

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

Using QTL to Identify Genes and Pathways Underlying the Regulation and Production of Milk Components in Cattle

Thomas Lopdell ¹0000-0002-7684-4870

¹ Livestock Improvement Corporation; thomas.lopdell@lic.co.nz

Abstract: Milk is a complex liquid, and the concentrations of many of its components are under genetic control. Many genes and pathways are known to regulate milk composition, and the purpose of this review is to highlight how the discoveries of quantitative trait loci (QTL) for milk phenotypes can elucidate these pathways. The main body of this review focusses primarily on QTL discovered in cattle (*Bos taurus*) as a model species for lactation biology, with occasional references to other ruminant species, and some comparisons with human milk composition are also presented. The following section describes a range of techniques that can be used to help identify the causative genes underlying QTL when the underlying mechanism involves the regulation of gene expression. As genotype and phenotype data bases continue to grow and diversify, new QTL will continue to be discovered, and, although proving the causality of underlying genes and variants remains difficult, these new data sets will further enhance our understanding of lactation biology.

Keywords: mammary biology; mammogenesis; lactation; lactogenesis; quantitative trait loci

1. Introduction

The composition of milk is complex, featuring an emulsion of fat globules and a colloidal dispersion of casein micelles in an aqueous solution of lactose (and other carbohydrates), whey proteins, and minerals. Although milk from different species contains the same basic constituents, their proportions can vary greatly. In cattle, the typically average percentages (g/100 g) of fat, caseins, whey proteins, and lactose are 3.9%, 2.6%, 0.6% and 4.6% respectively; in humans, the corresponding percentages are 4.5%, 0.4%, 0.5% and 7.1% [1]. Even more extreme differences can be seen in other species, with some seal species, for example, expressing little to no lactose, resulting in highly concentrated milk showing fat percentages of 50% [2].

Less extreme differences in milk composition are also visible within species. In many cases, the differences in composition among individuals are under partial genetic control. Regions of the genome where the genotypes of genetic variants are associated with phenotypes such as milk composition are known as quantitative trait loci (QTL). When they can be identified, the causative genes underlying these QTL can help elucidate the pathways involved in milk production. The aim of this review is to describe some of the major pathways required for milk production in terms of the QTL and genes that have helped to identify them, as well as to note some of the methods that can be used to identify causative genes underlying QTL. This will focus primarily on cattle (*Bos taurus*) as a model species; however, some comparisons with human milk composition are also presented.

2. QTL for major pathways involved in milk production

Many QTL have been observed for both milk yield and milk composition traits. The genes attributed to these QTL have a variety of functions. Some, such as the hormones prolactin and growth hormone, and their associated signalling pathways, are involved in mammogenesis: the development of the mammary gland during puberty and pregnancy. Other pathways, such as fat and protein synthesis, and ion channels, are involved directly

in milk production. The following sections list some of the pathways involved in mammo-
genesis and lactation, as identified by QTL for milk production where the candidate causal
gene sits within those pathways.

2.1. Milk proteins

QTL have been mapped to many milk proteins, i.e., those expressed directly into
milk. In cattle, the largest proportion of milk protein content (80% [3]) consists of the
four casein proteins, encoded by a cluster of genes mapping to BTA6: *CSN1S1*, encoding
 α_{S1} -casein, *CSN1S2* for α_{S2} -casein, *CSN2* for β -casein, and *CSN3* for κ -casein. The casein
proteins aggregate into micelles in the milk, sequestering calcium phosphate as a nutrient
source for the neonate. κ -casein seems to be particularly important for successful lactation,
as knockout mice deficient in κ -casein fail to lactate, due to destabilisation of the casein
micelles [4]. Like β -lactoglobulin, the various casein proteins all show a range of coding
variants, although, with the exception of *CSN1S1*G* [5], these have not been intrinsically
linked to lower rates of gene expression like the β -lactoglobulin B variants discussed above.
However, there is evidence that the SNP responsible for the A2 variant of β -casein is
associated with both milk yield and protein yield [6], and QTL for milk protein phenotypes
have also been detected at this locus in other studies [7,8]. These QTL can overlap with, but
are not in linkage disequilibrium with, other nearby QTL assigned to the *GC* gene [7,9,10]:
this gene encodes the protein Group-Specific Component (Vitamin D Binding), and maps
around one megabase from the casein gene cluster. Outside the mammary gland, casein
expression in human CD14⁺ monocytes has been observed to upregulate expression of of
the cytokines granulocyte-macrophage colony-stimulating factor (GM-CSF), interleukin-1 β
(IL-1 β), and IL-6 via the p38-MAPK pathway [11,12]. This suggests a possible role for
caseins in regulating the innate immune response in the neonate intestine following milk
consumption.

The major whey protein in cattle, accounting for around 50% of whey protein [3]
and 10% of total protein, is β -lactoglobulin, encoded by the gene *PAEP* on BTA11. A
QTL has been detected mapping to this gene, with pleiotropic effects on milk fat yield,
protein yield, and volume [7,13]. Other studies looking at individual milk proteins have also
shown that genetic variants associated with low β -lactoglobulin concentrations yield higher
concentrations of α -, β -, and κ -caseins [14]. Variation in milk β -lactoglobulin concentrations
are primarily driven by differences between the A and B protein variants [15,16], with
the B variant showing lower expression than the A variant. Other variants mapping
within the B variant background, such as B* [17] and B' [18], cause further reductions in β -
lactoglobulin expression. Milk with low β -lactoglobulin concentrations has potential uses in
infant formula (as human milk lacks this protein), and the associated higher concentrations
of caseins would also be expected to provide better properties for cheese making.

The dominant whey protein in humans, and second in cattle (18.5% of total whey
protein [3]), is α -lactalbumin, encoded by the gene *LALBA* on BTA5. QTL associated with
milk protein concentration have been identified at this locus in several dairy populations
and breeds [19–22]. The α -lactalbumin protein also comprises a regulatory subunit of
lactose synthase complex, so it is no surprise that a QTL for milk lactose concentration has
also been identified at this locus [23]. The importance of this is shown in α -lactalbumin
knockout mice, which produce highly viscous milk, with otherwise normal fat and protein
composition, that cannot be extracted by the pups from the mammary gland [24]. It has
been shown *in vitro* that multimeric α -lactalbumin is cytotoxic to stem cells and transformed
cell lines [25], suggesting that α -lactalbumin may also have a protective function in either
the mammary gland or neonate digestive system.

Another whey protein is the iron-binding antimicrobial protein Lactoferrin, encoded
by the *LTF* gene on BTA22. In cattle, this protein exists in bovine milk at low but variable
concentrations (average 115.4 μ g/mL in [26] but ranging from 31.8 to 485.6), but reaches
much higher concentrations in human milk at around 2 g/L in mature milk [27]. Interest-
ingly, another antimicrobial protein, lysozyme, which is present at lower levels in human

milk [3], is also relatively depleted in cattle [28], with concentrations of lysozyme in *Bovidae* reported to be as low as 1/1000th of other mammalian species [29]. Although lactoferrin expression can be induced by mastitic infection [30] it is also under partial genetic control [31], and genetic variants affecting expression have been identified at the *LTF* locus [32]. QTL at this locus have also been associated with casein number and lactose concentration [33].

A third antibacterial protein present in milk is lactoperoxidase, encoded by the *LPO* gene on BTA19. In contrast to lysozyme and lactoferrin, lactoperoxidase activity has been observed at around 20× higher in bovine milk compared to that of humans [34]. A *trans*-eQTL for *LPO*, overlapping a QTL for milk protein concentration, has been identified on BTA20 at the locus of the *C6* and *C7* genes [35]: these two genes encode proteins that comprise part of the complement pathway in the innate immune system, suggesting that lactoperoxidase may be co-regulated with this system.

2.2. Fat synthesis pathways

Fat is one of the major components of milk, forming membrane-bound droplets known as milk fat globules (MFG), mostly in the form of triglycerides. MFG membranes also contain a range of proteins, such as butyrophilin, adipophilin, mucin, lactadherin, lactoferrin, and xanthine oxidase [36]. Two of these, butyrophilin (*BTN1A1*) and xanthine oxidase (*XOR*), are required for enveloping the MFGs with the apical cell membrane and therefore for secretion of the MFG into milk [37]. Heterozygous knockout mice for *XOR* are unable to maintain lactation [38]. Many of the QTL for fat yield or milk concentration in cattle map to genes encoding fat synthesis or metabolic enzymes. One of the most prominent QTL detected in cattle [39] for milk volume, fat, and protein phenotypes in cattle maps to the *DGAT1* locus on BTA14 [40]. This gene encodes the enzyme diacylglycerol O-acyltransferase 1, responsible for the final stage triglyceride fat synthesis [41]. The causative variant for this QTL is a non-conservative lysine to alanine substitution at position 232 [40]; more recently, this same variant has been shown to cause an expression QTL for the *DGAT1* gene by disrupting an exon splice enhancer, which in turn alters the splicing efficiency of several introns in the transcript [42]. Another enzyme sitting earlier in the same triglyceride synthesis pathway is glycerol-3-phosphate acyltransferase 4, encoded by the *GPAT4* gene (formerly known as *AGPAT6*) on BTA27. Like *DGAT1*, a highly pleiotropic QTL has been detected at this locus for many milk phenotypes, including volume, fat, protein, and lactose traits [43]. The enzyme lipin 1 sits in a related pathway, catalysing the conversion of phosphatidate into diacylglycerol, as well as functioning as a transcriptional coactivator for genes involved in fatty acid oxidation [44]. QTL for milk protein and casein concentrations [33], as well as milk yield [45], have been detected at the *LPIN1* locus, which encodes this enzyme, on BTA11 in cattle.

Before they can be assembled into triglycerides, fatty acids first need to be obtained either from the diet or from *de novo* synthesis. Several genes involved in this latter pathway have also shown QTL for milk-related phenotypes. The rate-limiting step in fatty acid synthesis is the carboxylation of acetyl-CoA to malonyl-CoA, catalysed by the enzyme Acetyl-CoA carboxylase (ACC). The alpha form of this enzyme is encoded by the gene *ACACA* on BTA19, and QTL for fatty acid composition [46] and somatic cell score (a proxy phenotype for mastitis) [33] have been mapped to this locus. Another important enzyme in the fatty acid synthesis pathway is fatty acid synthase (FAS), encoded by the gene *FASN* on BTA19, dimers of which are responsible for synthesising the saturated C16 fatty acid palmitic acid from acetyl-CoA and malonyl-CoA [47]. QTL attributed to this gene have been identified for milk fat yield [7], fat concentration [8,48], and fatty acid composition [49] in cattle.

The fatty acid synthase gene will produce only saturated fatty acids. To generate monounsaturated or polyunsaturated fatty acids, desaturase enzymes are required. One cluster of fatty acid desaturase genes on BTA29 includes the two gene *FADS1* and *FADS2*; these enzymes are responsible for the final, rate-limiting step in omega-3 and -6 fatty acid

synthesis [50,51]. Variants mapping in cattle to these two genes have been associated with concentrations of a range of polyunsaturated fatty acids in milk [52], and similar associations are also observed in human milk [53,54]. A third desaturase enzyme involved in fatty acid synthesis is stearoyl-CoA desaturase, encoded by the gene *SCD* on BTA26 and responsible for oxidising the C16 and C18 saturated fatty acid compounds palmitoyl- and stearoyl-CoA into the monounsaturated compounds palmitoleoyl- and oleoyl-CoA respectively [55]. This enzyme is also important in regulating metabolism: knockout mice, as well as exhibiting lower levels of tissue triglycerides and low-density lipoproteins [56], also show an increase in insulin signalling and glucose uptake in muscle tissue [57]. In cattle, a QTL assigned to *SCD* has been identified for milk fat yield [7,48].

2.3. Hormones and signalling

Hormones, and the receptors and signalling pathways they activate, are important in most if not all biological functions, and milk production is no exception. For example, knocking out the *SCD* gene discussed in the previous section leads to an increase in tyrosine phosphorylation of the insulin receptor, with downstream effects on the PI3K-Akt signalling pathway, leading to increased levels of the glucose transporter GLUT4 in the plasma membrane and increased glucose uptake in muscle [57]. These knockout mice also showed lower levels of the hormone leptin, which is involved in regulating energy intake and partitioning. In cattle, leptin (encoded by the gene *LEP* on BTA4) has been associated with both milk yield and feed intake [58]. No milk phenotype QTL have been reported for the leptin receptor (*LEPR* on BTA3), though an association has been reported with body size [59]. In both humans [60] and mice [61], leptin resistance is associated with obesity. Another gene associated with obesity [62] is *FTO*, encoding the enzyme FTO alpha-ketoglutarate dependent dioxygenase. This enzyme is involved in DNA repair; specifically, it demethylates 3-methylthymidine [63]. It can also demethylate bases in RNA, including 3-methyluracil and 6-methyladenosine [64]. Via this latter mechanism, FTO can inhibit adipogenesis by demethylating cyclin A2 and cyclin-dependent kinase 2 mRNA, reducing expression of these genes and thereby prolonging the cell cycle [65]. In cattle, variants at the *FTO* locus on BTA18 have been associated with milk fat yield [66].

