A systematic review of the role of genomic copy number variation in cattle (*Bos taurus*) production and associated genes

by

Jessica Anne Old

Thesis presented in fulfilment of the requirements for the degree of
Master of Agricultural Sciences

at

Stellenbosch University
Department of Animal Sciences, Faculty of AgriSciences

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the author and are not necessarily to be attributed to the NRF

Supervisor: Prof. Kennedy Dzama
Co-supervisor: Dr. Annelin Molotsi

March 2023

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (unless to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third-party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2023

ABSTRACT

With an ever-increasing demand for milk and meat products, and with climate change posing a threat to the productive efficacy of cattle across the globe, there is a need for the application of genomic tools for animal breeding. The valuable genomic variation exhibited in cattle is important for identifying major genes and quantitative trait loci (QTL), to be used as molecular markers in the genomic selection of animals. One of the key steps of selective breeding in cattle is the characterisation of genetic variation responsible for phenotypic differences. Copy number variations (CNVs), such as duplications, deletions, and insertions, are increasingly being shown to be one of the main contributors to genomic diversity and subsequent phenotypic differences observed among animals. Copy number variation can cause major changes in gene expression, phenotypic traits, and evolutionary adaptation, through gene dosage and transcript structure alterations. Several CNV-based studies have identified gene variations that could potentially be responsible for the phenotypic differences, but the sheer number of studies on this topic makes it difficult for one to come to a definite conclusion. This systematic review summarised the relevant findings of cattle CNV research, and identified the important genes involved therewith, to help researchers remain up to date with current and ongoing research in cattle-CNV studies. This review revealed that cattle CNV research has increased considerably since 2008. The extent and distribution of the publications reflect the worldwide growing importance of understanding the cattle genome for genetic improvement of livestock. However, there is a lack of research in developing countries, a lack of emphasis on *Bos indicus* (12%) and *Bos taurus africanus* (4%) cattle, and a lack of standardised reporting across cattle CNV studies. Copy number variations can alter the gene expression and consequently influence phenotypic expression. This systematic review identified several important CNV-related genes in the published articles that influence economically important traits in cattle. These genes were related to adaptation and immunity (ABCC4, BOLA gene family, IGLL1, OR family, WC1, ZNF280B, BSP30A, DEFB, ULBP gene family, CATHL gene family, and HSP gene family), milk yield, milk composition and reproduction (DGAT1, IFNT, PAG, PRAME, PRL, AP3B1, IGLL1, SLC27A6, ITFG1, MTHFSD, PRP, and PTK2), meat yield, meat quality and growth (IGF2, PLA2G2D, CAST, IGF1R, APOL3, PTPRC, KCNJ12, CAPN1, AGBL3, CTNNA1, MSTN, ADRA1B, ATRN, LRRC49, MYH3, SORCS2, and TG), feed efficiency (PRKG1, FABP2, and EIF2S1), and coat colour, coat patterns and hair morphology (KIT, AP3B1, MC1R, PRLR, and FGF18). This knowledge is relevant from a molecular perspective to the practical application in animal breeding, and offers breeders the means to consider genomic selection of animals at a younger age. The productive efficacy of cattle is vulnerable, thus, the use of molecular assisted selective breeding is essential for overcoming current and future challenges in cattle productivity.

OPSOMMING

Met ’n al-toenemende aanvraag vir suiwel en vleis produkte, en met klimaatsverandering wat regoor die wêreld ’n bedreiging inhou vir die produktiewe doeltreffendheid van beeste, is daar ’n behoefte aan die toepassing van genomiese hulpbronne vir diere teling. Die waardevolle genomiese variasie wat in beeste waargeneem is, is belangrik vir die identifisering van hoof gene en kwantitatiewe eienskap loci, wat dan gebruik kan word as molekulêre merkers in die genomiese seleksie van diere. Een van die belangrikste stappe van selektiewe teling in beeste is die karakterisering van geneties variasie wat verantwoordelik is vir fenotipies verskille. Kopie Getal Variasies (KGV’s), soos duplikasies, verwyderde gene en invoegings, word toenemend geïdentifiseer as een van die hoof bydraers tot genomiese diversiteit en die daaropvolgende fenotipiese verskille wat tussen diere waargeneem word. KGV’s kan groot veranderinge in geen uitdrukking, fenotipiese eienskappe en evolusionêre aanpassing veroorsaak. Hierdie veranderinge is as gevolg van geen dosis en transkripsie struktuur veranderinge. Verskeie KGV-gebaseerde studies het geen variasies, wat moontlik vir die fenotipiese verskille verantwoordelik kan wees, geïdentifiseer. Maar die blote aantal studies op hierdie onderwerp maak dit moeilik om tot ’n definitiewe gevolgtrekking te kom. Hierdie sistematiese ontleding som die relevante bevindings van bees-KGV navorsing op, en identifiseer die belangrike gene wat daarin betrokke is, om navorsers te help om op datum te bly met huidige en aangaande navorsing in bees-KGV studies. Hierdie ontleding het aan die lig gebring dat bees-KGV navorsing aansienlik toegeneem het sedert 2008. Die omvang en die sprei van die publikasies reflekteer dat dit wêreldwyd toenemend belangrik is om die bees genoom te verstaan vir die genetiese verbetering van vee. Maar daar is egter ’n gebrek aan navorsing in ontwikkelende lande, ’n gebrek aan beklemtoning van *Bos indicus* (12%) en *Bos taurus africanus* (4%) bees, en ’n gebrek aan gestandaardiseerde verslagdoening regoor bees-KGV studies. KGV’s kan die geen uitdrukking verander en gevolglik die fenotipiese uitdrukking beïnvloed. Hierdie sistematiese ontleding het verskeie belangrike KGV-verwante gene, wat ekonomiese belangrike eienskappe in beeste beïnvloed, in die gepubliseerde artikels geïdentifiseer. Dié gene hou verband met aanpassing en immuniteit (ABCC4, BOLA geen familie, IGLL1, OR familie, WC1, ZNF280B, BSP30A, DEFB, ULBP geen familie, CATH geen familie, en die HSP geen familie), melk opbrengs, melk samestelling en voortplanting (DGAT1, IFNT, PAG, PRAME, PRL, AP3B1, IGLL1, SLC27A6, ITFG1, MTHFSD, PRP, en PTK2), vleis opbrengs, vleis kwaliteit en groei (IGF2, PLA2G2D, CAST, IGF1R, APOL3, PTPRC, KCNJ12, CAPN1, AGBL3, CTNNA1, MSTN, ADRA1B, ATRN, LRRC49, MYH3, SORCS2, en TG), voer doeltreffendheid (PRKG1, FABP2, en EIF2S1), en pels kleur, pels patrone en haar morfologie (KIT, AP3B1, MC1R, PRLR, en FGF18). Hierdie kennis is relevant vanaf ’n molekulêre perspektief tot die praktiese toepassing in diere teling, en bied telers ’n middel om die genomiese seleksie van diere op ’n jonger ouderdom te oorweeg. Die produktiewe doeltreffendheid van beeste is vatbaar, dus is die gebruik van molekulêre geassisteerde selektiewe teling noodsaaklik om die huidige en toekomstige uitdagings in bees produktiwiteit te oorkom.

ACKNOWLEDGEMENTS

I would like to express my gratitude and appreciation to the following persons:

- Thank you to both Prof K. Dzama and Dr A. Molotsi for being so willing to take me under their wing on such short notice. Thank you for not only being incredibly helpful and supportive, but for being so friendly, accommodating, and caring. You have both made this a very positive experience for me. I am forever grateful to both of you.

- Thank you, Dr H. Lambrechts, for ensuring that I continued with my Masters. I would not be where I am if it were not for you taking action and getting me back on board.

- Thank you, Elizabeth Will-Mollard, for your assistance with the planning of the systematic review and library queries.

- Thank you, Natasha, for your great ideas, guidance, time, and patience you have had with me, I am so very grateful.

- Thank you to my friends for all the laughs, all the cries, and all the words of encouragement. Thank you Coena for your technical editing, and Mari for help with translation.

- Thank you, Matthew, for your technical assistance and for your endless support and positivity, both emotionally and mentally.

- Thank you, mum and dad, for your constant love and care. Thank you for everything you do for me, without you this endeavour would not have been possible.

I would like to thank the following institutions for financial support and use of facilities:

- National Research Foundation (NRF)

- Department of Animal Sciences, Stellenbosch University.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

γδT	Gamma delta T
ADG	Average daily gain
AI	Artificial Insemination
AMPs	Antimicrobial peptides
ATP	Adenosine triphosphate
BAF	B allele frequency
BPI	Bactericidal/permeability-increasing
BTA	Bos Taurus Autosome
CGH	Comparative Genomic Hybridisation
CNV	Copy number variation
CNVR	Copy number variable region
DNA	Deoxyribonucleic acid
ECM	Extracellular matrix
FA	Fatty acid
FCR	Feed conversion ratio
FOSTES	Fork stalling and template switching
GIN	Gastrointestinal nematode
GTP	Guanosine triphosphate
GWAS	Genome Wide Association Study
Ig	Immunoglobulins
IMF	Intramuscular fat
LRR	Log R ratio
MAS	Marker-assisted selection
MT	Meat tenderness
NAHR	Non-allelic homologous recombination
NGF	Nerve growth factor
NGS	Next generation sequencing
NHEJ	Non-homologous end joining
PICO	Population, Intervention, Comparison, Outcome
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses

QTL	Quantitative trait loci
RNA	Ribonucleic acid
SA	South African
SCC	Somatic cell count
SFA	Saturated fatty acids
SNP	Single Nucleotide Polymorphism
SVs	Structural variants
UFA	Unsaturated fatty acids
USA	United States of America
VNTRs	Variable number tandem repeats
WES	Whole exome sequencing
WGS	Whole genome sequencing
YBP	Years before present

LIST OF GENES

ABCC4	ATP binding cassette subfamily C member 4
ADRA1B	Adrenoceptor alpha 1B
AGBL3	The ATP/GTP binding protein-like 3
AP	Adaptor-related protein
AP3B1	Adaptor related protein complex 3 subunit beta 1
APOL3	Apolipoprotein L3
ASIP	Agouti-signalling protein
ATRN	Attractin
BOLA	Bovine leukocyte antigen
BSP30A	Bovine salivary protein 30 kDa
CAPN1	Calpain 1
CAST	Calpastatin
CATHL	Cathelicidin
CATHL4	Cathelicidin 4
cGK	cGMP-dependent protein kinase
CTNNA1	Catenin alpha 1
DEFB	Beta-defensin
DGAT1	Diacylglycerol O-acyltransferase 1
EIF2S1	Eukaryotic translation initiation factor 2 subunit alpha
FABP2	Fatty acid binding protein 2
FATP	Fatty acid transport protein
FGF	Fibroblast growth factor
FGF18	Fibroblast growth factor 18
FGFR3	Fibroblast growth factor receptor 3
GBP2	Guanylate binding protein 2
GDF8	Growth and Differentiation Factor 8
HSF	Heat shock transcription factor
HSP	Heat shock protein
HSP70	70-kDa heat shock protein

IFABP	Intestinal fatty acid binding protein
IFN	Interferon
IFNT	Interferon tau
IGF	Insulin-like growth factor
IGF1R	Insulin-like growth factor 1 receptor
IGF2	Insulin-like growth factor 2
IGFBP	Insulin-like growth factor binding protein
IGLL1	Immunoglobulin lambda-like polypeptide 1
IL4	Interleukin 4
ITFG1	Integrin alpha FG-GAP repeat containing 1
KCNJ12	Potassium inwardly rectifying channel, subfamily J, member 12
KIT	KIT proto-oncogene, receptor tyrosine kinase
LRRC49	Leucine rich repeat containing 49
MC1R	Melanocortin 1 receptor
MHC	Major histocompatibility complex
MHCLA	Major Histocompatibility Complex Class I-like Gene Family A
MRP4	Multi-drug resistance protein 4
α-MSH	α-Melanocyte stimulating hormone
MSTN	Myostatin
MTHFSD	Methenyltetrahydrofolate Synthetase Domain Containing
MYH3	Myosin heavy chain 3
NKG2D	Natural killer group 2, member D
OR	Olfactory receptor
PAG	Pregnancy-associated glycoprotein
PLA2G2A	Phospholipase A2 group IIA
PLA2G2D	Phospholipase A2 group IID
PRAME	Preferentially expressed antigen in melanoma
PRAMEY	Y-linked preferentially expressed antigen in melanoma
PRKG1	Protein kinase cGMP-dependent type I
PRL	Prolactin
PRLR	Prolactin receptor

PRP	Prolactin-related proteins
PRRs	Pattern recognition receptors
PTK2	Protein tyrosine kinase 2
PTP	Protein tyrosine phosphatase
PTPRC	Protein tyrosine phosphatase receptor type C
SLC27A	Solute carrier 27A
SLC27A6	Solute carrier family 27 member 6
SORCS2	Sortilin related VPS10 domain containing receptor 2
TAP	Tracheal antimicrobial peptide
TCR	T cell antigen receptor
TG	Thyroglobulin
ULBP	UL16-binding protein
ULBP17	UL16-binding protein 17
ULBP2	UL16-binding protein 21
WC1	Workshop cluster 1
ZNF	Zinc finger
ZNF280B	Zinc finger protein 280B

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 GENERAL INTRODUCTION

1.1 Background

The domestication of cattle has provided many benefits to humans, such as meat, milk, leather and draught power for cultivation or transportation purposes (Felius *et al.*, 2014; Mei *et al.*, 2019). Cattle are therefore invaluable to the agricultural sector, in both developing and developed countries. After humans started livestock domestication, cattle populations were shaped through selective breeding in response to the owner's needs, and natural selection, for the adaptation to diverse environments (Mwai *et al.*, 2015). Domestic cattle can be classified into two main groups, namely, -*Bos taurus* (taurine) and *Bos indicus* (zebu) (Burt, 2009), with their cross being defined as *Bos taurus africanus*. Taurine cattle are mainly found in areas with a temperate climate, whereas zebu cattle are mainly found in more extreme climates to which they had to adapt (Pérez O'Brien *et al.*, 2014). Zebu cattle exhibit characteristics such as heat tolerance, parasite and disease resilience, and lowered nutritional requirements due to the long-term natural selection endured in these extreme climates (Canavez *et al.*, 2012; Utsunomiya *et al.*, 2019). Apart from having to adapt to these extreme environments, human herd management and several selection practices have brought about different patterns of variation in the genomes of these subspecies (Decker *et al.*, 2014). This genomic variation can be used as a molecular marker for marker-assisted selection (MAS) for the improvement of cattle breeds. Moreover, this variation provides a great opportunity for researchers to investigate the genes involved in cattle phenotypic trait differences, for the characterisation of cattle at the genomic level.

Cattle can occupy various environments with their innate ability to convert poor quality forage into good quality meat and milk products (Elsik *et al.*, 2009). This advantageous process has since been exploited by the livestock sector to provide many important products to humans. The commercial livestock sector is, however, constantly faced with new challenges. The increase in world population triggers an increase in requirement for livestock products. At the same time, climate change is causing, among other events, a rise in global temperatures and a reduction in annual rainfall (Silpa *et al.*, 2021). Consequently, an increase in temperature may cause heat stress in livestock, whereby beef cattle have reduced feed intake and growth performance, and dairy cattle a reduced milk yield (Dzama, 2016).

The agricultural sector will be challenged with the predicted temperature increase and scarcity of resources necessary for production (Escarcha *et al.*, 2018). Many studies have proven that cattle are able to adapt to harsh environments, and the genetic mechanisms behind this adaptation are continuously being investigated (Wang, 2016; Pickering, 2017; Fernandes Júnior *et al.*, 2020; Zhang *et al.*, 2020a; Guo *et al.*, 2021). Natural selection often leads to directional selection for adaptive

traits, such as thermotolerance, acclimatisation to severe environments and resistance to diseases and parasites, leading to better survival in a certain environment (Randhawa *et al.*, 2016). Similarly, artificial selection leads to the enhancement of economically important production traits through the genetic improvement of breeds (Randhawa *et al.*, 2016; Mei *et al.*, 2019).

Genetic diversity is a fundamental source of biodiversity (Hughes *et al.*, 2008), and refers to differences in allele frequencies and combinations, and thus genetic variation (Kantanen et al., 2015). The genetic variability among cattle breeds is a valuable genetic resource. Detecting DNA variants which directly impact an individual's phenotype is an important area of research in livestock genomics (Kijas *et al.*, 2011) as it assists researchers in uncovering genotype-phenotype associations in cattle (Zhan *et al.*, 2011). Genetic variation can be determined using microsatellites and single nucleotide polymorphisms (SNPs), which, until recently, were considered as the most important cause of variation. Copy number variations (CNVs) have since been reported to play a major role in genetic variation and thought to play a key role in adaptation and breed formation (Pierce *et al.*, 2018). They encompass both amplifications and deletions, are an essential source of genetic diversity, and have been discovered throughout every domain of life (Lauer & Gresham, 2019; Kommadath *et al.*, 2019). Due to the advancement of molecular genetics, and the wide use thereof, it is possible to find the CNV-related genes that are responsible for economically important traits, to use as molecular markers for selection (Yudin & Voevoda, 2015). Copy number variations contribute to rapid adaptive evolution, the development of diseases such as cancer, and population diversity, thus they are the subject of extensive research (Lauer & Gresham, 2019; Pös *et al.*, 2021).

Copy number variations have been identified and reported in numerous cattle breeds, and the genes detected within or near the CNV regions is understood to regulate aspects of phenotypic variation in cattle (Kijas *et al.*, 2011). By obtaining more information on the presence and prevalence of CNVs within the cattle genome, more understanding into the genetic mechanisms involved in economically important phenotypic traits can be discovered.

1.2 Problem Statement

Global beef production has more than doubled in the last 60 years (FAO, 2022), owing to the increase in world population as well as the increase in income in developing countries (Whitton *et al.*, 2021). Global milk production has increased alongside, with a total milk production of 718 million tonnes in 2020, increasing with approximately 404 million tonnes from 1961 (FAO, 2022). At the same time, climate change is shifting temperatures, increasing fire incidences, and changing rainfall patterns, thus posing a threat to the efficacy of cattle production throughout the world (Mwai *et al.*, 2015; Silpa *et al.*, 2021). Climate change can affect livestock directly, by influencing the animal's

performance, or indirectly, by impacting feed resources. Beef cattle reared on natural pastures and in feedlots are vulnerable because feed crops will be affected, with an expected grazing capacity loss of 30-50% (Rust & Rust, 2013). Moreover, loss in milk production owing to climate change could reach 15 million tons by 2050 (Silpa *et al.*, 2021).

These alarming statistics pose challenges to the livestock sector that will need to be overcome. This warrants the idea of developing a suitable mitigation approach to cope with both the negative effects of climate change and the escalating demand for meat, milk, and milk products. Although novel management and feeding regimes could solve the issue for a while, a more permanent strategy is needed. Cattle, an invaluable resource to humans, are able to perform in various environments, denoting their resilient nature. Therefore, investigation into the genetic mechanisms involved in this innate resilience is important. Using evolving genomic tools and advanced statistical models, important variable genomic regions responsible for phenotypic variation in cattle can be pinpointed. Several studies have been carried out to quantify this genetic variation and disentangle the numerous genes involved in various functions in cattle, but different data collection strategies and assessments used by researchers makes it difficult to compare studies and collate important data (Liu & Bickhart, 2012).

Moreover, there are still numerous open questions about the genes and mechanisms responsible for cattle phenotypic variation. Although CNV-based studies have identified several variable genes that could potentially be responsible for the phenotypic differences, the sheer number of studies on this topic makes it difficult for one to come to a definite conclusion.

1.3 Significance of research

The urge to have a comprehensive list of all the structural variants in a population is not exclusive to human genomics (Couldrey *et al.*, 2017). One of the key stages of selective breeding in cattle is the characterisation of the genetic variation responsible for phenotypic differences observed among cattle breeds. This observed variation can be associated with CNVs and then used to determine milk or meat production, feed conversion ratio, morphology, disease/parasite tolerance and, moreover, improve yields. This knowledge is very applicable, and important, not only from a molecular perspective, but also for practical application. Many CNV-based studies have been carried out to quantify this variation and, although this is a big step in the right direction, there are issues and challenges that have arisen. With the wave of CNV studies currently being reported, it is often difficult to compare studies and collate important data.

Aim of the study:

This study therefore aims to summarise the relevant findings of cattle CNV research, and the important genes involved therewith, for the purpose of elucidating the involvement of CNVs in important cattle production traits.

1.4 Objectives

1. The first objective of this study was to consolidate qualitative data of the published research to identify the state of knowledge and the progress in research surrounding cattle-CNV studies.

2. The second objective of this study was to perform a systematic review of published research to identify the main CNV-related genes being detected in research and their implication on economically important traits in cattle. These traits include adaptability and immune-related traits, milk production and reproduction-related traits, meat production and growth-related traits, feed efficiency traits, and coat characteristics.

Objective 1 was achieved in chapter 4 by identifying and discussing the bovine subspecies and production types that are predominantly studied, as well as which genotyping methodology is predominantly utilised. This was further mapped per country, to analyse in which regions publications are produced.

Objective 2 was achieved in chapter 5 by analysing the CNV results produced from the included publications, identifying the genes harboured within or near these CNV regions, and extracting the important candidate genes. These genes were then grouped according to which economical trait of importance they are related, thereafter the most frequent genes in each of the classification groups were reported and thoroughly discussed.

Attaining these objectives may help researchers remain up to date with current and ongoing research in cattle-CNV studies, and additionally, point out where gaps in knowledge exist.

1.5 Thesis layout

This thesis is structured in the form of a general introduction (chapter 1), a literature review (chapter 2), methodology (chapter 3), characterisation of the data (chapter 4), and an analysis of genes (chapter 5). A summary of results, recommendations and conclusion is presented at the end (chapter 6).

CHAPTER 2 LITERATURE REVIEW

2.1 History of cattle

Following the domestication of smaller and easier to manage livestock, such as sheep and goat, people began to domesticate cattle (Felius et al., 2014). Cattle, representing the suborder Ruminantia, were the ideal candidates for domestication due to their valuable traits, such as an herbivorous diet, rapid growth rate, temperate nature, and ability to breed in captivity (Elsik et al., 2009; Felius et al., 2014). Cattle are classified into two main groups, namely, *Bos taurus* (taurine) and *Bos indicus* (zebu), both are descendants from the wild aurochs, Bos primigenius (Burt, 2009). *Bos taurus africanus* are believed to comprise a cross between taurine and zebu cattle (Rewe et al., 2009; Gororo et al., 2018).

Archaeological and molecular data suggest that taurine and zebu cattle were domesticated separately. Taurine cattle were domesticated approximately 10 000 years before present (YBP) in Fertile Crescents, and zebu cattle were domesticated 8 000 YBP in the Indus Valley (Bollongino *et al.*, 2012; Felius *et al.*, 2014; Jang *et al.*, 2021). Both taurine and zebu cattle dispersed quickly after domestication to diverse environments. Taurine cattle made their way through Turkey into northern Italy, subsequently dispersing throughout Europe. Additionally, they may have also made their way along the northern coast of Africa and crossed over to the Iberian Peninsula (Figure 2.1) (Pitt *et al.*, 2019). Zebu cattle migrated into Africa (as recently as 3 000-4 000 YBP) and proceeded to central and southern parts of the continent. They also dispersed to South-East Asia and China (Figure 2.1) (Pitt *et al.*, 2019; Utsunomiya *et al.*, 2019). Cattle occupy various environments due to their innate ability to convert poor quality forage into good quality meat and milk products (Elsik *et al.*, 2009). This innate conversion process has been exploited since domestication, providing many products to humans, such as meat, milk, and leather, proving cattle to be an invaluable resource (Felius *et al.*, 2014).

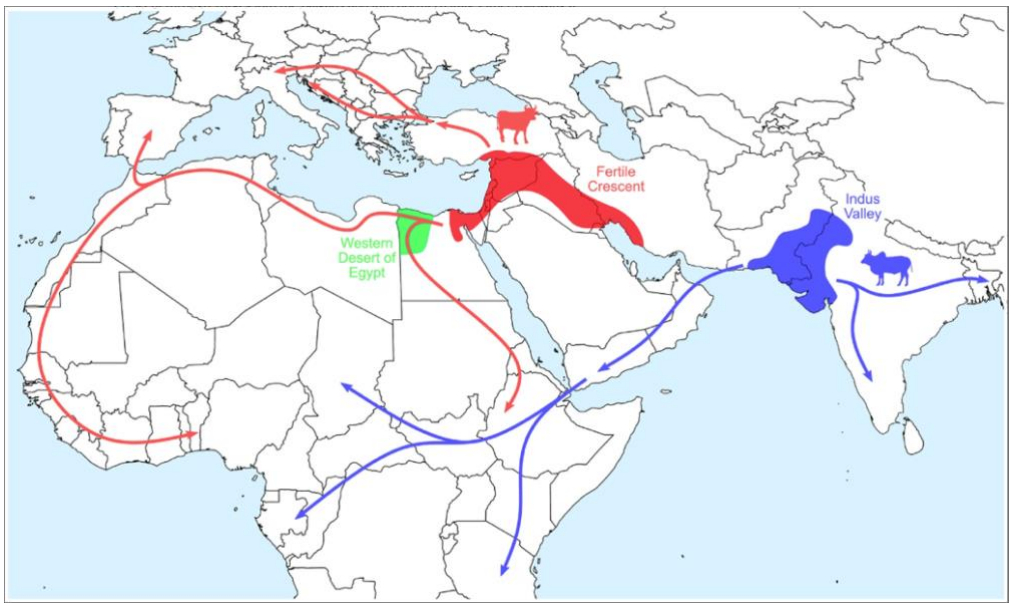

Figure 2.1: Approximate domestication sites and migration routes for *Bos taurus* (red) and *Bos indicus* (blue) cattle. Taken From Pitt *et al*. (2019).

2.1.1 Subspecies

Attributes of zebu cattle include the presence of a fatty hump, drooping ears, large dewlap and naval flap (Figure 2.2b) (Magee *et al.*, 2014). Zebu cattle also exhibit superior physiological characteristics such as heat tolerance, parasite resistance, disease resilience and lowered nutritional requirements, due to the long-term natural selection endured in the tropics (Canavez *et al.*, 2012; Utsunomiya *et al.*, 2019). Their body temperature is regulated by maintaining a lower metabolic rate during heat stress, moreover, they maintain a lower respiratory rate, rectal temperature and water requirement (Wang *et al.*, 2015).

Taurine cattle (Figure 2.2a), on the other hand, do not demonstrate these resilient characteristics. However, they do have excellent production characteristics (meat and milk quality) and are well adapted to cooler climates. Taurine cattle also present a better reproductive performance, such as earlier onset of puberty and shorter post-partum anoestrus, compared with zebu cattle (Utsunomiya *et al.*, 2019). However, the decreased reproductive performance of zebu cattle could be because they are often farmed in low-input systems, thus their poorer reproduction could be the result of inadequate nutrition (Utsunomiya *et al.*, 2019). Furthermore, copy number variations (CNVs) have been compared between *Bos taurus* and *Bos indicus* cattle (Hu *et al.*, 2020a). Although several CNV regions were shared between the two subspecies, there were considerable CNV differences detected to successfully set them apart. European breeds dominate the current genetic resources (Talenti *et al.*, 2022), with the primary reference genome stemming from a European taurine breed (Burt, 2009), along with the design of the first high-throughput microarrays of SNP markers (Utsunomiya *et al.*, 2019). Sanga-type cattle, such as the Nguni, Afrikaner, and Drakensberger of

South Africa, or the Nkone and Tuli of Zimbabwe, are believed to comprise of a cross between taurine and zebu cattle (Rewe *et al.*, 2009; Gororo *et al.*, 2018). These cervico-thoracically humped Sanga cattle reached Southern Africa 250-500 AD (Felius *et al.*, 2014). Sanga cattle are identified as the subspecies *Bos taurus africanus* (Strydom *et al.*, 2001).

In tropical areas with warmer climates, epidemic diseases and pathogens are a common occurrence, the environmental conditions are taxing and often food and water are scarce, thus the locally adapted breeds, such as Nguni, exhibit a great level of resistance and adaptation (Keba *et al.*, 2010). These indigenous breeds have not been well characterised or described, and rarely undergo structured breeding programmes that allow for improved performance (Nyamushamba *et al.*, 2017). Nguni cattle are recognised for their multicoloured, patterned hides as well as their small frame size (Figure 2.2c) (Wang *et al.*, 2015).

Figure 2.2: a) Taurine cattle (Hereford bull), b) Zebu cattle (Brahman bull), c) Sanga cattle (Nguni bull). From The Cattle site (2022) .

2.2 Cattle farming

After the domestication of cattle, human-oriented selection further contributed to the evolution of these animals (Porto-Neto *et al.*, 2013). The establishment of the concept of a breed in the 19th century triggered human breeders to select animals for their valuable production traits, thus creating distinct groups based on their phenotypes (Porto-Neto *et al.*, 2013). Globally, there are estimated 1.5 billion domesticated cattle (FAO, 2018), comprising over 1 000 domesticated breeds (Yurchenko *et al.*, 2018). Each breed has phenotypic characteristics, differentiating them from other breeds. Cattle form a crucial source of nutrition for the global human population (Canavez *et al.*, 2012). Thus, these breeds are reared for either beef or dairy production.

Populations have been consuming beef since the beginning of mankind (Hocquette *et al.*, 2018). After domestication, beef production advanced all over the world. In 2020, global beef production was 67.9 million tonnes, which is up 12.7 million tonnes from 2001 (55.2 MT), and approximately 40 million tonnes from 1961 (27.7 MT) (FAO, 2022). Thus, since 1961, meat production has increased by 145% (FAO, 2022). This increase in production is due to the growing population, as well as

increasing incomes in developing countries. Global milk production was 718 million tonnes in 2020, increased approximately 222 million tonnes from 2001 and 404 million tonnes from 1961 (FAO, 2022). In 2020, South Africa produced 1.04 million tonnes of beef and 3.8 million tonnes of milk.

Beef cattle are raised primarily for meat production. An important factor in beef production is the growth potential of the individual animal because the production of muscle and fat (meat) in beef cattle is a result of the growth function (Hozáková *et al.*, 2020). A beef animal should ideally have a moderate birth weight, a rapid growth rate and an early maturation for an early finishing for slaughter, however, both early (Jersey, Hereford) and late (Limousine, Charolais) maturing breeds have advantages and disadvantages (Hozáková *et al.*, 2020). The common goal among beef producers is to improve feed efficiency and growth rates, as this will be economically beneficial (Vickers & Stewart, 2019). Different beef breeds have distinctive characteristics, but overall, beef breeds have more muscle and fat, and have a stockier built than dairy cattle (Figure 2.3a).

Dairy cattle are raised primarily for milk production, producing milk in excess to what their calf requires. Cows must calve every year for milk production to continue. Naturally, dairy cattle have a very well-developed udder, while their body is often thin and bony, compared to beef cattle, (Figure 2.3b) as most of their energy is used for milk production. Different dairy breeds have distinguishing characteristics, such as the Holstein breed with the highest milk production (Seroussi *et al.*, 2010), or the Jersey breed, which is known to have low milk productivity, but higher protein and milk fat content (Lim *et al.*, 2020).

Moreover, dual purpose cattle (Figure 2.3c), such as the Fleckvieh cattle are acceptable for both beef and milk production. Dual-purpose systems are cattle production systems where both meat and milk are produced simultaneously, with lower productivity compared to systems based on exclusively on milk- or beef production (González-Quintero *et al.*, 2020).

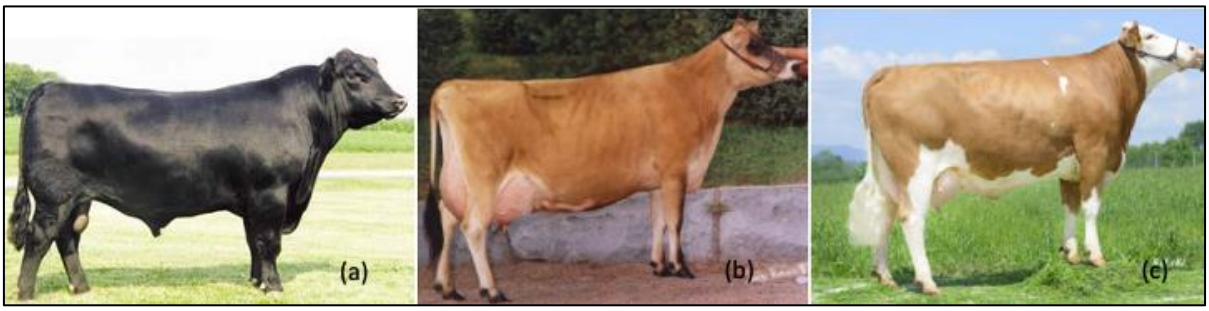

Figure 2.3: Typical (a) beef breed (Aberdeen Angus), (b) dairy breed (Jersey), (c) dual-purpose breed (Fleckvieh). From The Cattle site (2022) .

2.2.1 South African livestock industry

The South African livestock industry comprises beef, dairy and small stock herds farmed in all nine provinces, characterised by varied biomes, rainfall patterns, and temperatures (van Marle-Köster & Visser, 2018). The livestock sector consists of developed commercial sectors utilising advanced genetic technology as well as developing sectors such as emerging and smallholder farmers. There are many factors that influence the use of genomics in the livestock sector such as funding, socio-economic restraints and agricultural extension services (van Marle-Köster & Visser, 2018). Emerging farmers intend to join the commercial sector, but socio-economic factors such as financial support, land disputes and access to markets are hinderances (Khapayi & Celliers, 2016). Smallholder farmers tend to focus on subsistence and use more traditional farming methods. In South Africa, cattle production is the most vital livestock sub-sector, contributing approximately 20-30% to the total agricultural output per annum (Musemwa *et al.*, 2008). Smallholder cattle farms are often multi-sectoral, meaning animals are kept for purposes involved in more than purely economic endeavours, such as for meat, milk, manure, draught power and transport, hides and cultural purposes (Mapiye *et al.*, 2009). The main uses of these animals will be determined by the socio-economic and cultural factors of the community (Katiyatiya *et al.*, 2014). Cattle provide somewhat financial security during periods of poor cropping or crop failure, thus making households less vulnerable to adverse events (Gororo *et al.*, 2018). Millions of people who live in marginal production regions are dependent on cattle. This is due to cattle ownership forming a societal safety net and contributing to the resilience of livelihoods (Murungweni *et al.*, 2014; van Vliet *et al.*, 2015). The smallholder cattle farms are an underutilised resource for the production of beef in South Africa, thus efforts need to be made to increase production and off-take (Mapiye *et al.*, 2009).

The SA beef cattle industry is made up of various breeds of cattle and include indigenous, composite and exotic breeds (Abin *et al.*, 2016). Indigenous breeds, such as the Nguni cattle, can adapt and produce under various environments, while still producing high-quality meat (Strydom *et al.*, 2000; Muchenje *et al.*, 2008). However, they have been shown to have a smaller mature weight, poorer conformation, and a lower dressing percentage compared to imported breeds (Muchenje *et al.*, 2008). In beef cattle, meat quality, carcass traits and growth traits have a significant impact on product pricing, consumer satisfaction and overall profitability. Selection for growth traits have been a focus in several cattle breeds in South Africa (Abin *et al.*, 2016).

2.2.2 Impacts of climate change

The commercial livestock sector is continually facing new challenges. Population growth, changes in diet, increase in income and urbanisation triggers a rise in demand for livestock products (Escarcha *et al.*, 2018). Meanwhile, climate change inflicts adverse effects on meat and milk

production, reproduction, feed crop and forage quality, water availability, and livestock diseases (Rojas-Downing *et al.*, 2017). Climate change places increasing pressure and challenges on the long-term sustainability of livestock production, since agriculture is highly dependent on the climate (Mwai *et al.*, 2015; Dzama, 2016). Livestock systems operate under various environmental conditions, thus they are increasingly affected by climate change (Escarcha *et al.*, 2018). The predicted temperature increases and scarcity of resources necessary for production causes many issues in the livestock sector. Heat stressed beef cattle could incur reduced feed intake, lower growth performance and poorer carcass quality, while dairy cattle a reduced milk yield (Dzama, 2016; Chingala *et al.*, 2017; Silpa *et al.*, 2021). Moreover, reproductive performance of both beef and dairy of both sexes will be compromised, with reduced conception rates, impairment of embryo development, increased calving intervals and impaired sperm quality (Dzama, 2016; Rojas-Downing *et al.*, 2017). The extreme weather patterns could also indirectly impact cattle production. The quality and the quantity of the pastures, grains, and crops could be reduced. Moreover, the quantity and spread of pests and parasites such as flies, ticks and intestinal worms could rise, as temperatures increase coupled with changes in rainfall affects the quantity and spread of pests and parasites (Dzama, 2016; Rojas-Downing *et al.*, 2017). Thus, both intensive and extensive livestock systems will be affected. In South Africa, smallholder farmers are vulnerable to the effects of ongoing climate change because they are found predominantly in tropical areas that are drought or flood prone and experience extreme temperatures (Morton, 2007). Moreover, these farmers have limited resources and suboptimal management programmes and as a result livestock may be predisposed to tick, intestinal nematodes, heat stress and various diseases.

With climate change impacting all aspects of agriculture, there is a need to breed robust animals (van Marle-Köster & Visser, 2018). Cattle are dispersed all over the world, inhabiting diverse areas with different environments and exhibiting different traits. Their adaptation to these various environments, as well as artificial selection and introgression with other breeds, brought about genetic and phenotypic variations in modern cattle breeds (Decker *et al.*, 2014). This opens new avenues of research and exploration of the cattle genome.

2.3 The cattle genome

The genome is a collection of biological information for the formation of an organism (Goldman & Landweber, 2016). All living organisms have a genome, which consist of chromosomes. The cattle genome is made up of 30 pairs of chromosomes, of which 29 pairs are autosomal and 1 pair is sex-linked (Bae *et al.*, 2010). The bovine genome contains at least 22 000 genes and has an estimated size of 2.87 Gbp (The Bovine Genome Sequencing and Analysis Consortium, 2009). The first draft sequencing of the cattle genome began in December 2003, based on blood derived DNA from a

Hereford dam (L1 Dominette 01449), a well-known beef breed (Burt, 2009). At the same time, several single nucleotide polymorphisms (SNPs) were generated from the genome sequence of six other cattle breeds. This, together with the sequenced L1 Dominette, established an invaluable resource for marker assisted selection (MAS) of important traits in the beef and dairy breeding programmes (Burt, 2009).

In 2009, a complete draft bovine genome sequence was published (The Bovine Genome Sequencing and Analysis Consortium, 2009). This was one of the first mammalian genomes to be sequenced, likely due to the importance of cattle as a nutritional source for humans (Tellam *et al.*, 2009). A single inbred Hereford cow and her sire were sequenced using whole genome shotgun sequencing as well as hierarchical sequencing, whereafter the data were assembled into two bovine genome assemblies; Btau, conducted at Baylor College (Liu *et al.*, 2009), and UMD2, from the University of Maryland (Zimin *et al.*, 2009). This led to the resequencing of several bovine genomes, to understand the underlying genetic mechanisms responsible for phenotypic differences. The first single cattle whole genome resequencing project was carried out in a Fleckvieh bull and more than 2 million new SNPs were detected (Eck *et al.*, 2009).

The bovine genome has been updated since 2009, with higher sequence coverage and better annotation. The cattle reference genome assembly has facilitated new possibilities of research into bovine genomics, especially in beef and dairy sectors, through genome-enabled selection (Bickhart *et al.*, 2020). This allows for selection of favourable alleles that affect quantitative production traits. Most recent cattle population CNV studies have used the UMD3.1 and the Btau4.1 reference genome assemblies. These assemblies offered a basis for many studies, but still brought about mis-assemblies, errors and assembly gaps (Bickhart *et al.*, 2020). Recently, a new reference genome assembly, ARS-UCD1.2, has been released and reported to have better accuracy and improved continuity by 200-fold (Rosen *et al.*, 2020). The reference genome is a very important tool for comparative studies, but it cannot be fully utilised in genomic research without the ability to efficiently perform a genome-wide scan to detect variation in this species (Matukumalli *et al.*, 2009).

2.4 Genetic variation and genomic technologies

Genetic diversity, which is described as a measure that quantifies the immensity of a population's genetic variability, is a crucial source of biodiversity and thus a key constituent for the occurrence of adaptation (Hughes *et al.*, 2008). It is important to identify which DNA variants influence the phenotype of an individual (Kijas *et al.*, 2011). Linking phenotypic data and genetic variation will allow genotype-phenotype associations to be discovered and provide valuable information, however, in order to figure out the genetic mechanisms behind the phenotypic differences between animals,

all forms of genetic variation need to be understood (Liu & Bickhart, 2012). Genomic variability has been reported to exist in many forms, ranging from single nucleotide polymorphisms (SNPs), transposable elements, variable number of tandem repeats (VNTRs) and structural variants (SVs) (Freeman *et al.*, 2006). Types of SVs include translocations, inversions, segmental duplications and CNVs (Feuk *et al.*, 2006). Molecular markers are used to measure genome-wide genetic variation within and between individuals and populations.

2.4.1 Molecular Markers

2.4.1.1 Variable Number Tandem Repeats

Variable number tandem repeats (VNTRs) are short nucleotide sequences (longer than 3) that are repeated several times in tandem (Ellenbroek & Youn, 2016). They are subdivided into micro- and minisatellites. Microsatellites, also referred to as simple sequence repeats, consist of short tandem repeats (1-6bp) of nucleotides. They can be classified as mono-, di-, tri-, tetra-, penta- or hexanucleotide repeats (Ellegren, 2004). Longer repeats are classified as minisatellites.

Microsatellites are evenly distributed throughout the genome, and identified within gene coding regions, non-gene sequences and introns (Ellegren, 2004). Due to their even distribution, abundance, small locus size, high level of polymorphism and co-dominant mode of Mendelian inheritance, this marker has been very popular (Liu & Cordes, 2004). Microsatellite markers have been used to detect the within and between breed genetic diversity of important Zimbabwean Sanga cattle breeds (Gororo *et al.*, 2018), and investigate the population structure and genetic diversity in South African Nguni cattle (Sanarana *et al.*, 2016). Although microsatellites have many advantages, SNP markers are believed to be superior because they are mutationally more stable, thus conforming more strictly to Mendelian patterns (Liu & Cordes, 2004).

2.4.1.2 Single Nucleotide Polymorphisms

Single nucleotide polymorphisms (SNPs) are base substitutions, involving only a single nucleotide, that appear at a low frequency in the genome (Figure 2.4) (Freeman *et al.*, 2006). These polymorphisms are caused by point mutations at a given position within a locus, and occur as transitions or transversions. Single nucleotide polymorphisms are bi-allelic markers, inherited co-dominantly, and represent the most abundant polymorphism in any organism (Liu & Cordes, 2004), occurring every 300 bp in *Bos indicus* cattle, and every 700 bp in *Bos taurus* cattle (Seidel, 2010). Application of gene chip technology in the late 1990s allowed the rapid genotyping of large numbers of SNPs, rendering it a powerful tool for SNP characterisation and genetic selection (Liu & Cordes,

2004; Seidel, 2010). Genome-wide association studies (GWAS) using SNP data were employed to find genetic variants in the cattle genome and associate them with complex traits. Dikmen *et al.* (2013) performed a GWAS to identify SNPs associated with rectal temperature, while Iso-Touru *et al.* (2016) conducted a GWAS for production traits, and found several SNPs associated with milk, fat and protein yields. Single nucleotide polymorphisms have many advantages and were initially thought to be the main source of genetic variation (Feuk *et al.*, 2006), and be responsible for most phenotypic variation (Freeman *et al.*, 2006). However, a considerable development has been made in understanding other types of variation, namely, genomic structural variation (Liu *et al.*, 2010). Structural variants, such as copy number variations, account for a considerable amount of genetic variation (Freeman *et al.*, 2006).

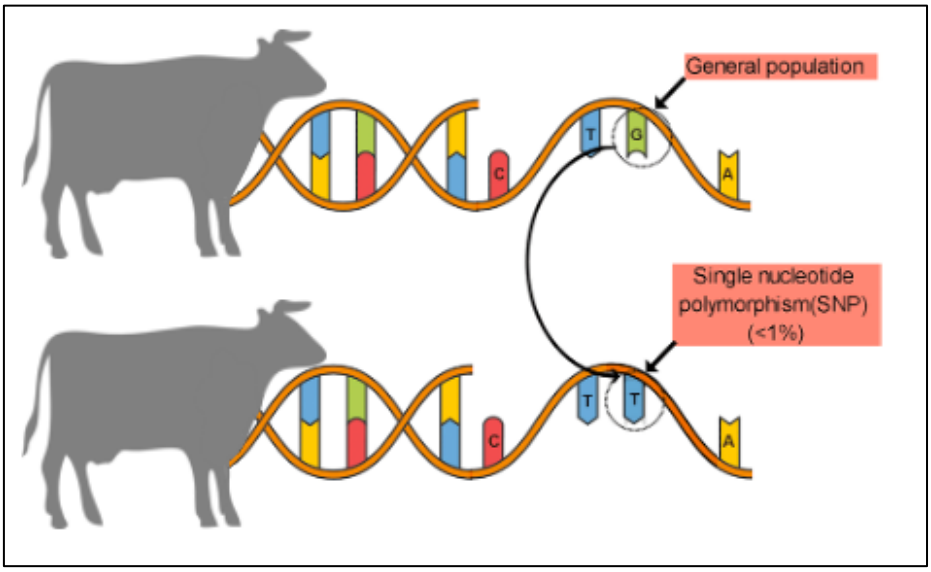

Figure 2.4: Single nucleotide polymorphism: a type of genetic variation that produces a new allele. Adapted from Lima *et al.* (2022).

2.4.2 Copy Number Variation

Copy number variations (CNVs) were reported as a novel form of genomic variation in 2004 (Iafrate *et al.*, 2004). In comparison to SNPs, CNVs influence a bigger portion of the genomic sequence and therefore have greater effects, such as alteration in gene dosage and structure, causing a change in gene regulation (Liu & Bickhart, 2012). Copy number variations refer to regions of the genome that differ in copy number in comparison to a reference genome (Redon *et al.*, 2006; Lauer & Gresham, 2019). They are structural variants (SVs) that range from 1 kilo base pair (Kbp) to several million base pairs (Mbp) comprising of duplications, deletions and insertions (Figure 2.5a) (Zhang *et al.*, 2009). Duplications are when a stretch of one or more nucleotides is replicated next to the original DNA sequence, deletions remove at least one nucleotide and insertions add at least one nucleotide, thereby changing the DNA structure. Two other classes of SVs, large-scale inversions and

translocations (Figure 2.5a), are classified as balanced events as their resulting copy number does not change in the affected individual, thus they can be particularly difficult to detect (Bickhart & Liu, 2014).

The merging of overlapping CNVs, identified in two or more different samples, generates copy number variable regions (CNVRs). The combining of CNVs into regions was first explained by Redon *et al.* (2006). The practice of combining CNVs into regions allows for population-wide CNV studies. CNVs can be compared on both the individual or population level (Zhou *et al.*, 2022). The development of genotyping arrays has enabled numerous studies to perform large-scale CNV analyses on different cattle breeds (Hou *et al.*, 2012a; Bagnato *et al.*, 2015; Zhou *et al.*, 2018b). Xu *et al.* (2016) revealed that CNVs can be used to investigate population genetics in cattle, and, furthermore, provided substantiation to support CNVs as genetic markers to be used in capturing the subspecies relationships and study diversity across the population. Most cattle CNVs affect genes that code for functions such as immunity and defence or receptor and signal recognition, and can, therefore, affect the functioning of the genes (Liu *et al.*, 2011). Studies indicate that cattle CNVRs contain between 300 and 1200 genes (Liu & Bickhart, 2012).

2.4.2.1 Mechanisms of CNV formation

Mechanisms for CNV formation include non-allelic Homologous Recombination (NAHR), non-homologous end joining (NHEJ), fork stalling and template switching (FOSTES) and retro-transposition (Figure 2.5b) (Gu *et al.*, 2008; Clop *et al.*, 2012; Bickhart & Liu, 2014). Non-allelic Homologous Recombination, one of the primary ways by which CNVs are formed, occurs during meiosis and mitosis. Copy number variations are generated when a segment of the genome, with high similarity to another (non-allelic) locus, cross over due to recombination (Figure 2.5b), leading to an increase in size of one chromosome (or a duplication) at the expense of another (Bickhart & Liu, 2014).

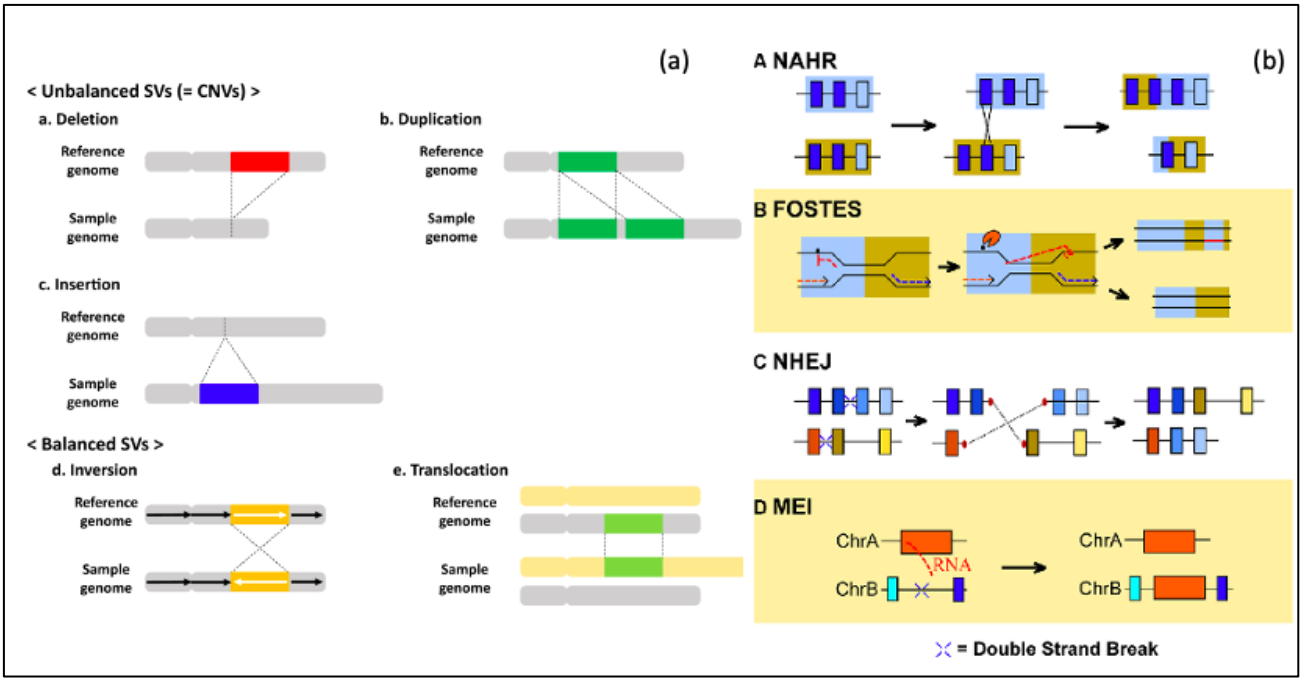

Figure 2.5: a) Different types of structural variation. From Nakatochi *et al*. (2021). b) Mechanisms of CNV formation. From Bickhart & Liu (2014).

2.4.2.2 Methods of detection

Methods used to detect CNVs include array-based approaches and sequence-based approaches (Liu & Bickhart, 2012). Array Comparative Genomic Hybridisation (CGH) and SNP arrays have been routinely used for CNV detection and their functioning widely reviewed (Lai *et al.*, 2005; Pinto *et al.*, 2012). More recently, next-generation sequencing (NGS) technologies, and their complementary analysis programs, have become available. Although NGS systems provide high coverage, resolution and accuracy in CNV detection, they are often very costly (Choi *et al.*, 2016). However, NGS platforms can be classified into second generation technologies (such as Illumina and Roche/454 sequencing) and third generation technologies (such as Oxford Nanopore and Pacific Biosciences sequencing), with the latter offering relatively longer read lengths at a reduced cost (Gatew & Tarekegn, 2018).

2.4.2.2.1 Array-based approaches

Array Comparative Genomic Hybridisation (CGH) can examine the whole genome in a single experiment by measuring the relative hybridisation intensity between the test and reference DNA samples and detect gains or losses (Figure 2.6a). Array CGH can be synthesised quickly and at a high uniformity and density, and be customised to target almost any area of interest (Liu & Bickhart, 2012). Liu *et al.* (2008) performed a CNV study on three different Holstein DNA samples using array

CGH and discovered 25 CNVs, while Kijas *et al.* (2011) discovered 109 CNV in nine cattle from three breeds, and Liu *et al.* (2010) identified 1 041 CNVs in 90 cattle from 17 breeds.

SNP genotyping arrays, such as the Illumina's Bovine SNP50 array (54 609 SNPs), Illumina's High-Density Bovine BeadChip Array (777 962 SNPs), or the Affymetrix Axiom Genome-Wide BOS 1 Array (648 874 SNPs), were initially designed for SNP genotyping, but their application has been expanded to include CNV detection (Rincon *et al.*, 2011; Liu & Bickhart, 2012). SNP arrays are advantageous because they can simultaneously measure the total signal intensities (Log R ratio, LRR) and the allelic intensity ratios (B allele frequency, BAF), which indicates allelic differences and copy number differences (Figure 2.6b) (Xu *et al.*, 2013a). The Log R ratio is a normalised measure of the total fluorescent intensity signal for two alleles at a SNP and the B allele frequency is a normalised measure of the relative ratio of the signals between two alleles at each SNP (de Araújo Lima & Wang, 2017). Several algorithms that use signal intensities (LRR and BAF) have been developed for CNV prediction from SNP data (Hou *et al.*, 2011a; Zhou *et al.*, 2018a). Research within the cattle community focuses on the use of SNPs for genome-wide studies, as well as to predict breeding values (Meuwissen, 2009), therefore SNP arrays are becoming the focal tool for research surrounding genetic variation in cattle (Hou *et al.*, 2011a).

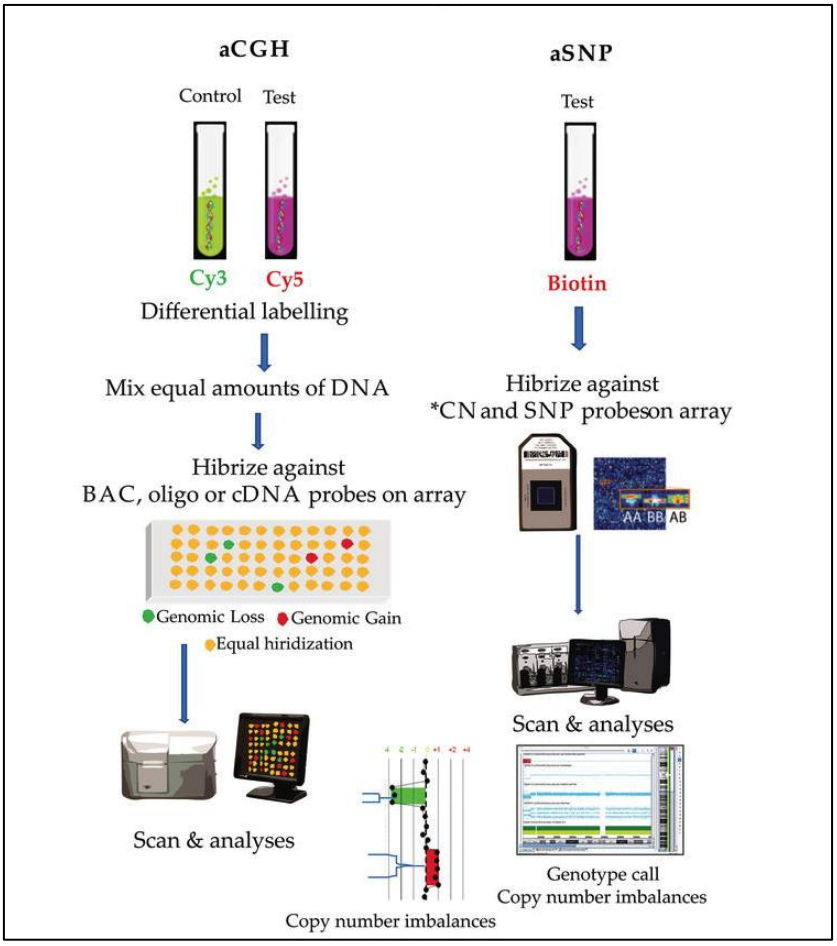

Figure 2.6: Basic concepts of chromosomal microarrays (CGH and SNP). Both technologies are able to detect sub-microscopic genomic variances such as CNVs. From Pinto *et al.* (2018).

A wide range of CNV detection software have been developed, such as PennCNV (Wang *et al.*, 2007), QuantiSNP (Colella *et al.*, 2007), Birdsuite (Korn *et al.*, 2008) and cnvPartition. Advantages and disadvantages of these algorithms have been reviewed and reported (Pinto *et al.*, 2012; Xu *et al.*, 2013a). Multiple algorithms should be used on a dataset to ensure the most reliable results (Winchester *et al.*, 2009). Jiang *et al.* (2012) detected CNVs in 2 047 Chinese Holsteins using three different programs, namely, PennCNV, GADA and cnvPartition. PennCNV is a Hidden Markov model based algorithm, cnvPartition is a Illumina BeadStudio plug-in software, and GADA is based on a Bayesian learning algorithm (Jiang *et al.*, 2012; Xu *et al.*, 2013a).

Both array platforms have their advantages and disadvantages, however, it is difficult to compare results across platforms. CGH arrays have improved sensitivity, resolution and signal to noise ratio compared to SNP arrays (Pinto *et al.*, 2012), but only report relative changes (Liu & Bickhart, 2012). Moreover, algorithms for CNV discovery from CGH arrays do not consider B allele frequency (BAF) information (Xu *et al.*, 2013a). Both array types have technical limitations, which could lead to several false negative and false positive calls, causing discrepancies between studies (Liu & Bickhart, 2012).

2.4.2.2.2 Sequence-based approaches

Next generation sequencing (NGS) is a rapidly advancing technology that has revolutionised DNA sequencing (Kumar *et al.*, 2019). NGS technology provides high coverage and accuracy and can significantly improve resolution in CNV detection (Choi *et al.*, 2016). NGS can simultaneously perform billions of individual sequencing reactions, referred to as massive parallel sequencing (Behjati & Tarpey, 2013). Next generation sequencing produces billions of bases of nucleotide sequences in short reads (Pirooznia *et al.*, 2015). Various enhanced strategies are available, all of which map the sequence reads to the reference genome, and then identify structural variants (Alkan *et al.*, 2011). There are four general tools available to systematically identify CNVs from sequencing data, namely, read depth, paired-end mapping /paired reads, split read and de novo assembly (Figure 2.7) (Liu & Bickhart, 2012). These CNV detection tools have been widely reviewed (Alkan *et al.*, 2011; Liu & Bickhart, 2012; Pirooznia *et al.*, 2015).

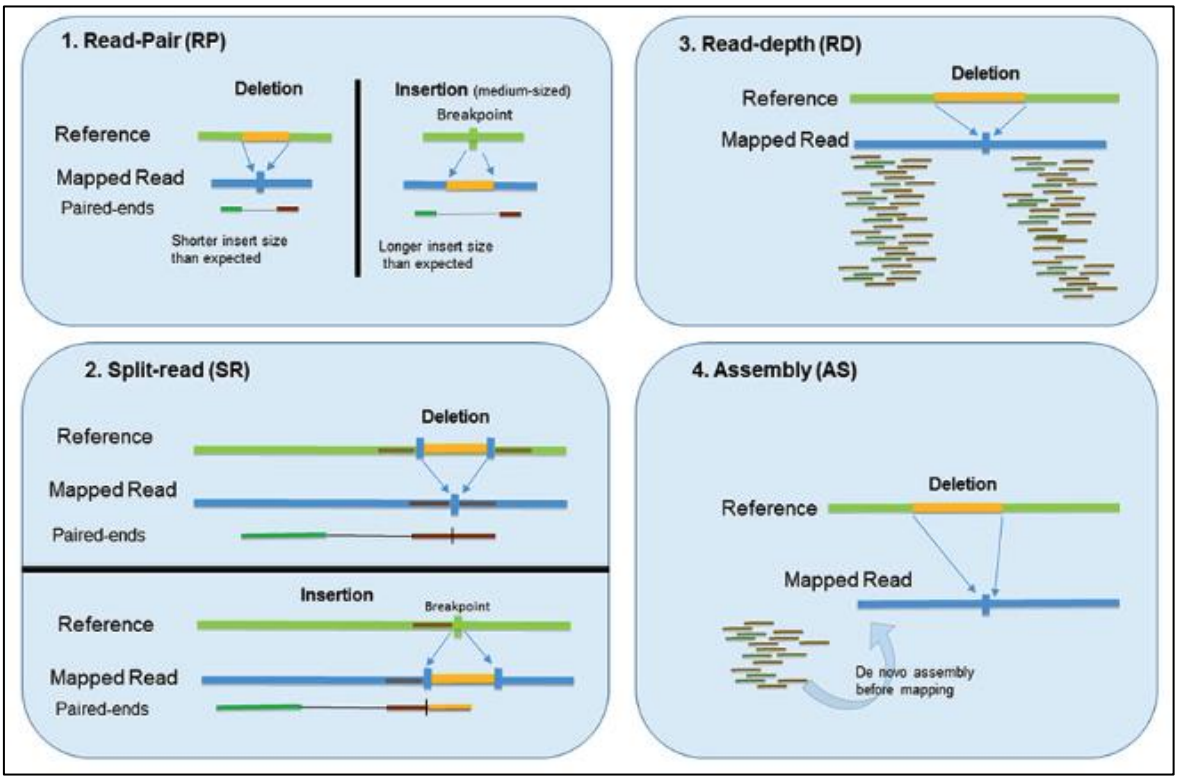

Figure 2.7: The four main tools available to systematically identify CNVs from sequencing data. From Pirooznia *et al.* (2015).

