muscular traits in all 5 breeds plus the meta analysis.

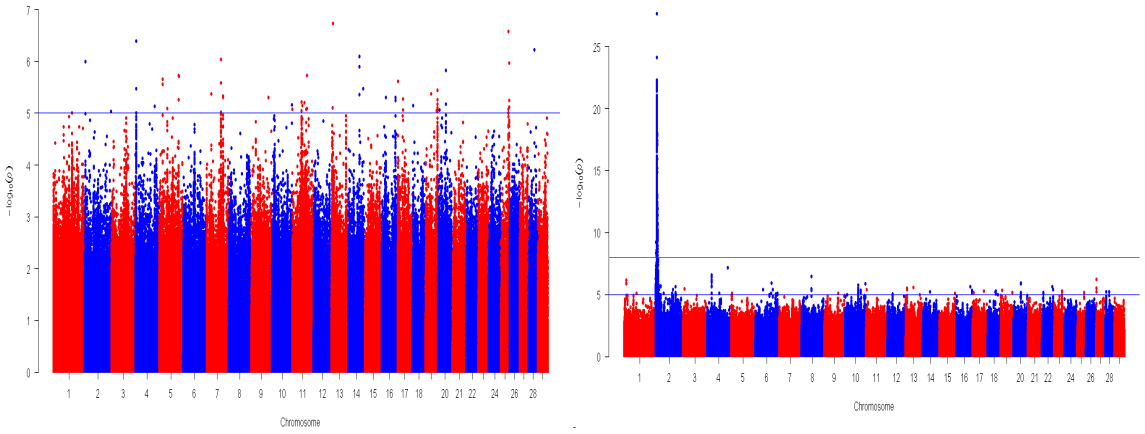


A4.1 Manhattan plots for development of hind quarters in a) Angus, b) Charolais, c) Hereford, d) Limousin, e) Simmental, and f) Meta-Analysis.

