Investigation into the genetic basis of

bovine horn development

Thesis submitted to The University of Adelaide in fulfilment

of the requirement for the degree of

Doctor of Philosophy

Johanna Ellen Aldersey

BSc (Animal Science) (Honours)

School of Animal and Veterinary Sciences

The University of Adelaide

June 2022

Table of Contents

Publication included within this thesis:

Aldersey J.E., Sonstegard T.S., Williams J.L. & Bottema C.D.K. (2020) Understanding the effects of the bovine POLLED variants. *Animal Genetics* **51**, 166-76.

Abstract

The presence of horns in ruminants has financial and welfare implications for the farming of cattle, sheep and goats worldwide. The genetic interactions that lead to horn development are not known. Hornless, or polled, cattle occur naturally, but the known causative DNA variants (Celtic, Friesian, Mongolian and Guarani) are in intergenic regions on bovine chromosome 1, and therefore, their functions are not known. The leading hypothesis is that horns are derived from cranial neural crest cells and the POLLED variants disrupt the migration or proliferation of these stem cells.

The bovine POLLED region was explored through bioinformatics analyses as horned animals may have genomic differences from hornless individuals or species near the POLLED DNA variants. The aim was to identify differences in genes synteny, lincRNA, and topologically associating domain (TAD) structure between horned and hornless individuals or species. Horned (n = 1) and polled (Celtic; n = 1) Hi-C sequences produced the same TAD structures. The POLLED genomic region was refined to a 520-kb region encompassing all four POLLED variants. *LOC526226* was unique to the bovine POLLED region and not conserved in the species analysed (water buffalo, sheep, goat, pig, horse, dog and human), and therefore, may be involved in horn development.

Histological analyses of cranial tissues from homozygous horned and polled fetuses at day 58 of development were conducted. The aims were to 1) determine the differences in the structure of horn bud region, and 2) compare immunohistochemistry staining of neural crest markers (SOX10 and NGFR) and RXFP2 between horned and polled tissues. Condensed cells were only observed in the horn bud mesenchyme of horned fetuses and may be progenitor cells. SOX10 and NGFR was not detected in these condensed cells, and therefore, these cells are not derived from the neural crest or have differentiated and no longer express neural crest markers. SOX10 and NGFR were detected in the peripheral nerves. RXFP2 was detected in peripheral nerves and in the horn bud epidermis.

Transcriptomic analyses of cranial tissues from the horned and polled fetuses at day 58 of development was also conducted. The aims were to 1) identify genes that may directly be affected by the polled variants, and 2) identify genes and pathways important for horn development. Near the POLLED region, three genes (*C1H21orf62, SON* and *EVA1C)* and one lincRNA (*LOC112447120*) were differentially expressed between horned and polled fetuses. Previously identified candidate genes, *RXFP2, TWIST2* and *ZEB2*, were also differentially expressed. New candidates for the horn development pathway were proposed based on the analyses (*MEIS2, PBX3, FZD8, CTNNB1* and *LEF1*). *LOC526226* was not differentially expressed in the horn bud. Differentially expressed genes had functions in axon guidance, cytoskeletal structure and the extracellular region, and therefore, these pathways may be vital for horn development.

Based on this research, it is now hypothesised that 1) horn stem cells are located in the mesenchyme and interact with the epidermis to initiate horn development, 2) the Celtic POLLED variant directly affects expression of *C1H21orf62*, *SON*, *EVA1C* and *LOC112447120*, and 3) the migration of horn stem cells is reduced by the effect of the POLLED variants upon *C1H21orf62*, *SON*, *EVA1C* and/or *LOC112447120* expression.

Thesis Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

The author acknowledges that copyright of published works contained within the thesis resides with the copyright holder(s) of those works.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Johanna Aldersey

31/05/2022

Acknowledgements

First and foremost, I would like to thank my supervisory panel, Dr Cynthia Bottema, Dr John Williams and Dr Tad Sonstegard. Cindy, I don't think I would have made it this far without you. I am extremely grateful for everything you have put into this project. In particular, for pushing me along when nothing seemed to be going right, and when I was having trouble pushing myself forward. You have truly gone above and beyond for me and this project. John, you are the reason this project was possible. I know that there were challenges and things did not go to plan, but I am happy with what we have achieved. I have valued your time, effort and feedback in putting my thesis together. Tad – I have enjoyed having you as a co-supervisor and have appreciated your interest in this topic. I valued your emails discussing new research in this field and potential mechanisms. I hope that you have found this research valuable and can't wait to see what comes from it! Thank you, John and Tad for the opportunity to visit Recombinetics laboratory in Minnesota. I learned so much and it is an experience I will always cherish.

Thank you, Dr Wayne Pitchford (panel) and Dr Gordon Howarth (Ph.D. coordinator), for your support and providing useful feedback on my project. I also acknowledge the University of Adelaide for providing me with the opportunity to pursue a Ph.D and for financially supporting me.

Thank you, Meat and Livestock Australia, for providing me with a top-up scholarship and the opportunity to participate in the postgraduate conferences and workshops throughout my program. I have thoroughly enjoyed participating in the program and value the skills I have learned and the people I met during this time.

Thank you, Dr Tong Chen, for teaching me everything in the lab from extracting DNA and RNA, running PCRs, cell culture and immunohistochemistry. I am sorry for all of my mistakes – I know that I probably gave you a few headaches! I appreciate your patience and

Chapter 1:

Literature review

1.1 Statement of Authorship

Title of Paper	Understanding the effects of the bovine POLLED variants		
Publication Status	X Published	☐	Accepted for Publication
	☐ Submitted for Publication	☐	Unpublished and Unsubmitted work written in manuscript style
Publication Details	Aldersey J.E., Sonstegard T.S., Williams J.L. & Bottema C.D.K. (2020) Understanding the effects of the bovine POLLED variants. Animal Genetics 51, 166-76.		

Principal Author

Name of Principal Author (Candidate)	Johanna E Aldersey		
Contribution to the Paper	Reviewed literature, wrote manuscript		
Overall percentage (%)	90%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	6/03/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

  i.   the candidate's stated contribution to the publication is accurate (as detailed above);

  ii. permission is granted for the candidate in include the publication in the thesis; and

  iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Tad S. Sonstegard		
Contribution to the Paper	Provided unpublished information in field Interpretation of information Critical review of article		
Signature		Date	May 6, 2022

Name of Co-Author	John L. Williams		
Contribution to the Paper	Interpretation of information Critical review of article		
Signature		Date	09/03/2022

Name of Co-Author	Cynthia D. K. Bottema		
Contribution to the Paper	Interpretation of information Critical review of article		
Signature		Date	08/03/2022

1.2 Publication

Understanding the effects of the bovine POLLED variants

J. E. Aldersey* [iD], T. S. Sonstegard[†] [iD], J. L. Williams* [iD] and C. D. K. Bottema* [iD]

*Davies Research Centre, School of Animal and Veterinary Sciences, University of Adelaide, Roseworthy Campus, Adelaide, SA5371, Australia. [†]Acceligen Inc., Eagan, 55121, MN, USA.

Summary

Horns are paired appendages on the head of bovine species, comprising an inner bony core and outer keratin sheath. The horn bud forms during early fetal development but ossification of the developing horn does not occur until approximately 1 month after birth. Little is known about the genetic pathways that lead to horn growth. Hornless, or polled, animals are found in all domestic bovids. Histological studies of bovine fetuses have shown that the horn bud does not form in polled individuals. There are currently four known genetic variants for polledness in cattle on BTA1. All of the variants are intergenic, but probably affect regulation of nearby genes or long non-coding RNAs. Transcriptomic studies suggest that the expression of two nearby long non-coding RNAs are affected by the Celtic POLLED variant, but further studies are required to confirm these data. Candidate genes located elsewhere in the genome are involved in regulating bone formation and epithelial-to-mesenchymal transition. Expression of one of these candidate genes, *RXFP2*, appears to be reduced in the fetal horn bud of polled animals carrying the Celtic variant compared with horned individuals. Investigating horn ontogenesis and the genetic pathway by which the POLLED variants prevent horn development has implications for cattle breeding. If the genetic basis of horn bud formation and polledness is better understood, then new targets may be identified for precision genome editing to create polled individuals.

Keywords Bovidae, cattle, Celtic, epithelial-to-mesenchymal transition, facial bone, horns, scurs

Introduction

Horns are cranial appendages of bovine species, which include antelope, goats, sheep and cattle. The primary function of horns is male competition for mates (Lundrigan 1996), but they are also used for protection against predators and to aid in competition for resources (Stankowich & Caro 2009), and they may be involved in thermoregulation (Pares-Casanova & Caballero 2014). However, domestic cattle with horns pose a risk to other cattle and handlers (Knierim *et al.* 2015), and can result in economic losses because of damaged hides and bruised tissue which must be trimmed when the meat is processed (Mendonca *et al.* 2016; Youngers *et al.* 2017).

In order to avoid issues related to horns, calves are disbudded using a hot iron, scoop dehorners or caustic paste

to prevent horn growth (Animal Health Australia 2014; Cozzi *et al.* 2015). The pain and distress caused to animals by disbudding and dehorning procedures is well documented (Knierim *et al.* 2015). Beyond this distress, there is the potential for the wound site to become infected and compromise animal growth, and the procedure is an additional labour cost to producers (Stafford & Mellor 2005; Bates *et al.* 2015; Bates *et al.* 2016).

Welfare guidelines recommend that preference should be given to breeding hornless, or polled, cattle over dehorning (Animal Health Australia 2014). However, introgression of the genetic variants for polled into specialised breeds (e.g. dairy, beef and tropically adapted breeds) leads to genetic loss of production traits (e.g. milk yield). This is because polledness is usually introduced into a herd by breeding with animals that have lower genetic merit or by crossing with another breed.

Advances in precision genome editing have the potential to introduce variants into the genome without compromising genetic gain. Gene editing allows a desirable phenotype to be introgressed into a population through a known DNA variant. Alternatively, a genetic target can be identified and altered (e.g. by an amino acid substitution) to observe the effect on the phenotype. Although genetic variants for

Address for correspondence

J. E. Aldersey, Davies Research Centre, School of Animal and Veterinary Sciences, University of Adelaide, Roseworthy Campus, Adelaide, SA 5371, Australia.
E-mail: johanna.aldersey@adelaide.edu.au

1

polled are known, the pathways that lead to horn formation and the mechanisms by which the complex bovine POLLED variants result in the hornless phenotype are unknown.

Horn morphology, development and inheritance

Horn and scur morphology

Horns of bovids are permanent, paired and symmetrical appendages that vary vastly in morphology between species and even breeds (Davis *et al.* 2011). Horns have two main parts: a 'dead' keratin outer sheath and a bony inner core of 'living' tissue (Zhu *et al.* 2016). Between the keratin sheath and bony core are several layers of tissue: the periosteum (tissue that lines the bones), subcutaneous connective tissue, dermis and epidermis (Davis *et al.* 2011). True horns have a bony core that is attached to the frontal bones and a frontal sinus that extends into the horn spike.

Scurs are horn-like appendages that can occur in bovids, but tend to be shorter than true horns (Capitan *et al.* 2011). The phenotype of scurs varies, ranging from small 'scabs' in the horn bud to appendages as long as 15 cm (Capitan *et al.* 2011). Scurs and horns have two main anatomical differences: (1) the scur is not anchored to the skull and the frontal sinus does not continue into the horn spike; and (2) the bony core of scurs is densely ossified compared with the pneumatised bony core of horns (Capitan *et al.* 2011).

Horn and scur inheritance

The inheritance of horns, polledness and scurs in cattle has been studied since the early 1900s. Understanding the pattern of inheritance was a challenging task for early researchers owing to the epistatic relationship that POLLED has with other loci, and the subsequent difficulties in inferring the genotype of an individual (reviewed by Prayaga 2007). Two types of scurs have been identified in cattle, Type I and Type II, and these have distinct inheritance patterns. In summary:

- Horned (p) is the wt state in cattle and is recessive to POLLED (P).
- Type I scurs are epistatic to POLLED and appear to be sex influenced; however, the inheritance pattern of scurs is unclear (White & Ibsen 1936; Blackwell & Knox 1958; Long & Gregory 1978; Wiedemar *et al.* 2014). Difficulty in determining the inheritance pattern of scurs is attributed to problems with phenotyping, inconsistent age of scur development, sex influence, epistasis with POLLED loci and genetic heterogeneity within breeds (Asai *et al.* 2004; Tetens *et al.* 2015; Grobler *et al.* 2018). Evidence supports the presence of a SCURS locus on BTA19 (Asai *et al.* 2004), and potential loci on BTA2, 9 and 10 (Tetens *et al.* 2015). Early research suggested that homozygous polled males could be polled or scurred (White & Ibsen 1936; Long & Gregory 1978); however, more recent studies have genotyped scurred cattle and found that they were always heterozygous polled (Wiedemar *et al.* 2014; Grobler *et al.* 2018). It was also assumed that scurred females were always homozygous at the SCURS locus (White & Ibsen 1936; Long & Gregory 1978); however, homozygosity mapping of BTA19 in scurred females did not identify a shared homozygous haplotype (Tetens *et al.* 2015).
- Type II scurs are the result of a mutation in *TWIST1* as observed in French Charolais cattle (Capitan *et al.* 2009; Capitan *et al.* 2011). The Type II scur phenotype is dominant over horns, but not over polled (A. Capitan, personal communication). Animals homozygous for the *TWIST1* mutation have not been identified, suggesting embryonic lethality.
- Horns in some zebu cattle breeds may be epistatic to POLLED in males rather than recessive (Smith 1927). In a cross between horned African zebu breeds and Angus, all female progeny were polled, but male progeny had one of three phenotypes: horned, scurred and polled (Smith 1927). This led to the suggestion that another gene is involved in this mode of inheritance, denoted as African horn (Ha) (White & Ibsen 1936). However, the existence of this locus has not been confirmed.

Development of horns

Originally, horn development was thought to be an outgrowth of the skull to form the horn spike. However, horns develop from a separate centre of ossification within the horn bud. Dove (1935) conducted a series of horn bud tissue transplants in young calves and goat kids to identify the origin of horn development and found that horn growth arises from the dermis and hypodermis, and not from the frontal bone. Bony processes develop in the horn bud, and as the neonate ages, the bone attaches to the skull and simultaneously grows outwards to produce the horn spike.

The horn bud was originally reported to be first visible in bovine fetuses at 60 days (Evans & Sack 1973). However, recently, the horn bud was observed by the authors at 58 days of development (Aldersey J.E., Sonstegard T.S., Williams J.L. & Bottema C., unpublished data). At 58 days, there is a ring of depressed tissue at the position where the horn bud develops, which is not visible in polled fetuses of the same age. At 70 days, the horn bud is reported to be well defined and appears as a small, yellowish spot on the fetal head (Wiener *et al.* 2015). By 90 days, the horn bud becomes slightly indented compared with the surrounding smooth skin (Wiener *et al.* 2015).

There are several histological differences between the horn bud and nearby frontal skin throughout bovine fetal development (Table 1) (Capitan *et al.* 2012; Allais-Bonnet *et al.* 2013; Wiener *et al.* 2015). Firstly, the epidermis of the horn bud is thicker than the epidermis of the frontal skin

(Wiener *et al.* 2015). Secondly, hair follicle development occurs later in the horn bud than surrounding tissue; hair follicles are present at 3–4 months of gestation in frontal skin but are not observed in the horn bud until 5–6 months of gestation (Wiener *et al.* 2015). Lastly, the horn bud has thick nerve bundles whereas nerve bundles are absent in frontal skin (Capitan *et al.* 2012; Allais-Bonnet *et al.* 2013; Wiener *et al.* 2015). Similar observations have been made for yak fetuses (Li *et al.* 2018). There is no evidence of ossification in the fetal horn bud (Wiener *et al.* 2015), and horn growth and ossification occur approximately 1 month after birth (Dove 1935). Thus, the horn bud differentiates during early fetal development but horn growth does not occur until after the calf is born.

POLLED genetic variants

The POLLED genetic locus for cattle was first localised to bovine chromosome 1 (BTA1) by linkage mapping (Georges *et al.* 1993), and the position was later refined to the centromeric region in several studies (Schmutz *et al.* 1995; Brenneman *et al.* 1996; Harlizius *et al.* 1997). Four DNA sequence variants have subsequently been identified on BTA1 that are associated with the polled phenotype: Celtic POLLED (P_C), Friesian POLLED (P_F), Mongolian POLLED (P_M), and Guarani POLLED (P_G) (Medugorac *et al.* 2012; Allais-Bonnet *et al.* 2013; Rothammer *et al.* 2014; Medugorac *et al.* 2017; Utsunomiya *et al.* 2019). All known variants are dominant, and cattle carrying a single POLLED variant will be either polled or scurred, depending on their genotype at the SCURS loci. It is likely that other unidentified POLLED variants exist in different populations and breeds (e.g. Shuxuan; Chen *et al.* 2017a).

Table 1 Bovine fetal development of frontal skin and horn bud tissue (Wiener *et al.* 2015).

Gestation length (months)	Frontal skin	Horn bud
2–3	Epidermis has three layers of vacuolated keratinocytes	Epidermis has seven layers of vacuolated keratinocytes
3–4	Epidermis has four layers	Epidermis has 12 layers
	Immature hair follicles present	No hair follicles present
	No nerve bundles present	Nerve bundles present
5–6	Epidermis has six layers	Epidermis has 12 layers
	Hair follicles and sebaceous glands present	Hair follicles and sebaceous glands present
	No nerve bundles present	Nerve bundles present and more pronounced
7–8	Keratinocytes are no longer vacuolated	Keratinocytes are no longer vacuolated
	Hair follicles and sebaceous glands present	Hair follicles and sebaceous glands present

Celtic POLLED variant

The Celtic POLLED variant was first identified in several European beef breeds originating from Celtic geographical areas. The variant is a complex insertion and deletion (indel). A 212 bp sequence (1 705 834–1 706 045 bp)[1] is duplicated and replaces a sequence of 10 bp (1 706 051–1 706 060-bp) that is 6 bp downstream of the original sequence (Fig. 1) (Medugorac *et al.* 2012). Independent association studies found that the indel was the only variant at this site that segregated completely with polledness (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014). The Celtic variant was found to be functionally responsible for polledness by gene editing the variant into wt (horned) crossbred Holstein fibroblasts, which were cloned to produce polled calves (Carlson *et al.* 2016). The progeny of horned dams and the gene-edited Holstein bulls produced from these fibroblasts, which were shown to only carry the Celtic allele and no other unintended edits, were also polled. The Celtic POLLED variant is located between the genes *IFNAR2* and *OLIG1* on BTA1 and does not appear to disrupt any known coding sequence, splice site or intronic region, or any known regulatory regions (Medugorac *et al.* 2012). The variant may interrupt a predicted *HAND1* enhancer site (Nguyen *et al.* 2018), although this is yet to be confirmed experimentally.

Friesian POLLED variant

First identified in Holstein-Friesian cattle, the Friesian POLLED variant is approximately 200 kb downstream of the Celtic variant and is an 80 128 bp duplication of the sequence between 1 909 352 and 1 989 480 bp (Fig. 1) (Medugorac *et al.* 2012; Allais-Bonnet *et al.* 2013; Rothammer *et al.* 2014). The duplicated segment is located immediately after the original sequence and is in the same orientation. It differs from the reference sequence by one T→A transversion at the third position and by a 2 bp deletion (TG) at the 45th position. Further research confirmed that this variant segregated in polled Holsteins that did not carry the Celtic POLLED allele (Wiedemar *et al.* 2014). As with the Celtic POLLED variant, the Friesian POLLED variant does not disrupt any known coding sequence, splice site or intronic region, or any known regulatory regions (Rothammer *et al.* 2014).

Mongolian POLLED variant

A third bovine POLLED variant has been discovered in Mongolian yaks and Mongolian Turano cattle (Medugorac *et al.* 2017). There are horned and polled individuals in these populations, and owing to their isolation, this

---

[1]All genomic locations refer to the UMD3.1 build. Updated coordinates are available on the OMIA website (omia.org/OMIA000483/9913).
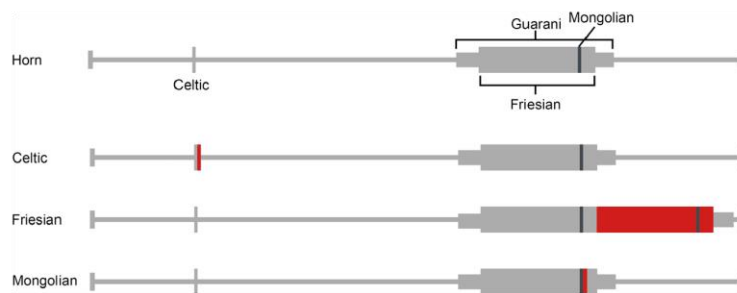
Figure 1 Celtic, Friesian and Mongolian polled variants. The grey rectangles represent the duplicated sequence, and red rectangles represent the insertion site of the duplications. The Guarani variant has not been completely defined at the time of publication.

POLLED variant was suspected to be a spontaneous mutation that had not previously been described.

Whole genome sequencing of a homozygous and heterozygous polled yak localised the Mongolian POLLED locus to an 800 kb region on BTA1 (Medugorac *et al.* 2017). The position of the locus was further refined by genotyping and two variants associated with the polled phenotype in Turano cattle and yaks were identified. The first variant was a 219 bp duplication–insertion 61 bp downstream from the original sequence (P_{219ID} beginning 1 976 128 bp) and the second was a 6 bp deletion and 7 bp insertion 621 bp upstream from P_{219ID} (Medugorac *et al.* 2017) (Fig. 1). Within the 219 bp duplicated sequence, an 11 bp motif (5′-AAAGAAGCAAA-3′) is entirely conserved among Bovidae, and therefore, may be functionally important (Medugorac *et al.* 2017). Intriguingly, the 219 bp sequence is also located within the Friesian variant and Guarani variant (see below), and therefore, the 219 bp sequence (and consequently, the 11 bp conserved motif) is duplicated in the Mongolian, Friesian and Guarani variants (Fig. 1). Haplotype analysis showed that the Mongolian variant is located on a bovine DNA segment, and the variant was introgressed from Turano cattle into Mongolian yaks (Medugorac *et al.* 2017).

Guarani POLLED variant

A fourth variant, Guarani POLLED (P_G), has been recently identified in Nellore cattle (*Bos indicus*) from Brazil (Utsunomiya *et al.* 2019). The polled phenotype in Nellore cattle was traced to a single polled bull, which implies that polledness in the breed is not the result of one of the previously discovered variants. Whole genome sequencing of polled Nellore bulls identified an approximately110 kb sequence (1 893 790–2 004 553 bp) within the POLLED region with increased coverage, indicating a copy number variation caused by an approximately 110 kb duplication. The insertion location of the duplication is yet to be determined. Intriguingly, SNP genotyping of the P_G region in the polled Nellore bulls confirmed that the Guarani variant originated from *Bos taurus* (Utsunomiya *et al.* 2019).

Phenotypes associated with POLLED

Polled fetuses carrying the Celtic variant do not develop horn buds, forming only smooth tissue that is histologically indistinguishable from frontal skin tissue (Allais-Bonnet *et al.* 2013; Wiener *et al.* 2015). Horn bud development is also absent in yak fetuses carrying the Mongolian variant (Li *et al.* 2018), but has not been investigated in fetuses homozygous for the Friesian and Guarani variants.

In addition to the complete absence of horn growth, the POLLED variants are associated with several other phenotypes. The skull morphology of polled cattle is characterised by a narrower and peaked poll (Dove 1935); however, it is unclear whether this phenotype is a result of the POLLED variants affecting skull development or due to the absence of horns, which would cause the outgrowth of the frontal sinus. Polled cattle that carry the Celtic or Friesian variants also have a second row of eyelashes (Fig. 2). Allais-Bonnet *et al.* (2013) examined 78 polled cattle and characterised the phenotype as additional eyelash growth and hypertrichosis (excessive hair growth) of the eyelid. There have been no reports regarding atypical eyelash growth for cattle carrying other variants. There are also no reports that this eyelash phenotype has any detrimental effects on polled individuals.

Bulls from Angus and other polled breeds are more likely to develop a spiral deviation of the penis, a so-called 'corkscrew penis' (Blockey & Taylor 1984). The corkscrew penis tends to occur in bulls at least 3 years old and reduces pregnancy rates owing to poor servicing (McDiarmid 1981; Blockey & Taylor 1984). A spiral deviation of the penis has been detected in 11–27% of polled breeds (Angus, Poll Hereford, Poll Shorthorn, Red Poll and Murray Grey) compared with 0–1% of horned Herefords (McDiarmid 1981; Blockey & Taylor 1984). However, it is not known if there is a direct association between the polled phenotype and corkscrew penis.

There have also been reports of preputial abnormalities (preputial prolapse) in polled (P_C) Charolais bulls, caused by poor development or absence of retractor muscles of the prepuce (Prayaga 2007; Allais-Bonnet *et al.* 2013). When assessed for preputial prolapse, two of two homozygous polled (P_C/P_C) Charolais bulls and 11 of 14 heterozygous

Figure 2 Eyelashes of horned cow (a) and double eyelashes of polled cow ($P_C P_C$) (b).

polled (P_C/p) Charolais bulls had this defect. However, this abnormality has not been observed in other breeds carrying the Celtic or Friesian POLLED variants or in horned animals. Therefore, the preputial defect appears to be a breed-specific loci interacting with or in LD the POLLED variant (Allais-Bonnet *et al.* 2013). The prepuce defect makes sheath cleaning prior to semen collection difficult; however, it does not appear to affect other reproductive traits or the health of the affected individuals (Allais-Bonnet *et al.* 2013). There have been no reports of other phenotypes associated with polledness in carriers of the Mongolian and Guarani variants.

Candidate genes

As the POLLED variants are not located in any known genes, long non-coding RNAs or microRNAs, it is postulated that the variants affect the expression of genes or non-coding RNAs by disrupting regulatory DNA elements, such as enhancers. The POLLED variants are located within one predicted topologically associating domain (TAD) (1 226 028–2 201 452 bp) containing 23 protein coding genes and non-coding RNAs (Fig. 3) (Wang *et al.* 2018). TADs are regions of a genome where there are more interactions between loci within a domain than between loci located in different domains (Dixon *et al.* 2012; Szalaj & Plewczynski 2018). There is evidence that TAD boundaries act as genetic insulators, ensuring appropriate enhancer–promoter interactions (Dixon *et al.* 2012; Krivega & Dean 2017). Disruption of TAD boundaries can lead to increased interactions between TADs, resulting in an altered phenotype (Yu & Ren 2017; Furlong & Levine 2018).

Two long intergenic non-coding RNAs (lincRNA) have been described within the POLLED predicted TAD, *LincRNA#1* and *LincRNA#2* (*LOC100848368* and *LOC112447133* respectively, in the ARS-UCD1.2 assembly)

(Allais-Bonnet *et al.* 2013). LincRNAs are defined as non-coding RNA longer than 200 nucleotides which do not occur within protein coding genes (Deniz & Erman 2017). LincRNAs are expressed at low levels and appear to be tissue or cell type specific (Deniz & Erman 2017). They can regulate gene expression by various methods, including binding to mRNA, miRNA and chromatin modifying complexes, and interacting with transcription factors (Deniz & Erman 2017).

Other candidate genes that may be involved in horn development, outside the POLLED TAD on BTA1, include genes that are (1) associated with the polled phenotype in other bovid species, (2) have variants associated with syndromes that include a polled phenotype or (3) have variants associated with distichiasis (abnormal eyelash growth) in other species (Table 2). One of these candidate genes, *FOXC2*, which is associated with distichiasis in humans, was identified as a horn-specific gene in a study of Bovidae transcriptomes (Wang *et al.* 2019). This study identified 624 horn-specific genes using transcriptomes from 16 tissues, including horn sprouts from goats and sheep, and fetal horn bud and frontal skin from sheep (Wang *et al.* 2019), but no other candidate genes (Table 2) were found to be horn-specific. *FOXC2* is highly expressed in horn tissue and bone (Wang *et al.* 2019). *FOXC2* was also found to be differentially expressed between the horn bud and frontal skin of horned (p/p) bovine fetuses at 90 days of development (Allais-Bonnet *et al.* 2013). These studies suggest that *FOXC2* may be involved in horn development.

Gene and protein expression in horned vs. polled horn bud

Gene expression studies of horn bud tissue from horned and polled cattle can be used to identify genetic pathways

15
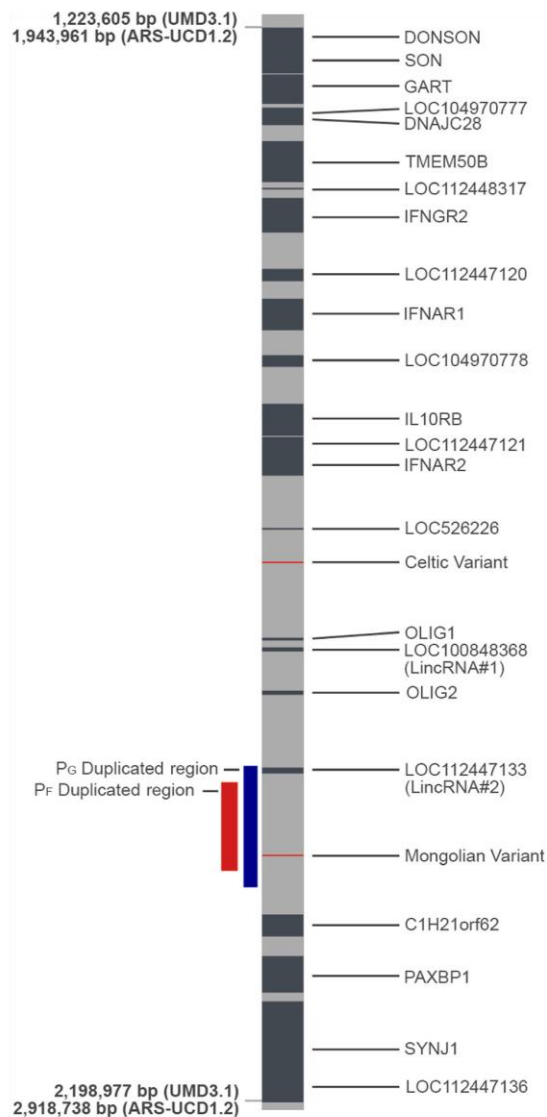
Figure 3 Gene map of predicted topologically associating domain containing POLLED variants on BTA1.

involved in normal horn development and provide clues about the mechanism by which POLLED variants prevent horns. Several studies have investigated gene expression and protein abundance in bovine fetal and neonatal horn bud tissue (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014; Li *et al.* 2018).

Gene expression in fetal horn bud tissue

The first study where gene expression was investigated in the fetal horn bud examined the genes and lincRNA 500 kb

upstream and downstream of the Celtic variant: *GART*, *TMEM50B*, *IFNGR2*, *IFNAR1*, *IL10RB*, *IFNAR2*, *OLIG1*, *LincRNA#1*, *OLIG2*, *LincRNA#2*, *C1H21orf62* and *PAXBP1* (Allais-Bonnet *et al.* 2013). Candidate genes *RXFP2*, *FOXL2*, *ZEB2*, *TWIST1*, *TWIST2* and *FOXC2* were also analysed. Biopsies from the horn bud and frontal skin regions of seven polled (P_c/p) and seven horned (p/p) fetuses at 90 days of pregnancy were examined using qRT-PCR. *RXFP2* and *LincRNA#1* were differentially expressed between horned and polled horn bud tissue (Table 3). Expression of *RXFP2* was lower in the polled fetuses than in horned fetuses ($P < 0.05$). The expression of *LincRNA#1* was slightly higher in the horn bud region of polled vs. horned fetuses ($P = 0.052$) (Allais-Bonnet *et al.* 2013). Although the differential expression of *LincRNA#1* was not quite significant between horned and polled horn bud tissue, it should be noted that lincRNAs are difficult to detect. This study also did not assess gene expression in homozygous polled fetuses, in which a larger effect on gene expression may be expected. In addition, differential expression of genes leading to horn bud formation is likely to occur before 90 days of development, as the horn bud is apparent before 60 days of gestation. Therefore, important expression differences in the genes may not have been observed.

An RNAseq study of one horned fetus (150 days post fertilisation) and one polled fetus (158 days post fertilisation) identified significant differences in the gene expression of *OLIG1*, *OLIG2*, *C1H21orf62*, *RXFP2*, *FOXL2* and *LincRNA#2* (Wiedemar *et al.* 2014). These RNAseq results were subsequently examined by qRT-PCR using horn bud and frontal skin biopsies from 21 fetuses that ranged from 70 to 175 days of fetal development. *LincRNA#2*, *RXFP2* and *FOXL2* appeared to be more highly expressed in horned fetuses than polled fetuses at all time points; however, these expression differences were not statistically significant.

The differences in fetal age and uncertainty arising from the small sample size makes it difficult to compare the results from Allais-Bonnet *et al.* (2013) and Wiedemar *et al.* (2014). *RXFP2* was reported to be differentially expressed in both studies, whereas *LincRNA#1*, *LincRNA#2* and *FOXL2* were only reported to be differentially expressed in one of the studies. *RXFP2* had reduced expression in polled horn bud tissue compared with wt horn bud tissue of the fetus. *RXFP2* is on BTA12, and therefore, the mechanism by which the Celtic variant affects *RXFP2* expression is not clear. Interestingly, an insertion in *RXFP2* has been linked with polledness in some European sheep breeds (Wiedemar & Drogemuller 2015; Luhken, *et al.* 2016) and SNPs in *RXFP2* have been associated with ovine horn size and shape (Pan *et al.* 2018). Thus, *RXFP2* may play a role in horn growth and shape, rather than horn bud formation *per se*.

A proteomic study of three polled (P_M) and three-horned yak fetuses at 80–90 days development investigated

Table 2 Candidate genes that may be involved in horn development, but are not located within the POLLED locus on BTA1.

Gene (location)	Function[1]	Association with polledness	Reference
RXFP2 (BTA12)	Relaxin family peptide receptor 2: encodes a G-coupled, 7-transmembrane receptor	Variants in RXFP2 associated with polledness and horn shape in sheep	Wiedemar & Drogemuller (2015); Luhken et al. (2016); Pan et al. (2018)
FOXL2 (BTA1)	Forkhead Box L2: may be involved in ovarian development and function	Loss of function of both FOXL2 alleles causes Polled Intersex Syndrome in goats	Boulanger et al. (2014)
ZEB2 (BTA2)	Zinc Finger E-Box Binding Homeobox 2: represses transcription by interacting with activated SMADs	Deletion including ZEB2 causes Polled and Multisystemic Syndrome in cattle	Capitan et al. (2012)
TWIST1 (BTA4)	Twist Family BHLH Transcription Factor 1: involved in embryonic development including cranial suture closure	Mutation causing frameshift in TWIST1 causes Type II scurs in cattle and haploinsufficiency causes craniosynostosis (premature fusion of skull)	Capitan et al. (2011)
TWIST2 (BTA3)	Twist Family BHLH Transcription Factor 2: may inhibit osteoblast maturation	Mutation in TWIST2 causes Setleis syndrome in humans involving abnormal skull morphology and distichiasis (eyelashes on inner eyelid)	Cervantes-Barragan et al. (2011)
FOXC2 (BTA18)	Forkhead Box C2: undetermined function but may be involved with mesenchymal tissue development	Mutations in FOXC2 cause syndromes with distichiasis in humans	Sargent et al. (2014); Zhang et al. (2016)

[1]Gene Cards Suite (2019).

Table 3 Summary of published differentially expressed genes revealed by qPCR comparison of wt and polled fetal horn bud tissue.

Gene	70 day old fetuses (Wiedemar & Drogemuller 2015)	90 day old fetuses (Allais-Bonnet et al. 2013)
GART	–	NDE
TMEM50B	–	NDE
IFNGR2	–	NDE
IFNAR1	–	NDE
IL10RB	–	NDE
IFNAR2	–	NDE
OLIG1	–	NDE
LincRNA#1	–	↑[1]
OLIG2	NDE	NDE
LincRNA#2	↓	NDE
C1H21orf62	NDE	NDE
PAXBP1	–	NDE
FOXL2	↓	NDE
RXFP2	↓	↓[2]
TWIST1	–	NDE
ZEB2	–	NDE
TWIST2	–	NDE
FOXC2	–	NDE

–, Not analysed by qPCR; NDE, not differentially expressed; ↓, decreased expression in polled vs. horned horn bud; ↑, increased expression in polled vs. horned horn bud.
[1]Significance = 0.052.
[2]Significance < 0.050.

differentially abundant proteins (DAPs) in tissue from the horn bud region (Li et al. 2018). This study identified 29 upregulated proteins and 71 downregulated proteins in the polled fetus compared with horned fetuses. Classification of proteins by Protein Analysis Thorough Evolutionary Relationships (PANTHER) showed that upregulated DAPs were related to metabolic activities, whereas downregulated DAPs were related to cell junction, cytoskeleton formation and cell component organisation. Overall, the DAPs had functions involving cell adhesion, cell motility, keratinocyte differentiation, cytoskeleton organisation, osteoblast differentiation and fatty acid metabolism. Although there were DAPs involved in osteoblast differentiation, bone development in the horn bud does not occur at this stage of fetal development. Proteins associated with cell structure and organisation may be differentially abundant owing to the structural differences between horned and polled fetal horn bud. For example, by 80–90 days of development nerve bundles are present in the wt horn bud and absent in the polled horn bud region.

Gene expression in neonatal horn bud tissue

Gene expression has been also examined in horn bud tissue of neonatal calves (Mariasegaram et al. 2010). A study of cDNA from the horn bud tissue of 1–2 week old Brahman calves with polled, scurred and horned phenotypes revealed no difference in expression of genes located within the predicted POLLED TAD region (Mariasegaram et al. 2010). The microarray used in the study included DONSON, SON, GART, TMEM50B, IFNGR2, IFNAR1, IL10RB, IFNAR2, OLIG1, OLIG2 and PAXBP1. However, there were no probes for LOC194970777, DNAJC28, LOC112448317, LOC112447120, LOC104970778, LincRNA#1, LincRNA#2 or C1H21orf62. The array included most functional candidate genes outside of the POLLED region, namely RXFP2, TWIST1, FOXL2 and FOXC2, but not ZEB2 and TWIST2. These functional candidate genes were not differentially expressed. However, the microarray analysis identified 93 other genes that were differentially expressed between horn and polled calves. Genes with greater expression in polled

calves were structural components of cell junctions, and genes with lower expression had functions relating to extracellular regions (Mariasegaram *et al*. 2010).

Candidate pathways

### Mammalian embryonic origins of bone and bone formation

As horns are partly bone, pathways involved in bone formation may be disrupted by the POLLED variants. Bone tissue is derived from the mesoderm and cranial neural crest. During embryo development, the mesoderm differentiates into paraxial, intermediate and lateral mesoderm. Only the paraxial and lateral mesoderm form bone; the former is the source of the axial skeleton (ribs, vertebrae and parietal bones of the skull) and the latter creates the appendicular skeleton (limbs) (Jin *et al*. 2016; Sheebaa *et al*. 2016). The cranial neural crest cells migrate to form the frontal and facial bones (Wu *et al*. 2017), and these cells are the most likely candidates to form horns in Bovidae species. In an immunohistochemistry study of sheep fetuses, cells expressing genetic markers for neural crest cells (SOX10 and NFGR) were found in the fetal horn bud at 90 days of development (Wang *et al*. 2019).

A gene within the predicted POLLED locus TAD, *PAXBP1*, potentially plays a role in facial bone development (Blake & Ziman 2014). In humans, PAXBP1 is a binding protein that links transcription factors PAX3 and PAX7 to histone methylation machinery (The UniProt Consortium 2019). The PAXBP1 and PAX3/PAX7 interaction is primarily associated with myogenesis, but there is evidence that PAX3/PAX7 is involved cranial facial development (Blake & Ziman 2014; Monsoro-Burq 2015). A missense mutation in *PAXBP1* leads to dysmorphia in facial bones of humans (Alharby *et al*. 2017) and PAX3 is involved in neural crest specification, delamination, cell survival during migration and differentiation (Monsoro-Burq 2015). Currently, there is no experimental evidence connecting *PAXBP1* to horn growth; however, gene expression of cranial neural crest tissue in horned and polled fetuses has not been assessed. Understanding the lineage of cells that form the horn bud would aid in determining which developmental pathways are disrupted by the POLLED variants.