Bickhart *et al.* (2012) used the Read Depth approach of NGS technology to examine copy number differences among taurine and indicine cattle. The read depth strategy is widely used due to its efficiency, ease, and accuracy in obtaining the exact copy number. However, in livestock studies, the sequencing depth is often limited by the funding available, which in turn affects the read depth strategy to obtain these accurate copy numbers (Hu *et al.*, 2020b). An important consideration in

NGS-based research is the sequencing coverage, which is described as the average number of reads in the reconstructed sequence (Xu *et al.*, 2016). High-coverage sequencing (20X or more) can find many variants with high confidence, but it is costlier than low-coverage sequencing. For example, Stothard *et al.* (2011) sequenced one Angus and one Holstein bull at approximately 20X coverage and discovered 790 CNVs. In the same year Zhan *et al.* (2011) sequenced one Holstein at 15X coverage and discovered 520 CNVRs. Keel *et al.* (2016) used low-coverage sequencing (2.9X mean coverage per bull) to sequence 154 beef cattle and found only 57 CNVRs. More often, lower coverage sequencing is used for large sample sizes (Xu *et al.*, 2016). This technology could resolve multiple technical drawbacks that arise from array-based approaches, as NGS is not limited to only oligonucleotide probed regions (Liu & Bickhart, 2012). Moreover, it can detect several types of CNVs in a single test.

Although the application of NGS for genotyping in livestock was limited by high costs due to the advanced computational infrastructure needed to store and analyse data (Gurgul *et al.*, 2019), the numerous advantages of NGS compensated for the associated costs (Alkan *et al.*, 2011). The cost of sequencing the human genome has decreased over the last 20 years (Figure 2.8). Moreover, the advent of long-read sequencing technologies, also called third generation sequencing technologies, such as the Pacific Biosystems sequencer (PacBio) and Oxford Nanopore offer longer read length and a cheaper sequencing cost (Gatew & Tarekegn, 2018). The PacBio platform is proving to be invaluable in closing assembly gaps and repairing assembly errors (Bickhart & Liu, 2014), while nanopore sequencing, the most recent third generation technology, offers long reads, fast results, and lower costs (Gatew & Tarekegn, 2018). However, both long-read sequencing platforms have high error rates (Adewale, 2020). The CNVs identified from longer read PacBio sequencing technology has been compared to CNVs detected from Illumina sequencing (Couldrey *et al.*, 2017).

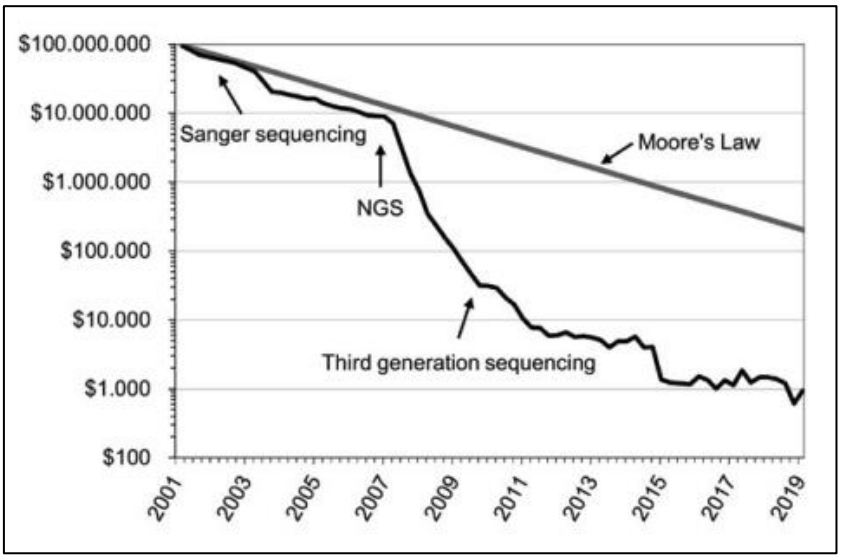

Figure 2.8: The decrease in sequencing costs from 2001 – 2019. From Furlani *et al.* (2021).

2.4.2.3 Variation in CNV Detection

Copy number variable fragments are ubiquitous in cattle genomes (Guo *et al.*, 2021), and determining the number of CNVs, and the functional consequences thereof is complicated. Moreover, the number of copies does not necessarily have a linear effect on the expression of a trait (Ben Sassi *et al.*, 2016). Although numerous studies have identified CNVs in various cattle breeds, there is often a significant difference in the number of CNVs discovered, making it difficult to compare studies and collate data. The number of CNVs detected is influenced by many variables, such as the methods used for detection, the sample group and size used, and the reference genome used.

2.4.2.3.1 Methodological

The variation in CNV discovery could be due to methodological reasons, such as the platform or algorithm used to identify CNVs (CGH vs. BovineSNP vs. sequencing). Not surprisingly, the increased resolution of an array platform increases the CNV count (Fadista *et al.*, 2010). Zhan *et al.* (2011) used three platforms to identify CNVs - sequencing, CGH arrays and SNP arrays - and found 520, 196 and 30 CNVs, respectively. There were large differences in the CNV discovery, but the genome coverage was similar. A low overlap of CNVs was detected across the different platforms, and this could be due to a few reasons. Each platform has different resolutions and size distributions, but CNV-seq has a higher coverage range than array platforms. Zhan *et al.* (2011) reported that all platforms have their advantages; CNV detection by sequencing allows for efficient detection of smaller CNV regions that are highly variable, whereas array platforms (CGH and SNP) are better at identifying the larger CNVs that have smaller copy number differences. An important factor to consider when examining the CNV output is that CNV counts are believed to be inversely proportional to their size (Fadista *et al.*, 2010). Research on Nellore cattle reported BTA1 displaying the highest number of CNVRs (692) with a mean length size of 46.68kb, whereas BTA25 displayed the lowest (131) with a mean length size of 79.37kb (Antunes de Lemos *et al.*, 2018b).

As mentioned, multiple algorithms should be used on a dataset to ensure the most reliable results (Winchester *et al.*, 2009). Jiang *et al.* (2012) used three algorithms to identify CNVs (PennCNV, GADA and cnvPartition), however, the next year, Jiang *et al.* (2013) used only PennCNV. Only one algorithm was utilised because using multiple algorithms results in only a small CNV output, since CNVs are only accepted if they are detected on all three algorithms. The researchers used a strict CNV calling criteria (a CNV had to comprise ten or more consecutive SNPs) to minimise the risk of false positives that could arise from using one algorithm. The stringency of CNV calling criteria differs between studies, and this too impacts the number of CNVs identified. For example, Jiang *et al.* (2013) inferred CNVs using a strict criterion (CNVs to contain 10 or more SNPs) as mentioned above, whereas Hou *et al.* (2012b) defined CNVs to contain three or more consecutive SNPs.

2.4.2.3.2 Sample Group size and composition

Differences in CNV output could be due to sample discrepancies, such as the number of animals or breeds used in the study (Hou *et al.*, 2011a). The size of the sample group is dependent on what will be tested, what the size of the project is, what animals are accessible, what methodology is going to be utilised, and what funding is available. This discrepancy causes CNVR output variation, given that CNVR count increases with a larger sample size (Antunes de Lemos *et al.*, 2018a). Sub-specie or breed also has an impact on the CNV output. It has been predicted (due to breed history and divergence) and reported that more CNV sites occur in *Bos indicus* cattle compared to European *Bos taurus* cattle (Liu *et al.*, 2010). Moreover, Kumar *et al.* (2021) observed that studies using a higher number of breeds had a higher CNVR count than studies using only one breed.

2.4.2.3.3 Reference Genome

The comparison of CNVs and CNVRs between studies is challenging because the mapping in each study is not necessarily based on the same reference genome (Fadista *et al.*, 2010). The UCSC Genome Browser LiftOver tool can be used to convert coordinates of variants from one genome assembly to another. Due to this fact, this tool has become popular due to reference genome assemblies continually being updated (Haeussler *et al.*, 2019; Luu *et al.*, 2020). Moreover, Zhou *et al.* (2022) reports SNP positions to differ between UMD3.1 and ARS-UCD1.2 maps of the SNP chip. They reported significant proportion of SNPs to be missing their location in UMD3.1 map file of the 150K BeadChip, and recommend the ARS-UCD1.2 reference genome be used for improved CNV detection. It is also important to note the risk of bias due to the reference genome if we are looking for CNVs in African cattle, since the primary reference genome stemmed from a European taurine breed (Figure 2.9). However, assemblies for two important African breeds, N'Dama and Ankole cattle, have recently been generated (Talenti *et al.*, 2022).

Figure 2.9: The first draft sequencing of the cattle genome was based on blood derived DNA from the L1 Dominette 01449 Hereford cow. From Rosen *et al.* (2020).

2.4.3 CNVs and important traits in the livestock industry

Copy number variations are linked to several complex traits and diseases in various cattle breeds (Huang *et al.*, 2014; Wang *et al.*, 2015; Da Silva *et al.*, 2016a; Upadhyay *et al.*, 2017; Liu *et al.*, 2020). This makes the study of these traits a very important undertaking, hence many advancements have been made in the field of molecular genetics. These advances have assisted in detecting several quantitative trait loci (QTL) regions and candidate genes that are found to be linked to important traits. Approximately 100 000 bovine QTLs for several traits have been described in the animal QTL database (Lee *et al.*, 2020). Genome-wide association studies were primarily focussed on SNP identification, but are now widely used for CNV discovery, to locate the genomic regions contributing to the genetic variation in phenotypic traits (Matukumalli *et al.*, 2009). This region can further be fine-mapped at a higher density to identify candidate genes responsible (Matukumalli *et al.*, 2009). Association studies have been used at large to record QTL of traits related to immunity (Pickering, 2017), growth and meat production (Da Silva *et al.*, 2016a; Zhou *et al.*, 2016a), milk production (Xu *et al.*, 2014a; Ben Sassi *et al.*, 2016), and feed efficiency (De Almeida Santana *et al.*, 2016). Therefore, CNVs such as deletions or duplications of a whole gene or a section of a gene, can disrupt gene expression, thus changing the phenotype (Lee *et al.*, 2020). Occurrences such as gene dosage modification, gene fusion, gene interruption, position effects (Guo *et al.*, 2021) and exposure of recessive alleles (Bickhart & Liu, 2014) can influence an organism's phenotype (Figure 2.10).

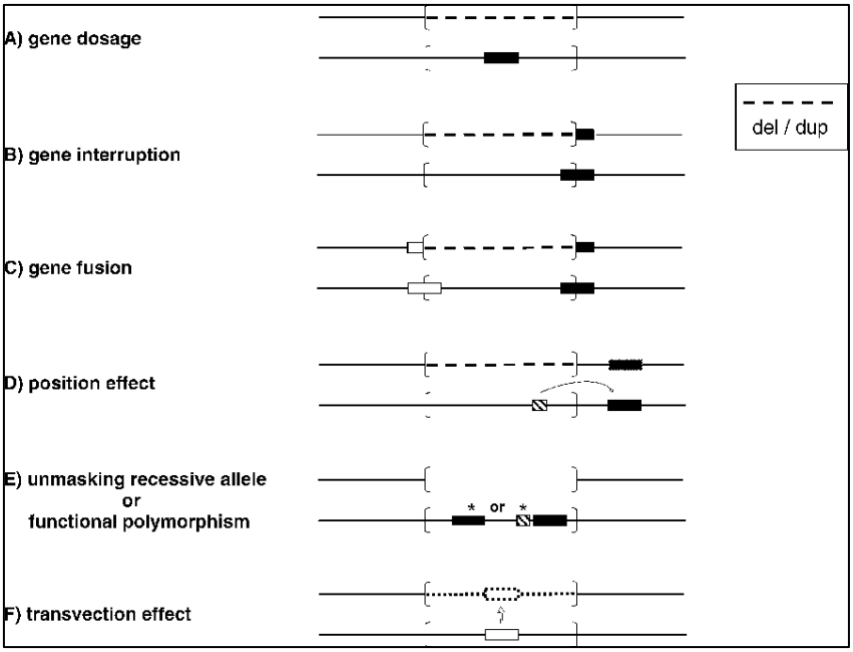

Figure 2.10: Molecular mechanisms showing how deletions or duplications influence gene expression and subsequent phenotype. From Lupski & Stankiewicz (2005).

2.4.3.1 Adaptation and immunity-related traits

Stress is crucial in all living organisms for homeostasis to be maintained. Animals cope with these stressors through mechanisms such as acclimation, acclimatisation, and adaptation (Archana, 2017). Acclimation and acclimatisation are involved in phenotypic change rather than genetic change; thus, the acclimation response goes once the stress is removed. Adaptation is the degree to which a population can become acclimatised to a range of diverse environments (Barker, 2009) involving physiological, morphological, anatomical and biochemical mechanisms (Key & Sneeringer, 2014). Adaptive fitness is distinguished by traits related to survival, health and reproduction (Keba *et al.*, 2010). Autochthonous breeds frequently exhibit notable adaptation to harsh conditions such as droughts, parasites, and disease (Wang *et al.*, 2015). In cattle, adaptation can be observed in the form of thermotolerance (Fernandes Júnior *et al.*, 2020), high altitude adaptation (Zhang *et al.*, 2020a; Guo *et al.*, 2021), and resistance to tick and nematode infections (Wang, 2016; Pickering, 2017).

Climate change is causing an increase in temperature, thus cattle are at risk of absorbing more heat than they are able to dissipate, resulting in heat stress (Mkize & Zishiri, 2020). The productivity and performance of animals is greatly influenced by age, nutrient availability, biomass productivity, water availability, and environmental conditions (Archana, 2017). The ability of an animal to acclimatise to fluctuating temperatures is essential for minimising heat stress. Zebu cattle show a higher heat tolerance compared to taurine cattle, which could be due to efficient heat dissipation strategies and hair follicle morphogenesis (Antunes de Lemos *et al.*, 2018a). Variation in copy number in heat shock

protein genes and transcription factor genes also play an important role in thermoregulation (Wang, 2016; Fernandes Júnior *et al.*, 2020). Heat Shock proteins (HSPs) are a special class of proteins that are important for the alleviation of heat stress in mammals. They are molecular chaperones that function by ensuring proper folding and refolding of proteins, instigating aggregations of proteins, and ensuring potentially damaging interactions are avoided (Mkize & Zishiri, 2020). The regulation of HSP production is crucial for cell survival (Basiricò *et al.*, 2011). There are many different types of HSPs studied in livestock, based on their molecular weight and functions, namely HSP110, HSP100, HSP90, HSP70, HSP60, HSP40, HSP10 and some smaller HSP families (Archana, 2017). The HSP70 is the most studied and reported to be the most abundant protein conferring thermotolerance (Basiricò *et al.*, 2011; Banerjee *et al.*, 2014).

Cattle with the dominant allele for the *slick* gene present a short and sleek hair coat, which is associated with an increased thermoregulatory ability to dissipate heat (Kava *et al.*, 2021). The *slick* haplotype causes short hair which promotes heat loss through conduction and convection by lessening the insulation in the hair coat (Berman, 2004), thus causing an increase in sensible heat loss. Originally, this gene was found in Senepol cattle, but its presence has since been shown in Holstein cattle. The Holstein breed is known to be a heat-sensitive cattle breed (Liu *et al.*, 2019a), thus studies have been done to identify biomarkers to use in breeding programs to develop climate-resilient cattle. A study done by Dikmen *et al.* (2014) confirmed that when exposed to heat stress, Holsteins with slick hair have a better thermoregulatory ability than those without. Moreover, the slick-haired cattle did not have such a large slump in milk yield in the hot summer months, thus showing key signs of resistance to heat stress. It is also suggested that this type of hair coat provides a natural protection to ectoparasites, increasing their tolerance against infestation (Ibelli *et al.*, 2012; Marufu, 2012).

In areas with a high altitude, such as the Tibetan Plateau with an average altitude above 4 000m, UV radiation is strong, temperatures are low and partial pressure of oxygen is reduced (Zhang *et al.*, 2020a). Populations native to these areas of high altitude have had to develop specific adaptive mechanisms to mitigate the environmental stress of hypoxia. Zhang *et al.* (2020) analysed the population structure and high-altitude adaptation of 25 Chinese indigenous cattle breeds. This study identified abundant CNV resources, which could be due to the diversity of distribution and environment, multiple origins of cattle and the lack of artificial selection.

In animal production, gastrointestinal nematode (GIN) infections pose a huge threat to the health and welfare of grazing animals (Sutherland & Leathwick, 2011). Infestations have subclinical and clinical effects on animals, and influences the economic and production gains of farmers (Gadberry & Powell, 2008; Mpetile *et al.*, 2017). As stated by Waller (1997), the global abundance of parasites found in the tropical or subtropical regions can be attributed to the optimal temperatures and rainfall

patterns which promote hatching and development of the parasites, resulting in death of young stock and therefore production losses. Almost all cattle carry a small number of worms, which could result in contaminated pastures and infected herd members. However, cattle younger than two years will have more internal parasites (Mwanza *et al.*, 2016) and are less resistant to large infection than adult cattle. Intestinal worms are mainly a disease of pastured cattle, therefore the need to control them will exist for as long as cattle are grazing pastures. Many countries have observed anthelmintic resistance in multiple species of gastrointestinal nematodes present in cattle, and due to the increasing number of cases reported in literature, anthelmintic resistance has, therefore, become an escalating concern (Sutherland & Leathwick, 2011). New approaches for treatment have been proposed such as state of the art vaccine application (Smith & Zarlenga, 2006), improved grazing management, biological control, and selective breeding to develop parasite-resistant animals (Waller, 1997; Hein *et al.*, 2001; Liu *et al.*, 2011). The genetic variability of the ruminant's immunity offers a viable way to control GIN infections without the use of anthelmintics (Sonstegard & Gasbarre, 2001). The immunobiology of GIN infections in cattle and the interactions between the parasite and the host immune system is very complex. Gasbarre *et al.* (2001) studied the distribution of GIN in cattle herds, this revealed host genetics to play a significant role in determining the immune state of the animal. There are a variety of manifestations of the immune response which results in superior herd immunity. Although the mechanisms behind these sorts of functional immunity is yet to be defined, GIN infections normally produce Th2-like (type II helper T cells) responses, which are characterised by elevated levels of cytokine Interleukin 4 (IL4), a multitude of mast cells, and elevated levels of antibodies (IgG1 and IgE) (Gasbarre *et al.*, 2001). However, different parasites stimulate different responses, and involve various regulated mechanisms.

Most cattle CNVs affect genes that code for functions such as immunity and defence or receptor and signal recognition and can therefore affect the functioning of genes (Liu *et al.*, 2011). Evidence has been provided that CNVs can act as genetic markers, which can be used to study diversity across populations and apprehend subspecies relationships (Xu *et al.*, 2016; Pierce *et al.*, 2018).

2.4.3.2 Meat Production and growth-related traits

In beef cattle, carcass traits and meat quality have a significant impact on the product pricing and consumer satisfaction. Carcass traits are related to the yield grade, carcass weight, backfat thickness, and rib eye area (Drake, 2004). The quality of the meat is based on organoleptic traits, namely, flavour, tenderness, colour and juiciness (Feitosa *et al.*, 2014). These traits are affected by the fat content, both subcutaneously and intramuscularly. Intramuscular fat (IMF) plays a major role in meat marbling, which determines the texture and flavour of the meat. Taurine breeds usually have better marbling scores than indicine cattle. Beef quality is also determined by breed, sex and age of

the animal, feeding and management schemes, and the genetic background of the animal (Figure 2.11).

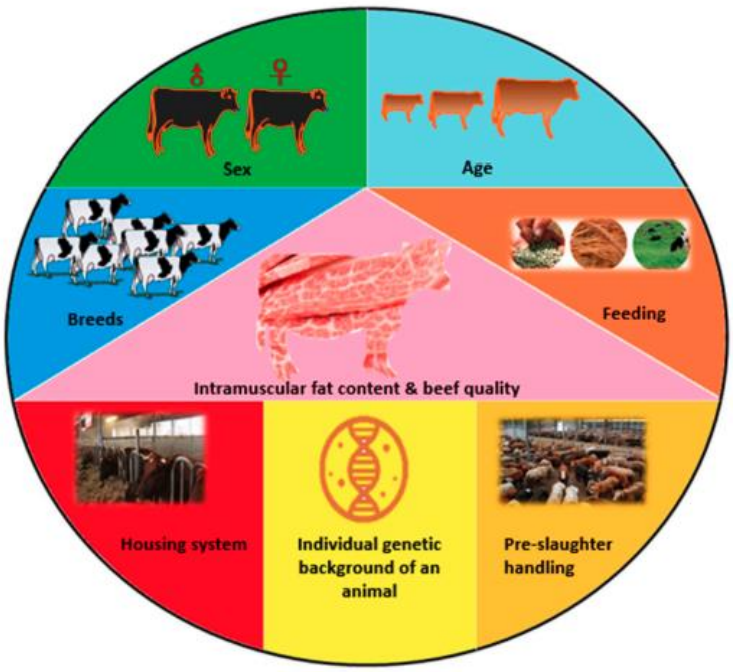

Figure 2.11: The determinants of meat quality in beef cattle. From Raza *et al.* (2019).

Growth traits, such as body weight measurements or visual scores of conformation, affect carcass selection and therefore beef production. Growth traits affect the development, structure and size of livestock (Guo *et al.*, 2020b). This complex, quantitative trait is controlled by nongenetic and genetic factors, however, genetic factors are main contributors to animal growth (Zhou *et al.*, 2016a; Guo *et al.*, 2020b). Zhang *et al.* (2018) reported a correlation between CNVs of guanylate binding protein 2 (GBP2) and growth traits in Chinese cattle. Genes such as the insulin-like growth factor-1 receptor (IGF1R) are also particularly important in embryo stages and growth after birth (Yang *et al.*, 2013). Average daily gain (ADG) is a performance measure that is closely monitored by beef producers as it is a very economically important trait. Xu *et al.* (2019) investigated ADG in Chinese Simmental cattle and found seven CNVs significantly associated with ADG, that were within or near several candidate genes. The meat tenderness phenotype, another economically important trait, was first studied in the 1920s, and is very well researched for various livestock species (Da Silva *et al.*, 2016a). Copy number variations have been linked in meat tenderness in Nellore cattle (Berton *et al.*, 2014). Meat tenderness is usually measured by shear force, a moderately heritable trait, thus selection for this trait could be used to improve tenderness (Zwambag *et al.*, 2013).

2.4.3.3 Milk Production and reproduction related traits

Milk production is an indispensable economic trait in cattle production. Globally, more than 6 billion people consume milk and milk products, most of whom live in developing countries (FAO, 2022). Milk production not only provides nutrition and food security, but is an important source of revenue for small scale farmers (Silpa *et al.*, 2021). In South Africa, over the past 30 years, the number of dairy farms have decreased, while the number of dairy cattle have more than doubled (Muller, 2017). This is due to the increased size of milking parlours allowing for a high throughput of cows. Moreover, computer programmes are able to collect important records such as milk yield or weight of the cow. This improvement has enabled dairy farmers to maximise milk yield. However, milk production is governed by environmental and, more importantly, genetic factors (Silpa *et al.*, 2021).

Milk production is polygenic; meaning that, it is affected by several genes (Yudin & Voevoda, 2015; Silpa *et al.*, 2021). Similarly, fertility is a complex trait affected by various factors (Muller *et al.*, 2018). Cow fertility traits are the most economically important traits affecting the productivity of the dairy industry (Liu *et al.*, 2017). However, there is an inverse, and therefore unfavourable, correlation between fertility and milk production (Liu *et al.*, 2008a). In Holstein cattle, milk production has increased while fertility has declined, likely because selection and breeding programmes in dairy herds focus on milk production and conformation traits, rather than fertility traits (Muller *et al.*, 2018). On dairy farms, the fertility of the cows affects both the financial sustainability and the genetic improvement of the herd.

Cows have two production phases, lactating or non-lactating. In a dairy herd, cows are the production unit, and the main source of income is milk, therefore, 80% of the cows should be lactating at any stage (Muller, 2017). To start a lactation period, cows must calve, whereafter they are in milk for approximately 300 days and then dried off to get ready for their next lactation. The introduction of artificial insemination (AI) as an Assisted Reproductive Technology to modern dairy herd management has improved reproductive efficiency (Valergakis *et al.*, 2007). In dairy farming, milk yield and milk composition are extremely important economic traits (Gao *et al.*, 2017; Nanaei *et al.*, 2020). The Holstein-Friesian is the highest producing dairy cattle breed in the world, largely owing to AI and use of semen collected from superior bulls (Seroussi *et al.*, 2010). Xu *et al.* (2014) performed a CNV-based GWAS for milk production traits in 26 362 Holstein cattle and identified 99 candidate CNVs, of which 34 were significantly associated with one (or more) milk production traits, namely, milk yield, protein yield, fat yield, fat percentage, and protein percentage.

Most dairy farms milk two or three times a day, therefore milking parlour management is very important for proper hygiene and udder care to reduce the incidence of health issues. Mastitis is one of the most common and costly diseases in dairy farming (Durán Aguilar *et al.*, 2017). Mastitis occurs

when microbes enter the teat canal and results in an inflammation reaction in the mammary gland of the cow. Mastitis leads to economic losses in the dairy industry due to decreases in milk yield, poor quality of milk and culling of animals (Cheng & Han, 2020). Efforts have been made to increase the resistance of cattle to mastitis, but low heritability of this trait has made it difficult. Correlated and measurable traits such as somatic cell count (SCC) can be an indicator of mastitis (Durán Aguilar *et al.*, 2017). Somatic cell count is the number of cells per millilitre of milk, therefore cows that are infected would have a higher SCC due to the immune response of white blood cells to the pathogen. Dairy farmers are financially penalised for high SCC in their herd. Szyda *et al.* (2019) aimed to identify the association of SNPs and CNVs with the occurrence of clinical mastitis. The study sequenced 32 Holstein cows and thereafter searched for genomic differences between mastitis resistant and mastitis-prone cows. A total of 191 CNVRs were found to be deleted in mastitis-prone cows and present in the mastitis resistant cows. These regions overlapped with important genes related to mastitis and the immune system. It was suggested that deleted genomic regions are more likely associated with an increase susceptibility to clinical mastitis (Szyda *et al.*, 2019).

Claw disorders (such as digital dermatitis, sole and toe ulcers or heel horn erosion) are the third main reason for culling in dairy cattle globally, after mastitis and reproductive issues (Heringstad *et al.*, 2018; Butty *et al.*, 2021). This is an expensive health issue and causes immense discomfort in cattle, thus being a major welfare concern. Butty *et al.* (2021) detected CNVs in 5 845 Holstein cattle to identify their association with hoof health traits and identified 14 regions that were significantly associated with hoof health traits, overlapping with 20 candidate genes.

2.4.3.4 Feed-efficiency traits

Feeding cattle is one of the main costs in both beef and dairy farming, and it has a huge effect on the profitability of the industry (De Almeida Santana *et al.*, 2016). Feed is the most variable cost in livestock production systems, thus making feed intake and efficiency very important economic traits (Sherman *et al.*, 2010). Sex, growth rate, worm burden, stress, nutrition, and feeding management all influence an animal's feed-use efficiency (Vickers & Stewart, 2019). Feed conversion ratio (FCR) is an important trait used to measure livestock production efficiency, by calculating the animal's capacity to convert consumed feed into the chosen output. de Almeida Santana *et al.* (2016) performed a GWAS to investigate the CNVs and genes involved in feed FCR in Nellore cattle. The researchers genotyped 2 253 animals and PennCNV identified 139 089 CNVs, merged into 2 667 CNVRs. The study identified many genes involved in lipid, energy, and protein metabolism. Residual feed intake and dry matter intake are other important traits related to feed efficiency and intake. Zhou *et al.* (2018b) performed a GWAS in 528 Holstein cows focusing on important production traits, such as feed intake related traits (residual feed intake and dry matter intake) and detected a few CNVs associated with these traits that overlapped with important genes.

An efficient animal can consume less feed, but still produce the same output, while maintaining body condition, health, and fertility (Houlahan *et al.*, 2021). The efficiency at which dairy cows convert feed to milk and beef cattle feed to meat influences farm costs and revenue, therefore the inclusion of feed efficiency traits in breeding programs is advantageous.

2.5 The systematic review

A systematic review is a research method whereby a researcher can synthesise the results of several relevant primary studies, using precise, systematic methods, to answer a certain research question (Lasserson *et al.*, 2022). When examining the hierarchy of evidence, systematic reviews and meta-analyses are at the top of the pyramid, indicating a high level of quality with a low volume of information (Figure 2.12). The Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guideline is a well-established, informative, and substantiated reporting guideline (Page *et al.*, 2021). The newly released PRISMA 2020 statement provides a detailed instruction guide to follow with explanations and elaborations, a checklist, and a flow diagram, to ensure reliable reporting (Page *et al.*, 2021).

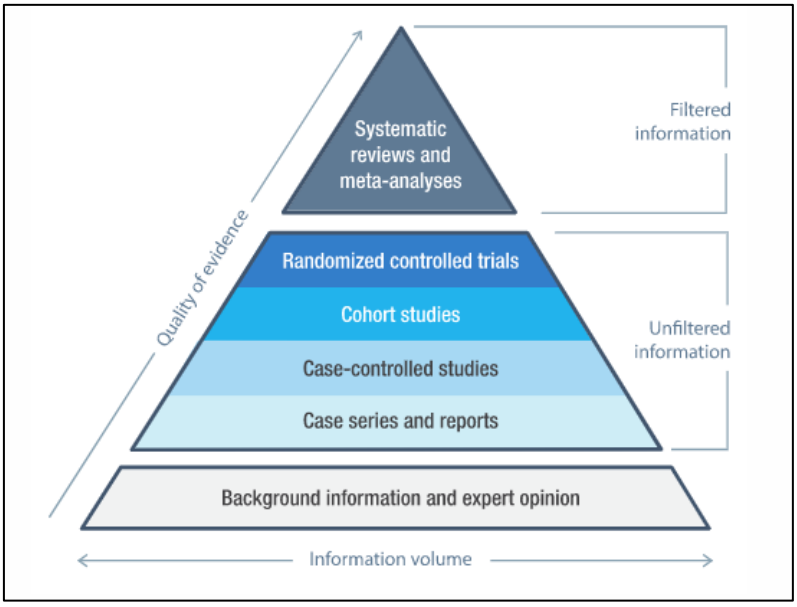

Figure 2.12: Hierarchy of evidence, with systematic review and meta analyses at the top, indicating a high level of quality with a low volume of information. Taken from openmd.com, Sep 2022.

A feature that differentiates a systematic review from a narrative review is the addition of a pre-specification criteria for including and excluding publications (Mckenzie *et al.*, 2019). When developing the protocol, research question, and eligibility criteria, a Population, Intervention, Comparison, Outcome (PICO) style approach is employed (Mckenzie *et al.*, 2019). Cattle CNV-related studies often have a cross-sectional study design. This type of study design can be

descriptive (when identifying the prevalence, experience, or incidence of a group) or analytic (when comparing different groups). This design aims to describe what is happening in a certain population as well as the frequency of that occurrence (Rezigalla, 2020), than testing experimental treatments and comparing them.

2.6 Conclusion

With climate change posing a huge threat to agricultural production, accompanied by the increase in world population, livestock farming is in jeopardy. The diversity of the cattle population provides researchers the opportunity to investigate the genetic reasons for this phenotypic variability. Encompassing many structural variations, CNVs are establishing their importance in genetic diversity. Cattle CNV research has developed considerably. The bovine reference genome is continuously being updated to have higher sequence coverage and better annotation, multiple methods and algorithms are available for usage, several genome wide CNV maps have been created, and numerous causal relationships have been established between CNVs and cattle phenotypes. Association studies linking CNVs and important phenotypes have focused on traits such as adaptation and immunity, growth, meat production, milk production, reproduction, and feed efficiency. Although this is a big step in the right direction, there are issues and challenges that have arisen. With the wave of CNV studies being reported, various data collection strategies and assessments are used, making it difficult to compare studies and collate important data. Nevertheless, CNVs are related to several complex traits and diseases in cattle, making the study of these structural variations, and their impact on important traits, a particularly important undertaking.

CHAPTER 3 METHODOLOGY

3.1 Introduction

A systematic review is a research method that aims to synthesise the results of several relevant primary studies, using precise, systematic methods, to answer a certain research question (Lasserson *et al.*, 2022). Primary research is growing at an ever-increasing rate, therefore this type of secondary research is of utmost importance, not only to keep researchers up to date with current and ongoing research, but also to point out where gaps in knowledge exist. The current chapter describes the data collection strategy to identify the role of copy number variations (CNVs) in cattle production. To ensure transparent and comprehensive reporting, this study follows an approach set out by the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines (Page *et al.*, 2021). This guideline was chosen for this review because it is well-established, informative, substantiated and provides a detailed instruction guide to follow. The PRISMA protocol also provides a 27-step checklist (Appendix A) and a flow diagram (Figure 3.1) to ensure reliable reporting.

This chapter provides information about the eligibility criteria, the data sources, the search approach, the selection and data collection process, and the strategy used to analyse the data.

Ethical approval was not required as no animal experimentation was done in this study.

3.2 Compilation of database

3.2.1 Eligibility criteria

A web-based literature approach was used to identify the scientific publications to be used in this study. A Population, Intervention, Comparison, Outcome (PICO) (Mckenzie *et al.*, 2019) style approach was employed. As previously mentioned, cattle CNV-related studies often have a cross-sectional study design, which aims to describe what is happening in a certain population as well as the frequency of that occurrence (Rezigalla, 2020), rather than testing experimental treatments and comparing them. Therefore, the *intervention* and *comparator* elements of PICO were not applicable for this study. The *population* describes the group of study participants the review will address, and their important characteristics and demography, such as age, gender and species (Aslam & Emmanuel, 2010; Mckenzie *et al.*, 2019). For this review, the population was outlined in a broad sense, with the intent that all possible studies could be considered. Moreover, to make comparisons between different population groups or subspecies. The population was defined as ‘cattle’ or ‘cow’ or ‘bull’ or ‘bovine’. The *outcome* describes what is being identified or measured in a study as well

as what could contribute to the study. For this review, the outcome was the identification of copy number variations (CNVs) and/or copy number variable regions (CNVRs), as well as gene content of these variable regions. The inclusion criteria is used to determine which studies will be included in the systematic review, while the exclusion criteria are the aspects of a study that make it ineligible for inclusion (Table 3.1).

Table 3.1: Inclusion and exclusion criteria for this review.

Inclusion criteria	Exclusion criteria
Studies that focus on cattle (of any age, breed, sex, region, or production type)[a]	Studies that focus on a different livestock species
Studies that identify CNVs and/or CNVRs, and state the number found	Studies that only identify SNPs, or do not state the number of CNVs found
Studies that identify CNVs at a genome-wide scale, using either an array or sequencing based approach	Studies that focus on a specific CNV genes or a specific chromosome encompassing a CNV
Original research articles, brief communications or conference abstracts	Review articles or novel methodology/software tool reports

[a]If a publication included 'eligible' and 'ineligible' participants (e.g. cattle and sheep), only data from the eligible participants were extracted.

3.2.2 Search strategy and databases consulted

As per the recommendations set out by Page *et al.* (2021) in the PRISMA guideline, the search strategy was peer reviewed by a Stellenbosch University librarian, Elizabeth Will-Mollard. The search string (CNV OR "copy number variants") AND (genome OR gene) AND (cattle OR cow OR bull OR bovine) was used in various online databases (Table 3.2).
Five databases were selected for the literature search. The databases accessed included:

Google Scholar - https://scholar.google.com/

PubMed - https://pubmed.ncbi.nlm.nih.gov/

Science direct - https://www.sciencedirect.com/

Scopus - https://www.scopus.com/home.uri

Web of Science - https://clarivate.com/webofsciencegroup/solutions/web-of-science/

It was decided that using all five databases would produce a broad and all-inclusive record of applicable studies. Although Google Scholar has a lower search precision, and does not have a peer review process, it is still a very comprehensive database and is useful for finding supplementary sources such as grey literature (Gusenbauer & Haddaway, 2020).

The search was limited to the period of 2000 – March 2022. The first draft sequencing of the cattle genome commenced in 2003 (Burt, 2009), making 2000 a suitable starting year, and the literature search was performed at the end of March 2022, thus it was the cut off date. The search string (CNV OR "copy number variants") AND (genome OR gene) AND (cattle OR cow OR bull OR bovine) was used for PubMed, Web of science and Scopus (Table 3.2). The search string was edited for two of the online databases, namely, Science direct and Google Scholar. This edit was done to narrow the search result output. Although a very narrow search string was utilised for Google Scholar, the search resulted in 1470 records. The search precision of Google Scholar is known to be lower than other search engines, therefore this was expected. The output yielded 'fuzzy' results after the first 100 records, thus, after discussion with a librarian, only the first 200 results were exported. A total of 1098 publications were exported to Mendeley for further evaluation (Table 3.2).

Table 3.2: Full search strategies for all databases, including the date of search.

Date	Search string	Google Scholar	PubMed	Science direct	Scopus	Web of science	Total
31-Mar-22	(cnv OR "copy number variants") AND (genome OR gene) AND (cattle OR cow OR bull OR bovine)		175			222	397
31-Mar-22	TITLE-ABS-KEY ((cnv OR "copy number variants") AND (genome OR gene) AND (cattle OR cow OR bull OR bovine))				189		189
31-Mar-22	(cnv OR "copy number variants") AND (genome OR gene) AND (cattle OR cow OR bull OR bovine) AND Adaptation			312			312
04-Apr-22	(cnv OR "copy number variants") AND "cattle" AND "Adaptation"	1470					200
Total							**1098**

3.2.3 Selection of publications

Following the literature search, publication selection began. The selection protocol is divided into three steps, identification, screening, and inclusion, based on the PRISMA protocol described by (Page *et al.*, 2021). The results of the search and selection process are shown with a flow diagram (Figure 3.1).

Identification: The identification step involves identifying publications from the internet sources (Table 3.2). A total of 1098 publication records were exported to Mendeley referencing management tool for evaluation. Thereafter, 490 records were removed, of which 464 were duplicate records and 26 were export errors (defective sources).

Screening: In the screening process, the publications were independently reviewed and assessed through the prespecified eligibility criteria (Table 3.1). An initial screening process was carried out by scanning the titles and abstracts of the 608 records, which resulted in 319 publications being excluded, as they were unrelated to the review question. A second screening process excluded a further 62 articles. A total of 227 articles remained for a detailed review of full text, whereafter a third screening process excluded a further 123 articles for various reasons (Appendix B). If it was unclear whether an article should be included in the review, it was settled through discussion and consensus with project supervisors.

Included: A total of 104 articles remained for qualitative synthesis (Appendix C). These studies were thoroughly examined to extract the necessary data.

This study aimed to include all eligible studies from all potential sources. Journal articles were the source of most of the data. This source has advantages such as ease of access and data extraction but are sometimes unavailable for full access and have a risk of biased reporting (Li *et al.*, 2019b). Conference abstracts were eligible for inclusion in this review, although they could contain limited information that may not be dependable, they can identify unpublished studies, and their inclusion results in more information, precision, and comprehensiveness (Scherer & Saldanha, 2019; Li *et al.*, 2019b). Doctoral and master's dissertations were eligible for inclusion in this review as these data sources are usually rigorously reviewed by academics. If a journal article had been published from a master's dissertation, these reports were collated. This is done so that the study was reported as one unit of interest, rather than multiple reports of the same study (Li *et al.*, 2019b). Doctoral dissertations usually include two or more different studies, published as separate journal articles, therefore, these were identified as separate reports.

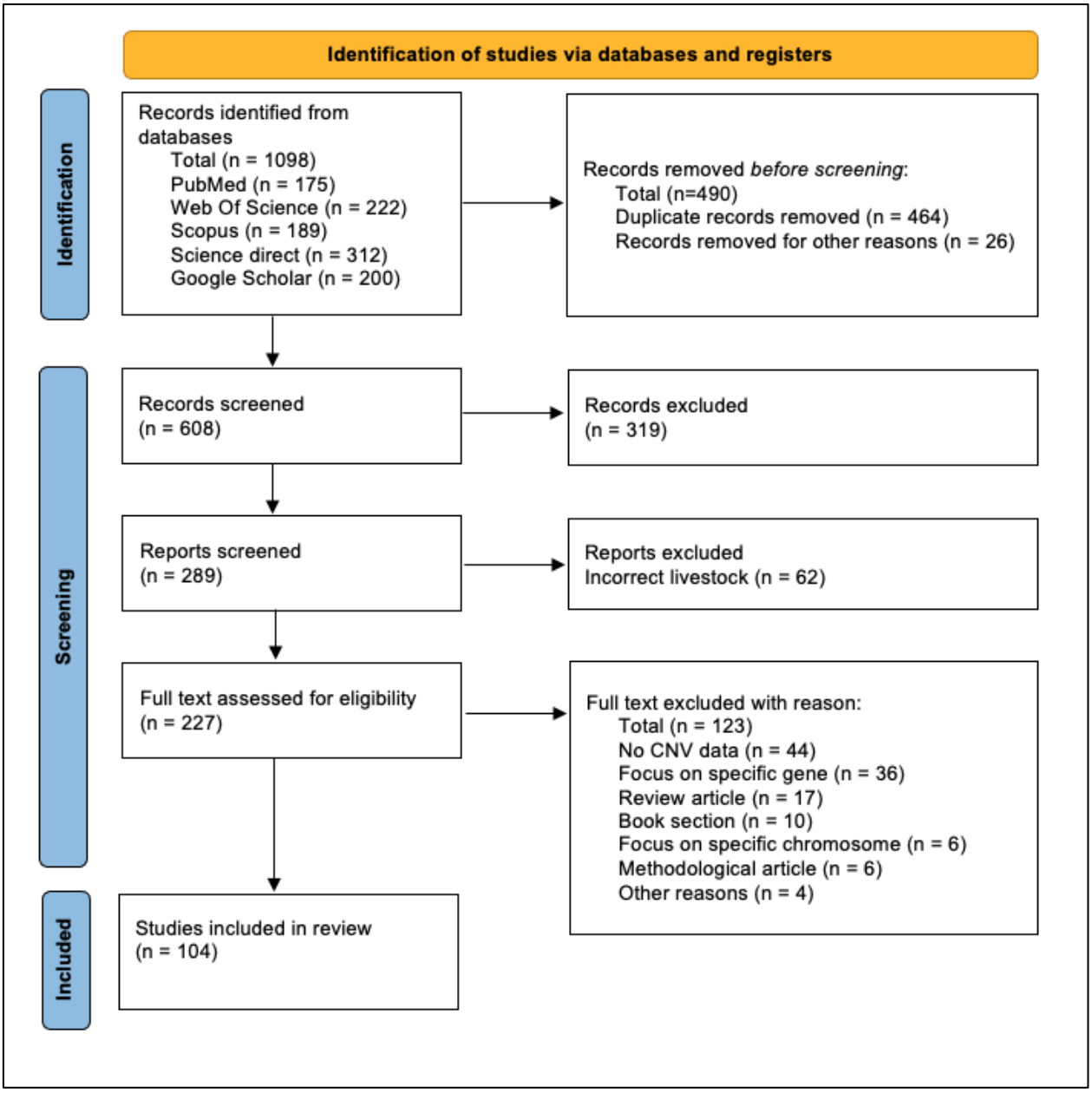

Figure 3.1: The PRISMA flow diagram. Adapted from Page *et al.* (2021).

3.3 Data extraction and transformation into datasets

For a systematic review, data extraction involves collecting all the data from the included studies pertaining to the participants, study methodology, interventions, outcomes, results, and other miscellaneous finds of importance (Li *et al.*, 2019b). The data extracted from each publication were combined on a database on Microsoft Excel for analysis. A data extraction template was compiled and trialled on a small number of articles before being completed. Data were extracted on specific study attributes of importance. These included year, region and journal of publication, bovine subspecies, breed, production type, study methodology, breed count, sample count, CNV count, CNVR count, reference genome used, important genes found, and for which important trait these

genes were responsible. When sufficient information was not supplied, supplementary data for each article was analysed and extracted. If supplementary data was not available, this was noted. Each original data source was assessed to ensure that unbiased analysis results be achieved. The aim, objectives, targeted populations, methods, and presentation of results were thoroughly evaluated. If uncertainties arose, this was settled through discussion and consensus with project supervisors.

3.3.1 Outcome domains

Data was extracted on specific study attributes of importance and tabulated in Microsoft Excel. The process used to select results and classify data is described below.

3.3.1.1 Publication details

The year, journal, and country of publication for each study was extracted. To determine the country affiliated with the publication, the region in which the ethical clearance was obtained for the study was used, and if unavailable, the country of the first author's associated university was used. If primary authors contributed equally to the work, but were from different regions, the region in which the DNA was from or the region where the animals were kept was considered. Tableau Desktop software (Professional Edition, version 2022.2.2) (Murray, 2013) was used to create a bar graph of the publications over the years, and to visually display the geographical distribution of the included studies.

3.3.1.2 Bovine Subspecies and breed

The breed of the animals used in each study was extracted and classified into a subspecies. This information was found in the 'methodology' chapter of each study. A publication focusing on a population from the subspecies *Bos indicus* was classified as 'Bos indicus'. A publication focusing on a population from the subspecies *Bos taurus* was classified as 'Bos taurus'. A publication focusing on both *Bos indicus* and *Bos taurus* populations was classified as 'Bos taurus & indicus'. Publications focusing on a population from the subspecies *Bos taurus africanus* was classified as 'Bos africanus'. The website Beef2Live was used to aid in the categorisation of subspecies of the various breeds (Beef2Live, 2022). Tableau Desktop software was used to visually display the proportions of bovine subspecies being studied by means of a pie chart as well as to map the dominant bovine subspecies per country.

3.3.1.3 Production type

The breed of the animals used in each study was further classified into production type. Publications that comprised milk production cattle were classified as 'Dairy'. Publications that comprised meat production cattle were classified as 'Beef'. Publications that comprised both meat production cattle

and milk production cattle were classified as 'Dairy & Beef'. Publications that comprised dual-purpose cattle (simultaneous dairy and beef production) were classified as 'Dual-purpose'. The Cattle Site website was used to aid in the categorisation of production type of the various breeds (The Cattle Site, 2022). Tableau Desktop software was used to visually display the proportions of cattle production type being studied by means of a pie chart as well as to map the dominant production type per country.

3.3.1.4 Methodology

To determine the platform used in each study, the 'materials and methods' chapter was examined. Studies that utilised Array Comparative Genomic Hybridization were classified as 'CGH'. Studies that utilised Illumina's Bovine SNP50 array were classified as 'SNP50'. Studies that utilised Illumina's High-Density Bovine BeadChip Array were classified and 'SNPHD'. If a study used a SNP chip that was not from the aforementioned arrays (such as the GeneSeek HD 77k BeadChip or the Affymetrix Axiom Genome-Wide BOS 1 Array) it was classified according to the SNP density it was more similar to. For example, the GeneSeek HD 77k BeadChip contains 76,999 SNPs thus it was classified as SNP50 (54,609 SNPs), whereas the Affymetrix Axiom Genome-Wide BOS 1 Array contains 648,315 SNPs, thus classified as SNPHD (777,962 SNPs). This was done to simplify the data characterisation. Studies that used whole genome sequencing (WGS) or whole exome sequencing (WES) based techniques were classified as 'Sequencing'. Tableau Desktop software was used to visually display the proportions of platforms being utilized by means of a pie chart, as well as to create a line graph and trend line graph to show the development of methodology of the included studies over the years.

3.3.1.5 Breed count and sample count

The breed count was classified as the total number of different cattle breeds each study used, while the sample count was the number of cattle (regardless of breed) used in the study. The sample tally was the number of cattle after quality control filtering.

3.3.1.6 CNV and CNVR count

The autosomal CNV and/or CNVR data were extracted from the publications. If a publication reported the total CNV number (autosomal + x), the supplementary data was consulted, and the autosomal CNV number was calculated. If supplementary data was not available or accessible, the total CNV number was reported, and this was noted on the excel sheet. If a publication used more than one software to identify CNVs, the consensus CNVs were extracted. The deletion- and gain CNVs were not extracted because many publications did not include this information. Similarly, the CNV- and CNVR lengths were not extracted due to missing data in many publications.

3.3.1.7 Reference genome

The mapping in each study was not based on the same reference genome, thus the reference genome used in each study was extracted. Although some studies use the UCSC Genome Browser LiftOver tool to convert coordinates of variants from one genome assembly to another, the original reference genome each study used was reported. The reference genomes used included UMD3.1, Btau_4.0 and ARS-UCD1.2.

3.3.1.8 Candidate genes

A second full text review was completed to analyse the CNV results produced from the included publications, and to identify the genes harboured in these CNV regions. Whereafter, the important candidate genes were extracted from each publication. Supplementary material was not consulted for gene extraction, only genes detailed in the results and/or discussion section of the study were extracted. This process was done to ensure that only genes of key importance be reported in this review. In order to extract genes efficiently and accurately, a machine learning technique on Alteryx software was trialled, but unfortunately it was unsuccessful. Thus, to ensure accuracy, a third full text review was completed to confirm that all the relevant and significant genes were extracted.

3.3.1.8.1 Transformation into datasets

The genes extracted from the publications were classified as either ‘adaptability and immune-related’, ‘milk production and fertility/reproduction related’, ‘meat production and growth-related’ or ‘feed-efficiency related’. If the included genes were related to another trait, such as coat colour or horn development, it was classified as ‘additional trait related’. These genes were then grouped according to their trait-related classification and reported on individual Excel worksheets in a separate Excel workbook (Figure 3.2). To ensure uniform gene nomenclature, each gene was subsequently searched on the National Centre for Biotechnology Information (NCBI) webpage for the identification of Gene ID aliases and gene families (NCBI, 2022). Because genes often hold several alternate names or are reported with a LOC symbol, this step was done to ensure uniform reporting of genes in each Excel worksheet, for accurate analysis. To prepare the data for analysis, the base data were extracted from the source worksheet, standardised, and cleaned to remove irrelevant characters (such as spaces and commas). The data were placed into a Tableau software for analysis and graphical representations of frequency of genes. The most frequent genes in each of the classification groups were reported and thoroughly discussed.

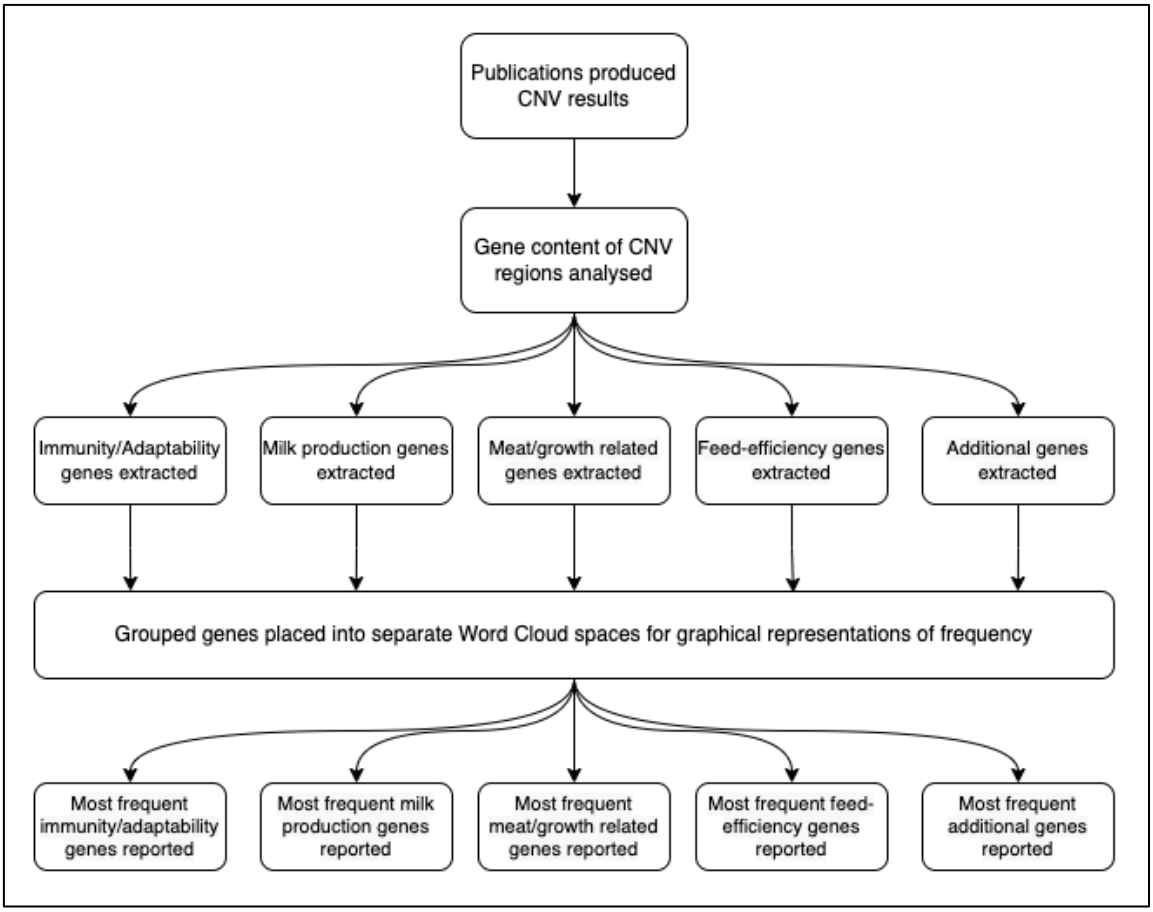

Figure 3.2: The extraction process used in this review.

3.3.2 Risk of bias assessment

Risk of bias refers to the possibility of the study results deviating from the truth, due to methodological faults in study design, conduct or analysis. Bias refers to a systematic error rather than a random error. There are two aspects of risk of bias that needs to be considered. Firstly, risk of bias is an important consideration in the results of the separate publications that are to be included in the review, and secondly, risk of bias is an important consideration in the outcomes of a review synthesis due to missing publications, or missing results within publications (Boutron *et al.*, 2019). Bias mitigation strategies were established for systematic reviews for health care studies, because of the high risk of bias as a result of the large number of people involved in these studies. Moreover, because the accuracy in medical care systematic reviews is of utmost importance as the findings are applied in evidence-based medical trials and could impact a patient's treatment regime and consequently their well-being. Cattle CNV-based studies often have a cross-sectional study design and are therefore less prone to bias reporting. Moreover, the reporting of methodology in agricultural science studies offers transparency and clarity to the research. In this review, an in-depth analysis of the important CNV-related genes was completed, therefore, a risk of bias assessment was not necessary.

3.4 Characterisation of data

Domains extracted from the publications used to characterise the dataset, such as the year, region and journal of publication, bovine subspecies, breed, production type, and study methodology were utilised for the narrative aspect of this review (Chapter 4). The data were prepared and sorted in Excel worksheets, whereafter it was transferred to Tableau software for visualisation.

CHAPTER 4 CHARACTERISATION OF DATA

4.1 Publication details

4.1.1 Year and journal of publication

The earliest study included in this review was published in 2008 (Figure 4.1) (Liu *et al.*, 2008b). CNV research has developed considerably since 2008 (Liu & Bickhart, 2012). In 2008 and 2009 only one study was published in each year, whereas in 2010, four studies were published. This big jump is likely due to the publishing of the complete draft bovine genome sequence in 2009 (Elsik *et al.*, 2009). This was one of the first mammalian genomes to be sequenced, owing to the importance of cattle as a nutritional source for humans (Tellam *et al.*, 2009), which then led to the resequencing of several bovine genomes. Due to the growing importance of understanding the cattle genome for genetic improvement of livestock, there has been a steady increase in the number of CNV-focused publications. The number of studies identified as relevant to this review has increased from an average of four per year (in 2008 to 2014) to an average of ten per year (in 2015 to 2021). This growing importance triggered advances in the efficiency of detection algorithms and software for cattle CNV-identification. Moreover, the rapid decrease in sequencing costs spurred efforts to advance cattle genome knowledge (Bickhart *et al.*, 2020).

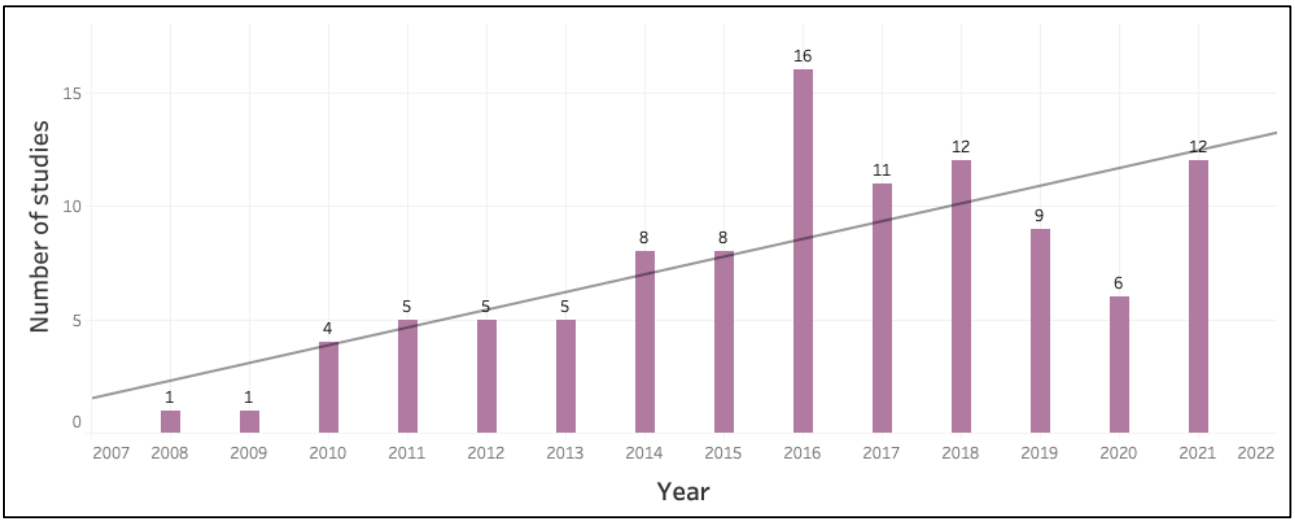

Figure 4.1: The year of publication of included studies from 2008-2021. The positive trend line indicates the increase in cattle CNV studies.

A drop in the included publications can be seen in 2020, almost certainly due to the coronavirus disease 2019 (COVID-19) pandemic. During this time, resources were allocated to COVID-19

research, causing non-COVID-19-related studies to be delayed or suspended, and research grants be diverted (Riccaboni & Verginer, 2022). Shan *et al.* (2020) investigated publication trends in journals during the COVID-19 pandemic and found that non-COVID-19-related studies decreased in volume as COVID-19-related studies increased concurrently. The data from 2022 has been excluded from the table because it does not represent a full year of publications and could falsely skew the output.

The studies included in this review were published in a total of 33 journals (Figure 4.2). The journal that published the most studies was “BMC Genomics” (24 publications), followed by “PLoS ONE” (7 publications), “Livestock Science” (6 publications) and “Frontiers in Genetics” (6 publications).

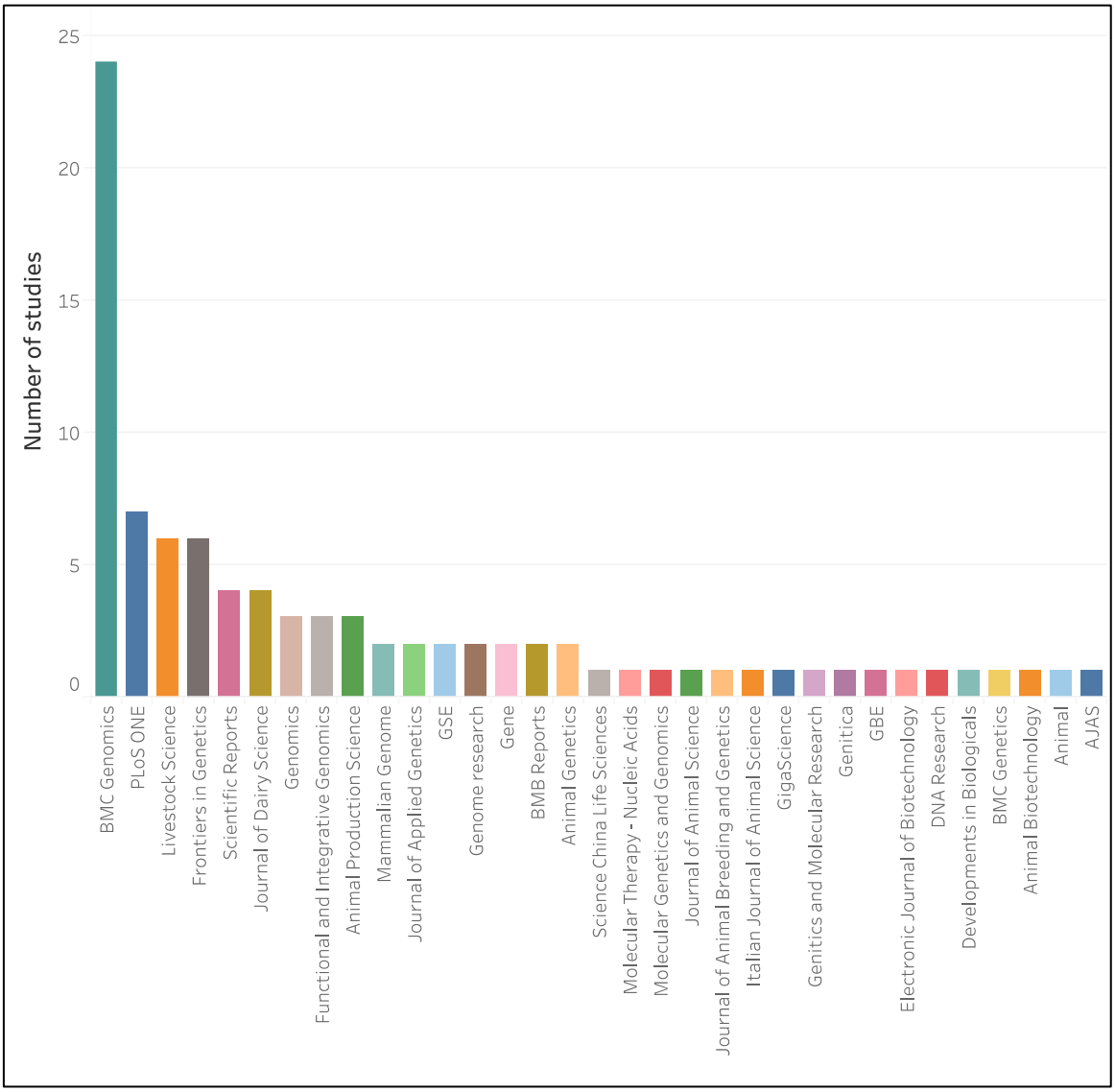

Figure 4.2: The journal of publication of the included studies. The journal that published the most studies was BMC Genomics, followed by PLoS ONE, Livestock Science, and Frontiers in Genetics.