One hormone of particular importance for lactation is prolactin, a peptide hormone secreted by the anterior pituitary gland. In cattle, this peptide is encoded by the *PRL* gene on BTA23. The importance of this hormone in cattle was underlined by the discovery of a dominant missense mutation that caused, among other phenotypes, a failure to lactate [67]. Prolactin promotes development of the mammary gland during pregnancy in conjunction with progesterone to generate ductal branching and alveolar buds [68]. Prolactin is detected by cells using prolactin receptor (*PRLR*; BTA20) and acts via the *PRLR*/*JAK2*/*STAT5* signalling pathway to promote mammary gland development and milk protein expression [69], as well as inducing expression of the enzyme UDP-glucose pyrophosphorylase 2 (*UGP2*) and the transporter UDP-galactose transporter 2 (*SLC35A2*), thereby promoting the synthesis of lactose [70]. In parallel, prolactin receptor also acts via the *PI3K*/*Akt* pathway to downregulate repressors of *PRLR*/*JAK2*/*STAT5* signalling [71], with *Akt1* also upregulating fat synthesis and glucose uptake into the cell for lactose production [72,73]. More recently, it has been shown that *Akt* signalling induces developing mammary epithelial cells to express prolactin, which in turn acts in an autocrine manner via *STAT5* to cause terminal differentiation of the mammary epithelium [69,74]. Given the importance of these pathways, it is not surprising that QTL for milk yield have been widely identified at the *PRLR* locus [33,75,76], as well as for somatic cell score [76]. Likewise, many studies have reported milk related QTL at the *STAT5* locus on BTA19 (genes *STAT5A* and *STAT5B*) [7,22,23,49,77].

Beyond the leptin and prolactin pathways, other hormones have also been linked to lactogenesis. One example is growth hormone, also known as somatotropin, which comprises part of the same family of protein hormones as prolactin [78]. Growth hormone acts in the lactating animal to partition nutrients and energy towards the mammary gland

[79], though to be mediated by increased serum insulin-like growth factor (IGF-1) levels [80]. In the cow, growth hormone is encoded by the gene *GH1* on BTA19, and its receptor, encoded by *GHR*, maps to BTA20. Both the *GH1* locus [33], and especially the *GHR* locus [8,58,81–83], have been associated with milk, fat, and protein yield phenotypes in a range of cattle populations.

Another important family of signalling molecules is the interleukin family of cytokines, a group of primarily immunomodulatory proteins that operate in a paracrine or autocrine manner to affect cell growth and differentiation during immune responses [84]. The gene *CSF2RB* on BTA5 encodes the β -chain of GM-CSF (also known as CSF2) receptor, and also forms a common subunit with the receptors for interleukin-3 (IL-3) and IL-5. QTL for milk, fat, and protein volume, as well as fat and protein concentrations, have been identified at this locus [7,20,85,86], with *CSF2RB* suggested as the most likely candidate causative gene, although the neighbouring genes *NCF4* [85,86] and *TST* [8] have also been put forward as candidates. The GM-CSF receptor operates via the JAK/STAT signalling pathway, specifically, JAK2 and STAT5 [84], the same pathway as used by prolactin signalling.

In contrast to STAT5, which is linked to increased milk protein and lactose expression, STAT3 has been identified as a mediator of involution and apoptosis in the mammary gland [87], primarily activated by the cytokine leukaemia inhibitory factor (LIF) [88,89]. STAT3 is also involved in leptin signalling, interacting with the long form of the leptin receptor [90]. Mice where either *STAT3* or *LIF* are knocked out show delayed involution and reduced levels of apoptosis [88], raising the possibility of improving efficiency in dairy herds, especially those milking once-a-day, by breeding for animals lacking one or more of these genes. It is difficult to assign QTL unambiguously to the *STAT3* locus, as it maps on BTA directly between the two *STAT5* genes *STAT5A* and *STAT5B*. Yet another STAT protein, *STAT1*, is believed to be involved in the development of the mammary gland [91]. A QTL mapping to the *STAT1* locus on BTA2 has been associated with milk, fat, and protein yield [91].

One group of proteins that is upregulated by the JAK/STAT pathway is the suppressors of cytokine signalling (SOCS) gene family [92]. Proteins in this family down-regulate JAK/STAT-mediated signalling. For example, *SOCS1* expression is induced by prolactin signalling, and in turn binds to JAK2, inhibiting its association with STAT5 and dampening signal transmission [68,93]. *SOCS2* negatively regulates GH signalling, with knockout mice showing increased weight and skeletal dimensions [94]. Because SOCS proteins ultimately downregulate milk protein expression via prolactin and growth hormone signalling pathways, there is the potential that knocking the genes coding them out could improve lactation phenotypes in dairy animals. For example, mice with homozygous *Socs1* knockout genotypes show enhanced alveolar development [93], and a point mutation in the ovine *Socs2* gene has been associated with higher milk production in dairy sheep, albeit with increased susceptibility to mastitis [95]. In cattle, several different SOCS genes have been associated with milk volume, fat, and protein phenotypes [92].

2.4. Transporters and ion channels

Another important mechanism affecting milk production involves trans-membrane transport and ion channels. All major milk components need to either be produced within the mammary epithelial cells or transported across them from the blood, and in both cases need to cross the apical membrane into the lumen. While some small molecules like urea can simply diffuse across the membrane, in most cases either passive or active transport channels are required.

The volume of water excreted into the milk is driven by osmotic pressure, in turn created by exporting lactose and ions across the cell membrane against their concentration gradients. This means that QTL for milk phenotypes frequently map to genes encoding transporters. One important group is the sugar transporters. The gene *SLC37A1* (on BTA1) encodes a glucose-6-phosphate : inorganic phosphate antiporter [96], responsible for importing glucose into the cell, and QTL for milk volume have been detected at this

locus in several populations [20,85,97]. Another glucose transporter, SWEET1, is encoded by the gene *SLC50A1* on BTA3. SWEET1 has been observed in the Golgi in mammary cells, and is possibly responsible for importing glucose into the Golgi for lactose synthesis [98]. QTL for lactose concentration [23] and protein concentration [85] have been identified near this gene, though the latter QTL has been assigned to *GBA1*, which encodes the lysosomal protein glucosylceramidase beta 1. Other glucose transporters have also been implicated in milk production, such as GLUT1 (*SLC2A1* on BTA3) [99], the Na⁺/glucose co-transporter SGLT1 (*SLC5A1* on BTA17) [100], and GLUT12 (*SLC2A12* on BTA9) [100].

While the mammary epithelial cells are using osmotic pressure to secrete milk, it is important that they maintain their own cell volumes correctly. One mechanism by which they can do this is using voltage-regulated anion channels (VRACs), which help regulate cell volume by exporting Cl⁻ ions, as well as small organic anions such as taurine [101,102]. VRACs are comprised of heteromers of leucine-rich repeat containing 8 (LRRC8) proteins A to E, encoded by the genes *LRRC8A–LRRC8E*. In cattle, the locus on BTA3 containing *LRRC8B–LRRC8D* has been associated with milk lactose concentration [23,77], with *LRRC8C* suggested as the likely causative gene on the basis of gene expression data. Interestingly, *LRRC8C* has been associated with adipocyte differentiation (under the name *fad158*) [103], and is also present in the membranes of MFGs [104], suggesting it may also have a role in the storage or export of fat.

As stated above, it is important that water be able to move across the cell membrane to osmotic balance as milk is expressed. However, the lipid bilayer of the cell membrane is impermeable to water, requiring channels to facilitate the crossing. These channels are provided by aquaporin proteins (AQPs). At least seven aquaporins are expressed in the mammary gland [105,106], where they are believed to play a role in gland development, and in transporting water to the lactating gland for milk synthesis and secretion. For example, *AQP1* is expressed in the capillary endothelial cells, and may be involved in oestrogen-mediated angiogenesis in the developing mammary gland [107], while *AQP5* is expressed in mammary epithelial cells, and the protein pores are moved from the cytoplasm to the apical cell membrane under the regulation of prolactin [108], suggesting a role in milk production. Genetic effects mapping to aquaporin genes have been reported, such as a study in sheep [109] that mapped QTL for milk fat and protein concentrations to a window on OAR3 that contains the genes *AQP2*, *AQP5*, and *AQP6*, albeit alongside *LALBA*, which also presents a strong candidate causal gene for milk-related traits.

Another important class of transporter is the potassium channel, a type of widely expressed ion channel found in the majority of cell types. A number of different families of potassium channel have been identified. The largest family, the voltage-gated potassium channels (K_v, reviewed in [110]), responds to voltage changes in the cell's membrane potential. One channel, K_v3.3 is encoded by the gene *KCNC3* on BTA18, has been associated with milk yield in cattle [22]. Another QTL for both milk volume and fat yield maps to the gene *KCNS2*, encoding K_v9.2, on BTA14 [83]. A third gene in the same family, *KCNH4*, encodes the channel K_v12.3, and is a potential candidate for a lactose concentration QTL on BTA19 [23], although the QTL also encompasses the genes *STAT3*, *STAT5A*, and *STAT5B*, which are also candidates. A second large family is the potassium inwardly rectifying channel family (K_{ir}, reviewed in [111]), comprising lipid-gated channels that are activated by phosphatidylinositol 4,5-bisphosphate (PIP₂). In cattle, a locus on BTA19 has been associated with QTL for fat and protein concentrations [22], lactose concentration [23,77], and milk yield [7]. This locus contains two genes encoding K_{ir} channels: *KCNJ2* (K_{ir}2.1) and *KCNJ16* (K_{ir}5.1), and both genes have been proposed as candidate causatives underlying the QTL. A third, smaller, family of channels is the two-pore-domain potassium channels (K_{2p}), known as leak channels [112]. A QTL for milk fat, protein, and volume has been identified on BTA26 at the locus of the *KCNK18* gene, encoding the potassium channel K_{2p}18.1 [83]. A fourth family is the calcium- and sodium-activated potassium channels [113], of which the most well-known member is K_{Ca}1.1, also known as BK. The activity of

K_{Ca}1.1 is modulated by auxiliary β and γ subunits [114,115], including γ 1, encoded by the gene *LRRC26* on BTA11, where a QTL for fat concentration has been detected [22].

The channels discussed so far are limited to moving solutes along an existing concentration gradient. Another class of transporter requires energy in the form of ATP to concentrate solutes to establish a gradient or membrane potential. Many of these belong to the ATP binding cassette family (ABC). One important example from this family is ATP binding cassette subfamily G member 2 (*ABCG2*). Initially identified as a xenobiotic drug transporter [116], *ABCG2* also functions as a riboflavin transporter into the milk [117] and a urate transporter in the kidneys [118]. QTL for several milk phenotypes have been mapped to the *ABCG2* locus on BTA6, including milk yield [7,20,83,119], fat and protein concentration [8,77,119], lactose concentration [23,77] and α _{S1}-CN concentration [97]. The causative variant at this QTL is generally believed to be a tyrosine to serine substitution (Y581S), identified by Cohen-Zinder et al. [119]; however the adjacent gene *SPP1*, encoding the protein osteopontin (involved in bone remodelling), has also been proposed as a candidate [120,121]. A second member of the ABC family is *SUR2* (encoded by *ABCC9* on BTA5), which forms a component of the ATP-sensitive potassium channel K_{ATP}, alongside *SUR1* (*ABCC8*) and the inward rectifying channels K_{ir}6.1 (*KCNJ8*) and K_{ir}6.2 (*KCNJ11*) [122], with the exact composition varying by tissue [123]. The channel is inhibited by ATP and activated by MgADP. The K_{ATP} transporter is important for glucose level sensing to control insulin release in pancreatic beta cells [124]: at low glucose levels, ATP levels are low and ADP levels raised, so the channel is open, while at high glucose levels, ATP closes the channel. This polarises the plasma membrane, a change that is detected by voltage-gated calcium channels, which then open, causing calcium to enter the cell and trigger the release of insulin. In cattle, a QTL for milk fat yield has been identified at the *ABCC9* locus [22,82]. A third ATP-binding transporter is the calcium transporter *SERCA2*, encoded by the gene *ATP2A2* on BTA17. This transporter pumps Ca²⁺ from the cytosol into the endoplasmic reticulum, whence it can be exported into milk [125]. A QTL at this locus has been detected for milk and protein yield, as well as for milk calcium [35].