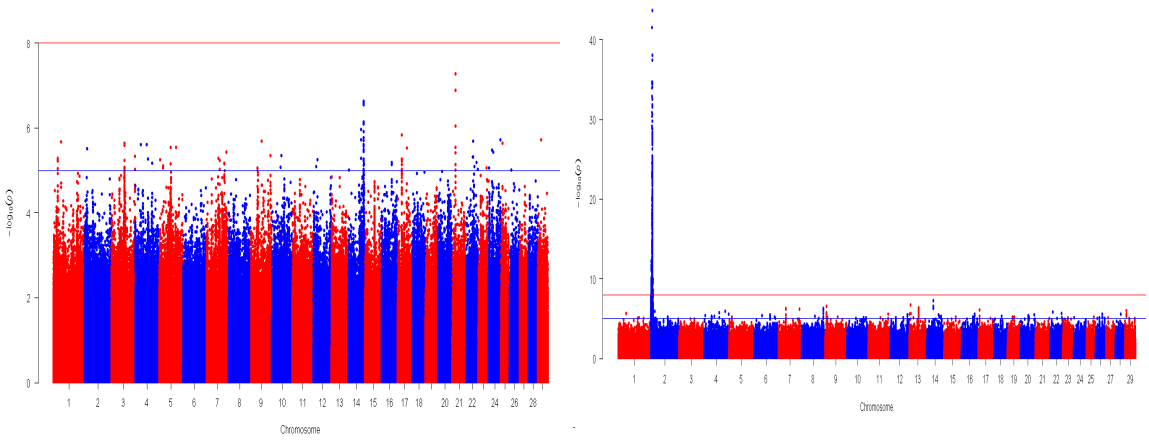
a) **b)**



c) **d)**

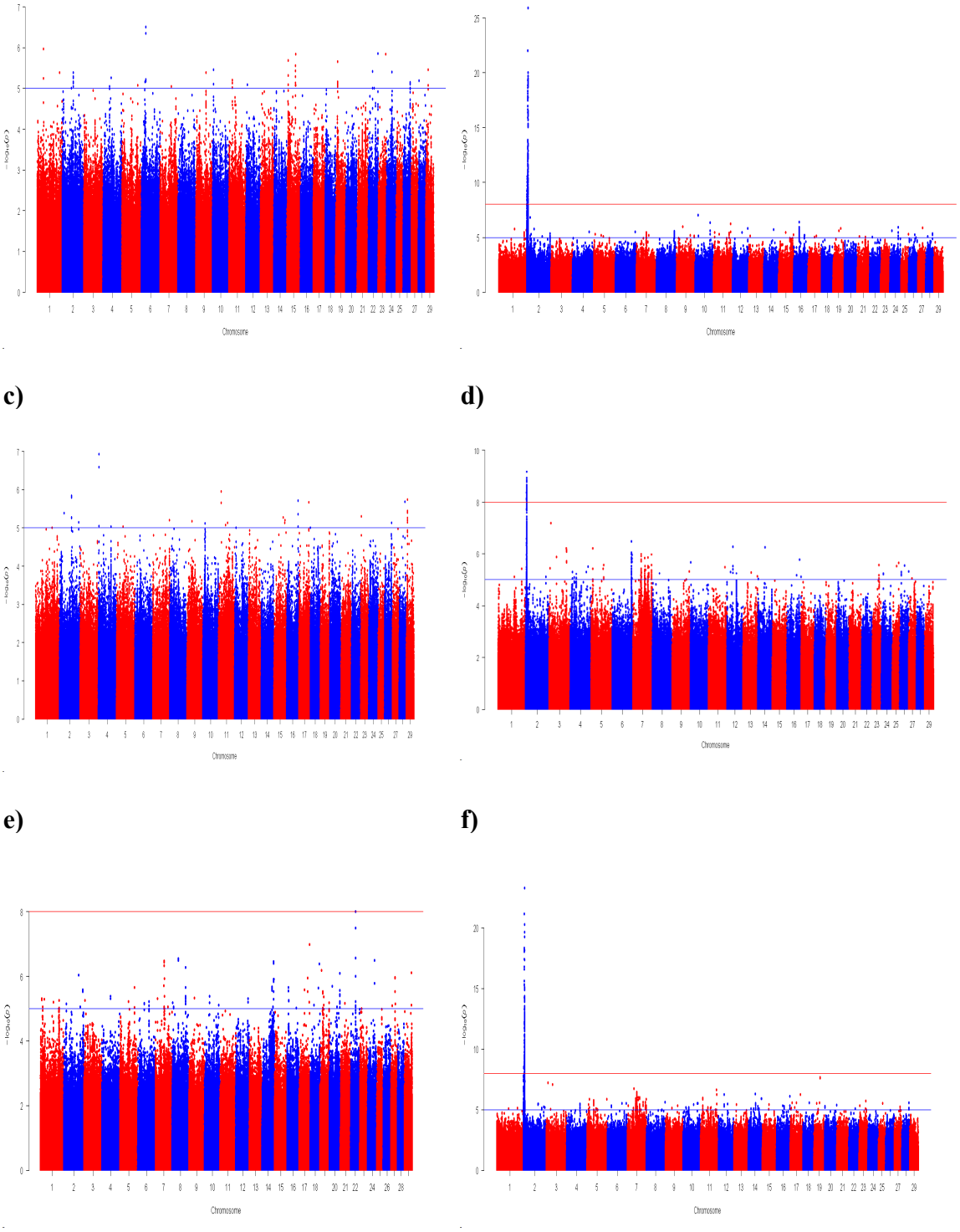


e) **f)**

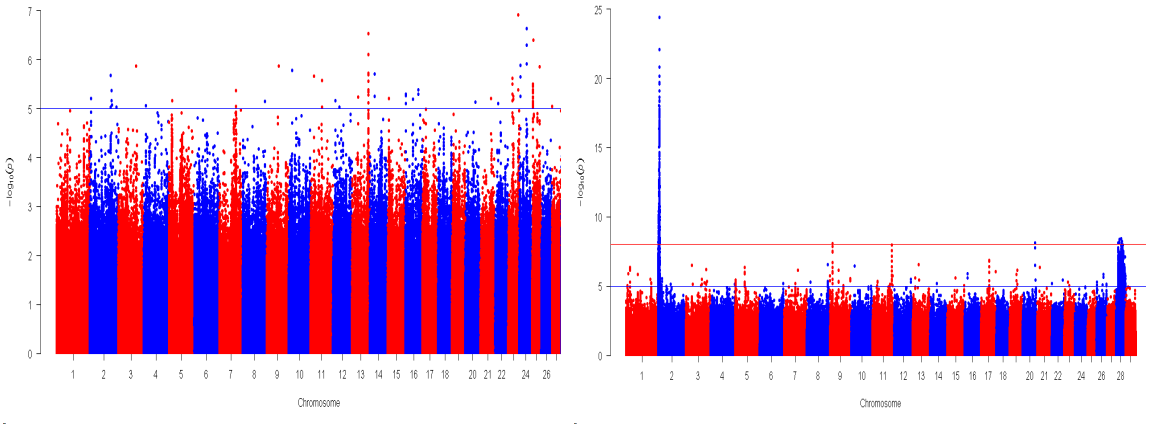


A4.2 Manhattan plots for development of inner thigh in a) Angus, b) Charolais, c) Hereford, d) Limousin, e) Simmental, and f) Meta-Analysis.

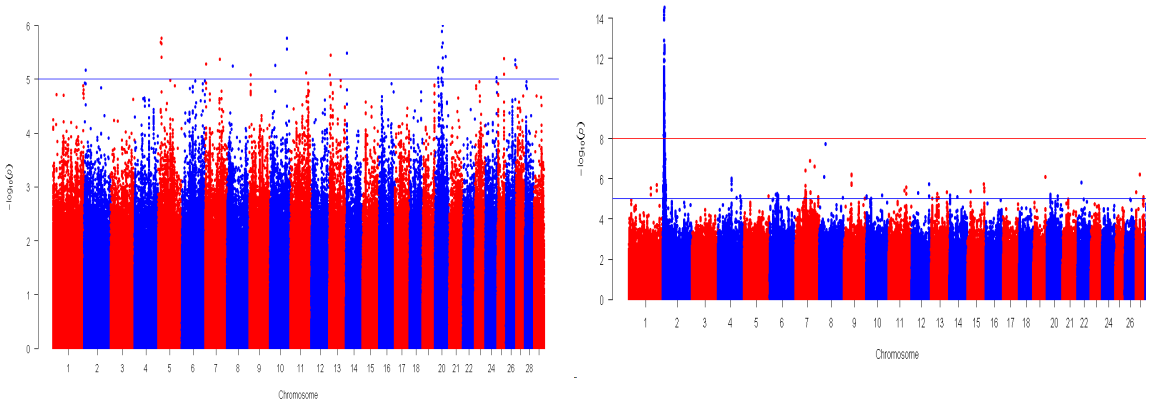
a) **b)**



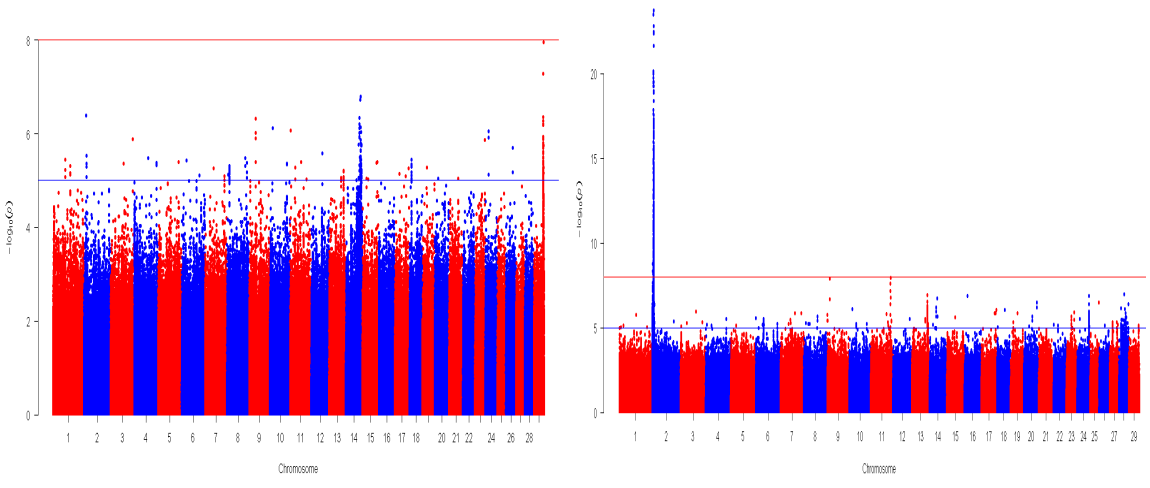
A4.3 Manhattan plots for development of loin in a) Angus, b) Charolais, c) Hereford, d) Limousin, e) Simmental, and f) Meta-Analysis.



c) **d)**

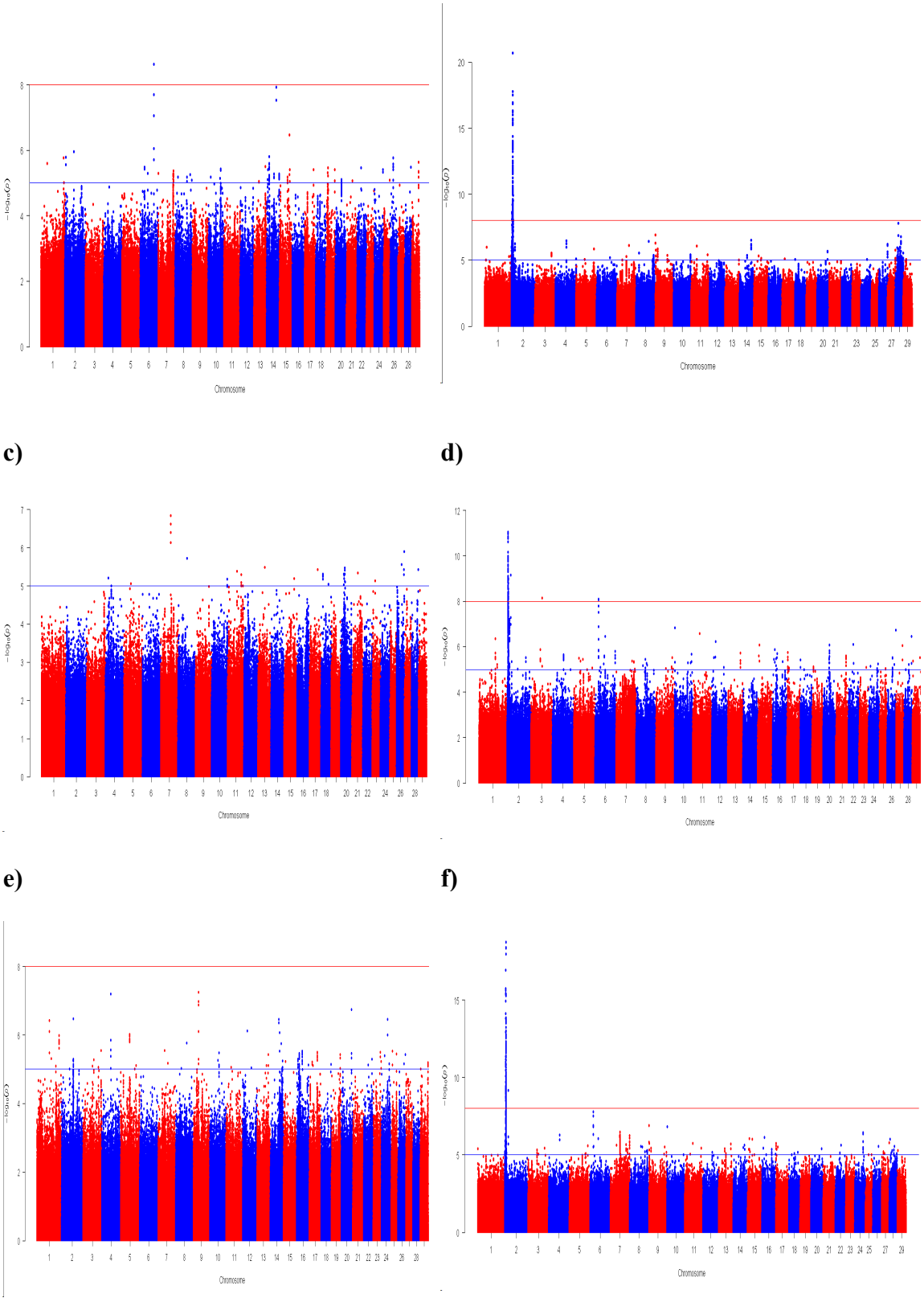


e) **f)**



A4.4 Manhattan plots for thigh width in a) Angus, b) Charolais, c) Hereford, d) Limousin, e) Simmental, and f) Meta-Analysis.