4.1.2 Country of publication

Research related to the review topic was done in 25 countries. Figure 4.3 displays the geographical distribution of the included studies. China produced the most publications (21), followed by the United States of America (USA) (20) and Brazil (12). These results are understandable considering all three of these countries are of the top 5 world's largest milk and meat producers (FAO, 2022), therefore considerable research is undertaken within these countries. Particularly in China, several studies have investigated genes related to growth traits in indigenous Chinese cattle breeds (Cheng *et al.*, 2019a; Guo *et al.*, 2020a; Hao *et al.*, 2020; Hu *et al.*, 2022; Huang *et al.*, 2022; Liang *et al.*, 2022). Moreover, China is rich in genomic resources, with 28 recognised breeds (Zhang *et al.*, 2015b). When analysing the cattle livestock counts (total number of live animals at a given time) for each country, Brazil has approximately 218.15 million cattle, followed by India, USA, Ethiopia, and China (FAO, 2022) (Figure 4.4).

Only one study identified as relevant to this review was published in India (Kumar *et al.*, 2021). This is interesting considering India has the second highest number of cattle (194.48 million cattle) and are the top milk producers globally (FAO, 2022). However, almost 50% of milk production is contributed by indigenous buffaloes (DAHD, 2021), which have a head count of 109.72 million (FAO, 2022), and buffalo data was not included in this review (Section 3.2.1). Livestock production in India used to be low, due to unorganised breeding programmes and lacking nutrition, but efforts were made by the government to improve production by cross breeding indigenous cattle with high-yielding exotic breeds and upgrading buffalo production (Saravanan *et al.*, 2022). Indigenous cattle have an average milk yield of 3.9 kg/cow/day which is comparatively low compared to exotic cattle (11.88 kg/cow/day) and indigenous buffalo (6.43 kg/milking buffalo/day) (DAHD, 2021). Although livestock genetics has improved in India, it is still lagging developed nations (Saravanan *et al.*, 2022). This was evident in this review, not just because only one study was published, but due to the study, published in 2021, was the first study to attempt to create a CNV atlas for Indian Thanparkar cattle (Kumar *et al.*, 2021). Very few studies relevant to this review have been published in Africa, even though there are tens of millions of heads of cattle in Africa (Figure 4.4) and over 150 different breeds (Wang *et al.*, 2016c). This could be attributed to the fact that African indigenous cattle are comparatively not as intensively studied at the genomic level as other cattle populations (Mwai *et al.*, 2015). Moreover, the implementation of CNV studies in Africa pose several challenges, such as limited resources and finances, practical sampling complications, transportation, storage difficulties of samples from remote areas, or minimal breeding and management records (Wang *et al.*, 2016c).

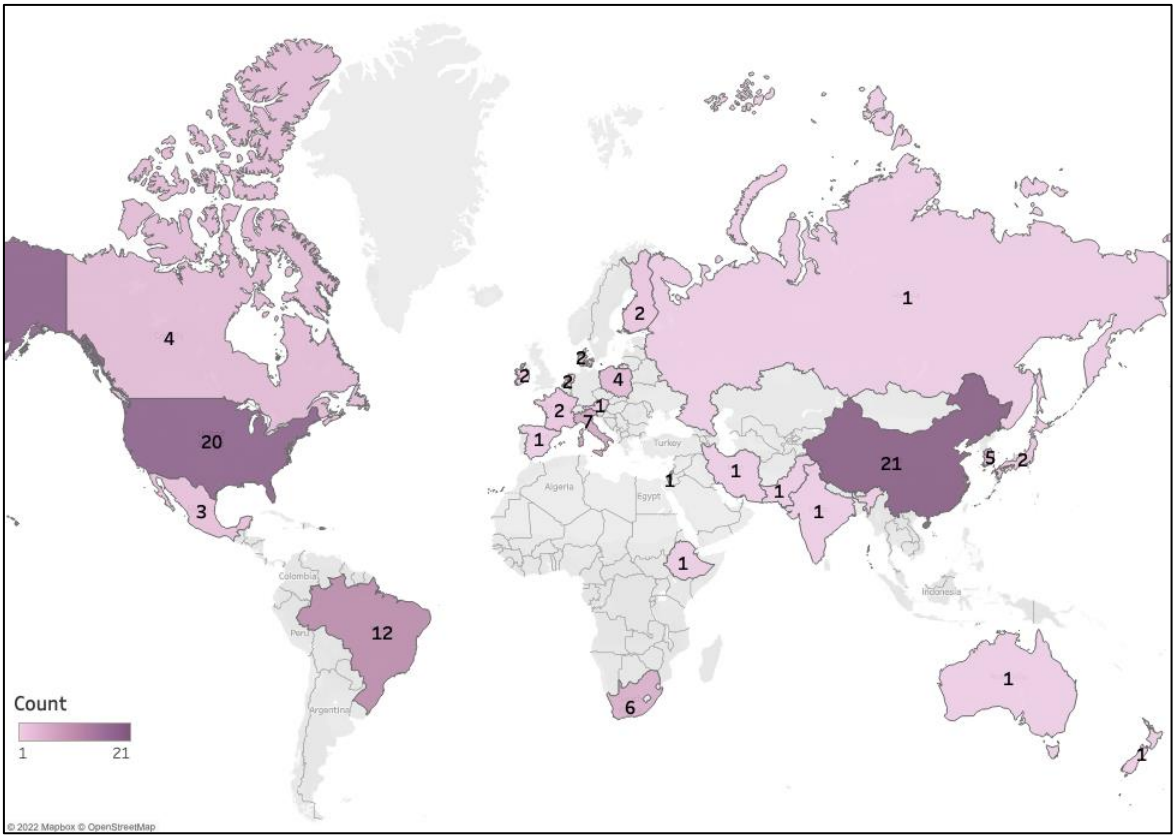

Figure 4.3: The Country- and number of publications of the included studies. China produced the most publications followed by the USA and Brazil.

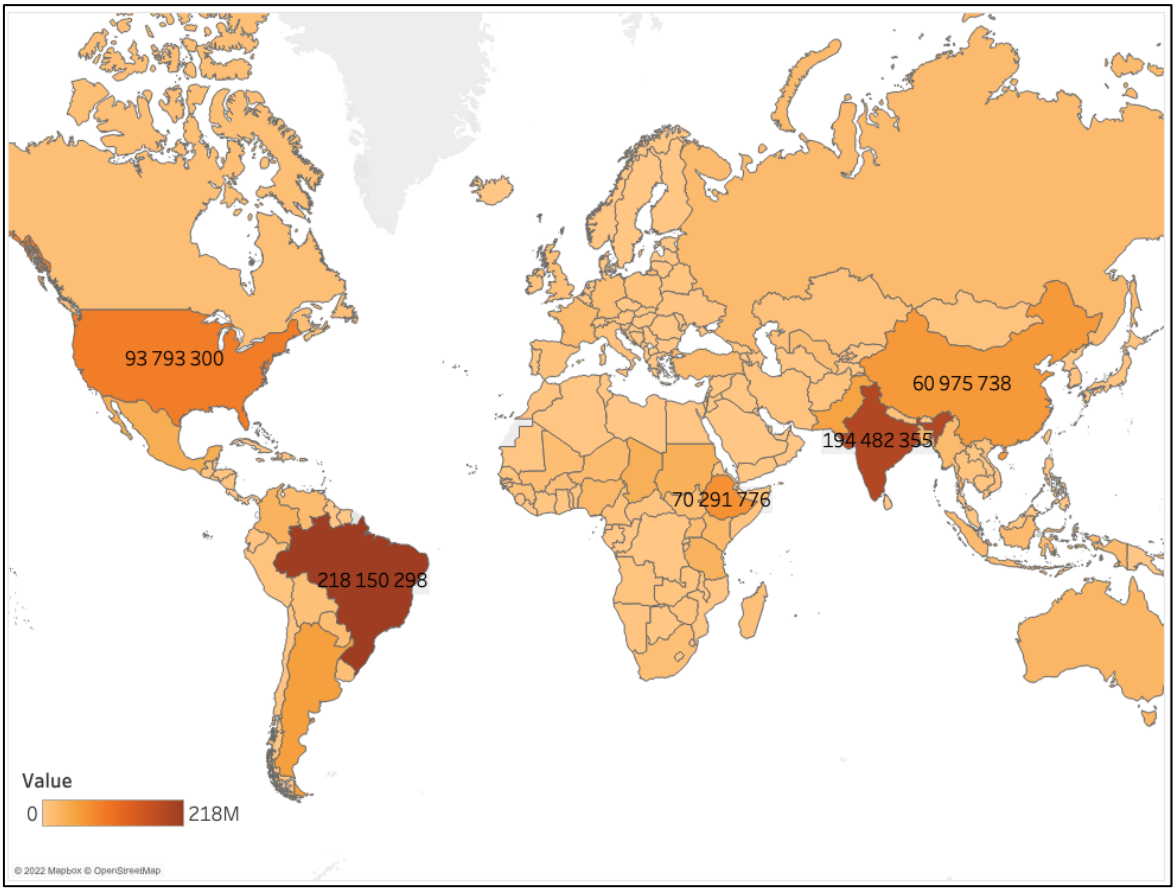

Figure 4.4: Global cattle counts, highlighting the countries with the highest head count.

4.2 Population details

4.2.1 Subspecies

Cattle are classified into two main groups, namely, Bos *taurus* (taurine) and *Bos indicus* (zebu) (Burt, 2009). Sanga-type cattle, such as the Nguni are believed to comprise a cross between taurine and zebu cattle (Rewe *et al.*, 2009; Gororo *et al.*, 2018), and are identified as the subspecies *Bos taurus africanus* (Strydom *et al.*, 2001). The population break down of the included studies (focusing on bovine subspecies) can be seen in Figure 4.5. Over 60% of the publications studied *Bos taurus* cattle, while only 12% studied *Bos indicus* cattle. This high percentage of taurine-focused studies could be due to several factors. European breeds dominate the current genetic resources sequenced and studied (Talenti *et al.*, 2022). This is substantiated by the fact that the primary reference genome was stemmed from a European taurine breed (Burt, 2009), along with the design of the first high-throughput microarrays of SNP markers (Utsunomiya *et al.*, 2019), and projects (like the 1000 bulls genome project) being skewed towards taurine cattle likely due to ease of sample accessibility and geographical distribution of cattle (Talenti *et al.*, 2022).

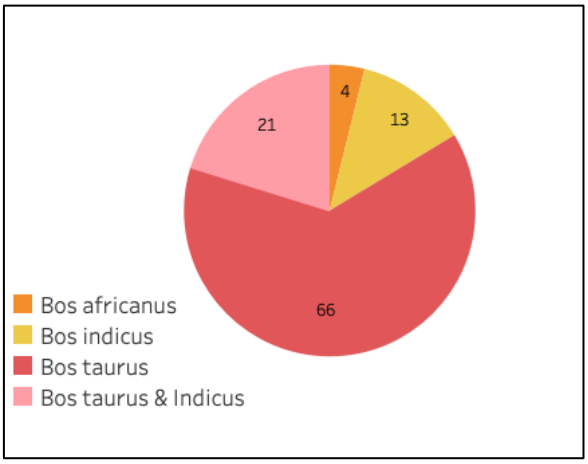

Figure 4.5: Pie chart showing count of studies by publication subspecies-focus.

The low percentage of zebu-focused studies could be because zebu cattle inhabit south and southwest Asia (for example, India has 41 breeds that are of *Bos indicus* origin) but when examining Figure 4.3, very few publications came from this region, thus causing this low percentage. Similarly, the small percentage of *Bos taurus africanus* publications could be a result of a small number of studies identified as relevant to this review have been published in Africa, and this continent is home to Sanga cattle (Mwai *et al.*, 2015). Moreover, these indigenous breeds have not been well characterised or described, and rarely undergo structured breeding programmes (Nyamushamba *et al.*, 2017). Considering genomic data from zebu breeds are limited compared to taurine breeds, this result can be expected (Saravanan *et al.*, 2022).

Figure 4.6 depicts the dominant bovine sub-specie for each country of the included publications. Zebu cattle are not predominant in modern-day Europe (Utsunomiya *et al.*, 2019), thus the results are understandable. Publications focusing on taurine cattle (red) predominate in the temperate zone, whereas publications focusing on zebu cattle (yellow) predominate in the subtropical and tropical zones. This is because zebu cattle exhibit characteristics such as heat tolerance, parasite resistance and disease resilience and are therefore able to adapt to the tropical environments (Utsunomiya *et al.*, 2019) while taurine cattle do not generally demonstrate these resilient characteristics. When considering the divergent history of these subspecies, it is believed taurine cattle were domesticated in Fertile Crescents (present-day southern Iraq, Lebanon, Syria, Israel, and Jordan), and made their way through Turkey into northern Italy, subsequently dispersing around Europe, they may have also made their way along the northern coast of Africa and crossed over to the Iberian Peninsula (Pitt *et al.*, 2019) (See Figure 2.6). Whereas zebu cattle were domesticated in the Indus Valley (present-day Pakistan) and migrated into Africa, East Asia, and southwestern Asia. Zebu cattle reached the Americas more recently, with most of the animals being directly imported from India (Utsunomiya *et al.*, 2019).

Figure 4.6 shows that publications from Brazil and India focussed on *Bos indicus* cattle (yellow), these results are expected considering Brazil has the largest cattle herd in the world, followed by India, and the bulk of the cattle found in these countries are *Bos indicus* cattle, or carry *Bos indicus* ancestry (Utsunomiya *et al.*, 2019).The studies concentrating on *Bos taurus africanus* cattle (orange) were all published in South Africa, as these studies focus on Nguni cattle. Nguni cattle exhibit excellent adaptation characteristics to the harsh environmental conditions in Southern Africa, thus there is much interest in this breed is (Wang, 2016).

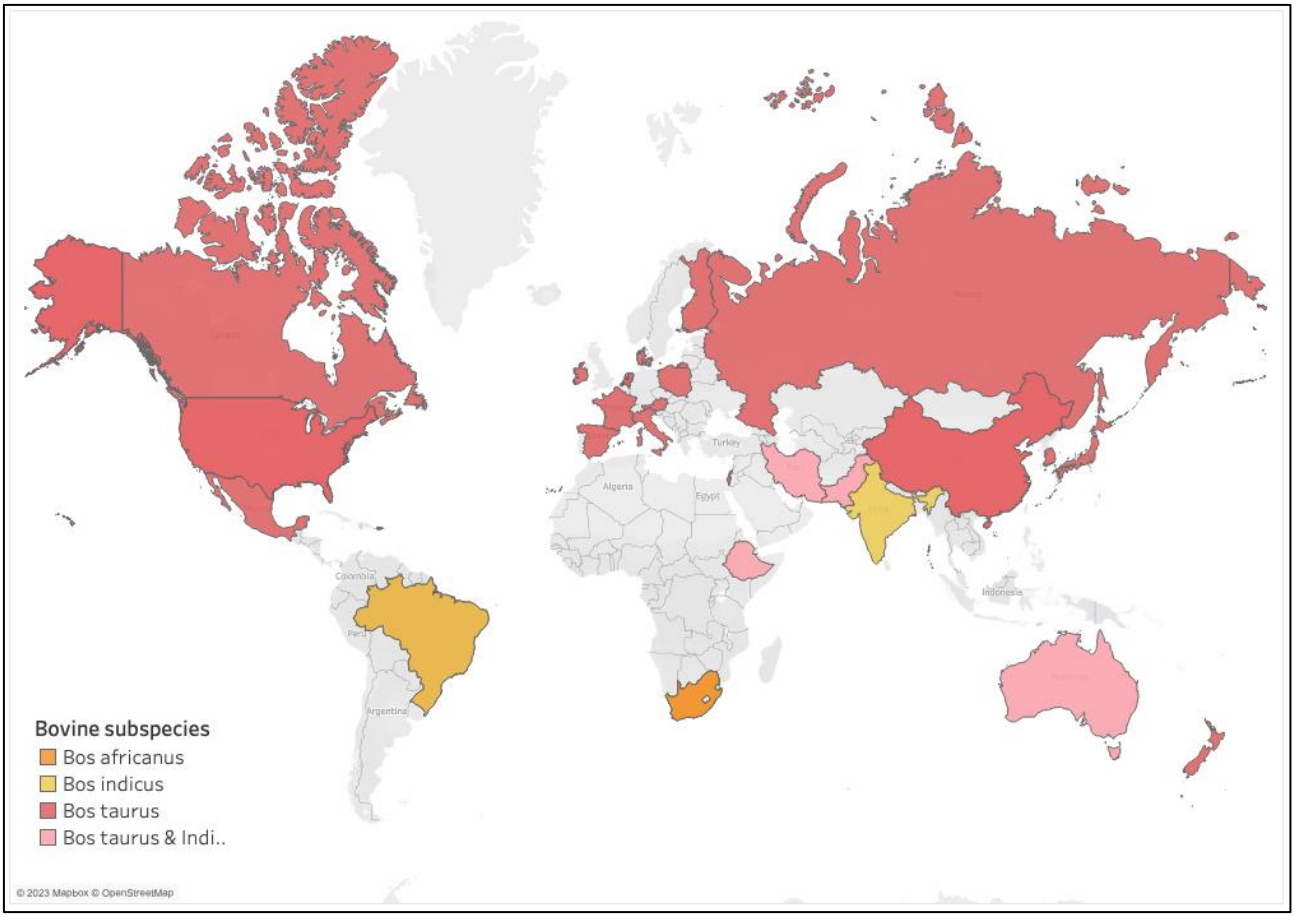

Figure 4.6: Dominant bovine subspecies in the data, mapped per country, with taurine cattle studies predominating in the temperate zone, and zebu cattle studies predominating in the subtropical and tropical zones.

4.2.2 Production type

Globally, there are estimated 1.5 billion cattle (FAO, 2022), comprising over 1000 domesticated breeds exhibiting various levels of meat quality, milk production, feed-efficiency, and other important traits (Yurchenko *et al.*, 2018). Each breed has phenotypic characteristics rendering them as beef or dairy breeds. Globally, there are an estimated 293 million beef cattle and 268 million dairy cattle (FAO, 2022). The population break down of the included studies (focusing on cattle production type) can be seen in Figure 4.7. Approximately 40% of the included publications focused on beef cattle and 28% of the publications focused on dairy cattle. The higher percentage of beef cattle publications could be because the countries with the highest number of publications (China, USA, Brazil) (Figure 4.3) predominantly focused on beef. The improvement of beef quality in Chinese cattle is currently a very important undertaking (Ma *et al.*, 2019). Dual-purpose systems are cattle production systems where both meat and milk are produced simultaneously, with lower productivity compared to systems based exclusively on milk- or beef production (González-Quintero *et al.*, 2020), there are less dual-production systems worldwide, explaining the low percentage of studies focusing on dual purpose cattle.

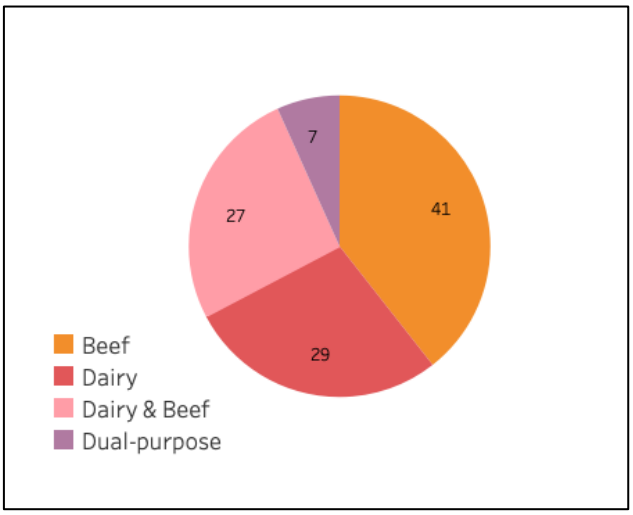

Figure 4.7: Pie chart showing count of studies by publication production type-focus.

Figure 4.8 depicts the dominant cattle production type for each country of the included publications. Publications focusing on beef cattle (orange) were mainly completed in Brazil, USA, and China. This result is foreseeable since these three countries are the leading beef producing countries worldwide (Statista, 2022), therefore much effort is put into the research of beef cattle. In China, indigenous beef cattle breeds have gained much attention for the detection of genes related to growth traits, and in Brazil, Nellore beef cattle are continually being studied to identify the genetic mechanisms involved in their environmental adaptation. Similarly, in South Africa, Nguni and Brahman cattle are beef breeds known for their superior hardiness, and thus often studied. Beef cattle studies were also completed in South Korea and Japan. In South Korea, the native Hanwoo cattle breed is a popular beef breed often studied, and in Japan much attention is given to Japanese Black cattle due to their abundant marbling of meat from intramuscular fat (Sasaki *et al.*, 2016b).

Publications focusing on dairy cattle (red) were concentrated in Europe (Italy, France, Denmark, Netherlands, Spain, Poland). The European Union (EU) is a significant producer of milk and milk products, with the total production being approximately 155 million tonnes (FAO, 2022). The major producers are Germany, France, the Netherlands, Italy, Poland, Ireland and Spain, accounting for about 70% of total EU milk production (Eurostat, 2020). The highest milk producing country is India, with 183 million tonnes in 2020 (FAO, 2022), although the single publication from this country did not focus on dairy cattle, but rather focussed on a high milk producing dual-purpose breed, Thanparkar cattle.

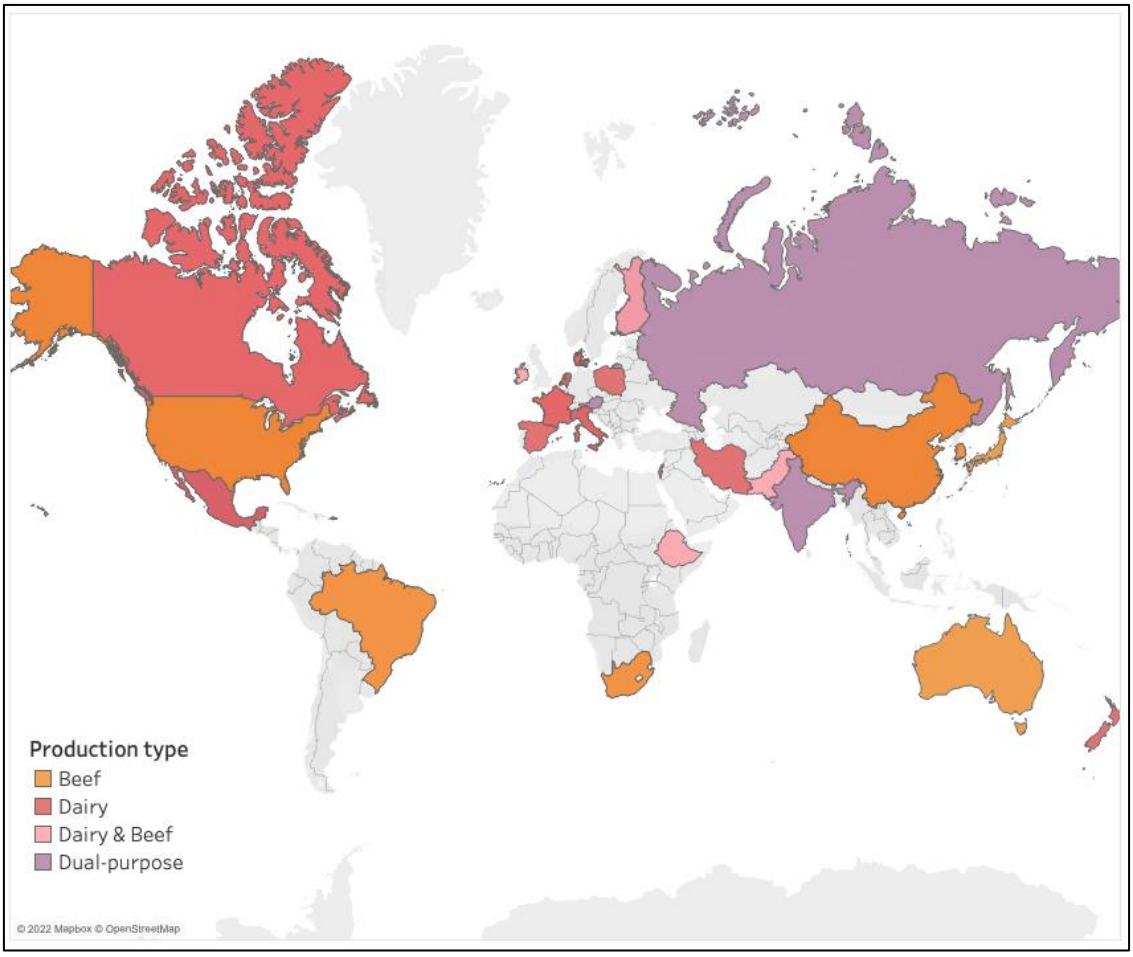

Figure 4.8: Dominant cattle production type in the data, mapped per country.

4.3 Method used for CNV detection

Different methods can be used to identify or genotype CNVs at a genome-wide scale (Clop *et al.*, 2012). Methods used to detect CNVs include array-based approaches and sequence-based approaches (Liu & Bickhart, 2012). Array based approaches such as Array Comparative Genomic Hybridization (CGH) and Single Nucleotide Polymorphism (SNP) arrays have been regularly used for the identification of CNVs, and their functioning widely reviewed (Lai *et al.*, 2005; Pinto *et al.*, 2012). SNP genotyping arrays, such as the Illumina's Bovine SNP50 array (54 609 SNPs) and Illumina's High-Density Bovine BeadChip Array (777 962 SNPs) were initially designed for SNP genotyping, but their application has been expanded to include CNV detection. More recently, next-generation sequencing (NGS) technologies, and their complementary analysis programs, have become available. NGS systems provide high coverage, resolution, and accuracy in CNV detection, but can be very costly (Choi *et al.*, 2016).

Figure 4.9 shows the breakdown of the platforms used in the included studies to detect cattle CNVs. SNPHD (dark green) is a widely used platform, as it has distinct advantages in terms of cost and throughput (Alkan *et al.*, 2011), therefore the high number of studies utilising this platform is not surprising. The SNPHD provides the opportunity to examine the genomic CNV landscape of large data sets due to its affordability and high-density markers (compared to SNP50). Sequencing (light blue) is closely behind SNPHD and will likely overtake array-based approaches (Alkan *et al.*, 2011), due to its numerous advantages. This technology resolves multiple technical drawbacks that arise from array-based approaches (Liu & Bickhart, 2012). Although sequencing has always been known to be a costly undertaking (Choi *et al.*, 2016), this has more recently been overcome by the decrease in sequencing costs (Bickhart *et al.*, 2020). Whole exome sequencing (WES) using NGS technology to identify CNVs is a cost-effective approach for large scale studies (Guo *et al.*, 2013).

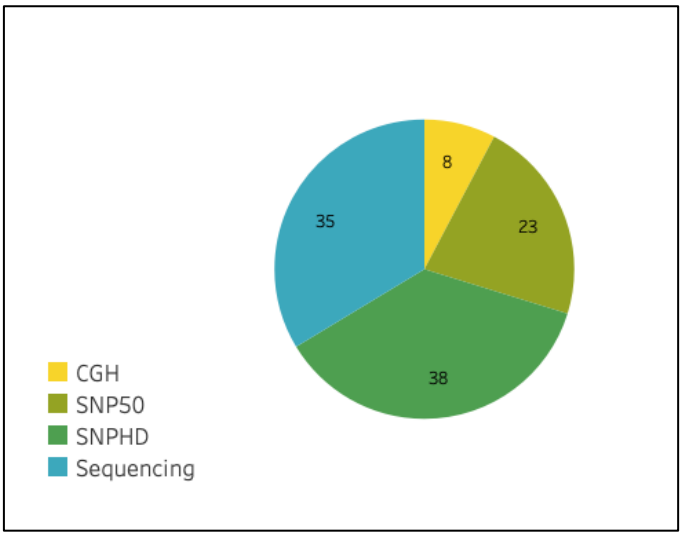

Figure 4.9: Pie chart showing count of studies by publication CNV detection method.

Figure 4.10 shows the CNV detection platforms used in each of the included studies from the years 2008 to 2021. The earliest publication included in this review implemented a CGH analysis to detect CNVs (Liu *et al.*, 2008b). After 2011 there was a steady decrease in the number of studies using Array CGH. Although CGH arrays have improved sensitivity, resolution and signal to noise ratio compared to SNP arrays (Pinto *et al.*, 2012), it only reports relative changes (Liu & Bickhart, 2012), and algorithms for CNV discovery from CGH arrays do not consider B allele frequency (BAF) information (Xu *et al.*, 2013a). The Bovine SNP50 array first became readily available in 2009 (Matukumalli *et al.*, 2009). Figure 4.10 shows that the usage of the Bovine SNP50 stayed quite stable, with low to medium usage from 2009 to 2021. The usefulness of the SNP50 is limited due to the low density of markers. Bovine SNP50 markers are spaced at an average of about 50-kb intervals, and it is estimated that in cattle, markers need to be spaced less than 10 kb apart (Rincon *et al.*, 2011). The SNPHD (BovineHD) platform, with over 777 000 evenly spaced SNPs, first became available in 2010 (Matukumalli *et al.*, 2011; Rincon *et al.*, 2011). It can be seen that the usage of the

BovineHD platform has an upward tendency from 2012 onwards, with peaks in 2016 and 2018, and lower usage from 2019 onwards (Figure 4.10). The drop in the usage of SNP genotyping arrays could be because these array-based methods are unable to detect CNV events that are balanced (such as translocation or inversions, see Figure 2.5a), nor can they detect small CNVs or exact breakpoints, because of their limited probe density (Yang *et al.*, 2021). Moreover, the increasing cost-effectiveness of sequencing technology opens new opportunities for this platform as an alternative. CNV detection and analysis by sequencing was first published in 2011 by Zhan *et al.* (2011) and Stothard *et al.* (2011), followed by Bickhart *et al.* (2012) the next year. This platform has a low usage from 2011 to 2015, and high usage in 2016, likely due to the above-mentioned decrease in cost. The usage dropped in 2018 but has been increasing gradually since. Sharma *et al.* (2017) reviewed NGS in livestock species and reported *Bos taurus* to be the most highly sequenced species.

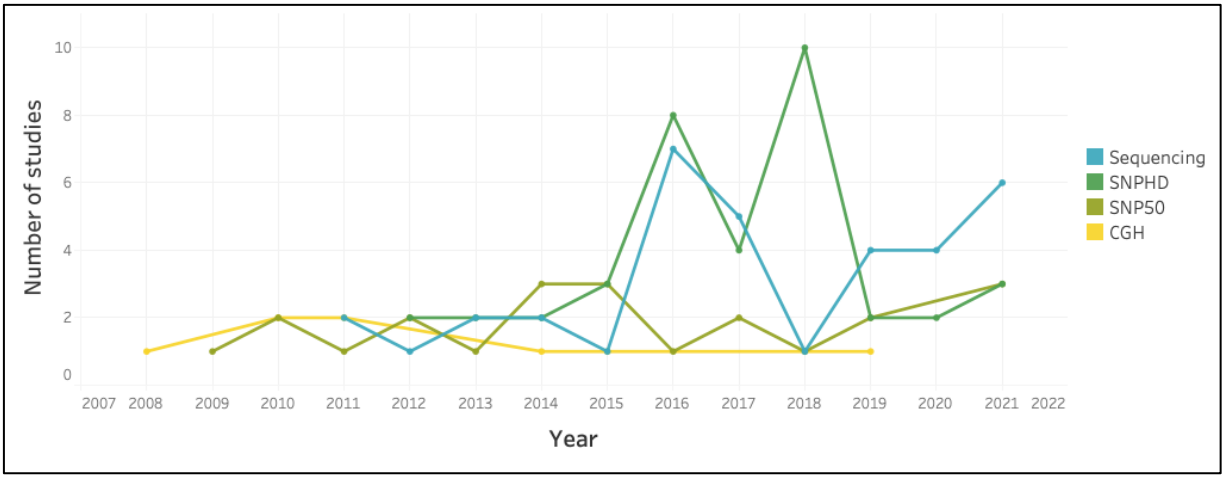

Figure 4.10: Line graph showing CNV detection platforms used in each of the included studies.

The trend lines of the CNV detection platforms used in each of the included studies from the years 2008 to 2021 can be seen in Figure 4.11. It can be seen in this graph that the sequencing, SNPHD and SNP50 platforms have a positive slope, whereas CGH has a negative slope. Moreover, the sequencing platform trend line is steeper than Bovine SNP trend lines, illustrating that it is likely to surpass array-based approaches over time. Microarray approaches have been the main platform for CNV detection (Xu *et al.*, 2013a), but the development of cost-effective sequencing approaches with enhanced resolution and accuracy (Yang *et al.*, 2021) could alter that. The advances in sequencing technology provides researchers with a plethora of opportunities to figure out the genetic mechanisms involved in complex traits to be used accordingly in livestock breeding programmes (Sharma *et al.*, 2017).

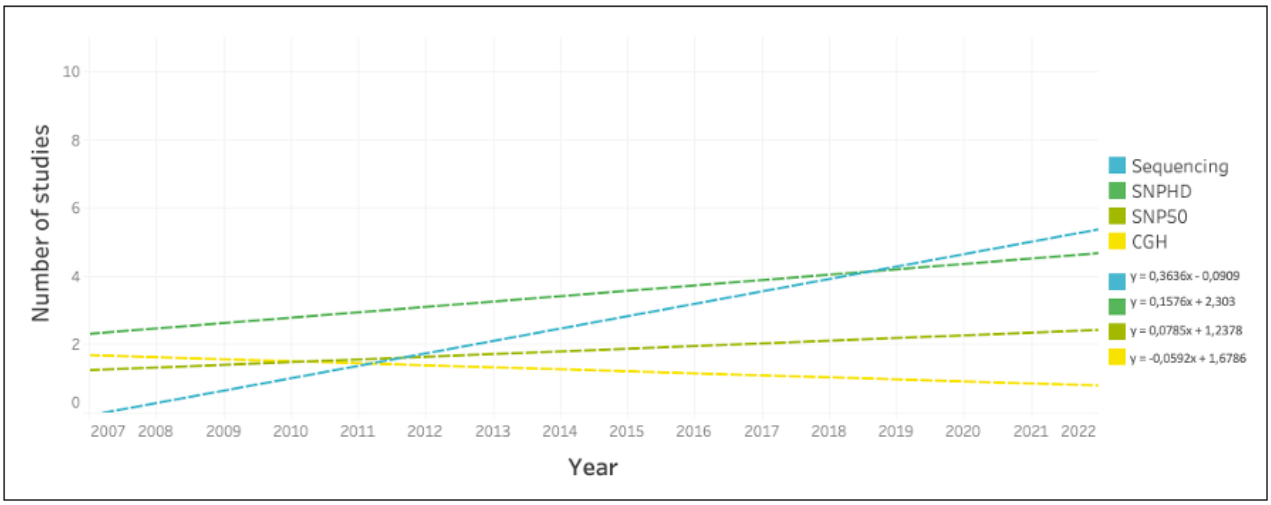

Figure 4.11: Trend lines of the CNV detection platforms used in each of the included studies.

CHAPTER 5 ANALYSIS OF GENES

5.1 Introduction

The analysis of DNA variations that affect the phenotypes of animals is a crucial area of research in livestock genetics. Copy number variations, such as duplications, deletions, and insertions of a whole gene or a section of a gene, shape phenotypic variation in cattle by altering gene dosage, gene expression regulation, gene fusion and interruption, and transcript structure (Bickhart & Liu, 2014). Therefore, copy number variation may be one of the main contributors to phenotypic differences and evolutionary adaptation observed in animals (Clop *et al.*, 2012).

A key objective in the use of genetic markers for livestock breeding improvement is the detection of genes that influence quantitative traits (Goshu *et al.*, 2018b). In this chapter, the most frequent CNV-related genes affecting adaptation, immunity, milk production, reproduction, meat production, growth, feed-efficiency, and coat characteristics extracted from the systematic review are reported. This was done by extracting the genes from each study, grouping the genes according to their trait-related classification, and reporting them on individual Excel worksheets in a separate Excel workbook. To ensure uniform gene nomenclature, each gene was subsequently searched on the National Centre for Biotechnology Information (NCBI) webpage for the identification of Gene ID aliases and gene families. To prepare the data for analysis, the base data were extracted from the source worksheet, standardised, and cleaned to remove irrelevant characters. The data were placed into a Tableau software for analysis and graphical representations of frequency of genes. A word cloud function in Tableau was used for graphical representation, whereby the size of each word (or gene in this case) indicates its frequency and importance.

5.2 CNV-related genes affecting adaptation and immunity traits

Adaptation is a feature of livestock involving physiological, morphological, anatomical, and biochemical mechanisms (Key & Sneeringer, 2014). Adaptive fitness is distinguished by traits related to survival, health, and reproduction (Keba *et al.*, 2010). In cattle, adaptation can be observed in the form of thermotolerance, high altitude adaption, disease resistance, and resistance to tick and nematode infections. In this review, the most frequent adaptability and immunity-related genes detected in the included publications is presented in Table 5.1, and a visual representation thereof can be seen in Figure 5.1. These genes include ABCC4, BOLA gene family, IGLL1, OR family, WC1 and ZNF280B, followed by BSP30A, DEFB, ULBP gene family, CATHL gene family, and the HSP gene family.

Table 5.1: The main adaptability and immunity-related genes detected in the included publications.

Gene	Count	References
ABCC4	7	(Matukumalli *et al.*, 2009), (Liu *et al.*, 2010), (Liu *et al*., 2011), (Hou *et al*., 2012b), (Hou *et al*., 2012c), (Dolezal *et al*., 2014), (Upadhyay *et al*., 2017)
BOLA	6	(Hou *et al.*, 2011a), (Hou *et al.*, 2012c), (Lee *et al.*, 2013b), (Wang *et al.*, 2014), (Prinsen *et al.*, 2016), (Zhou *et al.*, 2016b)
IGLL1	6	(Hou *et al.*, 2012b), (Hou *et al.*, 2012c), (Wang *et al.*, 2015), (Xu *et al.*, 2016), (Yang *et al.*, 2017b), (Pierce *et al.*, 2018)
OR	6	(Matukumalli *et al.*, 2009), (Liu *et al.*, 2010), (Seroussi *et al.*, 2010), (Hou *et al.*, 2011a), (Hou *et al.*, 2012b), (Butty *et al.*, 2020)
WC1	6	(Liu *et al.*, 2010), (Hou *et al.*, 2011a), (Bickhart *et al.*, 2012), (Hou *et al.*, 2012b), (Hou *et al.*, 2012c), (Zhang *et al.*, 2014)
ZNF280B	6	(Bickhart *et al.*, 2012), (Bickhart *et al.*, 2016), (Xu *et al.*, 2016), (Zhou *et al.*, 2016a), (Mustafa *et al.*, 2018), (Strillacci *et al.*, 2018)
BSP30A	5	(Liu *et al.*, 2010), (Liu *et al*., 2011), (Bickhart *et al.*, 2012), (Hou *et al*., 2012b), (Zhang *et al.*, 2014)
DEFB	5	(Liu *et al.*, 2010), (Hou *et al.*, 2011a), (Bickhart *et al.*, 2016), (Butty *et al.*, 2020), (Peripolli, 2021)
ULBP17	5	(Liu *et al.*, 2010), (Liu *et al*., 2011), (Bickhart *et al.*, 2012), (Bickhart *et al.*, 2016), (Da Silva *et al.*, 2016b)
CATHL4	3	(Bickhart *et al.*, 2012), (Da Silva *et al.*, 2016b), (Jang *et al.*, 2021)
HSP70	3	(Pierce *et al.*, 2018), (Hu *et al.*, 2020b), (Jang *et al.*, 2021)

Figure 5.1: Visual representation of adaptability and immunity-related gene prevalence in included publications.

Bovine leukocyte antigen gene family (BOLA)

Major histocompatibility complex (MHC) genes are particularly relevant in the livestock farming sector, because of their association with resistance or susceptibility to infectious and costly diseases. In cattle, the MHC is called the bovine leukocyte antigen (BOLA) system, is located on BTA23, and is involved in the immune response to pathogens (Behl *et al.*, 2012). The MHC class I molecules are expressed by every nucleated cell, and put forward endogenous antigens to cytolytic T-cells with $CD8^+$ receptors, while MHC class II molecules are expressed by antigen-presenting cells and put forward exogenous antigens to helper T-cells with $CD4^+$ receptors (Oprzadek *et al.*, 2018; Bohórquez *et al.*, 2020). Cattle express two class II proteins, DQ and DR, which encode proteins that process the antigen and then present it to the $CD4^+$ T-lymphocytes (Figure 5.2) (Norimine & Brown, 2005; Oprzadek *et al.*, 2018).

Within the adaptability and immunity gene classification, the BOLA gene family is the most frequently reported. Genes such as BOLA-DRB3, BLA-DQB, BOLA-DQA2, BOLA-DQA5, and BOLA-DQB, were identified in the included publications and classified as immune related genes. The BOLA-DR region comprises the BOLA-DRA locus and three DRB loci, however only the BoLA-DRB3 gene is functional (Figure 5.2) (Longeri *et al.*, 2021). The DRA gene encodes the α-chain, whereas the DRB gene (BOLA-DRB3) encodes the β-chain MHC class II molecules (Behl *et al.*, 2012). The BOLA-DRB3 gene is the most polymorphic, functional, and highly expressed gene of the bovine MHC, and has therefore been thoroughly researched for more than twenty years (Oprzadek *et al.*, 2018;

Longeri *et al.*, 2021). Studying polymorphism in the DRB3 locus is particularly important due to its presence in the antigen-presenting site, and because the variability therein could be linked to immune response variability (Behl et al., 2012). Various studies have reported the association of BOLA-DRB alleles with resistance or susceptibility to diseases such as clinical (Rupp *et al.*, 2007) and subclinical (Ramírez *et al.*, 2014) mastitis. Alleles of the BOLA-DRB3 have also been associated with resistance to tick infestation in cattle (Kim *et al.*, 2017). The BOLA-DQ consists of DQA and DQB loci (Figure 5.2), which may differ in number. Cattle comprise five DQA loci and five DQB loci, which emerged from gene duplication (Miyasaka *et al.*, 2012). Although DQ genes are not as polymorphic as DR genes, DRB3 is the only functional DRB gene, thus DR molecules alone could be lacking a sufficiently broad range of presented antigens, suggesting DQ molecules have equal importance to DR molecules for immunity (Miyasaka *et al.*, 2012).

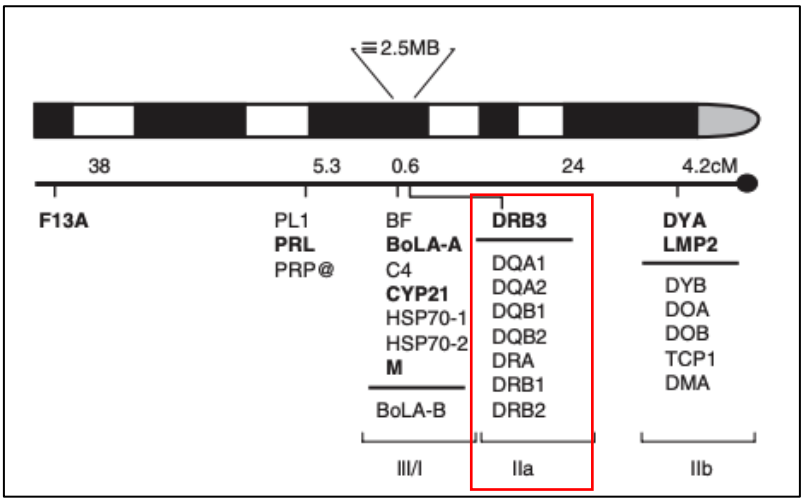

Figure 5.2: The map of structural genes located on BTA23. Adapted from Takeshima & Aida (2006).

In summary, BOLA genes are expressed by the cells of the immune system which process antigenic peptides to present to helper T-cells for an immune response against pathogens (Oprzadek *et al.*, 2018), thus making this gene a very important gene in cattle immunity.

ATP binding cassette subfamily C member 4 (ABCC4)

The ATP-binding cassette (ABC) genes, representing one of the largest known protein superfamilies (Kaminski *et al.*, 2006), can be separated into seven distinct subfamilies, ABC-A to ABC-G (Dean *et al.*, 2002; NCBI, 2022). The ABC transporters use energy from ATP hydrolysis to move various substrates across intra- and extra-cellular membranes (Farke *et al.*, 2008). These substrates include lipids and sterols, metabolic products, peptides, and drugs (Dean *et al.*, 2002). A few members of the ABC subfamily C (ABCC), such as MRP1-MRP5 (ABCC1-ABCC5), are drug transporters and are best characterised by their involvement in the transport of xenobiotics (foreign substances) (Li *et al.*, 2017).

Within the adaptability and immunity gene classification, the ABCC4 gene is one of the most frequently reported genes. Studies have reported the importance of ABCC4 genes in defence/innate immunity, drug detoxication and adaptive immunity (Liu *et al.*, 2010; Upadhyay *et al.*, 2017). This gene encodes the multi-drug resistance protein 4 (MRP4). As mentioned previously, the MRP4 gene belongs to the superfamily of ABC transporters (Lacroix-Pépin *et al.*, 2011), and acts like an export pump for cellular detoxication (Bögeholz *et al.*, 2021), by actively extruding diverse endogenous and xenobiotic compounds, together with their phase II metabolites, from the cell (Russel *et al.*, 2008), thereby granting resistance to several cytotoxic compounds, and protecting important tissues against them (Sodani *et al.*, 2012). Liu *et al.* (2011b) performed three array CGH experiments to analyse the CNVs and genes in Angus cattle selected for their resistance or susceptibility to GIN. In that study, two gain events in regions corresponded to ABCC4 genes, suggesting that it could play a role in multidrug resistance. A sequence based GWAS performed by Naval-Sánchez *et al.* (2020) revealed the ABCC4 gene to be directly involved in response to *Mycobacterium avium* subsp. *Paratuberculosis* infection. Moreover, Li & Gasbarre (2009) experimentally infected cattle with *Cooperia oncophora* and found an upregulated expression of ABCC4 in the resistant cattle, suggesting that these may relate to the over usage of anthelmintics for long periods of time.

Zinc finger protein 280B (ZNF280B)

Zinc finger (ZNF) proteins are transcription factors with finger-like binding domains, playing a major role in gene regulation (Li *et al.*, 2022b). The ZNF proteins are among the most abundant proteins, and exhibit a variety of molecular functions, such as DNA or RNA binding, transcriptional regulation, and numerous other processes (Laity *et al.*, 2001; Cassandri *et al.*, 2017). They have an important role in various diseases, especially human cancer, with zinc finger protein 280B (ZNF280B) moderating pro-growth and -survival functions in prostate cancer and displaying an association with gastric cancer (Zhai *et al.*, 2018). A study focussing on zebu cattle found ZNF280B in a CNV region and overlapped with a QTL region associated with clinical mastitis (Zhou *et al.*, 2016a).

Bovine salivary protein 30 kDa (BSP30A)

Although Bovine salivary protein 30 kDa (BSP30A) has been renamed to BPIFA2A (Bingle *et al.*, 2011), in this review it is referred to as BSP30A to keep uniform with previous publications reporting on this protein. The BSP30A gene belongs to the BSP30 gene family, a bactericidal/permeability-increasing (BPI)-like protein gene family believed to contribute to the innate immunity of the oral cavity and airways (Wheeler *et al.*, 2003; Bickhart *et al.*, 2012).

Due to the favourable environment of the oral cavity, such as elevated temperature and humidity, it serves as a major portal of entry for pathogens, which could potentially enter the lungs or gastrointestinal tract. Thus, the oral cavity comprises a robust host defence network to hinder the progress of pathogens, of which saliva is an essential element (Haigh *et al.*, 2008). Both BSP30A

and BSP30B are expressed in the salivary glands in abundance, and make up 15-30% of the bovine salivary protein content, rendering them essential in the first line of defence against orally ingested bacteria or parasites (Haigh *et al.*, 2008). In ruminants, it is suggested that duplications in the BSP30A region could be a response to evolutionary pressures from bacteria or parasites encountered in the soil while grazing (Bickhart *et al.*, 2012).

Immunoglobulin lambda-like polypeptide 1 (IGLL1)

Immunoglobins, through their antibody activity, are an important factor in disease resistance in the body (Zhao *et al.*, 2021). Immunoglobins (Ig) are the mediators for the adaptive humoral immunity response in jawed vertebrates (Ekman *et al.*, 2009). Immunoglobulin molecules include two heavy (IgH) chains and two light (IgL) chains, which are connected by disulphide bridges to form a "Y" shaped molecule with a twofold symmetry structure (Figure 5.3) (Cui *et al.*, 2021).

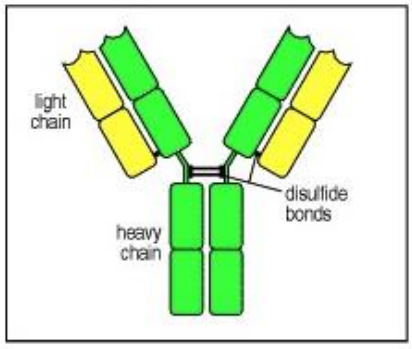

Figure 5.3: Immunoglobulin molecule composed of two types of protein chain, light (yellow) and heavy (green) chains. From Janeway *et al.* (2001).

Furthermore, IgL chains are classified into two types, namely, lambda (λ) and kappa (κ) (Chi *et al.*, 2020). No functional difference has been discovered between antibodies having λ or κ light chains, but the ratio of the two types differs between species. In humans, the average κ to λ ratio is 2:1, whereas in cattle it is 1:20 (Janeway *et al.*, 2001). The bovine immunoglobulin λ chain locus is located on chromosome 17 (Ekman *et al.*, 2009). In domestic cattle, IGLL1 forms part of IgL chain gene pool that are involved in B lymphocyte production (Ekman *et al.*, 2009). This gene was differentially expressed in cattle that were selected for resistance or susceptibility to GIN (Araujo *et al.*, 2009), and similarly, was associated with resistance to GIN in Angus cattle (Hou *et al.*, 2012b). The copy number variation displayed within this gene could explain the differences in adaptive immunity found among several cattle populations (Wang *et al.*, 2015). Interestingly, IGLL1 can also affect fertility traits, as described in section 5.2.

Olfactory receptors (ORs)

In mammals, olfactory receptors (ORs) are responsible for the detection of odorant molecules, which gives rise to the perception of smell (Lee *et al.*, 2013c). Olfaction is crucial for the avoidance of danger and the location of potential mates or feed (Connor *et al.*, 2018). The OR gene family is the largest gene family in the mammalian genome, and various OR genes were identified in the relevant publications, such as OR5H8, OR7A10, OR7A17, OR10AD1, OR5T2, OR4S1, and others (Butty *et al.*, 2020, 2021; Jang *et al.*, 2021).The nomenclature scheme for the OR genes was described by Glusman *et al.* (2000) (Figure 5.4). The root symbol “OR” indicates that it is an olfactory receptor, followed by an integer denoting the family of the gene, a letter denoting the subfamily of the gene and an integer denoting the individual gene within the subfamily (Glusman *et al.*, 2000). Thus, the OR5H8 gene, identified by Butty *et al.* (2020), can be identified as olfactory receptor family 5 subfamily H member 8.

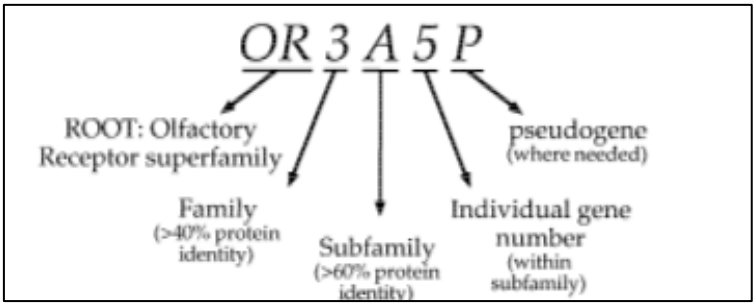

Figure 5.4: The nomenclature scheme for the olfactory receptor genes. From Glusman *et al.* (2000).

Olfactory receptor proteins belong to the family of G-protein-coupled receptors and have a seven-transmembrane domain structure (Glusman *et al.*, 2000). The function of the OR has evolved to warn animals about potential threats. Multiple odorant molecules can be detected and bound by one OR, and conversely, one odorant molecule can bind to several ORs (Figure 5.5) (Malnic *et al.*, 1999). This combinatorial receptor coding scheme allows animals to differentiate varied and complex odours (Malnic *et al.*, 1999), which is essential for animal survival. These chemical messages help animals locate feed, identify potential mates, detect chemical toxins, and detect the presence of predators (Connor *et al.*, 2018). In cattle there are 881 OR genes, and approximately 40% of OR loci are copy number variable, indicating that CNVs of OR genes are very common (Lee *et al.*, 2013c). Jang *et al.* (2021) called CNVs by a read depth strategy of NGS from 336 cattle, representing breeds from *Bos taurus*, *Bos indicus* and their African hybrids. The researchers found significant variations in the number of OR genes between populations, indicating that olfactory function could be greatly influenced by natural selection and OR CNVs could be candidate genes under selection. This multigene family is a vital genetic factor in the evolution of mammalian species (Fernandes Júnior *et al.*, 2020).

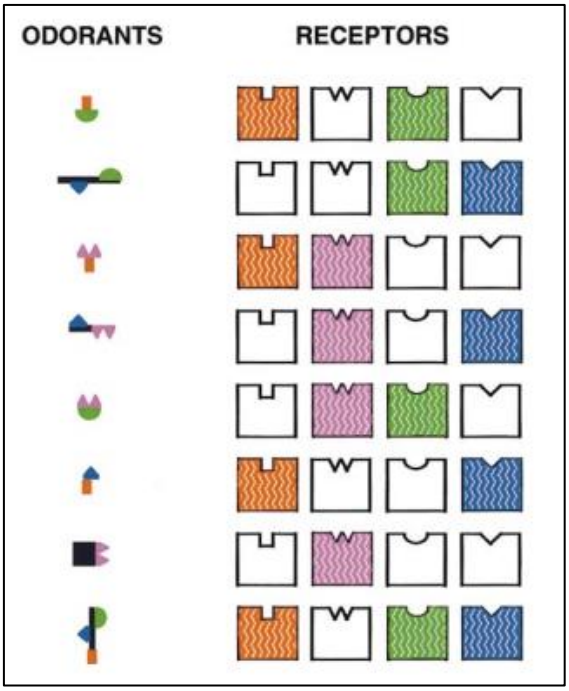

Figure 5.5: The combinatorial receptor coding scheme for odorants. From Malnic *et al.* (1999).

UL16-binding protein (ULBP) gene family

The cattle Major Histocompatibility Complex Class I-like Gene Family A (MHCLA) was first discovered in the spleen cDNA of cattle (Larson *et al.*, 2003). This MHC Class I superfamily has two paralogues, MHCLA1 and MHCLA2. These molecules have similar peptide sequences and are related to the human ULBP and RAET1 genes (Larson *et al.*, 2003). To keep uniform with previous publications reporting on this molecule, in this review, the cattle MHCLA1 and MHCLA2 genes are referred to as ULBP1 and ULBP2, respectively. The ULBP molecules interact with the natural killer group 2, member D (NKG2D) receptor, thus activating the effector cells in the immune system. This is important in natural killer cell-mediated cytotoxicity (Larson *et al.*, 2006). It is assumed that this gene family evolved and expanded due to selective pressure from viral pathogens (Larson *et al.*, 2003).

In cattle, the ULBP genes are involved in antiviral immunity (Larson *et al.*, 2006). Both UL16-binding protein 17 (ULBP17) and UL16-binding protein 21 (ULBP21) were identified in the included publications. The ULBP17 gene has been identified as a highly duplicated gene in cattle (Liu *et al.*, 2010; Hou *et al.*, 2011a). Bickhart *et al.* (2012) found duplications of ULBP17 in Nellore cattle, which could be in response to increased viral pathogens, while Peripolli (2021) identified variable regions, with functions associated with environmental resilience, harbouring important genes such as ULBP21.

Cathelicidin (CATHL) family

In mammals, defensins and cathelicidins (CATHLs) are two of the main antimicrobial peptide families (Tran *et al.*, 2002; Huan *et al.*, 2020). These small, cationic, antimicrobial peptides (AMPs) have been discovered in almost all living species (E *et al.*, 2015), and are a crucial component of the innate immune system, playing a vital role in host defence and disease resistance in mammals (Flores, 2011). Cathelicidins vary in amino acid sequence, structure, and size (Kościuczuk *et al.*, 2012). They have a highly conserved N-terminal cathelin-like pro sequence and a highly variable C-terminal antimicrobial domain which encodes the mature peptide (Flores, 2011). While the N-terminal (~100 residues) is highly conserved across many species, the C-terminal is heterogenous in structure (α-helical, β-hairpins, or specific amino acid enriched) and in length (12 - 80 or more residues) (Figure 5.6) (Kościuczuk *et al.*, 2012). Cathelicidin genes have a four exon/three intron arrangement (Figure 5.6); exons one, two and three code for the precursor cathelin-like domain, and exon four codes for the mature antimicrobial peptide (Whelehan *et al.*, 2014).

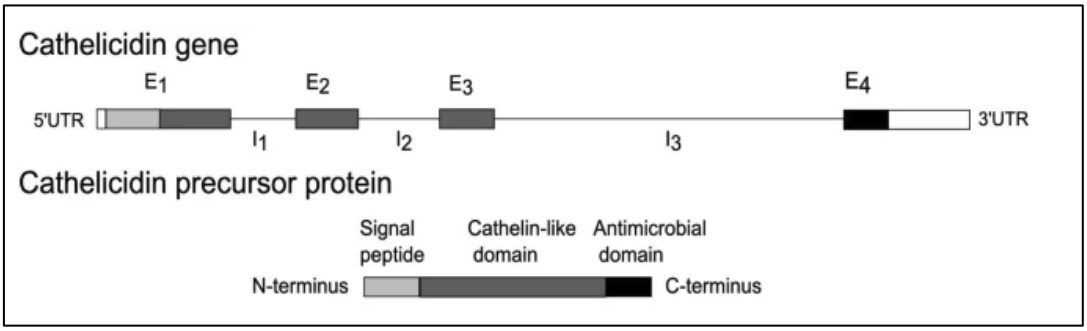

Figure 5.6: Cathelicidin structure, with a four exon/three intron arrangement. From Whelehan *et al.* (2014).

Cathelicidins have two structural characteristics, namely, the ability to fold into amphipathic structures, and a net-positive charge (Zanetti, 2004). This allows mature CATHLs to bind to microbial surfaces that are negatively charged, thereby disrupting their membranes and inactivating the invading microbe. However, to become active, these molecules must be freed from the N-terminal (Kościuczuk *et al.*, 2012). Bovine CATHLs are frequently α-helical and show effective antimicrobial activity against numerous types of bacteria. In cattle, the first CATHLs were isolated from bovine neutrophils, namely, bactenecins 5 and 7 (Bac5 and Bac7) (Kościuczuk *et al.*, 2012). In contrast to the human genome that only has one member of this gene, several CATHL peptides are found in cattle (Figure 5.7). There are seven protein-coding CATHL genes, marked CATHL1-CATHL7 (Whelehan *et al.*, 2014). Their structural diversity implies these molecules have distinct functional capabilities and a diversified role in host defence (Zanetti, 2004).

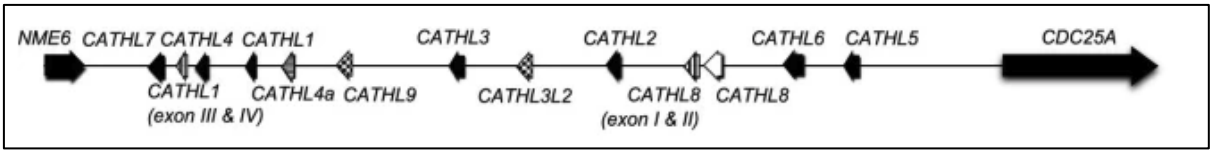

Figure 5.7: BTA22: Genomic organisation of the CATHL gene cluster. From Whelehan *et al.* (2014).