Another ion transporter, showing widely reported associations with milk phenotypes, is inorganic pyrophosphate transport regulator, encoded by the gene *ANKH* on BTA20. The *ANKH* transporter controls extracellular mineralisation by regulating the levels of inorganic pyrophosphate in the extracellular matrix. In humans and mice, mutations in this transporter have been associated with arthritis and bone growth disorders linked to tissue calcification [126,127]. This transporter has previously been assumed to be a pyrophosphate transporter; however, recent work has shown that the channel in fact transports ATP, and that the production of extracellular pyrophosphate from ATP requires the enzyme ectonucleotide pyrophosphatase/phosphodiesterase 1 (ENPP1) [128]. In cattle, QTL mapping to the *ANKH* locus have been identified for milk yield [7,20], milk lactose concentration [23,77], and α -lactalbumin concentration [97]. Another recent study [129] discovered that the ANK protein cycles between the plasma membrane and trans-Golgi network using clathrin-coated vesicles mediated by clathrin adaptors AP1 and AP2. The phosphatidylinositol binding clathrin assembly protein also interacts with AP2 [130], binding to the signalling molecule phosphatidylinositol and recruiting the AP2 complex to form clathrin-coated pits. This protein is encoded by the gene *PICALM* on BTA29, and is associated with QTL for α _{S1}-CN [97] and lactose concentration [23,77].

3. Identifying candidate causative genes

The previous section may have given the impression that identifying the causative genes or even variants underlying QTL is easy. However, in many cases, there will be either no obvious candidate genes under the QTL, possibly because the causative variant sits in a long-range regulatory element for a distant gene. In other cases, there may be many potential candidate genes, with different selection methods highlighting different candidates. For example, the window from 42.2 to 42.5 Mbp on BTA19 envelopes several candidate genes for milk phenotypes. A 2015 study by Raven et al. [85] highlighted the

genes *GHDC*, *STAT5A*, and *STAT5B* on the basis of differential expression and enrichment of significant variants, and proposed *STAT5A* as causative on the basis of knockout studies in mice. In contrast, later work by our group [23,77] used eQTL data and missense variants in strong LD with the top QTL variant to highlight the genes *DHX58*, *GHDC*, *KAT2A*, *KCNH4*, *STAT5A*, and *STAT5B*, with the two *STAT5* genes again proposed as the most likely candidates. As discussed above, potassium channels such as that encoded by *KCNH4* have been associated with milk at a number of loci, while the histone deacetylase and transcription activator *KAT2A2* is also a possible candidate, based on the high levels of gene expression required in the lactating mammary gland. It is possible that multiple QTL are segregating at this locus, and therefore that more than one gene is causative.

Another region with two strong candidate causal genes maps between 15.3 and 15.6 Mbp on BTA3. At this locus, Raven et al. [85] identified a QTL for milk protein concentration, and proposed the epithelial mucin gene *MUC1* as the best candidate. The mucin 1 protein coded by this gene is considered a “metabolic master regulator” [131], regulating tyrosine-kinase signalling and the expression of metabolic genes. Work by our group [23,77] as identified a QTL for milk lactose concentration at the same locus, and proposed the gene *SLC50A1*, encoding a sugar transporter. Again, it is possible that these are two separate QTL, and that both genes are causative for the corresponding phenotypes; however, another possibility is that the QTL is pleiotropic, and only one QTL is present controlling both phenotypes. Distinguishing between these two possibilities is difficult or impossible using purely statistical or bioinformatic means, and additional experiments are likely to be required.

3.1. Molecular phenotypes

Widely used phenotypes such as fat or protein are effectively aggregating signals from a large number of milk components: individual fatty acids and other lipids for fat, and a number of casein and whey proteins for protein, for example. It is likely that these different components are not all under the same genetic regulation, and some correlations may even be negative. For example, Stoop et al [132] observed genetic correlations of as low as -0.84 between milk fatty acid measures (C14:0 and C16:0), as well as correlations over 0.9. This suggests that using finer composition measures, such as individual fatty acids or proteins, should give cleaner, and possibly a larger number of, genetic signals compared to complex phenotypes such as fat or protein. This approach has been successfully applied to use milk minerals and individual protein measurements to identify novel QTL and to highlight *SLC37A1* as involved in lactation [35]. These phenotypes can be considered “molecular”, as measurements for a single molecule, for example using gas or liquid chromatography with mass spectroscopy, are providing the phenotype.

Molecular phenotypes in milk are often measured using Fourier-transform mid-infrared spectroscopy, as this method is commonly used for commercial herd testing, and is cheaper than to perform on a large scale than spectroscopy. This method uses the absorbance of around 900 different frequencies of infrared light (known as wavenumbers), then uses these values in a model to predict the phenotypes of interest. In commercial herd testing, the predicted phenotypes are typically fat, protein, and lactose concentrations; however, models have been developed for a range of other phenotypes, such as individual fatty acids [133] and proteins [134], as well as phenotypes more remote to milk such as methane production [135] and fertility [136]. However, because FT-MIR phenotypes are predicted rather than measured, their utility for QTL discovery can be variable. For example, our recent work [137] showed that both HPLC measured and FT-MIR predicted phenotypes could detect significant *cis*-QTL for α -casein, κ -casein, and β -lactoglobulin. However, FT-MIR was unable to identify the highly significant *cis*-QTL for lactoferrin that was identified using HPLC, and identified a QTL at the locus of the fat synthesis gene *DGAT1* for α -casein; this latter may be due to signal crossover from fat, which is highly correlated with protein in milk. This form of crossover, if present, would complicate the task of identifying pleiotropy.

In addition to predicting concentrations of milk components, it is also possible to use the individual wavenumber data itself as phenotypes [77]: different wavenumbers represent different chemical bonds that absorb MIR light at the corresponding frequencies, so wavenumber phenotypes are effectively measuring the concentrations in milk of compounds that contain specific chemical bonds. Wavenumber phenotypes have been shown to detect stronger and more numerous genetic signals compared to the predicted phenotypes such as fat and protein [77].

Examining the expression level of each gene in a relevant tissue sample provides an alternative to measuring or estimating protein concentrations. In milk, for example, expression levels of the genes encoding milk proteins in the lactating mammary epithelial tissue can give a proxy measurement for milk proteins. Traditionally, this could be done using techniques such as qPCR or expression microarrays; more recently, using RNA sequencing (RNA-seq) to sequence and count (relatively) mRNA molecules has become common. Because RNA-seq captures data for every expressed gene, it is possible to search for expression QTL (eQTL), as well as allele-specific expression (ASE), for every gene, including those not expressed in the milk, such as fat synthesis enzymes and hormone receptors. The presence of a QTL for a given gene can provide evidence to the causality (or otherwise) of the gene at an overlapping QTL: when the QTL is caused by an underlying eQTL, we expect that the causal variant or variants will be strongly associated with both the QTL and eQTL, with associations for other variants decaying away in proportion to linkage disequilibrium with the causal variants. This pattern can be identified by examining the correlations across variants between the QTL and eQTL variant effects [138] or p-values (on a logarithmic scale) [23,43], or correlations between local genomic estimated breeding values (GEBVs) and gene expression [139]. Other methods used to associate eQTL with GWAS results include Transcriptome-Wide Association Scan (TWAS) [140], Mendelian Randomisation [141,142] and Bayesian colocalisation methods [139,143].

Molecular phenotypes and their QTL have assisted in highlighting candidate causative genes in cases where no obvious candidates existed. For example, the gene *MGST1* on BTA5 encodes microsomal glutathione S-transferase 1, belonging to a family of detoxification enzymes [144], with no obvious role in milk production. Nevertheless, QTL for milk traits mapping to this locus have been reported in many different studies [20,85,97,145,146], with the neighbouring gene *EPS8* (Epidermal growth factor receptor kinase substrate 8) sometimes proposed as a candidate. However, gene expression data from lactating bovine mammary tissue have shown that the milk QTL at this locus co-segregate with an eQTL for *MGST1*, whereas *EPS8* is barely expressed in this tissue. Similarly, a QTL for milk fat percentage on the distal end of BTA11 has been linked to the gene encoding the ABO blood group, also named *ABO* (alpha 1-3-N-acetylgalactosaminyltransferase and alpha 1-3-galactosyltransferase), using RNA-seq data to show that a splice donor site mutation causes aberrant splicing of the *ABO* transcript. in turn causing an eQTL that co-segregates with the fat percentage QTL [77].

3.2. Chromatin structure phenotypes

Chromatin is the name given to the compound comprising DNA wound around nucleosomes, themselves composed of histone proteins. The three-dimensional folded structure of chromatin can have an important impact on gene expression. On a proximal scale, the structure of the chromatin region surrounding a gene can be in open or closed configurations, with the former providing access for transcription factors and RNA polymerase components to reach the promoter region and trigger gene expression, while the latter, closed, configuration blocks expression. Variants sitting within an open chromatin region in the appropriate cell type are more likely to have a regulatory impact on gene expression [138], and therefore are more likely to present good candidate causal variants for a given trait, compared to other non-coding variants.

One method commonly used to study chromatin state is chromatin immunoprecipitation (ChIP) [147], which can be used to identify transcription factor binding sites or

histone modifications. Histone proteins feature long tails, and a wide variety of post-translational modifications can be made to these tails to alter the chromatin state. For example, the modifications H3K4me2 (histone 3, lysine 4 dimethylation), H3K4Me3, and K3K27ac (lysine 27 acetylation) are associated with open chromatin in active promoters and enhancers, while closed, repressed regions can be marked by H3K9me2 or H3K27me3 [148]. More recent versions of ChIP, using DNA sequencing techniques to target the whole genome in a single experiment, include ChIP-seq [149,150] and CUT&RUN-seq [151]. Recent work [138] has identified *cis*-QTL for ChIP-seq phenotypes, and demonstrated that these QTL frequently exhibit strong correlation with nearby eQTL, and that eQTL tag variants are significantly enriched within ChIP-seq identified open chromatin windows; this illustrates the utility of of this type of data for understanding regulatory variation in the genome.

Open chromatin is accessible to transcription factors and other proteins required for gene expression. However, it is also more accessible than other regions to nuclease enzymes, and this fact is used in another approach to identifying open chromatin regions. One commonly-used nuclease is DNase 1, which can be used in conjunction with high-throughput sequencing in the technique DNase-seq [152]. The hypersensitivity sites identified using this technique highlight the positions of regulatory elements such as promoters, silencers, and enhancers. A similar technique, using the enzyme micrococcal nuclease, is called MNase-seq [153,154], and identifies the positions of nucleosomes. As nucleosomes are scarcer in open chromatin, these data give an inverse signal to hypersensitivity sites identified by DNase-seq. A more recent technique called Assay for Transposase Accessible Chromatin (ATAC-seq) [155] uses a modified hyperactive transposase enzyme Tn5 to fragment and load sequencing adaptors to open chromatin regions. ATAC-seq typically requires fewer cells and less time to perform compared to DNase-seq, although single-cell protocols have now been developed for both methods [156,157].

On a larger scale, the 3D folded structure allows distal enhancer and silencer elements to enter close contact with the gene to perform their respective functions on gene expression [158]. Identifying this structure allows linkages between genes and distal regulatory elements to be identified. These sorts of long-range interactions can be studied using chromosome conformation capture (3C) [159], and related methods such as 3C-on-chip (4C) [160] and 3C carbon copy (5C) [161]. These older techniques can study only small parts of the genome. The more recent technique Hi-C [162] could originally study the structure of the whole genome only at low resolution ($\approx 1\text{Mbp}$), but later refinements of the method [163,164] have allowed for improved resolution and signal-to-noise ratios.

3.3. From candidate genes to causative variants

The results of GWAS and similar experiments are typically genomic intervals containing several associated variants that are in strong LD with one another. Although candidate causal genes may be selected using known pathways, gene expression data, or chromatin structure as discussed above, for some applications it is useful or necessary to know the causal variant underlying the QTL. These applications include improving the accuracy of genomic selection, performing genetic testing, or creating gene-edited animals [165]. However, identifying the causal variant using purely statistical or bioinformatics approaches is typically not possible, as variants in strong LD can not be distinguished from one other statistically.