The cranial facial bones are produced via intramembranous ossification whereby bone tissue forms directly from the condensed mesenchymal cells (Ishii *et al*. 2015; Jin *et al*. 2016; Wu *et al*. 2017). Several candidate genes are involved in regulation of intramembranous ossification. The *TWIST* genes regulate ossification, and mutations in *TWIST1* often cause craniosynostosis, early closure of the cranial sutures (Hayashi *et al*. 2007; Connerney *et al*. 2008; Derderian & Seaward 2012; Huang *et al*. 2014). Additionally, there is evidence that the ligand of RXFP2, relaxin, induces oesteogenic differentiation through the activation of regulators of intramembranous ossification, namely, alkaline phosphatase, RUNX2 and BMP2 (Duarte *et al*. 2014). *RXFP2* expression is lowered in the horn bud region of polled fetuses. Thus, the formation of horn bud bone tissue could be prevented by reduced availability of the RXFP2 receptor. The POLLED variants may affect several other stages of horn bud formation, including blocking the bone precursor cells from successfully migrating to the horn bud or differentiating to bone tissue (Fig. 4). A comparison of the transcriptomes of cranial neural crest and horn bud tissue from horned and polled fetuses pre- and post-neural crest cell migration may resolve the affected pathways.

Epithelial-to-mesenchymal transition

Four of the candidate genes (*TWIST1*, *TWIST2*, *ZEB2* and *FOXC2*; Table 2) encode transcription factors that regulate the epithelial-to-mesenchymal transition (EMT). EMT occurs during embryo implantation, embryogenesis and organ development, and is one of the processes that results in the diversification of cell types and the development of tissues which create organs (Kalluri & Weinberg 2009). During EMT, epithelial cells undergo a series of biochemical changes to become mesenchymal cells (Kalluri & Weinberg 2009). For instance, as part of neural crest cell delamination, the epithelial cells of the neural crest change to migratory mesenchymal cells. Thus, altered gene expression of EMT related transcription factors may contribute to the polled phenotype. Reduced expression of *E-cadherin*, the protein that forms adhesion junctions between cells, is a key event in EMT. Interestingly, expression of the *E-cadherin* gene is directly repressed by the transcription factors encoded by *TWIST1*, *TWIST2*, *ZEB2* and *FOXC2* (Chen *et al*. 2017b). The
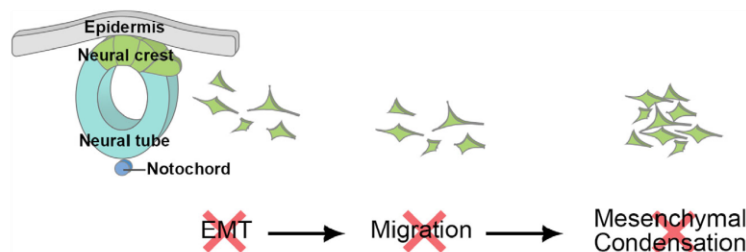
Figure 4 Hypothetical mechanisms whereby the POLLED variants may prevent horn bud formation. EMT, Epithelial-to-mesenchymal transition.

expression of these genes and various EMT markers (*E-cadherin*, *N-cadherin*, *occludin* and *vimentin*) has been examined in the horn bud for bovine fetuses at 90 days of development; however, the gene expression did not differ between horned and polled fetuses (Allais-Bonnet *et al.* 2013). This suggests that EMT is not occurring in the horn bud at 90 days of fetal development. To further explore the effect of the POLLED variants on EMT, expression of EMT candidate genes and markers should be assessed in cranial neural crest cells from the midbrain region in horned and polled fetuses. However, expression may need to studied before the horn bud is visible at 58 days, as the midbrain starts to form between 32 and 41 days of development in *Bos indicus* embryos (Assis *et al.* 2009).

Conclusions

There are four DNA sequence variants currently known to produce the polled phenotype in cattle; however, all of these variants are intergenic. Comparison of gene expression of horn bud tissue in polled and horned fetuses suggests that *LincRNA#1* and *LincRNA#2*, two long intergenic non-coding RNA located near the POLLED variants on BTA1, and *RXFP2* located on BTA12 could be involved in the development of horns. Based on these gene expression studies, the most likely hypothesis is that the POLLED variants affect the regulation of *LincRNA#1* and *LincRNA#2*. However, given there are phenotypic differences between horned and polled fetuses at 58 days of fetal development, the effect of the POLLED variants is likely to have occurred earlier. RNAseq and chromatin interaction studies of tissues from younger horned and polled fetuses would provide information on gene expression differences and regulatory DNA elements within the genomic POLLED region. This information would help to determine whether bone formation, cell migration, EMT and/or other processes are involved in the control of horn development in bovids.

Acknowledgements

References

Alharby E., Albalawi A.M., Nasir A. *et al.* (2017) A homozygous potentially pathogenic variant in the PAXBP1 gene in a large family with global developmental delay and myopathic hypotonia. *Clinical Genetics* **92**, 579–86.

Allais-Bonnet A., Grohs C., Medugorac I. *et al.* (2013) Novel insights into the bovine polled phenotype and horn ontogenesis in bovidae. *PLOS ONE* **8**, e63512.

Animal Health Australia (2014) *Australian Animal Welfare Standards and Guidelines for Cattle.* Commonwealth of Australia, Australian Animal Welfare Standards and Guidelines, Canberra. http://www.animalwelfarestandards.net.au/

Asai M., Berryere T.G. & Schmutz S.M. (2004) The scurs locus in cattle maps to bovine chromosome 19. *Animal Genetics* **35**, 34–9.

Assis N.A., Pereira F., Santos T., Ambrosio C., Leiser R. & Miglino M. (2009) Morpho-physical recording of bovine conceptus (*Bos indicus*) and placenta from days 20 to 70 of pregnancy. *Reproduction in Domestic Animals* **45**, 760–72.

Bates A.J., Eder P. & Laven R.A. (2015) Effect of analgesia and anti-inflammatory treatment on weight gain and milk intake of dairy calves after disbudding. *New Zealand Veterinary Journal* **63**, 153–7.

Bates A.J., Laven R.A., Chapple F. & Weeks D.S. (2016) The effect of different combinations of local anaesthesia, sedative and non-steroidal anti-inflammatory drugs on daily growth rates of dairy calves after disbudding. *New Zealand Veterinary Journal* **64**, 282–7.

Blackwell RL, Knox JH (1958) Scurs in a herd of Aberdeen-Angus cattle. *Journal of Heredity* **49**, 117–9.

Blake J.A. & Ziman M.R. (2014) Pax genes: regulators of lineage specification and progenitor cell maintenance. *Development* **141**, 737–51.

Blockey M.A.B. & Taylor E.G. (1984) Observations on spiral deviation of the penis in beef bulls. *Australian Veterinary Journal* **61**, 141–5.

Boulanger L., Pannetier M., Gall L. *et al.* (2014) FOXL2 is a female sex-determining gene in the goat. *Current Biology* **24**, 404–8.

Brenneman R.A., Davis S.K., Sanders J.O., Burns B.M., Wheeler T.C., Turner J.W. & Taylor J.F. (1996) The polled locus maps to BTA1 in a *Bos indicus* x *Bos taurus* cross. *Journal of Heredity* **87**, 156–61.

Capitan A., Grohs C., Gautier M. & Eggen A. (2009) The scurs inheritance: new insights from the French Charolais breed. *BMC Genetics* **10**, 33.

Capitan A., Grohs C., Weiss B., Rossignol M.N., Reverse P. & Eggen A. (2011) Newly described bovine type 2 scurs syndrome segregates with a frame-shift mutation in TWIST1. *PLOS ONE* **6**, e22242.

Capitan A., Allais-Bonnet A., Pinton A. *et al.* (2012) A 3.7 Mb deletion encompassing ZEB2 causes a novel polled and multisystemic syndrome in the progeny of a somatic mosaic bull. *PLOS ONE* **7**, e49084.

Carlson D.F., Lancto C.A., Zang B., Kim E.-S., Walton M., Oldeschulte D., Seabury C., Sonstegard T.S. & Fahrenkrug S.C. (2016) Production of hornless dairy cattle from genome-edited cell lines. *Nature Biotechnology* **34**, 479–81.

Cervantes-Barragan D.E., Villarroel C.E., Medrano-Hernandez A., Duran-McKinster C., Bosch-Canto V., Del-Castillo V., Nazarenko I., Yang A. & Desnick R.J. (2011) Setleis syndrome in Mexican-Nahua sibs due to a homozygous TWIST2 frameshift mutation and partial expression in heterozygotes: review of the focal facial dermal dysplasias and subtype reclassification. *Journal of Medical Genetics* **48**, 716–20.

Chen S.Y., Liu L.H., Fu M.Z., Zhang G.W., Yi J., Lai S.J. & Wang W. (2017a) Simultaneous introgression of three POLLED mutations into a synthetic breed of Chinese cattle. *PLOS ONE* **12**, e0186862.

Chen T., You Y.A., Jiang H. & Wang Z.Z. (2017b) Epithelial-mesenchymal transition (EMT): a biological process in the development, stem cell differentiation, and tumorigenesis. *Journal of Cellular Physiology* **232**, 3261–72.

Connerney J., Andreeva V., Leshem Y., Mercado M.A., Dowell K., Yang X.H., Lindner V., Friesel R.E. & Spicer D.B. (2008) Twist1

homodimers enhance FGF responsiveness of the cranial sutures and promote suture closure. *Developmental Biology* **318**, 323–34.

Cozzi G., Gottardo F., Brscic M. *et al.* (2015) Dehorning of cattle in the EU Member States: a quantitative survey of the current practices. *Livestock Science* **179**, 4–11.

Davis E.B., Brakora K.A. & Lee A.H. (2011) Evolution of ruminant headgear: a review. *Proceedings of the Royal Society B: Biological Sciences* **278**, 2857–65.

Deniz E. & Erman B. (2017) Long noncoding RNA (lincRNA), a new paradigm in gene expression control. *Functional & Integrative Genomics* **17**, 135–43.

Derderian C. & Seaward J. (2012) Syndromic craniosynostosis. *Seminars in Plastic Surgery* **26**, 64–75.

Dixon J.R., Selvaraj S., Yue F., Kim A., Li Y., Shen Y., Hu M., Liu J.S. & Ren B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–80.

Dove F.W. (1935) The physiology of horn growth: a study of the morphgenesis, the interaction of tissues, and the evolutionary processes of a Mendelian recessive character by means of transplantation of tissues. *Journal of Experimental Zoology* **69**, 347–405.

Duarte C., Kobayashi Y., Kawamoto T. & Moriyama K. (2014) RELAXIN enhances differentiation and matrix mineralization through relaxin/insulin-like family peptide receptor 2 (Rxfp2) in MC3T3-E1 cells in vitro. *Bone* **65**, 92–101.

Evans E.H. & Sack W.O. (1973) Prenatal development of domestic and laboratory mammals: growth curves, external features and selected references. *Anatomia, Hostologia, Embryologia* **2**, 11–45.

Furlong E.E.M. & Levine M. (2018) Developmental enhancers and chromosome topology. *Science* **361**, 1341–5.

Gene Cards Suite (2019) Gene cards human database. www.genecards.org

Georges M., Drinkwater R., King T. *et al.* (1993) Microsatellite mapping of a gene affecting horn development in *Bos taurus*. *Nature Genetics* **4**, 206–10.

Grobler R., Visser C., Capitan A. & Van Marle-Köster E. (2018) Validation of the POLLED Celtic variant in South African Bonsmara and Drakensberger beef cattle breeds. *Livestock Science* **217**, 136–9.

Harlizius B., Tammen I., Eichler K., Eggen A. & Hetzel D.J.S. (1997) New markers on bovine chromosome 1 are closely linked to the polled gene in Simmental and Pinzgauer cattle. *Mammalian Genome* **8**, 255–7.

Hayashi M., Nimura K., Kashiwagi K., Harada T., Takaoka K., Kato H., Tamai K. & Kaneda Y. (2007) Comparative roles of Twist-1 and Id1 in transcriptional regulation by BMP signaling. *Journal of Cell Science* **120**, 1350–7.

Huang Y.Y., Meng T., Wang S.Z., Zhang H., Mues G., Qin C.L., Feng J.Q., D'Souza R.N. & Lu Y.B. (2014) Twist1-and Twist2-haploin-sufficiency results in reduced bone formation. *PLOS ONE* **9**, e99331.

Ishii M., Sun J.J., Ting M.C. & Maxson R.E. (2015) The development of the calvarial bones and sutures and the pathophysiology of craniosynostosis. In: *Craniofacial Development* (Ed. by Y. Chai), pp. 131–56. Elsevier Inc, Waltham, MA.

Jin S.W., Sim K.B. & Kim S.D. (2016) Development and growth of the normal cranial vault: an embryologic review. *Journal of Korean Neurosurgical Society* **59**, 192–6.

Kalluri R. & Weinberg R.A. (2009) The basics of epithelial-mesenchymal transition. *Journal of Clinical Investigation* **119**, 1420–8.

Knierim U., Irrgang N. & Roth B.A. (2015) To be or not to be horned-consequences in cattle. *Livestock Science* **179**, 29–37.

Krivega I. & Dean A. (2017) CTCF fences make good neighbours. *Nature Cell Biology* **19**, 883–5.

Li M.N., Wu X.Y., Guo X., Bao P.J., Ding X.Z., Chu M., Liang C.N. & Yan P. (2018) Comparative iTRAQ proteomics revealed proteins associated with horn development in yak. *Proteome Science* **16**, 1–11.

Long C.R. & Gregory K.E. (1978) Inheritance of the horned, scurred, and polled condition in cattle. *Journal of Heredity* **69**, 395–400.

Luhken G., Krebs S., Rothammer S., Kupper J., Mioc B., Russ I. & Medugorac I. (2016) The 1.78-kb insertion in the 3'-untranslated region of RXFP2 does not segregate with horn status in sheep breeds with variable horn status. *Genetics Selection Evolution* **48**, 1–14.

Lundrigan B. (1996) Morphology of horns and fighting behaviour in the family *Bovidae*. *Journal of Mammology* **77**, 462–75.

Mariasegaram M., Reverter A., Barris W., Lehnert S.A., Dalrymple B. & Prayaga K. (2010) Transcription profiling provides insights into gene pathways involved in horn and scurs development in cattle. *BMC Genomics* **11**, 370.

McDiarmid J.J. (1981) "Corkscrew penis" and other breeding abnormalities in beef bulls. *New Zealand Veterinary Journal* **29**, 35–6.

Medugorac I., Seichter D., Graf A., Russ I., Blum H., Goepel K.H., Rothammer S., Foerster M. & Krebs S. (2012) Bovine polledness – an autosomal dominant trait with allelic heterogeneity. *PLOS ONE* **7**, e39477.

Medugorac I., Graf A., Grohs C. *et al.* (2017) Whole-genome analysis of introgressive hybridization and characterization of the bovine legacy of Mongolian yaks. *Nature Genetics* **49**, 470–5.

Mendonca F.S., Vaz R.Z., Leal W.S., Restle J., Pascoal L.L., Vaz M.B. & Farias G.D. (2016) Genetic group and horns presence in bruises and economic losses in cattle carcasses. *Semina-Ciencias Agrarias* **37**, 4265–73.

Monsoro-Burq A.H. (2015) PAX transcription factors in neural crest development. *Seminars in Cell & Developmental Biology* **44**, 87–96.

Nguyen Q.H., Tellam R.L., Naval-Sanchez M., Porto-Neto L.R., Barendse W., Reverter A., Hayes B., Kijas J. & Dalrymple B.P. (2018) Mammalian genomic regulatory regions predicted by utilizing human genomics, transcriptomics, and epigenetics data. *Gigascience* **7**, 1–17.

Pan Z.Y., Li S.D., Liu Q.Y. *et al.* (2018) Whole-genome sequences of 89 Chinese sheep suggest role of RXFP2 in the development of unique horn phenotype as response to semi-feralization. *Gigascience* **7**, 1–15.

Pares-Casanova P.M. & Caballero M. (2014) Possible tendency of polled cattle towards larger ears. *Revista Colombiana de Ciencias Pecuarias* **27**, 221–5.

Prayaga K.C. (2007) Genetic options to replace dehorning in beef cattle – a review. *Australian Journal of Agricultural Research* **58**, 1–8.

Rothammer S., Capitan A., Mullaart E., Seichter D., Russ I. & Medugorac I. (2014) The 80-kb DNA duplication on BTA1 is the only remaining candidate mutation for the polled phenotype of Friesian origin. *Genetics Selection Evolution* **46**, 1–5.

Sargent C., Bauer J., Khalil M., Filmore P., Bernas M., Witte M., Pearson M.P. & Erickson R.P. (2014) A five generation family with a novel mutation in FOXC2 and lymphedema worsening to hydrops in the youngest generation. *American Journal of Medical Genetics Part A* **164**, 2802–7.

Schmutz S.M., Marquess F.L.S., Berryere T.G. & Moker J.S. (1995) DNA marker-assisted selection of the polled condition in Charolais cattle. *Mammalian Genome* **6**, 710–3.

Sheebaa C.J., Andrade R.P. & Palmeirim I. (2016) Mechanisms of vertebrate embryo segmentation: common themes in trunk and limb development. *Seminars in Cell & Developmental Biology* **49**, 125–34.

Smith A.D.B (1927) The inheritance of horns in cattle some further data. *Journal of Genetics* **18**, 365–74.

Stafford K.J. & Mellor D.J. (2005) Dehorning and disbudding distress and its alleviation in calves. *Veterinary Journal* **169**, 337–49.

Stankowich T. & Caro T. (2009) Evolution of weaponry in female bovids. *Proceedings of the Royal Society B: Biological Sciences* **276**, 4329–34.

Szalaj P. & Plewczynski D. (2018) Three-dimensional organization and dynamics of the genome. *Cell Biology and Toxicology* **34**, 381–404.

Tetens J., Wiedemar N., Menoud A., Thaller G. & Drögemüller C. (2015) Association mapping of the scurs locus in polled Simmental cattle – evidence for genetic heterogeneity. *Animal Genetics* **46**, 224–5.

The UniProt Consortium (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research* **47**, D506–15.

Utsunomiya Y.T., Torrecilha R.B.P., Milanesi M., Paulan S.D.C., Utsunomiya A.T.H. & Garcia J.F. (2019) Hornless Nellore cattle (*Bos indicus*) carrying a novel 110 kbp duplication variant of the polled locus. *Animal Genetics* **50**, 187–188.

Wang M., Hancock T.P., Chamberlain A.J., Jagt C.J.V., Pryce J.E., Cocks B.G., Goddard M.E. & Hayes B.J. (2018) Putative bovine topological association domains and CTCF binding motifs can reduce the search space for causative regulatory variants of complex traits. *BMC Genomics* **19**, 395.

Wang Y., Zhang C.Z., Wang N.N. *et al.* (2019) Genetic basis of ruminant headgear and rapid antler regeneration. *Science* **364**, 1153–60.

White W.T. & Ibsen H.L. (1936) Horn inheritance in Galloway-Holstein cattle crosses. *Journal of Genetics* **32**, 33–49.

Wiedemar N. & Drogemuller C. (2015) A 1.8-kb insertion in the 3-UTR of RXFP2 is associated with polledness in sheep. *Animal Genetics* **46**, 457–61.

Wiedemar N., Tetens J., Jagannathan V., Menoud A., Neuenschwander S., Bruggmann R., Thaller G. & Drogemuller C. (2014) Independent polled mutations leading to complex gene expression differences in cattle. *PLOS ONE* **9**, e93435.

Wiener DJ, Wiedemar N, Welle MM, Droegemueller C (2015) Novel features of the prenatal horn bud development in cattle (*Bos taurus*). *PLOS ONE* **10**, e0127691.

Wu T.F., Chen G.Q., Tian F. & Liu H.X. (2017) Contribution of cranial neural crest cells to mouse skull development. *International Journal of Developmental Biology* **61**, 495–503.

Youngers M.E., Thomson D.U., Schwandt E.F., Simroth J.C., Bartle S.J., Siemens M.G. & Reinhardt C.D. (2017) Prevalence of horns and bruising in feedlot cattle at slaughter. *Professional Animal Scientist* **33**, 135–9.

Yu M. & Ren B. (2017) The three-dimensional organization of mammalian genomes. *Annual Review of Cell and Developmental Biology* **33**, 265–89.

Zhang L.L., He J., Han B., Lu L.N., Fan J.Y., Zhang H., Ge S.F., Zhou Y.X., Jia R.B. & Fan XQ (2016) Novel FOXC2 mutation in hereditary distichiasis impairs DNA-binding activity and transcriptional activation. *International Journal of Biological Sciences* **12**, 1114–20.

Zhu B., Zhang M. & Zhao J (2016) Microstructure and mechanical properties of sheep horn. *Microscopy Research and Technique* **79**, 664–74.

21

Chapter 2:

Bioinformatic analysis of the bovine POLLED genomic region

2.1 Introduction

Hornless, or polled, individuals have emerged in domesticated Bovidae (e.g., cattle, sheep, goat, yak), and this phenotype was subsequently selected. In cattle, there are four known dominant genetic variants for polledness, all located within 300 kb on chromosome 1 (BTA1). The Celtic POLLED variant was the first to be described and is the most widespread across breeds (Medugorac *et al.* 2012; Grobler *et al.* 2018; Koufariotis *et al.* 2018; Falomir-Lockhart *et al.* 2019). The Celtic variant consists of a 212 bp duplication/insertion into a 10 bp deletion (Medugorac *et al.* 2012). The Friesian POLLED variant is an ~80 kb tandem duplication (Medugorac *et al.* 2012) and the Mongolian POLLED variant includes a complex 219 bp duplication insertion with an additional 6 bp deletion/7 bp insertion located 621 bp upstream (Medugorac *et al.* 2017). In Nellore cattle, a ~110 kb duplication (Guarani variant) is associated with the polledness but the exact location of the duplication on BTA1 and nature of this variant is yet to be described (Utsunomiya *et al.* 2019). The clustering of the polled variants highlights the importance of this genomic region for horn development in cattle. Except for the Guarani variant, none of the mutations overlap known protein-coding genes, non-coding RNAs or regulatory elements (Allais-Bonnet *et al.* 2013). The Guarani duplication appears to include LincRNA#2. Therefore, despite the importance of polledness as a phenotype, the molecular mechanisms regulating horn development have eluded discovery.

2.1.1 Genomic features

Advances in genomic technologies and bioinformatics mean that researchers have access to better tools and data to investigate functional effects of causative variants. The functional effects of the Celtic and Friesian variants were considered when they were first discovered but the variants did not appear to disrupt any known coding sequences, splice sites, intronic regions, or known regulatory regions (Medugorac *et al.* 2012). However, as better annotation of this region becomes available, the new information may be shed light on the

mechanisms of horn development (e.g. unannotated protein coding genes, long non-coding RNA genes, microRNA, topologically associating domains [TADs] and enhancers). All of these need to be reconsidered in the context of the polled phenotype.

It is possible that the POLLED variants overlap unannotated protein-coding genes. A recent gene expression study using high depth RNAseq detected approximately 48,000 novel bovine transcripts when aligned to UMD3.1.1 (Foissac *et al.* 2019). These novel transcripts did not align to known reference transcripts, did not extend reference transcripts, or share an intron with reference transcripts. This suggests that there may be a substantial number of protein-coding genes that have not been yet identified.

Other methods of POLLED variant function might be through gene expression regulation. Horns could be a threshold trait for which a specific level of gene expression must be reached in order for horns to develop (Serpico 2020). So while the POLLED variants are not within any known genes, the variants could alter expression of a known or unknown gene by disturbing the function of regulatory elements. The affected gene(s) could be nearby (*cis* interacting) or further away (*trans* interacting) on the same chromosome or even on a different chromosome. If horns are a threshold trait, then the reduced expression of a specific gene would prevent horns from developing. Alternatively, the expression of a regulatory gene, such as a transcription factor, could be induced by the variants which leads to the altered expression of horn related genes. In this scenario, the variants could increase expression of a gene by disturbing the function of regulatory elements. The expression of the horn gene(s) would then be insufficient to overcome the effect of the inhibitor.

In addition to protein coding genes, non-coding genes such as long non-coding RNA (lncRNA) are important features of the genome. As the name suggests, lncRNA are transcribed, but generally are not translated (Murillo-Maldonado & Riesgo-Escovar 2019). Due to their single stranded structure, lncRNA can interact with DNA, RNA, and proteins, and can regulate expression levels during transcription, post-transcription, translation, or post-translation

(Cuevas-Diaz Duran *et al.* 2019). LncRNA are categorized by their position relative to protein-coding genes. *Intergenic* lncRNA (lincRNA) are located between genes, while *intragenic* overlap protein-coding genes (Cuevas-Diaz Duran *et al.* 2019).

The structure of lncRNAs is different from protein-coding genes. One study compared lncRNA and mRNA in different domestic animal species (cattle, chicken, goat and pig) (Foissac *et al.* 2019). The authors found that on average lncRNA are smaller (1800 bp vs 3600 bp) and have fewer exons (1.5 vs. 10) than mRNA, although their median exon length is longer (660 vs. 130 bp) (Foissac *et al.* 2019). LncRNA are expressed at low levels and are tissue-specific, making them difficult to detect (Cuevas-Diaz Duran *et al.* 2019; Foissac *et al.* 2019).

Two lncRNA (referred to as LincRNA#1 and LincRNA#2 by Allais-Bonnet *et al.* (2013)) near the POLLED variants could play a role in horn development as there are differences in the expression of these lincRNA between horned and polled fetuses early in fetal development (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014).

MicroRNA (miRNA) are short non-coding RNA molecules, 23-25 nucleotides in length, that interfere with gene expression by interacting with mRNA (Huntzinger & Izaurralde 2011; Shruti *et al.* 2011; O'Brien *et al.* 2018; Remsburg *et al.* 2019). This interference is achieved by partial or full sequence-specific binding to the 3' untranslated region of the mRNA, causing repression of translation or degradation of the transcript (Huntzinger & Izaurralde 2011; O'Brien *et al.* 2018). MicroRNAs are involved in various developmental processes including cellular proliferation, cell fate determination and apoptotic pathways (Shruti *et al.* 2011; Divisato *et al.* 2020). The role of miRNA in cell fate determination is demonstrated by tissue specific expression of miRNA during embryonic development. For example, there is sequential expression of specific miRNA during neurogenesis (Fiore *et al.* 2008). Given their role in embryo development, miRNA could participate in horn ontogenesis.

Enhancers are key regulatory elements that modulate when and where genes are expressed. Enhancers drive cell-type specific gene expression by interacting with the promoter

region of a gene via transcription factors (Lewis *et al.* 2019). They increase transcription irrespective of their orientation and can even interact with distant promoters (Andersson *et al.* 2014). Enhancers are typically associated with several genes, and a gene will have several enhancers. Clusters of enhancers occurring in close proximity to each other, referred to as super enhancers, can be activated in synergy to drive target gene transcription (Lewis *et al.* 2019). Enhancer regions for cattle are yet to be identified. Putative bovine regulatory elements have been mapped using human data, with some validated using bovine data (Nguyen *et al.* 2018). A HAND1 transcription factor binding site has been predicted to overlap the 10 bp deletion site of the Celtic variant (Wang *et al.* 2017). If this binding site is functional, then regulation of ontogenesis could be affected by this variant.

TADs are one of the hierarchical levels of chromatin structure. TADs are chromatin regions where there is a high frequency of interactions between the loci within the region, and fewer interactions across boundaries (i.e., between adjacent TADs). Binding sites for CTCFs, a DNA-bound transcription repressor molecule, are enriched at the TAD boundaries (Sanborn *et al.* 2015; Fudenberg *et al.* 2017; Krietenstein *et al.* 2020).

TAD boundaries act as genetic insulators by limiting inter-domain interactions (Gong *et al.* 2018; Foissac *et al.* 2019). Disruption of a boundary can alter the interactions between domains, and allow an enhancer-promoter interaction that would not otherwise occur (Lupiáñez *et al.* 2015). As enhancers control gene expression, it is plausible that the disruption of a TAD boundary may affect gene expression during development. For example, an enhancer located near a boundary may act upon the promoter of a gene in a different TAD when the boundary is removed (Yu & Ren 2017; Furlong & Levine 2018). Knowledge of the TAD structure in the genomic region surrounding the POLLED variants may help to identify genes potentially affected by the variants.

2.1.2 Bioinformatic analyses

2.1.2.1 Open reading frames

An open reading frame (ORF) is a potentially translatable sequence that begins with a start codon and ends with a stop codon (Andrews & Rothnagel 2014). *In silico* detection of ORFs identifies DNA sequences that may be translated into protein. Splicing in eukaryotes complicates ORF searches because the DNA sequence rarely reflects the mRNA sequence (Brown 2002). However, analyzing protein and DNA sequence conservation can help to determine if an ORF has protein-coding potential.

2.1.2.2 Expressed sequence tags

The POLLED variants may affect unannotated genes or long non-coding RNA (lncRNA) as the current annotation of the bovine genome is incomplete. This incompleteness of the annotation of the bovine genome (UMD3.1) in the Ensembl v90 release reference genome was demonstrated by the discovery of novel transcripts in a high-depth RNAseq study of bovine tissues (Foissac *et al.* 2019).

The presence of an expressed sequence tag (EST) may indicate unannotated genes. ESTs are sequence reads 200-800 bp long generated from cDNA libraries of poly-adenylated transcripts (Parkinson & Blaxter 2009). ESTs have traditionally been used to identify protein-coding genes but can be also used to identify poly-adenylated lncRNA. For example, a lncRNA associated with cancer metastasis in humans was identified in an RNAseq study and shown to overlap a known EST (Park *et al.* 2013). Therefore, ESTs that map near the POLLED variants may indicate unannotated genes or lncRNA.

2.1.2.3 Enhancer atlases and databases

The identification and annotation of functional elements is limited for the bovine genome, but significant progress has been made towards the functional annotation of the human genome, so the human is a good starting model for other mammalian species. Additionally, the

FANTOM project (Functional ANnoTation Of the Mammalian genome) is a worldwide collaborative effort to identify all functional elements in mammalian genomes (Andersson *et al.* 2014; FANTOM Consortium *et al.* 2014).

The FANTOM project has used Cap Analysis of Gene Expression (CAGE) to map sets of transcripts, transcription factors, promoters and enhancers that are active in the majority of mammalian primary cell types and a series of cancer cell lines and tissues (Andersson *et al.* 2014; FANTOM Consortium *et al.* 2014). Human enhancers and their activity within the analogous genomic region can be used to indicate their functional importance in other species.

Some studies have predicted bovine enhancers *in silico* using data from other species in an attempt to annotate the bovine genome (Wang et al. 2017; Nguyen et al. 2018). In the study by Nguyen et al. (2018), biochemical assays in human samples were used to predict regulatory elements in cattle, and then bovine specific data were used to identify high-confidence regulatory regions. One of these regulatory regions overlapped the 10 bp deletion site of the Celtic variant, suggesting that the variant interferes with this regulatory site (Nguyen et al. 2018). Wang *et al.* (2017) also detected potential bovine enhancers by aligning human enhancer sequences to the bovine genome and validated these enhancers using bovine ChIP-Seq data. To the best of the author's knowledge, the validated enhancers are not publicly available, and thus, cannot be used to assess if any are associated with the POLLED region.

2.1.2.4 Conservation of sequence

The sequence of many genomic features are constrained by their function, and therefore, the conservation of DNA sequence can indicate that a sequence is functionally important. By comparison of DNA sequences across species, common or distinct genomic regions can be discovered (Sanges *et al.* 2013). Common DNA sequences are usually derived from a common ancestor (Kawashima 2019). Exceptions to this are horizontal gene transfer and species convergence (Kawashima 2019). Since some features are not yet discovered or annotated in a

reference genome, analysing conservation can be used to predict the importance of a given sequence.

The genome structure is an important consideration when analysing conservation. The genome can be categorized into protein-coding genes, non-coding genes (such as long non-coding RNA), intergenic regions (including regulatory elements), the centromere and telomere (Kawashima 2019). The presence of a feature can explain why a given sequence is conserved. For example, some features, such as non-coding elements, rely on their conserved sequence to function correctly (Polychronopoulos *et al.* 2017; Wong *et al.* 2020), and thus, they tend not to vary between species. The evolutionary distance between species should be considered carefully to ensure that results are informative (Kawashima 2019). Closely related species may have few differences so that even intergenic regions are highly conserved, while more distantly related species may have very few conserved intergenic sequences. Sequences within the POLLED region that are conserved between horned Bovidae, but not conserved in non-horned species, may be functionally important and may provide clues to the molecular control of horn development.

2.1.2.5 Gene synteny

Gene synteny refers to the order in which two or more genes occur. Some homologous blocks of sequence, therefore gene synteny, has been conserved for 10s of millions of years (Farré *et al.* 2019). Gene synteny can be maintained because 1) a region is not structurally predisposed to chromosomal breaks and rearrangements, and/or 2) the genes and regulatory regions are interconnected. A study of the human and mouse genomes showed that regions have different susceptibility to chromosomal rearrangements (Pevzner & Tesler 2003; Peng *et al.* 2006). Major genomic rearrangements tend to occur at "fragile" chromosomal regions while conserved regions are "solid". Additionally, gene synteny can be affected in regions where genes share similar expression patterns and the genes are functionally related (Kikuta *et al.* 2007; Dávila López *et al.* 2010). An example of such case are the developmental HOX gene

clusters where the precise timing of transcription is linked to the gene configurations (Lee *et al.* 2006; Kikuta *et al.* 2007; Darbellay *et al.* 2019).

Currently, it is not known whether gene synteny is conserved around the POLLED variants. Gene synteny may be different between horned and hornless species, and differences could indicate genes important for horn development.

2.1.2.6 Conservation of lincRNAs

The sequence of lincRNAs is poorly conserved across species (Cuevas-Diaz Duran *et al.* 2019). For example, only 5.7%, 5.2% and 1.6% of sheep lincRNA aligned to goat, cattle and human transcripts respectively with high confidence (that is, > 50% identity and > 50% alignment length) (Bush *et al.* 2018). Therefore, conservation of lincRNA is often assessed by conservation of gene-lincRNA-gene synteny. By this definition, a lincRNA is conserved if 1) it is located between two orthologous protein-coding genes, 2) it is the only lincRNA located between the two genes, and 3) the orientation of the triplet is identical (Foissac *et al.* 2019).

Two lincRNA, LincRNA#1 and LincRNA#2, are found near the POLLED variants and their expression may be affected by the Celtic variant in horn bud tissue (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014). If these lincRNA are conserved in horned species but not in non-horned species, their function may be related to horn development.

2.1.2.7 Topologically associating domains

Hi-C sequencing is a technique that can reveal chromatin conformation, thereby facilitating the study of chromatin interactions (Dixon *et al.* 2012). Hi-C sequencing identifies where chromatin conformation in the nucleus brings DNA strands into close proximity and assesses all loci vs. all loci. This is the technique that led to the discovery of TADs (Dixon *et al.* 2012; Dixon *et al.* 2016).

To the best of the author's knowledge, TADs have not been analysed for the cattle genome using Hi-C data, and there is no public repository of known TADs for cattle. One study

has predicted the bovine TAD structure in cattle using human, mouse, dog and macaque data (Wang *et al.* 2018). The Hi-C data produced by Low *et al.* (2020) provided the opportunity to analyse bovine TADs and specifically those in the POLLED region, which may aid our understanding of the chromatin interactions in this region.

2.1.3 Aim

In order to better understand how the POLLED variants may be affecting the local chromosomal dynamics, the POLLED variant region was analysed to determine the level of conservation of the genomic region surrounding the POLLED variants and to characterise the TAD structures in this region. Specifically, sequence conservation and TAD structure between horned and non-horned species was compared in the defined genomic segment corresponding to the POLLED region. The coding potential of the variants was also investigated by searching for ORF and ESTs that overlap with the POLLED variants.

2.2 Materials and Methods

2.2.1 Genomic region analysed

The genomic region analysed was the bovine TAD predicted to include all of the POLLED variants (Wang et al. 2018). The boundaries of this single 975 kb TAD (chr1:1,226,028 – 2,201,452 bp; bovine assembly UMD3.1) were predicted using data from human, mice, dogs and macaques (Wang *et al.* 2018). All analyses focused on this region, which is referred to as the "putative POLLED TAD" herein.

All analyses were conducted using the most recent bovine genome (ARS-UCD1.2), unless otherwise stated. One kb sequences from the putative TAD boundaries were used to map the TAD from UMD3.1 coordinates to ARS-UCD1.2 (Table 2.1). For the POLLED variants that were originally described using UMD3.1, the sequences were aligned to the current genome using NCBI BLAST (Agarwala *et al.* 2018) to determine their position (Table 2.1). Sequence

from the beginning (chr1: 1,909,352 – 1,910,351) and end (chr1:1,989,480 – 1,990,479) of the

Friesian variant were used to identify the duplicated region in ARS-UCD2.1.

Table 2.1: Genome coordinates used for analyses.

Region	Gene Upstream	Gene Downstream	Description	ARS-UCD1.2. Coordinates used herein
Putative POLLED TAD	-	-	TAD region predicted using data from humans and other species (Wang et al. 2018)	Chr1:1,946,384 – 2,921,213
Celtic (P_C)	LOC526226	OLIG1	202 dup/ins: 212 bp duplication + 10 bp deletion	NC_037328.1:g.[2429327_2429336del;2429109_2429320dupins]
Mongolian (P_M)	OLIG2	C1H21orf62	1 del/ins: 7 bp deletion/6 bp insertion 219 dup/ins: a complex 219 insertion (158 bp duplicated + 61 bp unique sequence)	NC_037328.1:g.[2695261_2695267delinsTCTGAA;2695889_2696047dupins]
Friesian (P_F)	OLIG2	C1H21orf62	~80kb duplication	NC_037328.1:g.2629113_2709242dup
Nellore (P_G)	OLIG2	C1H21orf62	~110kb duplication	NC_037328.1:g.2614828_2724315dup

2.2.2 Open Reading Frame (ORF) search

NCBI ORF Finder (Wheeler *et al.* 2003) was used to search for open reading frames that overlap the Celtic and Mongolian variants. The sequences 500 bp up- and downstream of the Celtic and Mongolian variants were analysed and compared to the wild-type sequence. The shortest bovine reference gene is ~316 bp (Foissac *et al.* 2019; Supplementary Figure 12). Therefore, a 300 nucleotide cut off was used in the ORF search.

2.2.3 Expressed Sequence Tag (EST) search

NCBI BLAST (Agarwala *et al.* 2018) was used to align entries in the bovine EST database (National Centre for Biotechnology Information 2020; GP/9913.10708/ESTs; accessed: Feb 2020) to the predicted POLLED TAD region. The locations of the ESTs for genes and lincRNAs were identified using BEDTools/2.25.0-foss-2015b. ESTs that were located within an annotated gene or lincRNA were filtered from the dataset. Overlapping ESTs were grouped and the cluster coordinates were recorded. The cluster coordinates were used to obtain sequences from NCBI and aligned to ARS-UCD1.2 using NCBI discontiguous megablast (Agarwala *et al.* 2018) to ensure the sequence did not match genes elsewhere in the genome.

2.2.4 MicroRNA search

MicroRNA databases, miRBase (Kozomara *et al.* 2019; date accessed: Oct 2020) and RumimiR (Bourdon *et al.* 2019; date accessed: Mar 2021) and datasets employed by various genome browsers (NCBI, Ensembl and UCSC) were searched for *Bos taurus* miRNA in the putative POLLED region. The miRBase database uses coordinates from the assembly Btau_5.0.1 (GCA_000003205.6) so the putative POLLED region in this assembly was mapped by aligning 1 kb sequences from the TAD boundary and determined to be located at chr1:1,222,109- 2,196,408 (Btau_5.0.1).