Much research has revolved around the bovine CATHL4 antimicrobial peptide due to its small size and high tryptophan content (Flores, 2011). The CATHL4 gene encodes the antimicrobial peptide indolicidin which exerts activity against fungi such as *Candida albicans, Cryptococcus neoformans* and bacteria such as *S. aureus* and *E. coli* and can induce autophagic cell death of the protozoan pathogen *Leishmana donovani* (Kościuczuk *et al.*, 2012). Bickhart *et al.* (2012) found recent duplications of CATHL4 in Nellore cattle, which could be in response to increased bacterial or helminthic pathogens.

β-Defensins (DEFB)

β-defensins are cationic, cysteine-rich antimicrobial peptides (AMPs) belonging to the defensin family. β-defensins contain six cysteines that form specific disulphide bonds at C1-C5, C2-C4, and C3-C6 residues, and have a two-exon structure (Figure 5.8) (Roosen *et al.*, 2004). These multifunctional peptides can act against various types of bacteria, enveloped viruses, fungi, and other unicellular parasites. In addition to their antimicrobial activity, they are chemoattractants for immature dendritic cells, T-lymphocytes and monocytes (Roosen *et al.*, 2004).

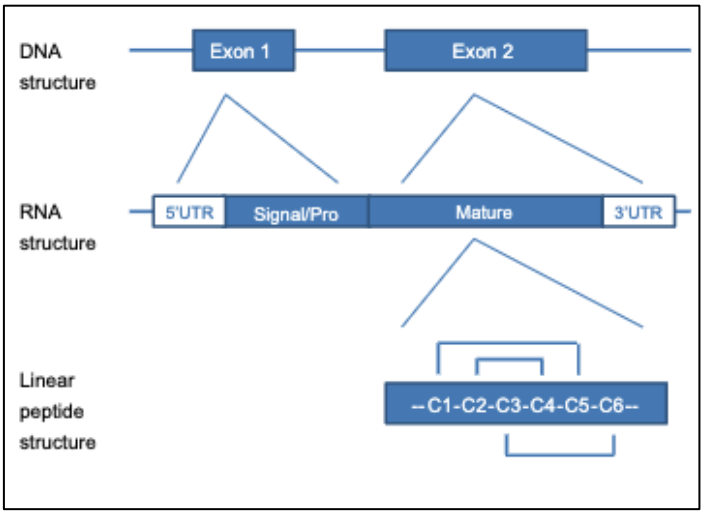

Figure 5.8: The β-Defensin gene structure. From Meade *et al.* (2014).

Defensins can be organised into three distinct subfamilies: α, β, and θ-defensins. The α-defensin was isolated from murine Paneth cells of the small intestine (Selsted *et al.*, 1992), and the θ-defensins have been discovered in Rhesus monkey genomes (Tran *et al.*, 2002). The first β-defensin was discovered in the bovine respiratory tract and was termed tracheal antimicrobial peptide (TAP)

(Diamond *et al.*, 1991), which later was associated with immediate reaction to inflammation and thereafter shown to be present in the cattle mammary gland (Meade *et al.*, 2014). Roosen *et al.* (2004) reported 18 bovine β-defensin peptide sequences, which was noticeably smaller than the number identified in other species. However, a systematic search increased the count of putative bovine β-defensin genes by identifying 57 open reading frames with resemblance to the distinguishing six-cysteine spacing (Cormican *et al.*, 2008). The clustering of these genes was done based on synteny analysis, and four clusters were identified in the bovine genome on chromosome 8 (cluster A), 13 (cluster B), 23 (cluster C), and 27 (cluster D) (Meade *et al.*, 2014).

The bovine β-defensins that are located on BTA27 (cluster D) consist of the most immunologically important genes for intramammary infections (Gurao *et al.*, 2017). In cattle, β-defensin variable genes are upregulated in reaction to inflammation in mammary, uterine or lung tissue. Wojdak-Maksymiec *et al.* (2006) associated β-defensins with reduced somatic cell count (SCC) in Jersey cattle. While Peripolli (2021) identified several β-defensin genes, such as DEFB1, DEFB4A, DEFB5, DEFB6, DEFB7, DEFB10, and DEFB13, which are essential for protection against viral, fungal, and bacterial infections, and for linking the innate and adaptive immune responses. Furthermore, Saravanan *et al.* (2021) identified candidate genes for resistance to mastitis (DEFB4, DEFB7, DEFB10) in Thanparkar cattle. Mastitis is one of the most frequent diseases in dairy cattle, thus the relation of β-defensins to reduced SCC and resistance to mastitis is very relevant and useful for dairy producers.

Heat shock proteins (HSP) family

In mammals, heat shock transcription factors (HSFs) regulate and mediate the cellular response to thermal stress. These HSFs are regulated by the expression of corresponding HSF genes. After activation, the HSFs bind with heat shock elements in the promoter region of the HSP genes, which then leads to increased transcription of HSP mRNA (Archana, 2017). Thermoregulation is a mechanism by which livestock species maintain their body temperature within certain boundaries, despite the wide variations in environmental temperature (Renaudeau *et al.*, 2011). High environmental temperatures for extended periods of time can lead to animals absorbing more heat than they are able to dissipate, resulting in heat stress (Mkize & Zishiri, 2020). When animals are subjected to this type of stress, there are proteins that are preferentially expressed, such as HSPs. Heat shock proteins are important for the alleviation of heat stress in mammals, as they are molecular chaperones that function by ensuring proper folding and refolding of proteins, instigating aggregations of proteins, and ensuring potentially damaging interactions are avoided (Mkize & Zishiri, 2020). The regulation of HSP production is very important for cell survival (Basiricò *et al.*, 2011).

There are many different types of HSPs studied in livestock, based on their activity, function, and molecular weight, namely HSP110, HSP100, HSP90, HSP70, HSP60, HSP40, HSP10 and some smaller HSP families (Archana, 2017). However, new guidelines were proposed for the nomenclature of HSP families, HSPH (HSP110), HSPC (HSP90), HSPA (HSP70), DNAJ (HSP40), and HSPB (small HSPs) (Kampinga *et al.*, 2009). Publications that were included in this review reported on HSPAs (HSP70) and HSPBs (small HSPs), thus the nomenclature proposed by Kampinga *et al.* (2009) is used in this review. Heat shock protein genes such as HSPA5, HSPA9, HSPB1, and HSPB8 were identified in the included publications.

The genes HSPA5 and HSPA9 form part of the HSPA (HSP70) family, with the latter family being the most studied and reported to be the most abundant protein conferring thermotolerance (Basiricò *et al.*, 2011; Banerjee *et al.*, 2014). The genes HSPB1 and HSPB8 form part of the small-molecular-weight HSPB family. The HSPA along with HSPB1 (and HSPC) proteins exhibit anti-apoptotic activity in mammalian cells (Garrido *et al.*, 2001). HSPB8 gene, mapped on BTA17, halts the accumulation of insoluble aggregates, as it prevents aggregation or promotes degradation of misfolded proteins. Therefore, HSPB8 encodes a protein with a cytoprotective role and thus is expressed in response to thermal stress (Verma *et al.*, 2016). Wang (2016) reported that HSPB8 was presented as a deletion in Nguni cattle. Nguni cattle can survive in harsh environments and are characterised by their heat tolerance capabilities. Hence, heat shock proteins are understood to play a role in the ability of cattle to tolerate heat, and thus in the climatic adaptabilities of different breeds (Wang, 2016).

Workshop cluster 1 co-receptors (WC1)

Cattle are the most extensively studied species in respect to this gene family. Workshop cluster 1 (WC1) molecules form part of the B scavenger receptor cysteine rich superfamily and are exclusively expressed on gamma delta T (γδT) cells (Herzig & Baldwin, 2009). WC1 molecules can act as pattern recognition receptors (PRRs) by binding components of pathogens, and function as signalling co-receptors for the T cell antigen receptor (TCR) by directing an immune response from γδT cell subgroups (Hsu *et al.*, 2015; Telfer & Baldwin, 2015). The PRRs can recognise pathogen-associated patterns from a range of viruses, bacteria, and protozoa (Chen *et al.*, 2014). Co-receptors are recognised to potentiate the activation of T cells (Hsu *et al.*, 2015).

Bovine γδT cells are separated into two main subgroups based on the surface expression of WC1 (Albarrak *et al.*, 2017). WC1− predominates in organs such as the spleen and intestine, whereas the WC1+ subset predominates in circulating blood (Blumerman *et al.*, 2006). There are 13 members in the WC1 gene family (WC1-1 to WC1-13), organised within two loci on BTA5 (Herzig & Baldwin, 2009). The WC1+ γδT cell subset can be further divided into three different groups based on their intron/exon structure and thus, antibody reactivity. The endodomain tail regions of Type I (WC1.1),

Type II (WC1.2) and Type III (WC1 – 11) genes are coded for by four, five, or six exons, respectively (Figure 5.9) (Herzig & Baldwin, 2009; Albarrak *et al.*, 2017). The expression of WC1+ molecules from the three groups is linked to the variation in γδT cell responses (Wang *et al.*, 2011).

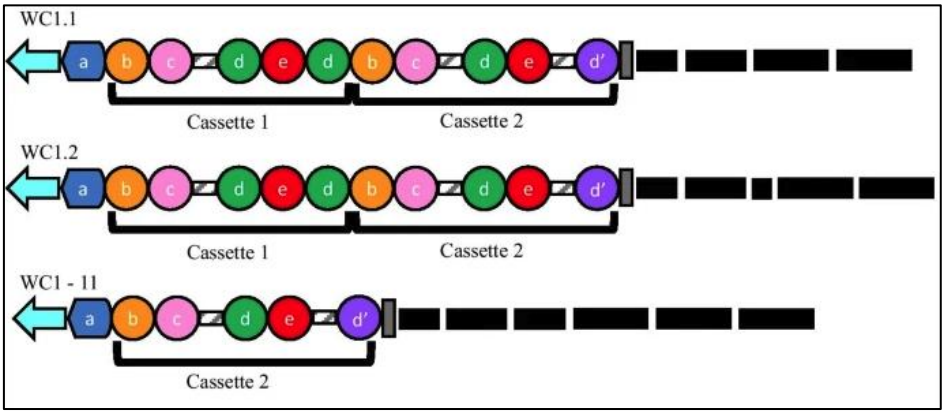

Figure 5.9: Structure of WC1 molecules, Type I (WC1.1), Type II (WC1.2) and Type III (WC1 – 11). From Loonie *et al.* (2021).

Cattle have a high percentage of γδT cells dispersed throughout their tissues and in their peripheral blood, compared with humans (Telfer & Baldwin, 2015). The γδT cells can make up 10-15% of the peripheral blood mononuclear cells in adult cattle and up to 40% in young calves, thus, this system is excellent for investigating γδT cells and the surface receptors needed for their activation (Pollock & Welsh, 2002; Wang *et al.*, 2011). It is speculated that the high percentage of γδT cells, along with the genetic diversity of γδ TCRs, could have evolved as a defence response to certain infectious agents to which cattle are frequently exposed (Pollock & Welsh, 2002). Increasing evidence indicates that WC1 γδT cells play an important immunological role in bovine tuberculosis (Kennedy *et al.*, 2002; Rusk *et al.*, 2017; Blanco *et al.*, 2021). This important animal and zoonotic disease is caused by *Mycobacterium bovis*. The γδT cells have a remarkable characteristic in that they are able to connect the innate and acquired immune responses, due to their capacity to respond directly to pathogen-associated molecular patterns because of their PRRs and TCRs (Blanco *et al.*, 2021).

5.3 CNV-related genes affecting milk production and reproduction related traits

Milk production is an indispensable economic trait in cattle production. Globally, more than 6 billion people consume milk and milk products (FAO, 2022). Milk production provides worldwide nutrition and food security, and is governed by environmental and genetic factors (Silpa *et al.*, 2021). In dairy farming, milk yield and milk composition are extremely important economic traits (Gao *et al.*, 2017; Nanaei *et al.*, 2020). Milk production is polygenic; therefore, it is affected by several genes (Yudin & Voevoda, 2015; Silpa *et al.*, 2021). Similarly, fertility is a complex trait affected by various factors (Muller *et al.*, 2018). Cow fertility traits are the most economically important traits affecting the productivity of the dairy industry (Liu *et al.*, 2017). Cows have to calve to begin a new milk production cycle (Muller *et al.*, 2018). However, there is an inverse, and therefore unfavourable, correlation between fertility and milk production (Liu *et al.*, 2008a). In this review, the most frequent milk production and fertility-related genes detected in the included publications is presented in Table 5.2, and a visual representation thereof can be seen in Figure 5.10.

Table 5.2: The main milk production and fertility-related genes detected in the included publications.

Gene	Count	References
DGAT1	4	(Choi *et al.*, 2013), (Xu *et al.*, 2014b), (Gurgul *et al.*, 2015), (Mielczarek *et al.*, 2017)
IFNT	4	(Liu *et al.*, 2010), (Hou *et al.*, 2011a), (Bickhart *et al.*, 2012), (Zhang *et al.*, 2014)
PAG	4	(Hou *et al.*, 2011a), (Bickhart *et al.*, 2012), (Wang *et al.*, 2015), (Bickhart *et al.*, 2016)
PRAME	4	(Xu *et al.*, 2016), (Mustafa *et al.*, 2018), (Strillacci *et al.*, 2018), (Butty *et al.*, 2020)
PRL	4	(Cicconardi *et al.*, 2013), (Bagnato *et al.*, 2015), (Gurgul *et al.*, 2015), (Gao *et al.*, 2017)
AP3B1	2	(Shin *et al.*, 2014), (Yang *et al.*, 2021)
IGLL1	2	(Xu *et al.*, 2016), (Yang *et al.*, 2017b)
SLC27A6	2	(Mei *et al.*, 2019), (Mei *et al.*, 2021)
ITFG1	2	(Mei *et al.*, 2019), (Mei *et al.*, 2021)
MTHFSD	2	(Ben Sassi *et al.*, 2016), (Upadhyay *et al.*, 2017)
PRP	2	(Cicconardi *et al.*, 2013), (Bickhart *et al.*, 2016)
PTK2	2	(Bagnato *et al.*, 2015), (Di Gerlando *et al.*, 2019)

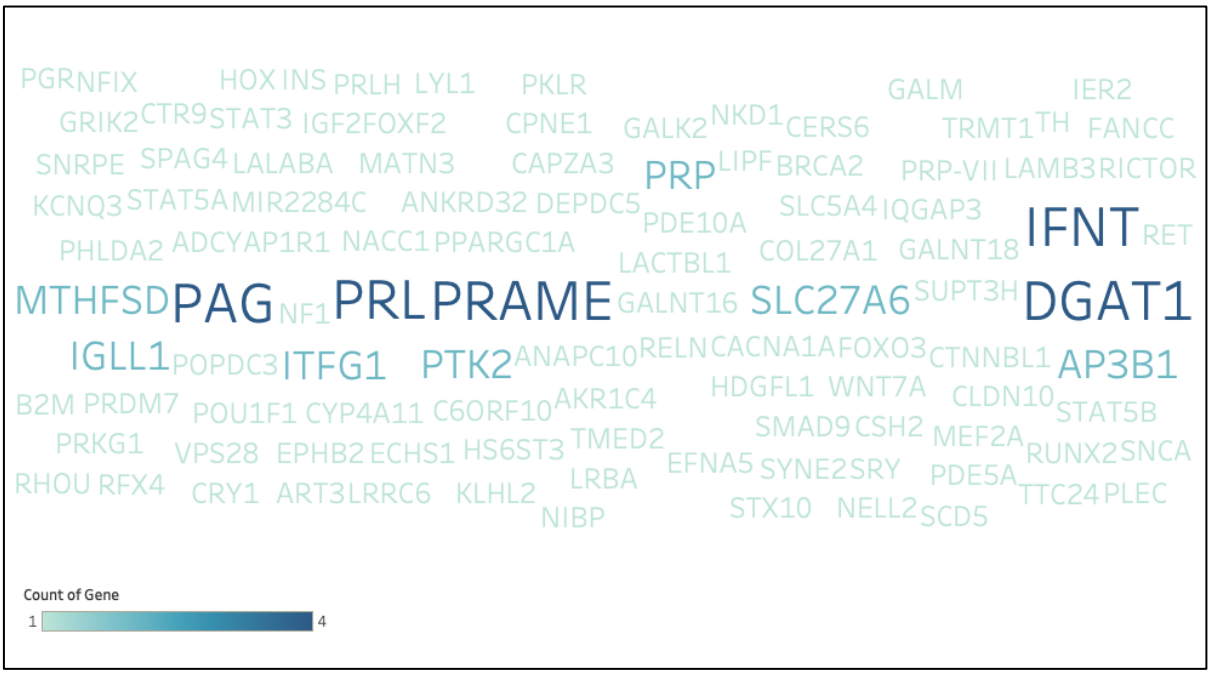

Figure 5.10: Visual representation of the milk production and fertility-related gene prevalence in included publications.

Diacylglycerol O-acyltransferase 1 (DGAT1)

The diacylglycerol O-acyltransferase (DGAT1) gene, found on BTA14, encodes a microsomal enzyme that catalyses the terminal step of triglyceride synthesis in eukaryotic cells (Winter *et al.*, 2002; Grisart *et al.*, 2004). The expression of DGAT1 has been reported in the liver, small intestine, mammary gland, and adipose tissue (Mohammed *et al.*, 2014). Triglycerides, the main components of fat, including the fat of secreted milk, are formed by covalently joining diacylglycerol to long chain fatty acyl CoAs (Winter *et al.*, 2002; Mohammed *et al.*, 2014). At least two enzymes catalyse this reaction, of which one is encoded by DGAT1 (Cases *et al.*, 2004). Thus, DGAT1 was proposed to be a functional candidate gene for milk production traits, after studies indicated that lactation impairment was observed in mice lacking both copies of this gene (Winter *et al.*, 2002). This gene has been studied extensively for production in dairy cattle. Moreover, genomic variation in the region harbouring DGAT1 was associated with milk fat content differences in breeds (Winter *et al.*, 2002). Additionally, DGAT1 polymorphisms were associated with milk fat yield and milk fat percentage in Italian Holstein cattle (Fontanesi *et al.*, 2014). Similarly, DGAT1 was associated with peak milk production (Ardicli *et al.*, 2018), thus genetic variation of DGAT1 can be used as a marker for milk production improvement in dairy cattle (Khan *et al.*, 2021).

Preferentially expressed antigen in melanoma (PRAME)

The preferentially expressed antigen in melanoma (PRAME) gene is a member of the cancer/testis (CT) genes, predominantly expressed in normal testis and in various tumours, thus understood to

be important for immunity and reproduction (Chang *et al.*, 2011). In the mammalian genome, this gene is one of the most amplified gene families, with approximately 90, 50, and 30 copies in the mouse, human, and bovine genome, respectively (Yue *et al.*, 2013). The PRAME gene plays a vital role in the male reproduction function in cattle (Chang *et al.*, 2011; Seleguim Chud *et al.*, 2018). Seleguim Chud *et al.* (2018) investigated copy number variation in the PRAME gene within Girolando, Gir and Holstein cattle breeds. In that study, only Gir bulls exhibited copy number variation in the PRAME region, suggesting that this region is variable only in the indicine lineage. However, Xu *et al.* (2016) compared European taurine and African taurine cattle and found the PRAME gene to be overlapping with CNV regions. Similarly, Strillacci *et al.* (2018) compared CNVs between Holstein and Brown Swiss cattle and found the PRAME gene to be harboured in a CNVR in both taurine breeds.

Chang *et al.* (2011) discovered a bovid-specific Y-linked PRAME family, PRAMEY, which was derived from a BTA17 transposition and underwent amplification on the Y chromosome during evolution. This Y-linked gene is likely to be important in spermatogenesis (Chang *et al.*, 2011). Yue *et al.* (2013) found the copy number of PRAMEY to be variable among cattle breeds, however, in Holstein cattle, PRAMEY was negatively correlated with percentage of normal sperm and testis size thus, indicating a low copy number of PRAMEY could be advantageous for bull fertility. The progression of genetic markers (such as CNVs of PRAMEY) could enable earlier prediction of male fertility and hasten genetic improvement for fertility traits.

Methenyltetrahydrofolate Synthetase Domain Containing (MTHFSD)

The methenyltetrahydrofolate Synthetase Domain Containing (MTHFSD) gene belongs to the folate metabolism genes. Folate metabolism genes are involved in the regulation of milk protein synthesis which is a complex biological process regulated at various levels within the mammary epithelial cells of dairy cattle (Menzies *et al.*, 2009). Variations in the MTHFSD gene have not been well explained in humans and animals and its function is indistinct. However, Ben Sassi *et al.* (2016) sought out to identify CNVRs associated with seven important traits in Spanish Holstein dairy cattle and found the MTHFSD gene to be associated with milk protein and fat yield.

Interferon tau (IFNT)

Interferons (IFNs) are signalling proteins that belong to the large class of proteins, namely cytokines (Walker & Roberts, 2009). Three subclasses of IFNs exist - type I, II and III. Type I IFN genes reside on BTA8 and are made up of at least nine subfamilies, namely, IFN-alpha, beta, delta, omega, epsilon, kappa, tau, zeta, and X (Walker & Roberts, 2009; Ealy & Wooldridge, 2017). Bickhart *et al.* (2012) found the interferon tau (IFNT) locus to be highly variable in copy number among animals. The function of IFNT is not immune related, but rather related to the maternal system for initiation and maintenance of pregnancy (Ealy & Wooldridge, 2017). During the peri-implantation window,

IFNT is the most abundant protein and transcript (Ealy & Wooldridge, 2017). Interferon tau is secreted exclusively by the trophectoderm cells of the ruminant conceptus and is recognised as the primary pregnancy recognition signal in ruminants (Shirasuna *et al.*, 2013; Forde & Lonergan, 2017). This is an indispensable activity for the continuation of pregnancy, and thus is essential for production. Moreover, IFNT stimulates interferon-stimulated genes and have increased expression for 21 days after insemination (Kowalczyk *et al.*, 2021). In most instances, embryo loss or lost pregnancies occur within the first four to five weeks of gestation. Thus, variation in IFNT could be a reliable candidate for a marker for early pregnancy detection (Kowalczyk *et al.*, 2021).

Pregnancy-associated glycoproteins (PAGs)

Pregnancy-associated glycoproteins (PAGs) are expressed products of the trophoblast cells in the placenta of even-toed ungulates (hoofed animals) (Telugu *et al.*, 2009). They are members of the large family of vertebrate aspartic peptidases, thus PAGs are directly linked to lysosomal enzymes such as cathepsin D, gastric enzymes such as chymosin and pepsin, or the enzyme renin, which is important for controlling blood pressure and maintaining sodium homeostasis (Telugu *et al.*, 2009). Thus, proteolytic activity in some PAGs is expected and their gene products are thought to play a functional role at the placenta-uterine interface. However, many bovine PAGs are incapable of being enzymatically active, adding to their complexity (Wallace *et al.*, 2015; Wiedemann *et al.*, 2018).

In cattle, at least 22 transcribed PAG genes have been reported, together with multiple variants (Telugu *et al.*, 2009). These genes exhibit various temporal and spatial patterns of expression (Wallace *et al.*, 2015). Phylogenetic analyses show that, in cattle, PAGs fall into two main groupings - ‘ancient’ and ‘modern’ PAGs, based on the time each group arose. Most PAGs form part of the modern group and are frequent in the Bovidae family (Telugu *et al.*, 2009; Wallace *et al.*, 2015). Bickhart *et al.* (2012) detected several PAG paralogs to be highly variable in copy number among individual animals (comprising members of the modern PAG groups). The authors suggest that this variability in the PAG family expansion could indicate important differences in reproductive traits among the individuals (Bickhart *et al.*, 2012). Filho *et al.* (2019) reported the potential of using the circulating concentration of day 24 PAG for early pregnancy diagnosis.

Prolactin (PRL) and prolactin-related proteins (PRPs)

Prolactin (PRL) is a multipurpose polypeptide hormone of the pituitary gland. The PRL gene has been mapped to BTA23 and is made up of five exons, coding for 199 amino acids, and four exons. Prolactin has important functions in reproduction, mammary gland development (mammogenesis), synthesis of milk secretion (lactogenesis), and maintenance of lactation (galactopoiesis) (Dong *et al.*, 2013). Moreover, PRL is also largely responsible for the synthesis of several milk components (such as lactose, lipids, and protein content), thus the PRL gene is an excellent candidate gene for milk traits (Patel & Chauhan, 2017). Dong *et al.* (2013) associated PRL with a higher milk yield at

305 days of lactation. A silent adenosine-guanine (A→G) mutation in the codon for amino acid 103 on exon 3 results in a polymorphic RsaI restriction site (Lewin *et al.*, 1992). Lewin *et al.* (1992) reported this locus to affect several milk production traits. The PRL-RsaI has become a popular genetic marker for the characterisation of cattle population using PCR-RFLP (Alipanah *et al.*, 2007). Alfonso *et al.* (2012) investigated the relationship between polymorphisms in the PRL gene and milk production traits in American Swiss cattle and concluded that animals with genotype PRL-RsaI AA had higher milk production. Similarly, Patel & Chauhan (2017) detected the allelic and genotypic frequencies of the PRL gene in Gir and Kankrej cattle for the analysis of milk production traits and found cattle with the PRL-RsaI AA genotype had a higher milk yield, but a lower milk fat percentage. Bayıl Oğuzkan & Bozkurt (2019) aimed to determine the relationship between daily milk production and exon 3 region of the PRL gene and considered it as a candidate gene in MAS. This gene is not only involved in lactation, but also in hair morphology and thermoregulation (Kim *et al.*, 2017), as seen in section 5.4.

The bovine placenta produces a range of proteins that are both functionally and structurally like pituitary prolactin (Soares, 2004). Bovine placental lactogen (PL), and various prolactin-related proteins (PRPs) have been identified (Soares, 2004; Ushizawa *et al.*, 2005). These molecules are expressed in the placenta of ruminants and are important for implantation and placentation (Yamada *et al.*, 2002). Ten bovine PRP genes have been identified, PRP-I – PRP-X (Takahashi *et al.*, 2008), and are highly expressed in the placental binucleate cells of bovine trophoblasts (Ushizawa *et al.*, 2005). Cicconardi *et al.* (2013) identified a breed specific CNVR that contained genes belonging to the PRP family (PRP 1,3,4,6,9, and PRP-VII). Furthermore, Bickhart *et al.* (2016) detected a cluster of PRP genes with a high degree of duplication (96% of animals) and suggested that these genes are likely subject to subfunctionalisation (duplication splits function) and neofunctionalisation (duplication generates a new function) (Figure 5.11).

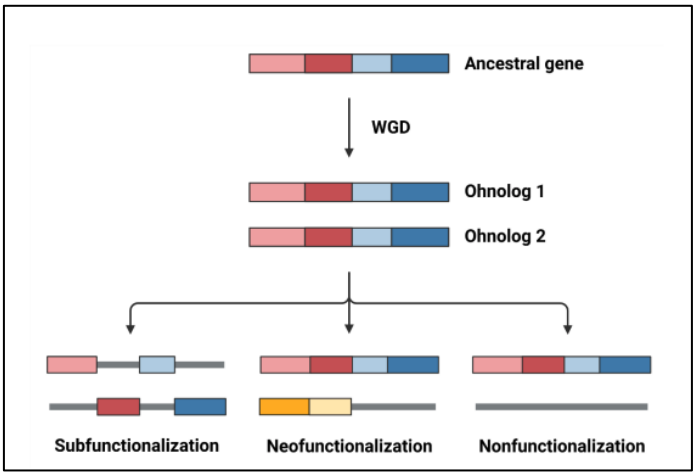

Figure 5.11: Fate of genes following duplication (subfunctionalisation, neofunctionalisation and nonfunctionalisation). From Sandholm (2022).

Solute carrier family 27 member 6 (SLC27A6)

The solute carrier 27A (SLC27A) gene consists of six members (SLC27A1- SLC27A6) that encode fatty acid transporter proteins (Pecka-Kiełb *et al.*, 2020). Isoform 6 is the predominant isoform in this gene family (Bionaz & Loor, 2008). Milk lipid biosynthesis occurs in the mammary gland and entails fatty acid (FA) uptake from plasma into the bovine mammary epithelial cells, followed by FA transport within the cells (Nafikov *et al.*, 2013). The uptake of these long-chain fatty acids is mediated by FA transport proteins (FATPs), specifically SLC27 proteins (Bionaz & Loor, 2008). Zhang *et al.* (2021) identified and characterised the genes related to FA metabolism throughout the lactation phases and found SLC27A6 to have a central coordinative role in FA metabolism in the bovine mammary epithelial cells.

Milk has many important nutritional components, one of which is the FA composition. Fatty acid composition is especially important due to its link with human health (Dixit *et al.*, 2015). A high content of saturated fatty acids (SFA) is associated with heart disease, obesity, and diabetes, whereas unsaturated fatty acids (UFA) are beneficial to health, especially for their impact of cholesterol levels (Arnould & Soyeurt, 2009). Nafikov *et al.* (2013) investigated the association of polymorphisms in SLC27A6 with FA composition of bovine milk, with the aim to develop genetic markers for the selection of animals producing milk containing a lower concentration of SFA and a higher concentration of UFA. In that study, polymorphisms in SLC27A6 could be utilised for the selection of animals with lower SFA and higher UFA.

Immunoglobulin lambda-like polypeptide 1 (IGLL1)

In domestic cattle, the immunoglobulin lambda-like polypeptide 1 (IGLL1) forms part of IgL chain gene pool that are involved in B lymphocyte production (Ekman *et al.*, 2009). Although immunoglobins, through their antibody activity, are an important factor in disease resistance in the body (Zhao *et al.*, 2021), this gene has also been shown to be differentially expressed in the mammary tissue in dairy cattle and to affect production and fertility traits in cattle. Li *et al.* (2016) identified IGLL1 as one of the top 20 genes expressed in the bovine mammary tissue during the non-lactating period in Chinese Holstein cows. Cerri *et al.* (2012) found the IGLL1 gene to be upregulated throughout the peripartum period in the endometrium in lactating dairy cows, which indicated the increase in B-lymphocyte and γδ T-cell activity affected fertility. Similarly, Minozzi *et al.* (2013) found IGLL1 to be associated with fertility in Italian Holstein cattle. It was suggested that the variant in the IGLL1 gene be a marker for fertility traits in dairy cattle (Frischknecht *et al.*, 2017).

Integrin alpha FG-GAP repeat containing 1 (ITFG1)

Integrins are heterodimers that consist of alpha (α) and beta (β) subunits and are expressed in various cells (Hynes, 1992). They are cell surface receptors that control adhesion to the extracellular matrix (ECM) and certain cell-ECM interactions (Taddei *et al.*, 2003). Integrins have been reported as important regulators in mammary gland development (Taddei *et al.*, 2003), however research on integrin α FG-GAP repeat containing 1 (ITFG1) gene in cattle is limited. The ITFG1 gene is found on BTA18 and is expressed in bovine mammary gland (Stella *et al.*, 2010). Stella *et al.* (2010) aimed to detect signatures of selection in cattle breeds selected for dairy production and discovered approximately 700 putative selection signatures, and ITFG1 was identified as a statistically significant signature for dairy production.

Adaptor related protein complex 3 subunit β-1 (AP3B1)

Adaptor-related protein (AP) complexes are involved in cargo selection and vesicle formation, and thus the trafficking of proteins in the intracellular membrane (Simpson *et al.*, 1997). Five AP complexes have been identified, AP-1 to AP-5, each of which have different functions and distinct localisations (Nakatsu & Ohno, 2003; Adamopoulos, 2018). The adaptor related protein complex 3 subunit β-1 (AP3B1) gene encodes a protein involved in organelle biogenesis linked to platelet dense granules, melanosomes, and lysosomes (Yang *et al.*, 2021). Cochran *et al.* (2013) identified AP3B1 to be related to daughter pregnancy rate in Holstein cattle and suggested that it may also affect the release of neurotransmitters, which control the hypothalamic pituitary complex. Ortega *et al.* (2016) found AP3B1 to be associated with three fertility traits, namely cow conception rate, daughter pregnancy rate, and heifer conception rate. The AP3B1 gene has been reported to be involved in coat colour in cattle and other domestic animals, as seen in section 5.4.

Protein tyrosine kinase 2 (PTK2)

Protein tyrosine kinase 2 (PTK2) is a focal adhesion-associated protein kinase that regulates cellular processes such as cell motility, assembly and disassembly of focal adhesions, cell proliferation, and apoptosis (Ahn & Park, 2010; Zhao & Guan, 2011). Although a few studies have identified this gene in variable regions (Bagnato *et al.*, 2015; Di Gerlando *et al.*, 2019), research surrounding the function of PTK2 in cattle is limited. However, Wang. *et al.* (2013) presented evidence for associations of PTK2 variants with milk production traits in Chinese Holstein cattle.

5.4 CNV-related genes affecting meat, growth, and feed-efficiency traits

In beef cattle, carcass traits and meat quality have a significant impact on product pricing and consumer satisfaction. Carcass traits are related to the yield grade, carcass weight, backfat thickness, and rib eye area (Drake, 2004). The quality of the meat is based on organoleptic traits, namely, flavour, tenderness, colour, and juiciness (Feitosa *et al.*, 2014). Intramuscular fat (IMF) plays a major role in meat marbling, which determines the texture and flavour of the meat. Tenderness is one of the main determinants in consumer approval of beef, thus improving meat quality is very important to the beef industry.

Growth traits, such as body weight measurements or visual scores of body conformation, affect carcass selection and therefore beef production. Growth traits affect the development, structure, and size of livestock (Guo *et al.*, 2020a). This complex, quantitative trait is controlled by nongenetic and genetic factors, however, genetic factors are main contributors to animal growth (Zhou *et al.*, 2016a; Guo *et al.*, 2020a).

Feed efficiency has been described in various ways (Berry & Crowley, 2013). Feed conversion ratio (feed:gain), or its inverse, feed conversion efficiency (gain:feed) were used at large previously (Kenny *et al.*, 2018). More recently, residual feed intake, defined as the difference between actual feed intake and the predicted intake required for maintenance of body weight and production (Hou *et al.*, 2012a), has become the favoured measurement (Kenny *et al.*, 2018). Feed is the most variable and unpredictable cost in livestock production systems, making feed intake and efficiency vital economic traits (Sherman *et al.*, 2010).

In this review, the most frequent meat production and growth-related genes detected in the included publications is presented in Table 5.3, and a visual representation thereof can be seen in Figure 5.12. These genes include IGF2, PLA2G2D and CAST, followed by IGF1R, APOL3, PTPRC and KCNJ12, and lastly, CAPN1, AGBL3, CTNNA1, MSTN, ADRA1B, ATRN, LRRC49, MYH3, SORCS2 and TG. The most frequently detected feed-efficiency related genes is presented Table 5.4, with a visual representation thereof in Figure 5.17, these genes include PRKG1, FABP2, and EIF2S1.

5.4.1 Meat and growth-related genes

Table 5.3: The main meat production and growth-related genes detected in the included publications.

Gene	Count	References
IGF2	7	(Hou *et al.*, 2012b), (Gurgul *et al.*, 2015), (Da Silva *et al.*, 2016b), (Da Silva *et al.*, 2016a), (Carmo *et al.*, 2019), (Di Gerlando *et al.*, 2019), (Berton *et al.*, 2021)
CAST	4	(Da Silva *et al.*, 2016a), (Kommadath *et al.*, 2019), (Mei *et al.*, 2020), (Guo *et al.*, 2021)
PLA2G2D	4	(Stothard *et al.*, 2011), (Zhang *et al.*, 2014), (Xu *et al.*, 2017), (Berton *et al.*, 2021)
APOL3	3	(Bickhart *et al.*, 2012), (Da Silva *et al.*, 2016b), (Cozzi *et al.*, 2019)
PTPRC	3	(Hou *et al.*, 2012c), (Shin *et al.*, 2014), (Yang *et al.*, 2021)
KCNJ12	3	(Xu *et al.*, 2016), (Zhou *et al.*, 2016a), (Huang *et al.*, 2021b)
AGBL3	2	(Choi *et al.*, 2013), (Da Silva *et al.*, 2016a)
CAPN1	2	(Da Silva *et al.*, 2016a), (Guo *et al.*, 2021)
CTNNA1	2	(Hu *et al.*, 2020b), (Jang *et al.*, 2021)
MSTN	2	(Rafter *et al.*, 2021b), (Yang *et al.*, 2021)
ADRA1B	2	(Wang *et al.*, 2016b), (Mei *et al.*, 2020)
ATRN	2	(Xu *et al.*, 2017), (Mei *et al.*, 2019)
TG	2	(Bagnato *et al.*, 2015), (Mei *et al.*, 2021)
MYH3	2	(Zhang *et al.*, 2014), (Xu *et al.*, 2017)
LRRC49	2	(Strillacci *et al.*, 2018), (Lee *et al.*, 2020)
SORCS2	2	(Da Silva *et al.*, 2016b), (Strillacci *et al.*, 2018)

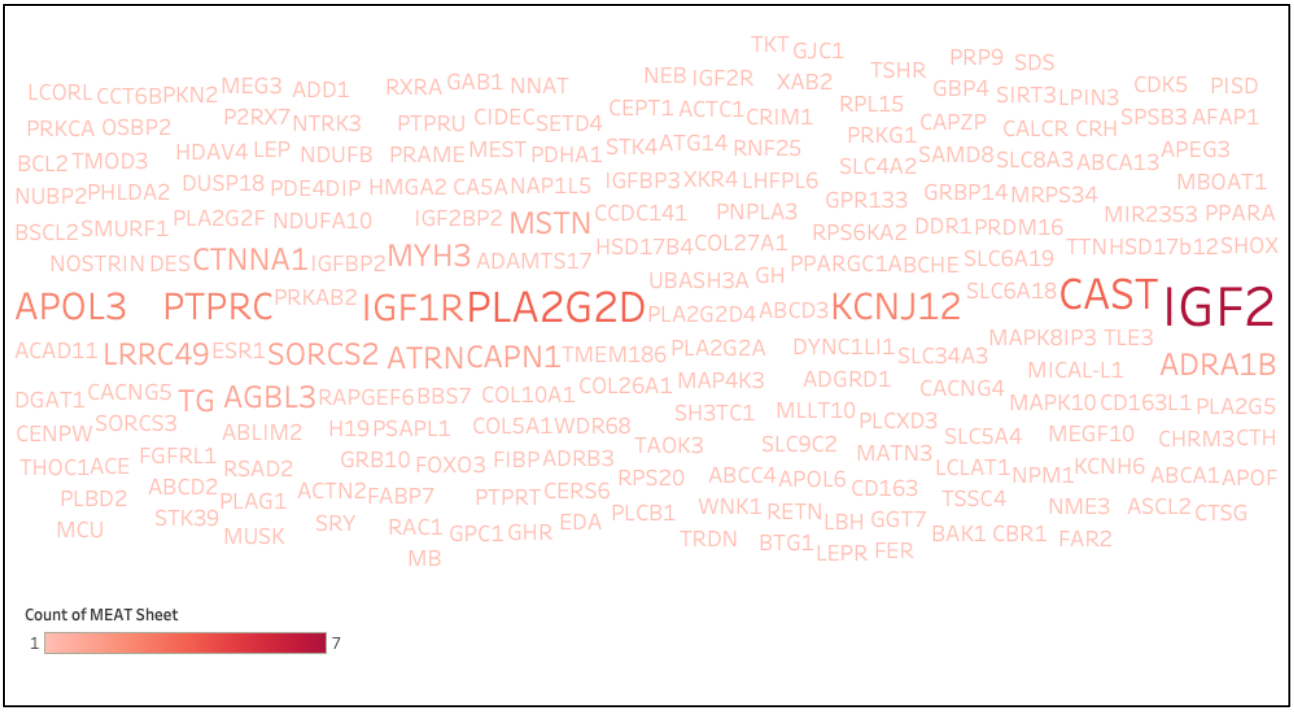

Figure 5.12: Visual representation of meat production and growth-related gene prevalence in included publications.

Insulin-like growth factor (IGF) family

The insulin-like growth factor (IGF) system is important for both pre- and postnatal growth and development (Baker *et al.*, 1993; Allan *et al.*, 2001). Components of the IGF system include two growth factors (IGF1 and IGF2), type 1 and 2 IGF receptors (IGF1R and IGF2R), a family of six major IGF binding proteins (IGFBP1 – IGFBP6), and IGFBP protease (Figure 5.13) (Allan *et al.*, 2001). Both IGF1 and IGF2 are expressed ubiquitously, thus they have strong growth promoting paracrine, autocrine, and endocrine actions in several pre- and postnasal tissues. Insulin-like growth factors undergo changes in gene expression throughout prenatal development and during postnasal growth (Ghanipoor-Samami *et al.*, 2018). Circulating IGF1 concentration is associated with heifer body weight during prepubertal growth (Lammers *et al.*, 1999), and studies have revealed an increase in circulating IGF1 as heifers approached puberty.

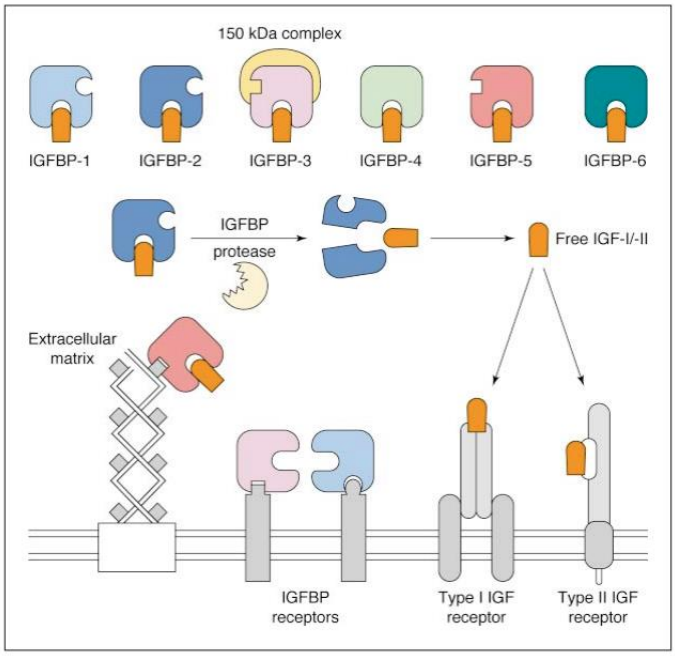

Figure 5.13: Components of the IGF system. From Allan *et al.* (2001).

Both IGF2 and IGF1R were identified at a high frequency in this review. Type 1 IGF receptor (IGF1R), the main receptor with which IGFs present biological effects (Yang *et al.*, 2013), regulates IGF half-life and activity. This receptor binds IGF1 with high affinity, and IGF2 with a lower affinity to initiate a repertoire of cellular responses, and is in involved with immune regulation, and muscle and bone growth (Adams *et al.*, 2000). Jang *et al.* (2021) suggested that copy number differences within IGF1R could contribute to differences in body size among populations. Ma *et al.* (2019) investigated CNVs in the IGF1R gene across four Chinese beef breeds. In that study, variation in IGF1R was significantly associated with body weight, body height, and hucklebone width in some breeds. Ma *et al.* (2019) proposed the IGF1R CNV to be a promising marker for the improvement of meat production in breeding programs.

The IGF2 gene, a paternally expressed imprinted gene located on BTA29, is expressed in most tissues affecting the lean muscle content in pigs, mice, and cattle (Goodall & Schmutz, 2007). It is involved in the regulation of cell proliferation, migration, differentiation, and survival (Berton *et al.*, 2021), and plays an important role in embryonic growth and development. This gene has been associated with rib eye area (Goodall & Schmutz, 2007; Berton *et al.*, 2021), meat tenderness in Landrace pigs (Rehfeldt *et al.*, 2012) and Nellore cattle (Da Silva *et al.*, 2016a), growth traits in Qinchuan cattle (Huang *et al.*, 2014), and backfat, ADG and FCR in beef cattle (Sherman *et al.*, 2008).

Phospholipase A2 group IID (PLA2G2D)

The family of phospholipase A2 (PLA2) are low-molecular-weight enzymes that hydrolyse at the sn-2 position of membrane phosphoglycerides, to form free fatty acids lysophospholipids (Figure 5.14A) (Golik *et al.*, 2006). The PLA2s consist of five types of enzymes, of which one is secretory PLA2 (sPLA2), more than one-third of the PLA2 enzymes belong to the sPLA2 family (Murakami *et al.*, 2015; Yang *et al.*, 2022a). Moreover, the sPLA2 family contains at least ten isozymes, organised according to their structural features. The genes for sPLA2-2A, -2C, -2D, -2E, -2F and 5 are grouped on the same chromosome and are thus referred to as group 2 subfamily secretory phospholipase A2 (PLA2G2) (Figure 5.14B) (Lappas & Rice, 2004).

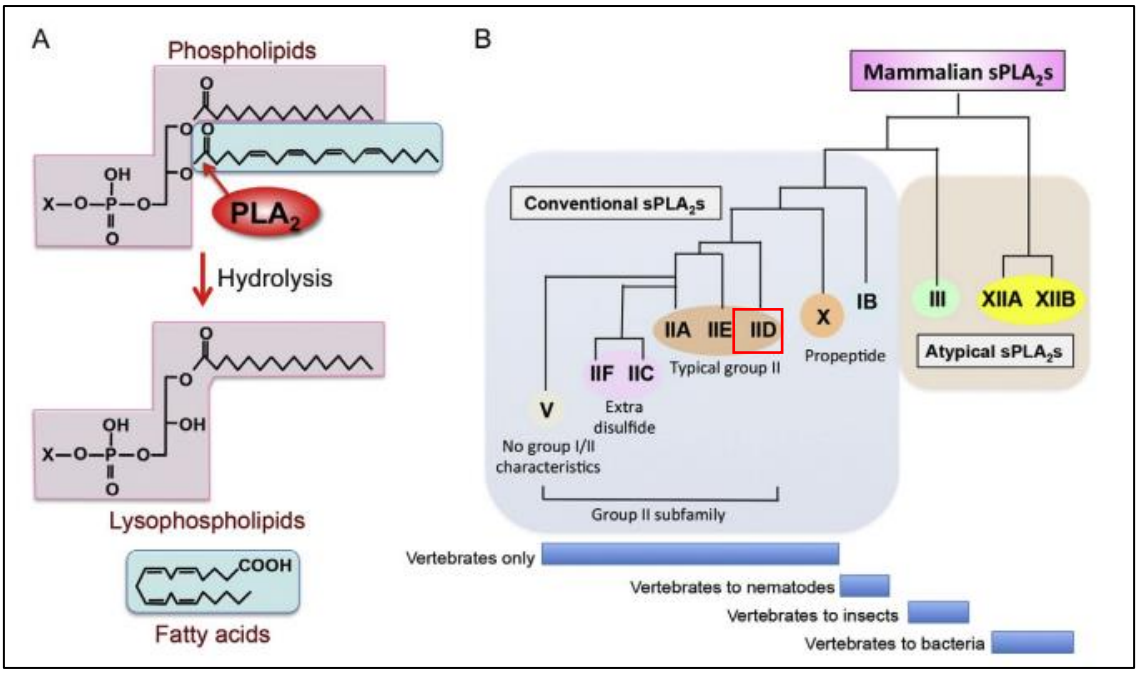

Figure 5.14: A) Hydrolysis of phospholipids, forming free fatty acids and lysophospholipids; B) Organisation of the sPLA2 family, highlighting group 2D. From Murakami *et al.* (2015).

Golik *et al.* (2006) characterised a group of bovine PLA2G2D-like genes on BTA2, whereafter Stothard *et al.* (2011) reported CNVRs covering the PLA2G2D, involved in fat deposition, lipid metabolism, and gonadotropin-releasing hormone signalling, to be highly duplicated in beef cattle breeds, thus signifying its association with meat quality traits. Berton *et al.* (2021) performed a GWAS between CNVRs and meat quality and carcass traits in Nellore cattle. The researchers found CNVRs harbouring candidate genes, such as the PLA2 family genes with functions related to lipid metabolism, to be associated with meat colour factors, such as lightness (L*).

The PLA2G2D gene is the most structurally similar to PLA2G2A (Murakami et al., 2016). Yang *et al.* (2022) discovered copy number variation of the PLA2G2A gene had a significant effect on growth traits in two breeds of Chinese cattle. Taye *et al.* (2017) identified PLA2G2A to be a candidate gene

for intramuscular fat in Ankole cattle. Zhang *et al.* (2014) detected 370 CNVRs in 24 taurine cattle from 12 breeds and identified a CNVR to have significant negative effects on cattle body measurements in PLA2G2D gene. This indicates that group 2 subfamily secretory phospholipase A2 can be used as a molecular marker for meat quality traits and growth-related traits.

Calpastatin (CAST) and calpain 1 (CAPN1)

Meat tenderness (MT) is primarily influenced by the amount of connective tissue, marbling, or intramuscular fat, and post-mortem myofibrillar protein degradation (Leal-Gutiérrez *et al.*, 2018). Two enzymes are responsible for myofibrillar protein degradation, namely calpain 1 (CAPN1), and its inhibitor, calpastatin (CAST) (Casas *et al.*, 2006a). The CAPN1 and CAST genes are located on BTA29 and BTA7, respectively (Bishop *et al.*, 1993; Smith *et al.*, 2000). The calpain-calpastatin system is the most widely studied enzyme system involved in the meat tenderization process (Koohmaraie & Geesink, 2006; Casas *et al.*, 2006b; Leal-Gutiérrez *et al.*, 2018; Lee *et al.*, 2019). Calpains are endogenous proteases primarily responsible for post-mortem muscle protein degradation (Bhat *et al.*, 2018). The CAPN gene cuts proteins into fragments, thus disrupting the structure of the muscle cells and contributing to meat tenderness, while CAST inhibits this degradative action (Bhat *et al.*, 2018). Elevated CAST activity is therefore associated with reduced MT in cattle (Tizioto *et al.*, 2014).

Moravčíková *et al.* (2019) completed a selection signature analysis in beef cattle and found CAST to be a candidate gene for MT. While Mei *et al.* (2020) sequenced six Chinese indigenous cattle breeds and identified several CNVR-related genes to be associated with meat production and quality, such as CAST. Lee *et al.* (2019) investigated SNPs in the CAPN1 and CAST genes in Hanwoo cattle and validated that mutations in these genes are strongly associated with Warner-Bratzler shear force. Similarly, Guo *et al.* (2021) investigated variants associated with meat traits in Qaidam cattle and suggested that the tenderness of Qaidam beef might be due to the discrepancy in copy number of the CAST and CAPN1 genes.

Potassium inwardly rectifying channel, subfamily J, member 12 (KCNJ12)

The KCNJ12 gene belongs to the inwardly rectifying K^+ (Kir) channel family, which allows K^+ to move more easily into the cell (Hibino *et al.*, 2010). This gene encodes the inward-rectifier K^+ channel protein, Kir2.2, and is linked to the cardiac and nerve inward rectifier current (Cheng *et al.*, 2019b). Interestingly, KCNJ12 is often expressed in cardiac myocytes (cells responsible for contractile force in the heart), and animal neuronal cells, which controls cell excitability (Zheng *et al.*, 2019). Motor neurons are essential for the regulation of the function and properties of skeletal muscle. The Kir channels are important for controlling muscle excitability, thus KCNJ12 is suggested to be involved in the excitation–contraction regulation (Cheng *et al.*, 2019b).

The KCNJ12 gene is located on BTA19 and reported to have a significant influence on growth traits in cattle and found to participate in the muscle contraction process, thus this could be the mechanism used for regulating muscle cell growth and contraction. Zhou *et al.* (2016) performed a CNV-based GWAS of growth traits in *Bos indicus* cattle and concluded that KCNJ12 could be a candidate gene for muscling. Zheng *et al.* (2019) examined the distribution and association of the KCNJ12-variable gene across four Chinese cattle breeds. The researchers concluded that differentiation in copy number of this gene could be used as a reliable genetic marker for the early selection of growth traits in breeding programmes.

Myostatin (MSTN)

Myostatin (MSTN), also recognised as Growth and Differentiation Factor 8 (GDF8), is a circulating factor secreted by muscle cells and functions as a regulator of muscle cell proliferation (Aiello *et al.*, 2018). This gene regulates skeletal muscle growth at crucial points during the process of pre-natal muscle development, thus determining the number of myofibers (Bi *et al.*, 2020). These key points include muscle precursor proliferation, myoblast proliferation, and differentiation (Figure 5.15).

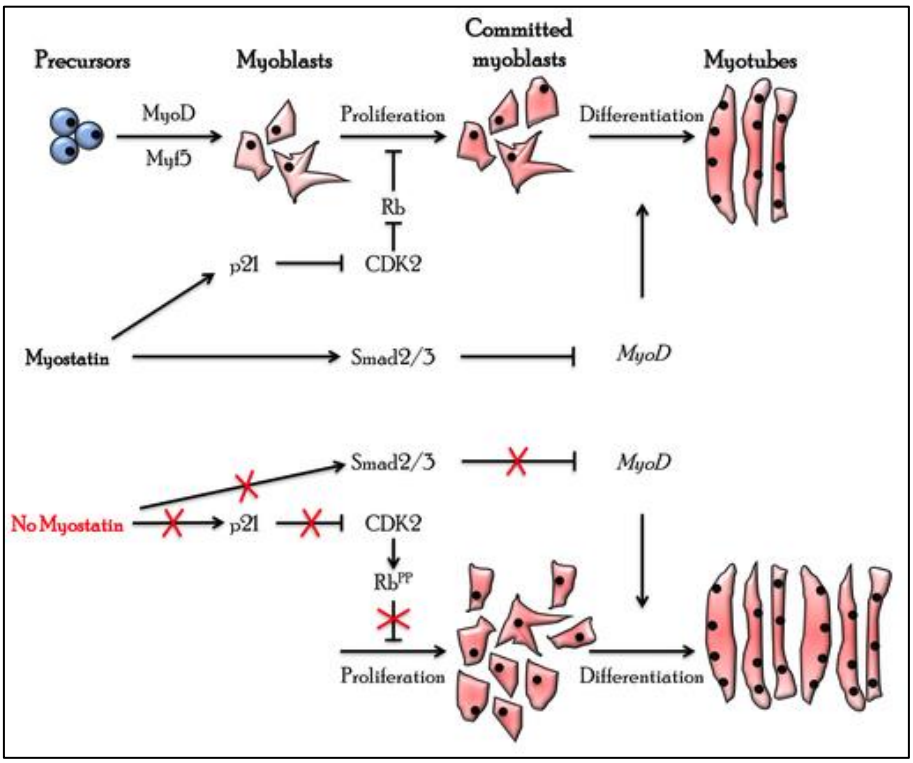

Figure 5.15: Above: Myostatin activity during proliferation and differentiation of myoblasts. Below: The absence of MSTN, resulting in increased proliferation and differentiation of myoblasts. From Aiello *et al.* (2018).

The bovine MSTN is found on BTA2, and mutations within this gene results in the loss of function, which leads to hyperplasia, more commonly known as 'double muscling' (Figure 5.16) (Grobet *et al.*, 1997). Hyperplasia is the increase in the number of muscle fibres, whereas hypertrophy is the enlargement of individual muscle fibres (Kambadur *et al.*, 1997). The double muscling phenotype is

a heritable condition. More than 20 different mutations (insertions, deletions, and SNPs) have been described in the cattle MSTN (Aiello *et al.*, 2018). Myostatin mutations in cattle can lead to high carcass yields and excellent conformation, thereby benefitting the meat industry (Haruna *et al.*, 2020). However, animals with the double muscling phenotype are more prone to respiratory diseases, lameness, nutritional stress, dystocia, and other reproductive issues (Bellinge *et al.*, 2005; Aiello *et al.*, 2018). Muscular hypertrophy has been most extensively analysed in the Piedmontese and Belgian Blue breeds (Kambadur *et al.*, 1997; McPherron & Lee, 1997; Grobet *et al.*, 1997; Miretti *et al.*, 2013; Jakaria *et al.*, 2021). Yang *et al.* (2021) sequenced 44 Simmental beef cattle and found a deletion CNV downstream of the MSTN gene and linked it to important traits such as muscle differentiation.

Figure 5.16: Belgian Blue cattle with the double muscling phenotype. From McPherron & Lee (1997).

Catenin alpha 1 (CTNNA1)

The α-catenin is an important protein found at cell adherens that provides connection between the cadherin-based adhesion complex and the actin cytoskeleton (Sheikh *et al.*, 2006). Catenin alpha 1 (CTNNA1) has been reported to be associated with myostatin expression level and function in the skeletal muscle of Holstein-Friesian bulls (Sadkowski *et al.*, 2008). Jang *et al.* (2021) found CTNNA1 to be overlapped with an indicine-specific deletion, indicating the probable likelihood of the sequence being specific to taurine cattle. Similarly, Hu *et al.* (2020b) analysed CNVs between taurine and indicine cattle and found all significant CNV segments (of which CTNNA1 was overlapped) to have a high ratio of deletion in zebu cattle. The researchers speculated that the CTNNA1 gene may be related to the difference in muscle development and meat productivity between taurine and zebu cattle (Hu *et al.*, 2020b; Jang *et al.*, 2021).

Thyroglobulin (TG)

Thyroglobulin (TG) is a homodimeric glycoprotein hormone that is produced in thyroid follicular cells (Gan *et al.*, 2008). Thyroglobulin is necessary to produce thyroid hormones such as triiodothyronine

(T3) and thyroxine (T4), which are involved in the regulation of metabolism and impacts adipocyte growth, differentiation, and the homeostasis of fat deposition (Zhang *et al.*, 2015a). The TG, located in the centromeric region of BTA14, is considered a positional and functional candidate for fat production (Casas *et al.*, 2006a). Thyroglobulin has been extensively studied and linked to carcass traits and fat distribution in cattle (Mears *et al.*, 2000; Casas *et al.*, 2007; Hou *et al.*, 2011b; Bennett *et al.*, 2013). Variation in the 5'-flanking region of thyroglobulin has been significantly associated with marbling score and can therefore be used in MAS programmes for the improvement of meat quality (Wood *et al.*, 2006; Gan *et al.*, 2008; Hou *et al.*, 2011b).

Myosin heavy chain 3 (MYH3)

Myosin proteins comprise heavy chains, associated with the speed of contractions, and light chains, whose function is not well defined, both of which are present in skeletal muscle (Wang *et al.*, 2013). Myosin is one of the main structural proteins of the thick filament of the sarcomere (Zhang *et al.*, 2011). The myosin heavy chain 3 (MYH3) gene, a member of the myosin heavy chain family, is a major contractile protein (Wang *et al.*, 2013) that converts chemical energy to mechanical energy through ATP hydrolysis (Niu *et al.*, 2013). The MYH3 gene is mostly expressed in skeletal muscle at different developmental stages and is important in the development of heart and skeletal muscles (Xu *et al.*, 2014c). This gene, found on BTA19, is involved in muscle development, differentiation, and contractions of striated muscles (Zhang *et al.*, 2011). Studies have shown that MYH3 is associated with meat quality and growth traits. Zhang *et al.* (2011) indicated that MYH3 may be an important gene involved in the muscle fibre property differences seen in intact and castrated male Qinchuan cattle. Wang *et al.* (2013) found that MYH3 gene polymorphisms affected growth and carcass traits in Qinchuan cattle. Whereafter, Xu *et al.* (2014b) found that CNVs of the MYH3 gene had a positive correlation with growth traits.

Sortilin related VPS10 domain containing receptor 2 (SORCS2)

Sortilin related VPS10 domain containing receptor 2 (SORCS2) belongs to the Vps10p-domain family of neutral receptors (Rezgaoui *et al.*, 2001). This proneurotrophin receptor is expressed as a single-chain protein that regulates dopaminergic axon guidance and has been shown to be important for pro-nerve growth factor (NGF) mediated growth cone collapse (Glerup *et al.*, 2014; Boggild *et al.*, 2016). However, in cattle, SORCS2 has been reported to be linked to meat quality traits. In this review, da Silva *et al.* (2016) identified a CNV region on BTA6 harbouring SORCS2 to be duplicated in over 1000 Nellore cattle and deleted in 20 Nellore cattle. This gene has been related to lipid metabolism in various mammalian species and has been suggested to be linked to backfat thickness in Nellore cattle (Júnior *et al.*, 2016). Backfat thickness influences the preservation of the carcass post slaughter and is plays a role in organoleptic characteristics evaluated by the consumer (Veneroni-Gouveia *et al.*, 2011).

Leucine rich repeat containing 49 (LRRC49)

The Leucine rich repeat containing 49 (LRRC49) gene codes a protein containing seven conserved leucine-rich repeats (de Souza Santos *et al.*, 2008). Leucine rich repeats are made up of 20-29 amino acids and are present in many proteins with various functions and are important for protein-protein interactions (Kobe & Kajava, 2001). In humans, LRRC49 has been reported to be involved in breast cancer (de Souza Santos *et al.*, 2008), while in cattle this gene has been associated with meat production traits. This gene has been associated with marbling score and subcutaneous fat in Canchim beef cattle (Mokry *et al.*, 2013). Strillacci *et al.* (2018) compared CNVRs between Valdostana Red Pied and Italian Brown Swiss and found a CNVR on BTA10 containing LRRC49. The Valdostana Red Pied is a dual-purpose breed with specific directional selection for meat quality (Strillacci *et al.*, 2018). Lee *et al.* (2020) identified CNVs in two dairy cattle breeds, Jersey and Holstein-Friesian, and identified LRRC49 in a variable region and related it to body size, suggesting that it could be linked to the size difference between these two breeds.