In some cases, variant annotation (using tools such as SnpEff [166] or Ensembl's Variant Effect Predictor [167]) can highlight missense or nonsense mutations that will make stronger candidates [39], but many QTL are driven by regulatory effects rather than coding ones. This means that other tools will be needed to resolve QTL with underlying regulatory effects. One commonly used technique for discovering *cis*-regulatory elements (CREs) is transcription factor binding site (TFBS) prediction. Several tools have been developed for this, generally using information from databases containing binding site motifs for large numbers of transcription factors, such as TRANSFAC [168] and JASPAR [169]. A second

approach to investigating regulatory effects is to use a reporter assay, where the effects of putative promoter variants can be tested against the expression of a reporter gene such as GFP. This method has recently been scaled up to test thousands of variants simultaneously using a massively parallel reporter assay (MPRA) [170]. A third method is to use CRISPR-Cas9 or other gene editing technologies to test putative regulatory sequences, either by deleting them using NHEJ, or by editing in alleles of interest using HDR, then observing the resulting effect on gene expression, often using single-cell RNA-seq. Methods incorporating CRISPR include Perturb-seq [171], CROP-seq [172], and HCR-Flowfish [173].

4. Conclusions

Over the last twenty years since the seminal work of Grisart et al [40] showed that the *DGAT1* gene underlies the large milk QTL on BTA14, a large number of milk related QTL have been identified in cattle, with candidate causative genes proposed for many of them. In coming years, we can expect that sequence-resolution data sets will continue both to grow, and to diversify to include additional breeds from around the world. These larger, more diverse, data sets will likely empower the discovery of many novel QTL. Although proving the causality of genes and variants underlying these QTL will likely remain difficult, the use of molecular phenotypes, massively parallel reporter assays, and CRISPR will extend the list of proven causatives. This new information should provide greater insight into the genes and pathways underlying the initiation and maintenance of lactation in mammals.

Institutional Review Board Statement: Not applicable

Conflicts of Interest: TL is an employee of Livestock Improvement Corporation, a commercial provider of bovine germplasm. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

BTA	<i>Bos taurus</i> chromosome
CRISPR	Clustered regularly interspaced short palindromic repeats
eQTL	Expression quantitative trait locus
FT-MIR	Fourier-transform mid-infrared
GM-CSF	Granulocyte-macrophage colony-stimulating factor
HDR	Homology-directed repair
HPLC	High-performance liquid chromatography
IL	Interleukin
LD	Linkage disequilibrium
Mbp	Million base pairs
MFG	Milk fat globules
MPRA	Massively parallel reporter assay
NHEJ	Non-homologous end joining
OAR	<i>Ovis aries</i> chromosome
qPCR	Quantitative polymerase chain reaction
QTL	Quantitative trait locus
TFBS	Transcription factor binding site
TWAS	Transcriptome-Wide association scan

References

1.

Jenness, R. Composition of Milk. In *Fundamentals of Dairy Chemistry*; Wong, N.P.; Jenness, R.; Keeney, M.; Marth, E.H., Eds.; Van Nostrand Reinhold Company: Boston, MA, USA, 1988; chapter 1, pp. 1–38. https://doi.org/10.1007/978-1-4615-7050-9_1.

545546

2.

Oftedal, O.T.; Iverson, S.J. Comparative Analysis of Nonhuman Milks. In *Handbook of Milk Composition*; Jensen, R.G., Ed.; Academic Press: San Diego, California, USA, 1995; chapter 10, pp. 749–789.

548549

3. Vegarud, G.E.; Langsrud, T.; Svenning, C. Mineral-binding milk proteins and peptides; occurrence, biochemical and technological characteristics. *British Journal of Nutrition* **2000**, *84*, 91–98. <https://doi.org/10.1017/S0007114500002300>. 550

4. Shekar, P.C.; Goel, S.; Rani, S.D.S.; Sarathi, D.P.; Alex, J.L.; Singh, S.; Kumar, S. κ -Casein-deficient mice fail to lactate. *Proceedings of the National Academy of Sciences* **2006**, *103*, 8000–8005. <https://doi.org/10.1073/pnas.0601611103>. 551

5. Caroli, A.; Chessa, S.; Erhardt, G. Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. *Journal of dairy science* **2009**, *92*, 5335–5352. <https://doi.org/10.3168/jds.2009-2461>. 552

6. Nilsen, H.; Olsen, H.G.; Hayes, B.; Sehested, E.; Svendsen, M.; Nome, T.; Meuwissen, T.; Lien, S. Casein haplotypes and their association with milk production traits in Norwegian Red cattle. *Genetics Selection Evolution* **2009**, *41*, 1–12. <https://doi.org/10.1186/1297-9686-41-24>. 553

7. Kemper, K.E.; Reich, C.M.; Bowman, P.J.; Vander Jagt, C.J.; Chamberlain, A.J.; Mason, B.A.; Hayes, B.J.; Goddard, M.E. Improved precision of QTL mapping using a nonlinear Bayesian method in a multi-breed population leads to greater accuracy of across-breed genomic predictions. *Genetics Selection Evolution* **2015**, *47*, 29. <https://doi.org/10.1186/s12711-014-0074-4>. 554

8. Pausch, H.; Emmerling, R.; Gredler-Grandl, B.; Fries, R.; Daetwyler, H.D.; Goddard, M.E. Meta-analysis of sequence-based association studies across three cattle breeds reveals 25 QTL for fat and protein percentages in milk at nucleotide resolution. *BMC genomics* **2017**, *18*, 853. <https://doi.org/10.1186/s12864-017-4263-8>. 555

9. MacLeod, I.; Bowman, P.; Vander Jagt, C.; Haile-Mariam, M.; Kemper, K.; Chamberlain, A.; Schrooten, C.; Hayes, B.; Goddard, M. Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC genomics* **2016**, *17*, 1–21. <https://doi.org/10.1186/s12864-016-2443-6>. 556

10. Olsen, H.G.; Knutsen, T.M.; Lewandowska-Sabat, A.M.; Grove, H.; Nome, T.; Svendsen, M.; Arnyasi, M.; Sodeland, M.; Sundsaasen, K.K.; Dahl, S.R.; et al. Fine mapping of a QTL on bovine chromosome 6 using imputed full sequence data suggests a key role for the *group-specific component* (GC) gene in clinical mastitis and milk production. *Genetics Selection Evolution* **2016**, *48*, 1–16. <https://doi.org/10.1186/s12711-016-0257-2>. 557

11. Vordenbäumen, S.; Braukmann, A.; Petermann, K.; Scharf, A.; Bleck, E.; von Mikecz, A.; Jose, J.; Schneider, M. Casein α s1 is expressed by human monocytes and upregulates the production of GM-CSF via p38 MAPK. *The Journal of Immunology* **2011**, *186*, 592–601. <https://doi.org/10.4049/jimmunol.1001461>. 558

12. Vordenbäumen, S.; Braukmann, A.; Altendorfer, I.; Bleck, E.; Jose, J.; Schneider, M. Human casein alpha s1 (CSN1S1) skews *in vitro* differentiation of monocytes towards macrophages. *BMC immunology* **2013**, *14*, 1–11. 559

13. Bovenhuis, H.; Weller, J.I. Mapping and analysis of dairy cattle quantitative trait loci by maximum likelihood methodology using milk protein genes as genetic markers. *Genetics* **1994**, *137*, 267–280. <https://doi.org/10.1093/genetics/137.1.267>. 560

14. Berry, S.; Lopez-Villalobos, N.; Beattie, E.; Davis, S.; Adams, L.; Thomas, N.; Ankersmit-Udy, A.; Stanfield, A.; Lehnert, K.; Ward, H.; et al. Mapping a quantitative trait locus for the concentration of β -lactoglobulin in milk, and the effect of β -lactoglobulin genetic variants on the composition of milk from Holstein-Friesian x Jersey crossbred cows. *New Zealand veterinary journal* **2010**, *58*, 1–5. <https://doi.org/10.1080/00480169.2010.65053>. 561

15. Farrell Jr, H.; Jimenez-Flores, R.; Bleck, G.; Brown, E.; Butler, J.; Creamer, L.; Hicks, C.; Hollar, C.; Ng-Kwai-Hang, K.; Swaisgood, H. Nomenclature of the proteins of cows' milk—Sixth revision. *Journal of dairy science* **2004**, *87*, 1641–1674. [https://doi.org/10.3168/jds.S0022-0302\(04\)73319-6](https://doi.org/10.3168/jds.S0022-0302(04)73319-6). 562

16. Bedere, N.; Bovenhuis, H. Characterizing a region on BTA11 affecting β -lactoglobulin content of milk using high-density genotyping and haplotype grouping. *BMC genetics* **2017**, *18*, 1–9. <https://doi.org/10.1186/s12863-017-0483-9>. 563

17. Braunschweig, M.; Leeb, T. Aberrant low expression level of bovine β -lactoglobulin is associated with a C to A transversion in the *BLG* promoter region. *Journal of dairy science* **2006**, *89*, 4414–4419. [https://doi.org/10.3168/jds.S0022-0302\(06\)72488-2](https://doi.org/10.3168/jds.S0022-0302(06)72488-2). 564

18. Davis, S.R.; Ward, H.E.; Kelly, V.; Palmer, D.; Ankersmit-Udy, A.E.; Lopdell, T.J.; Berry, S.D.; Littlejohn, M.D.; Tiplady, K.; Adams, L.F.; et al. Screening for phenotypic outliers identifies an unusually low concentration of a β -lactoglobulin B protein isoform in bovine milk caused by a synonymous SNP. *Genetics Selection Evolution* **2022**, *54*, 1–11. <https://doi.org/10.1186/s12711-022-00711-z>. 565

19. Daetwyler, H.D.; Schenkel, F.S.; Sargolzaei, M.; Robinson, J.A.B. A genome scan to detect quantitative trait loci for economically important traits in Holstein cattle using two methods and a dense single nucleotide polymorphism map. *Journal of dairy science* **2008**, *91*, 3225–3236. <https://doi.org/10.3168/jds.2007-0333>. 566

20. Kemper, K.; Vander Jagt, C.; Bowman, P.; Reich, C.; Mason, B.; Hayes, B.; Goddard, M. Mapping QTL in Australian dairy cattle using Genomic Selection Methodologies. In Proceedings of the 10th world congress on genetics applied to Livestock Production, 2014. 567

21. van den Berg, I.; Xiang, R.; Jenko, J.; Pausch, H.; Boussaha, M.; Schrooten, C.; Tribout, T.; Gjuvslund, A.B.; Boichard, D.; Nordbø, Ø.; et al. Meta-analysis for milk fat and protein percentage using imputed sequence variant genotypes in 94,321 cattle from eight cattle breeds. *Genetics Selection Evolution* **2020**, *52*, 1–16. <https://doi.org/10.1186/s12711-020-00556-4>. 568

22. Pedrosa, V.B.; Schenkel, F.S.; Chen, S.Y.; Oliveira, H.R.; Casey, T.M.; Melka, M.G.; Brito, L.F. Genomewide Association Analyses of Lactation Persistency and Milk Production Traits in Holstein Cattle Based on Imputed Whole-Genome Sequence Data. *Genes* **2021**, *12*, 1830. <https://doi.org/10.3390/genes12111830>. 569

23. Lopdell, T.J.; Tiplady, K.; Struchalin, M.; Johnson, T.J.; Keehan, M.; Sherlock, R.; Couldrey, C.; Davis, S.R.; Snell, R.G.; Spelman, R.J.; et al. DNA and RNA-sequence based GWAS highlights membrane-transport genes as key modulators of milk lactose content. *BMC genomics* **2017**, *18*, 1–18. <https://doi.org/10.1186/s12864-017-4320-3>. 570

24. Stinnakre, M.; Vilotte, J.; Soulier, S.; Mercier, J. Creation and phenotypic analysis of α -lactalbumin-deficient mice. *Proceedings of the National Academy of Sciences* **1994**, *91*, 6544–6548. <https://doi.org/10.1073/pnas.91.14.6544>. 609

25. Håkansson, A.; Zhivotovsky, B.; Orrenius, S.; Sabharwal, H.; Svanborg, C. Apoptosis induced by a human milk protein. *Proceedings of the National Academy of Sciences* **1995**, *92*, 8064–8068. <https://doi.org/10.1073/pnas.92.17.8064>. 610

26. Cheng, J.; Wang, J.; Bu, D.; Liu, G.; Zhang, C.; Wei, H.; Zhou, L.; Wang, J. Factors affecting the lactoferrin concentration in bovine milk. *Journal of dairy science* **2008**, *91*, 970–976. <https://doi.org/10.3168/jds.2007-0689>. 611

27. Villavicencio, A.; Rueda, M.S.; Turin, C.G.; Ochoa, T.J. Factors affecting lactoferrin concentration in human milk: how much do we know? *Biochemistry and Cell Biology* **2017**, *95*, 12–21. <https://doi.org/10.1139/bcb-2016-0060>. 612

28. Prieur, D.J. Tissue specific deficiency of lysozyme in ruminants. *Comparative Biochemistry and physiology. B, Comparative Biochemistry* **1986**, *85*, 349–353. [https://doi.org/10.1016/0305-0491\(86\)90011-8](https://doi.org/10.1016/0305-0491(86)90011-8). 613