2.2.5 Enhancer search

Putative bovine enhancers were sourced from Nguyen *et al.* (2018; downloaded Feb 2021) and aligned to the putative polled TAD position in ARS-UCD1.2 using NCBI BLAST (Agarwala *et al.* 2018).

The FANTOM 5 RIKEN human enhancer database (https://slidebase.binf.ku.dk/human_enhancers/; Accessed Oct 2020) was searched using SlideBase for enhancers within human chromosome 21 which contains the region homologous to the putative POLLED TAD (Andersson *et al.* 2014; FANTOM Consortium *et al.* 2014; Ienasescu *et al.* 2016). As the coordinates used herein did not correspond to the expected genomic region, the database was searched to find the coordinates that correspond with the region between *SYNJ1* and *DONSON* (inclusive of these genes; chr21:34,001,068–35,284,703), a reasonable proxy for the putative POLLED TAD. This is reasonable because downstream boundary of the putative TAD intersects *DONSON* and the upstream boundary is 2,475 bp from the 3` of *SYNJ1*.

Enhancers were identified and sequences of 400 bp centered on each enhancer were aligned to the bovine reference genome. The cell types and tissues in which the enhancers are active in humans was recorded.

2.2.6 Conservation analysis

The analysis described in this section was conducted by Dr Rick Tearle (Davies Research Centre, University of Adelaide). Genomes from 24 species were sourced from NCBI (July 2020) to analyse conservation of the putative POLLED TAD (Appendix Table A1). Ten of the species had horns (*Bos taurus*, *Bos indicus*, *Bos mutus*, *Bos grunniens*, *Bison bison, Bubalus bubalis*, *Ovis aries*, *Capra aegagrus*, *Capra hircus*, *Pantholops hodgsonii*), one species (*Cervus elaphus hippelaphus*) had antlers and the remaining thirteen species had no cranial appendages at the frontal bone (*Moschus moschiferus*, *Camelus bactrianus*, *Camelus dromedaries*, *Camelus ferus*, *Sus scrofa*, *Equus caballus*, *Canis lupus familiaris, Homo sapiens, Loxodonta Africana*,

Monodelphis domestica, *Ornithorhynchus anatinus*, *Choloepus hoffmanni*, *Dasypus novemcinctus*). Each species genome was aligned with the POLLED TAD against a repeat masked *Bos taurus* genome using LASTZ (Harris 2007). Variant sites were assessed by comparing conservation 1 kb and 10 kb up and downstream of the variant. An identity above 80% was considered conserved (Sanges *et al.* 2013).

2.2.7 Conservation of gene synteny

Annotation within the putative POLLED TAD (Wang *et al.* 2018) for genomes of horned species (cattle, water buffalo, goat, and sheep) and hornless species (horse, pig, dog, and human) were sourced from Ensembl and NCBI in .gff3 format (Nov 2019) (Appendix Table A2). The files were filtered for protein-coding genes using the filter function in the R package dplyr (version 0.8.4). To define the analogous region in each genome, 1-kb of bovine sequence from the TAD boundary in the ARS-UCD1.2 was aligned to the each of the genomes using NCBI BLAST. However, matches were only found for dog and human. As outlined in section 2.2.5, *DONSON* and *SYNJ1* were used as a proxy for the putative POLLED TAD for all species analysed. If a gene was not annotated in ENSEMBL, the NCBI genome was checked for an annotation and vice versa.

2.2.8 Conservation of lincRNA

As lincRNA sequences vary among species, the position of known lincRNAs in relation to protein-coding genes was used to determine if they were conserved (Bush *et al.* 2018; Foissac *et al.* 2019). Annotation files for cattle, water buffalo, goat, sheep, horse, pig, dog, and human were filtered for lincRNAs located within the POLLED TAD. Sheep and goat lincRNA from the catalog generated by Bush *et al.* (2018) were used to supplement the data. The number of lincRNAs annotated in this region ranged from one (sheep) to 13 (dog). The surrounding protein-coding genes for each lincRNA were determined to identify gene-lincRNA-gene syntenic triplets. A lincRNA was considered conserved if it was located between two orthologous protein-coding genes and transcribed in the same orientation. In some cases, there

were more than one lincRNA annotated between two genes that were transcribed in the same direction. These were not removed from the dataset at this stage as all lincRNAs identified were subsequently analyzed for sequence conservation.

Sequence conservation of bovine lncRNA were assessed using NCBI BLAST (Agarwala *et al.* 2018) to align them to water buffalo, sheep, goat, pig, horse, dog, and human. Query coverage and identity of matches were recorded, and sequences were considered highly conserved if both query coverage and identity were ≥ 50% (Bush *et al.* 2018).

2.2.9 Hi-C data analysis and TAD boundary identification

The analysis of Hi-C reads described in this section was conducted by Dr Ning Liu (Bioinformatics Hub, University of Adelaide). The Hi-C reads were generated by Low *et al.* (2020) and detailed methods are described in the publication. The raw Hi-C reads were produced from lung tissue of a F1 hybrid male fetus of Angus sire x Brahman dam at 90 days of development. The data were downloaded from NCBI SRR6691720 of PRJNA432857.

The Hi-C reads were separated into Brahman and Angus-specific sequence reads using k-mers specific to each breed as describe in Low *et al.* (2020). Sequencing adapters were trimmed from the raw sequence reads using AdapterRemoval (version 2.2.1a; Schubert *et al.* 2016). The trimmed Angus- and Brahman- specific reads were then aligned to the genomes of Angus (UOA_Angus_1), Brahman (UOA_Brahman_1), (Low *et al.* 2020), and Hereford (ARS-UCD1.2) (Rosen *et al.* 2020) using Bowtie2 with the HiC-Pro pipeline (version 2.9.0) (Servant *et al.* 2015) and with Bowtie2 indexed Brahman and Angus genomes (Low *et al.* 2020) and the bovine reference genome (ARS-UCD.1.2).

The reads were normalised using the iterative correction and eigenvector decomposition method, and normalised matrices of 40 kb bins from the HiC-Pro pipeline were used to identify Topologically-Associated Domains using Armatus (version 2.3) with the gamma-max set at 0.5 (Filippova *et al.* 2014).

2.2.9.1 CTCF binding motif prediction

CTCF binding motifs of the Brahman, Angus and ARS-UCD.1.2 genomes were predicted using Find Individual Motif Occurrences (FIMO) (version 4.12.0) from the MEME suite (Bailey *et al.* 2009) based on the CTCF position weight matrix from the JASPAR database (motif id: MA0139.1) (Fornes *et al.* 2019).

2.2.10 Identification of Celtic POLLED *variant interactions*

The analysis described in this section was conducted by Dr Ning Liu. Significant chromatin interactions within the Celtic region were identified using Angus and Brahman-specific reads aligned to ARS-UCD1.2. A 20-kb window (chr1:2420000-2440000) containing the Celtic variant was used as an anchor to identify potential chromatin interactions with this region. Significant chromatin interactions were identified from normalised interaction matrices of 10 kb resolution from the HiC-Pro pipeline using FitHiC2 (Kaul *et al.* 2020). The Benjamini-Hochberg procedure was used to compute the false discovery rate (FDR) based on the P-value from FitHiC2. Interactions with FDR larger than 0.05 or contact count < 2 were significant.

2.3 Results & Discussion

2.3.1 Open-reading frame analysis of the Celtic and Mongolian variants

When the Celtic variant was first discovered, it was reported that it did not overlap any known protein-coding gene or lncRNA (Allais-Bonnet *et al.* 2013). However, as genes that are lowly expressed or have a restricted tissue distribution may have been missed, the sequence ~500 bp up- downstream the Celtic and Mongolian variants was examined for ORFs. Only the Celtic and Mongolian variants were analysed as they affect small defined regions, whereas the Friesian and Guarani variants were not analysed due to their greater size.

The Mongolian variant comprises two rearrangements, a 6 bp insertion/7 bp deletion (in/del) and a complex 219 bp insertion located 621 bp downstream of this in/del (219 bp dup/ins: 158

bp duplicated + 61 bp unique sequence) (Medugorac *et al.* 2017). An ORF spanned the 219 bp dup/ins. The ORF size increases from 339 bp in the corresponding horned sequence to 558 bp in the Mongolian POLLED variant (

Figure 2.1). The insertion does not disrupt the start codon or stop codon nor alter the reading frame, as the insertion is divisible by three. There were no ORFs identified that were more than 300 nucleotides in length that overlapped the Celtic variant or the in/del of the Mongolian variant.

A
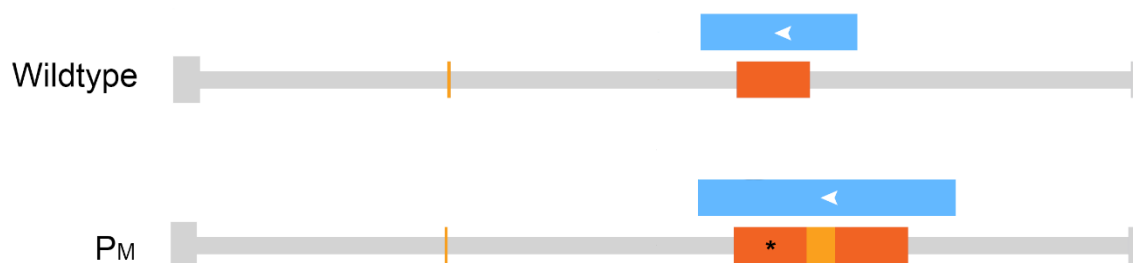
B

```
Wt          1    MNTGRILVVISRSAGLHGPRPLVPVSPRESHLLSTLIPGMPWNFASLISV
           51    PGVLLLCILTSFISNSCGINTCRQRVVSFCELPKVSCLPKVSCFLTLYRK
          101    PLKMLSFLFSYS

Mongolian   1    MNTGRILVVISRSAGLHGPRPLVPVSPRESHLLSTLIPGMPWNFASLISV
           51    PGVLLLCILTSFISNSCGINTCRQRVVSFCELPKVSCIVFCFLIHKVKSY
          101    ICCNWKFALIPGMPWNFASFISVPGVLLLCILTSFISNSCGINTCRQRVV
          151    SFCELPKVSCLPKVSCFLTLYRKPLKMLSFLFSYS
```

Figure 2.1: A) Open reading frame (ORF) spanning the Mongolian 219 bp insertion. P_M = Mongolian POLLED variant; ▇ = duplicated sequence; ▇ = insertion; ▇ = ORF; duplication indicated by asterisk (*). B) Amino acid sequence of the ORF for the wild-type allele and the Mongolian POLLED allele.

The amino acid sequence of the ORFs that overlapped the 219 bp dup/ins of the Mongolian rearrangement for both the wild-type and Mongolian allele was aligned using BLASTP to the Swiss-Prot database. No proteins were found to match the amino acid sequence.

NCBI BLAST (discontiguous megablast) showed that the wild-type ORF sequence was highly conserved in horned species with 100% query coverage and 93-97% sequence identity (water buffalo, sheep and goat). For horse, dog and human, the query coverage ranged from 43-44% and the identity ranged from 67-71% with the corresponding conserved region in ARS-UCD2.1 between 2,695,969 – 2,696,202 bp. The sequence was not conserved in pig. Since the sequence was highly conserved in closely related horned species but only partially conserved in non-horned species, this sequence may be functionally important for polled and perhaps protein-coding.

In eukaryotes, ORFs do not necessarily accurately represent the translated sequence because of the splicing of mRNA from multi-exon genes (Brown 2002). This means that not all of the sequence transcribed is translated to protein. Without knowledge of splice sites, further investigation into the ORF is difficult. However, the ORF search was followed by sequence and protein conservation analysis to see if there were any matches in other species or known proteins. Conserved DNA might suggest that the sequence is functionally important. Any conserved amino acid sequence might suggest protein functions such as DNA binding but no matches were found. The analysis that was conducted to assess the potential functionality of the predicted ORF was uninformative. However, this analysis can be only used as an indication of protein-coding potential and experimental data are required to confirm the ORF encodes a protein.

2.3.2 ESTs

ESTs that map to the POLLED region may identify unannotated protein-coding genes or lncRNAs. Bovine ESTs from the NCBI *Bos taurus* database (National Centre for Biotechnology Information 2020; GP/9913.10708/ESTs; accessed: Feb 2020) were aligned to the putative POLLED TAD.

Of the 1,579,753 ESTs in the *Bos taurus* EST database, 4,620 ESTs aligned within the putative POLLED TAD. The ESTs that aligned with annotated genes or lincRNA were

removed, leaving 1337 ESTs. Overlapping ESTs were grouped, resulting in 27 clusters of ESTs that had between 1 and 1218 ESTs. EST clusters that aligned to other regions of the genome were filtered from the dataset, leaving two clusters that uniquely aligned to the putative POLLED TAD. These two 'clusters' contained one EST each, CB166156.1 and DN819280.1.

Neither of these ESTs overlapped the POLLED variants. CB166156.1 is 146 bp long and is located between *LincRNA#1* and *OLIG2* (chr1: 2545442-2545587 bp). DN819280.1 is 105 bp long and is located between *OLIG2* and LincRNA#2 (chr1: 2576023-2576127). The conservation of these sequences was assessed by BLAST alignment (discontiguous megablast) to horned species (water buffalo, sheep, goat) and hornless species (pig, horse, dog and human). Both ESTs were conserved in all the horned species with 100% query coverage and 96-99% identity. CB166156.1 was also relatively conserved in pig (97% query coverage and 74% identity), but had low conservation in horse (31% coverage and 81% identity) and no matches were found for dog and human. There were no matches for DN819280.1 in the hornless species.

CB166156.1 and DN819280.1 could potentially represent two unannotated genes or polyadenylated lincRNAs in cattle. The high level of conservation in horned species indicates that these ESTs may be functional. Given that these ESTs could be potential genes or lincRNAs, and their proximity to the POLLED variants, they may be important for horn development. Therefore, these ESTs were included as annotations for the subsequent transcriptomic analysis of the horn bud (Chapter 4).

2.3.3 MicroRNAs

MicroRNAs are involved in developmental processes, and therefore, may affect horn ontogenesis. MicroRNAs within the putative POLLED TAD were identified using the NCBI genome browser and miRNA databases, miRbase and RumimiR.

One miRNA in the POLLED TAD was identified via the NCBI genome browser: bta-mir-6501 (chr1:1,978,829-1,978,892). RumimiR reports the location of bta-mir-6501 at chr1:

1,978,829-1,978,851 bp with an RNA sequence 5'-CCAGGGCAGCCUGUGGUAACAGU-3'. This places Bta-miR-6501 at the beginning of the putative POLLED TAD and overlapping with the third exon of the *SON* gene. The miRbase and RumimiR databases were also searched for miRNAs, but no additional miRNAs were found in this region.

The function of the one miRNA found in the putative POLLED region, bta-miR-6501, has not been determined. In cattle, this miRNA has been detected in milk and mammary gland tissue during mid-lactation (Li *et al.* 2015). There are two miRNAs transcribed from Hsa-mir-6501, MiR-6501-5p and MiR-6501-3p, and Bta-miR-6501 corresponds to the latter. MiR-6501 expression has been associated with colon adenocarcinomas and renal cancers (Yamaguchi *et al.* 2017; Chen *et al.* 2019). Overall, little is known about mir-6501, but miRNA expression studies may elucidate any involvement of miRNAs, such as mir-6501, in horn development.

2.3.4 Enhancers

2.3.4.1 Mapping predicted bovine enhancers to the POLLED region

The POLLED variants may affect the function of enhancers that are important for horn development. A previous study predicting bovine enhancers using human data noted that the Celtic variant overlapped an enhancer (Nguyen *et al.* 2018). Since this analysis, the Mongolian and Guarani variants have been reported. The regulatory regions identified by Nguyen *et al.* (2018) were mapped to the current reference genome to determine whether these polled variants overlapped any enhancers.

The bovine regulatory regions predicted by Nguyen *et al.* (2018) identified 163 putative enhancers in the putative POLLED region (defined in section 2.2.1). The analysis by Nguyen, *et al.* (2018) was conducted using an earlier reference bovine genome assembly (UMD3.1), so the sequences were aligned to the latest assembly, ARS-UCD1.2 using BLASTn. Thirteen sequences did not align to the current reference genome.

One regulatory region, 88Tis5dat3Type_ID415 (chr1: 2429321-2429457) overlapped the 10 bp deletion of the Celtic variant (Figure 2.2). Nine regulatory regions were located within the Guarani duplicated region, and six out of these nine regions were in the duplicated sequence of the Friesian variant. The Mongolian variant did not overlap any putative regulatory regions.
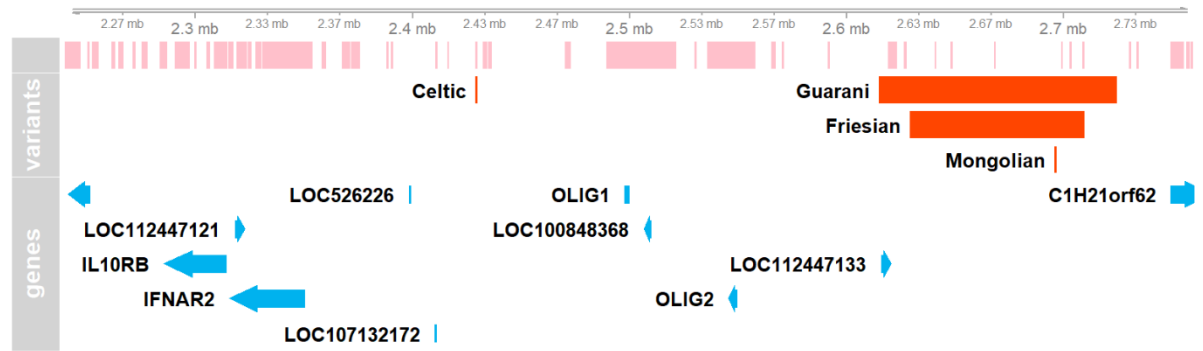
Figure 2.2: High confidence putative bovine enhancers predicted from human assays (Nguyen et al. 2018) between chr1:2,240,000-2,759,999 bp (ARS-UCD1.2). ▢ = Predicted enhancers; = polled variants; ▢ = genes.

Although putative regulatory DNA elements were found to overlap POLLED variants, transcription factor binding motif analysis is required to predict the function of these elements. This would give an understanding to which pathways are interrupted by the rearrangements causing polledness. For example, Nguyen *et al.* (2018) found that the regulatory region that overlaps the Celtic variant disrupts a *HAND1* (*heart and neural crest derivatives expressed 1*) transcription factor binding site. HAND1 is involved in craniofacial and limb development and morphology. Altering HAND1 expression leads to abnormalities to the craniofacial midline such as facial clefts (Barbosa *et al.* 2007; Bonilla-Claudio *et al.* 2012; Firulli *et al.* 2014) and affects the morphology of limbs (Laurie *et al.* 2016; Firulli *et al.* 2017; Funato *et al.* 2020). Overexpression of HAND1 was found to significantly reduce long bone length by decreasing

the expression of fibril-forming collagens via genes upstream of these collagens (Funato *et al.* 2020).

2.3.4.2 Human enhancers activity in comparative region of the putative POLLED TAD

Enhancers within the comparative human region of the putative POLLED TAD were found using SlideBase, the search engine for human enhancers from the FANTOM5 consortium (Andersson *et al.* 2014; Ienasescu *et al.* 2016). The activity of these enhancers in cells and organs was assessed to determine their function.

Based on the SlideBase analysis of the RIKEN Fantom 5 enhancer atlas, there are 34 human enhancers within the comparative putative polled TAD (chr21:34,001,068–35,284,703) (Andersson *et al.* 2014; FANTOM Consortium *et al.* 2014). Sequences at these enhancer sites were aligned to the bovine genome. Sixteen out of 34 enhancer sequences (47%) had a match in the bovine genome. Only eight of the 34 enhancer sequences (24%) aligned within the bovine putative POLLED TAD (Appendix Table A3). None of these enhancers overlapped the bovine variants (Appendix Figure A1) or the putative enhancers defined in section 2.3.4.1.

The activity of the 34 enhancers identified using SlideBase within the POLLED TAD was examined based on tissue (Figure 2.3) and cell types (Figure 2.4). SlideBase reports when the activity of an enhancer is significantly over-represented, i.e. the enhancer is active, in an above average number of tissues/cells within a category (e.g. brain). The enhancer activity within the POLLED TAD was significantly over-represented in the tissues from the brain, blood and testis. In cells, enhancer activity was most over-represented in monocytes, dendroctyes, basophils and neutrophils. This suggests that the enhancers in the putative POLLED TAD are involved in the function of the immune and nervous systems.

Figure 2.3: The proportion of the 34 enhancers from the comparative POLLED TAD (chr21:34,001,068–35,284,703) that have activity in a given tissue. Aqua indicates the proportion of enhancers that are active but that are not significantly over-represented. Red indicates enhancers that were active and significantly over-represented according to SlideBase. Only tissues in which enhancers were active are shown.

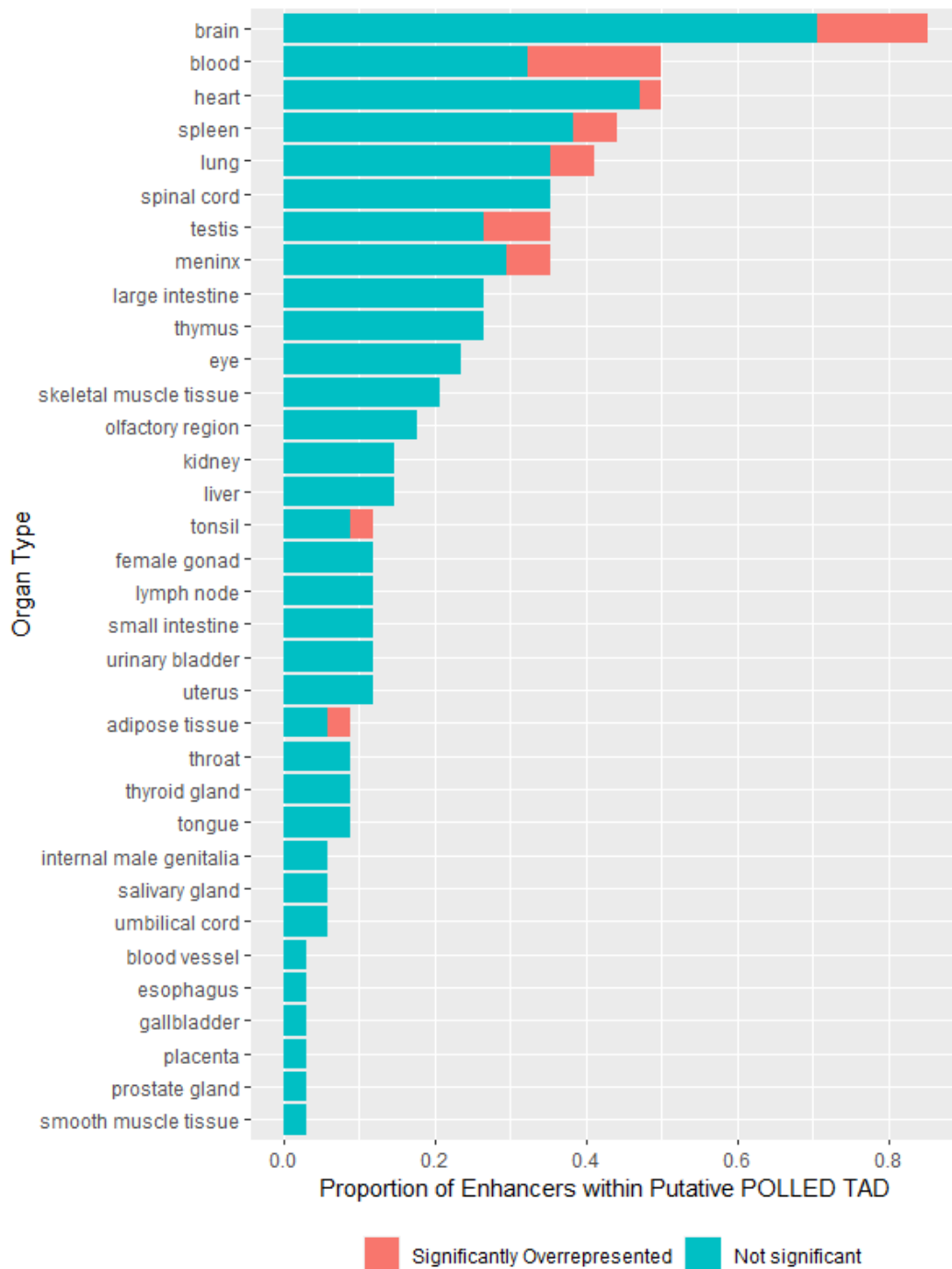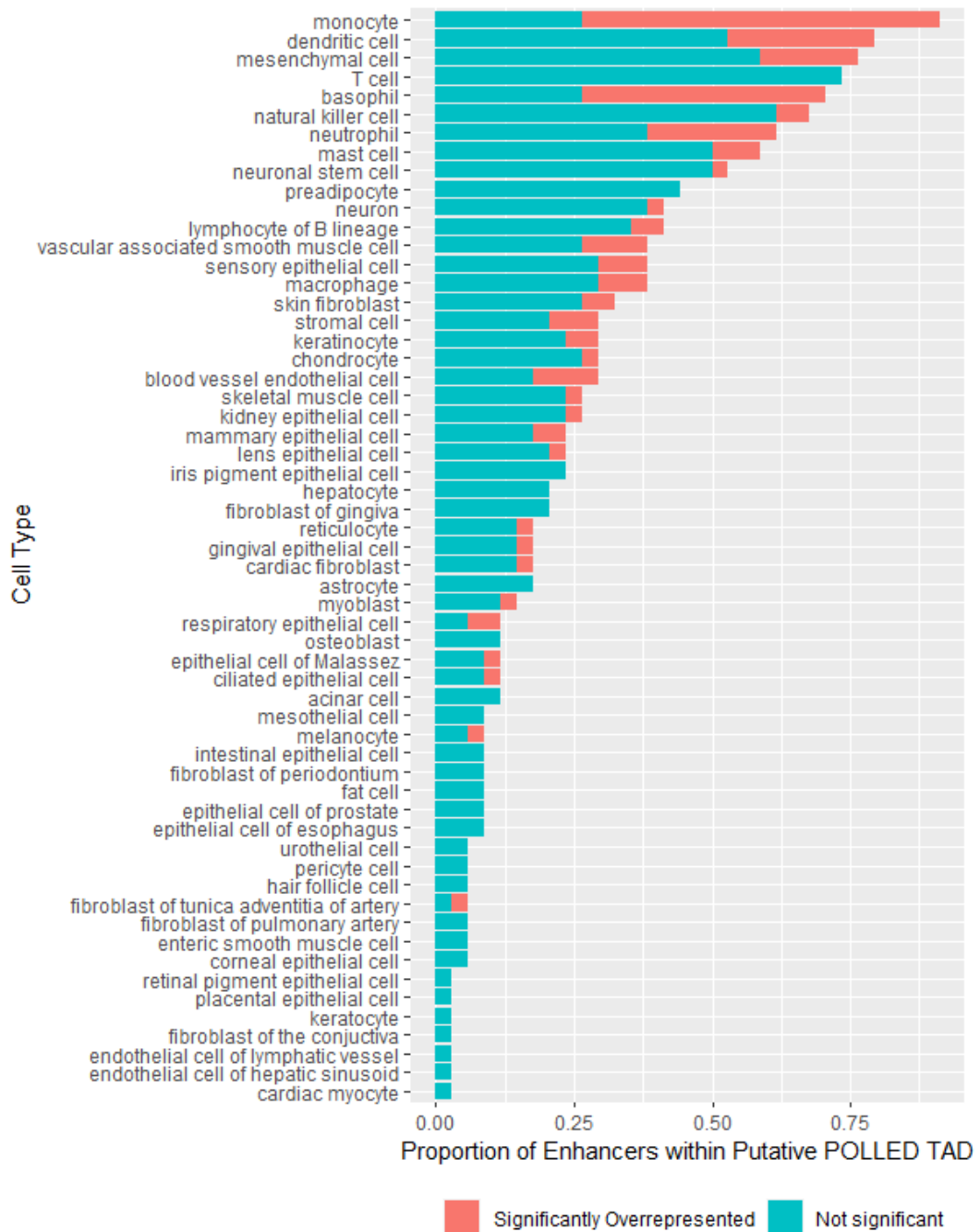Figure 2.4: The proportion of the 34 enhancers from the comparative POLLED TAD (chr21:34,001,068–35,284,703) that have activity in a given cell type. Aqua indicates the proportion of enhancers that are active but are not significantly over-represented. Red indicates enhancers that were active and significantly over-represented according to SlideBase. Only cell types in which enhancers were active are shown.

The cells and tissues in which the enhancers are active suggest that their function is important for gene expression relating to the nervous systems. Interestingly, some of the enhancers in the POLLED TAD are active in neuronal stem cells (18/34), neurons (13/34) and sensory epithelial cells (13/34). Assuming that the enhancers primarily interact with nearby genes, the activity of the enhancers should reflect the function of genes within this region. The role of the enhancers is interesting because bovine horns are innervated by sensory nerves (Godinho 1968; Madekurozwa 1996; Buda *et al.* 2011) and larger nerve bundles are present during early development (Wiener *et al.* 2015). Some of the enhancers in the polled region are also active in keratinocytes (10/34). During early fetal development, the epidermis is thickened at the horn bud by layers of keratinocytes (Wiener *et al.* 2015). Thus, enhancers in the region could regulate gene expression in cell types that are involved in horn ontogenesis. If enhancers are affected by the POLLED variants, then their activity might be altered in the tissues and cells presented here. However, unique bovine enhancers may affect transcription in different cell types.

2.3.5 Conservation

Regions within the putative POLLED TAD where sequence is conserved in horned species, but not in hornless species, were also identified. The genomic sequences of ten horned species, one antlered species (red deer) and 13 species without cranial appendages attached to the frontal bone were analysed for conservation of the putative POLLED TAD (Figure 2.5). The regions 1 kb down- and upstream of the Celtic and Mongolian POLLED variants were also considered. Musk deer were included because they represent a deer species that lost their antlers through evolution (Wang *et al.* 2019c). Sequences with greater than 80% identity were considered highly conserved.
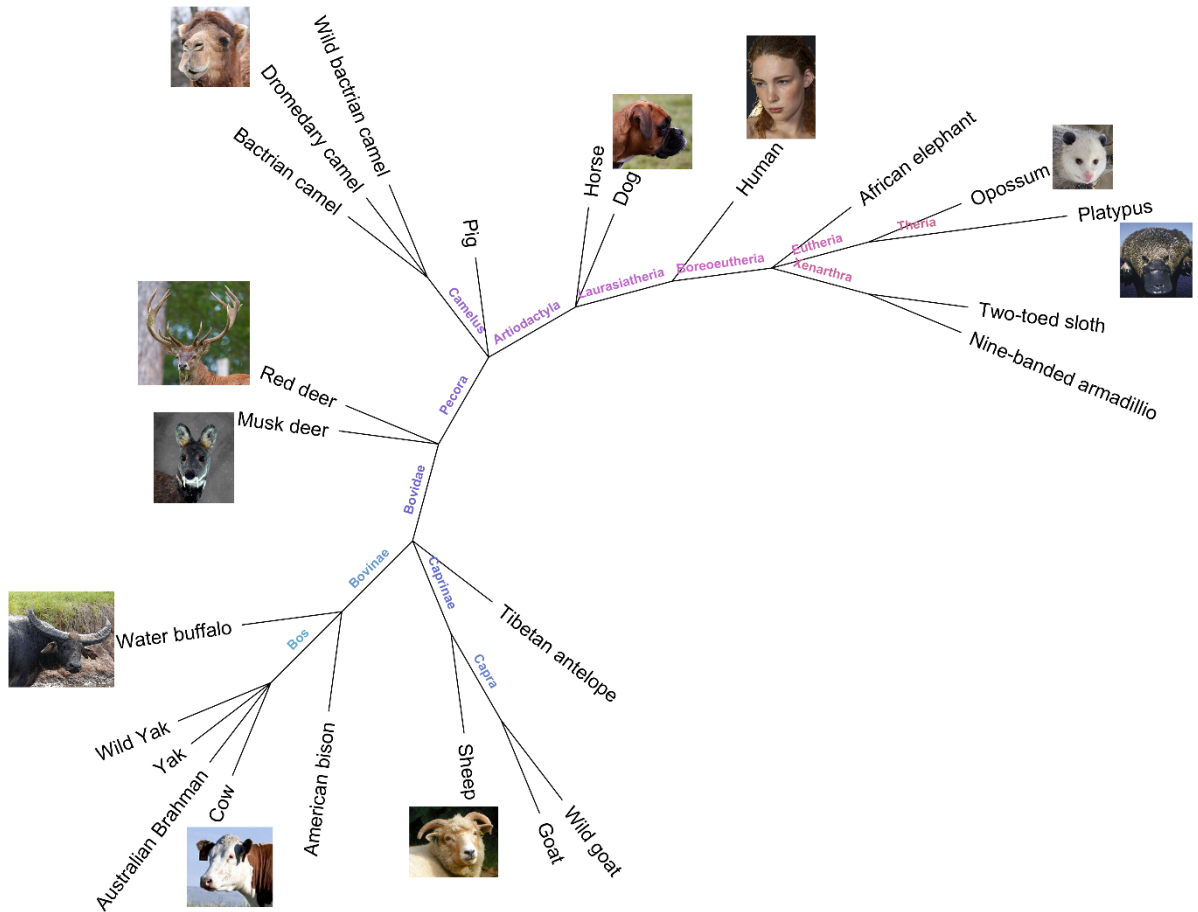
Figure 2.5: Phylogeny tree of species analysed for the sequence conservation of the putative POLLED TAD region. Phylogeny tree was generated using NCBI Lifemap (De Vienne 2016). Bovidae species have horns, musk deer are antlerless, red deer have antlers, and the remaining species lack headgear.

The sequence of the predicted POLLED TAD region was below 80% conservation for most regions in the Pecora species, which includes horned ruminants, antlered red deer and the musk deer (Figure 2.6). There was less overall sequence conservation with the other species that lack cranial appendages.