a) **b)**



A4.5 Manhattan plots for width of withers in a) Angus, b) Charolais, c) Hereford, d) Limousin, e) Simmental, and f) Meta-Analysis.

Appendix B

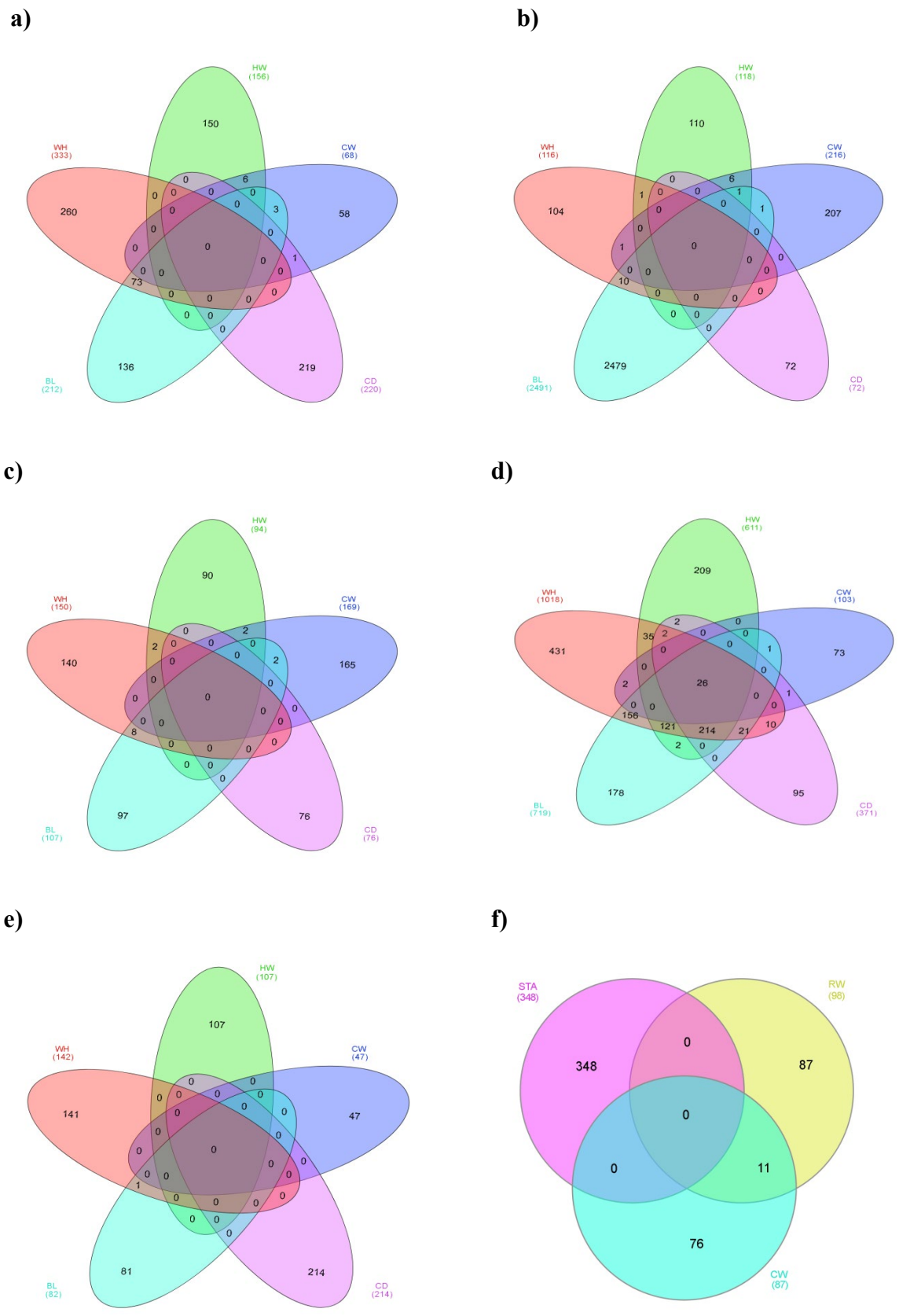
Genomic regions associated with skeletal type traits in beef and dairy cattle are common to regions associated with carcass traits, feed intake and calving difficulty

B1 The number of records, the mean, and the standard deviation of each linear type trait in each beef breed.

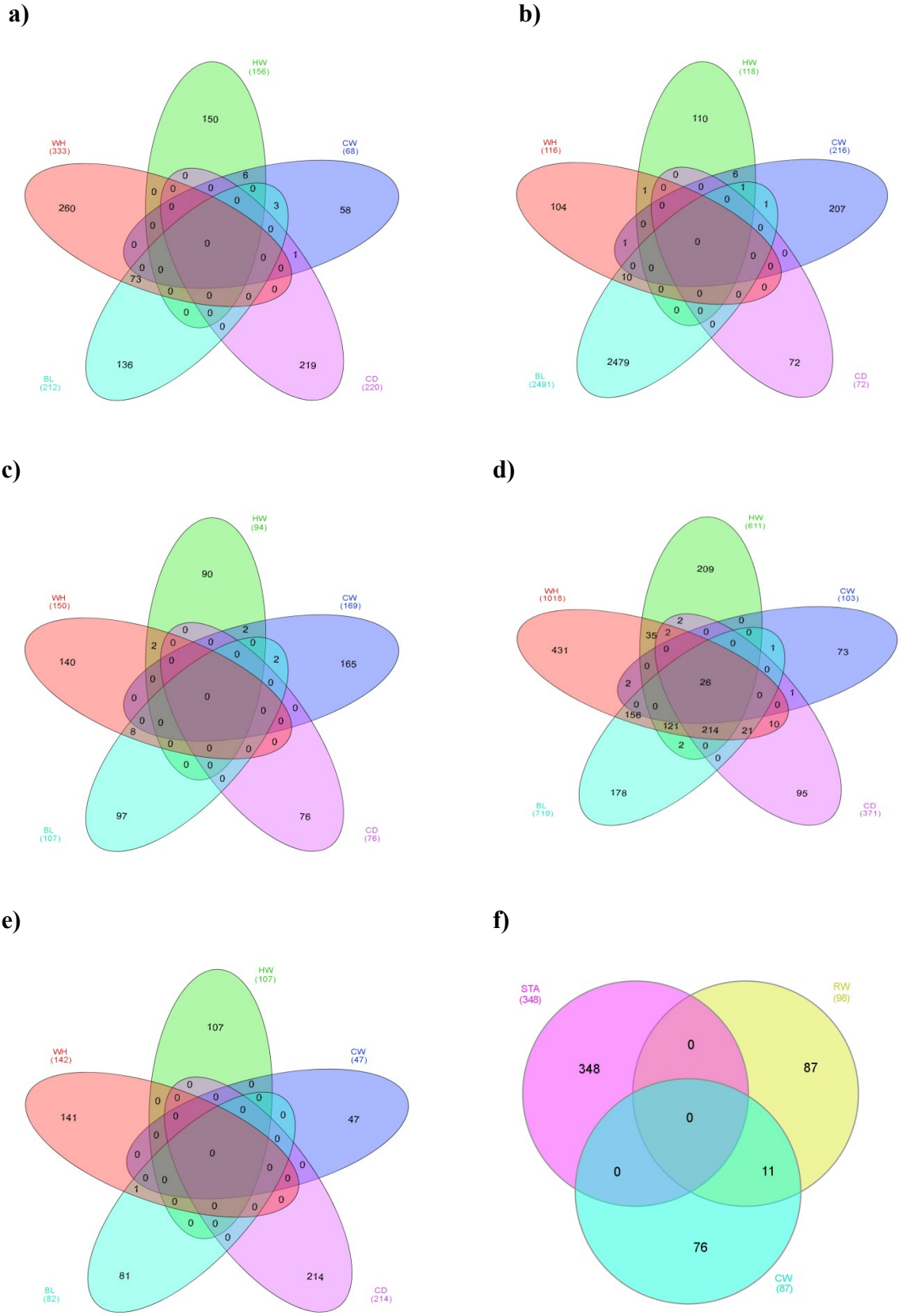
Trait	Scale 1-10	Angus			Charolais			Hereford			Limousin			Simmental		
		n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
Wither height	small - tall	1444	5.42	1.01	6432	5.98	1.13	1129	5.47	1.02	8745	5.74	1.09	1698	6.20	1.04
Chest width	narrow - wide	1434	5.63	0.96	6252	6.18	0.97	1128	5.64	0.88	8537	5.80	0.97	1619	6.06	0.95
Chest depth	shallow - deep	1433	6.35	0.90	6252	6.77	0.93	1128	6.36	0.88	8537	6.43	0.91	1619	6.84	0.87
Back length	short - long	1444	6.27	1.03	6432	6.75	1.07	1129	6.30	1.04	8745	6.59	1.08	1698	7.04	0.96
Hip width	narrow - wide	1444	5.46	0.95	6432	5.72	0.95	1129	5.66	0.89	8745	5.62	1.06	1698	5.96	0.95