29. Seyfert, H.M. Lactoferrin and lysozyme—two mammalian antibiotics. *Arch. Anim. Breed* **1999**, *42*, 22–27. 614

30. Carlsson, Å.; Björck, L.; Persson, K. Lactoferrin and lysozyme in milk during acute mastitis and their inhibitory effect in Delvotest P. *Journal of Dairy Science* **1989**, *72*, 3166–3175. [https://doi.org/10.3168/jds.S0022-0302\(89\)79475-3](https://doi.org/10.3168/jds.S0022-0302(89)79475-3). 615

31. Gaunt, S.; Raffio, N.; Kingsbury, E.; Damon Jr, R.; Johnson, W.; Mitchell, B. Variation of lactoferrin and mastitis and their heritabilities. *Journal of Dairy Science* **1980**, *63*, 1874–1880. [https://doi.org/10.3168/jds.S0022-0302\(80\)83154-7](https://doi.org/10.3168/jds.S0022-0302(80)83154-7). 616

32. Arnould, V.R.; Soyeurt, H.; Gengler, N.; Colinet, F.; Georges, M.; Bertozzi, C.; Portetelle, D.; Renaville, R. Genetic analysis of lactoferrin content in bovine milk. *Journal of Dairy Science* **2009**, *92*, 2151–2158. <https://doi.org/10.3168/jds.2008-1255>. 617

33. Cecchinato, A.; Ribeca, C.; Chessa, S.; Cipolat-Gotet, C.; Maretto, F.; Casellas, J.; Bittante, G. Candidate gene association analysis for milk yield, composition, urea nitrogen and somatic cell scores in Brown Swiss cows. *Animal* **2014**, *8*, 1062–1070. <https://doi.org/10.1017/S1751731114001098>. 618

34. Gothefors, L.; Marklund, S. Lactoperoxidase activity in human milk and in saliva of newborn infants. *Infection and immunity* **1975**, *11*, 1210–1215. <https://doi.org/10.1128/iai.11.6.1210-1215.1975>. 619

35. Kemper, K.; Littlejohn, M.; Lopdell, T.; Hayes, B.; Bennett, L.; Williams, R.; Xu, X.; Visscher, P.M.; Carrick, M.; Goddard, M. Leveraging genetically simple traits to identify small-effect variants for complex phenotypes. *BMC genomics* **2016**, *17*, 1–9. <https://doi.org/10.1186/s12864-016-3175-3>. 620

36. Cavaletto, M.; Giuffrida, M.G.; Conti, A. Milk fat globule membrane components—a proteomic approach. In *Bioactive Components of Milk*; Böszö, Z., Ed.; Springer: New York, USA, 2008; chapter II, pp. 129–142. 621

37. Robenek, H.; Hofnagel, O.; Buers, I.; Lorkowski, S.; Schnoor, M.; Robenek, M.J.; Heid, H.; Troyer, D.; Severs, N.J. Butyrophilin controls milk fat globule secretion. *Proceedings of the National Academy of Sciences* **2006**, *103*, 10385–10390. <https://doi.org/10.1073/pnas.0600795103>. 622

38. Vorbach, C.; Scriven, A.; Capecchi, M.R. The housekeeping gene xanthine oxidoreductase is necessary for milk fat droplet enveloping and secretion: gene sharing in the lactating mammary gland. *Genes & development* **2002**, *16*, 3223–3235. <https://doi.org/10.1101/gad.1032702>. 623

39. Reynolds, E.G.; Neeley, C.; Lopdell, T.J.; Keehan, M.; Dittmer, K.; Harland, C.S.; Couldrey, C.; Johnson, T.J.; Tiplady, K.; Worth, G.; et al. Non-additive association analysis using proxy phenotypes identifies novel cattle syndromes. *Nature Genetics* **2021**, *53*, 949–954. <https://doi.org/10.1038/s41588-021-00872-5>. 624

40. Grisart, B.; Coppieters, W.; Farnir, F.; Karim, L.; Ford, C.; Berzi, P.; Cambisano, N.; Mni, M.; Reid, S.; Simon, P.; et al. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine *DGAT1* gene with major effect on milk yield and composition. *Genome research* **2002**, *12*, 222–231. <https://doi.org/10.1101/gr.224202>. 625

41. Yen, C.L.E.; Stone, S.J.; Koliwad, S.; Harris, C.; Farese, R.V. DGAT enzymes and triacylglycerol biosynthesis. *Journal of lipid research* **2008**, *49*, 2283–2301. <https://doi.org/10.1194/jlr.R800018-JLR200>. 626

42. Fink, T.; Lopdell, T.J.; Tiplady, K.; Handley, R.; Johnson, T.J.; Spelman, R.J.; Davis, S.R.; Snell, R.G.; Littlejohn, M.D. A new mechanism for a familiar mutation – bovine *DGAT1* K232A modulates gene expression through multi-junction exon splice enhancement. *BMC genomics* **2020**, *21*, 1–13. <https://doi.org/10.1186/s12864-020-07004-z>. 627

43. Littlejohn, M.D.; Tiplady, K.; Lopdell, T.; Law, T.A.; Scott, A.; Harland, C.; Sherlock, R.; Henty, K.; Obolonkin, V.; Lehnert, K.; et al. Expression variants of the lipogenic *AGPAT6* gene affect diverse milk composition phenotypes in *Bos taurus*. *PloS one* **2014**, *9*, e85757. <https://doi.org/10.1371/journal.pone.0085757>. 628

44. Reue, K.; Brindley, D.N. Multiple roles for lipins/phosphatidate phosphatase enzymes in lipid metabolism. *Journal of lipid research* **2008**, *49*, 2493–2503. <https://doi.org/10.1194/jlr.R800019-JLR200>. 629

45. Han, B.; Yuan, Y.; Liang, R.; Li, Y.; Liu, L.; Sun, D. Genetic effects of *LPIN1* polymorphisms on milk production traits in dairy cattle. *Genes* **2019**, *10*, 265. <https://doi.org/10.3390/genes10040265>. 630

46. Matsumoto, H.; Sasaki, K.; Bessho, T.; Kobayashi, E.; Abe, T.; Sasazaki, S.; Oyama, K.; Mannen, H. The SNPs in the *ACACA* gene are effective on fatty acid composition in Holstein milk. *Molecular biology reports* **2012**, *39*, 8637–8644. <https://doi.org/10.1007/s11033-012-1718-5>. 631

47. Smith, S.; Witkowski, A.; Joshi, A.K. Structural and functional organization of the animal fatty acid synthase. *Progress in lipid research* **2003**, *42*, 289–317. [https://doi.org/10.1016/S0163-7827\(02\)00067-X](https://doi.org/10.1016/S0163-7827(02)00067-X). 632

48. Wang, T.; Chen, Y.P.P.; MacLeod, I.M.; Pryce, J.E.; Goddard, M.E.; Hayes, B.J. Application of a Bayesian non-linear model hybrid scheme to sequence data for genomic prediction and QTL mapping. *BMC genomics* **2017**, *18*, 618. <https://doi.org/10.1186/s12864-017-4030-x>. 633

49. Schennink, A.; Bovenhuis, H.; Léon-Kloosterziel, K.M.; Van Arendonk, J.A.; Visker, M.H. Effect of polymorphisms in the *FASN*, *OLR1*, *PPARGC1A*, *PRL* and *STAT5A* genes on bovine milk-fat composition. *Animal Genetics* **2009**, *40*, 909–916. <https://doi.org/10.1111/j.1365-2052.2009.01940.x>.

50. Cho, H.P.; Nakamura, M.; Clarke, S.D. Cloning, expression, and fatty acid regulation of the human Δ -5 desaturase. *Journal of Biological Chemistry* **1999**, *274*, 37335–37339. <https://doi.org/10.1074/jbc.274.52.37335>.

51. Park, W.J.; Kothapalli, K.S.; Lawrence, P.; Tyburczy, C.; Brenna, J.T. An alternate pathway to long-chain polyunsaturates: the *FADS2* gene product Δ 8-desaturates 20:2n-6 and 20:3n-3. *Journal of lipid research* **2009**, *50*, 1195–1202. <https://doi.org/10.1194/jlr.M800630-JLR200>.

52. Ibeagha-Awemu, E.M.; Akwanji, K.A.; Beaudoin, F.; Zhao, X. Associations between variants of *FADS* genes and omega-3 and omega-6 milk fatty acids of Canadian Holstein cows. *BMC genetics* **2014**, *15*, 1–9. <https://doi.org/10.1186/1471-2156-15-25>.

53. Golan, Y.; Assaraf, Y.G. Genetic and physiological factors affecting human milk production and composition. *Nutrients* **2020**, *12*, 1500. <https://doi.org/10.3390/nu12051500>.

54. Niwa, S.; Kawabata, T.; Shoji, K.; Ogata, H.; Kagawa, Y.; Nakayama, K.; Yanagisawa, Y.; Iwamoto, S.; Tatsuta, N.; Asato, K.; et al. Investigation of Maternal Diet and *FADS1* Polymorphism Associated with Long-Chain Polyunsaturated Fatty Acid Compositions in Human Milk. *Nutrients* **2022**, *14*, 2160. <https://doi.org/10.3390/nu14102160>.

55. Paton, C.M.; Ntambi, J.M. Biochemical and physiological function of stearoyl-CoA desaturase. *American Journal of Physiology-Endocrinology and Metabolism* **2009**, *297*, E28–E37. <https://doi.org/10.1152/ajpendo.90897.2008>.

56. Miyazaki, M.; Kim, Y.C.; Gray-Keller, M.P.; Attie, A.D.; Ntambi, J.M. The biosynthesis of hepatic cholesterol esters and triglycerides is impaired in mice with a disruption of the gene for stearoyl-CoA desaturase 1. *Journal of Biological Chemistry* **2000**, *275*, 30132–30138. <https://doi.org/10.1074/jbc.M005488200>.

57. Rahman, S.M.; Dobrzyn, A.; Dobrzyn, P.; Lee, S.H.; Miyazaki, M.; Ntambi, J.M. Stearoyl-CoA desaturase 1 deficiency elevates insulin-signaling components and down-regulates protein-tyrosine phosphatase 1B in muscle. *Proceedings of the National Academy of Sciences* **2003**, *100*, 11110–11115. <https://doi.org/10.1073/pnas.1934571100>.

58. Banos, G.; Woolliams, J.; Woodward, B.; Forbes, A.; Coffey, M. Impact of single nucleotide polymorphisms in leptin, leptin receptor, growth hormone receptor, and diacylglycerol acyltransferase (*DGAT1*) gene loci on milk production, feed, and body energy traits of UK dairy cows. *Journal of dairy science* **2008**, *91*, 3190–3200. <https://doi.org/10.3168/jds.2007-0930>.

59. Shi, T.; Xu, Y.; Yang, M.; Huang, Y.; Lan, X.; Lei, C.; Qi, X.; Yang, X.; Chen, H. Copy number variations at *LEPR* gene locus associated with gene expression and phenotypic traits in Chinese cattle. *Animal Science Journal* **2016**, *87*, 336–343. <https://doi.org/10.1111/asj.12531>.

60. Pan, H.; Guo, J.; Su, Z. Advances in understanding the interrelations between leptin resistance and obesity. *Physiology & behavior* **2014**, *130*, 157–169. <https://doi.org/10.1016/j.physbeh.2014.04.003>.

61. Caro, J.F.; Sinha, M.K.; Kolaczynski, J.W.; Zhang, P.L.; Considine, R.V. Leptin: the tale of an obesity gene. *Diabetes* **1996**, *45*, 1455–1463. <https://doi.org/10.2337/diab.45.11.1455>.

62. Loos, R.J.; Yeo, G.S. The bigger picture of *FTO*—the first GWAS-identified obesity gene. *Nature Reviews Endocrinology* **2014**, *10*, 51–61. <https://doi.org/10.1038/nrendo.2013.227>.

63. Han, Z.; Niu, T.; Chang, J.; Lei, X.; Zhao, M.; Wang, Q.; Cheng, W.; Wang, J.; Feng, Y.; Chai, J. Crystal structure of the *FTO* protein reveals basis for its substrate specificity. *Nature* **2010**, *464*, 1205–1209. <https://doi.org/10.1038/nature08921>.

64. Wei, J.; Liu, F.; Lu, Z.; Fei, Q.; Ai, Y.; He, P.C.; Shi, H.; Cui, X.; Su, R.; Klungland, A.; et al. Differential m⁶A, m⁶A_m, and m¹A demethylation mediated by *FTO* in the cell nucleus and cytoplasm. *Molecular cell* **2018**, *71*, 973–985. <https://doi.org/10.1016/j.molcel.2018.08.011>.