Adrenoceptor alpha 1B (ADRA1B)

Adrenergic receptors form part of the G-protein-coupled receptor family, located in the cell membrane (Inderwies *et al.*, 2003). Adrenergic receptors comprise alpha receptors (α_1 and α_2) and beta receptors (β_1, β_2 and β_3) (Inderwies *et al.*, 2003). Three α_1-adrenergic receptor subtypes exist, namely α_{1A}, α_{1B} and α_{1D}, which are involved in neurotransmission and regulation of the sympathetic nervous system in humans (Perez, 2020). Although the precise function of adrenoceptor alpha 1B (ADRA1B) in cattle is unclear, Lee *et al.* (2013a) identified ADRA1B in the regions of homozygosity in Hanwoo cattle and proposed it as candidate gene related to meat traits and disease resistance. Mei *et al.* (2020) identified CNVR-related genes in Japanese Black cattle and Red Angus cattle, such as ADRA1B, and related it to meat quality and production traits.

ATP/GTP binding protein-like 3 (AGBL3)

The ATP/GTP binding protein-like 3 (AGBL3) is a proteolysis-associated gene. Previous research on meat tenderness and skeletal muscle described GTP-related genes to be involved in myotube morphology and skeletal muscle myogenesis. The functional role of AGBL3 is unclear, however, the gene ontology term for AGBL3 is proteolysis (Choi *et al.*, 2013), which is the main contributing factor to meat tenderness in the muscle fibre. Choi *et al.* (2013) found this gene to be present in fewer copies in dairy cattle (Holstein) compared with beef cattle (Hanwoo), and due to beef and dairy breeds differing in meat tenderness, this suggests the AGBL3 locus may have an effect on tenderness. Da Silva *et al.* (2016) performed a GWAS to identify CNVs associated with meat tenderness and found a variable region downstream of AGBL3. The researchers recommend further investigation be done on whether variation the AGBL3 locus affect meat tenderness.

Protein tyrosine phosphatase receptor type C (PTPRC)

The protein tyrosine phosphatase receptor type C (PTPRC) gene encodes the CD45 receptor. The CD45 is a transmembrane glycoprotein found in the plasma of most haematopoietic cells and is important for regulating the antigen receptor signalling of T- and B-cells (Al Barashdi *et al.*, 2021). Members of the protein tyrosine phosphatase (PTP) superfamily control signalling pathways and are thus important in many cellular processes (Xie *et al.*, 2021). Twenty-one receptor-like PTP genes have been reported in the human genome (Alonso *et al.*, 2004). In cattle, this gene has been related to meat production and immunity traits. Yang *et al.* (2021) sequenced a population of Chinese Simmental beef cattle and found PTPRC located upstream of a deletion associated with tenderloin. The PTPRC gene has been identified in CNV regions in Angus beef cattle (Hou *et al.*, 2012c) and Hanwoo beef cattle (Shin *et al.*, 2014). This gene has also been reported to be related to porcine meat quality (Wan *et al.*, 2016). Protein tyrosine phosphatase receptor type K has been associated with tenderness and marbling scores in beef cattle (Gao *et al.*, 2014; Braz *et al.*, 2019). Interestingly, PTPRC has also been related to bovine immune response to ticks (Jonsson *et al.*, 2021) and intestinal nematodes (Araujo *et al.*, 2009). This is to be expected as CD45 plays a major role in the innate immune system (Al Barashdi *et al.*, 2021).

Attractin (ATRN)

Attractin (ATRN) encodes transmembrane proteins and secretory proteins and is widely expressed in the central nervous system (Kim *et al.*, 2005). Due to the wide distribution of this gene in the central nervous system, it is involved in a broad spectrum of functions such as the regulation of pigmentation, myelination, immune system and body weight regulation, and tumour susceptibility (Kim *et al.*, 2005). Previous studies in mice have indicated that there is an association between the ATRN locus products and growth and carcass characteristics (Nagle *et al.*, 1999; Gunn *et al.*, 2001). Attractin has been linked to body weight and fatness in porcine (Kim *et al.*, 2005). Agouti Signalling Protein (ASIP) has been identified as a promising candidate gene for bovine fat deposition (Albrecht *et al.*, 2012), and ATRN has been identified as a low-affinity receptor for this protein (He *et al.*, 2001). Bovine ATRN has been mapped to chromosome 13 (Edeal *et al.*, 2000; Graphodatskaya *et al.*, 2003). Lee *et al.* (2013a) identified ATRN and considered it a candidate gene associated with meat traits, while Liu *et al.* (2018) found ATRN to influence fat deposition in the bovine.

Apolipoprotein L3 (APOL3)

The apolipoprotein L3 (APOL3) gene belongs to the apolipoprotein L protein family, consisting of six members (APOL-I to-VI) (Page *et al.*, 2001; Peng *et al.*, 2020). The APOL3 gene plays an essential role in the transport of cholesterol and other cellular processes such as transcription modulation and signal transduction (Peng *et al.*, 2020). This lipid metabolism associated gene has been found to be highly duplicated in beef breeds (Bickhart *et al.*, 2012). Peng *et al.* (2020) analysed the CNV polymorphism of APOL3 in Xinan and Pinan cattle and found APOL3 CNV to be significantly

associated with traits such as hip height, cannon circumference, body slanting length, and hucklebone width. The results showed that multiple copies of this gene had positive effects on growth traits. Apolipoprotein L6 and apolipoprotein F have also been shown to influence lipid metabolism, thus influencing meat colour and quality (Taye *et al.*, 2017; Mei *et al.*, 2021).

5.4.2 Feed-efficiency related genes

Table 5.4: The main feed efficiency-related genes detected in the included publications.

Gene	Count	References
PRKG1	5	(Shin *et al.*, 2014), (Mielczarek *et al.*, 2017), (Mei *et al.*, 2020), (Mei *et al.*, 2021), (Yang *et al.*, 2021)
EIF2S1	2	(Hou *et al.*, 2012a), (Jang *et al.*, 2021)
FABP2	2	(Bickhart *et al.*, 2012), (Da Silva *et al.*, 2016b)

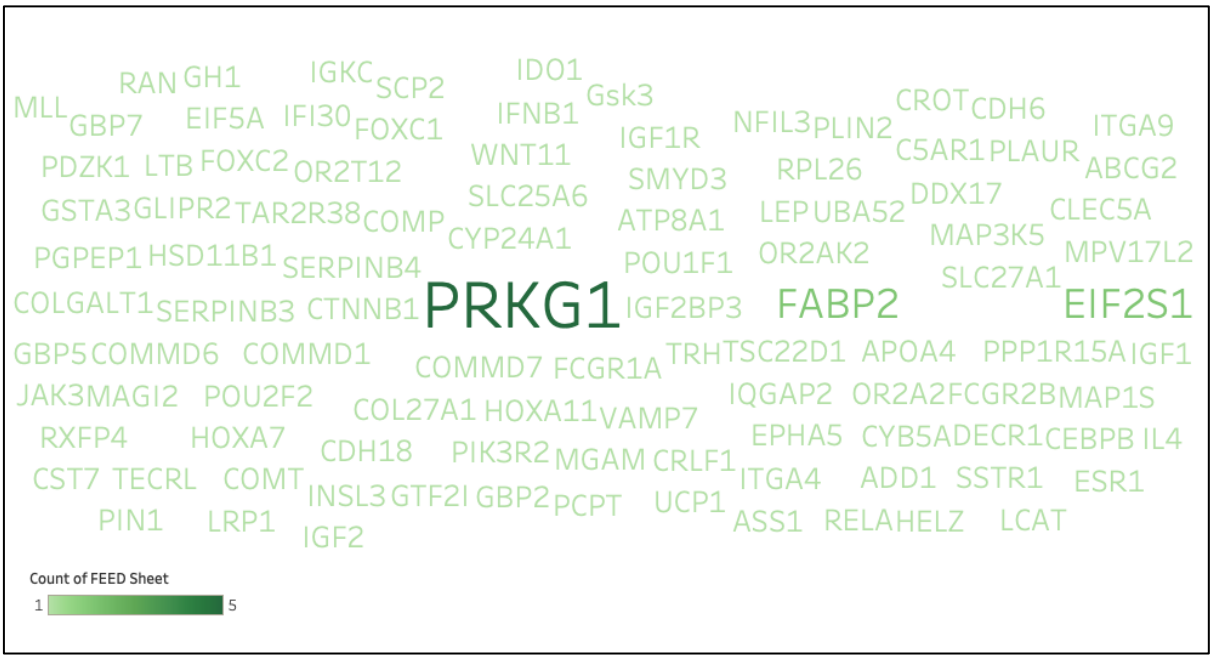

Figure 5.17: Visual representation of feed-efficiency gene prevalence in included publications.

Protein kinase cGMP-dependent type I (PRKG1)

The cGMP-dependent protein kinase (cGK) family includes two distinctive genes, PRKG1 and PRKG2 (Surks, 2007). Protein kinase cGMP-dependent type I (PRKG1) is involved in many pathways and plays different roles in these pathways. The PRKG1 proteins have been shown to be involved in the control of cardiovascular and neuronal functions, vascular smooth muscle contraction, deterrence of platelet aggregation, and cell growth modulation (Ørstavik *et al.*, 1997;

NCBI, 2022). Moreover, PRKG1 regulates lipolysis in fat cells to release glycerol and fatty acids through the hydrolysis of triacylglycerol (Shi *et al.*, 2019). Sherman *et al.* (2010) found a polymorphism in the PRKG1 gene to be significantly associated with dry matter intake in beef cattle. Yang *et al.* (2021) completed a comprehensive analysis of CNVs in Chinese Simmental beef cattle for association with economic traits of importance and identified deletion regions harbouring genes such as PRKG1. Taye *et al.* (2017) identified PRKG1 to be a candidate gene for residual feed intake, intramuscular fat, and tenderness in Ankole cattle. Moreover, polymorphisms in PRKG1 have been associated with residual feed intake in porcine (Onteru *et al.*, 2013). Interestingly, PRKG1 has also been related to milk fatty acids in Chinese Holstein dairy cattle (Shi *et al.*, 2019) and tick resistance in Nguni cattle (Mapholi *et al.*, 2016).

Eukaryotic translation initiation factor 2 subunit alpha (EIF2S1)

The eukaryotic translation initiation factor 2 (EIF2) protein complex, consisting of three subunits (alpha, beta, and gamma), have key functions in the initiation of protein synthesis (Green *et al.*, 1991). Eukaryotic translation initiation factor 2 subunit alpha (EIF2S1), found on BTA10, has been related to feed efficiency in cattle. Sainz *et al.* (2013) investigated the efficiency and estimated maintenance energy requirements of taurine and zebu cattle and concluded that the maintenance requirements of zebu cattle were lower compared with taurine cattle. Hou *et al.* (2012a) found EIF2S1 to overlap with a CNVR specific to a group of highly feed efficient Holstein cattle. Jang *et al.* (2021) found this gene to overlap with a taurine-specific duplication and suggested its contribution to different beef cattle feed efficiency between taurine and zebu cattle.

Fatty acid binding protein 2 (FABP2)

Fatty acid-binding proteins are intracellular polypeptides found in several tissues that play important roles in fatty acid transfer and metabolism (Gomez *et al.*, 2007). Fatty acid binding protein 2 (FABP2) codes for a small, intestinal fatty acid binding protein (IFABP) expressed in the proximal portion of the intestinal epithelial cells and typically facilitates the cellular uptake and transport of fatty acids across the cell membranes (Bickhart *et al.*, 2012). In humans, an Ala54Thr polymorphism in the FABP2 locus has been associated with lipid oxidation rates and insulin resistance (Formanack & Baier, 2004). In cattle, the FABP2 gene, found on BTA6, has been associated with lipid transport, metabolism, and feed efficiency. Bickhart *et al.* (2012) identified a CNV directly upstream from FABP2 in the beef breeds and hypothesised that it could be associated with feed efficiency and lipid transport. Thus, CNVs in or near the FABP2 locus could increase its expression in the intestinal epithelium, thereby increasing fatty acid sequestration from feed (Bickhart *et al.*, 2012).

5.5 CNV-related genes affecting coat characteristics

Coat patterns constitute one of the most important external phenotypic traits for breed identity (Fontanesi *et al.*, 2009). Coat colour in mammals is controlled by several genes – it is approximated that around 300 loci influence coat colour phenotypes, involving over 150 genes (Szczerbal *et al.*, 2017). In this review, the most frequent external phenotype-related genes detected in the included publications is presented in Table 5.5, and visually represented in Figure 5.18. These genes include KIT, followed by AP3B1, MC1R, PRLR, and FGF18.

Table 5.5: The main coat characteristic-related genes detected in the included publications.

Gene	Count	References
KIT	9	(Venhoranta *et al.*, 2013), (Khamzina, 2016), (Upadhyay *et al.*, 2017), (Yang *et al.*, 2017b), (Mielczarek *et al.*, 2018), (Kommadath *et al.*, 2019), (Mei *et al.*, 2020), (Guo *et al.*, 2021), (Mei *et al.*, 2021)
AP3B1	2	(Mei *et al.*, 2021), (Yang *et al.*, 2021)
FGF18	2	(Prinsen *et al.*, 2017), (Guo *et al.*, 2021)
MC1R	2	(Hou *et al.*, 2012b), (Mielczarek *et al.*, 2018)
PRLR	2	(Prinsen *et al.*, 2017), (Mei *et al.*, 2020)

Figure 5.18: Visual representation of coat characteristic related gene prevalence in included publications.

KIT proto-oncogene, receptor tyrosine kinase (KIT)

Colour-sidedness, a phenotype found in several cattle breeds, is determined by one of two different alleles, involving translocation (and subsequent duplication) of the bovine KIT locus (Durkin *et al.*, 2012; Brenig *et al.*, 2013). This proto-oncogene forms part of the tyrosine kinase receptor family. The KIT gene codes for a type III tyrosine kinase receptor protein, and is involved in melanogenesis, gametogenesis, erythropoiesis, haematopoiesis, and T-cell differentiation (Fontanesi *et al.*, 2009). Interestingly, KIT influences melanin deposition, and regulates the migration, proliferation, survival, and differentiation of cells in melanocytes (Hu *et al.*, 2020a). Furthermore, a KIT gene duplication on BTA6, and its translocation to BTA29 (and a subsequent translocation back to BTA6) causes white coat colour variation, known as colour-sidedness in cattle such as Belgian Blue and Brown Swiss cattle (Figure 5.19) (Durkin *et al.*, 2012).

Figure 5.19: Colour-sidedness in a) Belgian Blue and b) Brown Swiss cattle. From Durkin *et al.* (2012).

This colour-sidedness phenotype is generated in two steps (Figure 5.20). The first step involves the circular intermediate translocation of the BTA6 KIT locus to BTA29 in Belgian Blue cattle through a process of micro-homology mediated end-joining. The second step involves the circular intermediate translocation of a section of the novel BTA29-BTA6 fusion locus back to the original BTA6 wild-type locus via non-allelic homologous recombination (NAHR) (Durkin *et al.*, 2012). Interestingly, Brenig *et al.* (2013) discovered White Galloway and White Park cattle carried this allele, however, the effects of the translocated KIT locus resulted in mottled markings, instead of colour sidedness.

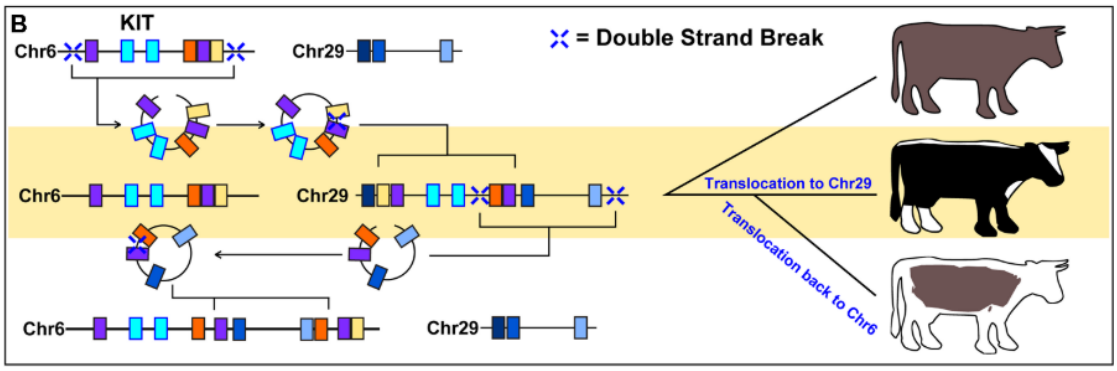

Figure 5.20: Translocation of the KIT locus. Colour sidedness in Belgian Blue cattle (middle) and Brown Swiss cattle (bottom). From Bickhart & Liu (2014).

Several previous studies have investigated mutations in the KIT gene and its association with coat colours in various cattle breeds such as Nguni, Pinzgauer, Gloucester, and Brown Swiss cattle (Szczerbal *et al.*, 2017; Küttel *et al.*, 2019; Artesi *et al.*, 2020; Häfliger *et al.*, 2020). Moreover, ectopic KIT CNV has also been associated with gonadal hypoplasia (small and underdeveloped gonads) in Swedish Mountain cattle and Northern Finncattle (Venhoranta *et al.*, 2013).

Adaptor related protein complex 3 subunit beta 1 (AP3B1)

Adaptor-related protein (AP) complexes are involved in vesicle formation and trafficking of proteins in the intracellular membrane (Simpson *et al.*, 1997). Five AP complexes have been identified, AP-1 to AP-5, each of which have different functions and distinct localisations (Nakatsu & Ohno, 2003; Adamopoulos, 2018). The AP3B1 gene encodes a protein involved in organelle biogenesis associated with platelet dense granules, melanosomes, and lysosomes (Yang *et al.*, 2021). Melanosomes are intracellular organelles generated by pigment cells in which melanins are synthesized and stored (Raposo & Marks, 2007). Melanin is the main pigment of the skin and hair in mammals. Mei *et al.* (2021) identified CNV-related genes in Chinese cattle breeds, such as AP3B1, and suggested its involvement in coat colour.

Melanocortin 1 receptor (MC1R)

The individual base coat colour of mammals is determined by the pigment-type switching system within melanocytes (Barsh *et al.*, 2000). Eumelanin (black to brown pigment) and pheomelanin (red to yellow pigment) are the two main constituents of mammalian pigments (Mohanty *et al.*, 2008; Walker & Gunn, 2010). The synthesis of each of these pigments is regulated by the level of tyrosinase (a rate-limiting enzyme) expression. High levels of tyrosinase results in the production of eumelanin, while low levels lead to the production of pheomelanin (Figure 5.21) (Gutiérrez-Gil *et al.*, 2007). Tyrosinase enzyme activity is regulated by the melanocortin 1 receptor (MC1R) (Gutiérrez-Gil *et al.*, 2007). The activation of MC1R, by α-melanocyte stimulating hormone, generates the production of brown/black eumelanin, while the inhibition of MC1R, occurring when the agouti-signalling protein (ASIP) is expressed, induces the production of red/yellow pheomelanin (Figure

5.21) (Hauser *et al.*, 2022). Interestingly, ASIP acts as an antagonist, by causing a blockage of α-melanocyte stimulating hormone action.

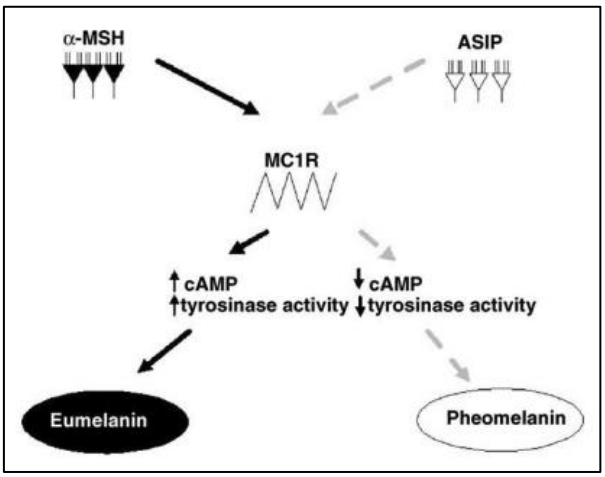

Figure 5.21: The alternate effects of binding α-melanocyte stimulating hormone (α-MSH) and agouti-signalling protein (ASIP) to MC1R. From Makova & Norton (2005).

In cattle, the MC1R gene, located on BTA18, is responsible for the switch of pheomelanin to eumelanin in the melanogenesis pathway and, thus, is responsible for differences in base coat colour (Gutiérrez-Gil *et al.*, 2007; Hulsman Hanna *et al.*, 2014). The MC1R gene in cattle has been the focus of many studies, whose permanent activation results in black (or dark) coat colour, and loss of function mutations result in red coat colour (Qanbari & Simianer, 2014). Hou *et al.* (2012c) investigated CNVs in Angus cattle and identified MC1R in a variable region, which could be associated with the black coat colour of Angus cattle. Mielczarek *et al.* (2018) identified a partial deletion of MC1R exon in Norwegian Red, Brown Swiss, and Simmental cattle, which are all red breeds.

Prolactin receptor (PRLR)

Prolactin is a multipurpose polypeptide hormone of the pituitary gland, and has important functions in reproduction, mammary gland development, synthesis of milk secretion, and maintenance of lactation (Dong *et al.*, 2013). However, a recent study has shown that PRL and its receptor (PRLR) is not only involved in lactation, but also in hair morphology and thermoregulation (Littlejohn *et al.*, 2014). Littlejohn *et al.* (2014) identified a dominant genetic syndrome in taurine cattle that presented unusual phenotypes such as lactation failure, long and ‘hairy’ coats, and heat stress (Figure 5.22 left). The authors found the ‘hairy’ locus to be located within the PRL gene (Figure 5.22 right).

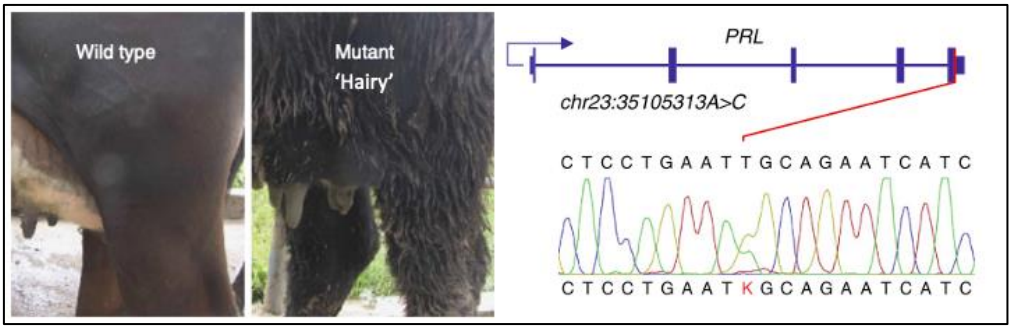

Figure 5.22: Left: 'hairy' syndrome in cattle, coat differences between wild-type and mutant half-sibs. Right: The PRL gene structure on BTA23, showing the location of mutation. Adapted from Littlejohn *et al.* (2014).

Moreover, Littlejohn *et al.* (2014) suggested mutations in the PRLR locus are responsible for the slick-coat phenotype. "Slick" cattle have a phenotype characterised by a very short, sleek hair coat and increased ability for thermoregulation (Figure 5.23) (Sosa *et al.*, 2022). The PRLR gene has been considered a positional candidate gene for the slick phenotype, thus carrying industrial importance.

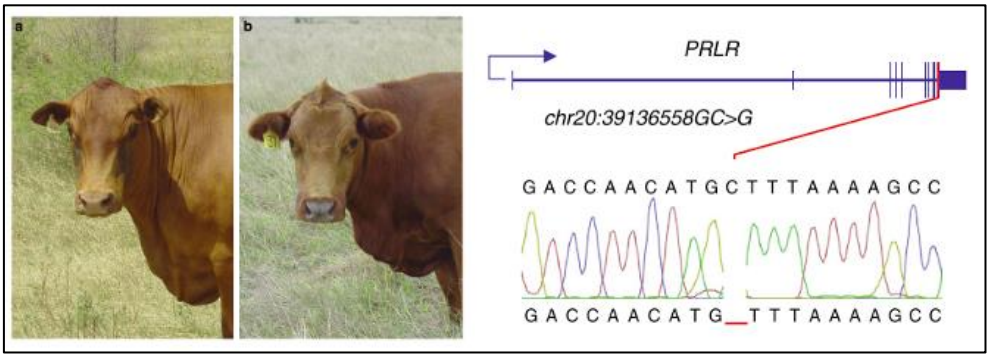

Figure 5.23: Left) Photographs of contrasting a. slick and b. non-slick Senepol crossbreeds. Right) The PRLR gene structure on BTA20, showing the location of mutation. Adapted from Littlejohn *et al.* (2014).

Fibroblast growth factor 18 (FGF18)

The family of fibroblast growth factors (FGFs) are multifunctional regulators involved in a variety of physiological events (Imamura, 2014). These signalling peptides are essential for regulating endochondral bone growth, tissue remodelling, organogenesis, and metabolism (Liu *et al.*, 2007; Imamura, 2014). There are 18 mammalian FGF subfamilies (FGF1-FGF10 and FGF16-FGF24) that exhibit various modes of action and mechanisms of secretion (Beenken & Mohammadi, 2009).

Various members of the FGF family are expressed in skin and are involved in hair follicle morphogenesis and regulation of the hair cycle (Kawano *et al.*, 2005; Ornitz & Itoh, 2015). Hair follicles cycle through growth, regression, and resting phases (Kimura-Ueki *et al.*, 2012). Previous studies have shown that FGF18 is highly expressed in hair follicles and is essential for the maintenance of the growth and resting phases (Kawano *et al.*, 2005; Kimura-Ueki *et al.*, 2012). Guo

et al. (2021) identified CNVR related genes in three Qaidam cattle, a local breed in northwest China known for having excellent adaptability traits, and identified KIT and FGF18 as candidate genes for hair colour and growth. The hair of Qaidam cattle is dense and long, which may have formed through long-term domestication of this breed to resist the cold.

Fibroblast growth factor 18 (FGF18) is involved in regulating chondrocyte proliferation and differentiation through FGF receptor 3 (FGFR3), and osteogenesis through other FGFRs (Liu *et al.*, 2007). Interestingly, Kim *et al.* (2017) found that only Ankole cattle had increased gene ontology categories involved in the FGF signalling pathway, and indicated that it may be related to the extreme horn development seen in this breed (Figure 5.24). This pathway included FGF18, which is important for differentiating osteoblasts during the development of carvarial bones (top part of skull) (Liu *et al.*, 2007), thus this gene might be connected to the distinctive morphology of the Ankole horn.

Figure 5.24: Ankole cattle, characterised by extreme horn development. From Farmer's weekly (2021).

5.6 Conclusion

Copy number variations can bring about major changes in gene expression, phenotypic traits, and evolutionary adaptation, through gene dosage and transcript structure alterations (Clop *et al.*, 2012; Bickhart & Liu, 2014). The distribution of CNVs amongst livestock could be due to demographic history, domestication and selection (Goshu *et al.*, 2018a). This review revealed that CNVs could be associated with adaptation and immunity (ABCC4, BOLA gene family, IGLL1, OR family, WC1, ZNF280B, BSP30A, DEFB, ULBP gene family, CATHL gene family, and HSP gene family), milk production and reproduction (DGAT1, IFNT, PAG, PRAME, PRL, AP3B1, IGLL1, SLC27A6, ITFG1, MTHFSD, PRP, and PTK2), meat production and growth (IGF2, PLA2G2D, CAST, IGF1R, APOL3, PTPRC, KCNJ12, CAPN1, AGBL3, CTNNA1, MSTN, ADRA1B, ATRN, LRRC49, MYH3, SORCS2, and TG), feed efficiency (PRKG1, FABP2, and EIF2S1) and coat characteristics (KIT, AP3B1, MC1R, PRLR, and FGF18) in cattle. This knowledge is relevant from a molecular perspective to the practical application in animal breeding and offers breeders the means to consider genomic selection of animals at a younger age.

CHAPTER 6 SUMMARY AND CONCLUSION

6.1 Synopsis of main results

This study aimed to summarise the relevant findings of cattle CNV research, and the important genes involved therewith, for the purpose of elucidating the involvement of CNVs in important cattle production traits. This was done was achieving two objectives. The first objective was to consolidate qualitative data of the published research to identify the state of knowledge, and the progress in research surrounding cattle-CNV studies. The second objective of this study was to perform a systematic review of published research to identify the main CNV-related genes being detected in research and their implication on economically important traits in cattle. Both of these objectives were met, and will be discussed below.

6.1.1 Objective 1: Consolidation of qualitative data of published research

The oldest relevant study included in this review was published in 2008, and since, there has been a steady increase in the number of cattle CNV-focused publications. Very few studies relevant to this review have been published in Africa. This may be because African indigenous cattle are comparatively not as intensively studied at the genomic level as other cattle populations (Mwai *et al.*, 2015) or because of the practical sampling challenges posed (Wang *et al.*, 2016a). Africa is richly endowed with 150 cattle breeds, comprising pure indicine lines, pure taurine lines, and a mixture of the two (Rewe *et al.*, 2009). These indigenous breeds display distinct physiological and anatomical characteristics that they have attained through the adaptation to abrasive conditions. Africa is richly endowed with cattle displaying enhanced attributes, thus more studies into this valuable genetic resource will be beneficial. Moreover, further CNV research into local South African breeds should be undertaken. Over 60% of the publications studied *Bos taurus* cattle, while only 12% studied *Bos indicus* cattle, and 4% focused on *Bos taurus africanus* cattle. Sanga breeds, such as Nguni cattle, are known for their resilience to parasite infestations, poor quality grazing, and harsh environments. Although indigenous South African cattle breeds possess an abundance of unique genetic material, research surrounding South African indigenous cattle is scarce. Therefore, the study of CNVs in South African cattle will not only be important for genomic selection of animals, but also for attaining insight into the adaptive mechanisms and unique traits of these local breeds. The methods used to detect CNVs include array-based approaches and sequence-based approaches (Liu & Bickhart, 2012). Approximately 37% of the publications utilised the Bovine SNPHD genotyping platform, while 34% of the publications used sequence-based approaches. The SNPHD is a widely used platform with distinct advantages in terms of cost and throughput (Alkan *et al.*, 2011), however, array-based methods are unable to detect CNV events that are balanced, nor can they detect small CNVs or exact breakpoints, because of their limited probe density. Sequencing is likely to overtake array-

based approaches (Alkan *et al.*, 2011), due to the development of cost-effective sequencing approaches with enhanced resolution and accuracy (Bickhart *et al.*, 2020). Utilizing sequence-based approaches for the identification of CNVs will have a positive impact on the livestock industry. Sequencing platforms allow researchers the opportunity to access comprehensive data on genetic markers responsible for important production traits (Gatew & Tarekegn, 2018), and thus important adaptability traits exhibited in local South African cattle. Moreover, WGS data has the potential to increase the accuracy of genomic prediction for traits that are low to moderately heritable. Next generation sequencing is also predicted to lower the overall cost of animal production, while improving the yield and quality of meat and milk products, the reproductive health, and the disease resistance of livestock (Sharma *et al.*, 2017; Gatew & Tarekegn, 2018). In South Africa, there is a lag in progress on the genomics front, mostly due to financial constraints (Lashmar *et al.*, 2019). However, with the decrease in the cost of certain sequencing platforms in recent years, there is potential for the use of sequencing based methods for CNV detection in South Africa in the future.

6.1.2 Objective 2: The predominant CNV-related genes and their implication on South African cattle

In cattle, ABCC4 has been reported to be important in defence/innate immunity, drug detoxication and adaptive immunity (Liu *et al.*, 2010, 2011; Upadhyay *et al.*, 2017) and is reported to be associated with the resistance or susceptibility to gastrointestinal nematodes (Li & Gasbarre, 2009; Liu *et al.*, 2011). Similarly, IGLL1 has been associated with resistance to intestinal nematodes (Hou *et al.*, 2012c). Several South African cattle farms comprise grazing animals, thus BSP30A, a BPI-like protein gene believed to contribute to the innate immunity of the oral cavity and airways (Wheeler *et al.*, 2003; Bickhart *et al.*, 2012) could be useful to hinder the progress of pathogens that cattle are likely to encounter in the soil while grazing. Both ULBP and CATHLs genes play an important role in antiviral immunity, host defence, and disease resistance in mammals (Flores, 2011).

Infestations of gastrointestinal nematodes have subclinical and clinical effects on animals and influence the economic and production gains of farmers (Gadberry & Powell, 2008; Mpetile *et al.*, 2017). The effects of parasites of the stomach and intestine are usually subclinical, such as reduced feed intake, poor feed conversion, decreased milk production, and decreased growth performance (Nyamushamba *et al.*, 2017). Moreover, internal parasitism is known to decrease production by 50% (Mwanza *et al.*, 2016). Gastrointestinal nematode infections pose a huge threat to the health and welfare of grazing animals (Sutherland & Leathwick, 2011). Therefore, in both sectors of the SA livestock industry, CNV-related genes such as ABCC4, IGLL1, BSP30A, ULBP17 and CATHL4 may have potential utility in DNA-based marker-assisted selection of pathogen- and parasite-resistance characteristics.

The South African dairy industry

Heat shock proteins are important for the alleviation of heat stress in mammals and are understood to play a role in the ability of cattle to tolerate heat, and thus in the climatic adaptabilities of different breeds (Wang, 2016). The HSP70 is reported to be the most abundant protein conferring thermotolerance (Basiricò *et al.*, 2011; Banerjee *et al.*, 2014). Interestingly, a recent study has shown that PRL and its receptor (PRLR) are not only involved in lactation, but also in hair morphology and thermoregulation (Littlejohn *et al.*, 2014). Mutations in the PRLR locus are responsible for the slick-coat phenotype. “Slick” cattle have a very short, sleek hair coat and an increased ability for thermoregulation. A study completed by Dikmen *et al.* (2014) confirmed that when exposed to heat stress, Holsteins with slick hair have a better thermoregulatory ability than those without. Moreover, the slick-haired cattle did not have such a large slump in milk yield in the hot summer months. The BOLA genes are expressed by the cells of the immune system which process antigenic peptides to present to helper T cells for an immune response against pathogens (Oprzadek *et al.*, 2018), and various members of the BOLA family have been reported to be associated with the resistance or susceptibility to mastitis (Yoshida *et al.*, 2009) and tick infestation (Kim *et al.*, 2017) in cattle. Bovine β-defensins that are located on BTA27 consist of the most immunologically important genes for intramammary infections (Gurao *et al.*, 2017), and have been associated with reduced somatic cell count (SCC) in Jersey cattle (Wojdak-Maksymiec *et al.,* 2006), and resistance to mastitis in Thanparkar cattle (Saravanan *et al.*, 2021).

Therefore, in the South African dairy industry, members of the BOLA gene family and β-defensin gene family could be used as potentially useful genetic markers for health and immunity traits in dairy cattle. Moreover, with the predicted temperature increase, HSP70 and PRLR could be a useful potential biomarker for thermotolerant traits in both dairy and beef cattle.

Milk production is an indispensable economic trait in cattle production and is affected by several genes (Yudin & Voevoda, 2015; Silpa *et al.*, 2021). Thus, the improvement of milk yield and composition could have positive effect on the SA dairy farms. The DGAT1 gene has been associated with milk fat content differences in various breeds (Winter *et al.*, 2002; Fontanesi *et al.*, 2014). Whereas, MTHFSD belongs to the folate metabolism genes, which are involved in the regulation of milk protein synthesis (Menzies *et al.*, 2009), and was found to be associated with milk protein and fat yield (Ben Sassi *et al.*, 2016). Prolactin is responsible for the synthesis of several milk components (such as lactose, lipids, and protein content). Additionally, ITFG1 and PTK2 have been related to milk production traits (Wang. *et al.*, 2013; Mei *et al.*, 2019, 2021). Thus, genes such as DGAT1, MTHFSD, PRL, ITFG1, and PTK2 have potential use in the MAS of milk production characteristics in South African dairy breeds.

The South African beef industry

The IGF system is important for both pre- and postnatal growth and development (Baker *et al.*, 1993; Allan *et al.*, 2001). Variance in IGF1R contributes to differences in body size among populations (Jang *et al.*, 2021), and was significantly associated with body weight, body height, and hucklebone width (Ma *et al.*, 2019). Whereas, IGF2 has been associated with rib eye area, meat tenderness and growth traits in cattle (Goodall & Schmutz, 2007; Huang *et al.*, 2014; Da Silva *et al.*, 2016a; Berton *et al.*, 2021). The KCNJ12 gene has been found to participate in the muscle contraction process and reported to have a significant influence on the growth traits of cattle (Zhou *et al.*, 2016a; Zheng *et al.*, 2019). Similarly, MYH3 is involved in muscle development, differentiation, and contractions of striated muscles (Zhang *et al.*, 2011), and affects growth and carcass traits (Wang *et al.*, 2013). Similarly, APOL3 has also been shown to affect growth traits (Peng *et al.*, 2020). The gene responsible for double muscling, MSTN, is a regulator of muscle cell proliferation (Aiello *et al.*, 2018), and mutations within this gene can lead to high carcass yields and excellent conformation (Haruna *et al.*, 2020). Interestingly, CTNNA1 has been reported to be associated with myostatin expression level and function (Sadkowski *et al.*, 2008), and may be related to the difference in muscle development and meat productivity between taurine and zebu cattle (Hu *et al.*, 2020b; Jang *et al.*, 2021). Growth traits have a huge impact on the value of the live animal, for breeding purposes and retail meat value purposes, therefore genes such as IGF1R, IGF2, KCNJ12, MYH3, APOL3, MSTN, and CTNNA1 could be used as potential genetic markers for body measurements and growth performance traits in South African cattle.

The improvement of meat quality components, such as meat tenderness, is a very important undertaking, as it is considered to be the most important factor for the consumer (Strydom *et al.*, 2000). The PLA2G2D gene has been associated with meat quality traits such as fat deposition, lipid metabolism, and meat colour factors (Stothard *et al.*, 2011; Berton *et al.*, 2021). While the SORCS2 gene has been related to lipid metabolism and suggested to be linked to backfat thickness in Nellore cattle (Júnior *et al.*, 2016). The LRRC49 gene has been associated with marbling score and subcutaneous fat in Canchim beef cattle (Mokry *et al.*, 2013). The two enzymes responsible for myofibrillar protein degradation, namely CAPN1 and CAST (Casas *et al.*, 2006a), have been associated with meat tenderness (Lee *et al.*, 2019; Moravčíková *et al.*, 2019; Guo *et al.*, 2021). Similarly, PTPRC has been associated with tenderness and marbling scores (Gao *et al.*, 2014; Braz *et al.*, 2019). In beef cattle, PRKG1 has been associated with dry matter intake (Sherman *et al.*, 2010) and identified to be a candidate gene for residual feed intake, intramuscular fat, and tenderness in Ankole cattle (Taye *et al.*, 2017). Meat quality traits hold much importance for consumer satisfaction, and consequently overall profitability. Similarly, feed is the major contributor to variable costs in livestock production systems, thus feed efficiency is very important for farm profitability. Accordingly, genes such as PLA2G2D, SORCS2, LRRC49, CAPN1, CAST, and PTPRC could be valuable candidate genes for meat quality traits in South African cattle.

6.2 Strengths, limitations, and future work

The strength of this study lies in the broad inclusion criteria chosen for this systematic review, which included peer-reviewed articles and grey literature, cattle of all breeds, sex or production types, and publications over a wide geographical area. This ensured all possible studies were considered. However, this broad search inclusion led to a large number of studies with different objectives and thus a limited degree of homogeneity. This systematic literature review excluded other molecular markers, such as SNPs. Since SNPs were initially thought to be the main source of genetic variation and be responsible for most phenotypic variation (Freeman *et al.*, 2006), there were several SNP-based GWA studies that had to be excluded. Therefore, the inclusion of only CNV-related literature could have been a limitation. Lastly, the screening for studies to be included in this review was only conducted by one person. It has been reported that single-reviewer screening missed 13.4% of relevant publications, whereas dual-reviewer screening only missed 2.5% of relevant publications (Gartlehner *et al.*, 2020). Although a rigorous protocol was employed, there is still a possibility that relevant studies were missed, thus single-reviewer abstract screening is recognised as a limitation (Gartlehner *et al.*, 2020).

Cattle CNV research has advanced considerably since 2008, but there are a few research gaps and needs that should be considered in future work:

<u>The need for more research in developing countries</u>

Although there is a wide geographical coverage of publications, the number of publications is not representative of the number of cattle present in Africa. This is likely due to the fact that African cattle populations are less intensively studied at a genomic level (Mwai *et al.*, 2015; Talenti *et al.*, 2022). The genetic diversity and distinctive features of indigenous African cattle represent a unique resource and opportunity to tackle livestock productivity challenges faced in developing countries. It is therefore clear that there is a great need for more cattle CNV research studies in Africa. Additionally, the number of publications is not representative of the number of cattle present in other developing countries such as India and Argentina, despite the fact that India is the top global milk-producing country and Argentina is one of the top meat-producing countries (FAO, 2022). Given this, it is imperative that more genetic research should be conducted on cattle in these countries. Overall, it is clear that more cattle CNV research in developing countries is needed.

<u>The need for more *Bos indicus* and *Bos taurus africanus* studies</u>

Publications included in this review focused mainly on *Bos taurus* cattle. *Bos indicus* and *Bos taurus africanus* cattle have physiological advantages over *Bos taurus* cattle, such as lower susceptibility to ticks and gastrointestinal nematodes and greater resistance to heat (Canavez *et al.*, 2012). Furthermore, it has been reported that more CNV sites occur in *Bos indicus* cattle compared to

European *Bos taurus* cattle (Liu *et al.*, 2010), thus there is a need for more research in zebu and Sanga cattle. The low number of publications on *Bos indicus* cattle is linked to the low number of publications in developing countries, thus increasing publications of both goes hand in hand.

The need for standardised reporting

The number of CNVs identified in each publication was greatly influenced by inconsistences in the studies such as the stringency of the CNV calling criteria, or the number of algorithms used to detect CNVs, as well as the reporting of findings. In future studies, CNV detection and reporting should be standardised for reliable comparisons between publications. Moreover, publications should make supplementary data more readily available, and provide more clarity on the reference genome used and the CNV/CNVR value reported.

6.3 Conclusion and recommendations

This systematic review reveals that cattle CNV research has increased considerably since 2008. The extent and distribution of the publications reflect the worldwide growing importance of understanding the cattle genome for genetic improvement of livestock. However, there is a lack of research in developing countries, a lack of emphasis on *Bos indicus* and *Bos taurus africanus* cattle, and a lack of standardised reporting across cattle CNV studies. It is clear that copy number variations can alter the gene expression and consequently influence phenotypic expression. This systematic review identified several important CNV-related genes that influence economically important traits in cattle such as adaptation and immunity, milk yield and composition, fertility and reproduction, meat quality and growth, feed efficiency, and coat colour, coat patterns and hair morphology. This knowledge is relevant from a molecular perspective to the practical application in cattle breeding and selection programs. Incorporating this information into breeding programmes could improve the production of both the developing and developed sectors of the South African beef and dairy industry. With the growing human population triggering an ever-increasing demand for milk and meat products, and with climate change posing a threat to the productive efficacy of cattle across the globe, the use of molecular assisted selective breeding is essential for overcoming current and future challenges in cattle productivity.

6.4 References

Abin, S., Theron, H. E., & Van Marle-Köster, E. 2016. Population structure and genetic trends for indigenous African beef cattle breeds in South Africa. South African J. Anim. Sci. 46, 152–156 https://doi.org/10.4314/sajas.v46i2.5.

Adamopoulos, P. 2018. Molecular cloning of novel transcripts of the adaptor-related protein complex 2 alpha 1 subunit (AP2A1) gene, using Next-Generation Sequencing. Gene 678.

Adams, T. E., Epa, V. C., Garrett, T. P. J., & Ward, C. W. 2000. Structure and function of the type 1 insulin-like growth factor receptor. Cell. Mol. Life Sci. 57, 1050–1093 https://doi.org/10.1007/PL00000744.

Adewale, B. A. 2020. Will long-read sequencing technologies replace short-read sequencing technologies in the next 10 years? Afr. J. Lab. Med. 9, 1–5 https://doi.org/10.1590/2318-0889202032E200032.

Aguiar, T. S., Torrecilha, R. B. P., Milanesi, M., Utsunomiya, A. T. H., Trigo, B. B., Tijjani, A., Musa, H. H., Lopes, F. L., Ajmone-Marsan, P., Carvalheiro, R., Neves, H. H. R., Do Carmo, A. S., Hanotte, O., Sonstegard, T. S., Garcia, J. F., & Utsunomiya, Y. T. 2018. Association of copy number variation at intron 3 of hmga2 with navel length in bos indicus. Front. Genet. 9 https://doi.org/10.3389/fgene.2018.00627.

Ahn, S., & Park, H. 2010. XIAP is essential for shear stress-enhanced Tyr-576 phosphorylation of FAK. Biochem. Biophys. Res. Commun. 399, 256–261 https://doi.org/10.1016/j.bbrc.2010.07.064.

Aiello, D., Patel, K., & Lasagna, E. 2018. The myostatin gene: an overview of mechanisms of action and its relevance to livestock animals. Anim. Genet. 49, 505–519 https://doi.org/10.1111/age.12696.

Albarrak, S. M., Waters, W. R., Stabel, J. R., & Hostetter, J. M. 2017. WC1+ γδ T cells from cattle naturally infected with Mycobacterium avium subsp. paratuberculosis respond differentially to stimulation with PPD-J. Vet. Immunol. Immunopathol. 190, 57–64 https://doi.org/10.1016/j.vetimm.2017.07.007.

Albrecht, E., Komolka, K., Kuzinski, J., & Maak, S. 2012. Agouti revisited: Transcript quantification of the ASIP gene in bovine tissues related to protein expression and localization. PLoS One 7 https://doi.org/10.1371/journal.pone.0035282.

Alfonso, E., Rojas, R., Herrera, J. G., Ortega, M. E., Lemus, C., Ruiz, J., Pinto, R., & Gómez, H. 2012. Polymorphism of the prolactin gene (PRL) and its relationship with milk production in American Swiss cattle. African J. Biotechnol. 11, 7338–7343 https://doi.org/10.5897/AJB11.1485.

Alipanah, M., Kalashnikova, L., & Rodionov, G. 2007. Association of genetic variants of the prolactin gene with milk production traits in Russian Red Pied cattle. Proc. Br. Soc. Anim. Sci. 2007, 156–156 https://doi.org/10.1017/s1752756200020597.

Alkan, C., Coe, B. P., & Eichler, E. E. 2011. Genome structural variation discovery and genotyping. Nat. Rev. Genet. 12, 363–376 https://doi.org/10.1038/nrg2958.

Allan, G. J., Flint, D. J., & Patel, K. 2001. Insulin-like growth factor axis during embryonic development. Reproduction 122, 31–39 https://doi.org/10.1530/rep.0.1220031.

De Almeida Santana, M. H., Junior, G. A. O., Cesar, A. S. M., Freua, M. C., da Costa Gomes, R., da Luz e Silva, S., Leme, P. R., Fukumasu, H., Carvalho, M. E., Ventura, R. V., Coutinho, L. L., Kadarmideen, H. N., & Ferraz, J. B. S. 2016. Copy number variations and genome-wide associations reveal putative genes and metabolic pathways involved with the feed conversion ratio in beef cattle. J. Appl. Genet. 57, 495–504 https://doi.org/10.1007/s13353-016-0344-7.

Alonso, A., Sasin, J., Bottini, N., Friedberg, I., Friedberg, I., Osterman, A., Godzik, A., Hunter, T., Dixon, J., & Mustelin, T. 2004. Protein tyrosine phosphatases in the human genome. Cell 117, 699–711 https://doi.org/10.1016/j.cell.2004.05.018.

Antunes de Lemos, M. V., Berton, M. P., Ferreira de Camargo, G. M., Peripolli, E., de Oliveira Silva, R. M., Ferreira Olivieri, B., Cesar, A. S. M. M., Pereira, A. S. C., de Albuquerque, L. G., de Oliveira, H. N., Tonhati, H., Baldi, F., de Lemos, M. V. A., Berton, M. P., de Camargo, G. M. F., Peripolli, E., Silva, R. M. D., Olivieri, B. F., Cesar, A. S. M. M., Pereira, A. S. C., de Albuquerque, L. G., de Oliveira, H. N., Tonhati, H., & Baldi, F. 2018a. Copy number variation regions in Nellore cattle: Evidences of environment adaptation. Livest. Sci. 207, 51–58 https://doi.org/10.1016/j.livsci.2017.11.008 WE - Science Citation Index Expanded (SCI-EXPANDED).

Antunes de Lemos, M. V., Peripolli, E., Berton, M. P., Loise, F., Feitosa, B., Olivieri, B. F., Stafuzza, N. B., Tonussi, R. L., Kluska, S., Lucas, H., Chiaia, J., Mueller, L., Ferrinho, A. M., Simone, A., Prereira, C., Oliveira, H. N. De, Albuquerque, L. G. De, Baldi, F., Feitosa, F. L. B., Olivieri, B. F., Stafuzza, N. B., Tonussi, R. L., Kluska, S., Chiaia, H. L. J., Mueller, L., Ferrinho, A. M., Prereira, A. S. C., de Oliveira, H. N., de Albuquerque, L. G., & Baldi, F. 2018b. Association study between copy number variation and beef fatty acid profile of Nellore cattle. J. Appl. Genet. 59, 203–223 https://doi.org/10.1007/s13353-018-0436-7.

de Araújo Lima, L., & Wang, K. 2017. PennCNV in whole-genome sequencing data. BMC Bioinformatics 18 https://doi.org/10.1186/s12859-017-1802-x.

Araujo, R. N., Padilha, T., Zarlenga, D., Sonstegard, T., Connor, E. E., Van Tassel, C., Lima, W. S., Nascimento, E., & Gasbarre, L. C. 2009. Use of a candidate gene array to delineate gene expression

patterns in cattle selected for resistance or susceptibility to intestinal nematodes. Vet. Parasitol. 162, 106–115 https://doi.org/10.1016/j.vetpar.2008.12.017.

Archana. 2017. Role of Heat Shock Proteins in Livestock Adaptation to Heat Stress. J. Dairy, Vet. Anim. Res. 5 https://doi.org/10.15406/jdvar.2017.05.00127.

Ardicli, S., Soyudal, B., Samli, H., Dincel, D., & Balci, F. 2018. Effect of STAT1, OLR1, CSN1S1, CSN1S2, and DGAT1 genes on milk yield and composition traits of Holstein breed. Rev. Bras. Zootec. 47, 1–9 https://doi.org/10.1590/rbz4720170247.

Armstrong, E., Almeida, F. A. R., McIntosh, M. M., Poli, M., Cibils, A. F., Martínez-Quintana, J. A., Félix-Portillo, M., & Estell, R. E. 2022. Genetic and productive background of Criollo cattle in Argentina, Mexico, Uruguay and the United States. J. Arid Environ. 200, 104722.

Arnould, V. M. R., & Soyeurt, H. 2009. Genetic variability of milk fatty acids. J. Appl. Genet. 50, 29–39 https://doi.org/10.1007/BF03195649.

Artesi, M., Tamma, N., Deckers, M., Karim, L., Coppieters, W., Van den Broeke, A., Georges, M., Charlier, C., & Durkin, K. 2020. Colour-sidedness in Gloucester cattle is associated with a complex structural variant impacting regulatory elements downstream of KIT. Anim. Genet. 51, 461–465 https://doi.org/10.1111/age.12932.

Aslam, S., & Emmanuel, P. 2010. Formulating a researchable question: A critical step for facilitating good clinical research. Indian J. Sex. Transm. Dis. 31, 47–50 https://doi.org/10.4103/0253-7184.69003.

Bae, J. S., Cheong, H. S., Kim, L. H., NamGung, S., Park, T. J., Chun, J.-Y., Kim, J. Y., Pasaje, C. F. A., Lee, J. S., & Shin, H. D. 2010. Identification of copy number variations and common deletion polymorphisms in cattle. BMC Genomics 11 https://doi.org/10.1186/1471-2164-11-232.

Bagnato, A., Strillacci, M. G., Pellegrino, L., Schiavini, F., Frigo, E., Rossoni, A., Fontanesi, L., Maltecca, C., Prinsen, R. R. T. M. M., & Dolezal, M. A. 2015. Identification and validation of copy number variants in Italian Brown Swiss dairy cattle using Illumina Bovine SNP50 Beadchip (R). Ital. J. Anim. Sci. 14, 552-+ https://doi.org/10.4081/ijas.2015.3900 WE - Science Citation Index Expanded (SCI-EXPANDED).

Baker, J., Liu, J. P., Robertson, E. J., & Efstratiadis, A. 1993. Role of insulin-like growth factors in embryonic and postnatal growth. Cell 75, 73–82.

Banerjee, D., Upadhyay, R. C., Chaudhary, U. B., Kumar, R., Singh, S., Ashutosh, Mohanarao, J. G., Polley, S., Mukherjee, A., Das, T. K., & De, S. 2014. Seasonal variation in expression pattern of genes under HSP70: Seasonal variation in expression pattern of genes under HSP70 family in heat- and cold-adapted goats (Capra hircus). Cell Stress Chaperones 19, 401–408 https://doi.org/10.1007/s12192-013-0469-0.

Al Barashdi, M. A., Ali, A., McMullin, M. F., & Mills, K. 2021. Protein tyrosine phosphatase receptor type C (PTPRC or CD45). J. Clin. Pathol. 74, 548–552 https://doi.org/10.1136/jclinpath-2020-206927.

Barker, J. S. F. 2009. Defining fitness in natural and domesticated populations.Pages 3-14 LB-B1 Chapter in a Scholarly Book in Adaptation and Fitness in Animal Populations. 1st ed. Springer Science + Business Media BV.

Barsh, G., Gunn, T., He, L., Schlossman, S., & Duke-Cohan, J. 2000. Biochemical and genetic studies of pigment-type switching. Pigment Cell Res. 13, 48–53 https://doi.org/10.1034/j.1600-0749.13.s8.10.x.

Basiricò, L., Morera, P., Primi, V., Lacetera, N., Nardone, A., & Bernabucci, U. 2011. Cellular thermotolerance is associated with heat shock protein 70.1 genetic polymorphisms in Holstein lactating cows. Cell Stress Chaperones 16, 441–448 https://doi.org/10.1007/s12192-011-0257-7.

Bayıl Oğuzkan, S., & Bozkurt, A. S. 2019. A Study on the Effect of Prolactin Gene Variants on Milk Production Traits of Holstein Cattle. Russ. J. Genet. 55, 480–486 https://doi.org/10.1134/S1022795419040082.

Beef2Live. 2022. Cattle Breeds: List Of Bos Taurus Cattle.

Beenken, A., & Mohammadi, M. 2009. The FGF family: biology, pathophysiology and therapy. Nat. Rev. Drug Discov. 8, 235–253 https://doi.org/10.1038/nrd2792.

Behjati, S., & Tarpey, P. S. 2013. What is next generation sequencing? Arch. Dis. Child. Educ. Pract. Ed. 98, 236–238 https://doi.org/10.1136/archdischild-2013-304340.

Behl, J. D., Verma, N. K., Tyagi, N., Mishra, P., Behl, R., & Joshi, B. K. 2012. The Major Histocompatibility Complex in Bovines: A Review. ISRN Vet. Sci. 2012, 1–12 https://doi.org/10.5402/2012/872710.

Bellinge, R. H. S., Liberles, D. A., Iaschi, S. P. A., O'Brien, P. A., & Tay, G. K. 2005. Myostatin and its implications on animal breeding: A review. Anim. Genet. 36, 1–6 https://doi.org/10.1111/j.1365-2052.2004.01229.x.

Ben-Jemaa, S., Mastrangelo, S., Lee, S.-H., Lee, J. H., & Boussaha, M. 2020. Genome-wide scan for selection signatures reveals novel insights into the adaptive capacity in local North African cattle. Sci. Rep. 10, 1–14.

Bennett, G. L., Shackelford, S. D., Wheeler, T. L., King, D. A., Casas, E., & Smith, T. P. L. 2013. Selection for genetic markers in beef cattle reveals complex associations of thyroglobulin and casein1-S1 with carcass and meat traits. J. Anim. Sci. 91, 565–571 https://doi.org/10.2527/jas.2012-5454.

Berman, A. 2004. Tissue and external insulation estimates and their effects on prediction of energy requirements and of heat stress. J. Dairy Sci. 87, 1400–1412 https://doi.org/10.3168/jds.S0022-0302(04)73289-0.

Bermingham, M. L., Bishop, S. C., Woolliams, J. A., Pong-Wong, R., Allen, A. R., McBride, S. H., Ryder, J. J., Wright, D. M., Skuce, R. A., McDowell, S. W., & Glass, E. J. 2014. Genome-wide association study identifies novel loci associated with resistance to bovine tuberculosis. Heredity (Edinb). 112, 543–551 https://doi.org/10.1038/hdy.2013.137.

Berry, D. P., & Crowley, J. J. 2013. Cell biology symposium: Genetics of feed efficiency in dairy and beef cattle. J. Anim. Sci. 91, 1594–1613 https://doi.org/10.2527/jas.2012-5862.

Berton, M. P., De Antunes Lemos, M. V., Seleguim Chud, T. C., Bonvino Stafuzza, N., Kluska, S., Amorim, S. T., Silva Ferlin Lopes, L., Cravo Pereira, A. S., Bickhart, D., Liu, G., Galvão De Albuquerque, L., & Baldi, F. 2021. Genome-wide association study between copy number variation regions and carcass- And meat-quality traits in Nellore cattle. Anim. Prod. Sci. 61, 731–744 https://doi.org/10.1071/AN20275.

Berton, M. P., Marcos Vinícius Antunes de Lemos[1], Aboujaoude[1], C., Camargo[1], G. M. F. de, Chardulo[1], L. A. L., Albuq, L. G. de, & (Jess, swop them around N. 2014. Association Between Copy Number Variation Regions and Meat Tenderness in Nelore Cattle.in 10th World Congress of Genetics Applied to Livestock Production Association.

Bhanuprakash, V., Chhotaray, S., Pruthviraj, D. R., Rawat, C., Karthikeyan, A., & Panigrahi, M. 2018. Copy number variation in livestock: a mini review. Vet. World 11, 535.

Bhat, Z. F., Morton, J. D., Mason, S. L., & Bekhit, A. E. D. A. 2018. Role of calpain system in meat tenderness: A review. Food Sci. Hum. Wellness 7, 196–204 https://doi.org/10.1016/j.fshw.2018.08.002.

Bi, Y., Feng, B., Wang, Z., Zhu, H., Qu, L., Lan, X., Pan, C., & Song, X. 2020. Myostatin (MSTN) gene indel variation and its associations with body traits in shaanbei white cashmere goat. Animals 10 https://doi.org/10.3390/ani10010168.

Bickhart, D. M., Hou, Y., Schroeder, S. G., Alkan, C., Cardone, M. F., Matukumalli, L. K., Song, J., Schnabel, R. D., Ventura, M., Taylor, J. F., Garcia, J. F., Van Tassell, C. P., Sonstegard, T. S., Eichler, E. E., & Liu, G. E. 2012. Copy number variation of individual cattle genomes using next-generation sequencing. Genome Res. 22, 778–790 https://doi.org/10.1101/gr.133967.111.

Bickhart, D. M., & Liu, G. E. 2014. The challenges and importance of structural variation detection in livestock. Front. Genet. 5 https://doi.org/10.3389/fgene.2014.00037.

Bickhart, D. M., McClure, J. C., Schnabel, R. D., Rosen, B. D., Medrano, J. F., & Smith, T. P. L. 2020. Symposium review: Advances in sequencing technology herald a new frontier in cattle genomics and genome-enabled selection. J. Dairy Sci. 103, 5278–5290.

Bickhart, D. M., Xu, L., Hutchison, J. L., Cole, J. B., Null, D. J., Schroeder, S. G., Song, J., Garcia, J. F., Sonstegard, T. S., Van Tassell, C. P., Schnabel, R. D., Taylor, J. F., Lewin, H. A., & Liu, G. E. 2016. Diversity and population-genetic properties of copy number variations and multicopy genes in cattle. DNA Res. 23, 253–262 https://doi.org/10.1093/dnares/dsw013.

Bingle, C. D., Seal, R. L., & Craven, C. J. 2011. Systematic nomenclature for the PLUNC/PSP/BSP30/SMGB proteins as a subfamily of the BPI fold-containing superfamily. Biochem. Soc. Trans. 39, 977–983 https://doi.org/10.1042/BST0390977.

Bionaz, M., & Loor, J. J. 2008. ACSL1, AGPAT6, FABP3, LPIN1, and SLC27A6 are the most abundant isoforms in bovine mammary tissue and their expression is affected by stage of lactation. J. Nutr. 138, 1019–1024 https://doi.org/10.1093/jn/138.6.1019.

Bishop, M. D., Koohmaraie, M., Killefer, J., & Kappes, S. 1993. Rapid communication: restriction fragment length polymorphisms in the bovine calpastatin gene. J. Anim. Sci. 71, 2277 https://doi.org/10.2527/1993.7182277x.

Blanco, F. C., Sabio y Garcia, J., & Bigi, F. 2021. Recent advances in non-specific immune memory against bovine tuberculosis. Comp. Immunol. Microbiol. Infect. Dis. 75, 101615 https://doi.org/10.1016/j.cimid.2021.101615.