B2 The number of records, the mean, and the standard deviation of each linear type trait in Holstein Friesian.

		Holstein Friesian		
Trait	Scale 1-9	n	Mean	SD
Stature	small- tall	4494	5.96	1.47
Chest width	narrow - wide	4494	5.15	1.48
Rump width	narrow - wide	4494	5.49	1.42



B3 Venn digrams of overlapping 1kb regions that contain at least one suggestive or significant SNP in each breed for a) wither height, b) hip width, c) chest width, d) chest depth, and e) back length.



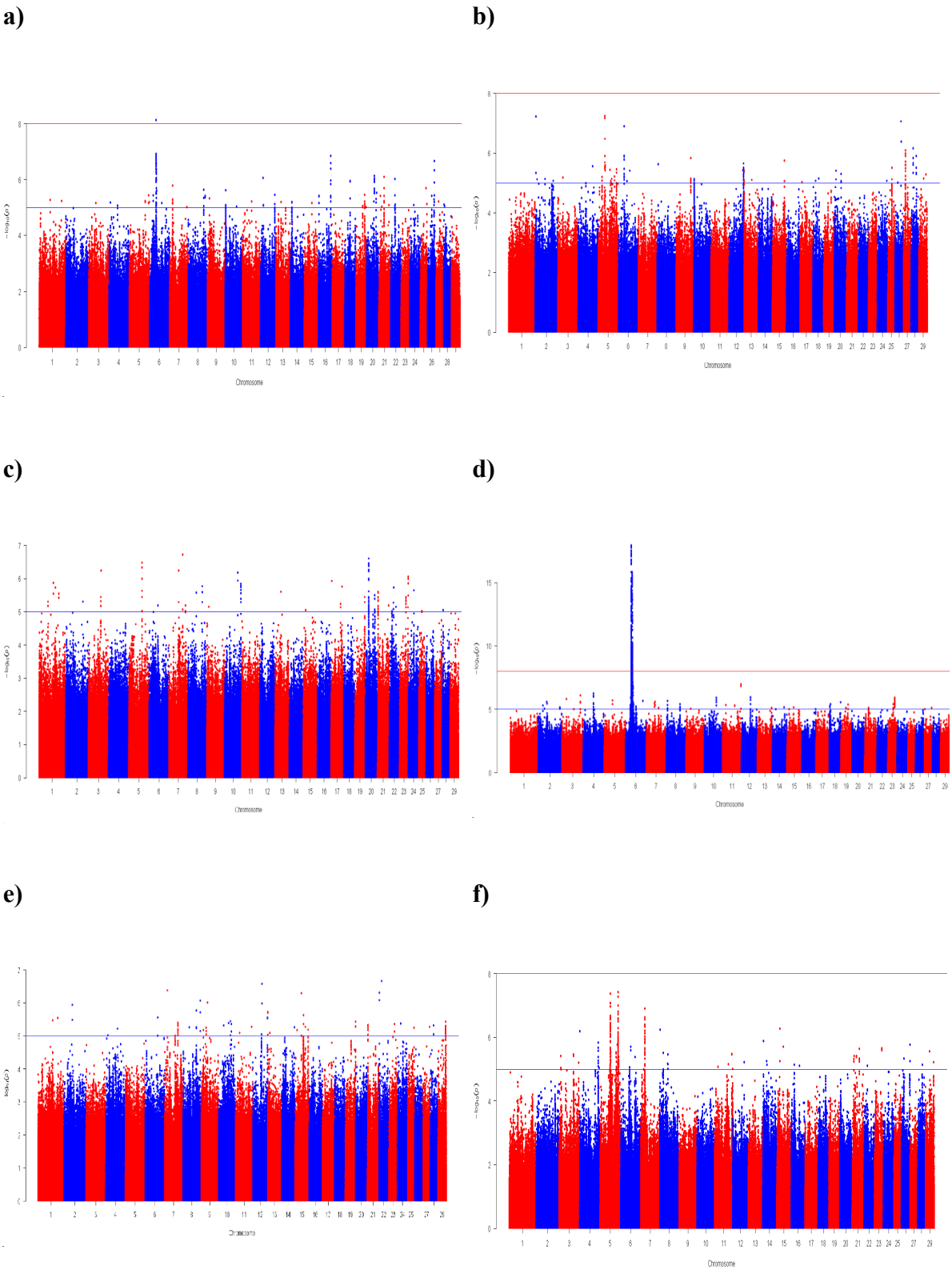
B4 Overlapping 1kb regions that contain at least one suggestive or significant SNP for the 5 muscular traits in a) Angus, b) Charolais, c) Hereford, d) Limousin, e) Simmental, and three muscular traits in the Holstein-Friesian (f). The muscular traits were wither height (WH), hip width (HW), chest width (CW), chest depth (CD), and back length (BL)

B5 The location of the top 5 most significant QTLs associated with each of the skeletal traits in the meta-analysis.