65. Wu, R.; Liu, Y.; Yao, Y.; Zhao, Y.; Bi, Z.; Jiang, Q.; Liu, Q.; Cai, M.; Wang, F.; Wang, Y.; et al. *FTO* regulates adipogenesis by controlling cell cycle progression via m⁶A-YTHDF2 dependent mechanism. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* **2018**, *1863*, 1323–1330. <https://doi.org/10.1016/j.bbalip.2018.08.008>.

66. Zielke, L.G.; Bortfeldt, R.H.; Reissmann, M.; Tetens, J.; Thaller, G.; Brockmann, G.A. Impact of variation at the *FTO* locus on milk fat yield in Holstein dairy cattle. *PLoS One* **2013**, *8*, e63406. <https://doi.org/10.1371/journal.pone.0063406>.

67. 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.; et al. Functionally reciprocal mutations of the prolactin signalling pathway define hairy and slick cattle. *Nature communications* **2014**, *5*, 1–8. <https://doi.org/10.1038/ncomms6861>.

68. Harris, J.; Stanford, P.M.; Sutherland, K.; Oakes, S.R.; Naylor, M.J.; Robertson, F.G.; Blazek, K.D.; Kazlauskas, M.; Hilton, H.N.; Wittlin, S.; et al. *Socs2* and *Elf5* mediate prolactin-induced mammary gland development. *Molecular endocrinology* **2006**, *20*, 1177–1187. <https://doi.org/10.1210/me.2005-0473>.

69. Oliver, C.H.; Watson, C.J. Making milk: A new link between *STAT5* and *Akt1*. *JAK-STAT* **2013**, *2*, e23228. <https://doi.org/10.4161/jkst.23228>.

70. Mohammad, M.A.; Hadsell, D.L.; Haymond, M.W. Gene regulation of UDP-galactose synthesis and transport: potential rate-limiting processes in initiation of milk production in humans. *American Journal of Physiology-Endocrinology and Metabolism* **2012**, *303*, E365–E376. <https://doi.org/10.1152/ajpendo.00175.2012>.

71. Chen, C.C.; Boxer, R.B.; Stairs, D.B.; Portocarrero, C.P.; Horton, R.H.; Alvarez, J.V.; Birnbaum, M.J.; Chodosh, L.A. *Akt* is required for *Stat5* activation and mammary differentiation. *Breast Cancer Research* **2010**, *12*, 1–14. <https://doi.org/10.1186/bcr2640>.

72. Boxer, R.B.; Stairs, D.B.; Dugan, K.D.; Notarfrancesco, K.L.; Portocarrero, C.P.; Keister, B.A.; Belka, G.K.; Cho, H.; Rathmell, J.C.; Thompson, C.B.; et al. Isoform-specific requirement for Akt1 in the developmental regulation of cellular metabolism during lactation. *Cell metabolism* **2006**, *4*, 475–490. <https://doi.org/10.1016/j.cmet.2006.10.011>. 726

73. Anderson, S.M.; Rudolph, M.C.; McManaman, J.L.; Neville, M.C. Key stages in mammary gland development. Secretory activation in the mammary gland: it's not just about milk protein synthesis! *Breast Cancer Research* **2007**, *9*, 204. <https://doi.org/10.1186/bcr1653>. 727

74. Chen, C.C.; Stairs, D.B.; Boxer, R.B.; Belka, G.K.; Horseman, N.D.; Alvarez, J.V.; Chodosh, L.A. Autocrine prolactin induced by the Pten–Akt pathway is required for lactation initiation and provides a direct link between the Akt and Stat5 pathways. *Genes & development* **2012**, *26*, 2154–2168. <https://doi.org/10.1101/gad.197343.112>. 728

75. Lü, A.; Hu, X.; Chen, H.; Dong, Y.; Zhang, Y.; Wang, X. Novel SNPs of the bovine *PRLR* gene associated with milk production traits. *Biochemical genetics* **2011**, *49*, 177–189. <https://doi.org/10.1007/s10528-010-9397-1>. 729

76. Meredith, B.K.; Kearney, F.J.; Finlay, E.K.; Bradley, D.G.; Fahey, A.G.; Berry, D.P.; Lynn, D.J. Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. *BMC genetics* **2012**, *13*, 21. <https://doi.org/10.1186/1471-2156-13-21>. 730

77. Tiplady, K.M.; Lopdell, T.J.; Reynolds, E.; Sherlock, R.G.; Keehan, M.; Johnson, T.J.; Pryce, J.E.; Davis, S.R.; Spelman, R.J.; Harris, B.L.; et al. Sequence-based genome-wide association study of individual milk mid-infrared wavenumbers in mixed-breed dairy cattle. *Genetics Selection Evolution* **2021**, *53*, 1–24. <https://doi.org/10.1186/s12711-021-00648-9>. 731

78. Wallis, M. The molecular evolution of pituitary growth hormone prolactin and placental lactogen: A protein family showing variable rates of evolution. *Journal of Molecular Evolution* **1981**, *17*, 10–18. <https://doi.org/10.1007/BF01792419>. 732

79. Bauman, D.E.; Currie, W.B. Partitioning of nutrients during pregnancy and lactation: a review of mechanisms involving homeostasis and homeorhesis. *Journal of dairy science* **1980**, *63*, 1514–1529. [https://doi.org/10.3168/jds.S0022-0302\(80\)83111-0](https://doi.org/10.3168/jds.S0022-0302(80)83111-0). 733

80. Shingu, H.; Hodate, K.; Kushibiki, S.; Ueda, Y.; Touno, E.; Shinoda, M.; Ohashi, S. Hormonal and lactational responses to growth hormone-releasing hormone treatment in lactating Japanese Black cows. *Journal of dairy science* **2004**, *87*, 1684–1693. [https://doi.org/10.3168/jds.S0022-0302\(04\)73321-4](https://doi.org/10.3168/jds.S0022-0302(04)73321-4). 734

81. Blott, S.; Kim, J.J.; Moio, S.; Schmidt-Küntzel, A.; Cornet, A.; Berzi, P.; Cambisano, N.; Ford, C.; Grisart, B.; Johnson, D.; et al. Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. *Genetics* **2003**, *163*, 253–266. <https://doi.org/10.1093/genetics/163.1.253>. 735

82. Jiang, J.; Ma, L.; Prakapenka, D.; VanRaden, P.M.; Cole, J.B.; Da, Y. A large-scale genome-wide association study in US Holstein cattle. *Frontiers in genetics* **2019**, *10*, 412. <https://doi.org/10.3389/fgene.2019.00412>. 736

83. Cai, Z.; Dusza, M.; Guldbrandtsen, B.; Lund, M.S.; Sahana, G. Distinguishing pleiotropy from linked QTL between milk production traits and mastitis resistance in Nordic Holstein cattle. *Genetics Selection Evolution* **2020**, *52*, 19. <https://doi.org/10.1186/s12711-020-00538-6>. 737

84. Bocker, C.; Thompson, D.; Matsumoto, A.; Nebert, D.W.; Vasiliou, V. Evolutionary divergence and functions of the human interleukin (IL) gene family. *Human genomics* **2010**, *5*, 1–26. <https://doi.org/10.1186/1479-7364-5-1-30>. 738

85. Raven, L.A.; Cocks, B.G.; Kemper, K.E.; Chamberlain, A.J.; Vander Jagt, C.J.; Goddard, M.E.; Hayes, B.J. Targeted imputation of sequence variants and gene expression profiling identifies twelve candidate genes associated with lactation volume, composition and calving interval in dairy cattle. *Mammalian genome* **2016**, *27*, 81–97. <https://doi.org/10.1007/s00335-015-9613-8>. 739

86. Lopdell, T.J.; Tiplady, K.; Couldrey, C.; Johnson, T.J.; Keehan, M.; Davis, S.R.; Harris, B.L.; Spelman, R.J.; Snell, R.G.; Littlejohn, M.D. Multiple QTL underlie milk phenotypes at the *CSF2RB* locus. *Genetics Selection Evolution* **2019**, *51*, 1–16. <https://doi.org/10.1186/s12711-019-0446-x>. 740

87. Li, M.; Liu, X.; Robinson, G.; Bar-Peled, U.; Wagner, K.U.; Young, W.S.; Hennighausen, L.; Furth, P.A. Mammary-derived signals activate programmed cell death during the first stage of mammary gland involution. *Proceedings of the National Academy of Sciences* **1997**, *94*, 3425–3430. <https://doi.org/10.1073/pnas.94.7.3425>. 741

88. Kritikou, E.A.; Sharkey, A.; Abell, K.; Came, P.J.; Anderson, E.; Clarkson, R.W.; Watson, C.J. A dual, non-redundant, role for LIF as a regulator of development and STAT3-mediated cell death in mammary gland. *Development* **2003**, *130*, 3459–3468. <https://doi.org/10.1242/dev.00578>. 742

89. Clarkson, R.W.; Boland, M.P.; Kritikou, E.A.; Lee, J.M.; Freeman, T.C.; Tiffen, P.G.; Watson, C.J. The genes induced by signal transducer and activators of transcription (STAT)3 and STAT5 in mammary epithelial cells define the roles of these STATs in mammary development. *Molecular endocrinology* **2006**, *20*, 675–685. <https://doi.org/10.1210/me.2005-0392>. 743

90. Bates, S.H.; Stearns, W.H.; Dundon, T.A.; Schubert, M.; Tso, A.W.; Wang, Y.; Banks, A.S.; Lavery, H.J.; Haq, A.K.; Maratos-Flier, E.; et al. STAT3 signalling is required for leptin regulation of energy balance but not reproduction. *Nature* **2003**, *421*, 856–859. <https://doi.org/10.1038/nature01388>. 744

91. Cobanoglu, O.; Zaitoun, I.; Chang, Y.; Shook, G.; Khatib, H. Effects of the signal transducer and activator of transcription 1 (*STAT1*) gene on milk production traits in Holstein dairy cattle. *Journal of dairy science* **2006**, *89*, 4433–4437. [https://doi.org/10.3168/jds.S0022-0302\(06\)72491-2](https://doi.org/10.3168/jds.S0022-0302(06)72491-2). 745

92. Arun, S.J.; Thomson, P.C.; Sheehy, P.A.; Khatkar, M.S.; Raadsma, H.W.; Williamson, P. Targeted analysis reveals an important role of JAK-STAT-SOCS genes for milk production traits in Australian dairy cattle. *Frontiers in Genetics* **2015**, *6*, 342. <https://doi.org/10.3389/fgene.2015.00342>. 746

93. Macias, H.; Hinck, L. Mammary gland development. *Wiley Interdisciplinary Reviews: Developmental Biology* **2012**, *1*, 533–557. <https://doi.org/10.1002/wdev.35>.

94. Greenhalgh, C.J.; Rico-Bautista, E.; Lorentzon, M.; Thaus, A.L.; Morgan, P.O.; Willson, T.A.; Zervoudakis, P.; Metcalf, D.; Street, I.; Nicola, N.A.; et al. SOCS2 negatively regulates growth hormone action in vitro and in vivo. *The Journal of clinical investigation* **2005**, *115*, 397–406. <https://doi.org/10.1172/JCI22710>.

95. Rupp, R.; Senin, P.; Sarry, J.; Allain, C.; Tasca, C.; Ligat, L.; Portes, D.; Woloszyn, F.; Bouchez, O.; Tabouret, G.; et al. A point mutation in suppressor of cytokine signalling 2 (Socs2) increases the susceptibility to inflammation of the mammary gland while associated with higher body weight and size and higher milk production in a sheep model. *PLoS genetics* **2015**, *11*, e1005629. <https://doi.org/10.1371/journal.pgen.1005629>.

96. Pan, C.J.; Chen, S.Y.; Jun, H.S.; Lin, S.R.; Mansfield, B.C.; Chou, J.Y. SLC37A1 and SLC37A2 are phosphate-linked, glucose-6-phosphate antiporters. *PloS one* **2011**, *6*, e23157. <https://doi.org/10.1371/journal.pone.0023157>.

97. Sanchez, M.P.; Govignon-Gion, A.; Croiseau, P.; Fritz, S.; Hozé, C.; Miranda, G.; Martin, P.; Barbat-Leterrier, A.; Letaïef, R.; Rocha, D.; et al. Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. *Genetics Selection Evolution* **2017**, *49*, 1–16. <https://doi.org/10.1186/s12711-017-0344-z>.

98. Chen, L.Q.; Hou, B.H.; Lalonde, S.; Takanaga, H.; Hartung, M.L.; Qu, X.Q.; Guo, W.J.; Kim, J.G.; Underwood, W.; Chaudhuri, B.; et al. Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* **2010**, *468*, 527–532. <https://doi.org/10.1038/nature09606>.