Blumerman, S. L., Herzig, C. T. A., Rogers, A. N., Telfer, J. C., & Baldwin, C. L. 2006. Differential TCR gene usage between WC1- and WC1+ ruminant γδ T cell subpopulations including those responding to bacterial antigen. Immunogenetics 58, 680–692 https://doi.org/10.1007/s00251-006-0122-5.

Bögeholz, A., Falker-Gieske, C., Guélat, M., Gurtner, C., Hunziker, S., Oevermann, A., Thaller, G., Drögemüller, C., & Tetens, J. 2021. GWAS hits for bilateral convergent strabismus with exophthalmos in holstein cattle using imputed sequence level genotypes. Genes (Basel). 12, 1–12 https://doi.org/10.3390/genes12071039.

Boggild, S., Molgaard, S., Glerup, S., & Nyengaard, J. R. 2016. Spatiotemporal patterns of sortilin and SorCS2 localization during organ development. BMC Cell Biol. 17, 1–18 https://doi.org/10.1186/s12860-016-0085-9.

Bohórquez, M. D., Ordoñez, D., Suárez, C. F., Vicente, B., Vieira, C., López-Abán, J., Muro, A., Ordóñez, I., & Patarroyo, M. A. 2020. Major Histocompatibility Complex Class II (DRB3) Genetic Diversity in Spanish Morucha and Colombian Normande Cattle Compared to Taurine and Zebu Populations. Front. Genet. 10, 1–13 https://doi.org/10.3389/fgene.2019.01293.

Bollongino, R., Burger, J., Powell, A., Mashkour, M., Vigne, J. D., & Thomas, M. G. 2012. Modern taurine cattle descended from small number of near-eastern founders. Mol. Biol. Evol. 29, 2101–2104

https://doi.org/10.1093/molbev/mss092.
Boussaha, M., Esquerré, D., Barbieri, J., Djari, A., Pinton, A., Letaief, R., Salin, G., Escudié, F., Roulet, A., Fritz, S., Samson, F., Grohs, C., Bernard, M., Klopp, C., Boichard, D., & Rocha, D. 2015. Genome-wide study of structural variants in bovine Holstein, Montbéliarde and Normande dairy breeds. PLoS One 10 https://doi.org/10.1371/journal.pone.0135931.
Boutron, I., Page, M. J., Higgins, J. P. T., Altman, D. G., Lundh, A., & Hróbjartsson, A. 2019. Considering bias and conflicts of interest among the included studies.Pages 177–204 in Cochrane Handbook for Systematic Reviews of Interventions.
Braz, C. U., Taylor, J. F., Bresolin, T., Espigolan, R., Feitosa, F. L. B., Carvalheiro, R., Baldi, F., De Albuquerque, L. G., & De Oliveira, H. N. 2019. Sliding window haplotype approaches overcome single SNP analysis limitations in identifying genes for meat tenderness in Nelore cattle. BMC Genet. 20, 1–12 https://doi.org/10.1186/s12863-019-0713-4.
Brenig, B., Beck, J., Floren, C., Bornemann-Kolatzki, K., Wiedemann, I., Hennecke, S., Swalve, H., & Schütz, E. 2013. Molecular genetics of coat colour variations in White Galloway and White Park cattle. Anim. Genet. 44, 450–453 https://doi.org/10.1111/age.12029.
Buggiotti, L., Yurchenko, A., Yudin, N., & Larkin, D. M. 2020. The genomes and mechanisms of adaptation to the cold climates in Russian native cattle breeds.Page 206 in BIOINFORMATICS OF GENOME REGULATION AND STRUCTURE/SYSTEMS BIOLOGY (BGRS/SB-2020).
Burt, D. W. 2009. The cattle genome reveals its secrets. J. Biol. 8 https://doi.org/10.1186/jbiol137.
Butty, A. M., Chud, T. C. S., Cardoso, D. F., Lopes, L. S. F., Miglior, F., Schenkel, F. S., Cánovas, A., Häfliger, I. M., Drögemüller, C., Stothard, P., Malchiodi, F., & Baes, C. F. 2021. Genome-wide association study between copy number variants and hoof health traits in Holstein dairy cattle. J. Dairy Sci. 104, 8050–8061 https://doi.org/10.3168/jds.2020-19879.
Butty, A. M., Chud, T. C. S. S., Miglior, F., Schenkel, F. S., Kommadath, A., Krivushin, K., Grant, J. R., Häfliger, I. M., Drögemüller, C., Cánovas, A., Stothard, P., & Baes, C. F. 2020. High confidence copy number variants identified in Holstein dairy cattle from whole genome sequence and genotype array data. Sci. Rep. 10, 1–13 https://doi.org/10.1038/s41598-020-64680-3.
Canavez, F. C., Luche, D. D., Stothard, P., Leite, K. R. M., Sousa-Canavez, J. M., Plastow, G., Meidanis, J., Souza, M. A., Feijao, P., Moore, S. S., & Camara-Lopes, L. H. 2012. Genome sequence and assembly of bos indicus. J. Hered. 103, 342–348 https://doi.org/10.1093/jhered/esr153.
Cao, X.-K., Huang, Y.-Z., Ma, Y.-L., Cheng, J., Qu, Z.-X., Ma, Y., Bai, Y.-Y., Tian, F., Lin, F.-P., Ma, Y.-L., & Chen, H. 2018. Integrating CNVs into meta-QTL identified GBP4 as positional candidate for adult cattle stature. Funct. Integr. Genomics 18, 559–567 https://doi.org/10.1007/s10142-018-0613-0.
Cardoso, D. F., de Albuquerque, L. G., Reimer, C., Qanbari, S., Erbe, M., Do Nascimento, A. V, Venturini, G. C., Scalez, D. C. B., Baldi, F., & de Camargo, G. M. F. 2018. Genome-wide scan reveals population stratification and footprints of recent selection in Nelore cattle. Genet. Sel. Evol. 50, 1–12.
Carmo, A. S., Santos, G., Teixeira, A. L., Silva, M., Moura, M. I., Mascioli, A., & Fioravanti, M. C. 2019. Identification of copy number variants in Curraleiro Pé Duro breed. J. Anim. Sci. 268 -269. Oxford University Press US.
Casas, E., White, S. N., Shackelford, S. D., Wheeler, T. L., Koohmaraie, M., Bennett, G. L., & Smith, T. P. L. 2007. Assessing the association of single nucleotide polymorphisms at the thyroglobulin gene with carcass traits in beef cattle. J. Anim. Sci. 85, 2807–2814 https://doi.org/10.2527/jas.2007-0179.
Casas, E., White, S. N., Wheeler, T. L., Shackelford, S. D., Koohmaraie, M., Riley, D. G., Chase, C. C., Johnson, D. D., & Smith, T. P. L. 2006a. Effects of calpastatin and μ-calpain markers in beef cattle on tenderness traits. J. Anim. Sci. 84, 520–525 https://doi.org/10.2527/2006.843520x.
Casas, E., White, S. N., Wheeler, T. L., Shackelford, S. D., Koohmaraie, M., Riley, D. G., Chase, C. C. J., Johnson, D. D., & Smith, T. P. L. 2006b. Effects of calpastatin and micro-calpain markers in beef cattle on tenderness traits. J. Anim. Sci. 84, 520–525 https://doi.org/10.2527/2006.843520x.
Cases, S., Zhou, P., Shillingford, J. M., Wiseman, B. S., Fish, J. D., Angle, C. S., Hennighausen, L., Werb, Z., & Farese Jr, R. V. 2004. Development of the mammary gland requires DGAT1 expression in stromal and epithelial tissues. Development 131, 3047–3055 https://doi.org/10.1242/dev.01158.
Cassandri, M., Smirnov, A., Novelli, F., Pitolli, C., Agostini, M., Malewicz, M., Melino, G., & Raschellà, G. 2017. Zinc-finger proteins in health and disease. Cell Death Discov. 3, 17071 https://doi.org/10.1038/cddiscovery.2017.71.
Cerri, R. L. A., Thompson, I. M., Kim, I. H., Ealy, A. D., Hansen, P. J., Staples, C. R., Li, J. L., Santos, J. E. P., & Thatcher, W. W. 2012. Effects of lactation and pregnancy on gene expression of endometrium of Holstein cows at day 17 of the estrous cycle or pregnancy. J. Dairy Sci. 95, 5657–5675 https://doi.org/10.3168/jds.2011-5114.
Chang, T., Yang, Y., Yasue, H., Bharti, A. K., Retzel, E. F., & Liu, W. 2011. The Expansion of the PRAME Gene Family in Eutheria. 6 https://doi.org/10.1371/journal.pone.0016867.
Chen, N., Fu, W., Zhao, J., Shen, J., Chen, Q., Zheng, Z., Chen, H., Sonstegard, T. S., Lei, C., & Jiang, Y. 2020. BGVD: An Integrated Database for Bovine Sequencing Variations and Selective Signatures.

Genomics, Proteomics Bioinforma. 18, 186–193 https://doi.org/10.1016/j.gpb.2019.03.007.
Chen, C., Hsu, H., Hudgens, E., Telfer, J. C., & Baldwin, C. L. 2014. Signal Transduction by Different Forms of the γδ T Cell–Specific Pattern Recognition Receptor WC1. J. Immunol. 193, 379–390 https://doi.org/10.4049/jimmunol.1400168.
Cheng, W. N., & Han, S. G. 2020. Bovine mastitis: risk factors, therapeutic strategies, and alternative treatments — A review. Asian-Australasian J. Anim. Sci. 33, 1699–1713 https://doi.org/10.5713/ajas.20.0156.
Cheng, J., Jiang, R., Cao, X. K., Liu, M., Huang, Y. Z., Lan, X. Y., Cao, H., Lei, C. Z., & Chen, H. 2019a. Association analysis of SSTR2 copy number variation with cattle stature and its expression analysis in Chinese beef cattle. J. Agric. Sci. 157, 365–374 https://doi.org/10.1017/S0021859619000674.
Cheng, J., Peng, W., Cao, X., Huang, Y., Lan, X., Lei, C., & Chen, H. 2019b. Differential expression of kcnj12 gene and association analysis of its missense mutation with growth traits in chinese cattle. Animals 9, 1–13 https://doi.org/10.3390/ani9050273.
Chi, X., Li, Y., & Qiu, X. 2020. V(D)J recombination, somatic hypermutation and class switch recombination of immunoglobulins: mechanism and regulation. Immunology 160, 233–247 https://doi.org/10.1111/imm.13176.
Chingala, G., Mapiye, C., Raffrenato, E., Hoffman, L., & Dzama, K. 2017. Determinants of smallholder farmers' perceptions of impact of climate change on beef production in Malawi. Clim. Change 142, 129–141 https://doi.org/10.1007/s10584-017-1924-1.
Choi, J. W. J.-W., Chung, W.-H. W. H., Lim, K.-S. K. S., Lim, W.-J. W. J., Choi, B. H. B.-H. B. H. B.-H., Lee, S.-S. S. S. H. S.-S. S.-H., Kim, H.-C. H. C., Lee, S.-S. S. S. H. S.-S. S.-H., Cho, E.-S. E. S., Lee, K.-T. K. T., Kim, N., Kim, J.-D. J. D. J.-B. J. B., Kim, J.-D. J. D. J.-B. J. B., Chai, H.-H. H. H., Cho, Y.-M. Y. M., Kim, T. H. T.-H., & Lim, D. 2016. Copy number variations in Hanwoo and Yanbian cattle genomes using the massively parallel sequencing data. Gene 589, 36–42 https://doi.org/10.1016/j.gene.2016.05.017.
Choi, J. W. J.-W., Lee, K. T. K.-T., Liao, X., Stothard, P., An, H. S. H.-S., Ahn, S., Lee, S.-Y. S. S. Y., Lee, S.-Y. S. S. Y., Moore, S. S., & Kim, T.-H. T. H. 2013. Genome-wide copy number variation in Hanwoo, Black Angus, and Holstein cattle. Mamm. Genome 24, 151–163 https://doi.org/10.1007/s00335-013-9449-z.
Cicconardi, F., Chillemi, G., Tramontano, A., Marchitelli, C., Valentini, A., Ajmone-Marsan, P., & Nardone, A. 2013. Massive screening of copy number population-scale variation in Bos taurus genome. BMC Genomics 14 https://doi.org/10.1186/1471-2164-14-124.
Clop, A., Vidal, O., & Amills, M. 2012. Copy number variation in the genomes of domestic animals. Anim. Genet. 43, 503–517 https://doi.org/10.1111/j.1365-2052.2012.02317.x.
Cochran, S. D., Cole, J. B., Null, D. J., & Hansen, P. J. 2013. Discovery of single nucleotide polymorphisms in candidate genes associated with fertility and production traits in Holstein cattle. BMC Genet. 14, 1 https://doi.org/10.1186/1471-2156-14-49.
Colella, S., Yau, C., Taylor, J. M., Mirza, G., Butler, H., Clouston, P., Bassett, A. S., Seller, A., Holmes, C. C., & Ragoussis, J. 2007. QuantiSNP: An objective bayes hidden-markov model to detect and accurately map copy number variation using SNP genotyping data. Nucleic Acids Res. 35, 2013–2025 https://doi.org/10.1093/nar/gkm076.
Connor, E. E., Zhou, Y., & Liu, G. E. 2018. The essence of appetite: Does olfactory receptor variation play a role? J. Anim. Sci. 96, 1551–1558 https://doi.org/10.1093/jas/sky068.
Cormican, P., Meade, K. G., Cahalane, S., Narciandi, F., Chapwanya, A., Lloyd, A. T., & O'Farrelly, C. 2008. Evolution, expression and effectiveness in a cluster of novel bovine β-defensins. Immunogenetics 60, 147–156 https://doi.org/10.1007/s00251-007-0269-8.
Couldrey, C., Keehan, M., Johnson, T., Tiplady, K., Winkelman, A., Littlejohn, M. D., Scott, A., Kemper, K. E., Hayes, B., Davis, S. R., & Spelman, R. J. 2017. Detection and assessment of copy number variation using PacBio long-read and Illumina sequencing in New Zealand dairy cattle. J. Dairy Sci. 100, 5472–5478 https://doi.org/10.3168/jds.2016-12199.
Cozzi, M. C., Martinez-Ruiz, C. P., Roman-Ponce, S. I., Murillo, V. E. V, Utrera, Á. R., Montaño-Bermúdez, M. M., Martinez-Velazquez, G., & Strillacci, M. G. 2019. Copy number variants reveal genomic diversity in a Mexican Creole cattle population. Livest. Sci. 229, 194–202 https://doi.org/10.1016/j.livsci.2019.09.030.
Cui, M., Huang, J., Zhang, S., Liu, Q., Liao, Q., & Qiu, X. 2021. Immunoglobulin Expression in Cancer Cells and Its Critical Roles in Tumorigenesis. Front. Immunol. 12, 1–10 https://doi.org/10.3389/fimmu.2021.613530.
Cymbron, T., Freeman, A. R., Isabel Malheiro, M., Vigne, J.-D., & Bradley, D. G. 2005. Microsatellite diversity suggests different histories for Mediterranean and Northern European cattle populations. Proc. R. Soc. B Biol. Sci. 272, 1837–1843.
Dean, M., Moitra, K., & Allikmets, R. 2002. The human ATP-binding cassette (ABC) transporter superfamily. Hum. Mutat. 43, 1162–1182 https://doi.org/10.1002/humu.24418.
Decker, J. E., McKay, S. D., Rolf, M. M., Kim, J. W., Molina Alcalá, A., Sonstegard, T. S., Hanotte, O., Götherström, A., Seabury, C. M., Praharani, L., Babar, M. E., Correia de Almeida Regitano, L., Yildiz, M.

A., Heaton, M. P., Liu, W. S., Lei, C. Z., Reecy, J. M., Saif-Ur-Rehman, M., Schnabel, R. D., & Taylor, J. F. 2014. Worldwide Patterns of Ancestry, Divergence, and Admixture in Domesticated Cattle. PLoS Genet. 10 https://doi.org/10.1371/journal.pgen.1004254.

Diamond, G., Zasloff, M., Eck, H., Brasseur, M., Lee Maloy, W., & Bevins, C. L. 1991. Tracheal antimicrobial peptide, a cysteine-rich peptide from mammalian tracheal mucosa: Peptide isolation and cloning of a cDNA. Proc. Natl. Acad. Sci. U. S. A. 88, 3952–3956 https://doi.org/10.1073/pnas.88.9.3952.

Dikmen, S., Cole, J. B., Null, D. J., & Hansen, P. J. 2013. Genome-Wide Association Mapping for Identification of Quantitative Trait Loci for Rectal Temperature during Heat Stress in Holstein Cattle. PLoS One 8, 1–7 https://doi.org/10.1371/journal.pone.0069202.

Dikmen, S., Khan, F. A., Huson, H. J., Sonstegard, T. S., Moss, J. I., Dahl, G. E., & Hansen, P. J. 2014. The SLICK hair locus derived from Senepol cattle confers thermotolerance to intensively managed lactating Holstein cows. J. Dairy Sci. 97, 5508–5520 https://doi.org/10.3168/jds.2014-8087.

Ding, X. T., Zhang, Z. J., Hu, R. F., Wen, Y. F., Huang, Y. Z., Shi, Q. T., Feng, Y. J., Wang, E. Y., Lei, C. Z., & He, H. 2022. A molecular marker of milk composition traits in NCAM2 gene of Chinese Holstein. Anim. Biotechnol. 33, 79–84 https://doi.org/10.1080/10495398.2020.1772802 WE - Science Citation Index Expanded (SCI-EXPANDED).

Dixit, S. P., Sivalingam, J., Tyagi, A. K., Saroha, V., Sharma, A., & Nagda, R. K. 2015. Association of novel SNPs in the candidate genes affecting caprine milk fatty acids related to human health. Meta Gene 4, 45–56 https://doi.org/10.1016/j.mgene.2015.01.004.

Dolezal, M. A., Bagnato, A., Schiavini, F., Santus, E., Holm, L. E., Bendixen, C., & Panitz, F. 2014. Copy Number Variation in Brown Swiss Dairy Cattle.in 10th World Congress on Genetics Applied to Livestock Production (WCGALP).

De Donato, M., Peters, S. O., Mitchell, S. E., Hussain, T., & Imumorin, I. G. 2013. Genotyping-by-Sequencing (GBS): A Novel, Efficient and Cost-Effective Genotyping Method for Cattle Using Next-Generation Sequencing. PLoS One 8 https://doi.org/10.1371/journal.pone.0062137.

Dong, C. H., Song, X. M., Zhang, L., Jiang, J. F., Zhou, J. P., & Jiang, Y. Q. 2013. New insights into the prolactin-RsaI (PRL-RsaI) locus in Chinese Holstein cows and its effect on milk performance traits. Genet. Mol. Res. 12, 5766–5773 https://doi.org/10.4238/2013.November.22.3.

Drake, D. J. 2004. Understanding and Improving Beef Cattle Carcass Quality. Underst. Improv. Beef Cattle Carcass Qual. https://doi.org/10.3733/ucanr.8130.

Duan, J., Zhang, J. G., Deng, H. W., & Wang, Y. P. 2013. Comparative Studies of Copy Number Variation Detection Methods for Next-Generation Sequencing Technologies. PLoS One 8, 1–12 https://doi.org/10.1371/journal.pone.0059128.

Duarte, I. N. H., Bessa, A. F. de O., Rola, L. D., Genuíno, M. V. H., Rocha, I. M., Marcondes, C. R., Regitano, L. C. de A., Munari, D. P., Berry, D. P., & Buzanskas, M. E. 2022. Cross-population selection signatures in Canchim composite beef cattle. PLoS One 17, e0264279 https://doi.org/10.1371/journal.pone.0264279.

Durán Aguilar, M., Román Ponce, S. I., Ruiz López, F. J., González Padilla, E., Vásquez Peláez, C. G., Bagnato, A., Strillacci, M. G., Aguilar, M. D., Ponce, S. I. R., Lopez, F. J. R., Padilla, E. G., Pelaez, C. G. V, Bagnato, A., & Strillacci, M. G. 2017. Genome-wide association study for milk somatic cell score in holstein cattle using copy number variation as markers. J. Anim. Breed. Genet. 134, 49–59 https://doi.org/10.1111/jbg.12238.

Durkin, K., Coppieters, W., Drögüller, C., Ahariz, N., Cambisano, N., Druet, T., Fasquelle, C., Haile, A., Horin, P., Huang, L., Kamatani, Y., Karim, L., Lathrop, M., Moser, S., Oldenbroek, K., Rieder, S., Sartelet, A., Sölkner, J., Stålhammar, H., Zelenika, D., Zhang, Z., Leeb, T., Georges, M., & Charlier, C. 2012. Serial translocation by means of circular intermediates underlies colour sidedness in cattle. Nature 482, 81–84 https://doi.org/10.1038/nature10757.

Dzama, K. 2016. Is the livestock sector in Southern Africa prepared for climate change. South African Inst. Int. Aff. Policy Brief., 1–4.

E, G. X., Na, R. S., Zhao, Y. J., Chen, L. P., Qiu, X. Y., & Huang, Y. F. 2015. Variability in the cathelicidin 6 (CATHL-6) gene in Tianzhu white yak from Tibetan area in China. Genet. Mol. Res. 14, 3129–3132 https://doi.org/10.4238/2015.April.10.23.

Ealy, A. D., & Wooldridge, L. K. 2017. The evolution of interferon-tau. https://doi.org/10.1530/REP-17-0292.

Eck, S. H., Benet-Pagès, A., Flisikowski, K., Meitinger, T., Fries, R., & Strom, T. M. 2009. Whole genome sequencing of a single Bos taurus animal for single nucleotide polymorphism discovery. Genome Biol. 10 https://doi.org/10.1186/gb-2009-10-8-r82.

Edeal, B., Rumph, J. M., Mass, R., Killinger, K., Jerez, N., Elnagar, S., McDaneld, T., Aljumaah, R., Pithpongsiriporn, U., Nephawe, K., Martinez, G., Gladney, C. D., Allan, M. F., & Pomp, D. 2000. Rapid communication: Linkage mapping of the Mahogany (attractin) locus in cattle and pigs. J. Anim. Sci. 78, 2479–2482 https://doi.org/10.2527/2000.7892479x.

Ekman, A., Niku, M., Liljavirta, J., & Iivanainen, A. 2009. Bos taurus genome sequence reveals the assortment of immunoglobulin and surrogate light chain genes in domestic cattle. BMC Immunol. 10, 1–11

https://doi.org/10.1186/1471-2172-10-22.
Ellegren, H. 2004. Microsatellites: Simple sequences with complex evolution. Nat. Rev. Genet. 5, 435–445 https://doi.org/10.1038/nrg1348.
Ellenbroek, B., & Youn, J. 2016. Chapter 2 - The Genetic Basis of Behavior.Pages 19–46 in Ellenbroek, B., Youn, J.B.T.-G.-E.I. in P., eds. Academic Press, San Diego.
Elsik, C. G., Tellam, R. L., & Worley, K. C. 2009. Genome Sequence of Taurine Cattle: A Window to Ruminant Biology and Evolution The Bovine Genome Sequencing and Analysis Consortium. Science (80-.). 324, 17 https://doi.org/10.1126/science.1169588.The.
Escarcha, J. F., Lassa, J. A., & Zander, K. K. 2018. Livestock under climate change: A systematic review of impacts and adaptation. Climate 6, 1–17 https://doi.org/10.3390/cli6030054.
Fadista, J., Thomsen, B., Holm, L., & Bendixen, C. 2010. Copy number variation in the bovine genome. BMC Genomics 11 https://doi.org/10.1186/1471-2164-11-284.
Farke, C., Meyer, H. H. D., Bruckmaier, R. M., & Albrecht, C. 2008. Differential expression of ABC transporters and their regulatory genes during lactation and dry period in bovine mammary tissue. J. Dairy Res. 75, 406–414 https://doi.org/10.1017/S002202990800335X.
Food and Agriculture Oraganization (FAO) of the United Nations (2018, 2019, 2020) *Data*. Available at: https://www.fao.org/faostat/en/#data (Accessed 14 November 2022).
Feitosa, F. L. B., Pereira, A. S. C., Venturini, G. C., Tonussi, R. L., Espigolan, R., Gordo, D. M., Camargo, G. M. F., Chardulo, L. A. L., Albuquerque, L. G., & Baldi, F. 2014. Genome wide association study between copy number variation regions with marbling score in Nelore cattle.in Proceedings, 10th World Congress of Genetics Applied to Livestock Production.
Felius, M., Beerling, M. L., Buchanan, D. S., Theunissen, B., Koolmees, P. A., & Lenstra, J. A. 2014. On the history of cattle genetic resources. Diversity 6, 705–750 https://doi.org/10.3390/d6040705.
Feng, X., Jiang, J., Padhi, A., Ning, C., Fu, J., Wang, A., Mrode, R., & Liu, J.-F. 2017. Characterization of genome-wide segmental duplications reveals a common genomic feature of association with immunity among domestic animals. BMC Genomics 18 https://doi.org/10.1186/s12864-017-3690-x.
Fernandes Júnior, G. A., de Oliveira, H. N., Carvalheiro, R., Cardoso, D. F., Fonseca, L. F. S., Ventura, R. V., & de Albuquerque, L. G. 2020. Whole-genome sequencing provides new insights into genetic mechanisms of tropical adaptation in Nellore (Bos primigenius indicus). Sci. Rep. 10, 1–7 https://doi.org/10.1038/s41598-020-66272-7.
Feuk, L., Carson, A. R., & Scherer, S. W. 2006. Structural variation in the human genome. Nat. Rev. Genet. 7, 85–97 https://doi.org/10.1038/nrg1767.
Filho, R. V. O., Franco, G. A., Reese, S. T., Dantas, F. G., Fontes, P. L. P., Cooke, R. F., Rhinehart, J. D., Thompson, K. W., & Pohler, K. G. 2019. Using pregnancy associated glycoproteins (PAG) for pregnancy detection at day 24 of gestation in beef cattle. Theriogenology https://doi.org/10.1016/j.theriogenology.2019.09.014.
Flores, E. G. 2011. Chacaterization of the bovine cathelicidin gene family.
Fontanesi, L., Calò, D. G., Galimberti, G., Negrini, R., Marino, R., Nardone, A., Ajmone-Marsan, P., & Russo, V. 2014. A candidate gene association study for nine economically important traits in Italian Holstein cattle. Anim. Genet. 45, 576–580 https://doi.org/10.1111/age.12164.
Fontanesi, L., Tazzoli, M., Russo, V., & Beever, J. 2009. Genetic heterogeneity at the bovine KIT gene in cattle breeds carrying different putative alleles at the spotting locus. Anim. Genet. 41, 295–303 https://doi.org/10.1111/j.1365-2052.2009.02007.x.
Forde, N., & Lonergan, P. 2017. Interferon-tau and fertility in ruminants. Reproduction 154, F33–F43 https://doi.org/10.1530/REP-17-0432.
Formanack, M. L., & Baier, L. J. 2004. Variation in the FABP2 promoter affects gene expression: Implications for prior association studies. Diabetologia 47, 349–351 https://doi.org/10.1007/s00125-003-1289-z.
Freeman, J. L., Perry, G. H., Feuk, L., Redon, R., McCarroll, S. A., Altshuler, D. M., Aburatani, H., Jones, K. W., Tyler-Smith, C., Hurles, M. E., Carter, N. P., Scherer, S. W., & Lee, C. 2006. Copy number variation: New insights in genome diversity. Genome Res. 16, 949–961 https://doi.org/10.1101/gr.3677206.
Frischknecht, M., Bapst, B., Seefried, F. R., Signer-Hasler, H., Garrick, D., Stricker, C., Fries, R., Russ, I., Sölkner, J., Bieber, A., Strillacci, M. G., Gredler-Grandl, B., & Flury, C. 2017. Genome-wide association studies of fertility and calving traits in Brown Swiss cattle using imputed whole-genome sequences. BMC Genomics 18, 1–13 https://doi.org/10.1186/s12864-017-4308-z.
Fukunaga, K., Yamashita, Y., & Yagisawa, T. 2020. Copy number variations in BOLA-DQA2, BOLA-DQB, and BOLA-DQA5 show the genomic architecture and haplotype frequency of major histocompatibility complex class II genes in Holstein cows. HLA 96, 601–609 https://doi.org/10.1111/tan.14086.
Furlani, B., Kouter, K., Rozman, D., & Paska, A. V. 2021. Sequencing of nucleic acids: From the first human genome to next generation sequencing in COVID-19 pandemic. Acta Chim. Slov. 68, 268–278 https://doi.org/10.17344/acsi.2021.6691.
Gadberry, S., & Powell, J. 2008. Internal Parasites in Beef and Dairy Cattle. Univ. Arkansas Div. Agric., 1–6.
Gan, Q. F., Zhang, L. P., Li, J. Y., Hou, G. Y., Li, H. De, Gao, X., Ren, H. Y., Chen, J. B., & Xu, S. Z. 2008.

Association analysis of thyroglobulin gene variants with carcass and meat quality traits in beef cattle. J. Appl. Genet. 49, 251–255 https://doi.org/10.1007/BF03195621.

Gao, Y., Jiang, J., Yang, S., Hou, Y., Liu, G. E., Zhang, S., Zhang, Q., & Sun, D. 2017. CNV discovery for milk composition traits in dairy cattle using whole genome resequencing. BMC Genomics 18 https://doi.org/10.1186/s12864-017-3636-3.

Gao, H., Wu, Y., Li, J., Li, H., Li, J., & Yang, R. 2014. Forward LASSO analysis for high-order interactions in genome-wide association study. Brief. Bioinform. 15, 552–561 https://doi.org/10.1093/bib/bbt037.

Garrido, C., Gurbuxani, S., Ravagnan, L., & Kroemer, G. 2001. Heat shock proteins: Endogenous modulators of apoptotic cell death. Biochem. Biophys. Res. Commun. 286, 433–442 https://doi.org/10.1006/bbrc.2001.5427.

Gartlehner, G., Affengruber, L., Titscher, V., Noel-Storr, A., Dooley, G., Ballarini, N., & König, F. 2020. Single-reviewer abstract screening missed 13 percent of relevant studies: a crowd-based, randomized controlled trial. J. Clin. Epidemiol. 121, 20–28 https://doi.org/10.1016/j.jclinepi.2020.01.005.

Gasbarre, L. C., Leighton, E. A., & Sonstegard, T. 2001. Role of the bovine immune system and genome in resistance to gastrointestinal nematodes. Vet. Parasitol. 98, 51–64 https://doi.org/10.1016/S0304-4017(01)00423-X.

Gatew, H., & Tarekegn, G. M. 2018. Next-generation sequencing platforms for latest livestock reference genome assemblies. African J. Biotechnol. 17, 1232–1240 https://doi.org/10.5897/ajb2018.16605.

Gautason, E., Schönherz, A. A., Sahana, G., & Guldbrandtsen, B. 2021. Genomic inbreeding and selection signatures in the local dairy breed Icelandic Cattle. Anim. Genet. 52, 251–262.

Di Gerlando, R., Sardina, M. T., Tolone, M., Sutera, A. M., Mastrangelo, S., & Portolano, B. 2019. Genome-wide detection of copy-number variations in local cattle breeds. Anim. Prod. Sci. 59, 815 https://doi.org/10.1071/AN17603.

Ghanipoor-Samami, M., Javadmanesh, A., Burns, B. M., Thomsen, D. A., Nattrass, G. S., Estrella, C. A. S., Kind, K. L., & Hiendleder, S. 2018. Atlas of tissue- and developmental stage specific gene expression for the bovine insulin-like growth factor (IGF) system. PLoS One 13, 1–28 https://doi.org/10.1371/journal.pone.0200466.

Glerup, S., Olsen, D., Vaegter, C. B., Gustafsen, C., Sjoegaard, S. S., Hermey, G., Kjolby, M., Molgaard, S., Ulrichsen, M., Boggild, S., Skeldal, S., Fjorback, A. N., Nyengaard, J. R., Jacobsen, J., Bender, D., Bjarkam, C. R., Sørensen, E. S., Füchtbauer, E. M., Eichele, G., Madsen, P., Willnow, T. E., Petersen, C. M., & Nykjaer, A. 2014. SorCS2 regulates dopaminergic wiring and is processed into an apoptotic two-chain receptor in peripheral glia. Neuron 82, 1074–1087 https://doi.org/10.1016/j.neuron.2014.04.022.

Glick, G., Shirak, A., Seroussi, E., Zeron, Y., Ezra, E., Weller, J. I., & Ron, M. 2011. Fine mapping of a QTL for fertility on BTA7 and its association with a CNV in the israeli holsteins. G3 Genes, Genomes, Genet. 1, 65–74 https://doi.org/10.1534/g3.111.000299.

Glusman, G., Bahar, A., Sharon, D., Pilpel, Y., White, J., & Lancet, D. 2000. The olfactory receptor gene superfamily: Data mining, classification, and nomenclature. Mamm. Genome 11, 1016–1023 https://doi.org/10.1007/s003350010196.

Goldman, A. D., & Landweber, L. F. 2016. What Is a Genome? PLoS Genet. 12, 1–7 https://doi.org/10.1371/journal.pgen.1006181.

Golik, M., Cohen-Zinder, M., Loor, J. J., Drackley, J. K., Band, M. R., Lewin, H. A., Weller, J. I., Ron, M., & Seroussi, E. 2006. Accelerated expansion of group IID-like phospholipase A2 genes in Bos taurus. Genomics 87, 527–533 https://doi.org/10.1016/j.ygeno.2005.12.015.

Gomez, L. C., Real, S. M., Ojeda, M. S., Gimenez, S., Mayorga, L. S., & Roqué, M. 2007. Polymorphism of the FABP2 gene: A population frequency analysis and an association study with cardiovascular risk markers in Argentina. BMC Med. Genet. 8, 1–7 https://doi.org/10.1186/1471-2350-8-39.

González-Quintero, R., Barahona-Rosales, R., Bolívar-Vergara, D. M., Chirinda, N., Arango, J., Pantévez, H. A., Correa-Londoño, G., & Sánchez-Pinzón, M. S. 2020. Technical and environmental characterization of dual-purpose cattle farms and ways of improving production: A case study in Colombia. Pastoralism 10 https://doi.org/10.1186/s13570-020-00170-5.

Goodall, J. J., & Schmutz, S. M. 2007. IGF2 gene characterization and association with rib eye area in beef cattle. Anim. Genet. 38, 154–161 https://doi.org/10.1111/j.1365-2052.2007.01576.x.

Gororo, E., Makuza, S. M., Chatiza, F. P., Chidzwondo, F., & Sanyika, T. W. 2018. Genetic diversity in Zimbabwean Sanga cattle breeds using microsatellite markers. S. Afr. J. Anim. Sci. 48, 128 https://doi.org/10.4314/sajas.v48i1.15.

Goshu, H. A., Chu, M., & Yan, P. 2018a. Applications of genomic copy number variations on livestock: A review. African J. Biotechnol. 17, 1313–1323.

Goshu, H. A., Wu, X. Y., Chu, M., Bao, P. J., Ding, X. Z., & Yan, P. 2018b. Copy Number Variations of KLF6 Modulate Gene Transcription and Growth Traits in Chinese Datong Yak (Bos Grunniens). ANIMALS 8 https://doi.org/10.3390/ani8090145 WE - Science Citation Index Expanded (SCI-EXPANDED).

Grandin, T., & Deesing, M. J. 2014. Chapter 1 - Behavioral Genetics and Animal Science.Pages 1–40 in

Grandin, T., Deesing, M.J.B.T.-G. and the B. of D.A. (Second E., eds. Academic Press, San Diego.
Graphodatskaya, D., Perelman, P., Serdukova, N., Joerg, H., & Stranzinger, G. 2003. Assignment of the bovine attractin (ATRN) gene to chromosome 13q21→q22 by in situ hybridization. Cytogenet. Genome Res. 103, 22–23 https://doi.org/10.1159/000076311.
Green, S. R., Fullekrug, J., Sauer, K., & Tuite, M. F. 1991. Isolation and characterization of a bovine cDNA encoding eukaryotic initiation factor 2α. Biochim. Biophys. Acta - Gene Struct. Expr. 1090, 277–280 https://doi.org/https://doi.org/10.1016/0167-4781(91)90119-7.
Grisart, B., Farnir, F., Karim, L., Cambisano, N., Kim, J. J., Kvasz, A., Mni, M., Simon, P., Frère, J. M., Coppieters, W., & Georges, M. 2004. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. Proc. Natl. Acad. Sci. U. S. A. 101, 2398–2403 https://doi.org/10.1073/pnas.0308518100.
Grobet, L., Martin, L. J., Poncelet, D., Pirottin, D., Brouwers, B., Riquet, J., Schoeberlein, A., Dunner, S., Ménissier, F., Massabanda, J., Fries, R., Hanset, R., & Georges, M. 1997. A deletion in the bovine myostatin gene causes the double-muscled phenotype in cattle. Nat. Genet. 17, 71–74 https://doi.org/10.1038/ng0997-71.
Gu, W., Zhang, F., & Lupski, J. R. 2008. Mechanisms for human genomic rearrangements. Pathogenetics 1, 1–17 https://doi.org/10.1186/1755-8417-1-4.
Gunn, T. M., Inui, T., Kitada, K., Ito, S., Wakamatsu, K., He, L., Bouley, D. M., Serikawa, T., & Barsh, G. S. 2001. Molecular and phenotypic analysis of Attractin mutant mice. Genetics 158, 1683–1695 https://doi.org/10.1093/genetics/158.4.1683.
Guo, X., Pei, J., Wu, X., Bao, P., Ding, X., Xiong, L., Chu, M., Lan, X., & Yan, P. 2020a. Detection of InDel and CNV of SPAG17 gene and their associations with bovine growth traits. Anim. Biotechnol. https://doi.org/10.1080/10495398.2020.1803342.
Guo, Y., Sheng, Q., Samuels, D. C., Lehmann, B., Bauer, J. A., Pietenpol, J., & Shyr, Y. 2013. Comparative Study of Exome Copy Number Variation Estimation Tools Using Array Comparative Genomic Hybridization as Control. Biomed Res. Int., 7409598 https://doi.org/10.1155/2017/7409598.
Guo, S., Wu, X., Pei, J., Wang, X., Bao, P., Xiong, L., Chu, M., Liang, C., Yan, P., & Guo, X. 2021. Genome-wide CNV analysis reveals variants associated with high-altitude adaptation and meat traits in Qaidam cattle. Electron. J. Biotechnol. 54, 8–16 https://doi.org/10.1016/j.ejbt.2021.07.006.
Guo, J. Z., Zhong, J., Liu, G. E., Yang, L., Li, L., Chen, G. L., Song, T. Z., & Zhang, H. P. 2020b. Identification and population genetic analyses of copy number variations in six domestic goat breeds and Bezoar ibexes using next-generation sequencing. BMC Genomics 21 https://doi.org/10.1186/s12864-020-07267-6 WE - Science Citation Index Expanded (SCI-EXPANDED).
Gupta, A. K., & Gupta, U. D. 2020. Chapter 20 - Next generation sequencing and its applications.Pages 395–421 in Verma, A.S., Singh, A.B.T.-A.B. (Second E., eds. Academic Press, Boston.
Gurao, A., Kashyap, S. K., & Singh, R. 2017. β-defensins: An innate defense for bovine mastitis. Vet. World 10, 990–998 https://doi.org/10.14202/vetworld.2017.990-998.
Gurgul, A., Jasielczuk, I., Szmatoła, T., Pawlina, K., Ząbek, T., Żukowski, K., & Bugno-Poniewierska, M. 2015. Genome-wide characteristics of copy number variation in Polish Holstein and Polish Red cattle using SNP genotyping assay. Genetica 143, 145–155 https://doi.org/10.1007/s10709-015-9822-9.
Gurgul, A., Miksza-Cybulska, A., Szmatoła, T., Jasielczuk, I., Piestrzyńska-Kajtoch, A., Fornal, A., Semik-Gurgul, E., & Bugno-Poniewierska, M. 2019. Genotyping-by-sequencing performance in selected livestock species. Genomics 111, 186–195 https://doi.org/10.1016/j.ygeno.2018.02.002.
Gusenbauer, M., & Haddaway, N. R. 2020. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. Res. Synth. Methods 11, 181–217 https://doi.org/10.1002/jrsm.1378.
Gutiérrez-Gil, B., Wiener, P., & Williams, J. L. 2007. Genetic effects on coat colour in cattle: Dilution of eumelanin and phaeomelanin pigments in an F2-Backcross Charolais × Holstein population. BMC Genet. 8, 1–12 https://doi.org/10.1186/1471-2156-8-56.
Haeussler, M., Zweig, A. S., Tyner, C., Speir, M. L., Rosenbloom, K. R., Raney, B. J., Lee, C. M., Lee, B. T., Hinrichs, A. S., Gonzalez, J. N., Gibson, D., Diekhans, M., Clawson, H., Casper, J., Barber, G. P., Haussler, D., Kuhn, R. M., & Kent, W. J. 2019. The UCSC Genome Browser database: 2019 update. Nucleic Acids Res. 47, D853–D858 https://doi.org/10.1093/nar/gky1095.
Häfliger, I. M., Hirter, N., Paris, J. M., Wolf Hofstetter, S., Seefried, F. R., & Drögemüller, C. 2020. A de novo germline mutation of KIT in a white-spotted Brown Swiss cow. Anim. Genet. 51, 449–452 https://doi.org/10.1111/age.12920.
Haigh, B., Hood, K., Broadhurst, M., Medele, S., Callaghan, M., Smolenski, G., Dines, M., & Wheeler, T. 2008. The bovine salivary proteins BSP30a and BSP30b are independently expressed BPI-like proteins with anti-Pseudomonas activity. Mol. Immunol. 45, 1944–1951 https://doi.org/10.1016/j.molimm.2007.10.032.
Hao, D., Wang, X., Thomsen, B., Kadarmideen, H. N., Wang, X., Lan, X., Huang, Y., Qi, X., & Chen, H. 2020. Copy number variations and expression levels of guanylate-binding protein 6 gene associated with growth traits of chinese cattle. Animals 10 https://doi.org/10.3390/ani10040566.

Haruna, I. L., Ekegbu, U. J., Ullah, F., Amirpour-Najafabadi, H., Zhou, H., & Hickford, J. G. H. 2020. Genetic variations and haplotypic diversity in the Myostatin gene of New Zealand cattle breeds. Gene 740, 144400 https://doi.org/10.1016/j.gene.2020.144400.

Hauser, M., Signer-Hasler, H., Küttel, L., Capitan, A., Guldbrandtsen, B., Hinrichs, D., Flury, C., Seefried, F. R., & Drögemüller, C. 2022. Identification of two new recessive MC1R alleles in red-coloured Evolèner cattle and other breeds. Anim. Genet. 53, 427–435 https://doi.org/10.1111/age.13206.

Hay, E. H. A., Utsunomiya, Y. T., Xu, L., Zhou, Y., Neves, H. H. R. R., Carvalheiro, R., Bickhart, D. M., Ma, L., Garcia, J. F., & Liu, G. E. 2018. Genomic predictions combining SNP markers and copy number variations in Nellore cattle. BMC Genomics 19, 441 https://doi.org/10.1186/s12864-018-4787-6.

Hayes, B. J., Lewin, H. A., & Goddard, M. E. 2013. The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation. Trends Genet. 29, 206–214 https://doi.org/https://doi.org/10.1016/j.tig.2012.11.009.

He, L., Gunn, T. M., Bouley, D. M., Lu, X. Y., Watson, S. J., Schlossman, S. F., Duke-Cohan, J. S., & Barsh, G. S. 2001. A biochemical function for attractin in agouti-induced pigmentation and obesity. Nat. Genet. 27, 40–47 https://doi.org/10.1038/83741.

Hein, W. R., Shoemaker, C. B., & Heath, A. C. G. 2001. Future technologies for control of nematodes of sheep. N. Z. Vet. J. 49, 247–251 https://doi.org/10.1080/00480169.2001.36239.

Heringstad, B., Egger-Danner, C., Charfeddine, N., Pryce, J. E., Stock, K. F., Kofler, J., Sogstad, A. M., Holzhauer, M., Fiedler, A., Müller, K., Nielsen, P., Thomas, G., Gengler, N., de Jong, G., Ødegård, C., Malchiodi, F., Miglior, F., Alsaaod, M., & Cole, J. B. 2018. Invited review: Genetics and claw health: Opportunities to enhance claw health by genetic selection. J. Dairy Sci. 101, 4801–4821 https://doi.org/10.3168/jds.2017-13531.

Herzig, C. T. A., & Baldwin, C. L. 2009. Genomic organization and classification of the bovine WC1 genes and expression by peripheral blood gamma delta T cells. BMC Genomics 10, 1–20 https://doi.org/10.1186/1471-2164-10-191.

Hibino, H., Inanobe, A., Furutani, K., Murakami, S., Findlay, I., & Kurachi, Y. 2010. Inwardly rectifying potassium channels: Their structure, function, and physiological roles. Physiol. Rev. 90, 291–366 https://doi.org/10.1152/physrev.00021.2009.

Hocquette, J.-F., Ellies-Oury, M.-P., Lherm, M., Pineau, C., Deblitz, C., & Farmer, L. 2018. Current situation and future prospects for beef production in Europe - A review. Asian-Australasian J. Anim. Sci. 31, 968–975 https://doi.org/10.5713/ajas.18.0201.

Horsburgh, K. A., Prost, S., Gosling, A., Stanton, J. A., Rand, C., & Matisoo-Smith, E. A. 2013. The genetic diversity of the Nguni breed of African Cattle (Bos spp.): complete mitochondrial genomes of haplogroup T1. PLoS One 8, 1–6 https://doi.org/10.1371/journal.pone.0071956.

Hou, Y., Bickhart, D. M., Chung, H., Hutchison, J. L., Norman, H. D., Connor, E. E., & Liu, G. E. 2012a. Analysis of copy number variations in Holstein cows identify potential mechanisms contributing to differences in residual feed intake. Funct. Integr. Genomics 12, 717–723 https://doi.org/10.1007/s10142-012-0295-y.

Hou, Y., Bickhart, D. M., Hvinden, M. L., Li, C., Song, J., Boichard, D. A., Fritz, S., Eggen, A., DeNise, S., Wiggans, G. R., Sonstegard, T. S., Van Tassell, C. P., & Liu, G. E. 2012b. Fine mapping of copy number variations on two cattle genome assemblies using high density SNP array. BMC Genomics 13, 1–10 https://doi.org/10.1186/1471-2164-13-376.

Hou, Y., Liu, G. E., Bickhart, D. M., Cardone, M. F., Wang, K., Kim, E. soo, Matukumalli, L. K., Ventura, M., Song, J., VanRaden, P. M., Sonstegard, T. S., & Van Tassell, C. P. 2011a. Genomic characteristics of cattle copy number variations. BMC Genomics 12 https://doi.org/10.1186/1471-2164-12-127.

Hou, Y., Liu, G. E., Bickhart, D. M., Matukumalli, L. K., Li, C., Song, J., Gasbarre, L. C., Van Tassell, C. P., & Sonstegard, T. S. 2012c. Genomic regions showing copy number variations associate with resistance or susceptibility to gastrointestinal nematodes in Angus cattle. Funct. Integr. Genomics 12, 81–92 https://doi.org/10.1007/s10142-011-0252-1.

Hou, G. Y., Yuan, Z. R., Zhou, H. L., Zhang, L. P., Li, J. Y., Gao, X., Wang, D. J., Gao, H. J., & Xu, S. Z. 2011b. Association of thyroglobulin gene variants with carcass and meat quality traits in beef cattle. Mol. Biol. Rep. 38, 4705–4708 https://doi.org/10.1007/s11033-010-0605-1.

Houlahan, K., Schenkel, F. S., Hailemariam, D., Lassen, J., Kargo, M., Cole, J. B., Connor, E. E., Wegmann, S., Oliveira, G. A., Miglior, F., Fleming, A., Chud, T. C. S., & Baes, C. F. 2021. Effects of incorporating dry matter intake and residual feed intake into a selection index for dairy cattle using deterministic modeling. Animals 11 https://doi.org/10.3390/ani11041157.

Hozáková, K., Vavrišínová, K., Neirurerová, P., & Bujko, J. 2020. Growth of beef cattle as prediction for meat production: A review. Acta Fytotech. Zootech. 23, 58–69 https://doi.org/10.15414/AFZ.2020.23.02.58-69.

Hsu, H., Chen, C., Nenninger, A., Holz, L., Baldwin, C. L., & Telfer, J. C. 2015. WC1 Is a Hybrid γδ TCR Coreceptor and Pattern Recognition Receptor for Pathogenic Bacteria. J. Immunol. 194, 2280–2288 https://doi.org/10.4049/jimmunol.1402021.

Hu, S., Chen, Y., Zhao, B., Yang, N., Chen, S., Shen, J., Bao, G., & Wu, X. 2020a. KIT is involved in

melanocyte proliferation, apoptosis and melanogenesis in the Rex Rabbit. PeerJ 2020, 1–16 https://doi.org/10.7717/peerj.9402.

Hu, Y., Xia, H., Li, M., Xu, C., Ye, X., Su, R., Zhang, M., Nash, O., Sonstegard, T. S., Yang, L., Liu, G. E., & Zhou, Y. 2020b. Comparative analyses of copy number variations between Bos taurus and Bos indicus. BMC Genomics 21 https://doi.org/10.1186/s12864-020-07097-6.

Hu, L., Yu, J., Huang, R., Yang, P., Zhang, Z., Chai, Y., Shi, Q., Chen, F., Liu, X., Li, Z., Ru, B., Wang, E., Lei, C., Peng, W., & Huang, Y. 2022. Copy number variation of the CCDC39 gene is associated with growth traits in Chinese cattle. Vet. Med. Sci. https://doi.org/10.1002/vms3.712.

Huan, Y., Kong, Q., Mou, H., & Yi, H. 2020. Antimicrobial Peptides: Classification, Design, Application and Research Progress in Multiple Fields. Front. Microbiol. 11, 1–21 https://doi.org/10.3389/fmicb.2020.582779.

Huang, C., Ge, F., Ren, W., Zhang, Y., Wu, X., Zhang, Q., Ma, X., Bao, P., Guo, X., Chu, M., Yan, P., & Liang, C. 2021a. Copy number variation of the HPGDS gene in the Ashidan yak and its associations with growth traits. Gene 772 https://doi.org/10.1016/j.gene.2020.145382.

Huang, Y., Li, Y., Wang, X., Yu, J., Cai, Y., Zheng, Z., Li, R., Zhang, S., Chen, N., Asadollahpour Nanaei, H., Hanif, Q., Chen, Q., Fu, W., Li, C., Cao, X., Zhou, G., Liu, S., He, S., Li, W., Chen, Y., Chen, H., Lei, C., Liu, M., & Jiang, Y. 2021b. An atlas of CNV maps in cattle, goat and sheep. Sci. China Life Sci. 64, 1747–1764 https://doi.org/10.1007/s11427-020-1850-x.

Huang, J., Li, R., Zhang, X., Huang, Y., Dang, R., Lan, X., Chen, H., & Lei, C. 2017. Copy number veriation regions detection in Qinchuan cattle. Livest. Sci. 204, 88–91 https://doi.org/10.1016/j.livsci.2017.08.016.

Huang, Y.-Z., Shi, Q. T., Shi, S.-Y., Yang, P., Zhang, Z.-J., Lyu, S.-J., Chen, F.-Y., Xu, J.-W., Liu, X., Li, Z., Ru, B., Cai, C., Xie, J., Lei, C., Chen, H., Xu, Z., & Wang, E. 2022. Association between copy number variation of SERPINA3-1 gene and growth traits in Chinese cattle. Anim. Biotechnol. https://doi.org/10.1080/10495398.2022.2038183.

Huang, Y. Z., Zhan, Z. Y., Li, X. Y., Wu, S. R., Sun, Y. J., Xue, J., Lan, X. Y., Lei, C. Z., Zhang, C. L., Jia, Y. T., & Chen, H. 2014. SNP and haplotype analysis reveal IGF2 variants associated with growth traits in Chinese Qinchuan cattle. Mol. Biol. Rep. 41, 591–598 https://doi.org/10.1007/s11033-013-2896-5.

Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Underwood, N., & Vellend, M. 2008. Ecological consequences of genetic diversity. Ecol. Lett. 11, 609–623 https://doi.org/10.1111/j.1461-0248.2008.01179.x.

Hulsman Hanna, L. L., Sanders, J. O., Riley, D. G., Abbey, C. A., & Gill, C. A. 2014. Identification of a major locus interacting with MC1R and modifying black coat color in an F2 Nellore-Angus population. Genet. Sel. Evol. 46, 1–8 https://doi.org/10.1186/1297-9686-46-4.

Hynes, R. O. 1992. Integrins: versatility, modulation, and signaling in cell adhesion. Cell 69, 11–25 https://doi.org/10.1016/0092-8674(92)90115-s.

Iafrate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y., Scherer, S. W., & Lee, C. 2004. Detection of large-scale variation in the human genome. Nat. Genet. 36, 949–951 https://doi.org/10.1038/ng1416.

Ibelli, A. M. G., Ribeiro, A. R. B., Giglioti, R., Regitano, L. C. A., Alencar, M. M., Chagas, A. C. S., Paço, A. L., Oliveira, H. N., Duarte, J. M. S., & Oliveira, M. C. S. 2012. Resistance of cattle of various genetic groups to the tick Rhipicephalus microplus and the relationship with coat traits. Vet. Parasitol. 186, 425–430 https://doi.org/https://doi.org/10.1016/j.vetpar.2011.11.019.

Imamura, T. 2014. Physiological functions and underlying mechanisms of Fibroblast Growth Factor (FGF) family members: Recent findings and implications for their pharmacological application. Biol. Pharm. Bull. 37, 1081–1089 https://doi.org/10.1248/bpb.b14-00265.

Inderwies, T., Pfaffl, M. W., Meyer, H. H. D., Blum, J. W., & Bruckmaier, R. M. 2003. Detection and quantification of mRNA expression of α- and β-adrenergic receptor subtypes in the mammary gland of dairy cows. Domest. Anim. Endocrinol. 24, 123–135 https://doi.org/10.1016/S0739-7240(02)00211-4.

Iqbal, N., Liu, X., Yang, T., Huang, Z., Hanif, Q., Asif, M., Khan, Q. M., & Mansoor, S. 2019. Genomic variants identified from whole-genome resequencing of indicine cattle breeds from Pakistan. PLoS One 14, e0215065.

Iso-Touru, T., Sahana, G., Guldbrandtsen, B., Lund, M. S., & Vilkki, J. 2016. Genome-wide association analysis of milk yield traits in Nordic Red Cattle using imputed whole genome sequence variants. BMC Genet. 17, 1–12 https://doi.org/10.1186/s12863-016-0363-8.

Jakaria, J., Ladhunka Nur Aliyya, W., Ismail, R., Yuni Siswanti, S., Fakhrul Ulum, M., & Priyanto, R. 2021. Discovery of SNPs and indel 11-bp of the myostatin gene and its association with the double-muscled phenotype in Belgian blue crossbred cattle. Gene 784, 145598 https://doi.org/10.1016/j.gene.2021.145598.

Janeway, C., Travers, P., Walport, M., & Shlomchik, M. 2001. The structure of a typical antibody molecule.in Immunobiology: The Immune System in Health and Disease. 5th edition. Garland Science.

Jang, J., Kim, K., Lee, Y. H., & Kim, H. 2021. Population differentiated copy number variation of Bos taurus, Bos indicus and their African hybrids. BMC Genomics 22 https://doi.org/10.1186/s12864-021-07808-7.

Jensen, P., & Wright, D. 2014. Chapter 2 - Behavioral Genetics and Animal Domestication.Pages 41–79 in

Grandin, T., Deesing, M.J.B.T.-G. and the B. of D.A. (Second E., eds. Academic Press, San Diego.
Jiang, L., Jiang, J., Wang, J., Ding, X., Liu, J., & Zhang, Q. 2012. Genome-Wide Identification of Copy Number Variations in Chinese Holstein. PLoS One 7 https://doi.org/10.1371/journal.pone.0048732.
Jiang, L., Jiang, J., Yang, J., Liu, X., Wang, J., Wang, H., Ding, X., Liu, J., & Zhang, Q. 2013. Genome-wide detection of copy number variations using high-density SNP genotyping platforms in Holsteins. BMC Genomics 14 https://doi.org/10.1186/1471-2164-14-131.
Jonsson, N. N., Cox, D. K., Piper, E. K., Valdivieso, E. F. M., Constantinoiu, C., Jackson, L. A., Stear, M. J., Ross, E. M., & Tabor, A. E. 2021. Allelic Variation in Protein Tyrosine Phosphatase Receptor Type-C in Cattle Influences Erythrocyte, Leukocyte and Humoral Responses to Infestation With the Cattle Tick Rhipicephalus australis. Front. Immunol. 12, 1–10 https://doi.org/10.3389/fimmu.2021.675979.
Júnior, G. A. F., Costa, R. B., De Camargo, G. M. F., Carvalheiro, R., Rosa, G. J. M., Baldi, F., Garcia, D. A., Gordo, D. G. M., Espigolan, R., Takada, L., Magalhães, A. F. B., Bresolin, T., Feitosa, F. L. B., Chardulo, L. A. L., De Oliveira, H. N., & De Albuquerque, L. G. 2016. Genome scan for postmortem carcass traits in nellore cattle. J. Anim. Sci. 94, 4087–4095 https://doi.org/10.2527/jas.2016-0632.
Kadri, N. K., Koks, P. D., & Meuwissen, T. H. 2012. Prediction of a deletion copy number variant by a dense SNP panel. Genet. Sel. Evol. 44, 7 https://doi.org/10.1186/1297-9686-44-7.
Kambadur, R., Sharma, M., Smith, T. P. L., & Bass, J. J. 1997. Mutations in myostatin (GDF8) in double-muscled Belgian Blue and Piedmontese cattle. Genome Res. 7, 910–916 https://doi.org/10.1101/gr.7.9.910.
Kaminski, W. E., Piehler, A., & Wenzel, J. J. 2006. ABC A-subfamily transporters: Structure, function and disease. Biochim. Biophys. Acta - Mol. Basis Dis. 1762, 510–524 https://doi.org/10.1016/j.bbadis.2006.01.011.
Kampinga, H. H., Hageman, J., Vos, M. J., Kubota, H., Tanguay, R. M., Bruford, E. A., Cheetham, M. E., Chen, B., & Hightower, L. E. 2009. Guidelines for the nomenclature of the human heat shock proteins. Cell Stress Chaperones 14, 105–111 https://doi.org/10.1007/s12192-008-0068-7.
Karimi, K., Esmailizadeh, A., Wu, D. D., & Gondro, C. 2018. Mapping of genome-wide copy number variations in the Iranian indigenous cattle using a dense SNP data set. Anim. Prod. Sci. 58, 1192–1200 https://doi.org/10.1071/AN16384.
Karrow, N. A., Sharma, B. S., Fisher, R. E., & Mallard, B. A. 2011. 4.33 - Epigenetics and Animal Health.Pages 390–403 in Moo-Young, M.B.T.-C.B. (Third E., ed. Pergamon, Oxford.
Katiyatiya, C. L. F., Muchenje, V., & Mushunje, A. 2014. Farmers' perceptions and knowledge of cattle adaptation to heat stress and tick resistance in the Eastern Cape, South Africa. Asian-Australasian J. Anim. Sci. 27, 1663–1670 https://doi.org/10.5713/ajas.2014.14174.
Kava, R., Peripolli, E., Berton, M. P., Lemos, M., Lobo, R. B., Stafuzza, N. B., Pereira, A. S. C., & Baldi, F. 2021. Genome-wide structural variations in Brazilian Senepol cattle, a tropically adapted taurine breed. Livest. Sci. 253 https://doi.org/10.1016/j.livsci.2021.104708.
Kawano, M., Komi-Kuramochi, A., Asada, M., Suzuki, M., Oki, J., Jiang, J., & Imamura, T. 2005. Comprehensive analysis of FGF and FGFR expression in skin: FGF18 is highly expressed in hair follicles and capable of inducing anagen from telogen stage hair follicles. J. Invest. Dermatol. 124, 877–885 https://doi.org/10.1111/j.0022-202X.2005.23693.x.
Keba, T. M., Dipl-Ing Johann Sölkner, U., Maria Wurzinger, D.-I., Dipl-Ing Alfons Willam, A., Haile, A., Rischkowsky, B., Okeyo, M., & Tibbo, M. 2010. Identifying breeding objectives of smallholders/pastoralists and optimizing community-based breeding programs for adapted sheep breeds in Ethiopia.
Keel, B. N., Keele, J. W., & Snelling, W. M. 2016a. Genome-wide copy number variation in the bovine genome detected using low coverage sequence of popular beef breeds. Anim. Genet. 48, 141–150 https://doi.org/10.1111/age.12519.
Keel, B. N., Lindholm-Perry, A. K., & Snelling, W. M. 2016b. Evolutionary and functional features of copy number variation in the cattle genome. Front. Genet. 7 https://doi.org/10.3389/fgene.2016.00207.
Keele, J. W., Kuehn, L. A., McDaneld, T. G., Tait, R. G., Jones, S. A., Keel, B. N., & Snelling, W. M. 2016. Genomewide association study of liver abscess in beef cattle. J. Anim. Sci. 94, 490–499 https://doi.org/10.2527/jas.2015-9887.
Kennedy, H. E., Welsh, M. D., Bryson, D. G., Cassidy, J. P., Forster, F. I., Howard, C. J., Collins, R. A., & Pollock, J. M. 2002. Modulation of immune responses to Mycobacterium bovis in cattle depleted of WC1+ γδ T cells. Infect. Immun. 70, 1488–1500 https://doi.org/10.1128/IAI.70.3.1488-1500.2002.
Kenny, D. A., Fitzsimons, C., Waters, S. M., & McGee, M. 2018. Invited review: Improving feed efficiency of beef cattle - The current state of the art and future challenges. Animal 12, 1815–1826 https://doi.org/10.1017/S1751731118000976.
Key, N., & Sneeringer, S. 2014. Potential effects of climate change on the productivity of U.S. dairies. Am. J. Agric. Econ. 96, 1136–1156 https://doi.org/10.1093/ajae/aau002.
Khamzina, A. 2016. Copy number variations in Eurasian cattle breeds.
Khan, M. Z., Ma, Y., Ma, J., Xiao, J., Liu, Y., Liu, S., Khan, A., Khan, I. M., & Cao, Z. 2021. Association of

DGAT1 With Cattle, Buffalo, Goat, and Sheep Milk and Meat Production Traits. Front. Vet. Sci. 8, 1–13 https://doi.org/10.3389/fvets.2021.712470.