In terms of the sequence conservation in the immediate vicinity of the POLLED variants, there was ~1.5 kb of sequence encompassing the Celtic variant site that was moderately well conserved in all species examined, except the African elephant, human, opossum and platypus (Figure 2.7). Similarly, there was ~500 bp of sequence encompassing the duplication/insertion site of the Mongolian variant that is highly conserved in all species

examined except human, opossum and platypus (Figure 2.8). However, the conservation was below the 80% threshold for the regions overlapping the Celtic and Mongolian variants, so neither can be considered highly conserved.
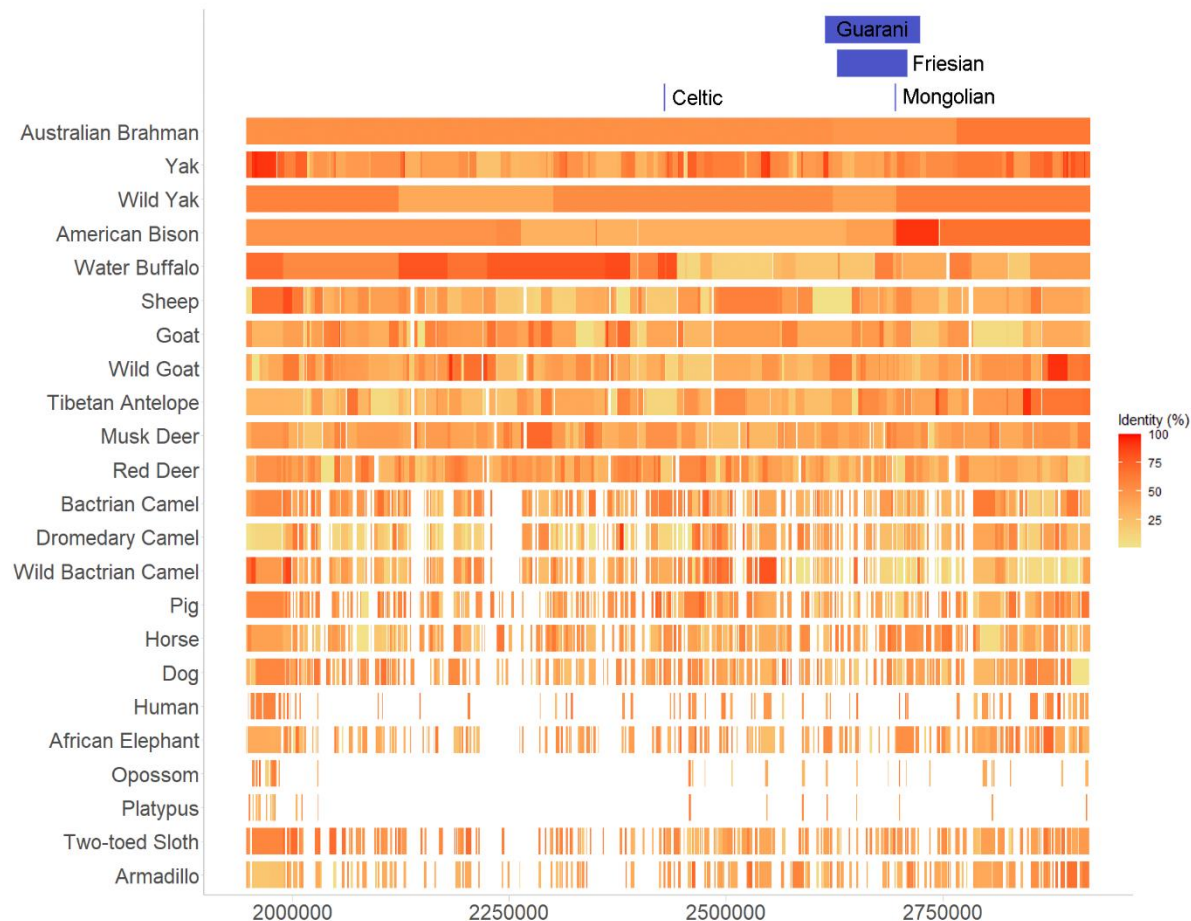
Figure 2.6: *Bos taurus* sequence conservation of the putative POLLED TAD. Blue blocks represent the location of the polled variants. First nine species are horned, musk deer are antlerless, red deer have antlers and the remaining 12 species lack headgear.
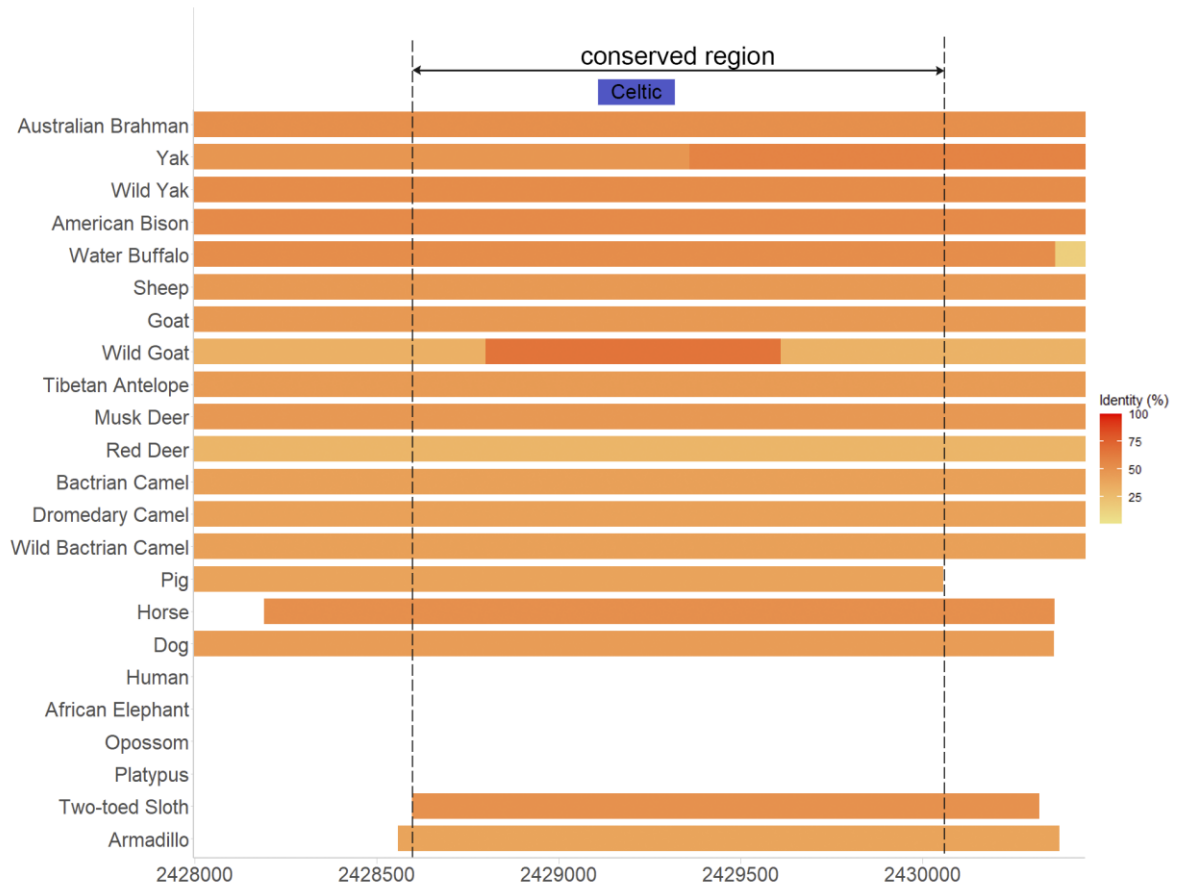
Figure 2.7: *Bos taurus* sequence conservation of the Celtic POLLED region (blue). Approximately 1.5 kb encompassing the Celtic variant is conserved across all species except African elephant, human, opossum and platypus. First nine species are horned, musk deer are antlerless, red deer have antlers and the remaining 12 species lack headgear.
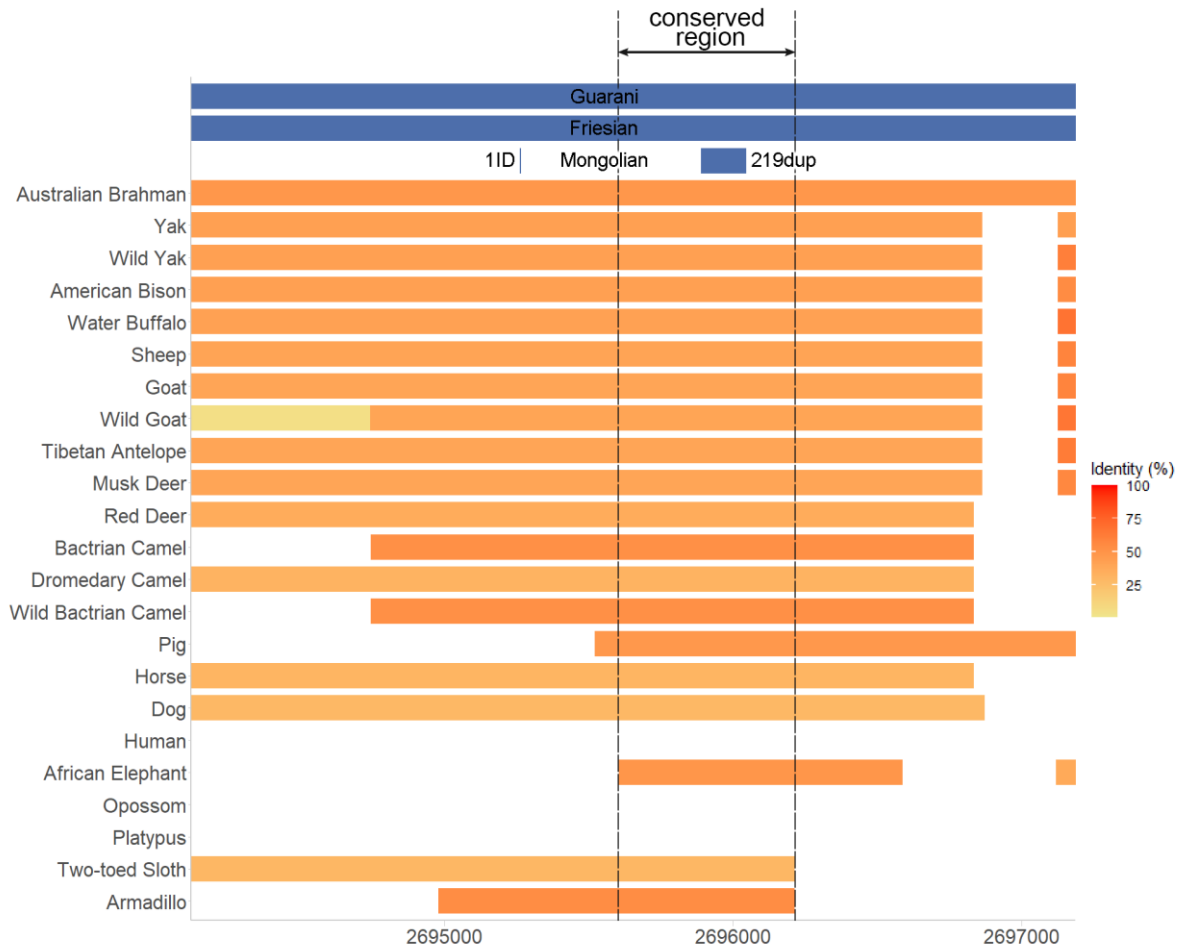
Figure 2.8: *Bos taurus* sequence conservation of the Mongolian POLLED region and sections of the duplicated sequence of the Guarani and Friesian variants (blue). Approximately 500 bp encompassing the 158 bp duplicated region is conserved across all species, except human, opossum and platypus. Blue blocks represent the location of the polled variants. First nine species are horned, musk deer are antlerless, red deer have antlers and the remaining 12 species lack headgear.

Analysis of the genomic sequence revealed that there was moderate conservation (< 80%) of the regions encompassing the location of the Celtic variant and the duplication/insertion of the Mongolian variant. These regions were conserved even in non-horned species that are evolutionarily distant such as armadillo and two-toed sloth. This suggests there may be some functional importance of these sequences, though they are not strictly protected from mutation. The conserved sequences overlapping the variants may

contain novel genes or lincRNA or regulatory DNA elements such as enhancers and promoters. For example, an enhancer that binds the HAND1 transcription factor is predicted from human data to overlap the 10-bp deletion of the Celtic variant as described above (Nguyen *et al.* 2018).

The hypothesis that there are conserved sequence unique to Pecora but in the other species analysed was not able to be addressed. There was not enough difference in conservation between the Pecora species to distinguish between regions that were conserved because of functional importance compared to regions conserved because there is less evolutionary divergence. Even for more distantly related Pecora species to cattle, such as red deer (*Cervidae*) and Siberian musk deer (*Moschidae*), the sequence in the putative POLLED TAD was conserved in large blocks and no gaps were present. As there were no gaps, specific highly conserved sequences and were not able to be identified. The five Pecora families (including Bovidae, Cervidae and Moschidae) diverged 23.3 to 20.8 million years ago (Wang *et al.* 2019c).

Interestingly, the conservation of the *Bos taurus* sequence within Australian Brahman (*Bos taurus indicus*) was lower than expected at ~ 51% identity. These are sub-species and it is estimated that ~ 10% of the Brahman genome is introgressed from *Bos taurus*. The low conservation identity may be an artefact of the alignment, as a large block of conservation was found compared to other species. This may have caused the total identity to be lower if the large block contained regions with high (> 85%) and lower conservation, but no gaps.

2.3.6 Conservation of gene synteny

The block of protein-coding genes in the putative POLLED TAD was compared across cattle, water buffalo, sheep, goat, pig, horse, dog, and human genomes to identify differences in synteny between horned and non-horned species. Gene synteny was also used to characterise conservation of lincRNAs via synteny of surrounding genes.

The analysis found conserved protein-coding gene synteny within the putative POLLED TAD between horned and hornless species which shows there were no species or headgear

specific chromosomal changes in this region that affected the gene order. The protein coding gene synteny was conserved for both position and direction for the four horned and four hornless species (Figure 2.9).

The gene synteny within the putative POLLED TAD may have been conserved for functional reasons (e.g. when a regulatory sequence of a given gene is located within a nearby protein-coding gene) (Engstrom *et al.* 2007; Kikuta *et al.* 2007; Irimia *et al.* 2012; Wong *et al.* 2020). Such genomic regions tend to retain genes that are important for expression of key developmental processes (for example, the HOX gene clusters) (Lee *et al.* 2006; Kikuta *et al.* 2007). An alternative explanation would be that the region is not prone to chromosomal rearrangements. There is a bias for chromosomal rearrangements to occur in regions that susceptible or "fragile" (Pevzner & Tesler 2003). Whereas, conserved regions may be affected less by rearrangements.

Note that the analysis of gene synteny is limited by the annotations available. Deep sequencing of RNA from four domestic species suggests that there are a significant number of unannotated protein-coding genes (Foissac *et al.* 2019). There was little difference in gene synteny between species in the present study, but incomplete gene annotation cannot be excluded.
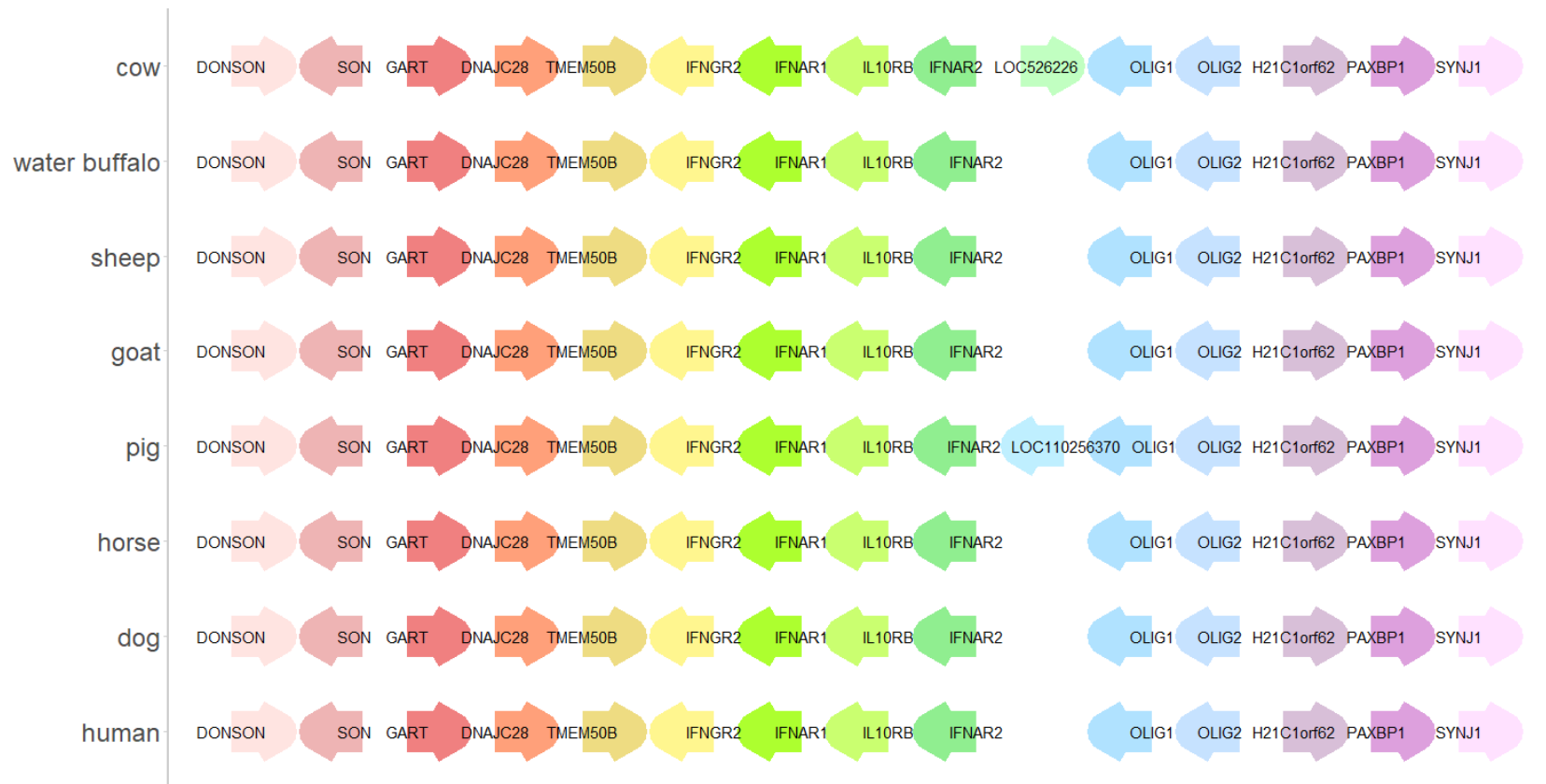
Figure 2.9: Gene synteny map of the putative POLLED TAD across horned (cattle, water buffalo, sheep, and goat) and hornless species (pig, horse, dog, and human). Synteny of genes (arrows) is conserved across species. Cattle and pig have additional protein-coding genes annotated within the TAD, LOC526226 and LOC110256370, respectively. The chromosomes and regions are listed in Appendix Table A2.

Despite the conserved gene synteny between the species, there were additional protein-coding genes annotated for cattle (*LOC526226*) and for pig (*LOC110256370*) between *IFNAR2* and *OLIG1* in the putative POLLED TAD (Figure 2.9). These loci, *LOC526226* and *LOC110256370*, are not found in the same location and did not share sequence with the other species.

The additional bovine gene *LOC526226* codes for histone H4 in a single exon and is 312 bp in length. *LOC526226* was aligned to the bovine reference genome and matched histone H4 annotations at three other locations in the genome, outside the polled region (Table 2.2). Differences in *histone H4* annotation between the cattle, water buffalo, sheep, goat, pig, horse, dog and human genomes indicate that there are different copy numbers of this gene in the various species (Table 2.2). These genes are located in different regions of each genome, but none of the *histone H4* genes are located in the POLLED TAD region with the exception of the copy in the bovine genome (Figure 2.9).

Table 2.2: Number of histone H4 annotations in horned and hornless species.

Species	Number of histone H4 annotations
Cattle	14
Water buffalo	6
Sheep	10
Goat	5
Pig	10
Horse	11
Dog	10
Human	13

Histone H4 is one of four core histones that associate with DNA to package the DNA into nucleosomes (Zhou & Bai 2019), but an alternative transcript from this gene in humans

and mice codes for the growth factor osteogenic growth peptide (OGP) (Bab *et al.* 1999b; Pigossi *et al.* 2016). OGP is a highly conserved, 14 amino acid long growth factor that is translated from the C-terminus of histone H4. A pre-OGP molecule is translated from an alternative start codon (amino acid 85-103) (Bab *et al.* 1999b). Pre-OGP is then converted to OGP by the removal of five terminal amino acids (90-103; NH2-ALKRQGRTLYGFGG-OH) (Bab *et al.* 1999b). The protein can be further cleaved to produce the physiologically active variant of OGP, OGP(10-14) (NH2-YGFGG-OH), which consists of only the last five amino acids of OGP (Bab *et al.* 1999a; Gabarin *et al.* 2001). Examination of *LOC526226* sequence shows that the alternative start site for OGP is conserved. If *LOC526226* were to be translated to OGP, the only amino acid change would be a substitution of arginine for cysteine at the 7th position.

OGP plays a role in bone formation by promoting osteoblast cell proliferation and osteoblastic differentiation of mesenchymal stem cells (Pigossi *et al.* 2016). It was initially isolated from bone marrow tissue and has been detected in human and mouse serum (Pigossi *et al.* 2016). The function of OGP and proximity of *LOC526226* to the Celtic variant (~30 kb), suggests that this gene could potentially have a role in horn ontogenesis. As the gene at this locus is unique to cattle, it may explain why the cattle POLLED variants are concentrated in this region on chromosome 1.

The *LOC110256370* identified in this region of the pig genome spans 57,789 bp and has four exons. It is annotated in NCBI but not Ensembl, and is classified by NCBI as a 'low-quality protein' like mRNA-40S ribosomal protein S2. The S2 protein is a subunit for ribosomes, the molecular machine that translates mRNA to protein (Kressler *et al.* 2017).

2.3.7 Conservation of LincRNA

LincRNAs in the POLLED region may be functionally important for the development of horns in cattle. *LincRNA#1* and *LincRNA#2*, which are located near the POLLED variants, were found to be differentially expressed in the bovine horn bud tissue during early fetal

development (70-90 days) in RNAseq and qPCR studies (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014). Therefore, bovine lincRNA within the putative POLLED TAD region were assessed for conservation across horned and non-horned species. Two methods were used to determine the conservation of bovine lincRNA, conservation of synteny (gene-lincRNA-gene triplets) and conservation of sequence.

Five long intergenic non-coding RNA are annotated in the bovine putative POLLED TAD (Table 2.3), including *LincRNA#1* (*LOC100848368*) and *LincRNA#2* (*LOC112447133*) (Allais-Bonnet *et al.* 2013). The genes flanking the five bovine lincRNA were identified to form triplets (gene, lincRNA, gene), taking strand of expression into account. These triplets were then compared across horned and hornless species (water buffalo, sheep, goat, pig, horse, dog, and human) (Appendix Table A4). Based on this comparison, the conservation of the LincRNAs was determined. The sequence conservation of lincRNA were also assessed (Table 2.4)

Table 2.3: Bovine lincRNA in the putative POLLED TAD.

LincRNA	Location	Strand	Accession numbers
LOC104970777	chr1:2,016,219-2,027,521	-	XR_804069.3
LOC112447120	chr1:2,162,947-2,173,433	-	XR_003035212.1
LOC104970778	chr1:2,241,318-2,251,416	-	XR_804070.3
LincRNA#1 (LOC100848368)	chr1:2,506,494-2,509,757	-	XR_003035217.1, XR_003035225.1, XR_233195.4
LincRNA#2 (LOC112447133)	chr1:2,615,875-2,620,559	+	XR_003035236.1

The sequence of *LOC104970777* was highly conserved in horned species but not in hornless species (Table 2.4). *LOC104970777* aligned to sequence in horse and dog, but did not

align with any sequence in the corresponding region in pig and human. Therefore, *LOC104970777* is potentially a Bovidae specific lincRNA based on the sequence conservation.

The syntenic triplet including *LOC112447120* was only conserved in one other species, pig (ENSSSCG00000051450), and < 1 kb from this lincRNA the sequence was conserved (query coverage = 43%, identity = 72%). The sequence of *LOC112447120* was highly conserved in horned species but not in hornless species (Table 2.4). Therefore, *LOC112447120* is potentially a Bovidae specific lincRNA based on the sequence conservation.

The sequence of *LOC104970778* was highly conserved in all horned species, pig and horse (Table 2.4). The syntenic triplet including *LOC104970778* was conserved in sheep and pig and the sequence was shared with the lincRNA in sheep and pig. Sequence conservation was low in dogs and human. *LOC104970778* does not appear to be Bovidae specific.

The syntenic triplet including LincRNA#1 was conserved in water buffalo, horse, dog, and human. The LincRNA#1 sequence was highly conserved in horned and hornless species (Table 2.4). The sequence aligned to the syntenic lincRNA of water buffalo, dog and human, and matched sequence < 1 kb from the lincRNA in horse. Therefore, LincRNA#1 is not Bovidae specific.

LincRNA#2 sequence was highly conserved in horned and hornless species (Table 2.4). The LincRNA#2 triplet was conserved in water buffalo, dog, and human. Only the sequence conserved in water buffalo corresponded to a syntenic LincRNA (LOC112583890). Based on sequence conservation, LincRNA#2 is not Bovidae specific.

In summary, the five bovine lincRNAs in the POLLED TAD region were examined for conservation of synteny and sequence with lincRNA annotated in other species. No lincRNAs showed conservation of synteny in horned species that were not conserved in hornless species. Four lincRNA had conserved synteny and sequence identity of > 75.6% with lincRNA in at least one other species (Table 2.5); the exception was *LOC104970777* which did not have any synteny in other species. The sequence of *LOC104970777* and *LOC112447120* was conserved

in horned species only, although *LOC104970777* shares syntenic conservation with dog. The function of these lincRNAs are not known.

Table 2.4: Alignment scores (NCBI discontiguous megablast) of bovine lincRNA across horned (water buffalo, sheep and goat) and hornless species (pig, horse, dog and human). Query = the percentage of sequence that was highly similar; Ident = the percentage of highly similar sequence that is identical. Grey = highly conserved (≥50% query coverage and ≥50% identity); Dark blue = highly conserved and aligns to lincRNA that has conserved synteny; Light blue= highly conserved and aligns to sequence <1kb from lincRNA with conserved synteny; Yellow = not highly conserved and aligns to sequence <1kb from lincRNA with conserved synteny.

ID	LOC104970777		LOC112447120		LOC104970778		LOC100848368 (LincRNA#1)						LOC112447133 (LincRNA#2)	
Accession	XR_804069.3		XR_003035212.1		XR_804070.3		XR_003035217.1		XR_003035225.1		XR_233195.4		XR_003035236.1	
	Query	*Ident*	*Query*	*Ident*	*Query*	*Ident*	*Query*	*Ident*	*Query*	*Ident*	*Query*	*Ident*	*Query*	*Ident*
Water buffalo	100	93	100	96	99	97	100	97	100	97	100	97	100	97
Sheep	71	89	99	91	99	91	99	97	99	97	99	97	99	96
Goat	71	84	99	91	99	91	99	97	99	97	99	97	99	96
Pig			43	72	65	76	64	87	65	87	60	87	67	81
Horse	22	68	46	69	78	75	80	87	82	87	70	87	80	80
Dog	19	78	21	73	32	74	79	81	80	82	74	82	61	79
Human			22	67	20	75	62	83	62	83	59	83	66	84

Table 2.5: Bovine lincRNA with conserved synteny and sequence in other species. The lincRNA annotated in cattle were analysed for synteny and sequence conservation in horned (water buffalo, sheep and goat) and hornless species (pig, horse, dog and human).

Gene Upstream	Gene Downstream	Species	LincRNA
IFNGR2	IFNAR1	Cattle	LOC112447120
		Pig	ENSSSCG00000051450*
IFNAR1	IL10RB	Cattle	LOC104970778
		Sheep	LOC114113258
		Pig	ENSSSCG00000049143
OLIG1	OLIG2	Cattle	LincRNA#1 (LOC100848368)
		Water buffalo	LOC102394079
		Horse	ENSECAG00000030232*
		Dog	LOC111093515 (ENSCAFG00000037762)
		Human	LINC00945
OLIG2	C1H21orf62	Cattle	LincRNA#2 (LOC112447133)
		Water buffalo	LOC112583890

*lincRNA did not share sequence but was < 1 kb from conserved sequence

LincRNA#1 had the most evidence for conservation across all the species, with shared synteny and sequence in both horned and hornless species. The human ortholog of lincRNA#1, LINC00945, is over-expressed in the testis, spinal cord (cervical c-1), and whole blood (Stelzer *et al.* 2016). This function of this lincRNA matches with the activity of the human enhancers (see section 2.3.4.2). These enhancers are active in the brain, blood and testis.

LincRNA#2 had high sequence conservation across all the species, but synteny was only conserved in water buffalo. The sequence conservation of LincRNA#2 could suggest that it has functional importance of this sequence in several species.

LincRNA#1 and LincRNA#2 trend towards being differentially expressed between horned and polled fetal horn bud (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014). More specifically, LincRNA#1 had decreased expression in the horned horn bud compared to polled, while LincRNA#2 had increased expression (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014).

However, the evidence in the present study suggests that these lincRNA are not specific to horned species. Even so, differences in regulation of these lincRNA could affect the development of horns.

LincRNAs are poorly annotated in some reference genomes. When a lincRNA is present but not annotated in a given species, synteny cannot not be addressed. The deficiency in the lincRNA annotation interferes with analysis. For example, sheep and goat had one and two lincRNAs annotated within the region, respectively, even though the other species had 5-13 lincRNAs annotated in the region (Appendix Table A5). Furthermore, the FR-AgEnCODE project detected 13,864 goat lncRNAs but only 16% were known (Foissac *et al.* 2019). In the present study, the sequence of bovine lincRNAs were highly conserved in sheep and goat. It is, therefore, likely that the analysis herein was limited by the annotation data available.

For species closely related to cattle, sequence conservation of lincRNAs may have occurred by chance rather than because of functional constraints. In other words, the lincRNA sequence may have high conservation just because the intergenic region itself maintains high conservation. However, the conservation analysis of this region shows that the identity of conserved segments ranges from high (>80%) to low (<20%) for some species (Appendix Figure A2-Figure A3, section 2.3.5). As some bovine lincRNA had > 90% identity in horned species which was greater than the general conservation of the putative POLLED region for a given species, the lincRNA may be more selectively conserved than others intergenic regions due to functional constraints. Thus, it seems likely that the sequences are functionally important, though there is still a chance that they are conserved due to evolutionary relatedness.

2.3.8 Topological associated domains

The TAD structure can be used to define a region of interest when assessing the function of a causative variant. Disruption of TAD structure can cause phenotypic changes by altering chromatin-chromatin interactions (Lupiáñez *et al.* 2015). To determine whether the TAD structure is affected by the Celtic variant, Hi-C data from a polled Angus genome carrying the

Celtic variant and from a Brahman horned animal carrying the ancestral horned allele were compared. The TAD structures for the Angus and Brahman sequences were identified (Appendix Table A6) and heat maps generated by aligning Angus and Brahman Hi-C sequences to the Angus and Brahman genomes, respectively (Appendix Figure A4), as well as to the bovine reference genome (Figure 2.10, Appendix Table A7).

The analysis revealed seven TADs within the previously predicted POLLED TAD (Figure 2.10) and identified boundaries ~13.5 kb and ~39 kb from the putative TAD boundaries (Wang *et al.* 2018). Insulator CTCF binding sites were found at the TAD boundaries (Appendix

Table A8) and were convergent in orientation, validating the Angus and Braham TADs. The exception was one boundary at 2,080,000 bp where there were no CTCF motifs. The TAD boundaries were compared between polled Angus and horned Brahman Hi-C sequences and no differences in TAD structure were found.

As there was no difference in TAD structure of the POLLED region between the polled and horned genomes, this suggests that the Celtic variant does not impact the TAD structure. While the position of the other POLLED variants can be mapped within the TAD structure, no

HiC data has been generated for genomes carrying these variants so their effect on TAD structure could not be addressed.
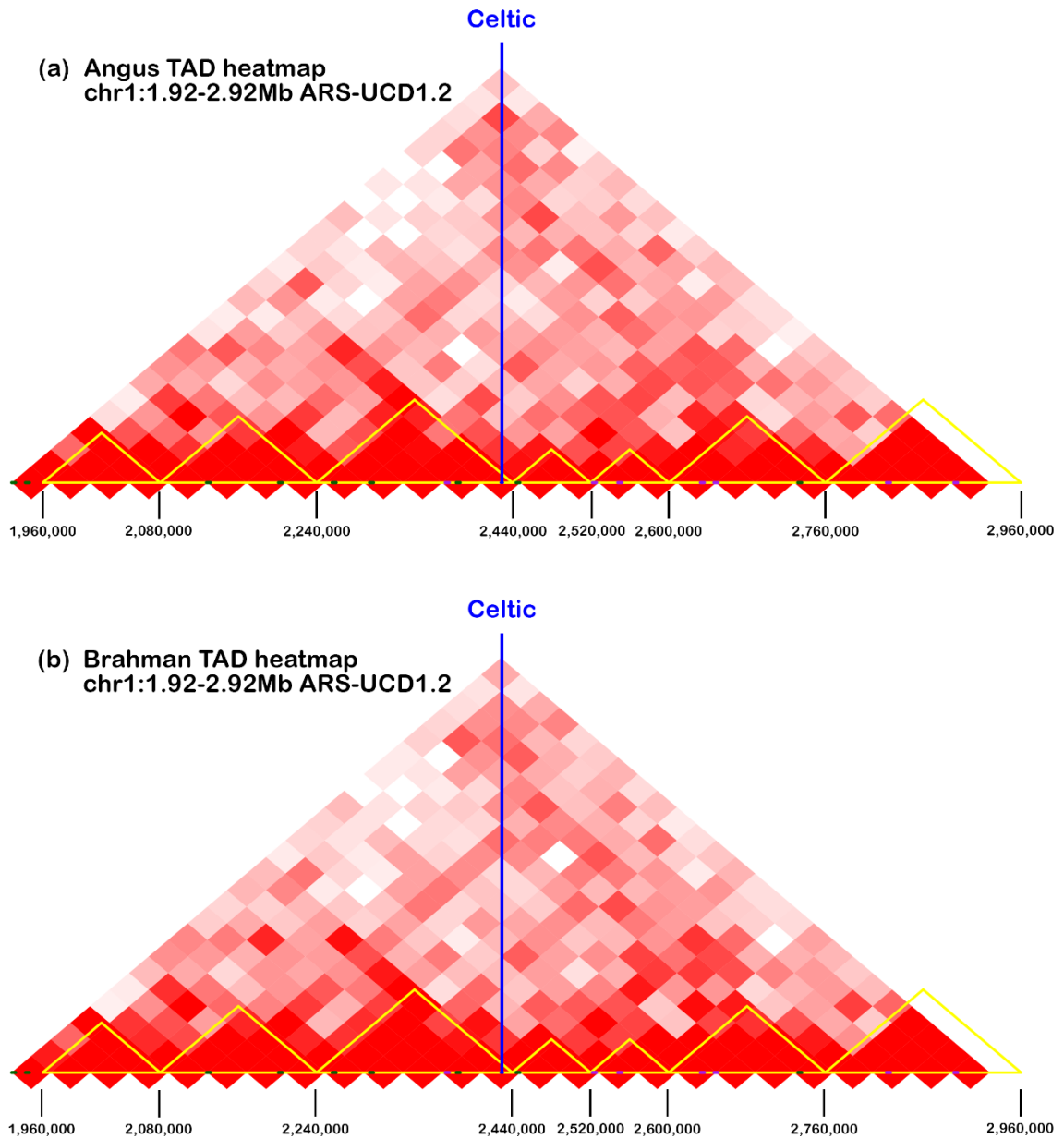
Figure 2.10: Heat map showing the TAD structure of the genomic region surrounding the POLLED variants. Hi-C data from (a) Angus and (b) Brahman were aligned to the *Bos taurus* reference genome (ARS-UCD1.2). The position of the Celtic variant is indicated by the blue line, and TADs are indicated by the yellow triangles along the x-axis.

TADs are regions of genome that are self-interacting. In other words, the loci within a domain are more likely to interact with other loci within the same domain than with loci in a different domain. Thus, the POLLED variants are more likely to be interacting with loci within the same domain and genes important for horn development are more likely to be within the four TADs within the POLLED region (510 kb) identified herein. The POLLED region contains six genes, four lincRNA, and one pseudogene.

Based on the analysis, the four POLLED variants were situated within a region that contained four TADs: two TADs containing the known variants were separated by two other TADs (Figure 2.11). These four TADs will be referred to as the "POLLED region" (chr1:2,240,000-2,759,999). As the POLLED variants are concentrated in this area, some loci of the molecular pathway that signals horn ontogenesis in cattle is likely to be present in this genomic region. The Celtic variant (P_C) is located 10,680 bp from the downstream boundary in a TAD that spans chr1:2,240,000–2,439,999 (TAD 1). TAD 1 contains three genes (*IL10RB, IFNAR2* and *LOC526226*), two lncRNAs (*LOC104970778* and *LOC112447121*) and a pseudogene (*LOC107132172*) (Table 2.6). The remaining variants (P_G, P_F, P_M) are located in a TAD that spans chr1:2,600,000–2,759,999 (TAD 4) which includes LincRNA#2 and intersects with C1H21orf62.
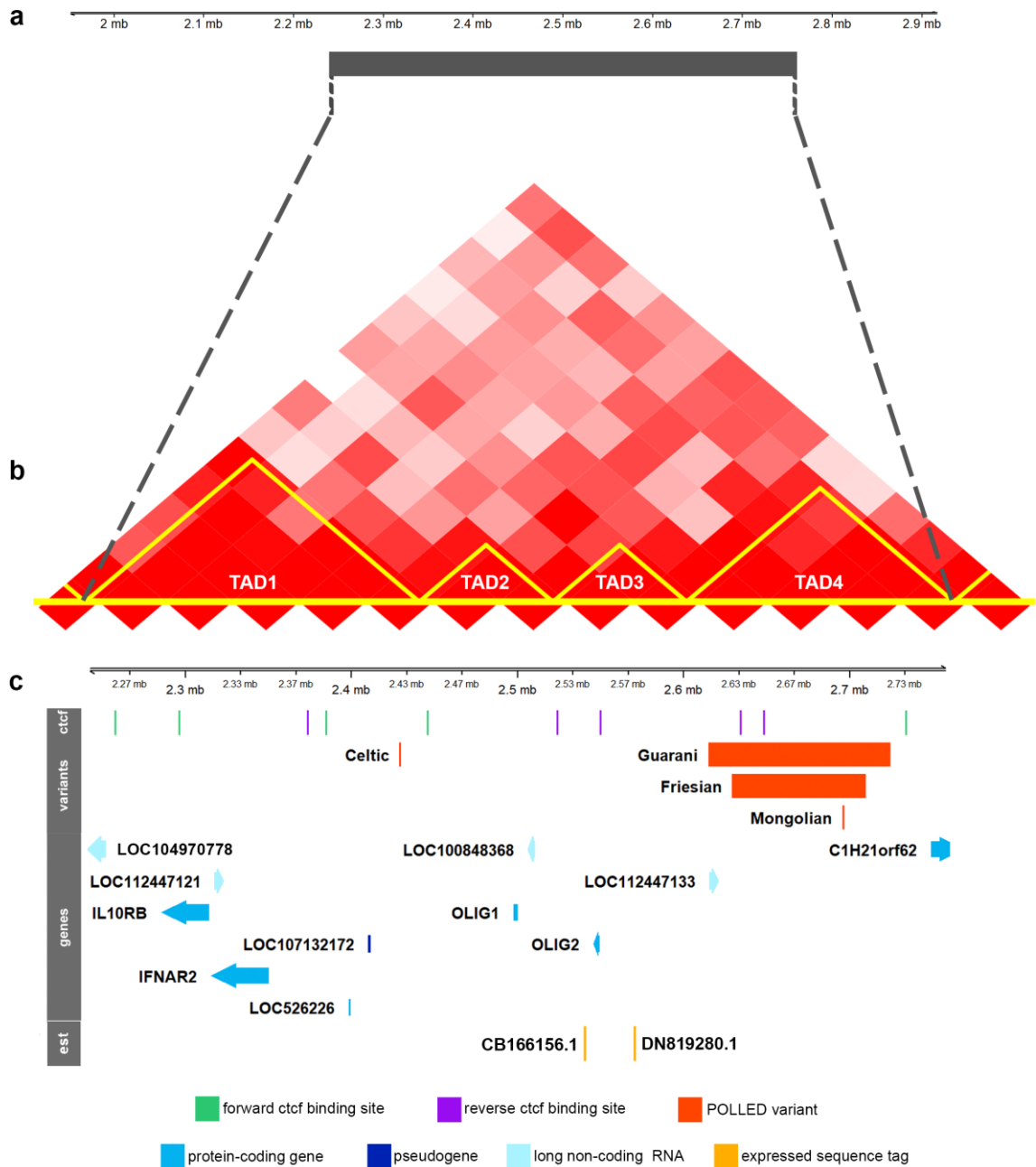
Figure 2.11: POLLED region includes four topologically associating domains (TADs). (a) The putative TAD (Chr1:1,946,384 – 2,921,213; ARS-UCD1.2) predicted Wang *et al.* (2018). (b) Hi-C heat map. The Celtic variant is located in a different TAD (TAD 1) to the Friesian, Mongolian and Guarani variants (TAD 4). (c) Annotations within the four TADs including the predicted CTCF binding sites, variants, genes, lncRNA and expressed sequence tags (ESTs). *LOC526226* encodes histone H4 and may alternatively transcribe the osteogenic growth peptide.

Table 2.6: Genes, non-coding RNA and POLLED variants in the TADs within the bovine POLLED region.

Number	TAD	Annotation	Biotype
TAD 1	Chr1:2,240,000 – 2,439,999	*LOC104970778*	LincRNA
		IL10RB	Protein-coding
		IFNAR2	Protein-coding
		LOC112447121	LncRNA
		LOC526226 (Histone H4/OGP)	Protein-coding
		LOC107132172	Pseudogene
		P_C	Variant
TAD 2	Chr1:2,440,000 – 2,519,999	*OLIG1*	Protein-coding
		LincRNA#1 (LOC100848368)	LincRNA
TAD 3	Chr1:2,520,000 – 2,599,999	*OLIG2*	Protein-coding
TAD 4	Chr1:2,600,000 – 2,759,999	P_G	Variant
		LincRNA#2 (LOC112447133)	LincRNA
		P_F	Variant
		P_M	Variant
		C1H21orf62 *	Protein-coding

* crosses TAD boundary

Gene expression studies that compared horned and polled (P_C) bovine fetuses have only analysed two genes, *IL10RB* and *IFNAR2,* within TAD 1, that carries the Celtic variant (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014). The expression of *IL10RB* and *IFNAR2* are not affected by the Celtic variant in 90 day old fetuses.

TAD2 contains *OLIG1* and LincRNA#1, and does not contain any POLLED variants. There is no evidence that *OLIG1* is differentially expressed between horned and polled fetal horn buds (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014). However, *OLIG1* was identified as a gene that might have been recruited to serve a function in horn development (Wang *et al.* 2019c). *LincRNA#1* had increased expression in horn bud region when compared to frontal skin in 90 day old both polled fetuses (Allais-Bonnet *et al.* 2013). *LincRNA#1* expression was almost significantly different (p = 0.052) between the horn bud region in horned and polled fetuses from the same study (Allais-Bonnet *et al.* 2013).

OLIG2 was the only annotation in TAD3. *OLIG2* was differentially expressed between horn bud tissue and frontal skin of both horned and polled fetuses (Allais-Bonnet *et al.* 2013). It was not differentially expressed between the horned horn bud and the horn bud region in polled fetuses (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014).

TAD4 includes *LincRNA#2* and part of *C1H21orf62*. LincRNA#2 may have lower expression in the polled horn bud compared to the horned horn bud in 70 day old fetuses (Wiedemar *et al.* 2014), although statistics was not conducted due to low sample numbers. Allais-Bonnet *et al.* (2013) did not report *LincRNA#2* as being differentially expressed. However, the authors described LincRNA#2 as being 73 kb (chr1:2,547,895-2,621,221bp), while the current annotation is only 4.7 kb. Therefore, primers used to amplify part of the LincRNA#2 might not have corresponded to the transcribed region. A comparison of transcriptomic expression from RNAseq found *C1H21orf62* to have lower expression in the horn bud of a horned fetus when compared to a polled fetus at ~5 months of development (Wiedemar *et al.* 2014). This was validated by qPCR analysis of horn bud and frontal skin samples from horned and polled fetus of various ages ranging from ~70-175 days of fetal development, though most age groups only had one horned and one polled fetus (Wiedemar *et al.* 2014). Wiedemar *et al.* (2014) concluded that expression of this gene was lower in the horn bud, irrespective of genotype (horned or polled). Therefore, more research is required determine if *C1H21orf62* is affected by POLLED variants.

Overall, the evidence from the fetal gene expression studies (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014) suggests that the Celtic variant potentially affects the expression of LincRNA#1, *OLIG2*, LincRNA#2 and *C1H21orf62*. The Celtic variant could be interacting with loci outside of its domain. Inter-TAD interactions are possible when the boundary between adjacent TADs is weak (Foissac *et al.* 2019). This is demonstrated in the heat map (Figure 2.10), where high frequency interactions across boundaries can be observed. Single cell analysis

of TAD structures show that boundaries may be fluid between cells (Luppino & Joyce 2020). The hypothesis that the Celtic variant interacts across domains cannot be excluded.