Trait	Chr	Start	End	No of suggestive and significant SNPs	Most significant SNP	P-Value	Candidate genes within this QTL
Wither Height	5	105280734	106302598	28	105783088[a]	4.00×10^{-7}	KCNA1, NDUFA9, FGF6, FGF23, TIGAR
	6	32458499	35344463	590	33574963[a]	1.60×10^{-8}	CCSER1
	6	36934944	40367479	233	38692833[a]	1.39×10^{-10}	PPM1K, ABCG2, PKD2, SPP1, MEPE, LAP3, NCAPG, LCORL
	15	16857667	17866073	3	17357667[a]	5.08×10^{-7}	ALKBH8, RAB39A
	22	1921471	3018467	32	2517667[a]	4.87×10^{-7}	CMC1
Chest Width	1	78722748	79819964	9	79281613[b]	1.78×10^{-7}	TPRG1, LPP
	2	84546041	85547642	17	85047257[a]	6.63×10^{-7}	SLC39A10, DNAH7, STK17B, HECW2
	8	111704513	112713771	5	112207970[b]	1.49×10^{-6}	CDK5RAP2, FBXW2, TRAF1
	16	34222491	35399155	7	34722616[b]	1.01×10^{-6}	AKT3, SDCCAG8, CEP170
	19	60882962	61882988	3	61382974[a]	3.64×10^{-7}	KCNJ2, KCNJ16, MAP2K6, ABCA5
Chest Depth	1	116367840	117409053	15	116884742[c]	7.63×10^{-7}	MBNL1, AADAC
	5	3711874	4919423	75	4314400[a]	1.44×10^{-7}	CAPS2
	5	80773564	81859095	13	81273564[a]	1.29×10^{-9}	CCDC91
	11	77828096	78855720	2	78355720[a]	5.74×10^{-7}	GDF7, RHOB, SDC1
	13	67722025	68774518	3	68222752[b]	8.81×10^{-9}	KIAA1755, ADIG, DHX35
Back Length	2	1	9238659	969	5535691[a]	1.69×10^{-19}	ASNSD1,ARHGEF4,MYO7B,NABI,MFSD6,MSTN,PMS1,ORMDL1,COL3A1,COL5A2,ANKAR,SLC40A1
	6	32966339	34249299	192	33471768[d]	1.66×10^{-8}	
	6	36399608	40835172	107	38672441[c]	4.04×10^{-10}	PPM1K, ABCG2, PKD2, SPP1, MEPE, LAP3, NCAPG, LCORL
	12	41360854	42774040	8	41860854[a]	3.67×10^{-7}	
	13	75166652	76244228	3	75166652[b]	4.86×10^{-7}	DNTTIP1, TNNC2, PLTP, NCOA5, CDH22, OCSTAMP
Hip Width	6	37042897	40544352	36	38648218[b]	2.58×10^{-7}	PPM1K, ABCG2, PKD2, SPP1, MEPE, LAP3, NCAPG, LCORL
	11	74549091	75686706	7	75078608[b]	3.54×10^{-7}	NCOA1, ITSN1, FKBP1B, ATAD2B, KLHL29
	15	7741333	8881109	3	8381102[a]	2.76×10^{-7}	ARHGAP42
	18	9307588	10781382	4	10281382[a]	3.42×10^{-7}	CDH13, HSBP1, MBTPS1, DNAAF1, TAF1C, ATP2C2, COTL1
	23	6956952	7970369	3	7456952[a]	5.27×10^{-7}	GCLC, DSB, TAP2, PSMB8, TAP1, PSMB9, COL11A2, HSD17B8, RING1, RPS18, DAXX, BAK1

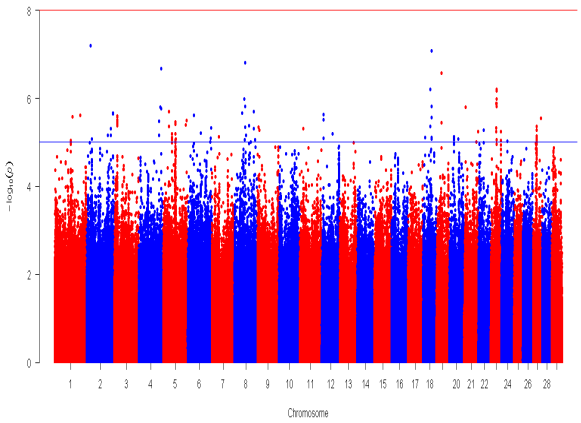
Superscript denotes SNP classification: [a]intergenic, [b]intron, [c]upstream gene variant, [d]downstream gene variant

B6 Manhattan plots for the skeletal traits

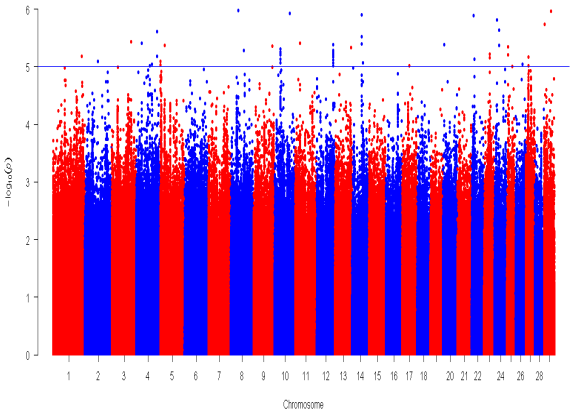


B6.1 Manhattan plots for wither height in a) Angus, b) Charolais, c) Hereford, d) Limousin, e) Simmental, and f) stature in Holstein-Friesian.

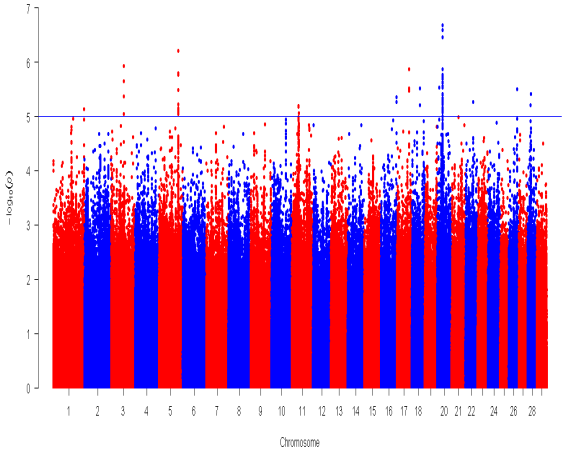
a)



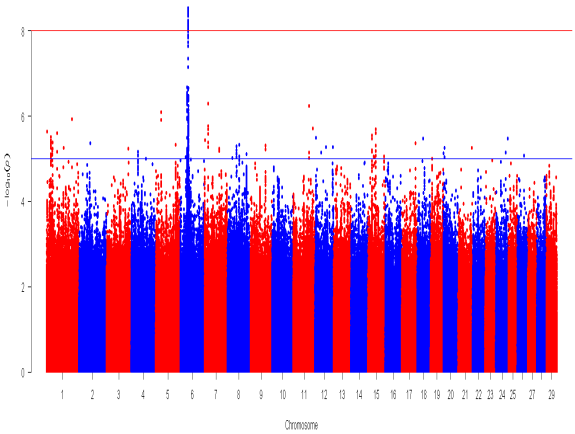
b)



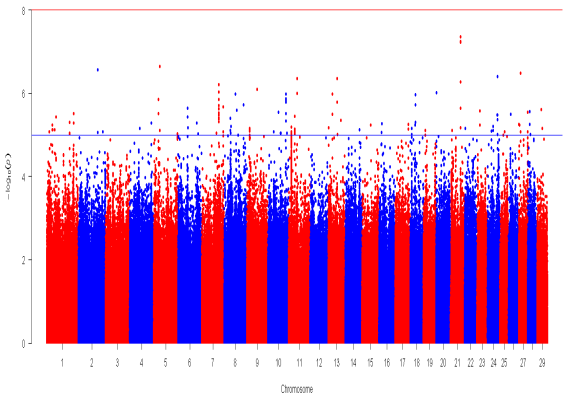
c)



d)