Khapayi, M., & Celliers, P. R. 2016. FACTORS LIMITING AND PREVENTING EMERGING FARMERS TO PROGRESS TO COMMERCIAL AGRICULTURAL FARMING IN THE KING WILLIAM'S TOWN AREA OF THE EASTERN CAPE PROVINCE, SOUTH AFRICA. 44, 25–41.

Kijas, J. W., Barendse, W., Barris, W., Harrison, B., McCulloch, R., McWilliam, S., & Whan, V. 2011. Analysis of copy number variants in the cattle genome. Gene 482, 73–77 https://doi.org/10.1016/j.gene.2011.04.011.

Kim, J., Hanotte, O., Mwai, O. A., Dessie, T., Salim, B., Diallo, B., Agaba, M., Kim, K., Kwak, W., Sung, S., Seo, M., Jeong, H., Kwon, T., Taye, M., Song, K. D., Lim, D., Cho, S., Lee, H. K. H. J., Yoon, D., Oh, S. J., Kemp, S., Lee, H. K. H. J., & Kim, H. 2017. The genome landscape of indigenous African cattle. Genome Biol. 18, 1–14 https://doi.org/10.1186/s13059-017-1153-y.

Kim, J. B., Zeng, Z. X., Nam, Y. J., Kim, Y., Yang, S. L., Wu, X., & Lee, C. 2005. Association of mahogany/attractin gene (ATRN) with porcine growth and fat. Asian-Australasian J. Anim. Sci. 18, 1383–1386 https://doi.org/10.5713/ajas.2005.1383.

Kimura-Ueki, M., Oda, Y., Oki, J., Komi-Kuramochi, A., Honda, E., Asada, M., Suzuki, M., & Imamura, T. 2012. Hair cycle resting phase is regulated by cyclic epithelial FGF18 signaling. J. Invest. Dermatol. 132, 1338–1345 https://doi.org/10.1038/jid.2011.490.

Kobe, B., & Kajava, A. V. 2001. The leucine-rich repeat as a protein recognition motif. Curr. Opin. Struct. Biol. 11, 725–732 https://doi.org/10.1016/S0959-440X(01)00266-4.

Kommadath, A., Grant, J. R., Krivushin, K., Butty, A. M., Baes, C. F., Carthy, T. R., Berry, D. P., & Stothard, P. 2019. A large interactive visual database of copy number variants discovered in taurine cattle. Gigascience 8 https://doi.org/10.1093/gigascience/giz073.

Koohmaraie, M., & Geesink, G. H. 2006. Contribution of postmortem muscle biochemistry to the delivery of consistent meat quality with particular focus on the calpain system. Meat Sci. 74, 34–43 https://doi.org/10.1016/j.meatsci.2006.04.025.

Korn, J. M., Kuruvilla, F. G., McCarroll, S. A., Wysoker, A., Nemesh, J., Cawley, S., Hubbell, E., Veitch, J., Collins, P. J., Darvishi, K., Lee, C., Nizzari, M. M., Gabriel, S. B., Purcell, S., Daly, M. J., & Altshuler, D. 2008. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat. Genet. 40, 1253–1260 https://doi.org/10.1038/ng.237.

Kościuczuk, E. M., Lisowski, P., Jarczak, J., Strzałkowska, N., Jóźwik, A., Horbańczuk, J., Krzyżewski, J., Zwierzchowski, L., & Bagnicka, E. 2012. Cathelicidins: family of antimicrobial peptides. A review. Mol. Biol. Rep. 39, 10957–10970 https://doi.org/10.1007/s11033-012-1997-x.

Koufariotis, L., Hayes, B. J., Kelly, M., Burns, B. M., Lyons, R., Stothard, P., Chamberlain, A. J., & Moore, S. 2018. Sequencing the mosaic genome of Brahman cattle identifies historic and recent introgression including polled. Sci. Rep. 8, 1–12.

Kowalczyk, A., Czerniawska-pi, E., & Wrzeci, M. 2021. Review Article The Importance of Interferon-Tau in the Diagnosis of Pregnancy. 2021, 1–6.

Kukučková, V., Moravčíková, N., & Kasarda, R. 2016. Evaluating signatures of selection through variation in linkage disequilibrium among different cattle breeds. Acta Fytotech. Zootech. 19.

Kumar, K. R., Cowley, M. J., & Davis, R. L. 2019. Next-Generation Sequencing and Emerging Technologies. Semin. Thromb. Hemost. 45, 661–673 https://doi.org/10.1055/s-0039-1688446.

Kumar, H., Panigrahi, M., Saravanan, K. A., Rajawat, D., Parida, S., Bhushan, B., Gaur, G. K., Dutt, T., Mishra, B. P., & Singh, R. K. 2021. Genome-wide detection of copy number variations in Tharparkar cattle. Anim. Biotechnol., 1–8.

Küttel, L., Letko, A., Häfliger, I. M., Signer-Hasler, H., Joller, S., Hirsbrunner, G., Mészáros, G., Sölkner, J., Flury, C., Leeb, T., & Drögemüller, C. 2019. A complex structural variant at the KIT locus in cattle with the Pinzgauer spotting pattern. Anim. Genet. 50, 423–429 https://doi.org/10.1111/age.12821.

Lacroix-Pépin, N., Danyod, G., Krishnaswamy, N., Mondal, S., Rong, P. M., Chapdelaine, P., & Fortier, M. A. 2011. The Multidrug Resistance-Associated Protein 4 (MRP4) appears as a functional carrier of prostaglandins regulated by oxytocin in the bovine endometrium. Endocrinology 152, 4993–5004 https://doi.org/10.1210/en.2011-1406.

Lai, C.-Q. 2012. Adaptive Genetic Variation and Population Differences.Pages 461–489 in Recent Advances in Nutrigenetics and Nutrigenomics. Bouchard, C., Ordovas, J.M.B.T.-P. in M.B. and T.S., eds. Academic Press.

Lai, W. R., Johnson, M. D., Kucherlapati, R., & Park, P. J. 2005. Comparative analysis of algorithms for identifying amplifications and deletions in array CGH data. Bioinformatics 21, 3763–3770 https://doi.org/10.1093/bioinformatics/bti611.

Laity, J. H., Lee, B. M., & Wright, P. E. 2001. Zinc finger proteins: new insights into structural and functional diversity. Curr. Opin. Struct. Biol. 11, 39–46 https://doi.org/https://doi.org/10.1016/S0959-440X(00)00167-6.

Lamb, H. J., Ross, E. M., Nguyen, L. T., Lyons, R. E., Moore, S. S., & Hayes, B. J. 2020. Characterization of

the poll allele in Brahman cattle using long-read Oxford Nanopore sequencing. J. Anim. Sci. 98 https://doi.org/10.1093/jas/skaa127.

Lammers, B. P., Heinrichs, A. J., & Kensinger, R. S. 1999. The effects of accelerated growth rates and estrogen implants in prepubertal Holstein heifers on growth, feed efficiency, and blood parameters. J. Dairy Sci. 82, 1746–1752 https://doi.org/10.3168/jds.S0022-0302(99)75405-6.

Lappas, M., & Rice, G. E. 2004. Phospholipase A2 isozymes in pregnancy and parturition. Prostaglandins Leukot. Essent. Fat. Acids 70, 87–100 https://doi.org/10.1016/j.plefa.2003.04.001.

Larson, J. H., Marron, B. M., Beever, J. E., Roe, B. A., & Lewin, H. A. 2006. Genomic organization and evolution of the ULBP genes in cattle. BMC Genomics 7, 1–14 https://doi.org/10.1186/1471-2164-7-227.

Larson, J. H., Rebeiz, M. J., Stiening, C. M., Windish, R. L., Beever, J. E., & Lewin, H. A. 2003. MHC class I-like genes in cattle, MHCLA, with similarity to genes encoding NK cell stimulatory ligands. Immunogenetics 55, 16–22 https://doi.org/10.1007/s00251-003-0551-3.

Lashmar, S. F., Muchadeyi, F. C., & Visser, C. 2019. Genotype imputation as a cost-saving genomic strategy for South African Sanga cattle: A review. S. Afr. J. Anim. Sci. 49, 262–280.

Lasserson, T., Thomas, J., & Higgins, J. 2022. Chapter 1: Starting a review.in Cochrane Handbook for Systematic Reviews of Interventions version 6.3. Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., Welch, V., eds. Cochrane.

Lauer, S., & Gresham, D. 2019. An evolving view of copy number variants. Curr. Genet. 65, 1287–1295.

Leal-Gutiérrez, J. D., Elzo, M. A., Johnson, D. D., Scheffler, T. L., Scheffler, J. M., & Mateescu, R. G. 2018. Association of μ-calpain and calpastatin polymorphisms with meat tenderness in a Brahman-Angus population. Front. Genet. 9, 1–10 https://doi.org/10.3389/fgene.2018.00056.

Lee, Y.-L., Bosse, M., Mullaart, E., Groenen, M. A. M., Veerkamp, R. F., & Bouwman, A. C. 2020. Functional and population genetic features of copy number variations in two dairy cattle populations. BMC Genomics 21 https://doi.org/10.1186/s12864-020-6496-1.

Lee, T., Cho, S., Seo, K. S., Chang, J., Kim, H., & Yoon, D. 2013a. Genetic variants and signatures of selective sweep of Hanwoo population (Korean native cattle). BMB Rep. 46, 346.

Lee, K.-T., Chung, W.-H., Lee, S.-Y., Choi, J.-W., Kim, J., Lim, D., Lee, S., Jang, G.-W., Kim, B., & Choy, Y. H. 2013b. Whole-genome resequencing of Hanwoo (Korean cattle) and insight into regions of homozygosity. BMC Genomics 14, 1–14.

Lee, H. J., Jin, S., Kim, H. J., Bhuiyan, M. S. A., Lee, D. H., Lee, S. H., Jang, S. B., Han, M. H., & Lee, S. H. 2019. Validation study of SNPs in CAPN1-CAST genes on the tenderness of muscles (Longissimus thoracis and semimembranosus) in Hanwoo (Korean cattle). Animals 9, 1–8 https://doi.org/10.3390/ani9090691.

Lee, K., Nguyen, D. T., Choi, M., Cha, S. Y., Kim, J. H., Dadi, H., Seo, H. G., Seo, K., Chun, T., & Park, C. 2013c. Analysis of cattle olfactory subgenome: The first detail study on the characteristics of the complete olfactory receptor repertoire of a ruminant. BMC Genomics 14, 1 https://doi.org/10.1186/1471-2164-14-596.

Lee, Y.-L., Takeda, H., Moreira, G. C. M., Karim, L., Mullaart, E., Coppieters, W., Appeltant, R., Veerkamp, R. F., Groenen, M. A. M., Georges, M., Bosse, M., Druet, T., Bouwman, A. C., & Charlier, C. 2021. A 12 kb multi-allelic copy number variation encompassing a GC gene enhancer is associated with mastitis resistance in dairy cattle. PLoS Genet. 17 https://doi.org/10.1371/journal.pgen.1009331.

Letaief, R., Rebours, E., Grohs, C., Meersseman, C., Fritz, S., Trouilh, L., Esquerré, D., Barbieri, J., Klopp, C., Philippe, R., Blanquet, V., Boichard, D., Rocha, D., & Boussaha, M. 2017. Identification of copy number variation in French dairy and beef breeds using next-generation sequencing. Genet. Sel. Evol. 49, 77 https://doi.org/10.1186/s12711-017-0352-z.

Lewin, H. A., Schmitt, K., Hubert, R., van Eijk, M. J. T., & Arnheim, N. 1992. Close linkage between bovine prolactin and BoLA-DRB3 genes: Genetic mapping in cattle by single sperm typing. Genomics 13, 44–48 https://doi.org/10.1016/0888-7543(92)90200-C.

Li, J., Bauer, M., Moe, B., Leslie, E. M., & Li, X. 2017. Multidrug Resistance Protein 4 (MRP4/ABCC4) Protects Cells from the Toxic E ff ects of Halobenzoquinones. 4 https://doi.org/10.1021/acs.chemrestox.7b00156.

Li, W., Bickhart, D. M., Ramunno, L., Iamartino, D., Williams, J. L., & Liu, G. E. 2018. Genomic structural differences between cattle and River Buffalo identified through comparative genomic and transcriptomic analysis. Data Br. 19, 236–239 https://doi.org/10.1016/j.dib.2018.05.015.

Li, W., Bickhart, D. M., Ramunno, L., Iamartino, D., Williams, J. L., & Liu, G. E. 2019a. Comparative sequence alignment reveals River Buffalo genomic structural differences compared with cattle. Genomics 111, 418–425 https://doi.org/10.1016/j.ygeno.2018.02.018.

Li, C., Cai, W., Zhou, C., Yin, H., Zhang, Z., Loor, J. J., Sun, D., Zhang, Q., Liu, J., & Zhang, S. 2016. RNA-Seq reveals 10 novel promising candidate genes affecting milk protein concentration in the Chinese Holstein population. Sci. Rep. 6, 1–11 https://doi.org/10.1038/srep26813.

Li, X., Ding, X., Liu, L., Yang, P., Yao, Z., Lei, C., Chen, H., Huang, Y., & Liu, W. 2022a. Copy number variation of bovine DYNC1I2 gene is associated with body conformation traits in chinese beef cattle. Gene 810 https://doi.org/10.1016/j.gene.2021.146060.

Li, R. W., & Gasbarre, L. C. 2009. A temporal shift in regulatory networks and pathways in the bovine small intestine during Cooperia oncophora infection. Int. J. Parasitol. 39, 813–824 https://doi.org/10.1016/j.ijpara.2008.11.007.

Li, X., Han, M., Zhang, H., Liu, F., Pan, Y., Zhu, J., Liao, Z., Chen, X., & Zhang, B. 2022b. Structures and biological functions of zinc finger proteins and their roles in hepatocellular carcinoma. Biomark. Res. 10, 2 https://doi.org/10.1186/s40364-021-00345-1.

Li, T., Higgins, J. P. T., & Deeks, J. J. 2019b. Collecting data.Pages 109–141 in Cochrane Handbook for Systematic Reviews of Interventions.

Liang, J., Liu, X., Yang, P., Yao, Z., Qu, K., Wang, H., Zhang, Z., Liang, H., Cheng, B., Li, Z., Ru, B., Zhang, J., Qi, Z., Wang, E., Lei, C., Chen, H., Huang, B., & Huang, Y. 2022. Copy number variation of GAL3ST1 gene is associated with growth traits of Chinese cattle. Anim. Biotechnol. https://doi.org/10.1080/10495398.2021.1996385.

Liao, X., Peng, F., Forni, S., McLaren, D., Plastow, G., & Stothard, P. 2013. Whole genome sequencing of Gir cattle for identifying polymorphisms and loci under selection. Genome 56, 592–598 https://doi.org/10.1139/gen-2013-0082.

Lim, D. H., Mayakrishnan, V., Lee, H. J., Ki, K. S., Kim, T. Il, & Kim, Y. 2020. A comparative study on milk composition of Jersey and Holstein dairy cows during the early lactation. J. Anim. Sci. Technol. 62, 565–576 https://doi.org/10.5187/JAST.2020.62.4.565.

Lima, L. S., Galiciolli, M. E. A., Pereira, M. E., Felisbino, K., Machado-Souza, C., de Oliveira, C. S., & Guiloski, I. C. 2022. Modification by genetic polymorphism of lead-induced IQ alteration: a systematic review. Environ. Sci. Pollut. Res. 29, 43435–43447 https://doi.org/10.1007/s11356-022-19981-7.

Littlejohn, M. D., Henty, K. M., Tiplady, K., Johnson, T., Harland, C., Lopdell, T., Sherlock, R. G., Li, W., Lukefahr, S. D., Shanks, B. C., Garrick, D. J., Snell, R. G., Spelman, R. J., & Davis, S. R. 2014. Functionally reciprocal mutations of the prolactin signalling pathway define hairy and slick cattle. Nat. Commun. 5, 1–8 https://doi.org/10.1038/ncomms6861.

Liu, Y., Albrecht, E., Schering, L., Kuehn, C., Yang, R., Zhao, Z., & Maak, S. 2018. Agouti signaling protein and its receptors as potential molecular markers for intramuscular and body fat deposition in cattle. Front. Physiol. 9, 1–13 https://doi.org/10.3389/fphys.2018.00172.

Liu, G. E., & Bickhart, D. M. 2012. Copy number variation in the cattle genome. Funct. Integr. Genomics 12, 609–624 https://doi.org/10.1007/s10142-012-0289-9.

Liu, G. E., Brown, T., Hebert, D. A., Cardone, M. F., Hou, Y., Choudhary, R. K., Shaffer, J., Amazu, C., Connor, E. E., Ventura, M., & Gasbarre, L. C. 2011. Initial analysis of copy number variations in cattle selected for resistance or susceptibility to intestinal nematodes. Mamm. Genome 22, 111–121 https://doi.org/10.1007/s00335-010-9308-0.

Liu, D., Chen, Z., Zhang, Z., Sun, H., Ma, P., Zhu, K., Liu, G., Wang, Q., & Pan, Y. 2019a. Detection of genome-wide structural variations in the Shanghai Holstein cattle population using next-generation sequencing. Asian-Australasian J. Anim. Sci. 32, 320–333 https://doi.org/10.5713/ajas.18.0204.

Liu, Z. J., & Cordes, J. F. 2004. DNA marker technologies and their applications in aquaculture genetics. Aquaculture 238, 1–37 https://doi.org/10.1016/j.aquaculture.2004.05.027.

Liu, M., Fang, L., Liu, S., Pan, M. G., Seroussi, E., Cole, J. B., Ma, L., Chen, H., & Liu, G. E. 2019b. Array CGH-based detection of CNV regions and their potential association with reproduction and other economic traits in Holsteins. BMC Genomics 20 https://doi.org/10.1186/s12864-019-5552-1.

Liu, G. E., Hou, Y., Zhu, B., Cardone, M. F., Jiang, L., Cellamare, A., Mitra, A., Alexander, L. J., Coutinho, L. L., Dell'Aquila, M. E., Gasbarre, L. C., Lacalandra, G., Li, R. W., Matukumalli, L. K., Nonneman, D., Regitano, L. C. D. A. L. C. D. A., Smith, T. P. L. L., Song, J., Sonstegard, T. S., Van Tassell, C. P., Ventura, M., Eichler, E. E., McDaneld, T. G., & Keele, J. W. 2010. Analysis of copy number variations among diverse cattle breeds. Genome Res. 20, 693–703 https://doi.org/10.1101/gr.105403.110.

Liu, Z., Jaitner, J., Reinhardt, F., Pasman, E., Rensing, S., & Reents, R. 2008a. Genetic evaluation of fertility traits of dairy cattle using a multiple-trait animal model. J. Dairy Sci. 91, 4333–4343 https://doi.org/10.3168/jds.2008-1029.

Liu, Z., Lavine, K. J., Hung, I. H., & Ornitz, D. M. 2007. FGF18 is required for early chondrocyte proliferation, hypertrophy and vascular invasion of the growth plate. Dev. Biol. 302, 80–91 https://doi.org/10.1016/j.ydbio.2006.08.071.

Liu, M., Li, B., Huang, Y., Yang, M., Lan, X., Lei, C., Qu, W., Bai, Y., & Chen, H. 2016. Copy number variation of bovine MAPK10 modulates the transcriptional activity and affects growth traits. Livest. Sci. 194, 44–50 https://doi.org/10.1016/j.livsci.2016.09.014.

Liu, M., Li, B., Shi, T., Huang, Y., Liu, G. E., Lan, X., Lei, C., & Chen, H. 2019c. Copy number variation of bovine SHH gene is associated with body conformation traits in Chinese beef cattle. J. Appl. Genet. 60, 199–207 https://doi.org/10.1007/s13353-019-00496-w.

Liu, A., Lund, M. S., Wang, Y., Guo, G., Dong, G., Madsen, P., & Su, G. 2017. Variance components and correlations of female fertility traits in Chinese Holstein population. J. Anim. Sci. Biotechnol. 8, 1–9 https://doi.org/10.1186/s40104-017-0189-x.

Liu, Y., Qin, X., Song, X. Z. H., Jiang, H., Shen, Y., Durbin, K. J., Lien, S., Kent, M. P., Sodeland, M., Ren, Y., Zhang, L., Sodergren, E., Havlak, P., Worley, K. C., Weinstock, G. M., & Gibbs, R. A. 2009. Bos taurus genome assembly. BMC Genomics 10, 1–11 https://doi.org/10.1186/1471-2164-10-180.

Liu, G. E., Van Tassell, C. P., Sonstegard, T. S., Li, R. W., Alexander, L. J., Keele, J. W., Matukumalli, L. K., Smith, T. P., & Gasbarre, L. C. 2008b. Detection of germline and somatic copy number variations in cattle.Pages 231–237 in Developments in Biologicals.

Liu, M., Woodward-Greene, J., Kang, X., Pan, M. G., Rosen, B., Van Tassell, C. P., Chen, H., & Liu, G. E. 2020. Genome-wide CNV analysis revealed variants associated with growth traits in African indigenous goats. Genomics 112, 1477–1480 https://doi.org/10.1016/j.ygeno.2019.08.018.

Liu, X., Yang, P., Sun, H., Zhang, Z., Cai, C., Xu, J., Ding, X., Wang, X., Lyu, S., Li, Z., Xu, Z., Shi, Q., Wang, E., Lei, C., Chen, H., Ru, B., & Huang, Y. 2022. CNV analysis of VAMP7 gene reveals variation associated with growth traits in Chinese cattle. Anim. Biotechnol. https://doi.org/10.1080/10495398.2021.2011741.

Longeri, M., Russo, V., Strillacci, M. G., Perillo, A., Carisetti, M., Cozzi, M. C., Neola, B., & Roperto, S. 2021. Association between bola-drb3.2 polymorphism and bovine papillomavirus infection for bladder tumor risk in podolica cattle. Front. Vet. Sci. 8 https://doi.org/10.3389/fvets.2021.630089.

Loonie, K., Gillespie, A., & Baldwin, C. 2021. The WC1 γδ T cell pathogen receptor of ruminants is preserved in the genome of ancient extinct auroch. Immunogenetics 73, 419–424 https://doi.org/10.1007/s00251-021-01211-y.

Low, W. Y., Tearle, R., Liu, R., Koren, S., Rhie, A., Bickhart, D. M., Rosen, B. D., Kronenberg, Z. N., Kingan, S. B., Tseng, E., Thibaud-Nissen, F., Martin, F. J., Billis, K., Ghurye, J., Hastie, A. R., Lee, J., Pang, A. W. C., Heaton, M. P., Phillippy, A. M., Hiendleder, S., Smith, T. P. L., & Williams, J. L. 2020. Haplotype-resolved genomes provide insights into structural variation and gene content in Angus and Brahman cattle. Nat. Commun. 11 https://doi.org/10.1038/s41467-020-15848-y.

Lupski, J. R., & Stankiewicz, P. 2005. Genomic disorders: Molecular mechanisms for rearrangements and conveyed phenotypes. PLoS Genet. 1, 0627–0633 https://doi.org/10.1371/journal.pgen.0010049.

Luu, P. L., Ong, P. T., Dinh, T. P., & Clark, S. J. 2020. Benchmark study comparing liftover tools for genome conversion of epigenome sequencing data. NAR Genomics Bioinforma. 2, 1–18 https://doi.org/10.1093/nargab/lqaa054.

Ma, Y.-L., Wen, Y.-F., Cao, X.-K., Cheng, J., Huang, Y.-Z., Ma, Y., Hu, L.-Y., Lei, C.-Z., Qi, X.-L., Cao, H., & Chen, H. 2019. Copy number variation (CNV) in the IGF1R gene across four cattle breeds and its association with economic traits. Arch. Anim. Breed. 62, 171–179 https://doi.org/10.5194/aab-62-171-2019.

Magee, D. A., MacHugh, D. E., & Edwards, C. J. 2014. Interrogation of modern and ancient genomes reveals the complex domestic history of cattle. Anim. Front. 4, 7–22 https://doi.org/10.2527/af.2014-0017.

Maiorano, A. M., Cardoso, D. F., Carvalheiro, R., Júnior, G. A. F., de Albuquerque, L. G., & de Oliveira, H. N. 2022. Signatures of selection in Nelore cattle revealed by whole-genome sequencing data. Genomics 114, 110304 https://doi.org/https://doi.org/10.1016/j.ygeno.2022.110304.

Makova, K., & Norton, H. 2005. Worldwide polymorphism at the MC1R locus and normal pigmentation variation in humans. Peptides 26, 1901–1908 https://doi.org/10.1016/j.peptides.2004.12.032.

Malnic, B., Hirono, J., Sato, T., & Buck, L. B. 1999. Combinatorial receptor codes for odors. Cell 96, 713–723 https://doi.org/10.1016/S0092-8674(00)80581-4.

Mapholi, N. O., Maiwashe, A. A., Matika, O., Riggio, V., Bishop, S. C., MacNeil, M. D., Banga, C. B., Taylor, J. F., & Dzama, K. 2016. Genome-wide association study of tick resistance in South African Nguni cattle. J. Anim. Sci. 94, 72–72 https://doi.org/10.2527/jas2016.94supplement472x.

Mapholi, N. O., Marufu, M. C., Maiwashe, A., Banga, C. B., Muchenje, V., MacNeil, M. D., Chimonyo, M., & Dzama, K. 2014. Towards a genomics approach to tick (Acari: Ixodidae) control in cattle: A review. Ticks Tick. Borne. Dis. 5, 475–483 https://doi.org/https://doi.org/10.1016/j.ttbdis.2014.04.006.

Mapiye, C., Chimonyo, M., Dzama, K., Raats, J. G., & Mapekula, M. 2009. Opportunities for improving Nguni cattle production in the smallholder farming systems of South Africa. Livest. Sci. 124, 196–204 https://doi.org/10.1016/j.livsci.2009.01.013.

van Marle-Köster, E., & Visser, C. 2018. Genetic improvement in South African livestock: Can genomics bridge the gap between the developed and developing sectors? Front. Genet. 9, 1–12 https://doi.org/10.3389/fgene.2018.00331.

Marufu, M. C. 2012. Mechanisms of resistance to Rhipicephalus ticks in Nguni cattle reared in the semiarid areas of South Africa.

Matukumalli, L. K., Lawley, C. T., Schnabel, R. D., Taylor, J. F., Allan, M. F., Heaton, M. P., O'Connell, J., Moore, S. S., Smith, T. P. L., Sonstegard, T. S., & Van Tassell, C. P. 2009. Development and Characterization of a High Density SNP Genotyping Assay for Cattle. PLoS One 4 https://doi.org/10.1371/journal.pone.0005350.

Matukumalli, L. K., Schroeder, S., DeNise, S. K., Sonstegard, T., Lawley, C. T., Georges, M., Coppieters, W., Gietzen, K., Medrano, J. F., & Rincon, G. 2011. Analyzing LD blocks and CNV segments in cattle: novel

genomic features identified using the BovineHD BeadChip. Illumina Inc. San Diego, CA.
McDaneld, T. G., Kuehn, L. A., Thomas, M. G., Pollak, E. J., & Keele, J. W. 2014. Deletion on chromosome 5 associated with decreased reproductive efficiency in female cattle. J. Anim. Sci. 92, 1378–1384 https://doi.org/10.2527/jas.2013-6821.
Mckenzie, J., Brennan, S., Ryan, R., Thomson, H., Johnston, R., & Thomas, J. 2019. Chapter 3: Defining the criteria for including studies and how they will be grouped for the synthesis.in Cochrane Handbook for Systematic Reviews of Interventions.
McPherron, A. C., & Lee, S. J. 1997. Double muscling in cattle due to mutations in the myostatin gene. Proc. Natl. Acad. Sci. U. S. A. 94, 12457–12461 https://doi.org/10.1073/pnas.94.23.12457.
Meade, K. G., Cormican, P., Narciandi, F., Lloyd, A., & O'Farrelly, C. 2014. Bovine β-defensin gene family: Opportunities to improve animal health? Physiol. Genomics 46, 17–28 https://doi.org/10.1152/physiolgenomics.00085.2013.
Mears, G. J., Mir, P. S., Bailey, D. R. C., & Jones, S. D. M. 2000. Effect of Wagyu genetics on marbling, backfat and circulating hormones in cattle. Can. J. Anim. Sci. 81, 65–73 https://doi.org/10.4141/A99-128.
Mei, C., Gui, L., Hong, J., Raza, S. H. A., Aorigele, C., Tian, W., Garcia, M., Xin, Y., Yang, W., Zhang, S., & Zan, L. 2021. Insights into adaption and growth evolution: a comparative genomics study on two distinct cattle breeds from Northern and Southern China. Mol. Ther. - Nucleic Acids 23, 959–967 https://doi.org/10.1016/j.omtn.2020.12.028.
Mei, C., Junjvlieke, Z., Raza, S. H. A., Wang, H., Cheng, G., Zhao, C., Zhu, W., & Zan, L. 2020. Copy number variation detection in Chinese indigenous cattle by whole genome sequencing. Genomics 112, 831–836 https://doi.org/10.1016/j.ygeno.2019.05.023.
Mei, C., Wang, H., Liao, Q., Khan, R., Raza, S. H. A., Zhao, C., Wang, H., Cheng, G., Tian, W., Li, Y., & Zan, L. 2019. Genome-wide analysis reveals the effects of artificial selection on production and meat quality traits in Qinchuan cattle. Genomics 111, 1201–1208 https://doi.org/10.1016/j.ygeno.2018.09.021.
Menzies, K. K., Lefèvre, C., Macmillan, K. L., & Nicholas, K. R. 2009. Insulin regulates milk protein synthesis at multiple levels in the bovine mammary gland. , 197–217 https://doi.org/10.1007/s10142-008-0103-x.
Meuwissen, T. H. E. 2009. Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. Genet. Sel. Evol. 41, 35 https://doi.org/10.1186/1297-9686-41-35.
Mielczarek, M., Frąszczak, M., Giannico, R., Minozzi, G., Williams, J. L., Wojdak-Maksymiec, K., & Szyda, J. 2017. Analysis of copy number variations in Holstein-Friesian cow genomes based on whole-genome sequence data. J. Dairy Sci. 100, 5515–5525 https://doi.org/10.3168/jds.2016-11987.
Mielczarek, M., Fraszczak, M., Nicolazzi, E., Williams, J. L., & Szyda, J. 2018. Landscape of copy number variations in Bos taurus: Individual - and inter-breed variability. BMC Genomics 19 https://doi.org/10.1186/s12864-018-4815-6.
Minozzi, G., Nicolazzi, E. L., Stella, A., Biffani, S., Negrini, R., Lazzari, B., Ajmone-Marsan, P., & Williams, J. L. 2013. Genome wide analysis of fertility and production traits in Italian Holstein cattle. PLoS One 8, 1–10 https://doi.org/10.1371/journal.pone.0080219.
Miretti, S., Martignani, E., Accornero, P., & Baratta, M. 2013. Functional effect of mir-27b on myostatin expression: A relationship in piedmontese cattle with double-muscled phenotype. BMC Genomics 14, 1–8 https://doi.org/10.1186/1471-2164-14-194.
Miyasaka, T., Takeshima, S. N., Sentsui, H., & Aida, Y. 2012. Identification and diversity of bovine major histocompatibility complex class II haplotypes in Japanese Black and Holstein cattle in Japan. J. Dairy Sci. 95, 420–431 https://doi.org/10.3168/jds.2011-4621.
Mkize, L. S., & Zishiri, O. T. 2020. Novel single nucleotide polymorphisms in the heat shock protein 70.1 gene in South African Nguni crossbred cattle. Trop. Anim. Health Prod. 52, 893–901 https://doi.org/10.1007/s11250-019-02088-6.
Mohammed, S., Mohammed, S. A., Rahamtalla, S. A., Ahmed, S. S., elamin, khalid mohammed, Dousa, B. M., mmm, E. A., & Ahmed, mohammed-khair A. 2014. DGAT1 Gene in Dairy Cattle: A Review. Glob. J. Anim. Sci. Res. 3, 191–198.
Mohanty, T. R., Seo, K. S., Park, K. M., Choi, T. J., Choe, H. S., Baik, D. H., & Hwang, I. H. 2008. Molecular variation in pigmentation genes contributing to coat colour in native Korean Hanwoo cattle. Anim. Genet. 39, 550–553 https://doi.org/10.1111/j.1365-2052.2008.01746.x.
Mokry, F. B., Higa, R. H., de Alvarenga Mudadu, M., Oliveira de Lima, A., Meirelles, S. L. C., Barbosa da Silva, M. V. G., Cardoso, F. F., Morgado de Oliveira, M., Urbinati, I., Méo Niciura, S. C., Tullio, R. R., Mello de Alencar, M., & Correia de Almeida Regitano, L. 2013. Genome-wide association study for backfat thickness in Canchim beef cattle using Random Forest approach. BMC Genet. 14, 1–11 https://doi.org/10.1186/1471-2156-14-47.
Moravčíková, N., Kasarda, R., Vostrý, L., Krupová, Z., Krupa, E., Lehocká, K., Olšanská, B., Trakovická, A., Nádaský, R., Židek, R., Belej, Ľ., & Golian, J. 2019. Analysis of selection signatures in the beef cattle genome. Czech J. Anim. Sci. 64, 491–503 https://doi.org/10.17221/226/2019-CJAS.
Morton, J. F. 2007. The impact of climate change on smallholder and subsistence agriculture.
Mpetile, Z., Dzama, K., & Cloete, S. W. P. 2017. Genetic variation in and relationships among faecal worm

eggs recorded in different seasons of the year at the Tygerhoek farm in South Africa. J. S. Afr. Vet. Assoc. 88, 1–7 https://doi.org/10.4102/jsava.v88i0.1484.

Muchenje, V., Dzama, K., Chimonyo, M., Raats, J. G., & Strydom, P. E. 2008. Meat quality of Nguni, Bonsmara and Aberdeen Angus steers raised on natural pasture in the Eastern Cape, South Africa. Meat Sci. 79, 20–28 https://doi.org/10.1016/j.meatsci.2007.07.026.

Mukherjee, A., Dass, G., G, J. M., Gohain, M., Brahma, B., Datta, T. K., & De, S. 2013. Absolute copy number differences of Y chromosomal genes between crossbred (Bos taurus × Bos indicus) and Indicine bulls. J. Anim. Sci. Biotechnol. 4 https://doi.org/10.1186/2049-1891-4-15.

Muller, C. J. C. 2017. The Dairy Farming Handbook.

Muller, C. J. C., Cloete, S. W. P., & Botha, J. A. 2018. Fertility in dairy cows and ways to improve it. S. Afr. J. Anim. Sci. 48, 858 https://doi.org/10.4314/sajas.v48i5.6.

Muniz, M. M. M., Fonseca, L. F. S., dos Santos Silva, D. B., de Oliveira, H. R., Baldi, F., Chardulo, A. L., Ferro, J. A., Cánovas, A., & de Albuquerque, L. G. 2021. Identification of novel mRNA isoforms associated with meat tenderness using RNA sequencing data in beef cattle. Meat Sci. 173, 108378 https://doi.org/https://doi.org/10.1016/j.meatsci.2020.108378.

Murakami, M., Sato, H., Miki, Y., Yamamoto, K., & Taketomi, Y. 2015. A new era of secreted phospholipase A2. J. Lipid Res. 56, 1248–1261 https://doi.org/10.1194/jlr.R058123.

Murray, D. G. 2013. Tableau your data!: fast and easy visual analysis with tableau software. Wiley.

Murungweni, C., van Wijk, M. T., Giller, K. E., Andersson, J. A., & Smaling, E. M. A. 2014. Adaptive livelihood strategies employed by farmers to close the food gap in semi-arid south eastern Zimbabwe. Food Secur. 6, 313–326 https://doi.org/10.1007/s12571-014-0348-2.

Musemwa, L., Mushunje, A., Chimonyo, M., Fraser, G., Mapiye, C., & Muchenje, V. 2008. Nguni cattle marketing constraints and opportunities in the communal areas of South Africa: Review.

Mustafa, H., Khan, W. A., Euisoo, K., Kuthu, Z. H., Ajmal, A., Ali, A., Javed, K., & Sonstegard, T. S. 2018. Genetic properties of copy number variations in some Pakistani cattle breeds. Kafkas Univ. Vet. Fak. Derg. 24, 599–602 https://doi.org/10.9775/kvfd.2018.19282.

Mwai, O., Hanotte, O., Kwon, Y. J., & Cho, S. 2015. - Invited review - African indigenous cattle: Unique genetic resources in a rapidly changing world. Asian-Australasian J. Anim. Sci. 28, 911–921 https://doi.org/10.5713/ajas.15.0002R.

Mwanza, M., Dlamini, Z. M., & Ramaili, T. 2016. Prevalent Bovine Internal Parasites Using Different Techniques in Communal Farms around Mafikeng. J. Hum. Ecol. 56, 160–165 https://doi.org/10.1080/09709274.2016.11907051.

Nafikov, R. A., Schoonmaker, J. P., Korn, K. T., Noack, K., Garrick, D. J., Koehler, K. J., Minick-Bormann, J., Reecy, J. M., Spurlock, D. E., & Beitz, D. C. 2013. Association of polymorphisms in solute carrier family 27, isoform A6 (SLC27A6) and fatty acid-binding protein-3 and fatty acid-binding protein-4 (FABP3 and FABP4) with fatty acid composition of bovine milk. J. Dairy Sci. 96, 6007–6021 https://doi.org/10.3168/jds.2013-6703.

Nagle, D. L., McGrail, S. H., Vitale, J., Woolf, E. A., Dussault, B. J., DiRocco, L., Holmgren, L., Montagno, J., Bork, P., Huszar, D., Fairchild-Huntress, V., Ge, P., Keilty, J., Ebeling, C., Baldini, L., Gilchrist, J., Burn, P., Carlson, G. A., & Moore, K. J. 1999. The mahogany protein is a receptor involved in suppression of obesity. Nature 398, 148–152 https://doi.org/10.1038/18210.

Nakatochi, M., Kushima, I., & Ozaki, N. 2021. Implications of germline copy-number variations in psychiatric disorders: review of large-scale genetic studies. J. Hum. Genet. 66, 25–37 https://doi.org/10.1038/s10038-020-00838-1.

Nakatsu, F., & Ohno, H. 2003. Adaptor protein complexes as the key regulators of protein sorting in the post-Golgi network. Cell Struct. Funct. 28, 419–429 https://doi.org/10.1247/csf.28.419.

Nanaei, H. A., Qanatqestani, M. D., & Esmailizadeh, A. 2020. Whole-genome resequencing reveals selection signatures associated with milk production traits in African Kenana dairy zebu cattle. Genomics 112, 880–885.

Nandolo, W., Utsunomiya, Y. T., Mészáros, G., Wurzinger, M., Khayadzadeh, N., Torrecilha, R. B. P., Mulindwa, H. A., Gondwe, T. N., Waldmann, P., Ferenčaković, M., Garcia, J. F., Rosen, B. D., Bickhart, D., Van Tassell, C. P., Curik, I., & Sölkner, J. 2018. Misidentification of runs of homozygosity islands in cattle caused by interference with copy number variation or large intermarker distances. Genet. Sel. Evol. 50 https://doi.org/10.1186/s12711-018-0414-x.

Naval-Sánchez, M., Porto-Neto, L. R., Cardoso, D. F., Hayes, B. J., Daetwyler, H. D., Kijas, J., Reverter, A., Naval-Sanchez, M., Porto-Neto, L. R., Cardoso, D. F., Hayes, B. J., Daetwyler, H. D., Kijas, J., & Reverter, A. 2020. Selection signatures in tropical cattle are enriched for promoter and coding regions and reveal missense mutations in the damage response gene HELB. Genet. Sel. Evol. 52 https://doi.org/10.1186/s12711-020-00546-6 WE - Science Citation Index Expanded (SCI-EXPANDED).

NCBI (2022) *Gene*. Available at: https://www.ncbi.nlm.nih.gov/gene (Accessed: 11 October 2022).

Nel, C. L. 2017. The comparison of immunological responses of Brahman, Nguni and Angus cattle infested with Rhipicephalus microplus and Rhipicephalus decoloratus.

Niu, F., Wang, L., Liu, X., Wang, H., Yang, J., Liu, Y., & Chen, L. 2013. Genetic diversity of MYH3 gene associated with growth and carcass traits in Chinese Qinchuan cattle. Mol. Biol. Rep. 40, 5635–5643 https://doi.org/10.1007/s11033-013-2665-5.

Norimine, J., & Brown, W. C. 2005. Intrahaplotype and interhaplotype pairing of bovine leukocyte antigen DQA and DQB molecules generate functional DQ molecules important for priming CD4+ T-lymphocyte responses. Immunogenetics 57, 750–762 https://doi.org/10.1007/s00251-005-0045-6.

Nyamushamba, G. B., Mapiye, C., Tada, O., Halimani, T. E., & Muchenje, V. 2017. Conservation of indigenous cattle genetic resources in Southern Africa's smallholder areas: Turning threats into opportunities - A review. Asian-Australasian J. Anim. Sci. 30, 603–621 https://doi.org/10.5713/ajas.16.0024.

Onteru, S. K., Gorbach, D. M., Young, J. M., Garrick, D. J., Dekkers, J. C. M., & Rothschild, M. F. 2013. Whole Genome Association Studies of Residual Feed Intake and Related Traits in the Pig. PLoS One 8 https://doi.org/10.1371/journal.pone.0061756.

Oprzadek, J. M., Brzozowska, A. M., Urtnowski, P., Rutkowska, K., & Lukaszewicz, M. 2018. Association of BoLA-DRB3 genotype with somatic cell count in milk of Polish Holstein cattle. Rev. Bras. Zootec. 47 https://doi.org/10.1590/rbz4720150290.

Ornitz, D. M., & Itoh, N. 2015. The fibroblast growth factor signaling pathway. Wiley Interdiscip. Rev. Dev. Biol. 4, 215–266 https://doi.org/10.1002/wdev.176.

Ørstavik, S., Natarajan, V., Taskén, K., Jahnsen, T., & Sandberg, M. 1997. Characterization of the human gene encoding the type Iα and type Iβ cGMP-dependent protein kinase (PRKG1). Genomics 42, 311–318 https://doi.org/10.1006/geno.1997.4743.

Ortega, M. S., Denicol, A. C., Cole, J. B., Null, D. J., & Hansen, P. J. 2016. Use of single nucleotide polymorphisms in candidate genes associated with daughter pregnancy rate for prediction of genetic merit for reproduction in Holstein cows. Anim. Genet. 47, 288–297 https://doi.org/10.1111/age.12420.

Page, N. M., Butlin, D. J., Lomthaisong, K., & Lowry, P. J. 2001. The human apolipoprotein L gene cluster: Identification, classification, and sites of distribution. Genomics 74, 71–78 https://doi.org/10.1006/geno.2001.6534.

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., Mcdonald, S., Mcguinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Mckenzie, J. E. 2021. PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. BMJ 372 https://doi.org/10.1136/bmj.n160.

Patel, J. B., & Chauhan, J. B. 2017. Polymorphism of the Prolactin Gene and Its Relationship with Milk Production in Gir and Kankrej Cattle. J. Nat. Sci. Biol. Med. 8, 167–170 https://doi.org/10.4103/jnsbm.JNSBM_303_16.

Pecka-Kiełb, E., Kowalewska-łuczak, I., Czerniawska-Piątkowska, E., & Zielak-Steciwko, A. E. 2020. Effects of single nucleotide polymorphisms in the SLC27A3 gene on the nutritional value of sheep milk. Animals 10, 1–11 https://doi.org/10.3390/ani10040562.

Pei, S. W., Qin, F., Li, W. H., Li, F. D., & Yue, X. P. 2019. Copy number variation of ZNF280AY across 21 cattle breeds and its association with the reproductive traits of Holstein and Simmental bulls. J. Dairy Sci. 102, 7226–7236 https://doi.org/10.3168/jds.2018-16063.

Peng, S.-J., Cao, X.-K., Dong, D., Liu, M., Hao, D., Shen, X.-M., Huang, Y.-Z., Lei, C.-Z., Ma, Y., Bai, Y.-Y., Hu, L.-Y., Qi, X.-L., Chaogetu, B., & Chen, H. 2020. Integrative analysis of APOL3 gene CNV for adult cattle stature. Anim. Biotechnol. 31, 440–446 https://doi.org/10.1080/10495398.2019.1615933.

Perez, D. M. 2020. α1-Adrenergic Receptors in Neurotransmission, Synaptic Plasticity, and Cognition. Front. Pharmacol. 11, 1–22 https://doi.org/10.3389/fphar.2020.581098.

Pérez O'Brien, A. M., Utsunomiya, Y. T., Mészáros, G., Bickhart, D. M., Liu, G. E., Van Tassell, C. P., Sonstegard, T. S., Da Silva, M. V., Garcia, J. F., & Sölkner, J. 2014. Assessing signatures of selection through variation in linkage disequilibrium between taurine and indicine cattle. Genet. Sel. Evol. 46, 1–14 https://doi.org/10.1186/1297-9686-46-19.

Peripolli, E. 2021. Identification of structural variants and selection signatures in cattle.

Pickering, L. H. L. . 2017. The relationship between copy number variations and tick resistance in South African Nguni cattle.

Pierce, M. D., Dzama, K., & Muchadeyi, F. C. 2018. Genetic diversity of seven cattle breeds inferred using copy number variations. Front. Genet. 9, 1–14 https://doi.org/10.3389/fgene.2018.00163.

Pinto, I. P., da Cruz, A., Costa, E., Pereira, S., Minasi, L., & da Cruz, A. 2018. Cytogenomic Microarray Testing.in Cytogenetics. Larramendy, M., Soloneski, S., eds. IntechOpen, Rijeka.

Pinto, D., Darvishi, K., Shi, X., Rajan, D., Rigler, D., Fitzgerald, T., Lionel, A. C., Thiruvahindrapuram, B., MacDonald, J. R., Mills, R., Prasad, A., Noonan, K., Gribble, S., Prigmore, E., Donahoe, P. K., Smith, R. S., Park, J. H., Hurles, M. E., Carter, N. P., Lee, C., Scherer, S. W., & Feuk, L. 2012. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants Dalila. Nat. Biotechnol. 29, 512–520 https://doi.org/10.1038/nbt.1852.Comprehensive.

Pirooznia, M., Goes, F., & Zandi, P. P. 2015. Whole-genome CNV analysis: Advances in computational approaches. Front. Genet. 6 https://doi.org/10.3389/fgene.2015.00138.

Pitt, D., Sevane, N., Nicolazzi, E. L., MacHugh, D. E., Park, S. D. E., Colli, L., Martinez, R., Bruford, M. W., & Orozco-terWengel, P. 2019. Domestication of cattle: Two or three events? Evol. Appl. 12, 123–136 https://doi.org/10.1111/eva.12674.

Pollock, J. M., & Welsh, M. D. 2002. The WC1+ γδ T-cell population in cattle: A possible role in resistance to intracellular infection. Vet. Immunol. Immunopathol. 89, 105–114 https://doi.org/10.1016/S0165-2427(02)00200-3.

Porto-Neto, L. R., Sonstegard, T. S., Liu, G. E., Bickhart, D. M., Da Silva, M. V. B., Machado, M. A., Utsunomiya, Y. T., Garcia, J. F., Gondro, C., & Van Tassell, C. P. 2013. Genomic divergence of zebu and taurine cattle identified through high-density SNP genotyping. BMC Genomics 14 https://doi.org/10.1186/1471-2164-14-876.

Pös, O., Radvanszky, J., Buglyó, G., Pös, Z., Rusnakova, D., Nagy, B., & Szemes, T. 2021. DNA copy number variation: Main characteristics, evolutionary significance, and pathological aspects. Biomed. J. 44, 548–559 https://doi.org/10.1016/j.bj.2021.02.003.

Prinsen, R. T. M. M. T. M. M., Rossoni, A., Gredler, B., Bieber, A., Bagnato, A., & Strillacci, M. G. 2017. A genome wide association study between CNVs and quantitative traits in Brown Swiss cattle. Livest. Sci. 202, 7–12 https://doi.org/10.1016/j.livsci.2017.05.011.

Prinsen, R. T. M. M., Strillacci, M. G., Schiavini, F., Santus, E., Rossoni, A., Maurer, V., Bieber, A., Gredler, B., Dolezal, M., & Bagnato, A. 2016. A genome-wide scan of copy number variants using high-density SNPs in Brown Swiss dairy cattle. Livest. Sci. 191, 153–160 https://doi.org/10.1016/j.livsci.2016.08.006.

Qanbari, S., & Simianer, H. 2014. Mapping signatures of positive selection in the genome of livestock. Livest. Sci. 166, 133–143 https://doi.org/https://doi.org/10.1016/j.livsci.2014.05.003.

Rafter, P., Gormley, I. C., Parnell, A. C., Kearney, J. F., & Berry, D. P. 2020. Concordance rate between copy number variants detected using either high-or medium-density single nucleotide polymorphism genotype panels and the potential of imputing copy number variants from flanking high density single nucleotide polymorphism haplotyp. BMC Genomics 21, 1–10 https://doi.org/10.1186/s12864-020-6627-8.

Rafter, P., Gormley, I. C., Parnell, A. C., Naderi, S., & Berry, D. P. 2021a. The Contribution of Copy Number Variants and Single Nucleotide Polymorphisms to the Additive Genetic Variance of Carcass Traits in Cattle. Front. Genet. 12 https://doi.org/10.3389/fgene.2021.761503.

Rafter, P., Gormley, I. C., Purfield, D., Parnell, A. C., Naderi, S., & Berry, D. P. 2021b. Genome-wide association analyses of carcass traits using copy number variants and raw intensity values of single nucleotide polymorphisms in cattle. BMC Genomics 22 https://doi.org/10.1186/s12864-021-08075-2.

Rafter, P., Purfield, D. C., Berry, D. P., Parnell, A. C., Claire Gormley, I., Kearney, J. F., Coffey, M. P., & Carthy, T. R. 2018. Characterization of copy number variants in a large multibreed population of beef and dairy cattle using high-density single nucleotide polymorphism genotype data. J. Anim. Sci. 96, 4112–4124 https://doi.org/10.1093/jas/sky302.

Ramírez, N., Montoya, A., Cerón-muñoz, M., Zoot, D., Villar, Luis, P., & Dvm. 2014. Association of BoLA-DRB3 and TLR4 alleles with subclinical mastitis in cattle from Colombia. Rev. Colomb. Ciencias Pecu. 27.

Randhawa, I. A. S., Khatkar, M. S., Thomson, P. C., & Raadsma, H. W. 2016. A meta-assembly of selection signatures in cattle. PLoS One 11, e0153013.

Raposo, G., & Marks, M. S. 2007. Melanosomes--dark organelles enlighten endosomal membrane transport. Nat. Rev. Mol. Cell Biol. 8, 786–797 https://doi.org/10.1038/nrm2258.

Raza, S. H. A., Khan, R., Abdelnour, S. A., El-Hack, M. E. A., Khafaga, A. F., Taha, A., Ohran, H., Mei, C., Schreurs, N. M., & Zan, L. 2019. Advances of molecular markers and their application for body variables and carcass traits in Qinchuan Cattle. Genes (Basel). 10, 1–24 https://doi.org/10.3390/genes10090717.

Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D., Fiegler, H., Shapero, M. H., Carson, A. R., Chen, W., Cho, E. K., Dallaire, S., Freeman, J. L., González, J. R., Gratacòs, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J. R., Marshall, C. R., Mei, R., Montgomery, L., Nishimura, K., Okamura, K., Shen, F., Somerville, M. J., Tchinda, J., Valsesia, A., Woodwark, C., Yang, F., Zhang, J., Zerjal, T., Zhang, J., Armengol, L., Conrad, D. F., Estivill, X., Tyler-Smith, C., Carter, N. P., Aburatani, H., Lee, C., Jones, K. W., Scherer, S. W., & Hurles, M. E. 2006. Global variation in copy number in the human genome. Nature 444, 444–454 https://doi.org/10.1038/nature05329.

Rehfeldt, C., Stabenow, B., Pfuhl, R., Block, J., Nurnberg, G., Otten, W., Metges, C. C., & Kalbe, C. 2012. Effects of limited and excess protein intakes of pregnant gilts on carcass quality and cellular properties of skeletal muscle and subcutaneous adipose tissue in fattening pigs. J. Anim. Sci. 90, 184–196 https://doi.org/10.2527/jas.2011-4234.

Renaudeau, D., Gourdine, J.-L., & St-Pierre, N. 2011. Meta-analysis of the effects of high ambient temperature on growth performance of growing-finishing pigs. J. Anim. Sci. 89, 2220–2230 https://doi.org/10.2527/jas.2010-3329.

Rewe, T. O., Herold, P., Kahi, A. K., & Valle Zárate, A. 2009. Breeding indigenous cattle genetic resources for

beef production in Sub-Saharan Africa. Outlook Agric. 38, 317–326 https://doi.org/10.5367/000000009790422205.

Rezgaoui, M., Hermey, G., Riedel, I. B., Hampe, W., Schaller, H. C., & Hermans-Borgmeyer, I. 2001. Identification of SorCS2, a novel member of the VPS10 domain containing receptor family, prominently expressed in the developing mouse brain. Mech. Dev. 100, 335–338 https://doi.org/10.1016/S0925-4773(00)00523-2.

Rezigalla, A. A. 2020. Observational Study Designs: Synopsis for Selecting an Appropriate Study Design. Cureus 12, 1–8 https://doi.org/10.7759/cureus.6692.

Riccaboni, M., & Verginer, L. 2022. The impact of the COVID-19 pandemic on scientific research in the life sciences. PLoS One 17, 1–16 https://doi.org/10.1371/journal.pone.0263001.

Rincon, G., Weber, K. L., Van Eenennaam, A. L., Golden, B. L., & Medrano, J. F. 2011. Hot topic: Performance of bovine high-density genotyping platforms in Holsteins and Jerseys. J. Dairy Sci. 94, 6116–6121 https://doi.org/10.3168/jds.2011-4764.

Robert, C. 2019. 4.31 - Recent Advances in Genomics-Based Animal Biotechnology: State of the Art of Selective Breeding and Embryo Genomics☆.Pages 365–375 in Moo-Young, M.B.T.-C.B. (Third E., ed. Pergamon, Oxford.

Rojas-Downing, M. M., Nejadhashemi, A. P., Harrigan, T., & Woznicki, S. A. 2017. Climate change and livestock: Impacts, adaptation, and mitigation. Clim. Risk Manag. 16, 145–163 https://doi.org/10.1016/j.crm.2017.02.001.

Roosen, S., Exner, K., Paul, S., Schröder, J. M., Kalm, E., & Looft, C. 2004. Bovine β-defensins: Identification and characterization of novel bovine β-defensin genes and their expression in mammary gland tissue. Mamm. Genome 15, 834–842 https://doi.org/10.1007/s00335-004-2387-z.

Rosen, B. D., Bickhart, D. M., Schnabel, R. D., Koren, S., Elsik, C. G., Tseng, E., Rowan, T. N., Low, W. Y., Zimin, A., Couldrey, C., Hall, R., Li, W., Rhie, A., Ghurye, J., McKay, S. D., Thibaud-Nissen, F., Hoffman, J., Murdoch, B. M., Snelling, W. M., McDaneld, T. G., Hammond, J. A., Schwartz, J. C., Nandolo, W., Hagen, D. E., Dreischer, C., Schultheiss, S. J., Schroeder, S. G., Phillippy, A. M., Cole, J. B., Van Tassell, C. P., Liu, G., Smith, T. P. L., & Medrano, J. F. 2020. De novo assembly of the cattle reference genome with single-molecule sequencing. Gigascience 9, 1–9 https://doi.org/10.1093/gigascience/giaa021.

Rosse, I. C., Assis, J. G., Oliveira, F. S., Leite, L. R., Araujo, F., Zerlotini, A., Volpini, A., Dominitini, A. J., Lopes, B. C., & Arbex, W. A. 2017. Whole genome sequencing of Guzerá cattle reveals genetic variants in candidate genes for production, disease resistance, and heat tolerance. Mamm. genome 28, 66–80.

Rowan, T. 2020. Leveraging large scale beef cattle genomic data to identify the architecture of polygenic selection and local adaptation.

Rupp, R., Hernandez, A., & Mallard, B. A. 2007. Association of bovine leukocyte antigen (BoLA) DRB3.2 with immune response, mastitis, and production and type traits in Canadian Holsteins. J. Dairy Sci. 90, 1029–1038 https://doi.org/10.3168/jds.S0022-0302(07)71589-8.

Rusk, R. A., Palmer, M. V., Waters, W. R., & McGill, J. L. 2017. Measuring bovine γδ T cell function at the site of Mycobacterium bovis infection. Vet. Immunol. Immunopathol. 193–194, 38–49 https://doi.org/10.1016/j.vetimm.2017.10.004.

Russel, F. G. M., Koenderink, J. B., & Masereeuw, R. 2008. Multidrug resistance protein 4 (MRP4/ABCC4): a versatile efflux transporter for drugs and signalling molecules. Trends Pharmacol. Sci. 29, 200–207 https://doi.org/10.1016/j.tips.2008.01.006.

Rust, J. M., & Rust, T. 2013. Climate change and livestock production: A review with emphasis on Africa. South African J. Anim. Sci. 43, 256–267 https://doi.org/10.4314/sajas.v43i3.3.

Sadkowski, T., Jank, M., Zwierzchowski, L., Siadkowska, E., Oprządek, J., & Motyl, T. 2008. Gene expression profiling in skeletal muscle of Holstein-Friesian bulls with single-nucleotide polymorphism in the myostatin gene 5′-flanking region. J. Appl. Genet. 49, 237–250 https://doi.org/10.1007/BF03195620.

Sainz, R. D., Cruz, G. D., Mendes, E., Magnabosco, C. U., Farjalla, Y. B., Araujo, F. R. C., Gomes, R. C., & Leme, P. R. 2013. Performance, efficiency and estimated maintenance energy requirements of Bos taurus and Bos indicus cattle. Energy protein Metab. Nutr. Sustain. Anim. Prod., 69–70 https://doi.org/10.3920/978-90-8686-781-3_8.

Salomón-Torres, R., González-Vizcarra, V. M., Medina-Basulto, G. E., Montaño-Gómez, M. F., Mahadevan, P., Yaurima-Basaldúa, V. H., Villa-Angulo, C., Villa-Angulo, R., Salomon-Torres, R., Gonzalez-Vizcarra, V. M., Medina-Basulto, G. E., Montano-Gomez, M. F., Mahadevan, P., Yaurima-Basaldua, V. H., Villa-Angulo, C., & Villa-Angulo, R. 2015. Genome-wide identification of copy number variations in holstein cattle from Baja California, Mexico, using high-density SNP genotyping arrays. Genet. Mol. Res. 14, 11848–11859 https://doi.org/10.4238/2015.October.2.18.

Salomon-Torres, R., Villa-Angulo, R., & Villa-Angulo, C. 2016. Analysis of copy number variations in Mexican Holstein cattle using axiom genome-wide Bos 1 array. Genomics Data 7, 97–100 https://doi.org/10.1016/j.gdata.2015.12.007.