Considering that the Celtic variant occurs in a different TAD from the other three polled variants, slightly different mechanisms or effects may disrupt horn ontogenesis. In support of this idea, carriers of the Celtic and Friesian variants appear to have different scur to polled phenotypic proportions (Gehrke *et al.* 2020a; Lyons & Randhawa 2020). Gehrke *et al.* (2020a) reported that the Friesian variant blocked the expression of an intermediate phenotype, scurs, in P_F/p animals more effectively than the Celtic variant (Gehrke *et al.* 2020a). This suggests that there are subtle phenotypic differences between the Celtic and Friesian variants. The TAD structure within this region may offer an explanation for the different phenotypes. The TAD structure for the Friesian, Mongolian and Guarani variants still needs to be examined.

The sample size is a major limitation of the TAD analysis. The analysis reported here used horned sequences and polled sequences, each from one individual. Therefore, there may not have been enough power to detect subtle differences between the horned and polled TADs. Future studies should compare chromatin structure between genomes carrying horned and POLLED variants in sufficient numbers to detect subtle differences. Characterising the effect of the Friesian and Guarani variants would be particularly interesting as they include large duplications.

2.3.9 Chromatin interactions of the Celtic variant

It has been proposed that the Celtic variant partially deletes an enhancer (Nguyen *et al.* 2018), which may change the chromatin-chromatin interactions. Using Angus and Brahman Hi-C sequence data, the location of the Celtic variant was examined to test if loci that significantly interacted with the position differed.

Horned (Brahman) and polled (Angus) Hi-C sequence data were used to identify significant interacting regions with the polled domain around the Celtic variant within 10 kb

bovine reference genome windows. Six interacting regions were identified for both the Brahman and Angus sequences, five of which were common between the two breeds (Appendix Table A9). In Brahman, the Celtic 20 kb window uniquely interacted with a window at chr1:3,090,000–3,100,000 bp that did not contain any annotated genes or other features. In Angus, the Celtic 20 kb window uniquely interacted with a window at 970000–980000 bp that maps to an intron of the gene Regulator of Calcineurin 1 (*RCAN1*). This interaction spans 1.44 Mb.
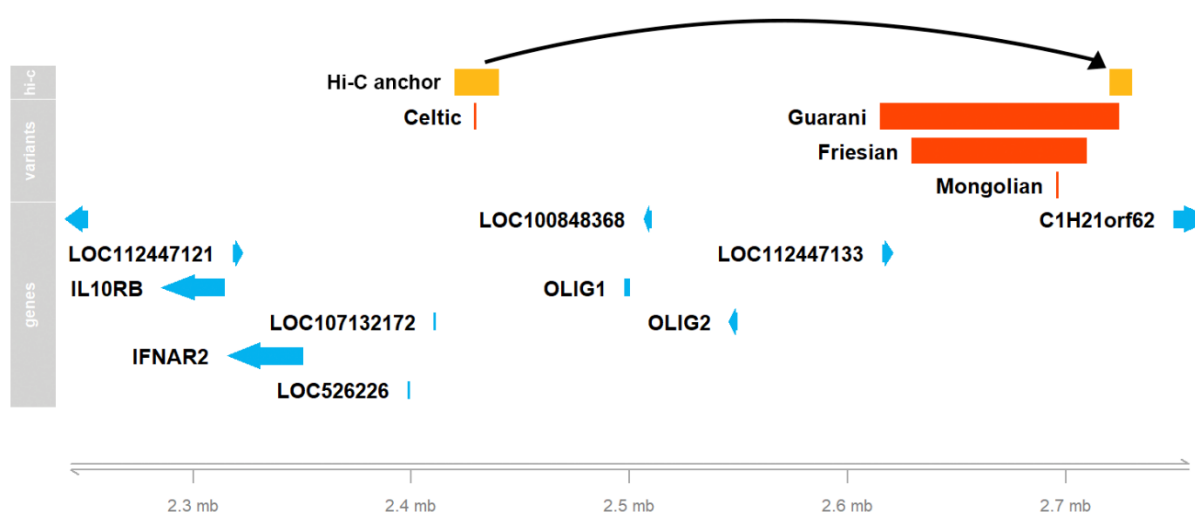
Figure 2.12: Gene map showing the significant interaction (yellow) between the Celtic region and the 3' end of the Guarani duplicated region in both breeds. The analysis was conducted using Hi-C data from fetal lung tissue of Brahman (horned) and Angus (polled).

RCAN1 (also called Down syndrome critical region 1 protein [DSCR1]) is a calicneruin inhibitor that binds to calcineurin molecules to suppress signaling pathways (Patel *et al.* 2015; Wang *et al.* 2016; Li *et al.* 2020). Transgenic mice with three copies of *Rcan1* had significantly fewer sympathetic neurons, and fewer sympathetic nerve fibers in target tissues (Patel *et al.* 2015). In contrast, increased axonal growth has also been reported for primary murine hippocampal neurons that overexpress *Rcan1* (Wang *et al.* 2016; Seo *et al.* 2019). Therefore,

Rcan1 expression is important for axon steering (Wang *et al.* 2016). Neural development may be key for the development of horns, which are innervated. The changes in the interactions with *RCAN1* identified in this analysis suggest a potential target gene for the Celtic POLLED variant.

In addition to the interaction detected in *RCAN1*, the Celtic region interacted significantly with a window ~20 kb upstream of *C1H21orf62* (chr1:2,749,168-2,768,577), which spans 320 kb in both horned and polled sequences, which suggests the Celtic variant does not affect the interaction (Figure 2.12). However, this interaction demonstrates a connection between the Celtic variant and TAD 4, which contains all the other variants. The 10 kb window in TAD4 contains the end of the duplicated region of the Guarani variant, suggesting that this variant also has the potential to directly alter this interaction.

While TAD structure is generally conserved between tissue types, the individual interactions are tissue specific, particularly during development. The Hi-C data used were collected from fetal lung tissue at 90 days of development. Therefore, the small-scale interactions observed here may not be observed in the tissues important for horn development. Different interactions may be observed in horn bud or the cranial neural crest, during a key period of horn ontogeny. The results here would need to be validated in horn specific tissue.

2.4 Conclusion

The mechanism of bovine horn development and the effects of the POLLED variants that result in the polled phenotype are unknown. The region encompassing the bovine POLLED variants was investigated *in silico* using bioinformatic methods to explore the mechanisms of horn development and the effects of the POLLED variants.

The putative POLLED TAD was studied by 1) investigating potentially overlapping, functional annotations (miRNA, ESTS, and enhancers), 2) comparing horned and polled sequences (ORF, TADs, chromatin interactions), 3) examining the conservation in horned and

non-horned species (sequence, gene synteny, lincRNA) and 4) the function of the putative POLLED TAD was inferred from human enhancer activity.

LincRNA#2 was found to overlap the region duplicated by the Guarani variant. However, no annotated or potential unannotated genes, lincRNA or miRNA were found to overlap the other POLLED variants. One miRNA and two ESTs were found within the putative POLLED TAD. One putative enhancer was found to overlap the 10 bp deletion site of the Celtic variant as reported by Nguyen *et al.* (2018). Using the same database (Nguyen *et al.* 2018), nine putative enhancers were found to overlap the Guarani variant, six of which also overlapped the Friesian variant.

Based on the TAD analysis, the POLLED region was further refined from a single TAD of 975 kb to four TADs covering 520 kb which encompasses all four POLLED variants. The TAD structure did not differ between genomes carrying horned and polled (P_C) sequences, although this needs to be validated as only one sample of each genotype was analysed. The analysis also showed that the Celtic variant was located in a different TAD from the other variants. A future study could investigate TAD structures in animals carrying other variants, particularly, the Friesian and Guarani variants because these are large duplications.

The sequence conservation analyses between horned and polled species showed that overall gene synteny is conserved, and most bovine lincRNA were conserved in non-horned species. Of note, cattle have an additional *histone H4* gene (LOC526226) within the POLLED region. *Histone H4* is a protein-coding gene that alternatively translates the osteogenic growth peptide, a protein involved in differentiation of mesenchymal cells to osteogenic cells. Given that the gene may be alternatively transcribed, long read RNA sequencing to identify expressed isotypes may be most appropriate to study expression levels of this gene. Based on sequence conservation, two lincRNA (*LOC104970777* and *LOC112447120*) may be conserved in horned species and not non-horned species, although their functions are unknown. Additionally, the

wild-type horned bovine sequences at the Celtic and Mongolian variant sites were conserved in distantly related mammals, suggesting functional importance of these sequences.

Overall, the *in silico* analyses identified genetic factors (LOC526226 and regulatory sites) that may be involved with horn development. The analyses were limited by data availability, sample size and tissue. Further analyses should consider these genes, interactions and expression data (e.g. including the ESTs). Lastly, regulatory sequences and chromatin interactions in the POLLED region should be further investigated in relevant tissues and developmental times to elucidate which are important for horns to develop.

Chapter 3:

Histological characterisation of the horn bud region in 58 day old bovine fetuses

3.1 Introduction

The bovine adult horn consists of a keratin outer sheath surrounding an inner core of bone. The horn is innervated by the corneal branch of the sensory trigeminal nerve (Buda *et al.* 2011). The embryonic origin of horn tissues is likely to be from the ectoderm (skin), neuroectoderm (bone and nerves) and possibly mesoderm (bone). The horn bud has been reported to be visible early in fetal development at 60 days of gestation (Evans & Sack 1973) and histological studies have been investigated horn bud histology from 70 days of gestation (Allais-Bonnet *et al.* 2013; Wiener *et al.* 2015; Schuster *et al.* 2020). The frontal skin is in the cranial region corresponding to the frontal bone and has been often used as a control in horn bud studies (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014; Wang *et al.* 2019c). These studies have shown that the epidermis at the horn bud is thicker than the epidermis of the frontal skin (Capitan *et al.* 2012; Allais-Bonnet *et al.* 2013; Wiener *et al.* 2015). From 115 days of gestation, thick nerve bundles are present beneath the horn bud, whereas only normal nerve fibres are found in the frontal skin (Wiener *et al.* 2015). A delay in hair follicle development in the horn bud region has been also observed in several studies (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014; Wiener *et al.* 2015; Schuster *et al.* 2020). No ossification is observed in the horn bud prior to birth. The structure of the horn bud earlier than 70 days of fetal development has not been described.

Very few genes are known to directly affect horn development especially as the bovine POLLED variants are intergenic. However, the *relaxin family peptide receptor 2* (*RXFP2*) gene has been associated with horn status and shape in sheep and with scurs in cattle (Kardos *et al.* 2015; Wiedemar & Drögemüller 2015; Lühken G *et al.* 2016; Pan *et al.* 2018; Wang & Gill 2021). A 1.8 kb insertion, originally thought to be located in the 3`-UTR of RXFP2, segregates with horn status in sheep breeds that are completely horned or completely polled (Wiedemar &

Drögemüller 2015; Lühken G *et al.* 2016). However, the 1.8 kb insertion does not segregate with polledness in breeds that have variable horn status (Lühken G *et al.* 2016) suggesting there are other genetic variants affecting horn status or there are additive genetic effects influencing horn status. Additionally, the insertion does not segregate in breeds where horn status is sex influenced (Lühken G *et al.* 2016). In more recent reference genomes, this 1.8 kb insertion is positioned just downstream of *RXFP2* within the gene LOC101110773 (Lühken G *et al.* 2016). Expression of *RXFP2* is higher in the fetal horn bud of cattle and sheep compared to the frontal skin or the horn bud region in polled embryos (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014; Wang *et al.* 2019c). *RXFP2* is also involved in antler development as it is expressed by antler stem cells collected from sika deer (Wang *et al.* 2019a). In musk deer species and Chinese water deer that do not have antlers, frameshift mutations have been found in *RXFP2* which may be responsible for the absence of antlers (Wang *et al.* 2019a; Wang *et al.* 2019c).

RXFP2 is a glycoprotein hormone receptor that binds insulin-like peptide-3 (INSL3) to release cAMP (Johnson *et al.* 2010; Petrie *et al.* 2015; Esteban-Lopez & Agoulnik 2020; Ivell *et al.* 2020). *RXFP2* expression has been observed in many tissues including the testis, brain and ovaries (Kumagai *et al.* 2002; Sedaghat *et al.* 2008; Ferlin *et al.* 2011; Pitia *et al.* 2015; Ferlin *et al.* 2017; Esteban-Lopez & Agoulnik 2020; Ivell *et al.* 2020). Therefore, it has a role in a variety of tissues. The role that RXFP2 may play in horn ontogenesis has not been investigated.

It is currently hypothesised that neural crest cells are responsible for horn ontogenesis (Wang *et al.* 2019c; Aldersey *et al.* 2020). After birth, the horn eventually connects to the underlying frontal bone which originates from neural crest cells (Ishii *et al.* 2015). *SRY-box transcription factor 10* (*SOX10*) and *neural growth factor receptor* (*NGFR*, also known as p75NGFR or p75NTR) have been used as markers to detect neural crest derived cells in the ovine horn buds at day 90 of development (Wang *et al.* 2019c; Rapizzi *et al.* 2020). Cells expressing SOX10 and NGFR were detected in the ovine fetal horn bud (Wang *et al.* 2019c).

SOX10 is a key regulator of neural crest cell development and maintains them in a stem cell state (Kim *et al.* 2003; Horikiri *et al.* 2017). *SOX10* is expressed in some cells derived from neural crest cells, including melanocytes, Schwann precursor cells, immature Schwann cells, pro-myelinating Schwann cells, myelinating Schwann cells and sensory ganglia of cranial nerves (Edgar *et al.* 2013; Table 3.1). NGFR is a member of the tumor necrosis receptor superfamily, and regulates neuronal processes such as neuron cell survival, degradation and apoptosis (Goncharuk *et al.* 2020). NGFR binds to neurotrophic factors, such as nerve growth factor (NGF), brain derived neurotrophic factor (BNDF), neurotrophin-3 (NT-3), and neurotrophin-4 (NT-4), and to neurotophin precursors (e.g. pro-ngf and pro-bdnf) (Arévalo & Wu 2006). *NGFR* is expressed in cells derived from the neural crest, namely Schwann precursor cells, immature Schwann cells and non-myelinating Schwann cells (Edgar *et al.* 2013; Table 3.1). However, in humans, *NGFR* is expressed in a wide range of tissues, including keratinocytes and mesenchymal cells at lower levels (Uhlén *et al.* 2015), and alone, may not be the best marker for detecting neural crest-derived cells. Schwann precursor cells and immature Schwann cells are the only cell types other than neural crest cells that co-express *SOX10* and *NGFR*.

Table 3.1: Expression of SOX10 and NGFR in neural crest cells and their derivatives in humans (Edgar et al. 2013).

Cell type	*SOX10*	*NGFR*
Neural crest cells	+	
Melanocytes	+	
Schwann precursor cells	+	+
Immature Schwann cells	+	+
Non-myelinating Schwann cells		x
Pro-myelinating Schwann cells	+	
Myelinating Schwann cells	x	
Sensory gangli of cranial nerves	x	

+ = gene is expressed in this cell type and is considered a marker, x = gene is expressed in this cell type but not considered a marker.

The structure of the cattle horn bud has not been characterized before ~70 days of gestation. The aim of this study was to 1) characterise the bovine horn bud structure at the earliest time that they are visible in the embryo, 58 days of development, by histomorphometric analysis, 2) determine the location of RXFP2 expression in the horn bud, and 3) investigate the lineage of horn bud cells using SOX10 and NGFR as markers. The hypothesis being tested was that there are differences between horned and polled fetuses in tissue structure in the region where the horn bud develops and that the neural crest cells are involved in the developmental origin of the horn bud.

3.2 Materials & Methods

3.2.1 Animals

In total, 12 horned and 12 polled Hereford heifers were used in this study, which was approved by the University of Adelaide Animal Ethics Committee (Project Approval No. S-2018-105).

3.2.2 Determination of tissue collection age

A preliminary trial was conducted to determine the best fetal age for collection and to test the dissection techniques. Heifers were genotyped by PCR for the Celtic polled variant (described in section 3.2.4). Six homozygous horned (pp) and six homozygous polled (PP) Hereford heifers from a mixed phenotype herd were synchronised. The heifers were synchronised by inserting an intrauterine progesterone controlled internal drug release (CIDR) device, for seven days. When the CIDR was inserted, the heifer also received 1 ml of gonadotropin releasing hormone (GnRH) (Ovurelin, Bayer Australia Ltd) via intramuscular injection (IM). When the CIDR was removed on day seven, the heifer received a 2 ml IM dose of prostaglandin F2α (Ovuprost, Bayer Australia Ltd). After 3 days, the heifers were

inseminated and given a second 1 ml IM dose of GnRH. The horned heifers were artificially inseminated with semen from a homozygous horned sire and polled heifers inseminated with semen from a homozygous Celtic polled sire. After ~ 30 days, the heifers were pregnancy tested using ultrasound, and tested again 1-2 days before surgery.

Two horned and three polled heifers were scanned pregnant via ultrasound. One horned and one polled fetus was surgically removed at 58 days, and one horned and two polled fetuses were surgically removed at 60 days by a veterinarian. The fetuses were collected via laparotomy carried out under local anaesthesia (distal paravertebral block with inverted L in the left flank). The fetuses were placed in sterile containers and immediately transported on ice to the laboratory, approximately 3 minutes away.

The dissection area in the laboratory was prepared by spraying the bench, containers and dissection equipment with RNAaseZAP (Sigma-Aldrich) and 70% ethanol. The fetuses were rinsed with phosphate buffered saline (PBS) and then dissected on a plate placed over ice. Samples were taken using 3 mm biopsies punches in the region of the horn bud (HB) and frontal skin (FS) (Figure 3.1). The samples were placed in Eppendorf tubes and immersed in formalin at room temperature for at least 48 hrs. The samples were then embedded, sectioned using a microtome (as described in section 3.2.5). The sections were stained with haematoxylin and eosin, and assessed for quality. The horn bud was visible at 58 days, which was selected as the time for collection of subsequent fetuses and processed as described below.
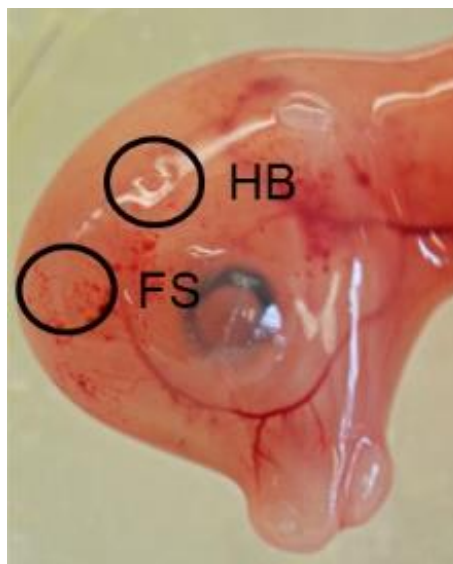
Figure 3.1: Horn bud (HB) and frontal skin (FS) are sampled for histology.

3.2.3 Generation of fetuses for analysis

The remaining 19 non-pregnant heifers were allocated into two groups for synchronisation. The first group included five horned and three polled heifers, while the second group included five horned and four polled heifers. Synchronisation groups were offset by two days to allow for all fetuses in each group to be surgically removed at 58 days and processed. The first group had five pregnant heifers (four horned, one polled) and the second group had six pregnant heifers (three horned, three polled). At 58 days of gestation, the fetuses were surgically removed (using the methods described in section 3.2.2), placed in sterile containers and immediately transported on ice to the laboratory.

The laboratory bench and equipment were prepared as above (Section 3.2.2). In the laboratory, the fetal heads were cut in half, rostral-caudally. One half of each head was preserved in formalin fixative solution. The other half of the head was preserved using RNAlater (as described in section 4.2.3). After fixation for more than 48 hours at room temperature, the samples were dehydrated in a graded ethanol series at 25% and 50% for 24 hours each and then 75%. Then, the brain tissue was removed from the cranial cavity and a 4

mm biopsy punch was used to sample the horn bud (HB) in horned animals, horn bud region (HBR) in polled animals and frontal skin (FS) region (Figure 3.1).

3.2.4 Genotyping for Celtic variant

All heifers and fetuses were genotyped and confirmed to be homozygous for wild-type sequence or Celtic POLLED sequence. DNA was extracted from tissue samples using the Qiagen DNeasy Blood and Tissue kit. Primers encompassing the Celtic variant location were used for PCR amplification (btHP-F1: 5'-GAAGGCGGCACTATCTTGATGGAA; btHP-R1: 5'-GGCAGAGATGTTGGTCTTGGGTGT). The PCR assay was conducted using the KAPA Taq ReadyMix PCR Kit (Kapa Biosystems, Inc.). Briefly, an initial melt was conducted (95°C, 3 m) before denaturation (95°C for 20 s), annealing (62°C for 20 s) and extension (72°C for 20s) for 34 cycles, followed by a 1 min extension at 72°C. Electrophoresis was conducted using a 1.5% agarose gel with GelRed (Fisher Biotec) at 62 V for ~2 hrs to separate the PCR products. A 389 bp product was amplified from the horned sequence and a 591 bp product was amplified from the Celtic polled sequence.

3.2.5 Tissue processing for histochemical and immunochemical analyses

Tissue processing was carried out by the Histology Services at the Adelaide Health and Medical Sciences (AHMS). The samples were dehydrated in ethanol and cleared with xylene, then embedded in paraffin with the tissue processor Tissue-Tek VIP 6 AI (Sakura). Samples were sectioned at 4 µm thickness using a microtome (RM2235, Leica) as this was the thickness used by Wiener *et al.* (2015). Four sections were placed per slide, and every other slide was stained with haematoxylin and eosin. The unstained slides were retained for immunohistochemistry. Samples collected as positive control (described in section 3.2.7.5) were sectioned at 5 µm, and had 2 sections per slide.

3.2.6 Haematoxylin and eosin staining

Haematoxylin and eosin (H&E) staining was carried out by Histology Services at the AHMS (Appendix Table B1) using the Dako CoverStainer (Agilent). Briefly, slides were dipped in xylene, absolute ethanol, and 70% ethanol before staining with H&E. The slides were then washed with absolute ethanol, xylene and histoclear. The slides were allowed to dry before a coverslip was applied.

3.2.7 Immunohistochemistry

3.2.7.1 Deparaffinization and rehydration

Deparaffinization was performed by incubating the slides in xylene. The slides were then rehydrated by serial washes in 100%, 95% and 75% ethanol solutions with the final wash in distilled water.

3.2.7.2 Antigen retrieval

Initially, a citrate buffer was used for antigen retrieval (Figure 3.2A). A 0.1 M citric acid solution was made by dissolving 10.5 g of citrate acid (Sigma, St Louis, USA) in 500 ml of RO water. A sodium citrate solution was made by dissolving 14.7 g of tri-sodium citrate (Ajax-Finechem) into 500 ml of RO water and 28.5 ml 0.1 M citric acid was combined with 124 ml 0.1 M sodium citrate. Then 250 ml of RO water was added, and using a pH meter (Lab-CHEM benchtop, TPS Ltd), the pH adjusted to 6.0. RO water was added to 300 ml. When the citrate buffer was used there was poor 3,3'-diaminobenzidine (DAB) staining of the cells. This may have been due to calcium ions from the tissues interfering with antigen retrieval (Morgan *et al.* 1994). When an EDTA buffer was used, stronger DAB staining was achieved (Figure 3.2B).

The EDTA buffer (0.5M) was pre-heated to boiling point (100°C) in a microwave (Samsung, model: ME6144W) for 3 min at ~1000 watts. The slides were microwaved in the heated EDTA for 10 min at power level of 30 (presumed to be 300 watts per the manual) (Figure

3.2B). Once cool, the slides were washed in 20mM Tris buffered saline with 0.1% sodium azide (TBS-azide) for 10 min.
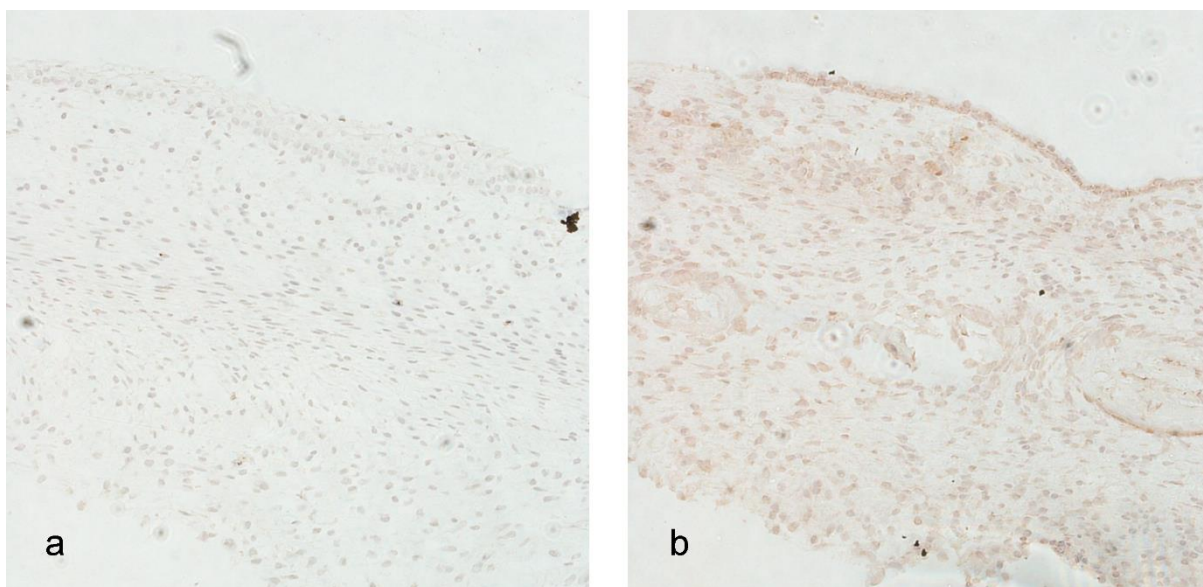
Figure 3.2: Staining with DAB using different buffers for antigen retrieval. (a) Poor staining was observed using citrate buffer. (b) Stronger staining was observed when EDTA buffer was used.

3.2.7.3 Block peroxidase activity

The slides were removed from the TBS-azide and the glass around the sections was carefully dried. The sections were covered with a few drops of 1% H_2O_2-50% methanol solution and incubated for 10 mins at room temperature, followed by a wash in TBS-azide for 10 mins.

3.2.7.4 Blocking non-specific binding sites

The glass around the sections was dried again and the sections were covered with a few drops of blocking solution (20% normal horse serum in TBS-azide) and incubated for 60 min

at room temperature. The samples were incubated within a humidity chamber which consisted of a container holding ultrapure water to prevent the sections from drying out.

3.2.7.5 Primary antibodies

The slides were dried around the sections and the primary antibody was applied at several dilutions specific for each antibody (Table 3.2). Antibody dilutions were prepared using antibody diluent solution (1% normal horse serum in TBS-azide). One section was incubated with only the antibody diluent solution (no antibody) as a negative control. The sections were incubated overnight (~ 20 hrs) at room temperature within the humidity chamber.

Table 3.2: Primary antibody dilutions used for immunohistochemistry.

Antibody	Dilution	Supplier (Product code)
Sox10	1:200	Santa Cruz Biotechnology, Inc. (sc-365692)
P75ngfr	1:300, 1:500	Santa Cruz Biotechnology, Inc. (sc-271708)
RXFP2	1:20	Santa Cruz Biotechnology, Inc. (sc-374293)

The primary antibodies were tested using tissues known to express the proteins. *SOX10* is expressed by oligodendrocytes, a type of glial cell, in the brain (Li *et al.* 2007). A local veterinarian provided formalin-fixed calf brain tissue sourced from recently deceased dystocia calves. The temporal lobe was sampled from the brain and used as a positive control for the SOX10 antibody.

Adult bovine testes were chosen as the positive control for the RXFP2 and NGFR antibodies. *RXFP2* is expressed in Leydig cells, seminiferous germ cells, spermocytes and spermatids (Pitia *et al.* 2017). *NGFR* has been shown to be expressed in stem Leydig cells and endothelial cells of blood vessels in adult human testes (Zhang *et al.* 2017; Eliveld *et al.* 2020). Adult bovine testes were sourced from a local abattoir and fixed in formalin.

3.2.7.6 Secondary antibody

After incubation with the primary antibody, the sections were washed in TBS-azide, dried, and incubated with a biotin goat anti-mouse IgG Cross-Adsorbed Secondary Antibody (ThermoFisher Scientific, #62-6540 diluted in antibody diluent solution) for 90 min at room temperature within the humidity chamber (Table 3.3). The sections were then washed in TBS-azide.

Table 3.3: Secondary antibody dilutions for immunohistochemistry.

Primary antibody	Secondary antibody dilution
SOX10	1:200
NGFR	1:200
RXFP2	1:100

3.2.7.7 Staining

The slides were dried again, and then incubated with avidin and biotin solution (ABC) for 60 min (VECTASTAIN ABC Kit, Vector Laboratories, Inc). The sections were washed in TBS. Slides were incubated for 3-5 min with a DAB substrate solution. The DAB substrate solution was prepared by dissolving a SIGMAFAST™ 3,3'-Diamino-benzidine tablet (SIGMA-ALDRICH Co.; product code: 1002771577) in 5 ml distilled water and 3.5 µl hydrogen peroxide. After incubation, the reaction was stopped by rinsing the slides in TBS-azide.

3.2.7.8 Counterstain

The slides were counterstained with haematoxylin (Mayer's Hematoxylin Solution, product number: MHS1, Sigma Aldrich) (Table 3.4). The slides were dried and coverslips applied.

Table 3.4: Counter-staining protocol for immunohistochemistry.

Solution	Minutes: Seconds
H2O	-
Haematoxylin	0:30
Running tap water	1:00
Acid alcohol	0:02
Running tap water	1:00
Absolute alcohol 1	0:10
Absolute alcohol 2	0:10
Absolute alcohol 3	0:10
Xylene 1	2:00
Xylene 2	2:00

3.2.8 Image analysis

3.2.8.1 Imaging

Slides were scanned using a NanoZoomer 2.0-HT slide scanner (Hamamatsu; model: C9600-01). The high-resolution scans (as .ndpi files) were viewed using NDP.view2 software (Hamamatsu). Images of H&E stained slides were exported at 20x magnification as JPEG images (300dpi).

Images of IHC stained slides were exported at 5x magnification as JPEG images (2000 dpi). Images were then manually processed in Photoshop (Adobe Creative Cloud 2018) to select the area of interest for machine learning. The epithelium and developing cranial bone were excluded from analysis. The .jpeg images were converted to .czias files using ZEN 3.4 (Zen lite; Carl Zeiss Microscopy GmbH).

Sections were grouped based on their genotype (horned or polled) and their location in the horn bud (outer or inner) using the depth of the epithelial cells as an indicator of position (Table 3.5). For the horn bud tissues, sections with a depth of 1-2 epithelial cell layers were considered to be on the outside of the horn bud (OuterHB, $n = 6$), while sections with a depth > 7 epithelium cells were considered to be at the centre of the horn bud (InnerHB, $n = 6$). Due to very small numbers of samples from the polled horn bud region at 58 days of development and from the polled FS at 60 days of development, these samples were grouped together (Polled HBR+FS, n=3). The polled horn bud and frontal skin sections were indistinguishable histologically.

Measurements were taken of the H&E sections using image processing software, FIJI (ImageJ) (Schindelin *et al.* 2012). The scale was set to measure images in micrometres (µm). Four tissue measurements were taken, in addition to epithelium cell depth, which were total depth, epithelial depth, mesenchyme depth and condensed cell depth (Table 3.5). The first section of four sections from every H&E stained slide was measured. Overall, this meant that every eighth section was measured, with ~28 µm between each section measured.

Table 3.5: Description of measurements obtained for H&E samples.

Measurement	Description
Epithelium Cell Depth (number)	Number of epithelium cells from the basal layer to outer layer of the epithelium at the thickest part of epithelium.
Total Depth (µm)	Measurement from the top of the epithelium to the bottom of the dense cell layer at the thickest part of epithelium.
Epithelium Depth (µm)	Measurement from the basement membrane to outer layer at the thickest part of epithelium.
Mesenchyme Depth (µm)	Measurement from the bottom of the basement membrane to the top of the dense cell layer at the thickest part of epithelium.
Condensed Cell Depth (µm)	Measurement from the bottom of the mesenchyme to the bottom of the condensed cell layer at the thickest part of epithelium.
Tissue Proportions (ratio)	Epithelium, mesenchyme and condense layer measurements are divided by the total depth to determine the proportion that these layers contribute to the tissue.

3.2.8.3 Image segmentation analysis

Image analysis was conducted on IHC stained sections. Images were segmented using a supervised machine learning approach in Intellesis (ZEISS) (Figure 3.3). Prior to training, the images were grouped based on staining strength (medium, light or dark). Images were removed from analysis if the section was damaged or if the stain was too light or too dark. For each sample, three stained sections per antibody were randomly selected ('analysis set'). Medium stained sections were prioritised over light and dark sections to maximise the number of sections that were analysed under the same machine learning model. Using the approach of Nesbit *et al.* (2021), six to ten sections were allocated to the 'training set' to create the image segmentation models. The model was then tested on the analysis set and the accuracy was scored as poor or good. Training and testing was repeated until the majority of sections had good segmentation. Some sections consistently did not fit the model, thus, separate models were trained to analyse these sections. That is, the section was used for both training and analysis.
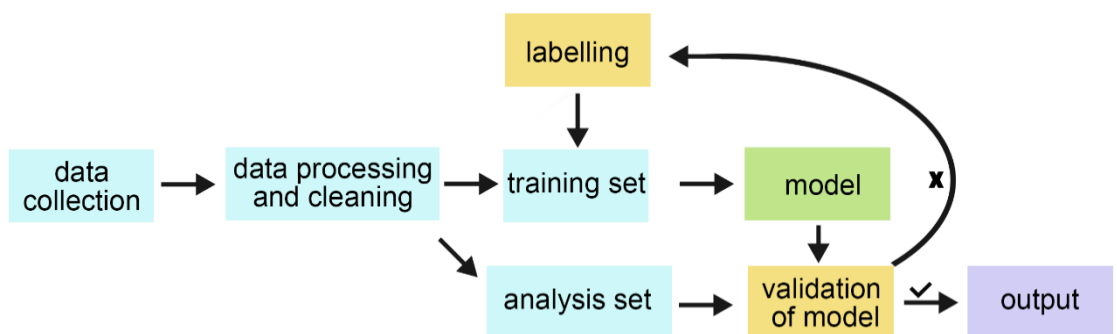
Figure 3.3: Flow chart of image segmentation via machine learning. Blue = input data, green = trained model, yellow = researcher input/decision making, purple = output data

To train the model, images were labelled for the following regions: "background", "negative nuclei" and "positively stained regions" (Figure 3.4B). For the anti-SOX10 labelled sections, the positively stained nuclei were detected as an additional region (Figure 3.4D). Small

88

areas were often mis-labelled and were removed from analyses. Minimum areas were set for "positively stained regions" at 200 pixels and "nuclei" at 200 pixels. For anti-SOX10 stained sections, minimum areas were set to 250 pixels for "positively stained nuclei", 300 pixels "positively stained regions" and 150 pixels for "negative nuclei". These cut-offs were used because differentiation between "positively stained nuclei" and "positively stained regions" was more difficult to achieve, and therefore, there was slightly more mis-labelling with this model.
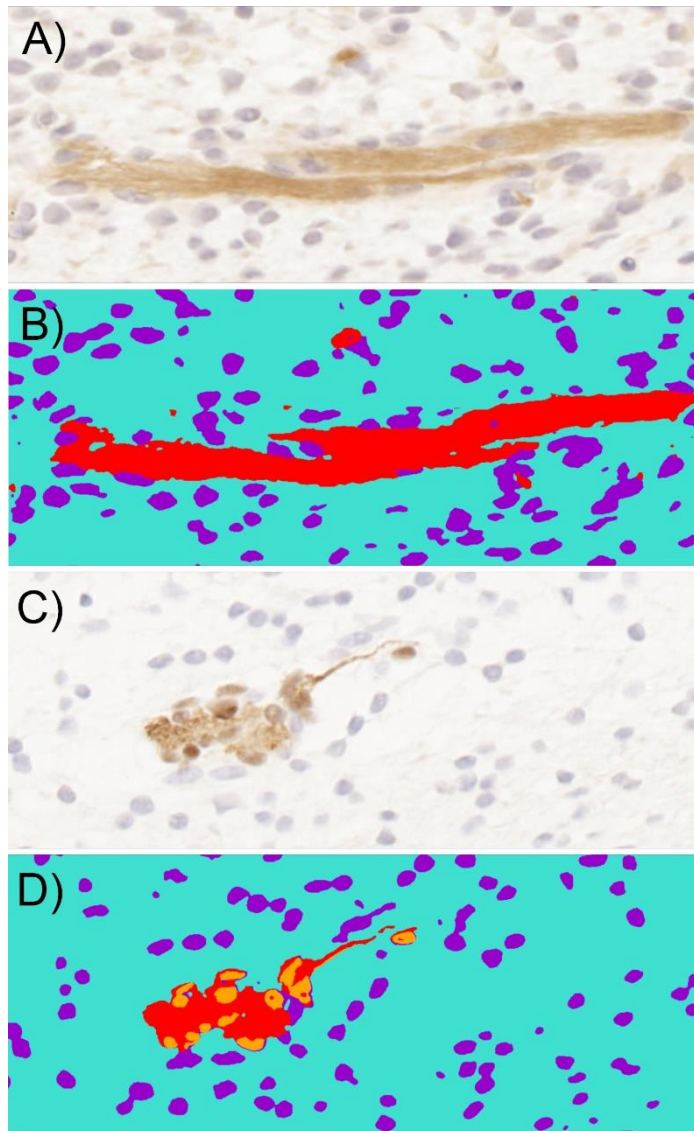
Figure 3.4: Machine learning was used to segment images for analysis. (A) Section stained with RXFP2 antibody. (C) Section stained with SOX10 antibody. (B, D) segmented images of A and C, respectively. Red = positively stained regions, orange = positively stained nuclei, purple = negative nuclei, cyan = background.

3.2.9 Statistical analysis

Statistical analysis was conducted in R (version 4.1.0). Measurements obtained from H&E stained slides were not normal (Appendix Table B3; Appendix Figure B3-Figure B6). Therefore, the paired Wilcoxon rank sum test were used to compare the horned measurements (InnerHB and OuterHB), and unpaired Wilcoxon rank sum tests were used to compare between

the horned and polled samples (code available at https://figshare.com/: <DOI: 10.25909/19971098>).

Measures collected from IHC sections were normally distributed according to the Q-Q normality plot and Shapiro-Wilks Normality Test (Appendix Table B4; Appendix Figure B9-Figure B18). Three sections were measured per sample to find a sample mean. The Fisher's exact test was used to compare between horned HB and polled HBR+FS (code available at https://figshare.com/: <DOI: 10.25909/19972088>). For all antibodies, the average area in pixels for Horned HB and Polled HBR were categorised, based on the image segmentation, as 'Positive' or 'Not positive'. 'Positive' represented the pixels that were assigned to positive staining for a given tissue, whereas 'Not positive' included the background and negatively labelled nuclei. For SOX10, additional comparisons were made between 'Positive nuclei' and 'Not positive nuclei', and 'Positive nerve and nuclei' and 'Not positive'.

3.3 Results

3.3.1 Confirmation of fetus genotype

In total, 16 fetuses were collected at 58 days of development (5 in the pilot study and 11 in the main experiment) (Table 3.6). Based on PCR genotyping, 15 fetuses aligned with their expected genotype, and one fetus was heterozygous for the Celtic variant (Figure 3.5). In total, there were eight homozygous horned, seven homozygous polled and one heterozygous fetuses. Samples were not collected and analysed from the heterozygous fetus.