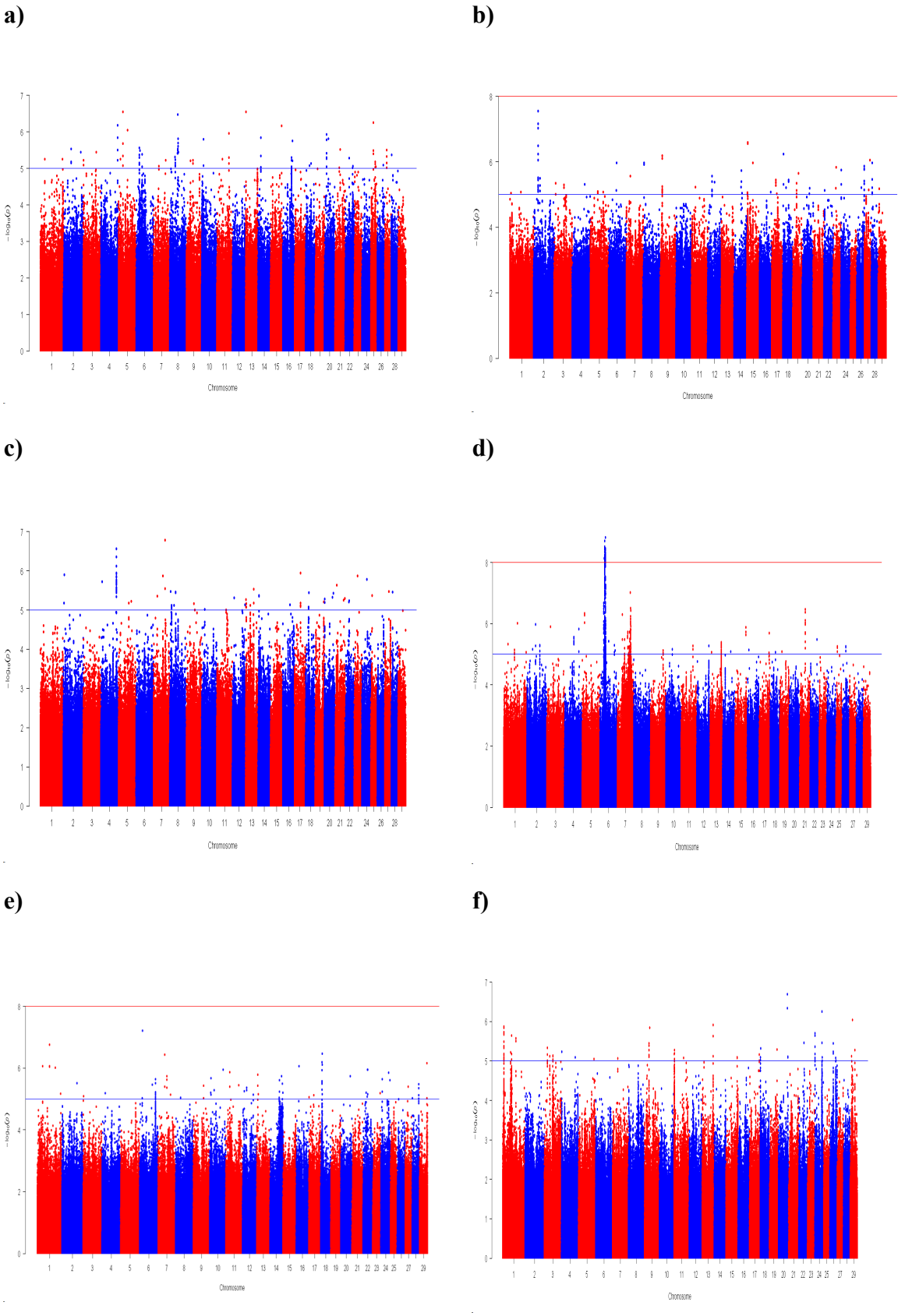


e)



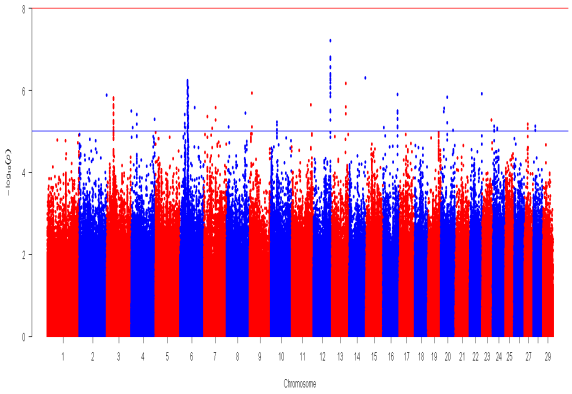
f)

B6.2 Manhattan plots for chest depth in a) Angus, b) Charolais, c) Hereford, d) Limousin, and e) Simmental

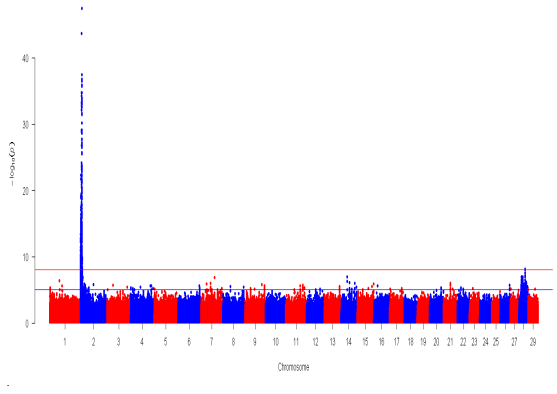


B6.3 Manhattan plots for hip width in a) Angus, b) Charolais, c) Hereford, d) Limousin, e) Simmental, and f) rump width in Holstein-Friesian.

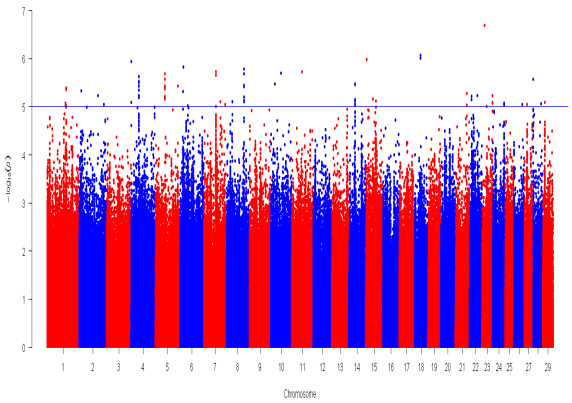
a)



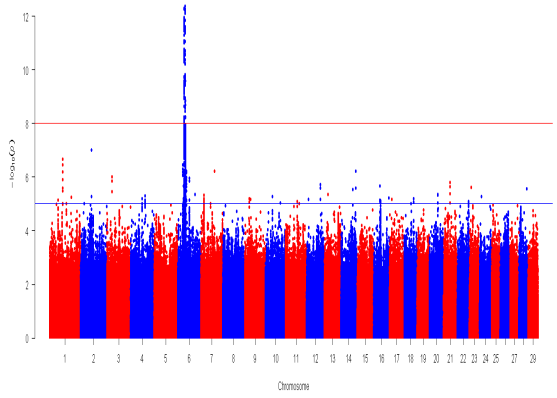
b)



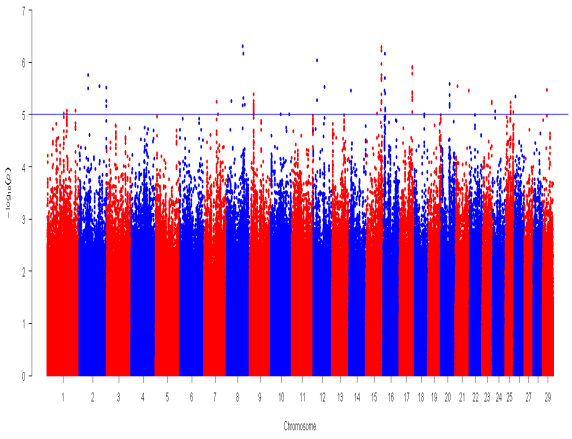
c)