Sanarana, Y., Visser, C., Bosman, L., Nephawe, K., Maiwashe, A., & van Marle-Köster, E. 2016. Genetic diversity in South African Nguni cattle ecotypes based on microsatellite markers. Trop. Anim. Health

Prod. 48, 379–385 https://doi.org/10.1007/s11250-015-0962-9.
Sánchez-Molano, E., Bay, V., Smith, R. F., Oikonomou, G., & Banos, G. 2019. Quantitative Trait Loci Mapping for Lameness Associated Phenotypes in Holstein–Friesian Dairy Cattle. Front. Genet. 10, 1–9 https://doi.org/10.3389/fgene.2019.00926.
Sandholm, R. M. 2022. Evolution of Alternative Splice Variation and Exon Usage Following Whole Genome Duplication.
Saravanan, K. A., Panigrahi, M., Kumar, H., Bhushan, B., Dutt, T., & Mishra, B. P. 2020. Selection signatures in livestock genome: A review of concepts, approaches and applications. Livest. Sci. 241, 104257 https://doi.org/https://doi.org/10.1016/j.livsci.2020.104257.
Saravanan, K. A., Panigrahi, M., Kumar, H., Nayak, S. S., Rajawat, D., Bhushan, B., & Dutt, T. 2022. Progress and future perspectives of livestock genomics in India: a mini review. Anim. Biotechnol. 0, 1–9 https://doi.org/10.1080/10495398.2022.2056046.
Saravanan, K. A., Panigrahi, M., Kumar, H., Parida, S., Bhushan, B., Gaur, G. K., Dutt, T., Mishra, B. P., & Singh, R. K. 2021. Genomic scans for selection signatures revealed candidate genes for adaptation and production traits in a variety of cattle breeds. Genomics 113, 955–963 https://doi.org/https://doi.org/10.1016/j.ygeno.2021.02.009.
Sasaki, S., Ibi, T., Akiyama, T., Fukushima, M., & Sugimoto, Y. 2016a. Loss of maternal ANNEXIN A10 via a 34-kb deleted-type copy number variation is associated with embryonic mortality in Japanese Black cattle. BMC Genomics 17 https://doi.org/10.1186/s12864-016-3312-z.
Sasaki, S., Miki, Y., Ibi, T., Wakaguri, H., Yoshida, Y., Sugimoto, Y., & Suzuki, Y. 2021. A 44-kb deleted-type copy number variation is associated with decreasing complement component activity and calf mortality in Japanese Black cattle. BMC Genomics 22 https://doi.org/10.1186/s12864-021-07415-6.
Sasaki, S., Watanabe, T., Nishimura, S., & Sugimoto, Y. 2016b. Genome-wide identification of copy number variation using high-density single-nucleotide polymorphism array in Japanese Black cattle. BMC Genet. 17, 1–9 https://doi.org/10.1186/s12863-016-0335-z.
Ben Sassi, N., González-Recio, Ó., de Paz-del Río, R., Rodríguez-Ramilo, S. T., & Fernández, A. I. 2016. Associated effects of copy number variants on economically important traits in Spanish Holstein dairy cattle. J. Dairy Sci. 99, 6371–6380 https://doi.org/10.3168/jds.2015-10487.
Scherer, R. W., & Saldanha, I. J. 2019. How should systematic reviewers handle conference abstracts? A view from the trenches. Syst. Rev. 8, 4–9 https://doi.org/10.1186/s13643-019-1188-0.
Seidel, G. E. 2010. Brief introduction to whole-genome selection in cattle using single nucleotide polymorphisms. Reprod. Fertil. Dev. 22, 138–144 https://doi.org/10.1071/RD09220.
Seleguim Chud, T., Bickhart, D., Cole, J., Baldi Rey, F., da Silva, M., & Munari, D. 2018. Copy number variation of PRAME gene across dairy breeds. J. Anim. Sci. 121
Selsted, M. E., Miller, S. I., Henschen, A. H., & Ouellette, A. J. 1992. Enteric defensins: Antibiotic peptide components of intestinal host defense. J. Cell Biol. 118, 929–936 https://doi.org/10.1083/jcb.118.4.929.
Seroussi, E., Glick, G., Shirak, A., Yakobson, E., Weller, J. I., Ezra, E., & Zeron, Y. 2010. Analysis of copy loss and gain variations in Holstein cattle autosomes using BeadChip SNPs. BMC Genomics 11, 673 https://doi.org/10.1186/1471-2164-11-673.
Shan, J., Ballard, D., & Vinson, D. R. 2020. Publication Non Grata: The Challenge of Publishing Non-COVID-19 Research in the COVID Era. Cureus 12, 10–14 https://doi.org/10.7759/cureus.11403.
Sharma, A., Park, J.-E., Chai, H.-H., Jang, G.-W., Lee, S.-H., & Lim, D. 2017. Next generation sequencing in animal science - A review. Anim. Sci. Pap. Reports 35, 205–224.
Sheikh, F., Chen, Y., Liang, X., Hirschy, A., Stenbit, A. E., Gu, Y., Dalton, N. D., Yajima, T., Lu, Y., Knowlton, K. U., Peterson, K. L., Perriard, J. C., & Chen, J. 2006. α-E-catenin inactivation disrupts the cardiomyocyte adherens junction, resulting in cardiomyopathy and susceptibility to wall rupture. Circulation 114, 1046–1055 https://doi.org/10.1161/CIRCULATIONAHA.106.634469.
Sherman, E. L., Nkrumah, J. D., & Moore, S. S. 2010. Whole genome single nucleotide polymorphism associations with feed intake and feed efficiency in beef cattle. J. Anim. Sci. 88, 16–22 https://doi.org/10.2527/jas.2008-1759.
Sherman, E. L., Nkrumah, J. D., Murdoch, B. M., Li, C., Wang, Z., Fu, A., & Moore, S. S. 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 meri. J. Anim. Sci. 86, 1–16 https://doi.org/10.2527/jas.2006-799.
Shi, L., Lv, X., Liu, L., Yang, Y., Ma, Z., Han, B., & Sun, D. 2019. A post-GWAS confirming effects of PRKG1 gene on milk fatty acids in a Chinese Holstein dairy population. BMC Genet. 20, 1–14 https://doi.org/10.1186/s12863-019-0755-7.
Shi, T., Xu, Y., Yang, M., Huang, Y., Lan, X., Lei, C., Qi, X., Yang, X., & Chen, H. 2016. Copy number variations at LEPR gene locus associated with gene expression and phenotypic traits in Chinese cattle. Anim. Sci. J. 87, 336–343 https://doi.org/10.1111/asj.12531.
Shin, D.-H., Lee, H.-J., Cho, S., Kim, H. J., Hwang, J. Y., Lee, C.-K., Jeong, J., Yoon, D., & Kim, H. 2014. Deleted copy number variation of Hanwoo and Holstein using next generation sequencing at the

population level. BMC Genomics 15 https://doi.org/10.1186/1471-2164-15-240.
Shirasuna, K., Matsumoto, H., Matsuyama, S., Kimura, K., Bollwein, H., & Miyamoto, A. 2013. Possible role of interferon tau on the bovine corpus luteum and neutrophils during the early pregnancy. , 1–9 https://doi.org/10.1530/REP-15-0085.
Silpa, M. V., Konig, S., Sejian, V., Malik, P. K., Reshma Nair, M. R., Fonsêca, V. D. F. C., Maia, A. S. C., & Bhatta, R. 2021. Climate resilient dairy cattle production: Applications of genomic tools and statistical models. Front. Vet. Sci. 8, 327.
Da Silva, V. H., De Almeida Regitano, L. C., Geistlinger, L., Pértille, F., Giachetto, P. F., Brassaloti, R. A., Morosini, N. S., Zimmer, R., & Coutinho, L. L. 2016a. Genome-wide detection of CNVs and their association with meat tenderness in Nelore cattle. PLoS One 11, e0157711 https://doi.org/10.1371/journal.pone.0157711.
Da Silva, J. M., Giachetto, P. F., da Silva, L. O., Cintra, L. C., Paiva, S. R., Yamagishi, M. E. B., & Caetano, A. R. 2016b. Genome-wide copy number variation (CNV) detection in Nelore cattle reveals highly frequent variants in genome regions harboring QTLs affecting production traits. BMC Genomics 17, 1–14 https://doi.org/10.1186/s12864-016-2752-9.
Silva, J. M. da, Giachetto, P. F., Silva, L. O. C. da, Cintra, L. C., Paiva, S. R., Caetano, A. R., & Yamagishi, M. E. B. 2015. Genomic variants revealed by invariably missing genotypes in Nelore cattle. PLoS One 10, e0136035.
Simpson, F., Peden, A. A., Christopoulou, L., Robinson, M. S., Darnell, R. B., Robinson, M. S., & Cell, J. 1997. Characterization of the Adaptor-related Protein Complex, AP-3. 137, 835–845.
Smith, T. P. L., Casas, E., Rexroad, C. E., Kappes, S. M., & Keele, J. W. 2000. Bovine CAPN1 maps to a region of BTA29 containing a quantitative trait locus for meat tenderness. J. Anim. Sci. 78, 2589–2594 https://doi.org/10.2527/2000.78102589x.
Smith, W. D., & Zarlenga, D. S. 2006. Developments and hurdles in generating vaccines for controlling helminth parasites of grazing ruminants. Vet. Parasitol. 139, 347–359 https://doi.org/10.1016/j.vetpar.2006.04.024.
Soares, M. J. 2004. The prolactin and growth hormone families: Pregnancy-specific hormones/cytokines at the maternal-fetal interface. Reprod. Biol. Endocrinol. 2, 1–15 https://doi.org/10.1186/1477-7827-2-51.
Sodani, K., Patel, A., Kathawala, R. J., & Chen, Z. S. 2012. Multidrug resistance associated proteins in multindrung resistance. Chin. J. Cancer 31, 58–72 https://doi.org/10.5732/cjc.011.10329.
Sonstegard, T. S., & Gasbarre, L. C. 2001. Genomic tools to improve parasite resistance. Vet. Parasitol. 101, 387–403 https://doi.org/10.1016/S0304-4017(01)00563-5.
Sosa, F., Santos, J. E. P., Rae, D. O., Larson, C. C., Macchietto, M., Abrahante, J. E., Amaral, T. F., Denicol, A. C., Sonstegard, T. S., & Hansen, P. J. 2022. Effects of the SLICK1 mutation in PRLR on regulation of core body temperature and global gene expression in liver in cattle. Animal 16, 100523 https://doi.org/10.1016/j.animal.2022.100523.
de Souza Santos, E., de Bessa, S. A., Netto, M. M., & Nagai, M. A. 2008. Silencing of LRRC49 and THAP10 genes by bidirectional promoter hypermethylation is a frequent event in breast cancer. Int. J. Oncol. 33, 25–31 https://doi.org/10.3892/ijo.33.1.25.
Stella, A., Ajmone-Marsan, P., Lazzari, B., & Boettcher, P. 2010. Identification of selection signatures in cattle breeds selected for dairy production. Genetics 185, 1451–1461 https://doi.org/10.1534/genetics.110.116111.
Stothard, P., Choi, J.-W., Basu, U., Sumner-Thomson, J. M., Meng, Y., Liao, X., & Moore, S. S. 2011. Whole genome resequencing of black Angus and Holstein cattle for SNP and CNV discovery. BMC Genomics 12 https://doi.org/10.1186/1471-2164-12-559.
Strillacci, M. G., Gorla, E., Cozzi, M. C., Vevey, M., Genova, F., Scienski, K., Longeri, M., & Bagnato, A. 2018. A copy number variant scan in the autochthonous Valdostana Red Pied cattle breed and comparison with specialized dairy populations. PLoS One 13 https://doi.org/10.1371/journal.pone.0204669.
Strydom, P. E., Naudé, R. T., Smith, M. F., Kotzé, A., Scholtz, M. M., & Van Wyk, J. B. 2001. Relationships between production and product traits in subpopulations of Bonsmara and Nguni cattle. South African J. Anim. Sci. 31, 181–194 https://doi.org/10.4314/sajas.v31i3.3801.
Strydom, P. E., Naude, R. T., Smith, M. F., Scholtz, M. M., & Van Wyk, J. B. 2000. Characterisation of indigenous African cattle breeds in relation to meat quality traits. Meat Sci. 55, 79–88 https://doi.org/10.1016/S0309-1740(99)00128-X.
Surks, H. K. 2007. cGMP-dependent protein kinase I and smooth muscle relaxation: A tale of two isoforms. Circ. Res. 101, 1078–1080 https://doi.org/10.1161/CIRCRESAHA.107.165779.
Sutherland, I. A., & Leathwick, D. M. 2011. Anthelmintic resistance in nematode parasites of cattle: A global issue? Trends Parasitol. 27, 176–181 https://doi.org/10.1016/j.pt.2010.11.008.
Szczerbal, I., Robinson, T. J., Duran, A., Davey, S. C., Andersson, M., & Switonski, M. 2017. Ectopic position of duplicated KIT gene in African Nguni cattle, associated with color sidedness, confirms its shared ancestry with the Bos taurus lineage. Anim. Genet. 48, 122–123 https://doi.org/10.1111/age.12495.
Szyda, J., Mielczarek, M., Fraęszczak, M., Minozzi, G., Williams, J. L., & Wojdak-Maksymiec, K. 2019. The

genetic background of clinical mastitis in Holstein-Friesian cattle. Animal 13, 2156–2163 https://doi.org/10.1017/S1751731119000338.

Taddei, I., Faraldo, M. M., Teulière, J., Deugnier, M. A., Thiery, J. P., & Glukhova, M. A. 2003. Integrins in mammary gland development and differentiation of mammary epithelium. J. Mammary Gland Biol. Neoplasia 8, 383–394 https://doi.org/10.1023/B:JOMG.0000017426.74915.b9.

Takahashi, T., Yamada, O., Soares, M. J., & Hashizume, K. 2008. Bovine prolactin-related protein-I is anchored to the extracellular matrix through interactions with type IV collagen. J. Endocrinol. 196, 225–234 https://doi.org/10.1677/JOE-07-0069.

Takeshima, S. N., & Aida, Y. 2006. Structure, function and disease susceptibility of the bovine major histocompatibility complex. Anim. Sci. J. 77, 138–150 https://doi.org/10.1111/j.1740-0929.2006.00332.x.

Talenti, A., Powell, J., Hemmink, J. D., Cook, E. A. J., Wragg, D., Jayaraman, S., Paxton, E., Ezeasor, C., Obishakin, E. T., Agusi, E. R., Tijjani, A., Marshall, K., Fisch, A., Ferreira, B. R., Qasim, A., Chaudhry, U., Wiener, P., Toye, P., Morrison, L. J., Connelley, T., & Prendergast, J. G. D. 2022. A cattle graph genome incorporating global breed diversity. Nat. Commun. 13, 1–14 https://doi.org/10.1038/s41467-022-28605-0.

Tan, W. (Spring), Carlson, D. F., Walton, M. W., Fahrenkrug, S. C., & Hackett, P. B. 2012. Chapter Two - Precision Editing of Large Animal Genomes.Pages 37–97 in Advances in Genetics. Friedmann, T., Dunlap, J., Goodwin, S.B.T.-A. in G., eds. Academic Press.

Tang, J., Shen, X., Yang, Y., Yang, H., Qi, A., Yang, S., Qu, K., Lan, X., Huang, B., & Chen, H. 2022. Two different copy number variations of the clcn2 gene in chinese cattle and their association with growth traits. Animals 12, 1–9 https://doi.org/10.3390/ani12010041.

Taye, M., Kim, J., Yoon, S. H., Lee, W., Hanotte, O., Dessie, T., Kemp, S., Mwai, O. A., Caetano-Anolles, K., Cho, S., Oh, S. J., Lee, H. K., & Kim, H. 2017. Whole genome scan reveals the genetic signature of African Ankole cattle breed and potential for higher quality beef. BMC Genet. 18 https://doi.org/10.1186/s12863-016-0467-1.

Telfer, J. C., & Baldwin, C. L. 2015. Bovine gamma delta T cells and the function of gamma delta T cell specific WC1 co-receptors. Cell. Immunol. 296, 76–86 https://doi.org/10.1016/j.cellimm.2015.05.003.

Tellam, R. L., Lemay, D. G., Van Tassell, C. P., Lewin, H. A., Worley, K. C., & Elsik, C. G. 2009. Unlocking the bovine genome. BMC Genomics 10, 193 https://doi.org/10.1186/1471-2164-10-193.

Telugu, B., Walker, A. M., & Green, J. A. 2009. the promoter and expression of selected genes. https://doi.org/10.1186/1471-2164-10-185.

The Cattle Site (2022) *Breeds*. Available at: https://www.thecattlesite.com/knowledge-centre/breeds (Accessed: 15 December 2022).

Tijjani, A., Salim, B., da Silva, M. V. B., Eltahir, H. A., Musa, T. H., Marshall, K., Hanotte, O., & Musa, H. H. 2021. Genomic signatures of desert adaptation at gene-rich regions in zebu cattle from the African drylands. bioRxiv.

Tizioto, P. C., Gromboni, C. F., de Araujo Nogueira, A. R., de Souza, M. M., de Alvarenga Mudadu, M., Tholon, P., Rosa, A. Ô. do N., Tullio, R. R., Medeiros, S. R., Nassu, R. T., & De Almeida Regitano, L. C. 2014. Calcium and potassium content in beef: Influences on tenderness and associations with molecular markers in Nellore cattle. Meat Sci. 96, 436–440 https://doi.org/10.1016/j.meatsci.2013.08.001.

Toro, M. A. 2010. Future trends in Animal Breeding due to new genetic technologies. Adv. Anim. Biosci. 1, 546–557 https://doi.org/10.1017/s2040470010005431.

Tran, D., Tran, P. A., Tang, Y. Q., Yuan, J., Cole, T., & Selsted, M. E. 2002. Homodimeric θ-defensins from Rhesus macaque leukocytes. Isolation, synthesis, antimicrobial activities, and bacterial binding properties of the cyclic peptides. J. Biol. Chem. 277, 3079–3084 https://doi.org/10.1074/jbc.M109117200.

Trigo, B. B., Utsunomiya, A. T. H., Fortunato, A. A. A. D., Milanesi, M., Torrecilha, R. B. P., Lamb, H., Nguyen, L., Ross, E. M., Hayes, B., & Padula, R. C. M. 2021. Variants at the ASIP locus contribute to coat color darkening in Nellore cattle. Genet. Sel. Evol. 53, 1–12.

Tsai, T., & St. John, J. C. 2016. The role of mitochondrial DNA copy number, variants, and haplotypes in farm animal developmental outcome. Domest. Anim. Endocrinol. 56, S133–S146 https://doi.org/https://doi.org/10.1016/j.domaniend.2016.03.005.

Upadhyay, M., Derks, M. F. L. L., Andersson, G., Medugorac, I., Groenen, M. A. M. M., & Crooijmans, R. P. M. A. M. A. 2021. Introgression contributes to distribution of structural variations in cattle. Genomics 113, 3092–3102 https://doi.org/https://doi.org/10.1016/j.ygeno.2021.07.005.

Upadhyay, M., da Silva, V. H., Megens, H.-J., Visker, M. H. P. W., Ajmone-Marsan, P., Bâlteanu, V. A., Dunner, S., Garcia, J. F., Ginja, C., Kantanen, J., Groenen, M. A. M., Crooijmans, R. P. M. A., Maulik, U., da Silva Vinicus, H., Hendrik-Jan, M., HPW, V. M., Paolo, A.-M., Susana, D., Catarina, G., Juha, K., AM, G. M., & PMA, C. R. 2017. Distribution and functionality of copy number variation across European cattle populations. Front. Genet. 8 https://doi.org/10.3389/fgene.2017.00108.

Ushizawa, K., Takahashi, T., Hosoe, M., Kaneyama, K., & Hashizume, K. 2005. Cloning and expression of two new prolactin-related proteins, prolactin-related protein-VIII and -IX, in bovine placenta. Reprod. Biol.

Endocrinol. 3 https://doi.org/10.1186/1477-7827-3-68.
Utsunomiya, Y. T., Milanesi, M., Fortes, M. R. S., Porto-Neto, L. R., Utsunomiya, A. T. H., Silva, M. V. G. B., Garcia, J. F., & Ajmone-Marsan, P. 2019. Genomic clues of the evolutionary history of Bos indicus cattle. Anim. Genet. 50, 557–568 https://doi.org/10.1111/age.12836.
Vajana, E., Barbato, M., Colli, L., Milanesi, M., Rochat, E., Fabrizi, E., Mukasa, C., Del Corvo, M., Masembe, C., & Muwanika, V. B. 2018. Combining landscape genomics and ecological modelling to investigate local adaptation of indigenous Ugandan cattle to East Coast fever. Front. Genet. 9, 385.
Valergakis, G. E., Arsenos, G., & Banos, G. 2007. Comparison of artificial insemination and natural service cost effectiveness in dairy cattle. Animal 1, 293–300 https://doi.org/10.1017/S1751731107340044.
Veneroni-Gouveia, G., Meirelles, S. L., Grossi, D. A., Santiago, A. C., Sonstegard, T. S., Yamagishi, M. E. B., Matukumalli, L. K., Coutinho, L. L., Alencar, M. M., Oliveira, H. N., & Regitano, L. C. A. 2011. Whole-genome analysis for backfat thickness in a tropically adapted, composite cattle breed from Brazil. Anim. Genet. 43, 518–524 https://doi.org/10.1111/j.1365-2052.2011.02286.x.
Venhoranta, H., Pausch, H., Wysocki, M., Szczerbal, I., Hänninen, R., Taponen, J., Uimari, P., Flisikowski, K., Lohi, H., Fries, R., Switonski, M., & Andersson, M. 2013. Ectopic KIT Copy Number Variation Underlies Impaired Migration of Primordial Germ Cells Associated with Gonadal Hypoplasia in Cattle (Bos taurus). PLoS One 8 https://doi.org/10.1371/journal.pone.0075659.
Verma, N., Gupta, I. D., Verma, A., Kumar, R., Das, R., & M.R, V. 2016. Novel SNPs in HSPB8 gene and their association with heat tolerance traits in Sahiwal indigenous cattle. Trop. Anim. Health Prod. 48, 175–180 https://doi.org/10.1007/s11250-015-0938-9.
Vickers, M., & Stewart, K. 2019. Feeding Growing and Finishing Cattle for Better Returns.
Vineeth, M. R., Surya, T., Sivalingam, J., Kumar, A., Niranjan, S. K., Dixit, S. P., Singh, K., Tantia, M. S., & Gupta, I. D. 2020. Genome-wide discovery of SNPs in candidate genes related to production and fertility traits in Sahiwal cattle. Trop. Anim. Health Prod. 52, 1707–1715.
van Vliet, J., Schut, A. G. T., Reidsma, P., Descheemaeker, K., Slingerland, M., Ven, G. W. J., & Giller, K. 2015. De-mystifying family farming: Features, diversity and trends across the globe. Glob. Food Sec. 5, 11–18 https://doi.org/10.1016/j.gfs.2015.03.001.
Walker, W. P., & Gunn, T. M. 2010. Shades of meaning: The pigment-type switching system as a tool for discovery. Pigment Cell Melanoma Res. 23, 485–495 https://doi.org/10.1111/j.1755-148X.2010.00721.x.
Walker, A. M., & Roberts, R. M. 2009. Characterization of the bovine type I IFN locus : rearrangements , expansions , and novel subfamilies. 15, 1–15 https://doi.org/10.1186/1471-2164-10-187.
Wallace, R. M., Pohler, K. G., Smith, M. F., & Green, J. A. 2015. Placental PAGs: Gene origins, expression patterns, and use as markers of pregnancy. Reproduction 149, R115–R126 https://doi.org/10.1530/REP-14-0485.
Waller, P. J. 1997. Nematode parasite control of livestock in the Tropics/subtropics: The need for novel approaches. Int. J. Parasitol. 27, 1193–1201 https://doi.org/10.1016/S0020-7519(97)00117-3.
Wan, X., Wang, D., Xiong, Q., Xiang, H., Li, H., Wang, H., Liu, Z., Niu, H., Peng, J., Jiang, S., & Chai, J. 2016. Elucidating a molecular mechanism that the deterioration of porcine meat quality responds to increased cortisol based on transcriptome sequencing. Sci. Rep. 6, 1–14 https://doi.org/10.1038/srep36589.
Wang., H., Jiang, L., Liu, X., Yang, J., Wei, J., Xu, J., Zhang, Q., & Liu, J. F. 2013. A post-GWAS replication study confirming the ptk2 gene associated with milk production traits in Chinese Holstein. PLoS One 8, 1–12 https://doi.org/10.1371/journal.pone.0083625.
Wang, M. D. 2016. Copy number variations in South African Nguni cattle: prevalence, characterization and genetic diversity.
Wang, M. D., Dzama, K., Hefer, C. A., & Muchadeyi, F. C. 2015. Genomic population structure and prevalence of copy number variations in South African Nguni cattle. BMC Genomics 16, 1–16 https://doi.org/10.1186/s12864-015-2122-z.
Wang, M. D., Dzama, K., Hefer, C. A., & Muchadeyi, F. C. 2016a. Whole Genome Sequencing Of South African Nguni Cattle: Copy Number Variation Prevalence And Genetic Diversity. Copy Number Var. South African Nguni Cattle Prevalence, Charact. Genet. Divers., 87.
Wang, M. D., Dzama, K., Rees, J., Hefer, C., & Muchadeyi, F. C. 2016b. Screening and characterization of copy number variation in South African nguni cattle using next-generation sequencing data. J. Anim. Sci. 61. Oxford University Press.
Wang, M. D., Dzama, K., Rees, J., & Muchadeyi, F. C. 2014. Preliminary identification and characterization of copy number variations in the genome of South African Nguni cattle.
Wang, M. D., Dzama, K., Rees, D. J. G. G., & Muchadeyi, F. C. 2016c. Tropically adapted cattle of Africa: Perspectives on potential role of copy number variations. Anim. Genet. 47, 154–164 https://doi.org/10.1111/age.12391.
Wang, F., Herzig, C. T. A., Chen, C., Hsu, H., Baldwin, C. L., & Telfer, J. C. 2011. Scavenger receptor WC1 contributes to the γδ T cell response to Leptospira. Mol. Immunol. 48, 801–809 https://doi.org/10.1016/j.molimm.2010.12.001.
Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S. F. A., Hakonarson, H., & Bucan, M. 2007.

PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res. 17, 1665–1674 https://doi.org/10.1101/gr.6861907.

Wang, L., Liu, X., Niu, F., Wang, H., He, H., & Gu, Y. 2013. Single nucleotide polymorphisms, haplotypes and combined genotypes in MYH3 gene and their associations with growth and carcass traits in Qinchuan cattle. Mol. Biol. Rep. 40, 417–426 https://doi.org/10.1007/s11033-012-2076-z.

Wang, M. D., Muchadeyi, F. C., Makgahela, M., Mdyogolo, S., & Maiwashe, A. 2018. Copy number variations and quantitative trait loci in South African Brahman cattle.

Weldenegodguad, M., Popov, R., Pokharel, K., Ammosov, I., Ming, Y., Ivanova, Z., & Kantanen, J. 2019. Whole-genome sequencing of three native cattle breeds originating from the northernmost cattle farming regions. Front. Genet. 9, 728.

Wen, Y. F., He, H., Liu, H. B., An, Q. M., Wang, D. H., Ding, X. T., Shi, Q. T., Feng, Y. J., Wang, E. Y., Lei, C. Z., Zhang, Z. J., & Huang, Y. Z. 2020. Copy number variation of the USP16 gene and its association with milk traits in Chinese Holstein cattle. Anim. Biotechnol. 33, 98–103 https://doi.org/10.1080/10495398.2020.1777148.

Wheeler, T. T., Hood, K., Oden, K., Mccracken, J., & Morris, C. A. 2003. Bovine parotid secretory protein : structure , expression and relatedness to other BPI. https://doi.org/10.1042/BST0310781.

Whelehan, C. J., Barry-Reidy, A., Meade, K. G., Eckersall, P. D., Chapwanya, A., Narciandi, F., Lloyd, A. T., & O'Farrelly, C. 2014. Characterisation and expression profile of the bovine cathelicidin gene repertoire in mammary tissue. BMC Genomics 15, 1–13 https://doi.org/10.1186/1471-2164-15-128.

Whitton, C., Bogueva, D., Marinova, D., & Phillips, C. J. C. 2021. Are we approaching peak meat consumption? Analysis of meat consumption from 2000 to 2019 in 35 countries and its rela-tionship to gross domestic product. Animals 11 https://doi.org/10.3390/ani11123466.

Wiedemann, I., Krebs, T., Momberg, N., & Knorr, C. 2018. Original Article mRNA expression profiling in cotyledons reveals significant up-regulation of the two bovine pregnancy- associated glycoprotein genes boPAG-8 and boPAG-11 in early gestation. , 341–350 https://doi.org/10.1002/vms3.123.

Winchester, L., Yau, C., & Ragoussis, J. 2009. Comparing CNV detection methods for SNP arrays. Briefings Funct. Genomics Proteomics 8, 353–366 https://doi.org/10.1093/bfgp/elp017.

Winter, A., Krämer, W., Werner, F. A. O., Kollers, S., Kata, S., Durstewitz, G., Buitkamp, J., Womack, J. E., Thaller, G., & Fries, R. 2002. Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:Diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content. Proc. Natl. Acad. Sci. U. S. A. 99, 9300–9305 https://doi.org/10.1073/pnas.142293799.

Wojdak-Maksymiec, K., Kmieć, M., & Zukiewicz, A. 2006. Associations between defensin polymorphism and somatic cell count in milk and milk utility traits in Jersey dairy cows. J. Vet. Med. Ser. A Physiol. Pathol. Clin. Med. 53, 495–500 https://doi.org/10.1111/j.1439-0442.2006.00899.x.

Wood, I. A., Moser, G., Burrell, D. L., Mengersen, K. L., & Hetzel, D. J. S. 2006. A meta-analytic assessment of a Thyroglobulin marker for marbling in beef cattle. Genet. Sel. Evol. 38, 479–494 https://doi.org/10.1051/gse:2006016.

Wu, Y., Fan, H., Jing, S., Xia, J., Chen, Y., Zhang, L., Gao, X., Li, J., Gao, H., & Ren, H. 2015. A genome-wide scan for copy number variations using high-density single nucleotide polymorphism array in Simmental cattle. Anim. Genet. 46, 289–298 https://doi.org/10.1111/age.12288.

Xie, F., Dong, H., & Zhang, H. 2021. Regulatory Functions of Protein Tyrosine Phosphatase Receptor Type O in Immune Cells. Front. Immunol. 12, 1–7 https://doi.org/10.3389/fimmu.2021.783370.

Xing, Y., & Gill, C. A. 2016. Comparison of three methods to discover copy number variants in Nellore and Angus cattle. J. Anim. Sci. 102. Texas A&M Univ, Interdisciplinary Grad Program Genet, College Stn, TX USA.

Xu, L., Bickhart, D. M., Cole, J. B., Schroeder, S. G., Song, J., Van Tassell, C. P., Sonstegard, T. S., & Liu, G. E. 2015. Genomic signatures reveal new evidences for selection of important traits in domestic cattle. Mol. Biol. Evol. 32, 711–725 https://doi.org/10.1093/molbev/msu333.

Xu, L., Cole, J. B., Bickhart, D. M., Hou, Y., Song, J., VanRaden, P. M., Sonstegard, T. S., Van Tassell, C. P., & Liu, G. E. 2014a. Genome wide CNV analysis reveals additional variants associated with milk production traits in Holsteins. BMC Genomics 15, 1–10 https://doi.org/10.1186/1471-2164-15-683.

Xu, L., Hou, Y., Bickhart, D., Song, J., & Liu, G. 2013a. Comparative Analysis of CNV Calling Algorithms: Literature Survey and a Case Study Using Bovine High-Density SNP Data. Microarrays 2, 171–185 https://doi.org/10.3390/microarrays2030171.

Xu, L., Hou, Y., Bickhart, D. M., Song, J., Van Tassell, C. P., Sonstegard, T. S., & Liu, G. E. 2014b. A genome-wide survey reveals a deletion polymorphism associated with resistance to gastrointestinal nematodes in Angus cattle. Funct. Integr. Genomics 14, 333–339 https://doi.org/10.1007/s10142-014-0371-6.

Xu, L., Hou, Y., Bickhart, D. M., Zhou, Y., Hay, E. H. A., Song, J., Sonstegard, T. S., Van Tassell, C. P., & Liu, G. E. 2016. Population-genetic properties of differentiated copy number variations in cattle. Sci. Rep. 6 https://doi.org/10.1038/srep23161.

Xu, Y., Jiang, Y., Shi, T., Cai, H., Lan, X., Zhao, X., Plath, M., & Chen, H. 2017. Whole-genome sequencing

reveals mutational landscape underlying phenotypic differences between two widespread Chinese cattle breeds. PLoS One 12 https://doi.org/10.1371/journal.pone.0183921.

Xu, Y., Shi, T., Cai, H., Zhou, Y., Lan, X., Zhang, C., Lei, C., Qi, X., & Chen, H. 2014c. Associations of MYH3 gene copy number variations with transcriptional expression and growth traits in chinese cattle. Gene 535, 106–111 https://doi.org/10.1016/j.gene.2013.11.057.

Xu, L., Yang, L., Bickhart, D. M., Li, J., & Liu, G. E. 2018. Analysis of population-genetic properties of copy number variations.Pages 179–186 in Copy Number Variants. Springer.

Xu, L., Yang, L., Wang, L., Zhu, B., Chen, Y., Gao, H., Gao, X., Zhang, L., Liu, G. E., & Li, J. 2019a. Probe-based association analysis identifies several deletions associated with average daily gain in beef cattle. BMC Genomics 20 https://doi.org/10.1186/s12864-018-5403-5.

Xu, Y., Zhang, L., Shi, T., Zhou, Y., Cai, H., Lan, X., Zhang, C., Lei, C., & Chen, H. 2013b. Copy number variations of MICAL-L2 shaping gene expression contribute to different phenotypes of cattle. Mamm. Genome 24, 508–516 https://doi.org/10.1007/s00335-013-9483-x.

Xu, J.-W., Zheng, L., Li, L.-J., Yao, Y.-F., Hua, H., Yang, S.-Z., Wen, Y.-F., Song, C.-C., Cao, X.-K., Liu, K.-P., Zhang, G.-M., Yang, J.-M., Hao, D., Dang, R.-H., Lan, X.-Y., Lei, C.-Z., Qi, X.-L., Chen, H., & Huang, Y.-Z. 2019b. Novel copy number variation of the KLF3 gene is associated with growth traits in beef cattle. Gene 680, 99–104 https://doi.org/10.1016/j.gene.2018.08.040.

Yamada, O., Todoroki, J., Kizaki, K., Takahashi, T., Imai, K., Patel, O. V., Schuler, L. A., & Hashizume, K. 2002. Expression of prolactin-related protein I at the fetomaternal interface during the implantation period in cows. Reproduction 124, 427–437 https://doi.org/10.1530/rep.0.1240427.

Yang, P., Cai, C., Niu, M., Liu, X., Wang, H., Liang, H., Cheng, B., Zhang, Z., Chen, F., Xie, J., Qi, Z., Yang, G., Shi, Q., Wang, E., Lei, C., Chen, H., Ru, B., & Huang, Y. 2022a. Effect of copy number variation of PLA2G2A gene to growth traits in Chinese cattle. Gene 809 https://doi.org/10.1016/j.gene.2021.146014.

Yang, M., Lv, J., Zhang, L., Li, M., Zhou, Y., Lan, X., Lei, C., & Chen, H. 2017a. Association study and expression analysis of CYP4A11 gene copy number variation in Chinese cattle. Sci. Rep. 7 https://doi.org/10.1038/srep46599.

Yang, L., Niu, Q., Zhang, T., Zhao, G., Zhu, B., Chen, Y., Zhang, L., Gao, X., Gao, H., Liu, G. E., Li, J., & Xu, L. 2021. Genomic sequencing analysis reveals copy number variations and their associations with economically important traits in beef cattle. Genomics 113, 812–820 https://doi.org/10.1016/j.ygeno.2020.10.012.

Yang, L., Xu, L., Zhu, B., Niu, H., Zhang, W., Miao, J., Shi, X., Zhang, M., Chen, Y., Zhang, L., Gao, X., Gao, H., Li, L., Liu, G. E., & Li, J. 2017b. Genome-wide analysis reveals differential selection involved with copy number variation in diverse Chinese Cattle. Sci. Rep. 7 https://doi.org/10.1038/s41598-017-14768-0.

Yang, W. C., Yang, L. G., Riaz, H., Tang, K. Q., Chen, L., & Li, S. J. 2013. Effects in cattle of genetic variation within the IGF1R gene on the superovulation performance and pregnancy rates after embryo transfer. Anim. Reprod. Sci. 143, 24–29 https://doi.org/10.1016/j.anireprosci.2013.10.008.

Yang, H., Yue, B., Yang, Y., Tang, J., Yang, S., Qi, A., Qu, K., Lan, X., Lei, C., Wei, Z., Huang, B., & Chen, H. 2022b. Distribution of Copy Number Variation in SYT11 Gene and Its Association with Growth Conformation Traits in Chinese Cattle. Biology (Basel). 11, 1–10 https://doi.org/10.3390/biology11020223.

Yao, Z., Li, J., Zhang, Z., Chai, Y., Liu, X., Li, J., Huang, Y., Li, L., Huang, W., Yang, G., Chen, F., Shi, Q., Ru, B., Lei, C., Wang, E., & Huang, Y. 2022. The relationship between MFN1 copy number variation and growth traits of beef cattle. Gene 811 https://doi.org/10.1016/j.gene.2021.146071.

Yao, Y.-F., Lyu, S., Wang, X., Zhang, Z., Qu, K., Xu, J., Cai, C., Li, Z., Xie, J., Ru, B., Xu, Z., Wang, E., Lei, C., Chen, H., Huang, B., & Huang, Y. 2021. The combination between NCSTN gene copy number variation and growth traits in Chinese cattle. Anim. Biotechnol. 32, 683–687 https://doi.org/10.1080/10495398.2020.1741382.

Yoshida, T., Mukoyama, H., Furuta, H., Kondo, Y., Takeshima, S. N., Aida, Y., Kosugiyama, M., & Tomogane, H. 2009. Association of BoLA-DRB3 alleles identified by a sequence-based typing method with mastitis pathogens in Japanese Holstein cows. Anim. Sci. J. 80, 498–509 https://doi.org/10.1111/j.1740-0929.2009.00663.x.

Yudin, N. S., & Voevoda, M. I. 2015. Molecular genetic markers of economically important traits in dairy cattle. Russ. J. Genet. 51, 506–517 https://doi.org/10.1134/S1022795415050087.

Yue, X. P., Chang, T. C., DeJarnette, J. M., Marshall, C. E., Lei, C. Z., & Liu, W.-S. 2013. Copy number variation of PRAMEY across breeds and its association with male fertility in Holstein sires. J. Dairy Sci. 96, 8024–8034 https://doi.org/10.3168/jds.2013-7037.

Yue, X.-P., Dechow, C., & Liu, W.-S. 2015. A limited number of Y chromosome lineages is present in North American Holsteins. J. Dairy Sci. 98, 2738–2745 https://doi.org/10.3168/jds.2014-8601.

Yurchenko, A. A., Daetwyler, H. D., Yudin, N., Schnabel, R. D., Vander Jagt, C. J., Soloshenko, V., Lhasaranov, B., Popov, R., Taylor, J. F., & Larkin, D. M. 2018. Scans for signatures of selection in Russian cattle breed genomes reveal new candidate genes for environmental adaptation and

acclimation. Sci. Rep. 8, 1–16 https://doi.org/10.1038/s41598-018-31304-w.
Zanetti, M. 2004. Cathelicidins, multifunctional peptides of the innate immunity. J. Leukoc. Biol. 75, 39–48 https://doi.org/10.1189/jlb.0403147.
Zhai, J., Yang, Z., Cai, X., Yao, G., An, Y., Wang, W., Fan, Y., Zeng, C., & Liu, K. 2018. ZNF280B promotes the growth of gastric cancer in vitro and in vivo. Oncol Lett 15, 5819–5824 https://doi.org/10.3892/ol.2018.8060.
Zhan, B., Fadista, J., Thomsen, B., Hedegaard, J., Panitz, F., & Bendixen, C. 2011. Global assessment of genomic variation in cattle by genome resequencing and high-throughput genotyping. BMC Genomics 12 https://doi.org/10.1186/1471-2164-12-557.
Zhang, L. P., Gan, Q. F., Hou, G. Y., Gao, H. J., Li, J. Y., & Xu, S. Z. 2015a. Investigation of TG gene variants and their effects on growth, carcass composition, and meat quality traits in Chinese steers. Genet. Mol. Res. 14, 5320–5326 https://doi.org/10.4238/2015.May.22.2.
Zhang, F., Gu, W., Hurles, M. E., & Lupski, J. R. 2009. Copy number variation in human health, disease, and evolution. Annu. Rev. Genomics Hum. Genet. 10, 451–481 https://doi.org/10.1146/annurev.genom.9.081307.164217.
Zhang, G. W., Guan, J. Q., Luo, Z. G., Zhang, W. X., Wang, L., Luo, X. L., & Zuo, F. Y. 2016. A tremendous expansion of TSPY copy number in crossbred bulls (Bos taurus × Bos grunniens). J. Anim. Sci. 94, 1398–1407 https://doi.org/10.2527/jas.2015-9983.
Zhang, Y., Hu, Y., Wang, X., Jiang, Q., Zhao, H., Wang, J., Ju, Z., Yang, L., Gao, Y., Wei, X., Bai, J., Zhou, Y., & Huang, J. 2020a. Population Structure, and Selection Signatures Underlying High-Altitude Adaptation Inferred From Genome-Wide Copy Number Variations in Chinese Indigenous Cattle. Front. Genet. 10, 1–12 https://doi.org/10.3389/fgene.2019.01404.
Zhang, L., Jia, S., Plath, M., Huang, Y., Li, C., Lei, C., Zhao, X., & Chen, H. 2015b. Impact of parental bos taurus and bos indicus origins on copy number variation in traditional Chinese cattle breeds. Genome Biol. Evol. 7, 2352–2361 https://doi.org/10.1093/gbe/evv151.
Zhang, L., Jia, S., Yang, M., Xu, Y., Li, C., Sun, J., Huang, Y., Lan, X., Lei, C., Zhou, Y., Zhang, C., Zhao, X., & Chen, H. 2014. Detection of copy number variations and their effects in Chinese bulls. BMC Genomics 15 https://doi.org/10.1186/1471-2164-15-480.
Zhang, Q., Ma, Y., Wang, X., Zhang, Y., & Zhao, X. 2015c. Identification of copy number variations in Qinchuan cattle using BovineHD Genotyping Beadchip array. Mol. Genet. Genomics 290, 319–327 https://doi.org/10.1007/s00438-014-0923-4.
Zhang, X., Qu, K., Jia, P., Zhang, J., Liu, J., Lei, C., & Huang, B. 2021a. Assessing Genomic Diversity and Productivity Signatures in Dianzhong Cattle by Whole-Genome Scanning. Front. Genet., 1864.
Zhang, H., Shen, Z., Yang, Z., Jiang, H., Chu, S., Mao, Y., Li, M., Chen, Z., Aboragah, A., Loor, J. J., & Yang, Z. 2021b. Abundance of solute carrier family 27 member 6 (: SLC27A6) in the bovine mammary gland alters fatty acid metabolism. Food Funct. 12, 4909–4920 https://doi.org/10.1039/d0fo03289a.
Zhang, G. W., Wu, Y., Luo, Z., Guan, J., Wang, L., Luo, X., & Zuo, F. 2019. Comparison of Y-chromosome-linked TSPY, TSPY2, and PRAMEY genes in Taurus cattle, yaks, and interspecific hybrid bulls. J. Dairy Sci. 102, 6263–6275 https://doi.org/10.3168/jds.2018-15680.
Zhang, Z. J., Yang, P., He, P., Xu, J. W., Lyu, S. J., Liu, X., Cai, C. C., Li, H. M., Li, Z. M., Ru, B. R., Xie, J. L., Lei, C. Z., Chen, H., Wang, E. R., & Huang, Y. Z. 2020b. Distribution and association study of PLAG1 gene between copy number variation and Chinese cattle populations. Anim. Biotechnol. https://doi.org/10.1080/10495398.2020.1793769 WE - Science Citation Index Expanded (SCI-EXPANDED).
Zhang, Y. Y., Zan, L. S., Wang, H. B., Qing, L., Wu, K. X., Quan, S. A., Li, C. Q., Zhong, X., & Wang, C. J. 2011. Differentially expressed genes in skeletal muscle tissues from castrated Qinchuan cattle males compared with those from intact males. Livest. Sci. 135, 76–83 https://doi.org/10.1016/j.livsci.2010.06.070.
Zhang, G.-M., Zheng, L., He, H., Song, C.-C., Zhang, Z.-J., Cao, X.-K., Lei, C.-Z., Lan, X.-Y., Qi, X.-L., Chen, H., & Huang, Y.-Z. 2018. Associations of GBP2 gene copy number variations with growth traits and transcriptional expression in Chinese cattle. Gene 647, 101–106 https://doi.org/10.1016/j.gene.2018.01.004.
Zhao, X., & Guan, J. L. 2011. Focal adhesion kinase and its signaling pathways in cell migration and angiogenesis. Adv. Drug Deliv. Rev. 63, 610–615 https://doi.org/10.1016/j.addr.2010.11.001.
Zhao, J., Peng, H., Gao, J., Nong, A., Hua, H., Yang, S., Chen, L., Wu, X., Zhang, H., & Wang, J. 2021. Current insights into the expression and functions of tumor-derived immunoglobulins. Cell Death Discov. 7, 1–12 https://doi.org/10.1038/s41420-021-00550-9.
Zheng, L., Xu, J.-W., Li, J.-C., Wang, D.-H., An, Q.-M., Xu, L.-N., Ma, Y.-L., Wang, J., Peng, S.-J., Lei, C.-Z., Lan, X.-Y., Chen, H., Huo, L.-J., & Huang, Y.-Z. 2019. Distribution and association study in copy number variation of KCNJ12 gene across four Chinese cattle populations. Gene 689, 90–96 https://doi.org/10.1016/j.gene.2018.12.019.
Zheng, L., Xu, J., Liu, X., Zhang, Z., Zhong, J., Wen, Y., Yao, Z., Yang, P., Wang, E., Chen, F., Huang, W.,

Qi, Z., Yang, G., Lei, C., Chen, H., Ru, B., & Huang, Y. 2022. The copy number variation of DMBT1 gene effects body traits in two Chinese cattle breeds. 3 Biotech 12 https://doi.org/10.1007/s13205-022-03124-3.

Zhou, Y., Bickhart, D. M., & Liu, G. E. 2018a. Statistical detection of genome differences based on CNV segments.Pages 49–59 in Methods in Molecular Biology.

Zhou, Y., Connor, E. E., Wiggans, G. R., Lu, Y., Tempelman, R. J., Schroeder, S. G., Chen, H., & Liu, G. E. 2018b. Genome-wide copy number variant analysis reveals variants associated with 10 diverse production traits in Holstein cattle. BMC Genomics 19, 1–9 https://doi.org/10.1186/s12864-018-4699-5.

Zhou, J., Liu, L., Reynolds, E., Huang, X., Garrick, D., & Shi, Y. 2022. Discovering Copy Number Variation in Dual-Purpose XinJiang Brown Cattle. Front. Genet. 12, 747431 https://doi.org/10.3389/fgene.2021.747431.

Zhou, Y., Utsunomiya, Y. T., Xu, L., Hay, E. H. abdel, Bickhart, D. M., Alexandre, P. A., Rosen, B. D., Schroeder, S. G., Carvalheiro, R., de Rezende Neves, H. H., Sonstegard, T. S., Van Tassell, C. P., Ferraz, J. B. S., Fukumasu, H., Garcia, J. F., & Liu, G. E. 2016a. Genome-wide CNV analysis reveals variants associated with growth traits in Bos indicus. BMC Genomics 17, 1–9 https://doi.org/10.1186/s12864-016-2461-4.

Zhou, Y., Utsunomiya, Y. T., Xu, L., Hay, E. H. A., Bickhart, D. M., Sonstegard, T. S., Van Tassell, C. P., Garcia, J. F., & Liu, G. E. 2016b. Comparative analyses across cattle genders and breeds reveal the pitfalls caused by false positive and lineage-differential copy number variations. Sci. Rep. 6 https://doi.org/10.1038/srep29219.

Zimin, A. V., Delcher, A. L., Florea, L., Kelley, D. R., Schatz, M. C., Puiu, D., Hanrahan, F., Pertea, G., Van Tassell, C. P., Sonstegard, T. S., Marçais, G., Roberts, M., Subramanian, P., Yorke, J. A., & Salzberg, S. L. 2009. A whole-genome assembly of the domestic cow, Bos taurus. Genome Biol. 10 https://doi.org/10.1186/gb-2009-10-4-r42.

Zwambag, A., Kelly, M., Schenkel, F., Mandell, I., Wilton, J., & Miller, S. 2013. Heritability of beef tenderness at different aging times and across breed comparisons. Can. J. Anim. Sci. 93, 307–312 https://doi.org/10.4141/CJAS2012-100.

Zwane, A. A., Nxumalo, K. S., Makgahlela, M. L., Van Marle-Koster, E., & Maiwashe, N. 2021. Gene-set enrichment analysis of selective sweeps reveals phenotypic traits in Nguni cattle. S. Afr. J. Anim. Sci. 51, 761–777.

Zwane, A. A., Schnabel, R. D., Hoff, J., Choudhury, A., Makgahlela, M. L., Maiwashe, A., Marle-Koster, V., & Taylor, J. F. 2019. Genome-wide SNP discovery in indigenous cattle breeds of South Africa. Front. Genet. 10, 273.

APPENDIX A

Table A 1: The 27-step PRISMA checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	✓Title page
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	✓Page ii
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	✓ Page 2
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	✓ Page 4
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	✓ Table 3.1
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	✓ Page 31-32
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	✓ Table 3.2
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	✓ Page 33
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	✓ Page 33
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	✓ Page 34
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	✓ Page 34
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	✓ Page 38
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	✓ Page 38

Section and Topic	Item #	Checklist item	Location where item is reported
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	✓ Page 34-37
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	✓
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	✓
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	✓ Page 37
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	✓ Discussed on page 18
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	n/a
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	n/a
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	n/a
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	✓ Figure 3.1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	✓ Appendix B
Study characteristics	17	Cite each included study and present its characteristics.	✓ Appendix C
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	✓ Page 38
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Excel workbook
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	✓ Chpt 4
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	✓ Chpt 4&5
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Excel workbook
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	n/a
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	n/a
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	✓ Chpt 4&5
DISCUSSION			

Section and Topic	Item #	Checklist item	Location where item is reported
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	✓ Chpt 4&5
	23b	Discuss any limitations of the evidence included in the review.	✓ Page 93-94
	23c	Discuss any limitations of the review processes used.	✓ Page 93
	23d	Discuss implications of the results for practice, policy, and future research.	✓ Page 89-93
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	n/a
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	n/a
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	n/a
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	✓
Competing interests	26	Declare any competing interests of review authors.	n/a
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	n/a

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

APPENDIX B

Table B 1: Excluded publications with reasons

	Reference	Reason for exclusion
1	(Cymbron *et al.*, 2005)	No CNV data
2	(Liu *et al.*, 2009)	No CNV data
3	(Toro, 2010)	Review article
4	(Karrow *et al.*, 2011)	Book section
5	(Glick *et al.*, 2011)	Focused on specific chromosome
6	(Rincon *et al.*, 2011).	Review article
7	(Lai, 2012)	Book section
8	(Tan *et al.*, 2012)	Book section
9	(Kadri *et al.*, 2012)	Focused on specific chromosome
10	(Clop *et al.*, 2012)	Review article
11	(Liu & Bickhart, 2012)	Review article
12	(Khamzina, 2013)	Book section
13	(Mukherjee *et al.*, 2013)	Focused on specific chromosome
14	(Xu *et al.*, 2013a)	Focusing on specific gene
15	(De Donato *et al.*, 2013)	Methodological article
16	(Duan *et al.*, 2013)	Methodological article
17	(Xu *et al.*, 2013b)	Methodological article
18	(Horsburgh *et al.*, 2013)	No CNV data
19	(Lee *et al.*, 2013a)	No CNV data
20	(Liao *et al.*, 2013)	No CNV data
21	(Porto-Neto *et al.*, 2013)	No CNV data
22	(Hayes *et al.*, 2013)	Review article
23	(Grandin & Deesing, 2014)	Book section
24	(Jensen & Wright, 2014)	Book section
25	(McDaneld *et al.*, 2014)	Focused on specific chromosome
26	(Xu *et al.*, 2014c)	Focusing on specific gene
27	(Bermingham *et al.*, 2014)	No CNV data
28	(Qanbari & Simianer, 2014)	No CNV data
29	(Pérez O'Brien *et al.*, 2014)	No CNV data
30	(Bickhart & Liu, 2014)	Review article
31	(Mapholi *et al.*, 2014)	Review article
32	(Silva *et al.*, 2015)	No CNV data
33	(Xu *et al.*, 2015)	No CNV data
34	(Yue *et al.*, 2015)	No CNV data
35	(Yudin & Voevoda, 2015)	Review article
36	(Liu *et al.*, 2016)	Focusing on specific gene
37	(Shi *et al.*, 2016)	Focusing on specific gene
38	(Zhang *et al.*, 2016)	Focusing on specific gene
39	(Sasaki *et al.*, 2016a)	Focusing on specific gene
40	(Salomon-Torres *et al.*, 2016)	Methodological article
41	(Randhawa *et al.*, 2016).	No CNV data
42	(Keele *et al.*, 2016)	No CNV data
43	(Kukučková *et al.*, 2016)	No CNV data

44	(Wang *et al.*, 2016c)	Review article
45	(Tsai & St. John, 2016)	Review article
46	(Yang *et al.*, 2017a)	Focusing on specific gene
47	(Feng *et al.*, 2017)	No CNV data
48	(Kim *et al.*, 2017)	No CNV data
49	(Nel, 2017)	No CNV data
50	(Rosse *et al.*, 2017)	No CNV data
51	(Xu *et al.*, 2018)	Book section
52	(Zhou *et al.*, 2018a)	Book section
53	(Aguiar *et al.*, 2018)	Focusing on specific gene
54	(Cao *et al.*, 2018)	Focusing on specific gene
55	(Zhang *et al.,* 2018)	Focusing on specific gene
56	(Li *et al.*, 2018)	Incorrect Species
57	(Cardoso *et al.*, 2018)	No CNV data
58	(Gororo *et al*., 2018)	No CNV data
59	(Vajana *et al.*, 2018)	No CNV data
60	(Koufariotis *et al.*, 2018)	No CNV data
61	(Bhanuprakash *et al.*, 2018)	Review article
62	(Goshu *et al.*, 2018b)	Review article
63	(Cao *et al.*, 2018)	Uses a meta approach
64	(Robert, 2019)	Book section
65	(Cheng *et al.*, 2019b)	Focusing on specific gene
66	(Liu *et al.*, 2019c)	Focusing on specific gene
67	(Ma *et al.*, 2019)	Focusing on specific gene
68	(Pei *et al.*, 2019)	Focusing on specific gene
69	(Xu *et al.*, 2019b)	Focusing on specific gene
70	(Zhang *et al.*, 2019)	Focusing on specific gene
71	(Zheng *et al.*, 2019)	Focusing on specific gene
72	(Li *et al.*, 2019a)	Incorrect Species
73	(Iqbal *et al.*, 2019)	No CNV data
74	(Küttel *et al.*, 2019)	No CNV data
75	(Weldenegodguad *et al.*, 2019)	No CNV data
76	(Sánchez-Molano *et al.*, 2019)	No CNV data
77	(Zwane *et al.*, 2019)	No CNV data
78	(Lauer & Gresham, 2019)	Review article
79	(Gupta & Gupta, 2020)	Book section
80	(Peng *et al.*, 2020)	Focused on specific chromosome
81	(Fukunaga *et al.*, 2020)	Focusing on specific gene
82	(Guo *et al.*, 2020b).	Focusing on specific gene
83	(Hao *et al.*, 2020)	Focusing on specific gene
84	(Low *et al.*, 2020)	Focusing on specific gene
85	(Wen *et al.*, 2020)	Focusing on specific gene
86	(Yang *et al.*, 2022a)	Focusing on specific gene
87	(Zhang *et al.*, 2020b)	Focusing on specific gene
88	(Chen *et al.*, 2020)	Methodological article
89	(Rafter *et al.*, 2020)	Methodological article
90	(Ben-Jemaa *et al.*, 2020)	No CNV data
91	(Fernandes Júnior *et al.*, 2020)	No CNV data

92	(Nanaei *et al.*, 2020;)	No CNV data
93	(Rowan, 2020)	No CNV data
94	(Vineeth *et al.*, 2020)	No CNV data
95	(Lamb *et al.*, 2020)	No CNV data
96	(Muniz *et al.*, 2021)	No CNV data
97	(Bickhart *et al.*, 2020).	Review article
98	(Saravanan *et al.*, 2020)	Review article
99	(Lee *et al.*, 2021)	Focused on specific chromosome
100	(Yao *et al.*, 2021)	Focusing on specific gene
101	(Gautason *et al.*, 2021)	No CNV data
102	(Saravanan *et al.*, 2021)	No CNV data
103	(Tijjani *et al.*, 2021)	No CNV data
104	(Trigo *et al.*, 2021)	No CNV data
105	(Upadhyay *et al.*, 2021)	No CNV data
106	(Zhang *et al.*, 2021a)	No CNV data
107	(Zwane *et al.*, 2021)	No CNV data
108	(Silpa *et al.*, 2021)	Review article
109	(Rafter *et al.*, 2021a)	Uses data from included article
110	(Ding *et al.*, 2022)	Focusing on specific gene
111	(Hu *et al.*, 2022)	Focusing on specific gene
112	(Huang *et al.*, 2022)	Focusing on specific gene
113	(Li *et al.*, 2022a)	Focusing on specific gene
114	(Liang *et al.*, 2022)	Focusing on specific gene
115	(Liu *et al.*, 2022)	Focusing on specific gene
116	(Tang *et al.*, 2022)	Focusing on specific gene
117	(Yang *et al.*, 2022b)	Focusing on specific gene
118	(Yang *et al.*, 2022a)	Focusing on specific gene
119	(Yao *et al.*, 2022)	Focusing on specific gene
120	(Zheng *et al.*, 2022)	Focusing on specific gene
121	(Duarte *et al.*, 2022)	No CNV data
122	(Maiorano *et al.*, 2022)	No CNV data
123	(Armstrong *et al.*, 2022)	Review article

APPENDIX C

Table C 1: List of included publications

	Study reference
1	(Liu *et al.*, 2008b)
2	(Matukumalli *et al.*, 2009)
3	(Bae *et al.*, 2010)
4	(Fadista *et al.*, 2010)
5	(Liu *et al.*, 2010)
6	(Seroussi *et al.*, 2010)
7	(Hou *et al.*, 2011a)
8	(Kijas *et al.*, 2011)
9	(Liu *et al.*, 2011)
10	(Stothard *et al.*, 2011)
11	(Zhan *et al.*, 2011)
12	(Bickhart *et al.*, 2012)
13	(Hou *et al.*, 2012a)
14	(Hou *et al.*, 2012b)
15	(Hou *et al.*, 2012c)
16	(Jiang *et al.*, 2012)
17	(Choi *et al.*, 2013)
18	(Cicconardi *et al.*, 2013)
19	(Jiang *et al.*, 2013)
20	(Lee *et al.*, 2013b)
21	(Venhoranta *et al.*, 2013)
22	(Berton *et al.*, 2014)
23	(Dolezal *et al.*, 2014)
24	(Feitosa *et al.*, 2014)
25	(Shin *et al.*, 2014)
26	(Wang *et al.*, 2014)
27	(Xu *et al.*, 2014a)
28	(Xu *et al.*, 2014b)
29	(Zhang *et al.*, 2014)
30	(Bagnato *et al.*, 2015)
31	(Boussaha *et al.*, 2015)
32	(Gurgul *et al.*, 2015)
33	(Salomón-Torres *et al.*, 2015)
34	(Wang *et al.*, 2015)
35	(Wu *et al.*, 2015)
36	(Zhang *et al.*, 2015b)
37	(Zhang *et al.*, 2015c)
38	(Ben Sassi *et al.*, 2016)
39	(Bickhart *et al.*, 2016)
40	(Choi *et al.*, 2016)

41 (Da Silva *et al.*, 2016b)
42 (Da Silva *et al.*, 2016a)
43 (De Almeida Santana *et al.*, 2016)
44 (Keel *et al.*, 2016a)
45 (Keel *et al.*, 2016b)
46 (Khamzina, 2016)
47 (Prinsen *et al.*, 2016)
48 (Sasaki *et al.*, 2016b)
49 (Wang *et al.*, 2016b)
50 (Xing & Gill, 2016)
51 (Xu *et al.*, 2016)
52 (Zhou *et al.*, 2016b)
53 (Zhou *et al.*, 2016a)
54 (Couldrey *et al.*, 2017)
55 (Durán Aguilar *et al.*, 2017)
56 (Gao *et al.*, 2017)
57 (Huang *et al.*, 2017)
58 (Letaief *et al.*, 2017)
59 (Mielczarek *et al.*, 2017)
60 (Pickering, 2017)
61 (Prinsen *et al.*, 2017)
62 (Upadhyay *et al.*, 2017)
63 (Xu *et al.*, 2017)
64 (Yang *et al.*, 2017b)
65 (Antunes de Lemos *et al.*, 2018a)
66 (Antunes de Lemos *et al.*, 2018b)
67 (Hay *et al.*, 2018)
68 (Karimi *et al.*, 2018)
69 (Mielczarek *et al.*, 2018)
70 (Mustafa *et al.*, 2018)
71 (Nandolo *et al.*, 2018)
72 (Pierce *et al.*, 2018)
73 (Rafter *et al.*, 2018)
74 (Strillacci *et al.*, 2018)
75 (Wang *et al.*, 2018)
76 (Zhou *et al.*, 2018b)
77 (Carmo *et al.*, 2019)
78 (Cozzi *et al.*, 2019)
79 (Di Gerlando *et al.*, 2019)
80 (Kommadath *et al.*, 2019)
81 (Liu *et al.*, 2019a)
82 (Liu *et al.*, 2019b)
83 (Mei *et al.*, 2019)
84 (Szyda *et al.*, 2019)
85 (Xu *et al.*, 2019a)

86 (Buggiotti *et al.*, 2020)
87 (Butty *et al.*, 2020)
88 (Hu *et al.*, 2020b)
89 (Lee *et al.*, 2020)
90 (Mei *et al.*, 2020)
91 (Zhang *et al.*, 2020a)
92 (Berton *et al.*, 2021)
93 (Butty *et al.*, 2021)
94 (Guo *et al.*, 2021)
95 (Huang *et al.*, 2021b)
96 (Jang *et al.*, 2021)
97 (Kava *et al.*, 2021)
98 (Kumar *et al.*, 2021)
99 (Mei *et al.*, 2021)
100 (Peripolli, 2021)
101 (Rafter *et al.*, 2021b)
102 (Sasaki *et al.*, 2021)
103 (Yang *et al.*, 2021)
104 (Zhou *et al.*, 2022)