Results for the initial polled genotyping assay for fetus #456 was unclear (Figure 3.5a), however, a second PCR test confirmed that the genotype was heterozygous (Figure 3.5b). It was noted that there is an additional band visible at ~ 500 bp. This may be due to non-specific binding of the primers. The heterozygous fetus, which was expected to be horned, did not have a visible horn bud. If one of the bulls was heterozygous for a POLLED allele, then half of the

horned fetuses would have been expected to be heterozygous. Therefore, it is suspected that the heifer was mis-genotyped or was accidently inseminated with the semen from the polled bull.
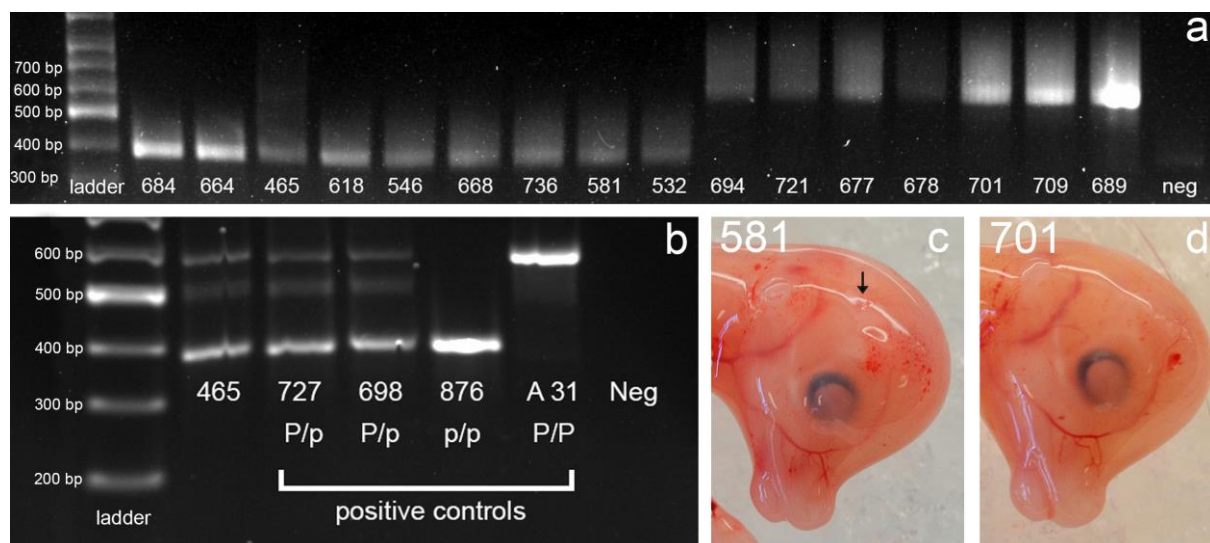
Figure 3.5: (a) PCR genotyping of fetuses collected. Eight were homozygous horned, seven were homozygous polled and one was heterozygous (465) for the Celtic variant. (b). Heterozygosity of #465 was confirmed by a second PCR using the same primers (c) Horned fetus (581) with an arrow indicating the horn bud. (d) Polled fetus (701) with no visible horn bud.

3.3.2 Tissue integrity

The protocol from Allais-Bonnet *et al.* (2013) used 3 mm biopsy punches to sample small areas of tissue from specific regions on the head for histological analyses and this protocol was initially trialled in a pilot study. The protocol was used for fetuses #678 and #664 at 58 days of development (Table 3.6) to determine the best age to sample. However, it was not possible to take 3 mm biopsy punches from unfixed samples with this approach as the tissue was extremely fragile.

A second protocol tested in the pilot study was to biopsy the fetal head after fixing in formalin and this method was performed on fetuses #684, #694 and #721 (Table 3.6). The fixed tissue was slightly more durable, and 3 mm punches were taken and processed. However, when

the protocol was applied to fetus #532, it was difficult to sample the frontal skin. Although formalin fixed samples gave better results than unfixed sampled, some fixed samples were not good enough for analysis. This protocol may have been successful on fetuses #684, #694 and #721 because the fetuses were collected at 60 days compared to fetus #532 collected at 58 days.

Therefore, the method used for the remainder of the fetuses in the main study included a tissue dehydration step and used a 4 mm biopsy punch (described in section 3.2.3). Six horn buds and five frontal skin collected from horned fetuses, and five horn bud regions and five frontal skins from polled fetuses were processed for histological analyses using this method (Figure 3.6).

Table 3.6: Metadata of fetuses collected for analysis and quality of histology samples.

Fetus ID	Trial	Age (days)	Genotype Horned (pp), polled (PP), or heterozygous (Pp)	Stain		Quality	
				H&E	IHC	HB/HBR	FS
678	Pilot	58	Polled	Y	N	1	1
664	Pilot	58	Horned	Y	N	1	1
684	Pilot	60	Horned	N	N	N/A	N/A
694	Pilot	60	Polled	Y	Y	1	3
721	Pilot	60	Polled	N	N	N/A	N/A
465	Main	58	Heterozygous	N	N	N/A	N/A
618	Main	58	Horned	Y	Y	2	2
546	Main	58	Horned	Y	Y	3	1
736	Main	58	Horned	Y	Y	3	1
532	Main	58	Horned	Y	Y	3	N/A
668	Main	58	Horned	Y	Y	3	1
581	Main	58	Horned	Y	Y	3	1
667	Main	58	Polled	Y	Y	3	1
701	Main	58	Polled	Y	Y	1	1
689	Main	58	Polled	Y	Y	1	1
709	Main	58	Polled	Y	Y	2	1

Y = yes; N = no; 1 = poor quality; 2 = moderate quality; 3 = good quality; N/A = not available; H&E = haematoxylin and eosin; IHC = immunohistochemistry; HB = horned bud; HBR = horned bud region, FS = frontal skin

The frontal skin and polled samples were more fragile than the horn bud samples from the horned fetuses. The fragility resulted in damaged sections during processing. The samples that had structural damage to the tissue, such as missing epidermis, could not be used for accurate measurements (Table 3.6). Only samples with good or moderate quality could be measured. All six horn bud samples from horned fetuses were intact, but only one frontal skin sample from the horned fetuses had moderate quality (Table 3.6). Of the polled samples, two polled horn bud samples and one polled frontal skin were undamaged (Table 3.6).
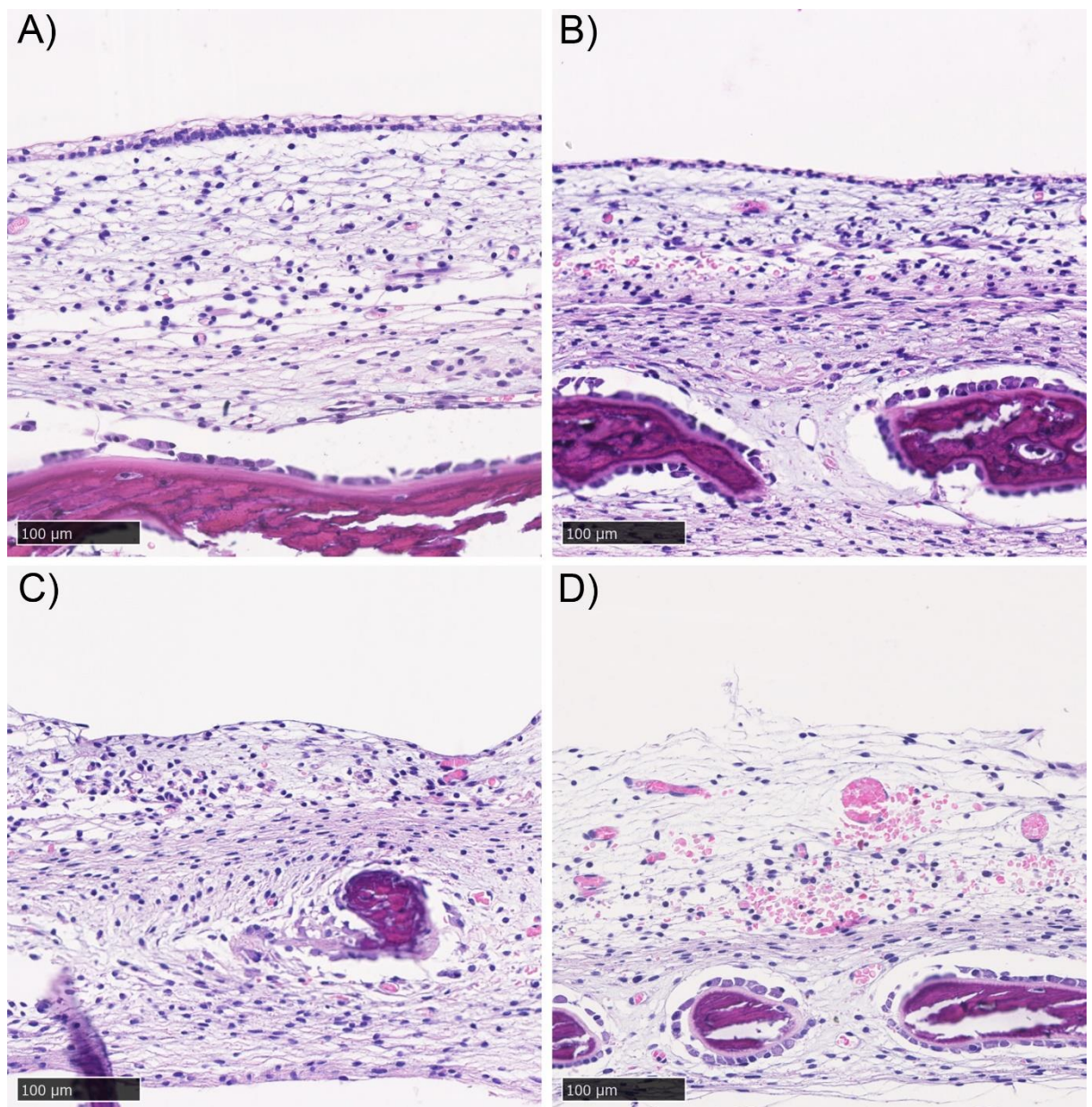

Figure 3.6: Histological example slides from polled fetuses. Good quality samples from fetus #694 FS (polled) (A) and fetus #667 HBR (B). Poor quality samples from sample fetus #701

HBR (C) and fetus #689 FS (D). HBR = horn bud region. FS = frontal skin. Magnification = 20x.

3.3.3 Macroscopic appearance and position of the horn bud

The horn bud was macroscopically visible as a slightly depressed ring of skin, the centre of which was raised, on the heads of the day 58 horned fetuses, but was absent in the polled fetuses. The position of the horn bud was at the right angle of perpendicular lines drawn from the eye and ear bud (Figure 3.7). The skin was smooth in polled fetuses at the same position. The positioning of the horn bud in horned fetuses was used as a guide for sampling the horn bud region of the polled fetuses.

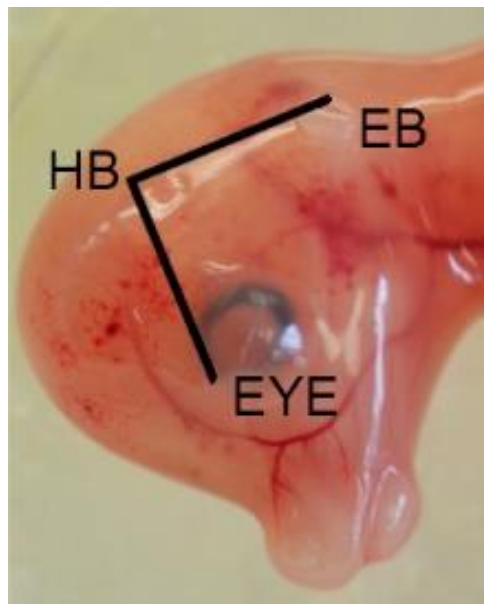Figure 3.7: Position of the horn bud (HB) in relation to the eye (EYE) and ear bud (EB). The horn bud was positioned at the right angle of perpendicular lines drawn from the eye and ear bud.

3.3.4 Microscopic tissue structure of the horn bud

The horn bud at 58 days in the horned fetuses was characterised by a thickened epidermis of up to 10 cells in depth, while surrounding tissue and frontal skin only were only

1-2 cells thick. Five tissue layers were observed in the horn bud sections of the horned fetuses (Figure 3.8):

1. *Epidermis*: The outer epithelial layer was comprised of keratinocytes and the depth of the layer increased towards the centre of the horn bud. The epidermis was separated from the mesenchyme by the basal membrane.

2. *Mesenchyme*: In the mesenchyme layer, loosely packed cells with round nuclei were observed, and staining with eosin, which is an acidic dye and binds to basic proteins such as those in the cytoplasm and connective tissue fibres (Bancroft & Layton 2019), indicated that there were fewer basic proteins than in the condensed cell layer.

3. *Condensed cell layer*: Densely packed cells with oval nuclei, with greater eosin staining than the mesenchyme making the condensed cell layer distinct from the mesenchyme layer (Figure 3.9). The condensed cells are possibly horn progenitor cells.

4. *Ephrin cell layer*: A layer of 1-3 cells with oval/elongated nuclei demarcated the edge of the condensed cell layer. These cells are possibly Ephrin/eph expressing cells, which are involved with segmentation of tissues in embryo development. The Ephrins/eph system directs cell migration by providing short-range repulsive cues to migrating cells (Klein 2004). This layer creates a cellular barrier between the mesenchyme and osteogenic cells in the cranium to ensure cells migrate to appropriate regions (Ting *et al.* 2009; Ishii *et al.* 2015). The ephrin cell layer also repels developing nerves to guide them towards the correct destinations. In the present study, developing bones and nerves were not observed to cross between the mesenchyme and developing skull vault, supporting the hypothesis that these cells are Ephrin/eph expressing cells.

5. *Developing skull vault:* Ossification was observed in the developing skull vault layer as samples were taken over the frontal bone. Frontal bone develops through intra-membranous ossification. The ossifying tissue was observed in both horned and polled fetuses. Some osteoid was present in this layer of all the sections.

96
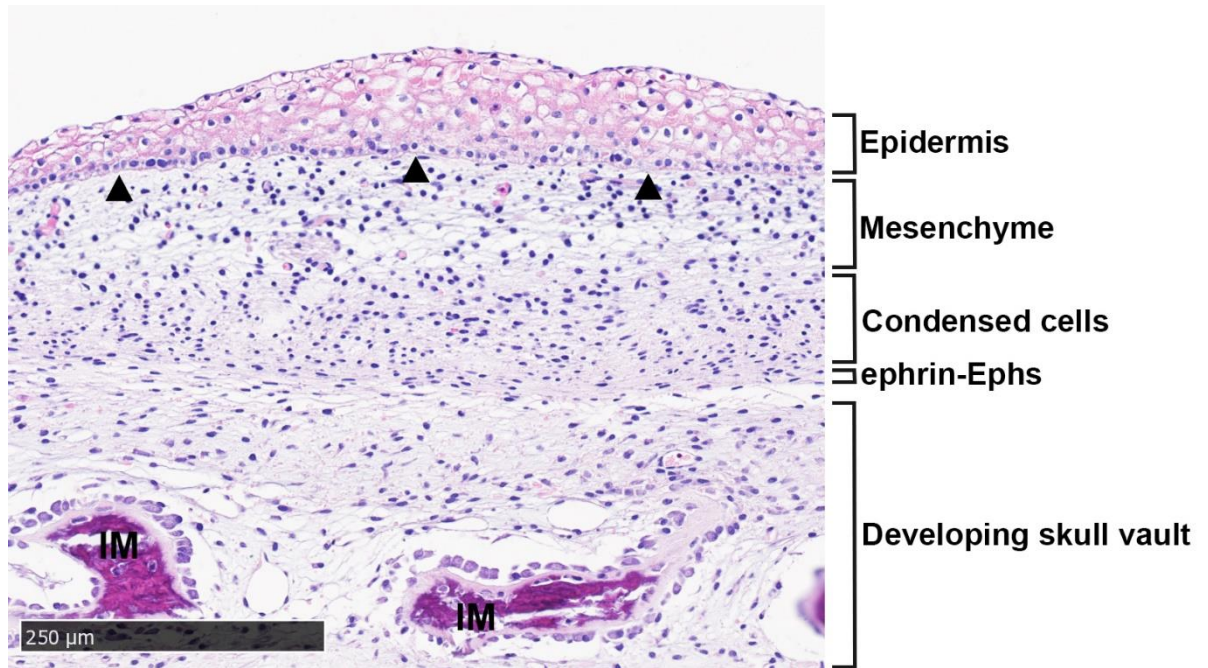
Figure 3.8: Tissue layers in the horn bud of horned fetus #546 at 58 days of development. The epidermis and mesenchyme were separated by the basal membrane (black arrowhead). Cells presumed to be expressing ephrin and ephrin receptors (Ephs) form a barrier between the mesenchyme and developing skull vault (Ishii *et al.* 2015). H&E stained tissue section at magnification = 10x.
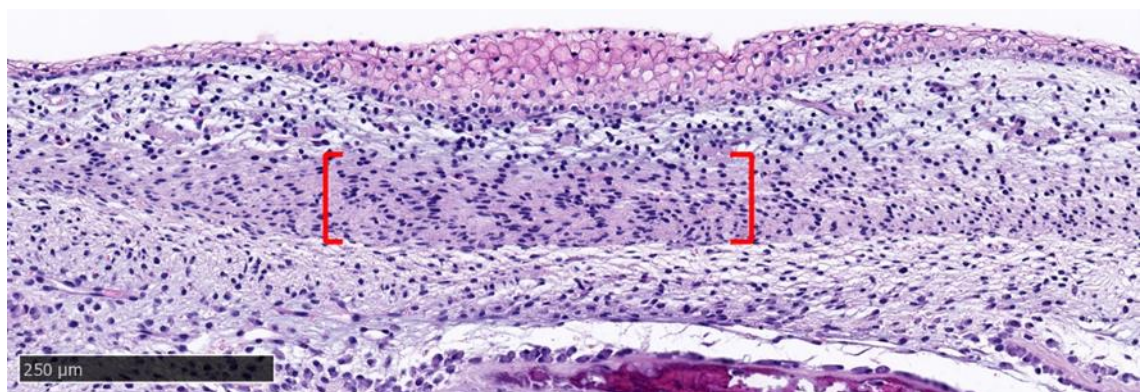
Figure 3.9: Horn bud section from horned fetus #668 stained with haematoxylin and eosin (H&E). Condensed cells were located below the thickened epithelium of the horn bud (red bracket). Magnification = 10x.

3.3.5 Microscopic tissue structure outside the horned horn bud

The structure of tissue outside the horn bud had an epithelial cell depth of 1-2 cells. The same tissue layers were present as within the horn bud, but the condensed cell depth was less apparent or absent. The tissues appeared similar between the outer horn bud (Figure 3.10A) and horn frontal skin (Figure 3.10B).
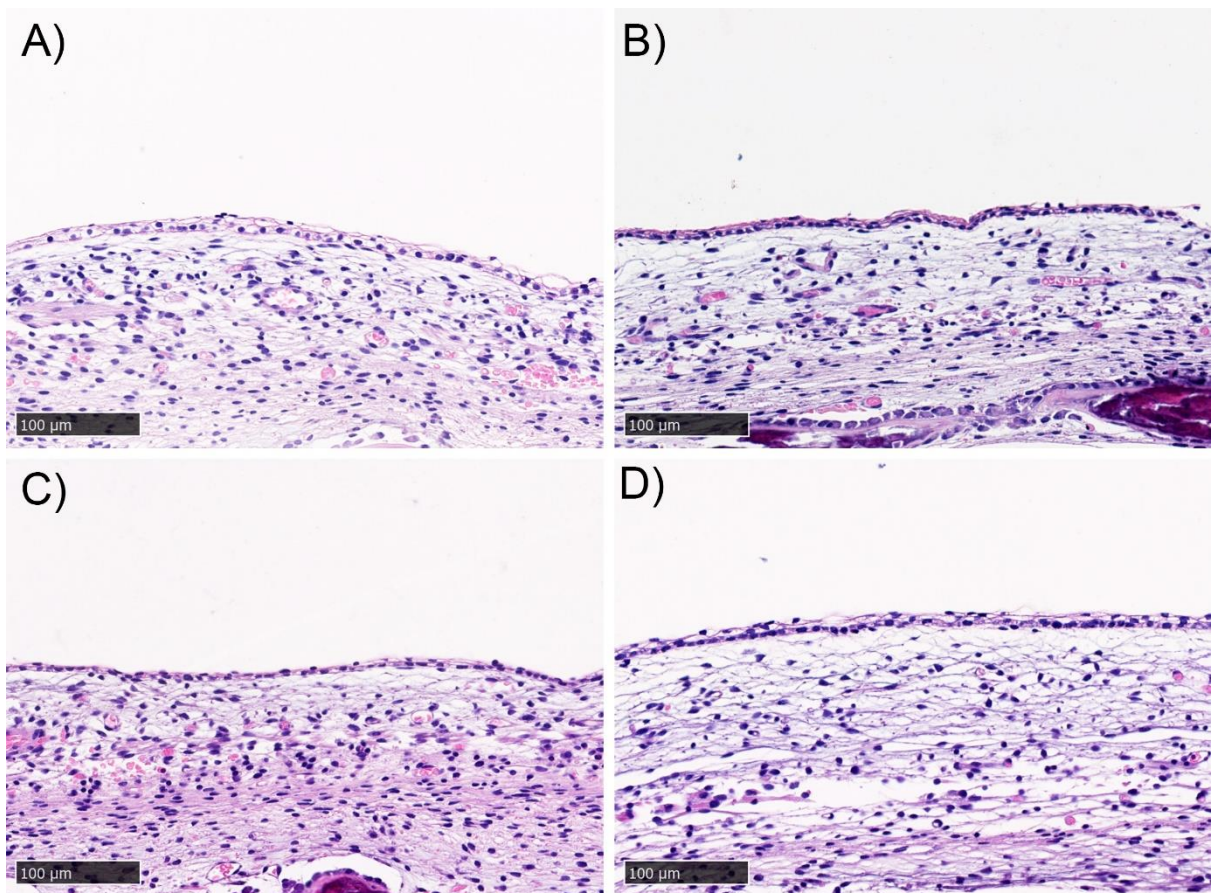
Figure 3.10: Histological similarities between A) the outer horn bud (fetus #532), B) horned frontal skin (fetus #618), C) polled horn bud (fetus #667) and D) polled frontal skin (fetus #694) stained with H&E. Magnification = 20x.

3.3.6 Microscopic tissue structure horn bud region and frontal skin of polled

The structure of tissue from the polled fetuses had an epithelial cell depth of 1-2 cells, irrespective of site (horn bud region or frontal skin) (Figure 3.10C -D). The same tissue layers observed horn bud samples were present in the polled samples, but like the horned frontal skin, the condensed cell layer had visibly reduced depth in most sections (2-3 cells thick) or was completely absent in other sections (Figure 3.10). There were no obvious differences between the polled samples and the horned samples from regions outside of the horn bud.

3.3.7 Comparison of the inner horn bud, outer horn bud and polled

To explore the differences between tissues, measurements of the tissue layers were taken from H&E sections (described in section 3.2.8.2; data available at https://figshare.com/: <DOI: 10.25909/21753890>). The measurements that were taken from sections with the same epithelium cell depth were grouped together and averaged (Table 3.7). For horned samples, an epithelium cell depth of 1-2 cells in the horn bud region were classified as outer horn bud (OuterHB) and epithelium cell depth > 7 cells was defined as the centre of the horn bud and classified as inner horn bud (InnerHB). The only polled sections that could be measured were from two horn bud regions and one frontal skin. (Note: the polled FS sample was collected at day 60 of development.) There was no discernible difference between the polled horn bud region and the frontal skin samples. Therefore, these sections were grouped together for measurement comparisons (Polled HBR+FS = polled horn bud region + polled frontal skin).

Measurements from the InnerHB, OuterHB and PolledHBR+FS were compared (Table 3.8; Appendix Figure B7). The measurements from the InnerHB were highly significantly different from the OuterHB and from the PolledHBR+FS samples (Figure 3.11), indicating that the tissue at these sites were structurally different. Overall, the InnerHB had a greater total depth, epithelium depth, and condensed cell depth, while the mesenchyme depth was greater in the OuterHB and PolledHBR+FS. There was a correlation between epithelium depth and total

depth (Appendix Figure B8). The increase in epithelium depth most likely contributed to a greater total depth in the InnerHB sections.

Table 3.7: Number of measurements collected from each sample.

Genotype	Fetus	Number of sections measured per sample	Number of sections [a]	
			1-2 cells	7-10 cells
Horned	532HB	39	10	5
	668HB	31	2	4
	546HB	53	11	7
	581HB	21	7	1
	736HB	33	0	3
	618HB	43	8	14
	Total	**220**	**38**	**34**
Polled	667HBR	32	29	0
	709HBR	26	13	0
	694FS	20	2	0
	Total	**78**	**44**	**0**

[a] Data from horned fetuses were grouped by epithelium cell depth to indicate position in the horned horn bud: 1-2 cell depth (OuterHB) and 7-10 cell depth (InnerHB). Polled samples had an epithelial cell depth between 1-2 cells. HB = horn bud; HBR = horn bud region; FS = frontal skin

Table 3.8: Descriptive statistics of measurements obtained from H&E horn bud (HB) and frontal skin (FS) samples. Median and SD of tissue depths are determined for the Inner Horned HB (InnerHB), Outer Horned HB (OuterHB) and Polled (PolledHB+FS) fetal samples.

	n	Total (µm)		Epithelium (µm)		Mesenchyme (µm)		Condensed cell (µm)	
		Median	SD	Median	SD	Median	SD	Median	SD
InnerHB	6	228.2	42.5	83.3	19.5	71.2	20.1	65.3	16.9
OuterHB	6	151.3	22.2	11.9	4.4	114.1	20.6	14.5	9.8
PolledHB+FS	3	163.0	41.5	9.0	3.1	141.8	47.4	12.7	10.1

n = number of samples
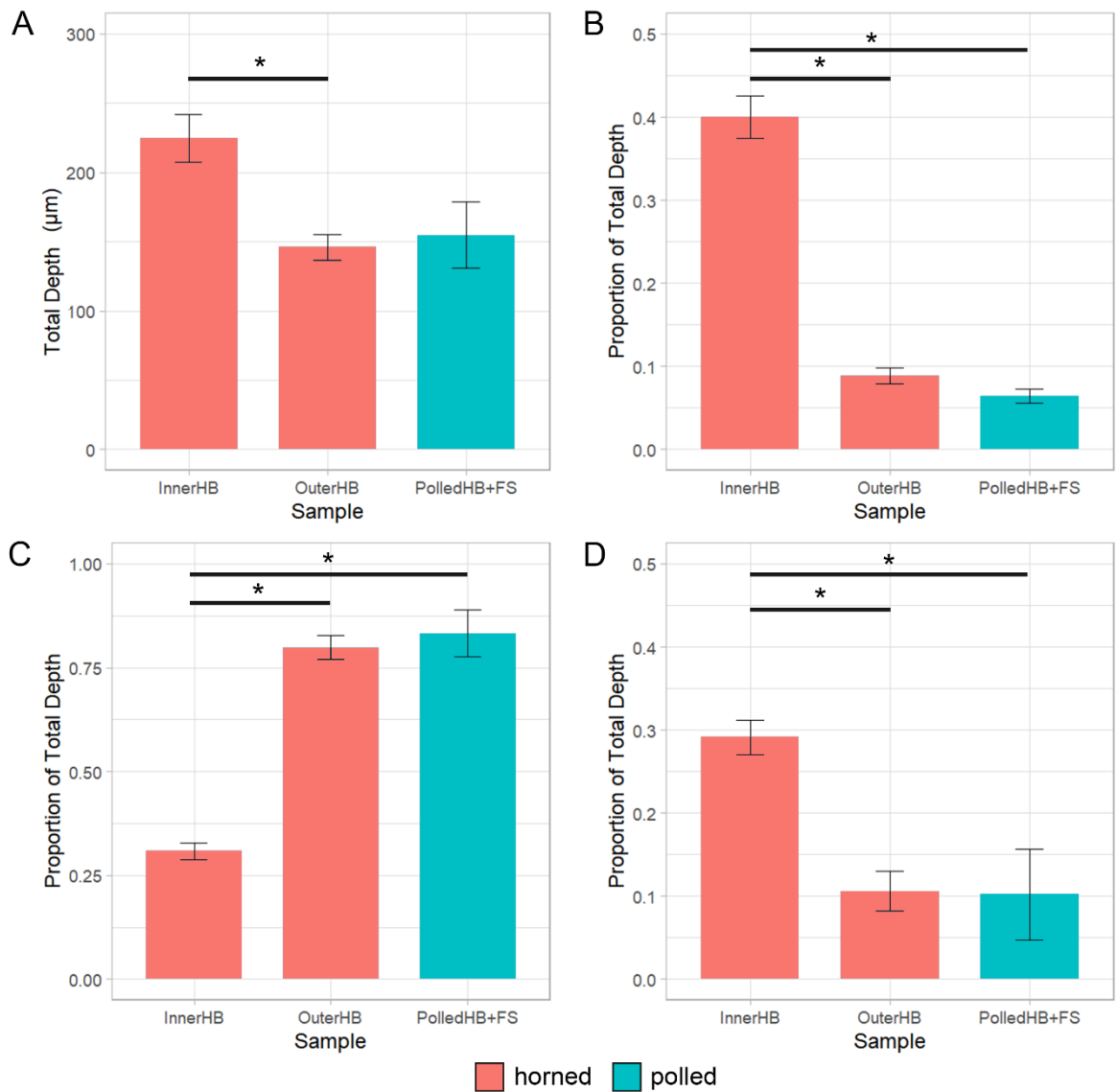
Figure 3.11: The total depth (A), epithelium depth (B), mesenchyme depth (C) and condensed cell depth (D) were compared between the inner horn bud (InnerHB), outer horn bud (OuterHB) and tissue from the polled horn bud region and frontal skin (PolledHB+FS). Horned fetuses = red, polled fetuses = blue. Epithelium depth, mesenchyme depth and condensed cell depth reported as proportions of the total depth. Columns and error bars represent the mean proportion and standard error. Paired and unpaired Wilcoxon rank sum tests were conducted to compare medians between groups. * Significant at $p < 0.05$.

3.3.8 Neural crest cells within the horn bud

3.3.8.1 Specificity of SOX10 and NGFR antibodies

The SOX10 antibody was validated by staining calf brain tissue from the temporal lobe where glial cells express *SOX10*. Brown staining was observed in the glial cells (Figure 3.12A), indicating that the antibody is most likely binding to the SOX10 antigen. Staining was absent in the negative control, which was not exposed to the primary antibody (Figure 3.12B).
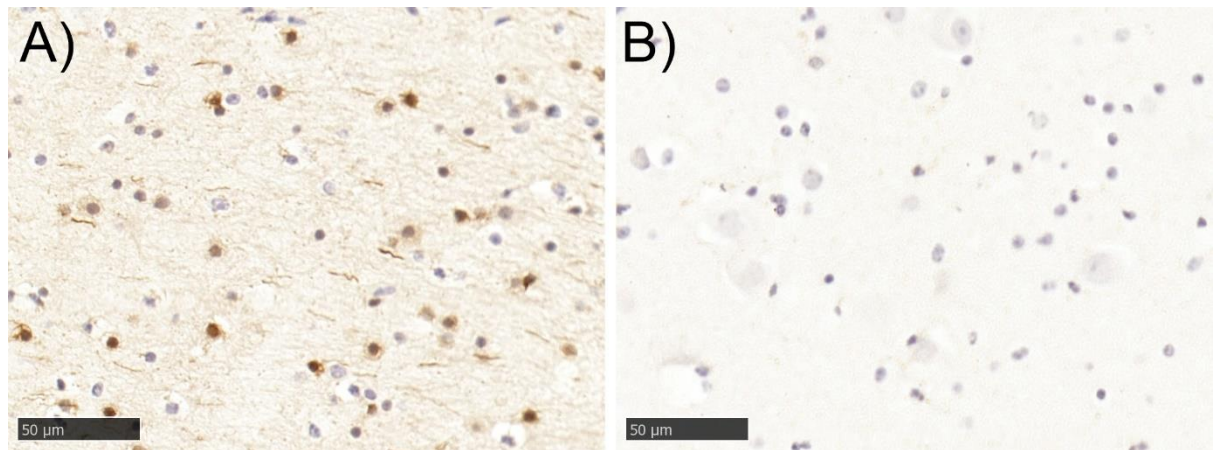
Figure 3.12: Glial cells stained with SOX10 antibody and counterstained with haematoxylin. A) Positive control: Brown stained glial cells indicating the antigen, SOX10, is present in these cells. B) Negative control: Glial cells unstained when only the secondary antibody is used. Magnification = 80x.

The NGFR antibody was validated by the detection of positive staining of spermatogonia and spermatocytes in bovine testis (Figure 3.13). *NGFR* has been shown to be expressed in various testicular cells from different species. In humans, expression has been observed in endothelial cells of testicular blood vessels (Eliveld et al. 2020). The Human Protein Atlas (date accessed: Dec 2021) reported that NGFR is expressed at high levels in the peritubular cells and at low levels in spermatogonia (Uhlén et al. 2015). In rats, *NGFR* expression is localised to the spermatogonia (Levanti et al. 2006). In wild ground squirrel, *NGFR* is expressed in spermatagonia and spermatocytes during breeding season (Zhang et al.

2015). The precise localisation of *NGFR* expression in bovine testis has not been determined although expression of NGFR in testes has been reported in the cattle Genotype-Tissue Expression atlas (Liu *et al.* 2020b) (https://cgtex.roslin.ed.ac.uk [accessed July 2021]).
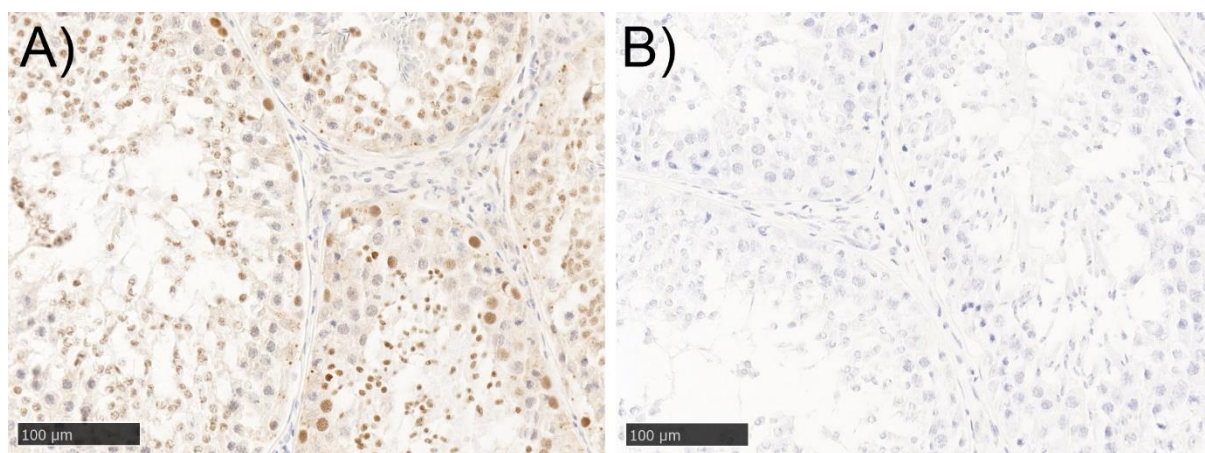
Figure 3.13: Testis stained with NGFR antibody and counterstained with haematoxylin. A) Positive control: Brown stained spermatogonia and spermocytes, indicating the antigen, NGFR, is present in these cells. B) Negative control: Cells unstained when not exposed to the primary antibody. Magnification = 20x.