99. Pradeep, J.; Monika, S.; Ankita, S.; Umesh, K.; Amit, K.; Ashok, M.; Mishra, B.; Sandeep, M.; Kataria, R.; Kaushik, J.; et al. Expression analysis of solute carrier (SLC2A) genes in milk derived mammary epithelial cells during different stages of lactation in sahiwal (*Bos indicus*) cows. *Journal of Advances in Dairy Research* **2014**, *2*, 117. <https://doi.org/10.4172/2329-888X.1000117>.

100. Zwierzchowski, L.; Ostrowska, M.; Żelazowska, B.; Bagnicka, E. Single nucleotide polymorphisms in the bovine SLC2A12 and SLC5A1 glucose transporter genes—the effect on gene expression and milk traits of Holstein Friesian cows. *Animal Biotechnology* **2021**, *2021 Aug 6*, 1–11. <https://doi.org/10.1080/10495398.2021.1954934>.

101. Voss, F.K.; Ullrich, F.; Münch, J.; Lazarow, K.; Lutter, D.; Mah, N.; Andrade-Navarro, M.A.; von Kries, J.P.; Stauber, T.; Jentsch, T.J. Identification of LRRC8 heteromers as an essential component of the volume-regulated anion channel VRAC. *Science* **2014**, *344*, 634–638. <https://doi.org/10.1126/science.1252826>.

102. Syeda, R.; Qiu, Z.; Dubin, A.E.; Murthy, S.E.; Florendo, M.N.; Mason, D.E.; Mathur, J.; Cahalan, S.M.; Peters, E.C.; Montal, M.; et al. LRRC8 proteins form volume-regulated anion channels that sense ionic strength. *Cell* **2016**, *164*, 499–511. <https://doi.org/10.1016/j.cell.2015.12.031>.

103. Tominaga, K.; Kondo, C.; Kagata, T.; Hishida, T.; Nishizuka, M.; Imagawa, M. The novel gene fad158, having a transmembrane domain and leucine-rich repeat, stimulates adipocyte differentiation. *Journal of Biological Chemistry* **2004**, *279*, 34840–34848. <https://doi.org/10.1074/jbc.M312927200>.

104. Lu, J.; Argov-Argaman, N.; Anggrek, J.; Boeren, S.; van Hooijdonk, T.; Vervoort, J.; Hettinga, K.A. The protein and lipid composition of the membrane of milk fat globules depends on their size. *Journal of dairy science* **2016**, *99*, 4726–4738. <https://doi.org/10.3168/jds.2015-10375>.

105. Matsuzaki, T.; Machida, N.; Tajika, Y.; Ablimit, A.; Suzuki, T.; Aoki, T.; Hagiwara, H.; Takata, K. Expression and immunolocalization of water-channel aquaporins in the rat and mouse mammary gland. *Histochemistry and cell biology* **2005**, *123*, 501–512. <https://doi.org/10.1007/s00418-005-0753-x>.

106. Kendall, B.; Maxwell, J.; German, A.; Marples, D.; Royal, M.; Mobasher, A. Immunohistochemical localisation of aquaporin water channels in the bovine mammary gland. In Proceedings of the Proceedings of the British Society of Animal Science. Cambridge University Press, British Society of Animal Science, 2007, Vol. 2007, pp. 1–1.

107. Mobasher, A.; Barrett-Jolley, R. Aquaporin water channels in the mammary gland: from physiology to pathophysiology and neoplasia. *Journal of mammary gland biology and neoplasia* **2014**, *19*, 91–102. <https://doi.org/10.1007/s10911-013-9312-6>.

108. Kaihoko, Y.; Tsugami, Y.; Suzuki, N.; Suzuki, T.; Nishimura, T.; Kobayashi, K. Distinct expression patterns of aquaporin 3 and 5 in ductal and alveolar epithelial cells in mouse mammary glands before and after parturition. *Cell and tissue research* **2020**, *380*, 513–526. <https://doi.org/10.1007/s00441-020-03168-y>.

109. Sutura, A.M.; Tolone, M.; Mastrangelo, S.; Di Gerlando, R.; Sardina, M.T.; Portolano, B.; Pong-Wong, R.; Riggio, V. Detection of genomic regions underlying milk production traits in Valle del Belice dairy sheep using regional heritability mapping. *Journal of Animal Breeding and Genetics* **2021**, *138*, 552–561. <https://doi.org/10.1111/jbg.12552>.

110. Gutman, G.A.; Chandy, K.G.; Grissmer, S.; Lazdunski, M.; Mckinnon, D.; Pardo, L.A.; Robertson, G.A.; Rudy, B.; Sanguinetti, M.C.; Stühmer, W.; et al. International Union of Pharmacology. LIII. Nomenclature and molecular relationships of voltage-gated potassium channels. *Pharmacological reviews* **2005**, *57*, 473–508. <https://doi.org/10.1124/pr.57.4.10>.

111. Kubo, Y.; Adelman, J.P.; Clapham, D.E.; Jan, L.Y.; Karschin, A.; Kurachi, Y.; Lazdunski, M.; Nichols, C.G.; Seino, S.; Vandenberg, C.A. International Union of Pharmacology. LIV. Nomenclature and molecular relationships of inwardly rectifying potassium channels. *Pharmacological reviews* **2005**, *57*, 509–526. <https://doi.org/10.1124/pr.57.4.11>.

112. Goldstein, S.A.; Bayliss, D.A.; Kim, D.; Lesage, F.; Plant, L.D.; Rajan, S. International Union of Pharmacology. LV. Nomenclature and molecular relationships of two-P potassium channels. *Pharmacological reviews* **2005**, *57*, 527–540. <https://doi.org/10.1124/pr.57.4.12>.

113. Kaczmarek, L.K.; Aldrich, R.W.; Chandy, K.G.; Grissmer, S.; Wei, A.D.; Wulff, H. International union of basic and clinical pharmacology. C. Nomenclature and properties of calcium-activated and sodium-activated potassium channels. *Pharmacological reviews* **2017**, *69*, 1–11. <https://doi.org/10.1124/pr.116.012864>. 843
114. Yan, J.; Aldrich, R.W. LRRC26 auxiliary protein allows BK channel activation at resting voltage without calcium. *Nature* **2010**, *466*, 513–516. <https://doi.org/10.1038/nature09162>. 844
115. Gonzalez-Perez, V.; Xia, X.M.; Lingle, C.J. Functional regulation of BK potassium channels by $\gamma 1$ auxiliary subunits. *Proceedings of the National Academy of Sciences* **2014**, *111*, 4868–4873. <https://doi.org/10.1073/pnas.1322123111>. 845
116. Doyle, L.A.; Yang, W.; Abruzzo, L.V.; Krogmann, T.; Gao, Y.; Rishi, A.K.; Ross, D.D. A multidrug resistance transporter from human MCF-7 breast cancer cells. *Proceedings of the National Academy of Sciences* **1998**, *95*, 15665–15670. <https://doi.org/10.1073/pnas.95.26.15665>. 846
117. van Herwaarden, A.E.; Wagenaar, E.; Merino, G.; Jonker, J.W.; Rosing, H.; Beijnen, J.H.; Schinkel, A.H. Multidrug transporter ABCG2/breast cancer resistance protein secretes riboflavin (vitamin B₂) into milk. *Molecular and cellular biology* **2007**, *27*, 1247–1253. <https://doi.org/10.1128/MCB.01621-06>. 847
118. Woodward, O.M.; Kottgen, A.; Coresh, J.; Boerwinkle, E.; Guggino, W.B.; Kottgen, M. Identification of a urate transporter, ABCG2, with a common functional polymorphism causing gout. *Proceedings of the National Academy of Sciences* **2009**, *106*, 10338–10342. <https://doi.org/10.1073/pnas.0901249106>. 848
119. Cohen-Zinder, M.; Seroussi, E.; Larkin, D.M.; Loo, J.J.; Everts-Van Der Wind, A.; Lee, J.H.; Drackley, J.K.; Band, M.R.; Hernandez, A.; Shani, M.; et al. Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome research* **2005**, *15*, 936–944. <https://doi.org/10.1101/gr.3806705>. 849
120. Leonard, S.; Khatib, H.; Schutzkus, V.; Chang, Y.; Maltecca, C. Effects of the osteopontin gene variants on milk production traits in dairy cattle. *Journal of Dairy Science* **2005**, *88*, 4083–4086. [https://doi.org/10.3168/jds.S0022-0302\(05\)73092-7](https://doi.org/10.3168/jds.S0022-0302(05)73092-7). 850
121. Schnabel, R.D.; Kim, J.J.; Ashwell, M.S.; Sonstegard, T.S.; Van Tassell, C.P.; Connor, E.E.; Taylor, J.F. Fine-mapping milk production quantitative trait loci on BTA6: Analysis of the bovine osteopontin gene. *Proceedings of the National Academy of Sciences* **2005**, *102*, 6896–6901. <https://doi.org/10.1073/pnas.0502398102>. 851
122. Fatehi, M.; Carter, C.R.; Youssef, N.; Hunter, B.E.; Holt, A.; Light, P.E. Molecular determinants of ATP-sensitive potassium channel MgATPase activity: diabetes risk variants and diazoxide sensitivity. *Bioscience reports* **2015**, *35*, e00238. <https://doi.org/10.1042/BSR20150143>. 852
123. Campbell, J.D.; Sansom, M.S.; Ashcroft, F.M. Potassium channel regulation: Structural insights into the function of the nucleotide-binding domains of the human sulphonylurea receptor. *EMBO reports* **2003**, *4*, 1038–1042. <https://doi.org/10.1038/sj.embor.7400003>. 853
124. Ashcroft, F.; Gribble, F. ATP-sensitive K⁺ channels and insulin secretion: their role in health and disease. *Diabetologia* **1999**, *42*, 903. 854
125. VanHouten, J.N.; Neville, M.C.; Wysolmerski, J.J. The calcium-sensing receptor regulates plasma membrane calcium adenosine triphosphatase isoform 2 activity in mammary epithelial cells: a mechanism for calcium-regulated calcium transport into milk. *Endocrinology* **2007**, *148*, 5943–5954. <https://doi.org/10.1210/en.2007-0850>. 855
126. Ho, A.M.; Johnson, M.D.; Kingsley, D.M. Role of the mouse *ank* gene in control of tissue calcification and arthritis. *Science* **2000**, *289*, 265–270. <https://doi.org/10.1126/science.289.5477.265>. 856
127. Vijen, S.; Hawes, C.; Runions, J.; Russell, R.G.G.; Wordsworth, B.P.; Carr, A.J.; Pink, R.C.; Zhang, Y. Differences in intracellular localisation of ANKH mutants that relate to mechanisms of calcium pyrophosphate deposition disease and craniometaphyseal dysplasia. *Scientific reports* **2020**, *10*, 7408. <https://doi.org/10.1038/s41598-020-63911-x>. 857
128. Szeri, F.; Niaziorimi, F.; Donnelly, S.; Fariha, N.; Tertyshnaia, M.; Patel, D.; Lundkvist, S.; van de Wetering, K. The mineralization regulator ANKH mediates cellular efflux of ATP, not pyrophosphate. *Journal of Bone and Mineral Research* **2022**, *37*, 1024–1031. <https://doi.org/10.1002/jbmr.4528>. 858
129. Seifert, W.; Posor, Y.; Schu, P.; Stenbeck, G.; Mundlos, S.; Klaassen, S.; Nürnberg, P.; Haucke, V.; Kornak, U.; Kühnisch, J. The progressive ankylosis protein ANK facilitates clathrin- and adaptor-mediated membrane traffic at the trans-Golgi network-to-endosome interface. *Human molecular genetics* **2016**, *25*, 3836–3848. <https://doi.org/10.1093/hmg/ddw230>. 859
130. Tian, Y.; Chang, J.C.; Fan, E.Y.; Flajolet, M.; Greengard, P. Adaptor complex AP2/PICALM, through interaction with LC3, targets Alzheimer’s APP-CTF for terminal degradation via autophagy. *Proceedings of the National Academy of Sciences* **2013**, *110*, 17071–17076. <https://doi.org/10.1073/pnas.1315110110>. 860
131. Mehla, K.; Singh, P.K. MUC1: a novel metabolic master regulator. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* **2014**, *1845*, 126–135. <https://doi.org/10.1016/j.bbcan.2014.01.001>. 861
132. Stoop, W.; Van Arendonk, J.; Heck, J.; Van Valenberg, H.; Bovenhuis, H. Genetic parameters for major milk fatty acids and milk production traits of Dutch Holstein-Friesians. *Journal of Dairy Science* **2008**, *91*, 385–394. <https://doi.org/10.3168/jds.2007-0181>. 862
133. Soyeurt, H.; Dardenne, P.; Dehareng, F.; Lognay, G.; Veselko, D.; Marlier, M.; Bertozzi, C.; Mayeres, P.; Gengler, N. Estimating fatty acid content in cow milk using mid-infrared spectrometry. *Journal of dairy science* **2006**, *89*, 3690–3695. [https://doi.org/10.3168/jds.S0022-0302\(06\)72409-2](https://doi.org/10.3168/jds.S0022-0302(06)72409-2). 863

134. Sanchez, M.P.; Ferrand, M.; Gelé, M.; Pourchet, D.; Miranda, G.; Martin, P.; Brochard, M.; Boichard, D. Genetic parameters for milk protein composition predicted using mid-infrared spectroscopy in the French Montbéliarde, Normande, and Holstein dairy cattle breeds. *Journal of dairy science* **2017**, *100*, 6371–6375. <https://doi.org/10.3168/jds.2017-12663>.