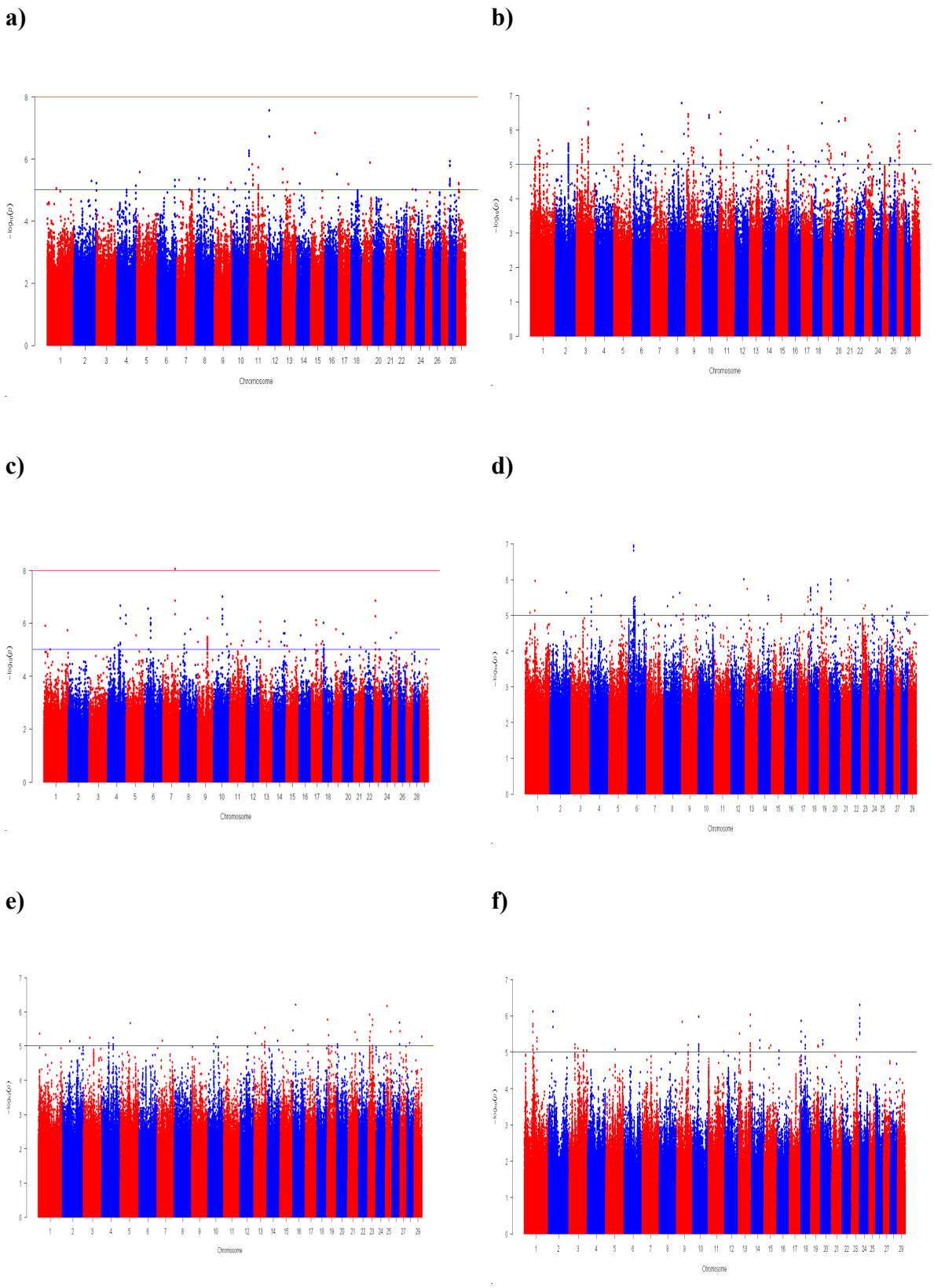
d)



e)



B6.4 Manhattan plots for back length in a) Angus, b) Charolais, c) Hereford, d) Limousin, and e) Simmental.



B6.5 Manhattan plots for chest width in a) Angus, b) Charolais, c) Hereford, d) Limousin, e) Simmental and f) Holstein-Friesian.

Appendix C

Identification of genomic regions that exhibit sexual dimorphism for size and muscularity in cattle

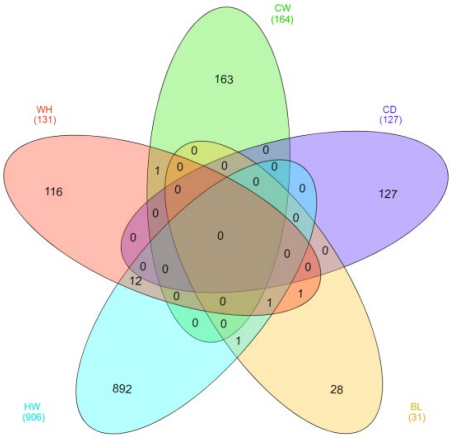
C1.The mean (μ) and phenotypic standard deviation (SD) of each linear type trait by breed and sex.

		Angus				Charolais				Hereford				Limousin				Simmental			
		Male		Female		Male		Female		Male		Female		Male		Female		Male		Female	
Traits	Scale	μ	SD	μ	SD	μ	SD	μ	SD	μ	SD	μ	SD	μ	SD	μ	SD	μ	SD	μ	SD
Muscular	*1 to 15*																				
Development of hind quarter	low - high	7.96	1.12	7.07	1.44	9.83	1.28	8.57	1.22	7.64	1.24	6.82	1.36	10.38	1.24	9.34	1.26	9.90	1.21	8.55	1.23
Development of loin	low - high	8.19	1.24	7.37	1.30	10.00	1.42	8.58	1.27	8.21	1.24	7.39	1.11	10.00	1.39	8.87	1.31	9.86	1.39	8.56	1.40
Thigh width	narrow - wide	7.86	1.28	6.99	1.41	9.68	1.51	8.34	1.30	7.71	1.36	6.78	1.20	9.62	1.56	8.60	1.52	9.69	1.45	8.35	1.38
Development of inner thigh	low - high	7.71	1.33	6.85	1.50	9.57	1.49	8.32	1.28	7.37	1.49	6.53	1.40	10.09	1.39	8.92	1.32	9.64	1.27	8.20	1.28
Wither width	narrow - wide	8.29	1.51	7.40	1.52	9.92	1.55	8.53	1.38	8.06	1.49	7.31	1.20	10.00	1.49	8.85	1.36	9.78	1.56	8.35	1.62
Skeletal	*1 to 10*																				
Chest width	narrow - wide	5.78	0.92	5.22	0.94	6.35	0.97	5.76	0.85	5.78	0.89	5.39	0.80	5.97	0.95	5.48	0.93	6.33	0.90	5.70	0.90
Chest depth	shallow - deep	6.45	0.88	6.09	0.88	6.93	0.91	6.34	0.85	6.50	0.87	6.10	0.82	6.59	0.90	6.12	0.86	7.08	0.86	6.53	0.78
Wither height	small- tall	5.55	0.97	5.08	1.02	6.10	1.14	5.66	1.05	5.64	1.01	5.16	0.98	5.89	1.08	5.45	1.06	6.49	1.00	5.82	0.97
Back length	short - long	6.42	0.98	5.88	1.04	6.88	1.07	6.41	0.99	6.47	1.04	6.00	0.97	6.72	1.06	6.34	1.06	7.28	0.92	6.73	0.93
Hip width	narrow - wide	5.48	0.91	5.40	1.06	5.80	0.95	5.51	0.94	5.65	0.87	5.68	0.91	5.71	1.04	5.46	1.07	6.10	0.90	5.76	0.98

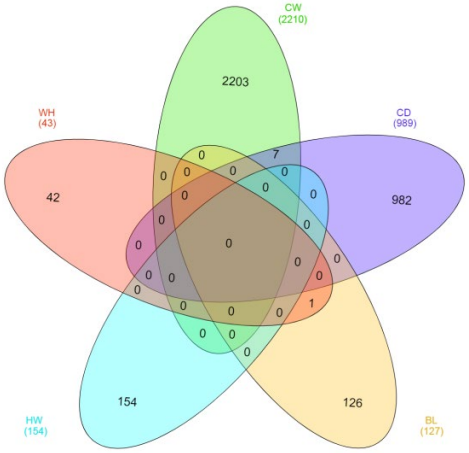
C2. The number of phenotyped and genotyped animals by sex and breed and along with the number of SNPs removed during quality control and included in the final analysis.