3.3.8.1 SOX10 and NGFR detection in the horn bud of horned fetuses

Horned and polled samples were stained with antibodies for neural crest markers, SOX10 and NGFR, to determine if the horn progenitor cells in the condensed cell layer were derived from neural crest cells. Neither SOX10 nor NGFR were detected in the predicted horn bud progenitor cells in the condensed layer (Figure 3.14H, M). Instead, SOX10 and NGFR expression was detected in developing peripheral nerves in the horn bud of horned fetuses (Figure 3.14I; Figure 3.15A-B). SOX10 staining was observed in glial cells associated with peripheral nerves. Growth of peripheral nerves into the epidermis was observed for five out of six horned HB samples (Figure 3.15C-D). No other cells expressed SOX10. In contrast to SOX10, NGFR was widely expressed in the horn bud. In addition to expression in peripheral nerves (Figure 3.14N), NGFR was expressed in cells in the epidermis, mesenchyme and developing cranial vault (Figure 3.15K-L, O).
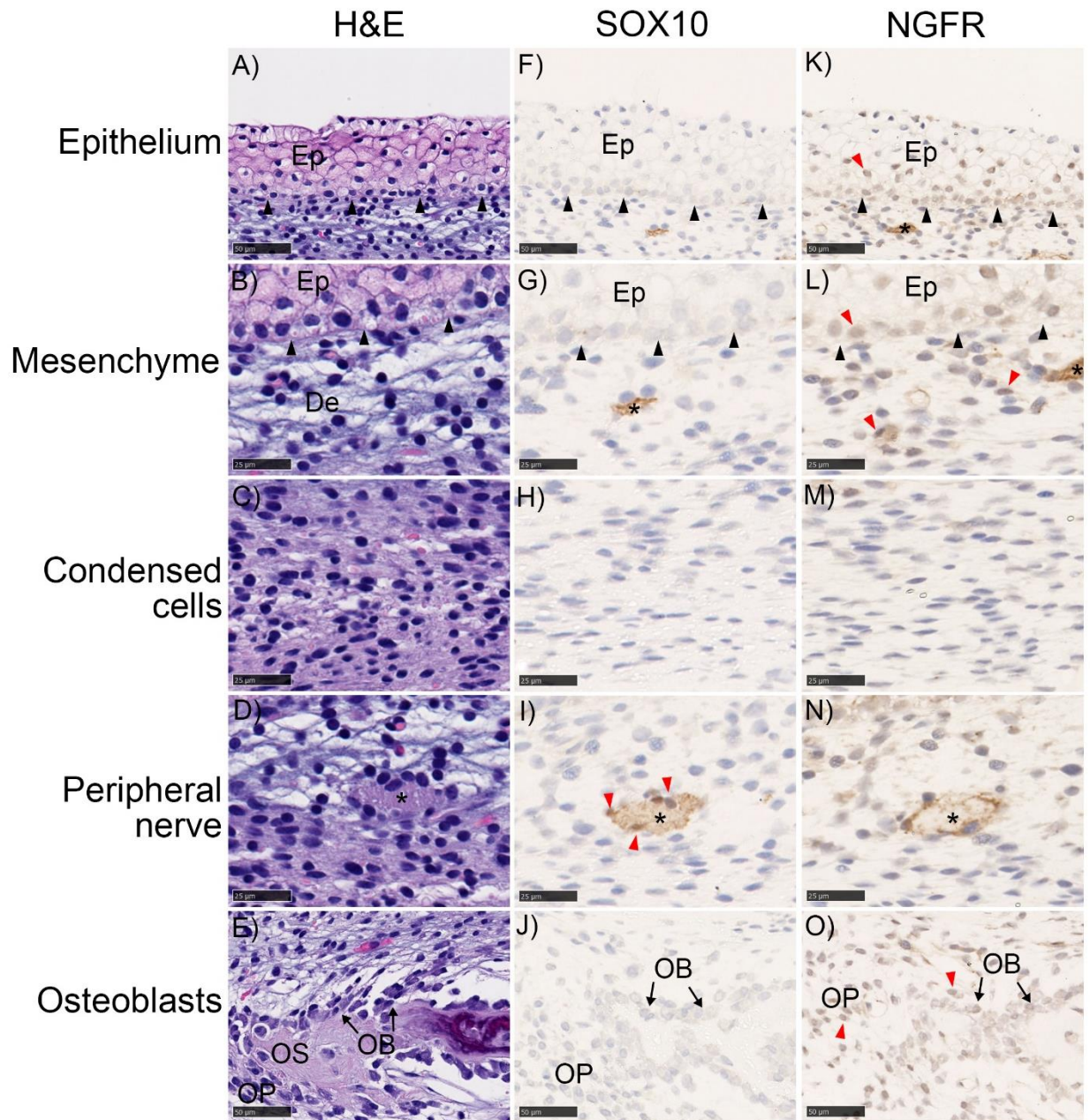
Figure 3.14: SOX10 and NGFR antibody staining of cells in the horn bud of horned fetuses. The horn bud was stained with haematoxylin and eosin (A-E), SOX10 antibody (F-J) and NGFR antibody (K-O). Positively staining nuclei observed for SOX10 (glial cells) and NGFR (epithelium, mesenchyme and osteoblasts) are indicated with red arrows. The basal epithelium is indicated with black arrows. The nerve tissue is indicated with an asterisk (*). Ep = epithelium, OB = osteoblast, OP = osteoprogenitor. Magnification = 40x (A, F, K, E, J and O) and 80x (B-D, G-I, L-N).
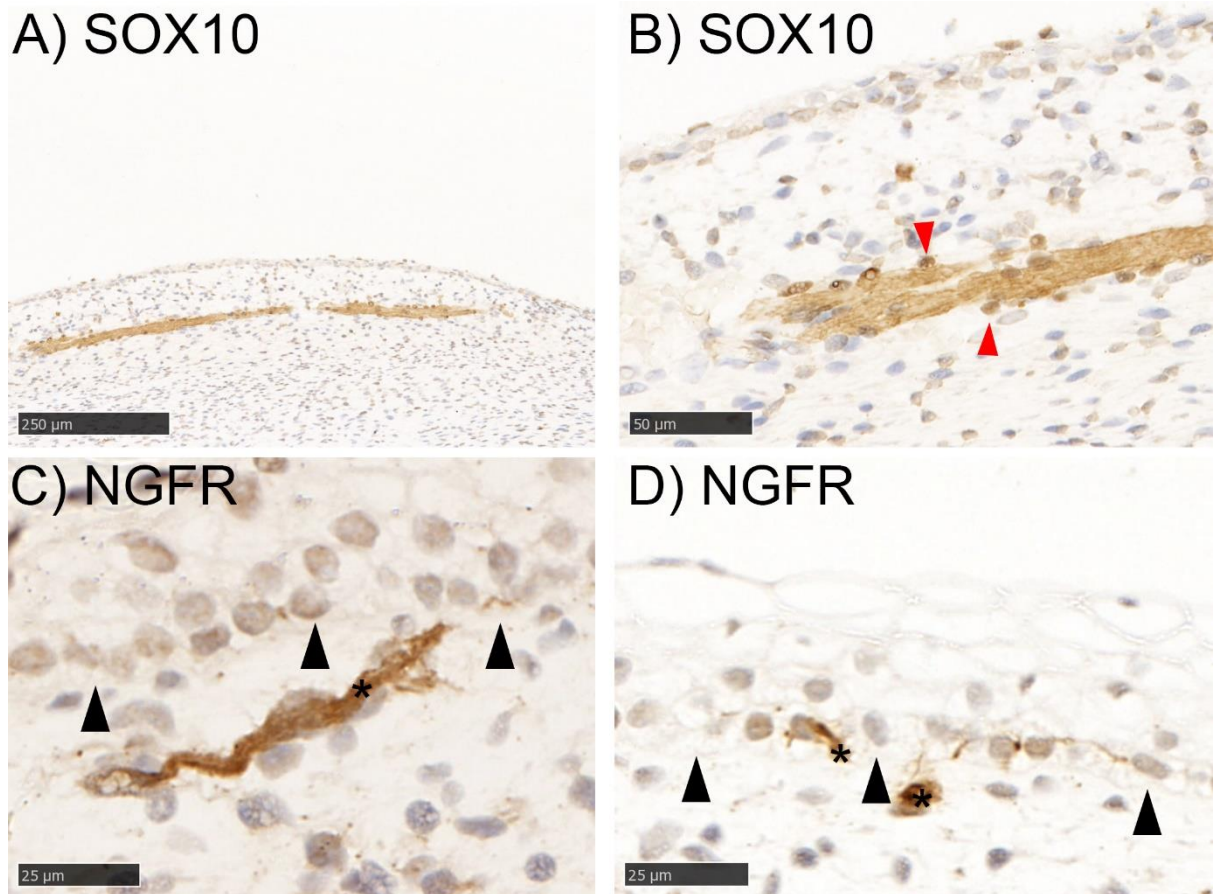
Figure 3.15: Staining of nerves with SOX10 and NGFR antibodies and counterstaining with haematoxylin. (A-B) SOX10 was detected in cells and structures presumed to be Schwann cell precursors and peripheral nerves in the fetal horn bud. (C-D) Nerve growth into the epidermis was observed in the horn bud of horned fetuses. Red arrow = positively stained nuclei; black arrow = basal membrane; * = nerve. Magnification = 10x (A), 40x (B), and 80x (C-D).

3.3.8.3 SOX10 and NGFR detection in polled tissues

The SOX10 and NGFR antibody staining of the polled tissues reflected the same pattern observed in the horn bud of the horned fetuses (Figure 3.16). SOX10 was detected in the glial cells and developing bundles of nerves, although the nerves appeared generally smaller than those observed in the horn bud (Figure 3.16C). Individual cells expressing SOX10 and unassociated with nerves were apparent in some sections (Figure 3.16A). This may have occurred from the section including nuclei but not the whole nerve. NGFR was detected in the developing nerves, mesenchymal cells epithelium cells and developing cranial vault (Figure

3.16B-D). No nerves appeared to be within the epithelium as was observed in the horn bud. The polled frontal skin sample was similar to the polled horn bud region samples.
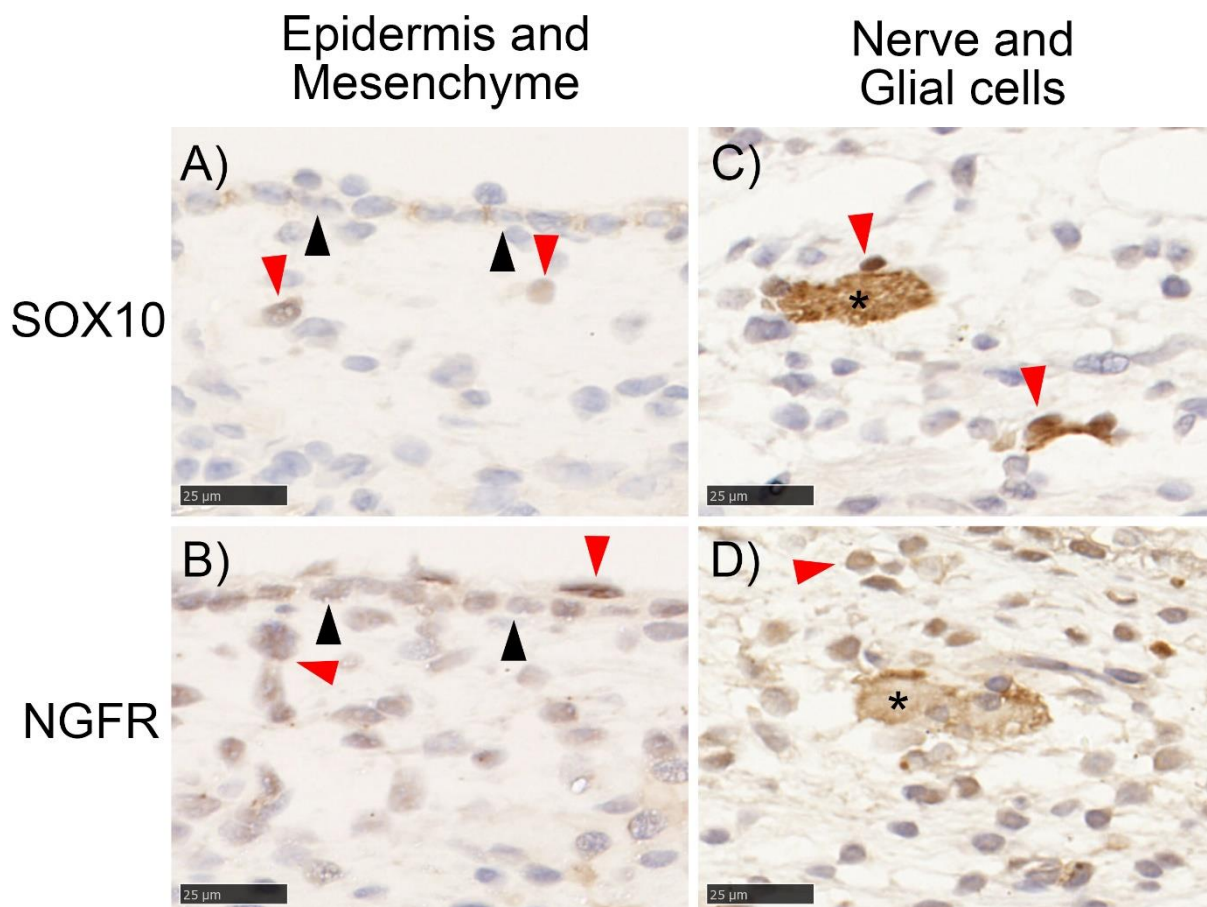
Figure 3.16: SOX10 (A, C) and NGFR (B, D) antibody staining and haematoxylin counterstaining of the horn bud region from polled fetuses. Epidermis and mesenchyme stained with SOX10 (A) and with NGFR (B). Nerve and glial cells stained with SOX10 (C) and NGFR (D). Red arrows indicate positively stained nuclei. Black arrows indicate the basal membrane. * = positively stained peripheral nerves. Magnification = 80x.

3.3.8.4 Nerves are larger in the horn bud of horned fetuses compared to polled tissues

To test whether there was more nerve tissue in the horn bud, machine learning was used to segment the section images into nerve, nuclei and background, and then to measure the areas (pixels). Additionally, positive staining nuclei were detected in the SOX10 stained sections.

Nerves identified by this positive staining were compared between horned and polled tissues (Appendix Figure B19-Figure B20).

There were 3.15 times (odds ratio) more nerves in the horn bud from horned fetuses compared to polled tissues (HB+FS) (Figure 3.17C; Fisher's Exact Test; $p < 2.2e\text{-}16$, confidence interval [CI] = 3.12 -3.18). There were 1.67 times more pixels assigned to positively stained cells for SOX10 in the horn bud of horned fetuses than in polled HB and FS (Figure 3.17B; Fisher's Exact Test; $p < 2.2e\text{-}16$, CI = 1.635-1.71). When the area of positive cells and nerves were combined, there was 1.67 times more positive staining in horn bud tissues compared to polled tissues (Figure 3.17A; Fisher's Exact Test: $p < 2.2e\text{-}16$, CI = 1.64-1.71). These results show that there is significantly greater innervation developing at the horn bud of horned fetuses compared to horn bud regions and frontal skin of polled fetuses.

The nerve tissue was also measured in NGFR stained sections. In the NGFR stained sections, there were 1.37 times more nerves in the horn bud of horned fetuses compared to polled HB+FS (Figure 3.18; Fisher's Exact test: $p < 2.2e\text{-}16$, CI = 1.36-1.38).
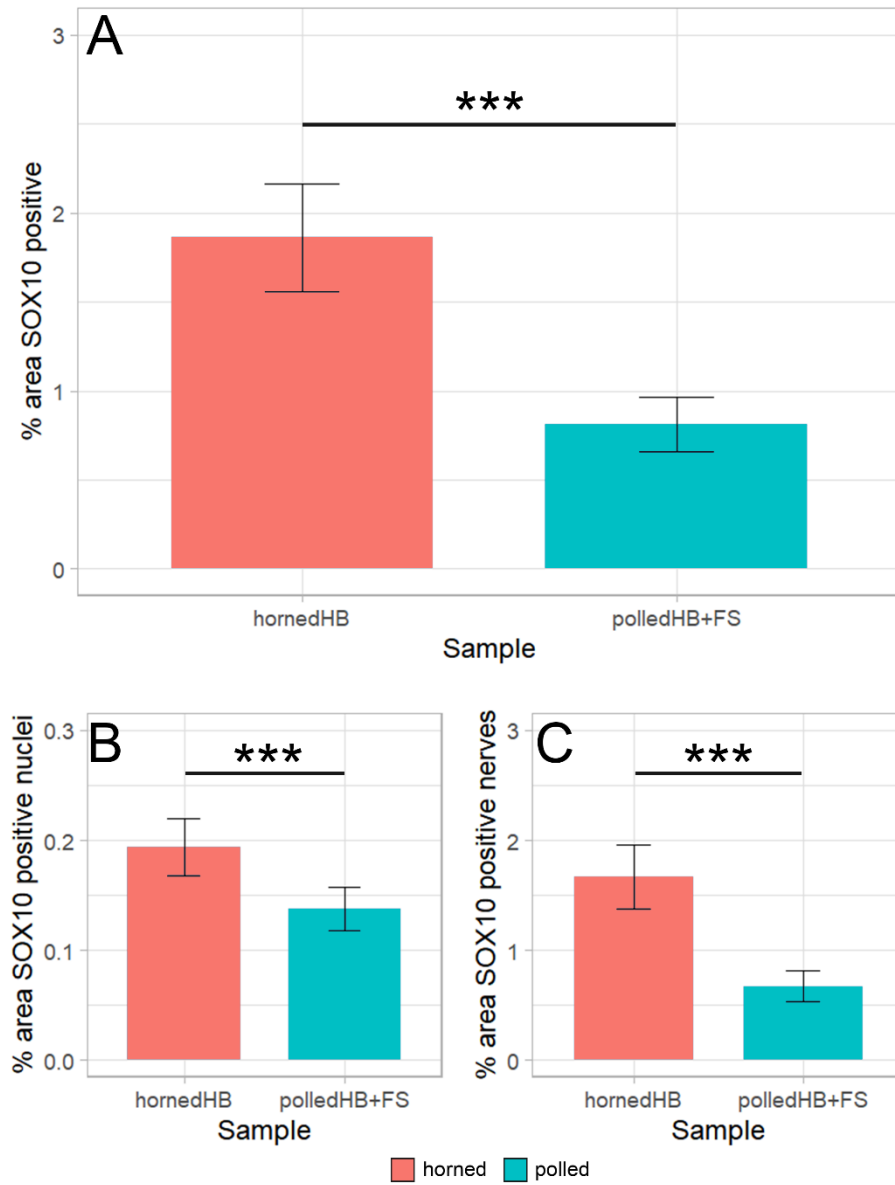
Figure 3.17: Positive staining of SOX10 in the horn bud of horned fetuses (HB, n = 5) and polled tissues (HB+FS; n = 3). Total positive staining (A), positively stained nuclei (B), and positively stained nerve tissue (C) presented as a proportion of the total tissue area. Columns and error bars represent the mean proportion and standard error. *** $p < 0.001$.

Figure 3.18: Positive staining of NGFR in the horn bud of horned fetuses (HB, n = 5) and polled tissues (HB+FS; n = 3). Positively stained nerve tissue presented as a proportion of the total tissue area. Columns and error bars represent the mean proportion and standard error. *** p < 0.001.

3.3.9 Localisation of RXFP2

The RXFP2 antibody was validated by the positive staining in Leydig cells from adult bovine testis (Figure 3.19A). RXFP2 is expressed in Leydig cells, seminiferous germ cells, spermocytes and spermatids (Pitia et al. 2017). There was no staining when the tissue was not exposed to the primary antibody (Figure 3.19B).
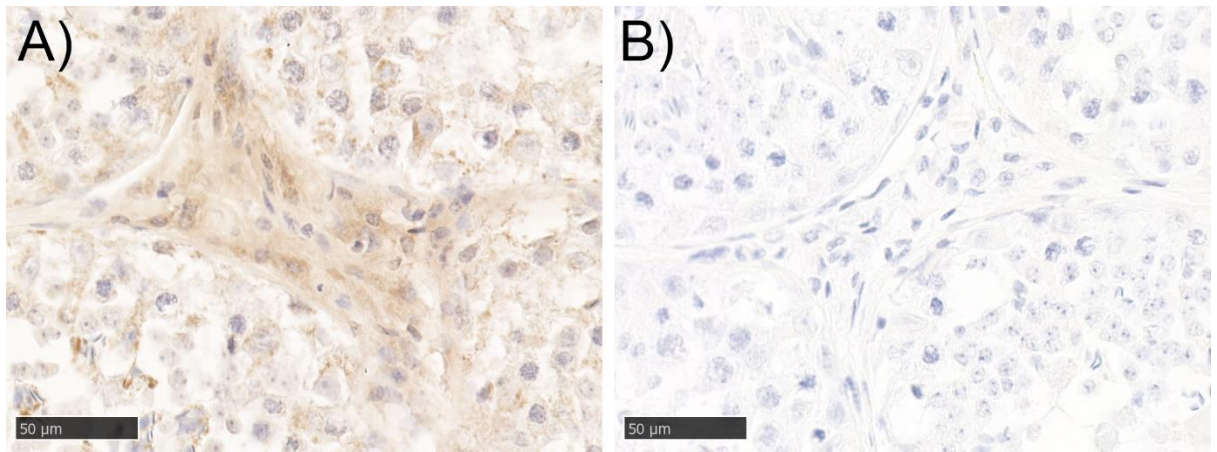
Figure 3.19: Staining of Leydig cells with RXFP2 antibody and counterstaining with haematoxylin. (A) Positive control: Brown stained Leydig cells indicates presence of RXFP2 antigen. (B) Negative control: Unstained Leydig cells when the tissue was not expose to the primary antibody.

RXFP2 has been associated with the development of horns, scurs and antlers (Johnston *et al.* 2011; Kardos *et al.* 2015; Wiedemar & Drögemüller 2015; Duijvesteijn *et al.* 2018; Wang *et al.* 2019c; He *et al.* 2020; Wang & Gill 2021), but the role that this gene plays in bovine horn development is not known. In the horn bud of the horned fetuses, RXFP2 was detected in the epithelium, developing peripheral nerves and osteoblasts (Figure 3.20). There was also low expression in endothelial cells of blood vessels. In the polled tissues, RXFP2 staining was observed in nerve tissues, osteoblasts and osteoprogenitor cells (Figure 3.20C-D). However, in contrast to the horned horn bud sections, RXFP2 was not located in the epithelium of polled fetuses (Figure 3.20A). In the horn bud of fetuses with the horned genotype, there was 1.25 times more positive staining for RXFP2 compared to the polled tissues (Polled HB+FS) (Figure 3.21; Fisher's Exact test: $p < 2.2e-16$, CI = 1.23-1.26; Appendix Figure B21). This is consistent with the results from staining with the SOX10 and NGFR antibodies which also showed that there was more nerve tissue in the horn bud of horned fetuses compared to the polled tissues.
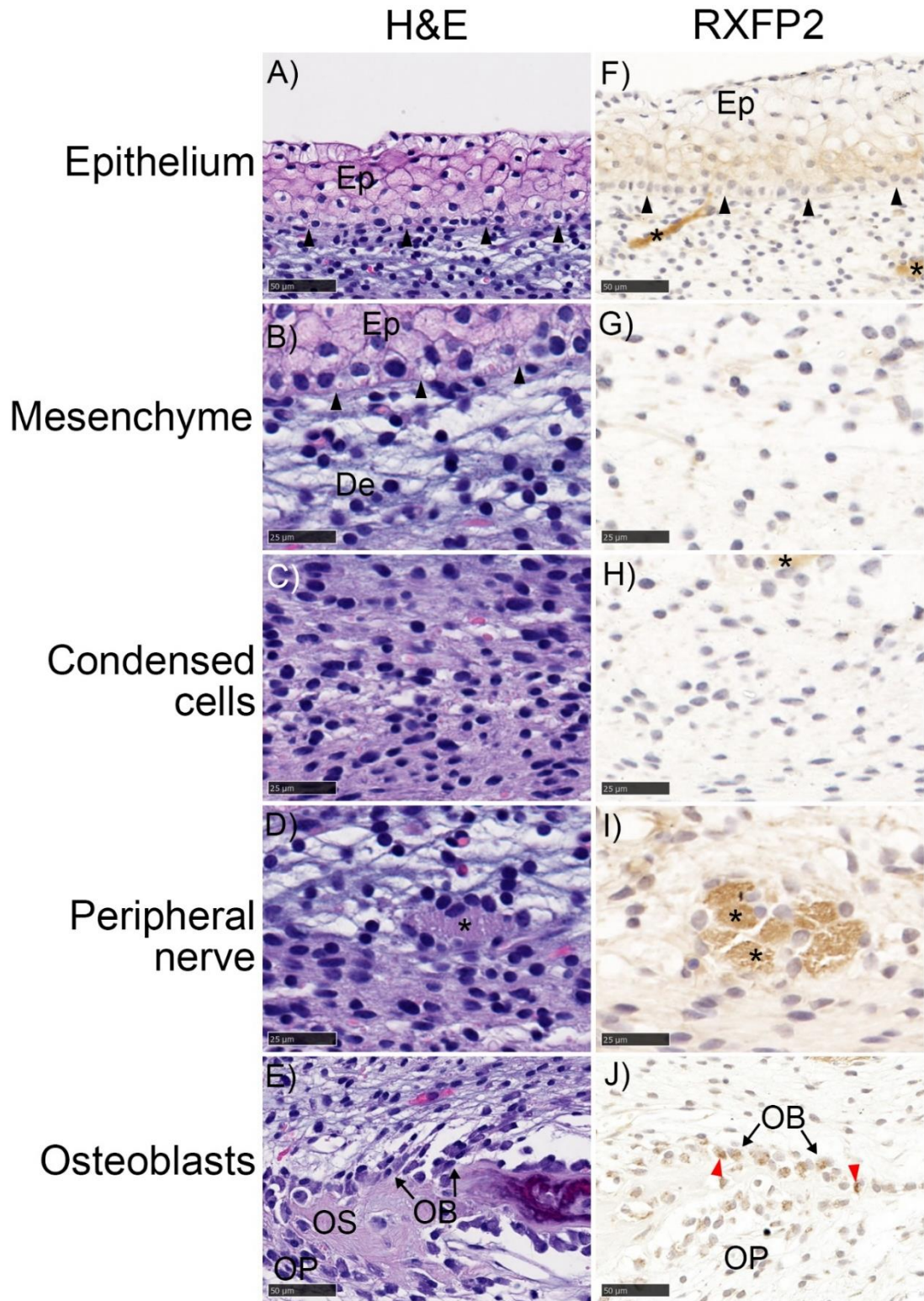
Figure 3.20: RXFP2 antibody staining of cells in the horn bud of horned fetuses. The horn bud was stained with haematoxylin and eosin (A-E) and RXFP2 antibody with haematoxylin counterstaining (F-J). Ep= epithelium, OB = osteoblasts (OB) and OP = osteoprogenitor cells, * = positively stained nerve tissue, red arrows = positively stained nuclei, black arrows = basal membrane. Magnification = 40x (A, E-F, J) and 80x (B-D, G-I)
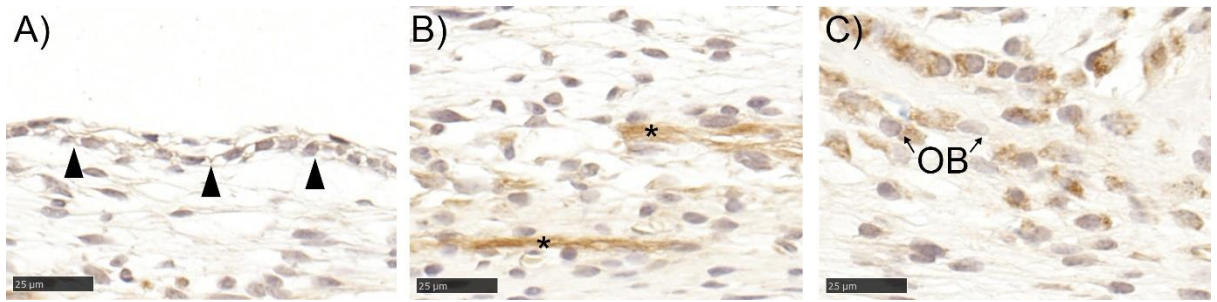
Figure 3.21: RXFP2 antibody staining and counterstaining with haematoxylin in the horn bud region of polled fetuses (A-C). Nuclei were not stained for RXFP2 in the epithelium and mesenchyme (A-B). Osteoblasts were positively stained for RXFP2 (C). * = nerves. OB = osteoblast. Black arrows = basal membrane. Magnification = 80x.
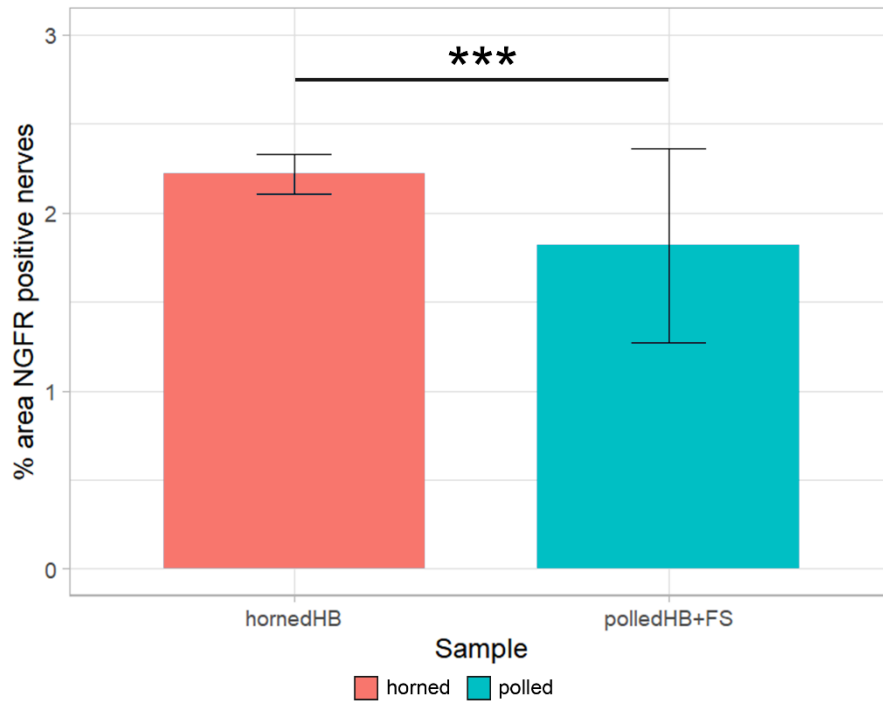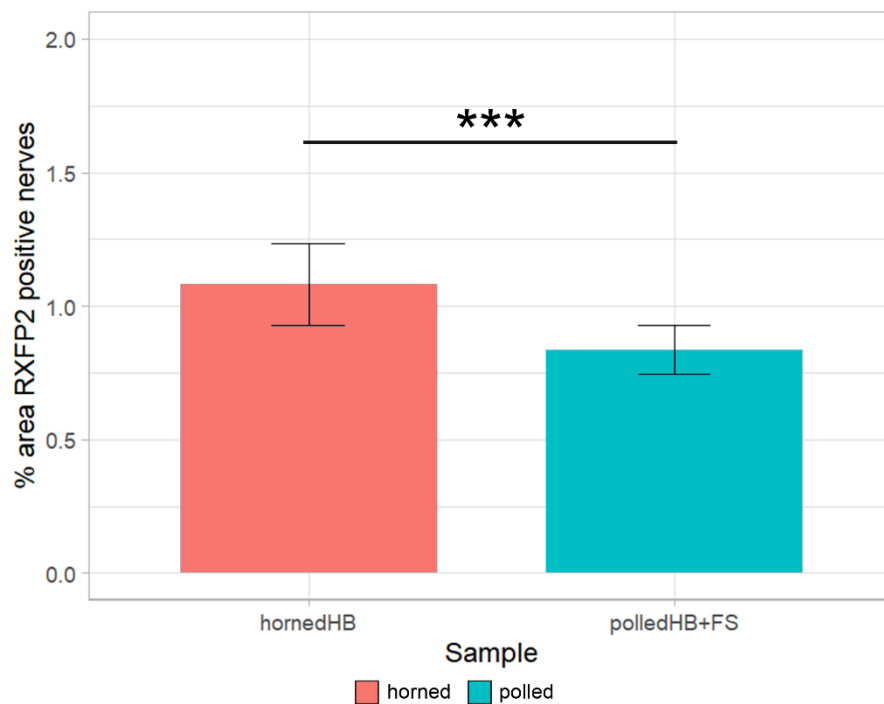

Figure 3.22: Positive staining of RXFP2 in the horn bud of horned fetuses (HB, n = 5) and polled tissues (HB+FS; n = 3). Positively stained nerves presented as a proportion of the total tissue area. Columns represent the mean proportion and standard error. *** $p < 0.001$

3.4 Discussion

The structure of the bovine horn bud before ~70 days of development has not been characterised, and therefore, the first aim of this study was to characterise the bovine horn bud structure at 58 days of development by histomorphometric analysis. Secondly, it was hypothesised that the horn progenitor cells are derived from the cranial neural crest, and to test this, the lineage of horn bud cells was investigated using SOX10 and NGFR as markers. Lastly, the location of the horn-associated protein, RXFP2, was determined in the horn bud.

At the horn bud of horned fetuses, a thickened epidermis and layer of condensed cells in the mesenchyme were observed (Figure 3.8). Previous studies also found that the epithelium layer was thicker in the horn bud compared to frontal skin in horned fetuses > 70 days of development (Allais-Bonnet *et al.* 2013; Wiener *et al.* 2015). Condensed mesenchymal cells were also observed in the horn bud of horned fetuses and this layer of cells was significantly reduced or absent outside the horn bud and in polled tissues. This has not been observed in other studies.

The identity of the condensed cells is not known, but they could be horn progenitor cells that have aggregated to undergo further differentiation. Aggregating cells often initiate organogenesis, including the development of bone, hair and teeth (Fuchs 2007; Li *et al.* 2016; Puthiyaveetil *et al.* 2016; Salhotra *et al.* 2020). In the case of hair and teeth, interactions between the epithelium and condensed cells are an essential part of that development (Fuchs 2007; Puthiyaveetil *et al.* 2016). As teeth develop, the mechanical compression caused by the expanding epithelium seems to lead to the aggregation of underlying mesenchymal cells (Svandova *et al.* 2020). However, the inverse mechanism is observed during avian follicle development for feathers. During feather follicle development, an aggregation of mesenchymal cells compress the epidermis, which initiates the differentiation of mesenchymal cells to primordial cells (the earliest stage of development of an organ or tissue) (Shyer *et al.* 2017). In

the present study, epithelial and mesenchymal changes were observed, although the mechanism causing these tissue changes needs to be explored further by studying the initial differentiation of cells in the horn bud.

A hypothesis tested here was that cells important for horn development are derived from the cranial neural crest. The neural crest markers, SOX10 and NGFR, were not detected in 'horn progenitor' cells identified in the condensed cell layer (section 3.3.4). However, it is possible that the cells may have differentiated and no longer express these markers. Therefore, investigation of neural crest markers at earlier time points than day 58 are essential to determine cellular origin of the horn progenitor cells.

SOX10 was detected in structures presumed to be peripheral nerves in the fetal horn bud, which was supported by the detection of NGFR in the peripheral nerve. NGFR, which binds to neurotrophins, has functions involving neuronal survival, neurogenesis, neurite outgrowth and apoptosis (Lu *et al.* 2005). SOX10 was also detected in glial cells. SOX10 is expressed in Schwann cells through multiple stages of development and regulates the myelination of peripheral nerves (Britsch *et al.* 2001; Finzsch *et al.* 2010; Srinivasan *et al.* 2012). Only pro-myelinating Schwann cells are reported to express SOX10, but not NGFR, whereas Schwann cell precursors and immature Schwann express both markers (Edgar *et al.* 2013). The cells identified in the horn bud are likely to be pro-myelinating Schwann cells.

The pattern of staining of SOX10 in the bovine horn bud was different from the expression observed in the horn bud of ovine fetuses (Wang *et al.* 2019c). In the ovine horn bud at approximately 90 days of development, positively stained cells were located in the epithelium and nearby in the mesenchyme (Wang *et al.* 2019c). This is in contrast to the present study, where strong SOX10 was detected in very few cells. The difference in staining may be because of the developmental stage of the fetuses studied. This may suggest that the SOX10

expressing cells may migrate to the horn bud epithelium after 58 days of development, but this will need further investigation.

The innervation in the horn bud of horned fetuses and the tissues from polled fetuses (horn bud region and frontal skin) was quantified and found to be significantly different. Peripheral nerves in the horn bud were larger than those detected in the polled tissues. Wiener *et al.* (2015) observed 'thick nerve bundles' in the horn bud of older fetuses and comparatively smaller nerves in polled fetuses. These nerves are likely to be the corneal branch of the sensory trigeminal nerve (Buda *et al.* 2011). The trigeminal nerve originates from the pons in the brain and passes through the forehead towards the horns (Figure 3.23).
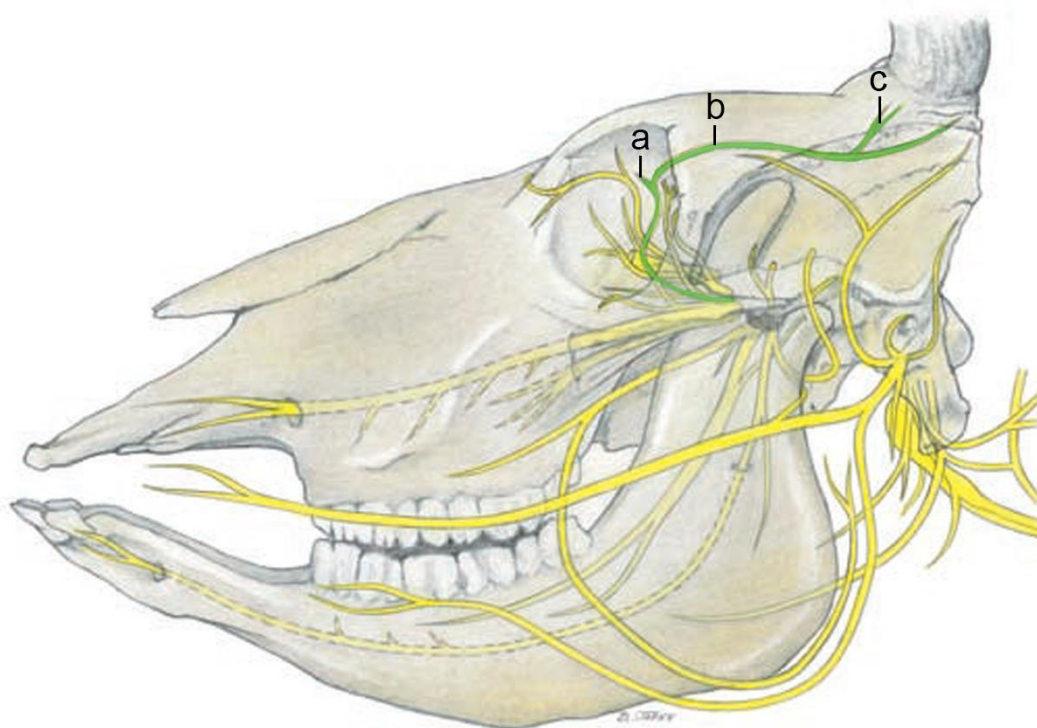
Figure 3.23: Innervation of the bovine horn. The trigeminal nerve, originating from the pons travels through the orbitorotundum foramen (not shown) and branches to the (a) lacrimal nerve, (b) zygomaticotemporal branch and (c) cornual branch. Figure adapted from Buda *et al.* (2011).

The finding that RXFP2 is present in peripheral nerves indicates that these nerves may have a role in the development of horns in the embryo. When RXFP2 interacts with its ligand, INSL3, in rat male fetal gubernacular cells, the pathways involved with neurogenesis were altered (Johnson *et al.* 2010). Nervous system development, neurogenesis, neuron differentiation, axon guidance and neuron projection morphogenesis were among the significant Gene Ontology (GO) pathways altered in the rat gubernacular cells, despite the gubernacular cells having no neuronal function (Johnson *et al.* 2010). Additionally in this rat study, *NGFR* had a 2.4 fold increase in expression in response to exposure to 10 nM of INSL3, suggesting a potential connection between RXFP2/INSL3 and *NGFR* expression (Johnson *et al.* 2010). As *RXFP2* is expressed in the horn bud nerves, nerve growth may be an important feature for horn ontogenesis. This is supported by a study of antler growth which demonstrated that a denervated antler is smaller than an innervated antler in the same deer during the first season of antler growth (Li *et al.* 1993).

RXFP2 was also detected in the horn bud epithelium, and therefore, may be expressed by the epithelial cells. In mice, RXFP2 has been detected in the oral epithelium during upper molar development from the bud stage (E13.5) to the bell stage (E17.5) (Duarte *et al.* 2014). Similarly, through spatial localization, RXFP2-like transcripts were detected in the oral epithelium of zebrafish larvae (Donizetti *et al.* 2015). Zebrafish lack teeth so the purpose of RXFP2 expression in zebrafish oral epithelium is not known (Abbate *et al.* 2006). Overall, it appears that RXFP2 plays a role in epithelium development and bud development of epithelial structures in embryos.

RXFP2 is expressed by the epithelium and peripheral nerves, and in addition can pass along nerves to other regions of the brain (Sedaghat *et al.* 2008). Therefore, RXFP2 expressed in the brain may be transported to the horn bud through the nerves. However, gene expression studies have detect *RXFP2* mRNA in the horn bud of day 70+ bovine fetuses, which suggests there is local synthesis of the RXFP2 (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014).

RXFP2 is differentially expressed between horned and polled bovine and ovine fetuses at 90 days (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014), has been associated with horn status and shape in sheep (Kardos *et al.* 2015; Wiedemar & Drögemüller 2015; Duijvesteijn *et al.* 2018; He *et al.* 2020), and is linked to scurs in sheep (Johnston *et al.* 2011) and cattle (Wang & Gill 2021). The evolutionary loss of antlers in Musk deer and Chinese water deer is attributed to nonsense and missense mutations in *RXFP2* (Wang *et al.* 2019c). Combining the knowledge of the association of *RXFP2* with horn and antler development, and that RXFP2 locally occurs in developing peripheral nerve and horn bud epithelium, it is most likely that the development of these structures is modulated by RXFP2, and that this modulation can lead to phenotypic differences in horn status.

RXFP2 is a G-protein coupled receptor (GPCR) and can result in the production of cAMP depending upon the associated G-protein. Through GPCRs, G-proteins transmit a signal from outside the cell to inside. There is evidence of RXFP2 coupling with G-proteins $G\alpha_s$ and $G\alpha_{oB}$ (Halls *et al.* 2006; Esteban-Lopez & Agoulnik 2020). Many studies have reported an increase in cAMP production after the addition of INSL3 and this has been observed in HEK293T cells, human and rat gubernacular cells, mouse Leydig cells and human primary osteoblasts (Kumagai *et al.* 2002; Halls *et al.* 2007; Johnson *et al.* 2010; Ferlin *et al.* 2011; Pathirana *et al.* 2012). In contrast, a study has reported decreases in cAMP in rat ovarian theca and testicular Leydig cells when exposed to INSL3 (Kawamura *et al.* 2004). The differences in cAMP effects are most likely due to the associated G-proteins. For example, RXFP2 coupling with $G\alpha_s$ showed an increase of cAMP through adenylate cyclase, while RXFP2 coupling with $G\alpha_{oB}$ decrease cAMP (Halls *et al.* 2006).