135. Vanlierde, A.; Soyeurt, H.; Gengler, N.; Colinet, F.G.; Froidmont, E.; Kreuzer, M.; Grandl, F.; Bell, M.; Lund, P.; Olijhoek, D.; et al. Development of an equation for estimating methane emissions of dairy cows from milk Fourier transform mid-infrared spectra by using reference data obtained exclusively from respiration chambers. *Journal of Dairy Science* **2018**, *101*, 7618–7624. <https://doi.org/10.3168/jds.2018-14472>.

136. Ho, P.; Bonfatti, V.; Luke, T.; Pryce, J. Classifying the fertility of dairy cows using milk mid-infrared spectroscopy. *Journal of dairy science* **2019**, *102*, 10460–10470. <https://doi.org/10.3168/jds.2019-16412>.

137. Tiplady, K.M.; Lopdell, T.J.; Sherlock, R.G.; Johnson, T.J.; Spelman, R.J.; Harris, B.L.; Davis, S.R.; Littlejohn, M.D.; Garrick, D.J. Comparison of the genetic characteristics of directly measured and Fourier-transform mid-infrared-predicted bovine milk fatty acids and proteins. *Journal of Dairy Science* **2022**, *105*, 9763–9791. <https://doi.org/10.3168/jds.2022-22089>.

138. Prowse-Wilkins, C.P.; Lopdell, T.J.; Xiang, R.; Vander Jagt, C.J.; Littlejohn, M.D.; Chamberlain, A.J.; Goddard, M.E. Genetic variation in histone modifications and gene expression identifies regulatory variants in the mammary gland of cattle. *BMC Genomics* **2022**, *23*, 815. <https://doi.org/10.1186/s12864-022-09002-9>.

139. Van Den Berg, I.; Hayes, B.; Chamberlain, A.; Goddard, M. Overlap between eQTL and QTL associated with production traits and fertility in dairy cattle. *BMC genomics* **2019**, *20*, 291. <https://doi.org/10.1186/s12864-019-5656-7>.

140. Gusev, A.; Ko, A.; Shi, H.; Bhatia, G.; Chung, W.; Penninx, B.W.; Jansen, R.; De Geus, E.J.; Boomsma, D.I.; Wright, F.A.; et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nature genetics* **2016**, *48*, 245–252. <https://doi.org/10.1038/ng.3506>.

141. Zhu, Z.; Zhang, F.; Hu, H.; Bakshi, A.; Robinson, M.R.; Powell, J.E.; Montgomery, G.W.; Goddard, M.E.; Wray, N.R.; Visscher, P.M.; et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature genetics* **2016**, *48*, 481–487.

142. Zhao, T.; Hu, Y.; Zang, T.; Wang, Y. Integrate GWAS, eQTL, and mQTL data to identify Alzheimer’s disease-related genes. *Frontiers in genetics* **2019**, *10*, 1021. <https://doi.org/10.3389/fgene.2019.01021>.

143. Hormozdiari, F.; Van De Bunt, M.; Segre, A.V.; Li, X.; Joo, J.W.J.; Bilow, M.; Sul, J.H.; Sankararaman, S.; Pasaniuc, B.; Eskin, E. Colocalization of GWAS and eQTL signals detects target genes. *The American Journal of Human Genetics* **2016**, *99*, 1245–1260. <https://doi.org/10.1016/j.ajhg.2016.10.003>.

144. Sheehan, D.; Meade, G.; Foley, V.M.; Dowd, C.A. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochemical journal* **2001**, *360*, 1–16. <https://doi.org/10.1042/bj3600001>.

145. Wang, X.; Wurmser, C.; Pausch, H.; Jung, S.; Reinhardt, F.; Tetens, J.; Thaller, G.; Fries, R. Identification and dissection of four major QTL affecting milk fat content in the German Holstein-Friesian population. *PloS one* **2012**, *7*, e40711. <https://doi.org/10.1371/journal.pone.0040711>.

146. Littlejohn, M.D.; Tiplady, K.; Fink, T.A.; Lehnert, K.; Lopdell, T.; Johnson, T.; Couldrey, C.; Keehan, M.; Sherlock, R.G.; Harland, C.; et al. Sequence-based association analysis reveals an *MGST1* eQTL with pleiotropic effects on bovine milk composition. *Scientific reports* **2016**, *6*, 25376. <https://doi.org/10.1038/srep25376>.

147. Hebbes, T.R.; Thorne, A.W.; Crane-Robinson, C. A direct link between core histone acetylation and transcriptionally active chromatin. *The EMBO journal* **1988**, *7*, 1395–1402. <https://doi.org/10.1002/j.1460-2075.1988.tb02956.x>.

148. Zhou, V.W.; Goren, A.; Bernstein, B.E. Charting histone modifications and the functional organization of mammalian genomes. *Nature Reviews Genetics* **2011**, *12*, 7–18. <https://doi.org/10.1038/nrg2905>.

149. Barski, A.; Cuddapah, S.; Cui, K.; Roh, T.Y.; Schones, D.E.; Wang, Z.; Wei, G.; Chepelev, I.; Zhao, K. High-resolution profiling of histone methylations in the human genome. *Cell* **2007**, *129*, 823–837. <https://doi.org/10.1016/j.cell.2007.05.009>.

150. Mikkelsen, T.S.; Ku, M.; Jaffe, D.B.; Issac, B.; Lieberman, E.; Giannoukos, G.; Alvarez, P.; Brockman, W.; Kim, T.K.; Koche, R.P.; et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **2007**, *448*, 553–560. <https://doi.org/10.1038/nature06008>.

151. Skene, P.J.; Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* **2017**, *6*, e21856. <https://doi.org/10.7554/eLife.21856.001>.

152. Boyle, A.P.; Davis, S.; Shulha, H.P.; Meltzer, P.; Margulies, E.H.; Weng, Z.; Furey, T.S.; Crawford, G.E. High-resolution mapping and characterization of open chromatin across the genome. *Cell* **2008**, *132*, 311–322. <https://doi.org/10.1016/j.cell.2007.12.014>.

153. Schones, D.E.; Cui, K.; Cuddapah, S.; Roh, T.Y.; Barski, A.; Wang, Z.; Wei, G.; Zhao, K. Dynamic regulation of nucleosome positioning in the human genome. *Cell* **2008**, *132*, 887–898. <https://doi.org/10.1016/j.cell.2008.02.022>.

154. Kuan, P.F.; Huebert, D.; Gasch, A.; Keles, S. A non-homogeneous hidden-state model on first order differences for automatic detection of nucleosome positions. *Statistical applications in genetics and molecular biology* **2009**, *8*, 29. <https://doi.org/10.2202/1544-6115.1454>.

155. Buenrostro, J.D.; Giresi, P.G.; Zaba, L.C.; Chang, H.Y.; Greenleaf, W.J. Transposition of native chromatin for multimodal regulatory analysis and personal epigenomics. *Nature methods* **2013**, *10*, 1213–1218. <https://doi.org/10.1038/nmeth.2688>.

156. Jin, W.; Tang, Q.; Wan, M.; Cui, K.; Zhang, Y.; Ren, G.; Ni, B.; Sklar, J.; Przytycka, T.M.; Childs, R.; et al. Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. *Nature* **2015**, *528*, 142–146. <https://doi.org/10.1038/nature15740>.

157. Buenrostro, J.D.; Wu, B.; Litzenburger, U.M.; Ruff, D.; Gonzales, M.L.; Snyder, M.P.; Chang, H.Y.; Greenleaf, W.J. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **2015**, *523*, 486–490. <https://doi.org/10.1038/nature14590>.

158. Sproul, D.; Gilbert, N.; Bickmore, W.A. The role of chromatin structure in regulating the expression of clustered genes. *Nature Reviews Genetics* **2005**, *6*, 775–781. <https://doi.org/10.1038/nrg1688>.

159. Dekker, J.; Rippe, K.; Dekker, M.; Kleckner, N. Capturing chromosome conformation. *Science* **2002**, *295*, 1306–1311. <https://doi.org/10.1126/science.1067799>.

160. Simonis, M.; Klous, P.; Splinter, E.; Moshkin, Y.; Willemsen, R.; De Wit, E.; Van Steensel, B.; De Laat, W. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nature genetics* **2006**, *38*, 1348–1354. <https://doi.org/10.1038/ng1896>.

161. Dostie, J.; Richmond, T.A.; Arnaout, R.A.; Selzer, R.R.; Lee, W.L.; Honan, T.A.; Rubio, E.D.; Krumm, A.; Lamb, J.; Nusbaum, C.; et al. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome research* **2006**, *16*, 1299–1309. <https://doi.org/10.1101/gr.5571506>.

162. Lieberman-Aiden, E.; Van Berkum, N.L.; Williams, L.; Imakaev, M.; Ragoczy, T.; Telling, A.; Amit, I.; Lajoie, B.R.; Sabo, P.J.; Dorschner, M.O.; et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *science* **2009**, *326*, 289–293. <https://doi.org/10.1126/science.1181369>.

163. Belaghal, H.; Dekker, J.; Gibcus, J.H. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods* **2017**, *123*, 56–65. <https://doi.org/10.1016/j.ymeth.2017.04.004>.

164. Lafontaine, D.L.; Yang, L.; Dekker, J.; Gibcus, J.H. Hi-C 3.0: improved protocol for genome-wide chromosome conformation capture. *Current Protocols* **2021**, *1*, e198. <https://doi.org/10.1002/cpz1.198>.

165. Meuwissen, T.; Hayes, B.; MacLeod, I.; Goddard, M. Identification of Genomic Variants Causing Variation in Quantitative Traits: A Review. *Agriculture* **2022**, *12*, 1713. <https://doi.org/10.3390/agriculture12101713>.

166. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w¹¹¹⁸; iso-2; iso-3. *Fly* **2012**, *6*, 80–92. <https://doi.org/10.4161/fly.19695>.

167. McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl variant effect predictor. *Genome biology* **2016**, *17*, 122. <https://doi.org/10.1186/s13059-016-0974-4>.

168. Wingender, E.; Dietze, P.; Karas, H.; Knüppel, R. TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucleic acids research* **1996**, *24*, 238–241. <https://doi.org/10.1093/nar/24.1.238>.

169. Sandelin, A.; Alkema, W.; Engström, P.; Wasserman, W.W.; Lenhard, B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic acids research* **2004**, *32*, D91–D94. <https://doi.org/10.1093/nar/gkh012>.

170. Tewhey, R.; Kotliar, D.; Park, D.S.; Liu, B.; Winnicki, S.; Reilly, S.K.; Andersen, K.G.; Mikkelsen, T.S.; Lander, E.S.; Schaffner, S.F.; et al. Direct identification of hundreds of expression-modulating variants using a multiplexed reporter assay. *Cell* **2016**, *165*, 1519–1529. <https://doi.org/10.1016/j.cell.2016.04.027>.

171. Dixit, A.; Parnas, O.; Li, B.; Chen, J.; Fulco, C.P.; Jerby-Arnon, L.; Marjanovic, N.D.; Dionne, D.; Burks, T.; Raychowdhury, R.; et al. Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* **2016**, *167*, 1853–1866. <https://doi.org/10.1016/j.cell.2016.11.038>.

172. Datlinger, P.; Rendeiro, A.F.; Schmidl, C.; Krausgruber, T.; Traxler, P.; Klughammer, J.; Schuster, L.C.; Kuchler, A.; Alpar, D.; Bock, C. Pooled CRISPR screening with single-cell transcriptome readout. *Nature methods* **2017**, *14*, 297–301. <https://doi.org/10.1038/nmeth.4177>.

173. Reilly, S.K.; Gosai, S.J.; Gutierrez, A.; Mackay-Smith, A.; Ulirsch, J.C.; Kanai, M.; Mouri, K.; Berenzy, D.; Kales, S.; Butler, G.M.; et al. Direct characterization of cis-regulatory elements and functional dissection of complex genetic associations using HCR-FlowFISH. *Nature genetics* **2021**, *53*, 1166–1176. <https://doi.org/10.1038/s41588-021-00900-4>.