	Angus		Charolais		Hereford		Limousin		Simmental	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Phenotyped animals	1,812	1,544	16,145	14,904	1,582	1,422	17,930	17,229	4,489	4,143
Genotyped animals	1,044	400	4,641	1,792	727	402	5,772	2,973	956	742
Sequence SNPs	42,920,277	42,920,277	42,920,277	42,920,277	42,920,277	42,920,277	42,920,277	42,920,277	42,920,277	42,920,277
SNPs removed during minor allele frequency edit	25,815,321	26,967,481	24,288,291	24,898,964	25,107,580	25,862,892	24,082,484	24,528,104	23,970,492	24,525,773
Removed due to poor imputation accuracy	563,043	550,636	577,712	572,365	571,545	562,481	780,880	624,936	692,610	580,207
SNPs included in analysis	16,541,913	15,402,160	18,054,274	17,448,948	17,241,152	16,494,904	18,056,913	17,767,237	18,257,175	17,814,297
Total number of SNPs present in both sexes	15,008,408		17,227,625		15,991,751		17,482,131		17,319,250	

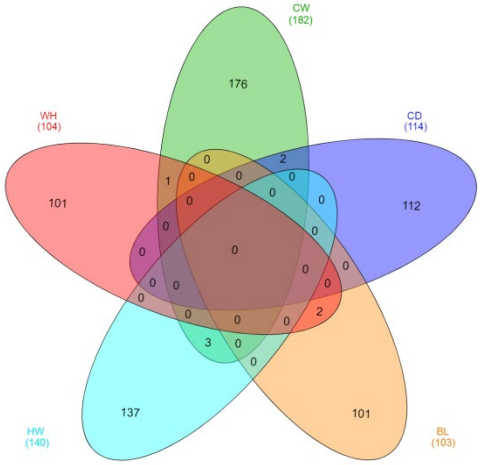
a)



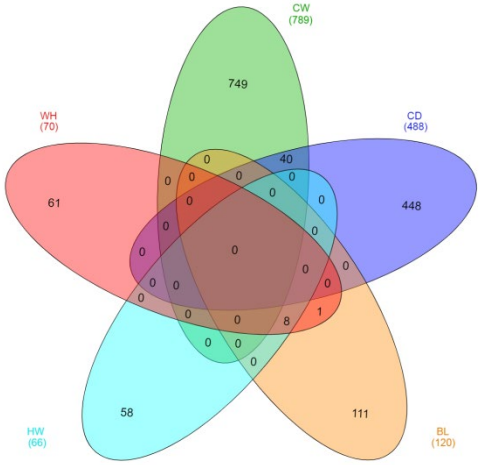
b)



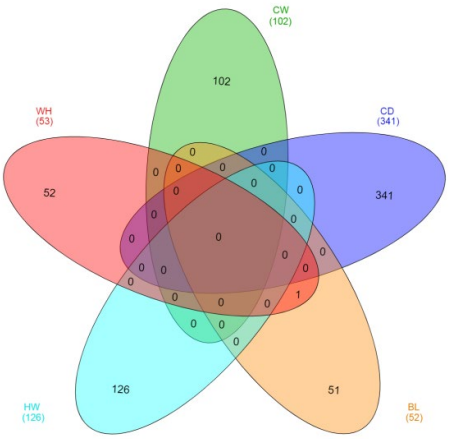
c)



d)

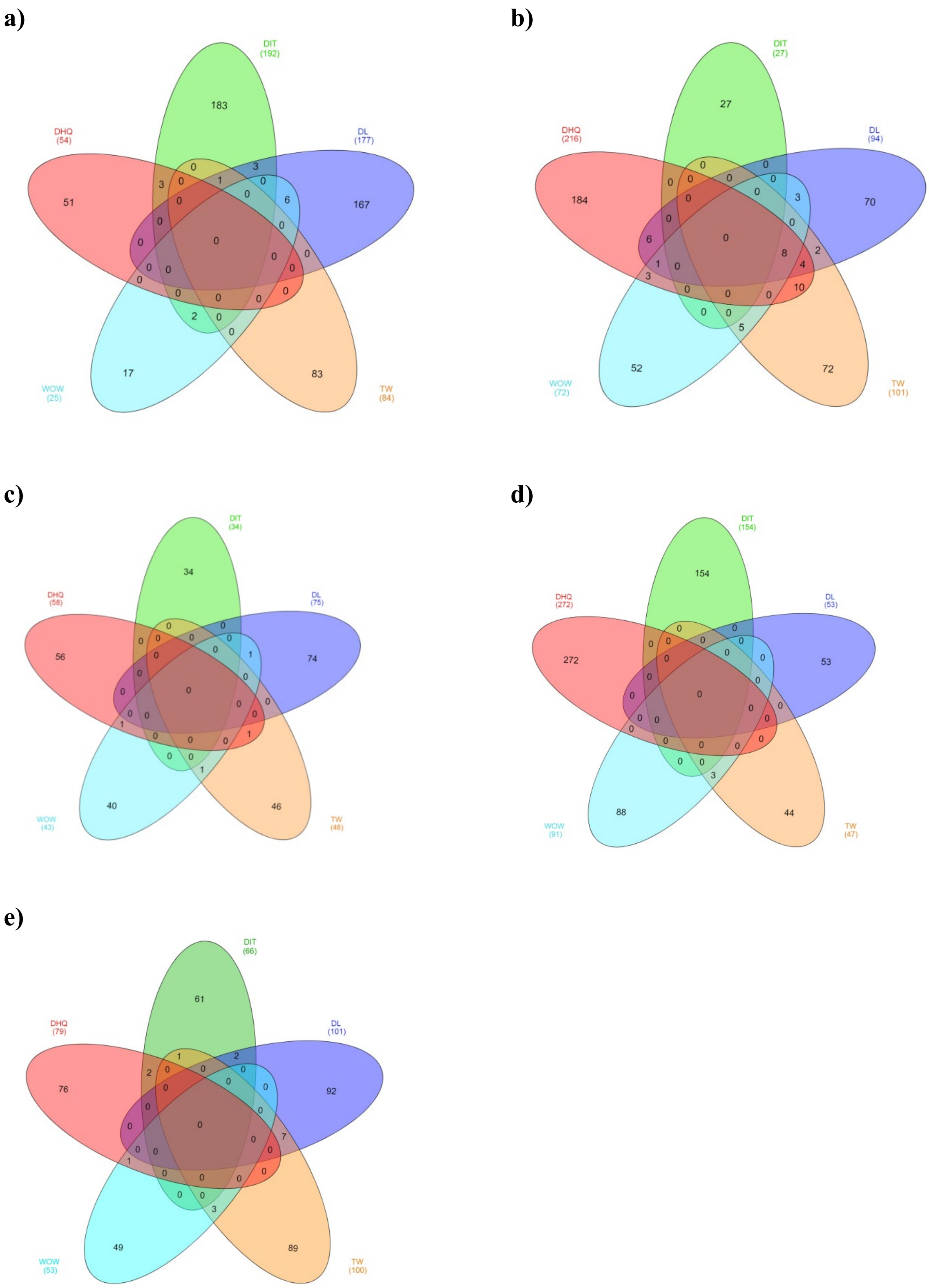


e)



C3. Overlapping 1kb regions that contain at least one suggestively or significantly dimorphic SNP for the 5 skeletal traits in a) Angus, b) Charolais, c) Hereford, d) Limousin, and e) Simmental

*trait abbreviations: CW = chest width, CD = chest depth, BL = back length, HW = hip width, WH = wither height



C4. Overlapping 1kb regions that contain at least one suggestively or significantly dimorphic SNP for the 5 muscular traits in a) Angus, b) Charolais, c) Hereford, d) Limousin and e) Simmental

*trait abbreviations: DHQ = development of hind quarter, DIT = development of inner thigh, DL = development of loin, TW = thigh width, WOW = width of wither.