Some studies have investigated downstream signalling of RXFP2 activation by INSL3. In a microarray study of rat gubernacular cells (Johnson *et al.* 2010), exposure to INSL3 led to increased expression of genes that drive hepatocyte growth factor (HGF). HGF/MET Proto-Oncogene, Receptor Tyrosine Kinase (MET) interactions which contribute to morphogenesis,

cell survival, proliferation and motility HGF/MET signalling occur in various cell types including Schwann cells, epithelial cells and keratinocytes (Krasnoselsky *et al.* 1994; Brinkmann *et al.* 1995; Mildner *et al.* 2002) to promote axonal growth and survival of sensory nerves (Maina *et al.* 1997). HGF/MET signalling can activate STAT3, PI3K/AKT, MEK/ERK (a pathway involved in the MAPK cascade) and JNK signalling (Liu *et al.* 2020c). Some of these pathways are induced in other cell types in response to the INSL3/RXFP2 system. Phosphorylation of MEK/ERK signals, indicating activity, was induced by INSL3/RXFP2 in human primary osteoblasts (Ferlin *et al.* 2011). This INSL3/RXFP2 caused osteoblast proliferation and differentiation (Ferlin *et al.* 2011; Ferlin *et al.* 2017). In myotubules, INSL3/RXFP2 promoted the AKT/mTOR/S6 pathway which lead protein synthesis (Ferlin *et al.* 2018).

Expression of genes involved in Wingless-Type MMTV Integration Site Family (WNT) signalling are also increased in response to INSL3 in rat gubernacular cells (Johnson *et al.* 2010). WNT signalling is involved in cell fate specification, proliferation, 3D organization, migration and differentiation (Wiese *et al.* 2018). One of the canonical WNT signalling pathways allows for the accumulation of β-catenin in the cytoplasm (Kestler & Kuhl 2008). β-catenin interacts with the transcription factors TCF and LEF in the nucleus to activate transcription (Kestler & Kuhl 2008). Strikingly, Lef1-null mice show loss of ectodermal appendages, including mammary glands, hair follicles and teeth (Van Genderen *et al.* 1994; Wiese *et al.* 2018). In early tooth development, *LEF1* is expressed in both the epithelium and mesenchyme (Sasaki *et al.* 2005) and WNT genes are expressed in the epithelium (Sarkar & Sharpe 1999). The Lef1-null mice also lacked trigeminal nerve nuclei in the brain (Van Genderen *et al.* 1994). The trigeminal nerve eventually branches to the cornual nerve that innervates bovine horns (Buda *et al.* 2011). Changes in WNT signalling have led to phenotypic changes that parallel with those observed in polled animals (loss of ectodermal appendage).

Thus, if WNT signalling occurs in the horn bud, it may be initiated by the RXFP2/INSL3 system.

Bone morphogenetic protein (BMP) signaling genes were also increased in rat gubernacular cells in response to INSL3 (Johnson *et al.* 2010). BMP signaling is conserved across species and has roles in patterning, differentiation and apoptosis essential for early development (Wang *et al.* 2014a; Montanari *et al.* 2021; Yan & Wang 2021). In the canonical pathway, BMP binds to its receptor, which is a heterotetrameric complex consisting of dimers of type I and type II serine/threonine kinase receptors (Wang *et al.* 2014a). This interaction leads to the phosphorylation of receptor-regulated SMADs (Wang *et al.* 2014a). The receptor-regulated SMADs associate with the co-mediator, SMAD4, to act as nuclear transcription factors (Wang *et al.* 2014a). Knockout studies of BMPs, BMP receptors and SMADs often lead to embryonic lethal defects (Wang *et al.* 2014a). Of note, mutations in the BMP pathway affect craniofacial development, the cells of which are derived from the cranial neural crest (Graf *et al.* 2016). There is also cross talk between BMP and WNT signaling (Wang *et al.* 2014a).

The HGF, WNT and BMP signaling pathways are all activated by RXFP2 which suggest that RXFP2 may promote cell proliferation and differentiation, and epithelial-mesenchymal interactions. RXPF2 may be also important for sensory nerve development axon growth and survival. INSL3/RXFP2 binding also leads to the expression of genes involved in axon guidance in rat fetal male gubernacular cells, a cell type that is not neuronal. Thus, RXFP2 activation in the horn bud epithelium may cause gene expression that guides nerves to the horn bud.

Due to the difficulties in obtaining high quality samples, the final sample numbers were lower than expected. Nevertheless, the sample numbers herein are greater than those of previous histological studies of horn development, where only a single fetus was observed per time-point and genotype (Allais-Bonnet *et al.* 2013; Wiener *et al.* 2015). The low sample numbers in these studies reflect the difficulties in obtaining bovine fetal tissues.

To overcome the low sample numbers when comparing the tissue layers histological herein, the centre of the horn bud was compared to sections outside the horn bud. This was not possible for the immunohistochemistry analysis though, because the stained regions corresponded to the nerve area and nerves presumably travel through the outer regions towards the inner regions. Therefore, it did not make sense to separate the inner and outer horn bud for the immunohistochemistry experiments. Instead, the two horn bud region samples and single frontal skin samples from polled fetuses were combined and assumed to be the same based on histology. To improve the evidence obtained from this experiment, measurements were taken from multiple sections for each sample, which reduced the likelihood of significance differences by chance. The differences were apparent but additional fetal samples for histological and immunohistochemical analyses would be ideal.

3.5 Conclusion

Macroscopic and microscopic differences between horned and polled tissues of the horn bud region were identified at day 58 of bovine fetal development. In addition to a thickened epidermis which was observed in earlier studies, an aggregation of cells was observed in the mesenchyme that may represent horn progenitor cells. Furthermore, the contribution of cranial neural crest cells to horn ontogenesis was explored, but at this stage of fetal development, the suspected horn progenitor cells did not express the neural crest markers SOX10 and NGFR. Instead, developing peripheral nerves stained with both NGFR and SOX10 but glial cells, likely to be pro-myelinating, Schwann cells only stained with SOX10. RXFP2 was found to be localised in the horn bud epithelium and peripheral nerves. The role RXFP2 may play in the developing horn bud was explored and canonical pathways which may be affected by RXFP2 were identified specifically HGF/MET, WNT and BMP signalling that have been associated with morphological variation.

The time-point of 58 days, while earlier than other studies, was histologically more developed than expected. Future studies could study horn bud histology at earlier time points, potentially as early as day 50, although taking such samples will be challenging. Studying earlier fetal stages will contribute to the understanding of horn bud formation.

Other markers may be also assessed, such as those in the WNT pathway. For example, LEF1, which is known to be involve in the development of skin appendages, may be important for horn bud formation as well, and therefore, localisation of this protein may contribute to our knowledge of horn ontogeny.

Chapter 4:

Horn bud transcriptomic analysis of bovine fetuses at 58 days

4.1 Background

Although horns have economic and welfare implications for ruminant production, the genetic basis for horn development is not known. Adult horns consist of bone and epidermal tissue with nerve and vascular systems. The presence of these tissues suggests ectodermal (skin), neural crest (peripheral nerve, bone, blood vessels) and potentially mesodermal (bone, blood vessels) cells as possible origins of horns.

A genomic region on cattle chromosome 1 (BTA1) has four intergenic variants (Celtic, Friesian, Mongolian and Guarani) that cause cattle to be hornless or polled (Aldersey *et al.* 2020). The Celtic variant does not overlap with the Friesian, Mongolian or Guarani variants yet they have similar phenotypes. The function and effect of the variants on horn development is not understood, however, the variants may affect the expression of nearby genes.

Horn development in ruminants is also impacted by mutations in various genes elsewhere in the genome (e.g. *ZEB2*, *TWIST1*, *RXFP2*, *MTX2*, *HOXD*, *FOXL2*, *KCNJ15* and *ERG*). Deletions and frameshift mutations in *ZEB2* cause polledness in cattle. Cattle with these *ZEB2* variants also have syndromes that affect growth and female fertility (Capitan *et al.* 2012; Gehrke *et al.* 2020b). A frameshift mutation in *TWIST1* causes the formation of Type II scurs in cattle (Capitan *et al.* 2011). Scurs are horn-like headgear that range from small 'scabs' to longer appendages that are very rarely attached to the skull, unlike true horns. Variants in *RXFP2* are associated with polledness in sheep (Johnston *et al.* 2011; Wiedemar & Drögemüller 2015; Duijvesteijn *et al.* 2018) and may be associated with loss of antlers in musk deer (Wang *et al.* 2019c). *RXFP2* is also associated with horn shape and the occurrence of scurs in sheep (Wang *et al.* 2014b; Pan *et al.* 2018). A genome-wide association study (GWAS) found *MTX2* and the *HOXD* cluster to be associated with the four-horn phenotype in some sheep breeds (He *et al.* 2016; Kijas *et al.* 2016). *FOXL2, KCNJ15* and *ERG* are associated with the Polled Intersex Syndrome in goats (Pailhoux *et al.* 2001; Guo *et al.* 2021).

123

Polled cattle have additional eyelashes on the eyelid giving a bushy appearance (Allais-Bonnet *et al.* 2013). Mutations in *TWIST2* cause distichiasis in humans, a condition where eyelashes arise from the inner eyelid margin (Cervantes-Barragan *et al.* 2011), and therefore, it has been suggested that TWIST2 may be involved in horn development (Allais-Bonnet *et al.* 2013). Interestingly, frameshift mutations in *TWIST2* cause Setleis syndrome in humans, where scar-like lesions occur on the temporal regions of the head (Tukel *et al.* 2010).

Quantitative PCR studies have compared the horn bud and frontal skin of horned fetuses and the horn bud region between horned and polled bovine fetuses at various ages (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014). Only one gene (*OLIG2*) and one lincRNA (*LincRNA#1*) located near the POLLED variants have been shown to be significantly differentially expressed in these tissues (Allais-Bonnet *et al.* 2013). However, *OLIG2* was found to be differentially expressed between the horn bud and frontal skin of both horned and polled fetuses, indicating that the Celtic POLLED variant does not affect transcription of this gene (Gehrke *et al.* 2020b). *LincRNA#1* has greater expression in the horn bud region compared to the frontal skin in polled fetuses and compared to the horn bud of horned fetuses (Allais-Bonnet *et al.* 2013). Further studies have shown potential differential expression of other genes and lincRNA near the cattle POLLED region (*OLIG1*, *C1H21orf62* and *LincRNA#2*), but the results were not validated (Wiedemar *et al.* 2014). These studies were limited by sample number (Wiedemar *et al.* 2014) and number of genes analysed (Allais-Bonnet *et al.* 2013). The detection of all expressed genes is now more readily accessible via RNAseq technology.

Of the candidate genes for polledness, *RXFP2* and related pathways are most likely to be involved in horn development. *RXFP2* has been shown to increase the expression of genes that drive Wingless-Type MMTV Integration Site Family (WNT), hepatocyte growth factor (HGF), and bone morphogenetic protein (BMP) signalling (described in Appendix Section C1-Section C3). As these pathways can be activated by RXFP2, which has a role in horn development, they were considered potential candidates for horn development.

Interestingly, four of the polled candidate genes (*TWIST1*, *TWIST2*, *ZEB2* and *FOXC2)* are transcription factors that regulate epithelial-to-mesenchymal transition (EMT) which may occur during horn development (Chapter 1). Markers of EMT are the genes that encode E-cadherin (*CDH1),* N-cadherin (*CDH2*), occludin (*OCLN*) and vimentin (*VIM*). The expression of the EMT genes (*TWIST1*, *TWIST2*, *ZEB2*, *FOXC2*, *CDH1, CDH2, OCLN* and *VIM*) have been measured in 90 dpc bovine fetuses and the results showed that EMT did not occur at 90 dpc in the horn bud (Allais-Bonnet *et al.* 2013). However, as these genes were only assessed at day 90, EMT may still occur but at a different time-point in horn bud formation. Therefore, the role of EMT should be considered another potential candidate pathway for horn development.

In cattle, the horn bud has been reported to appear at about day 60 of fetal development (Evans & Sack 1973). Previous bovine transcriptomic studies collected data from fetuses that were 70 days and older (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014). Therefore, effects of the POLLED variants on gene expression in the POLLED region were possibly undetected because the tissue had already differentiated substantially. In this study, gene expression in the horn bud of horned and horn bud region of polled bovine fetuses at day 58 was compared. Particular attention was given to the expression of the genes within the POLLED region and the expression of the polled candidate genes and their pathways (Uhlén *et al.* 2015). The aim was to uncover the genes directly affected by the Celtic POLLED variant, and thereby, reveal the pathways important for horn ontogenesis.

4.2 Materials & Methods

4.2.1 Animals

The twelve horned and 12 polled Hereford heifers acquired for this study and used to produce fetuses were as described in Chapter 3. This additional study on 10 of the fetuses was approved by the University of Adelaide Animal Ethics Committee (Project Approval No. S-2018-105). The collection of the 10 fetuses was as described in Section 3.2.

4.2.2 Sample collection

Within 15 minutes of fetus collection, the fetal head was bisected cranio-rostrally and one half the head was stored in 5 ml of RNAlater for 24-48 hours at 4°C. Horn bud (HB), frontal skin (FS), forebrain skin (FB) and midbrain skin (MB) samples were collected using a 3 mm biopsy punch (Figure 4.1). Note: horn bud region (HBR) will be used to describe the 'horn bud' of polled fetuses. The biopsies included the layers of tissue described in Chapter 3, including the underlying developing cranial bone. The biopsies were stored in cryogenic vials at -80°C. In total, samples from six horned and four polled fetuses were collected.
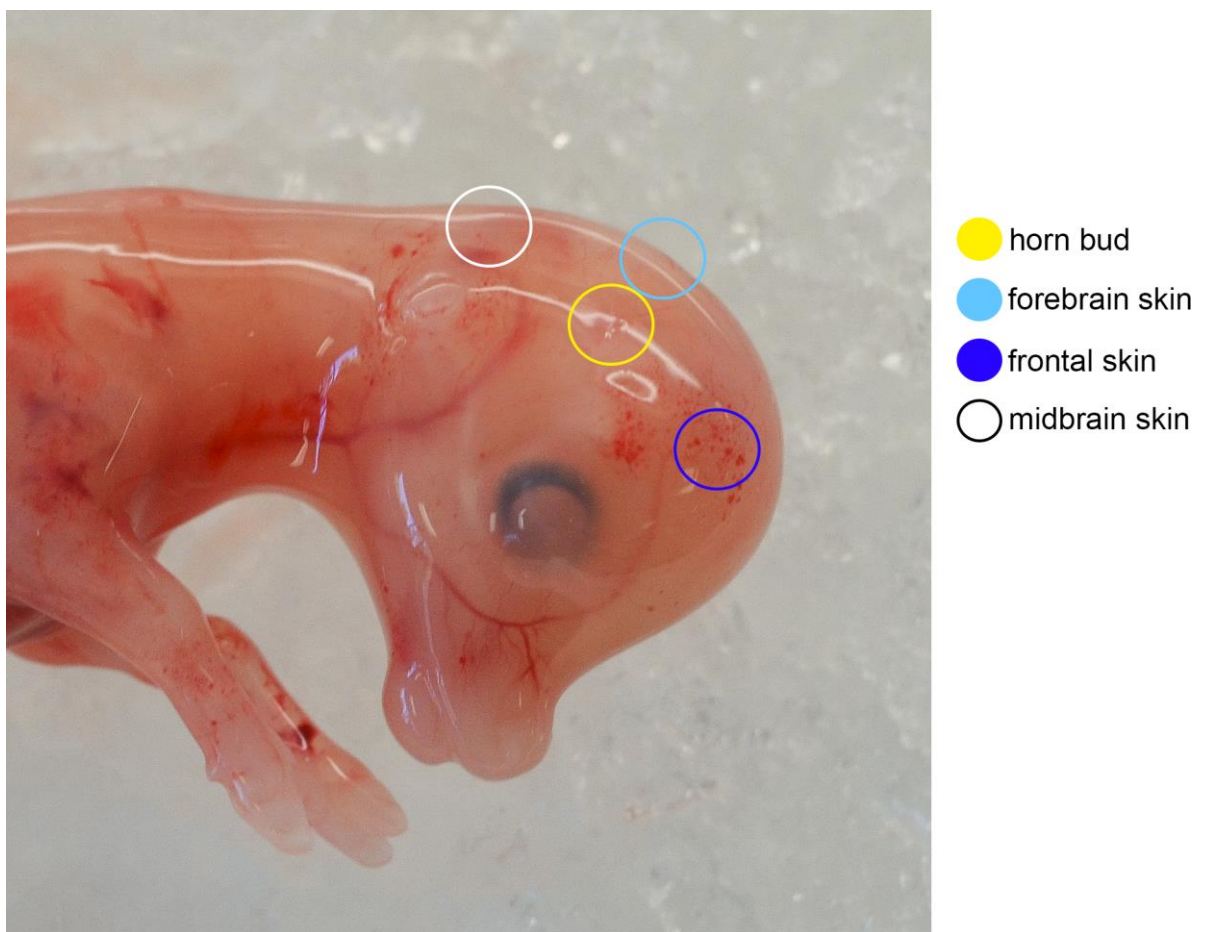
Figure 4.1: Biopsy sites for RNA sequencing.

4.2.3 RNA extraction and quality check

RNA was extracted from six horned (546, 618, 668, 736, 532 and 581) and four polled fetuses (667, 698, 709 and 701). The tissues were homogenised in 900 ul of QIAzol (Qiagen), and then, incubated for 5 minutes at room temperature. RNA was extracted using a RNeasy kit (Qiagen) according to the manufacturer's protocol. The RNA was eluted using 50 µl of RNAse-free water. The optical density (OD) at 260/280 was measured using a Nanodrop One spectrophotometer (ThermoFisher Scientific). The RNA concentration and RNA integrity number (RIN) were measured using a 2100 Bioanalyser Instrument (Aligent Technologies). Concentrations were confirmed using a Qubit fluorometer (Qubit 4, ThermoFisher Scientific) and RNA quality determined (Appendix Section C4). The quality of two horned samples (#532 and #581) and one polled sample (#701) was potentially too low for RNAseq analysis, and these samples were not sequenced.

4.2.4 RNAseq

The laboratory work described in this section was carried out by the Australian Cancer Research Foundation. Stranded mRNA libraries were prepared for four horned fetuses (#546, 618, 668 and 736) and three polled fetuses (#667, 698 and709) using the SMART-Seq Stranded kit (Takara Bio USA Inc.) with 100 ng of total RNA according to the manufacturer's instructions. RNA was incubated with a shearing mix for four minutes to fragment the RNA, as recommended for RIN scores of 5-7, and converted to cDNA. Adapters and indexes compatible with Illumina sequencing were ligated to the cDNA with five PCR cycles and purified using AMPure beads. Ribosomal cDNA was depleted with scZapR and scR-Probes. The library was further enriched in a second round of PCR (10 cycles). The final libraries were purified again with AMPure beads. Library sizes and yields were assessed with an Agilent Bioanalyser and diluted to 4 nM stocks. Libraries were pooled in equimolar ratios and sequenced on an Illumina NextSeq500 using two 300 cycle high output kits.

4.2.5 Bioinformatics analysis

Kelly Ren (Davies Research Centre, University of Adelaide) conducted the following methods in this section 4.2.5 under the guidance of Wai Low (Davies Research Centre, University of Adelaide) to prepare the results for interpretation.

4.2.5.1 Quality control

The initial quality of raw RNA-seq reads was checked using FASTQC v0.11.4 (Andrews 2010), and TrimGalore v0.4.2 (Krueger 2015) was used to trim the reads with Phred scores of 10. Sequencing adapters and reads shorter than 100 bp were removed by AdapterRemoval v2.2.1 (Schubert *et al.* 2016). The cleaned reads were checked by FASTQC again, and the number of reads were visualized (Figure 4.2).
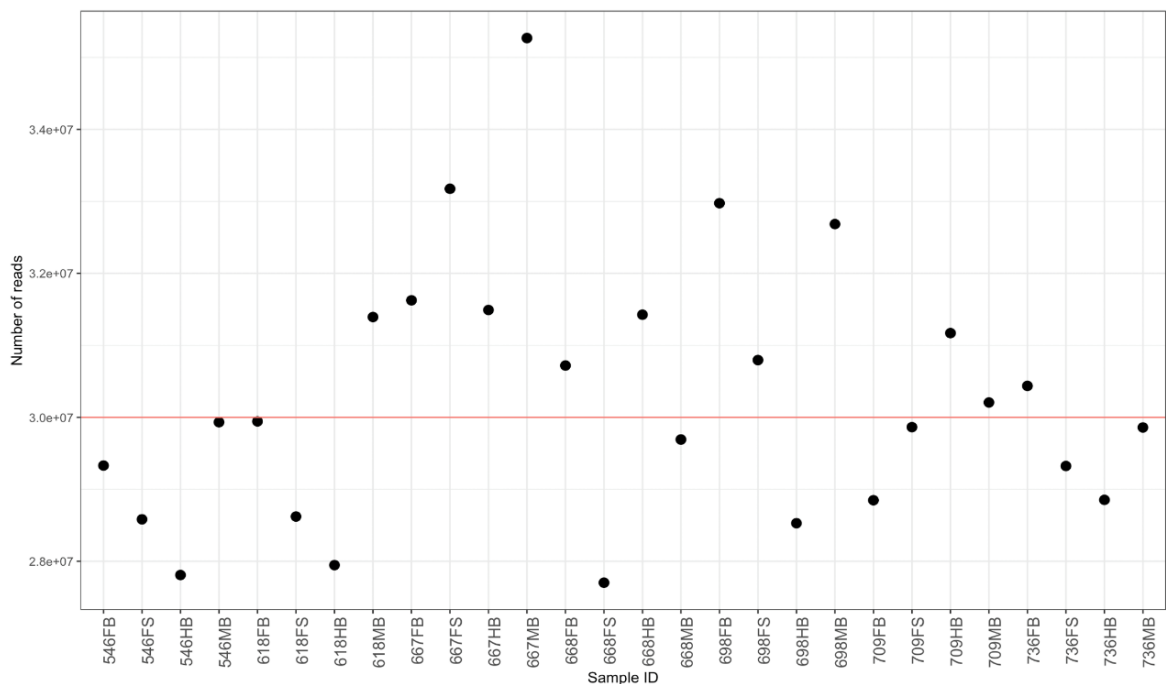
Figure 4.2: Number of reads for each sample. Red line shows the reference of 3.0e+07 reads.

4.2.5.2 Alignment and feature counts

The cleaned reads were mapped to the *Bos taurus* reference genome ARS-UCD 1.2 using Hisat2 (v2.1.0) (Kim *et al.* 2015) and 10 additional customised annotations (Table 4.1).

All samples had a mapping rate above 80%, indicating the libraries were of good quality (Figure 4.3). The mapped reads were sorted using SAMtools (v1.8) (Li *et al.* 2009) and summarised using FeatureCounts (v1.5.2) (Liao *et al.* 2014).

Table 4.1: Additional annotations included for mapping reads. Unknown loci (LOC) are predicted lincRNA in the POLLED region from the NCBI database (not in the Ensembl database). Inclusion of these annotations allow the identification of transcripts that aligned with these features.

Feature name	Description	Start	Strand
Celtic	Region overlapping the Celtic variant	chr1:2,429,109-2,429,320	+
Mongolian	Region overlapping the Mongolian variant	chr1:2,695,889-2,696,046	+
est_1	EST identified in Chapter Two	chr1:2,545,442-2,545,587	+
est_2	EST identified in Chapter Two	chr1:2,576,023-2,576,127	+
LOC104970777	Predicted lincRNA from NCBI database	chr1:2,016,219-2,027,521	-
LOC112447120	Predicted lincRNA from NCBI database	chr1:2,162,947-2,173,433	-
LOC104970778	Predicted lincRNA from NCBI database	chr1:2,241,318-2,251,416	-
LOC112447121	Predicted lincRNA from NCBI database	chr1:2,318,146-2,322,682	+
LOC112447133	Predicted lincRNA from NCBI database	chr1:2,318,146-2,322,682	+
LOC112447136	Predicted lincRNA from NCBI database	chr1:2,891,259-2,898,070	-
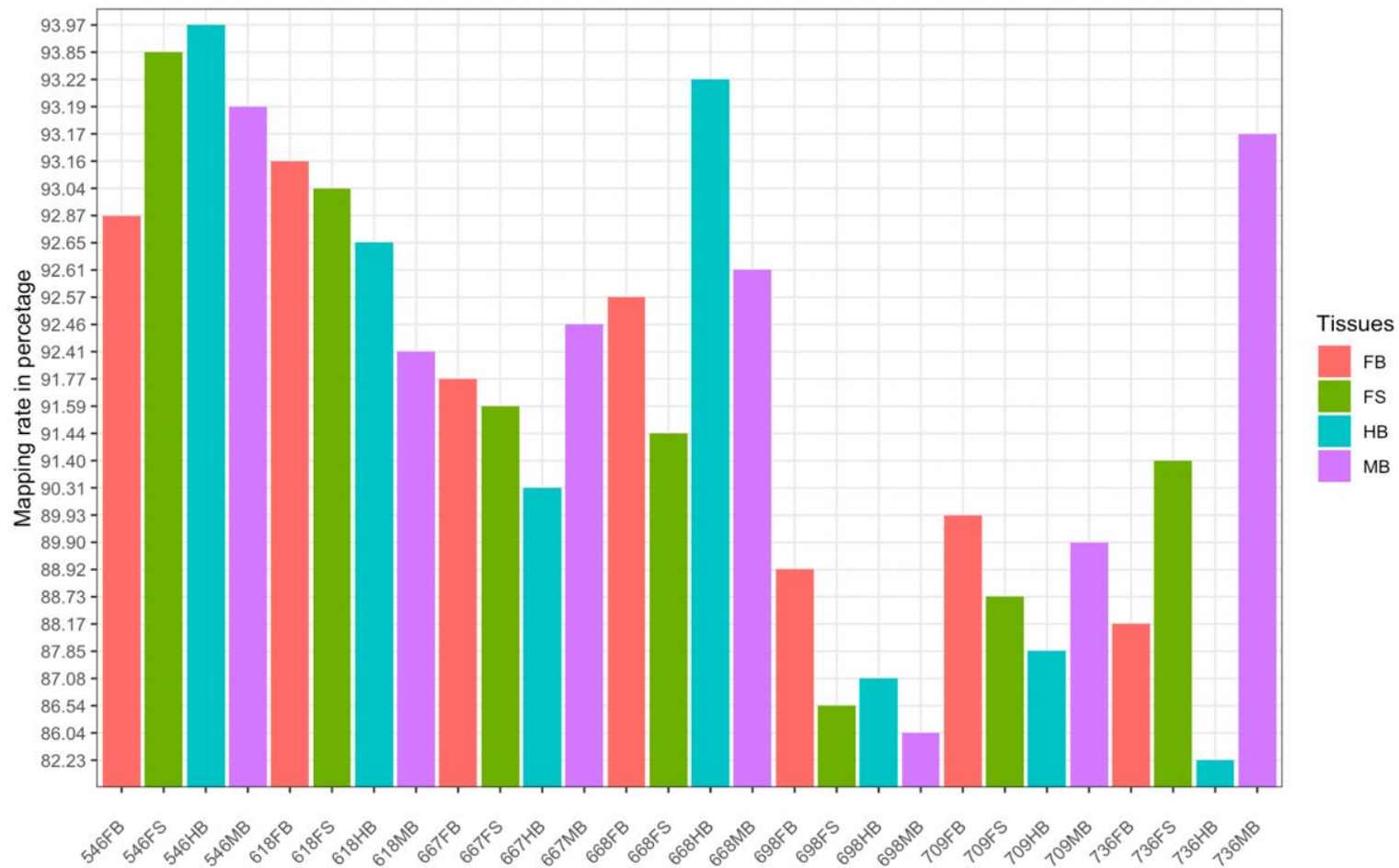
Figure 4.3: Mapping rate of all samples to the *Bos taurus* reference genome ARS-UCD 1.2 using Hisat2. Samples from #546, 618, 668 and 736 were from horned fetuses. Samples from #667, 698 and 709 were from polled fetuses.

4.2.5.3 Filtering and normalisation

Genes were selected and filtered for those with counts per million greater than 1 for at least three samples. The counts were normalised using the trimmed mean of M-values method (Schubert *et al.* 2016), where the M-values were weighted according to inverse variances by default. A combination of the voom method (Liu *et al.* 2015) with estimates of sample quality was used to account the heterogeneity in expression values. The multi-dimensional scaling plot (MDS) was used to observe tissue clustering after removing the noise (Figure 4.4).

4.2.5.4 Differential gene expression

The differential expression levels were analysed using the R package limma (version 3.44.3) linear model in R (available at https://figshare.com/: <DOI: 10.25909/19335278>). To check for similarity among tissues from the same fetus, the correlation between tissues was calculated using the duplicateCorrelation function. Gene expression was compared between tissues within a genotype (e.g. horned HB vs horned FS) and between genotypes (e.g. polled HB vs horned HB). In total, there were ten comparisons considered (Table 4.2). The p-values were adjusted by false discovery rate (FDR). Significantly differentially expressed genes had an FDR p-value < 0.05 and log fold change > 1.

4.2.5.5 GO and KEGG pathway analysis

The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis for differentially expressed genes was conducted using the *Bos taurus* cattle reference annotation ARS-UCD 1.2. The functions goana and kegga from limma (v3.44.3) were applied to summarise related GO and KEGG terms, respectively (available at https://figshare.com/: <DOI: 10.25909/19335314>). The p-value for each term was adjusted by the FDR method and filtered for a p-value < 0.05 for enriched terms. Differentially expressed genes from significant pathways were manually downloaded in R (available at https://figshare.com/: <DOI: 10.25909/19335236>).
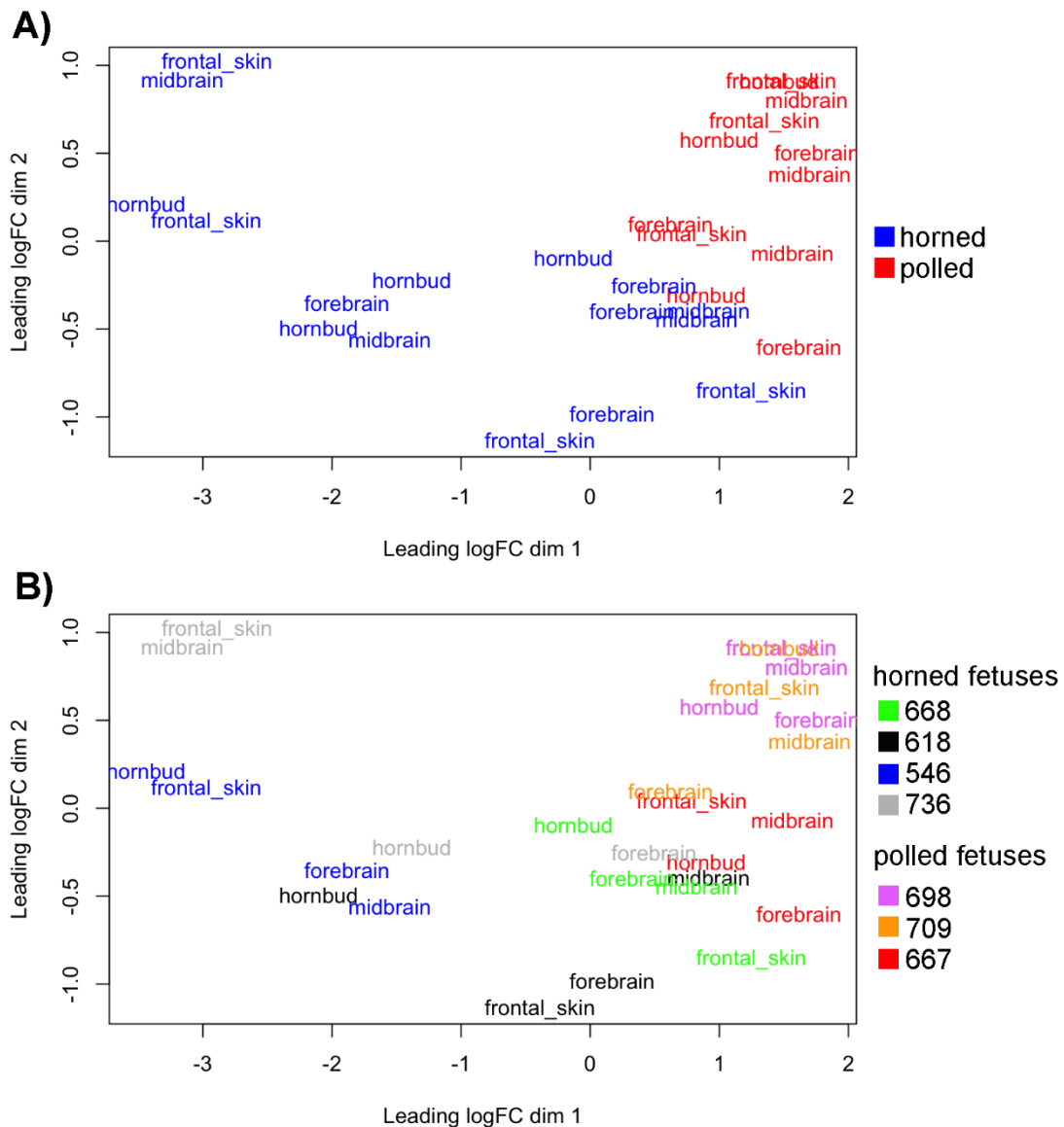
Figure 4.4: Multi-dimensional scaling (MDS) plot showing sample clustering. A) MDS plot with samples coloured by genotype. B) MDS plot with samples coloured by fetus ID. Note: "forebrain" and "midbrain" refer to the skin covering the corresponding brain regions rather than brain tissue itself.

Table 4.2: Differentially expressed gene lists obtained from ten comparisons.

Comparison	Genotype 1	Tissue 1	Genotype 2	Tissue 2
1	Horned	Horn bud	Horned	Frontal skin
2				Forebrain skin
3				Midbrain skin
4	Polled	Horn bud region	Polled	Frontal skin
5				Forebrain skin
6				Midbrain skin
7	Horned	Horn bud	Polled	Horn bud
8		Frontal skin		Frontal skin
9		Forebrain skin		Forebrain skin
10		Midbrain skin		Midbrain skin

4.2.5.6 Visualisation

GO terms were visualised using the R package, Rgraphviz (version 2.32.0) (Hansen K. D. *et al.* 2021). Other plots were visualised with ggplot2 (version 3.3.3) (Wilkinson 2011) (available at https://figshare.com/: <DOI: 10.25909/19374014>).

4.2.5.7 Determination of fetus sex

To determine the sex of the fetuses, the cleaned data were aligned to the UOA_Angus_1 *Bos taurus* cattle assembly (Low *et al.* 2020), which contains the Y chromosome. Gene counts on the Y chromosome were selected and plotted in a boxplot after filtering by counts-per-million > 1 in at least three samples.

4.3 Results

4.3.1 Number of differentially expressed genes

The overall consensus correlation for all comparisons between tissues from the same fetus were low (r = 0.25), suggesting that the tissues have different gene expression patterns. Gene expression was then compared between tissues and genotypes to identify differentially expressed (DE) genes (Table 4.3; Appendix Table C1-Table C10). The tissue comparisons

within horned fetuses yielded a greater number of DE genes than the equivalent comparisons in polled fetuses (Appendix Figure C2 A-F). When the same tissues were compared between the horned fetuses and the polled fetuses, the horn bud regions had the largest number of DE genes and the midbrain skin samples had the lowest (Table 4.3). There were more genes with decreased expression in the tissues from the horned fetuses compared to the same tissues in the polled fetuses (Table 4.3; Appendix Figure C2). Comparing across genotypes, there were few differentially expressed genes in common (Figure 4.5). There were no common genes for the HB/HBR vs FS comparisons, 52 common DE genes for the HB/HBR vs FB comparisons and 180 common DE genes for the HB/HBR vs MB comparisons.

Table 4.3: Number of differentially expressed genes for each tissue comparison. HB = horn bud, FS = frontal skin, FB = forebrain skin, MB = midbrain skin.

Comparison	Number of genes with increased expression [a]	Number of genes with decreased expression [b]	Total
Horned HB vs Horned FS	54	43	97
Horned HB vs Horned FB	740	4488	5228
Horned HB vs Horned MB	772	5523	6295
Polled HB vs Polled FS	2	0	2
Polled HB vs Polled FB	114	2	116
Polled HB vs Polled MB	176	104	280
Horned HB vs Polled HB	1214	6372	7586
Horned FS vs Polled FS	1122	5890	7012
Horned FB vs Polled FB	1480	3114	4594
Horned MB vs Polled MB	1451	1949	3400

[a] increased expression in horn bud or horned fetal tissue

[b] decreased expression in horn bud or horned fetal tissue
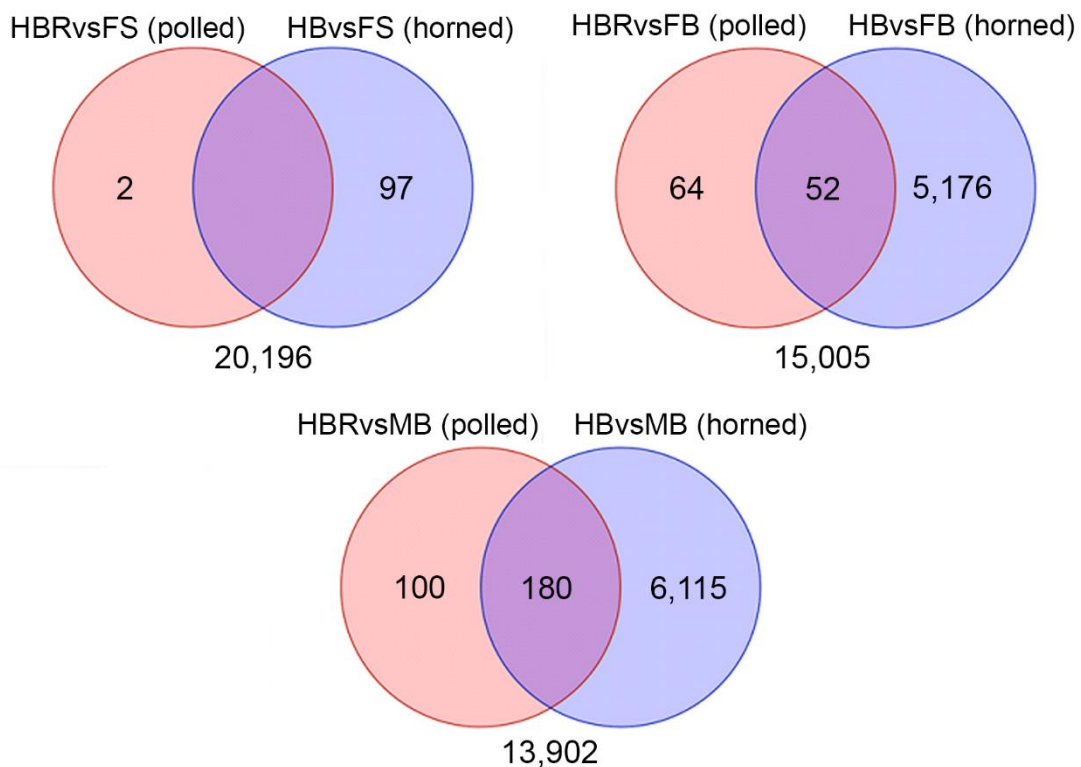
Figure 4.5: Venn diagram comparison of differentially expressed genes between the horn bud or horn bud region and frontal skin, forebrain skin and midbrain skin, between horned (n = 4) and polled fetuses (n = 3) collected at 58 days of development. HBR = horn bud region (polled), HB = horn bud (horned), FS = frontal skin, FB = forebrain skin, MB = midbrain skin.

The expression profile of the horn bud of horned fetuses was closer to the expression profile of frontal skin than to the forebrain skin. When the horn bud and frontal skin were compared in the horned fetuses, only 97 genes were differentially expressed, whereas 5528 genes were differentially expressed between the horn bud and the forebrain skin (Table 4.3; Appendix Figure C1). When the list of DE genes from the horned versus polled comparison of the horn bud region were compared with the list of DE genes from the horned versus polled comparison of the frontal skin, 79% of the genes were in common (Figure 4.6). In contrast, only 45% of the DE genes were in common between the horned versus polled comparisons of the

horn bud and of the forebrain skin. This is an interesting result because the horn bud and the forebrain skin samples were adjacent (Figure 4.1).
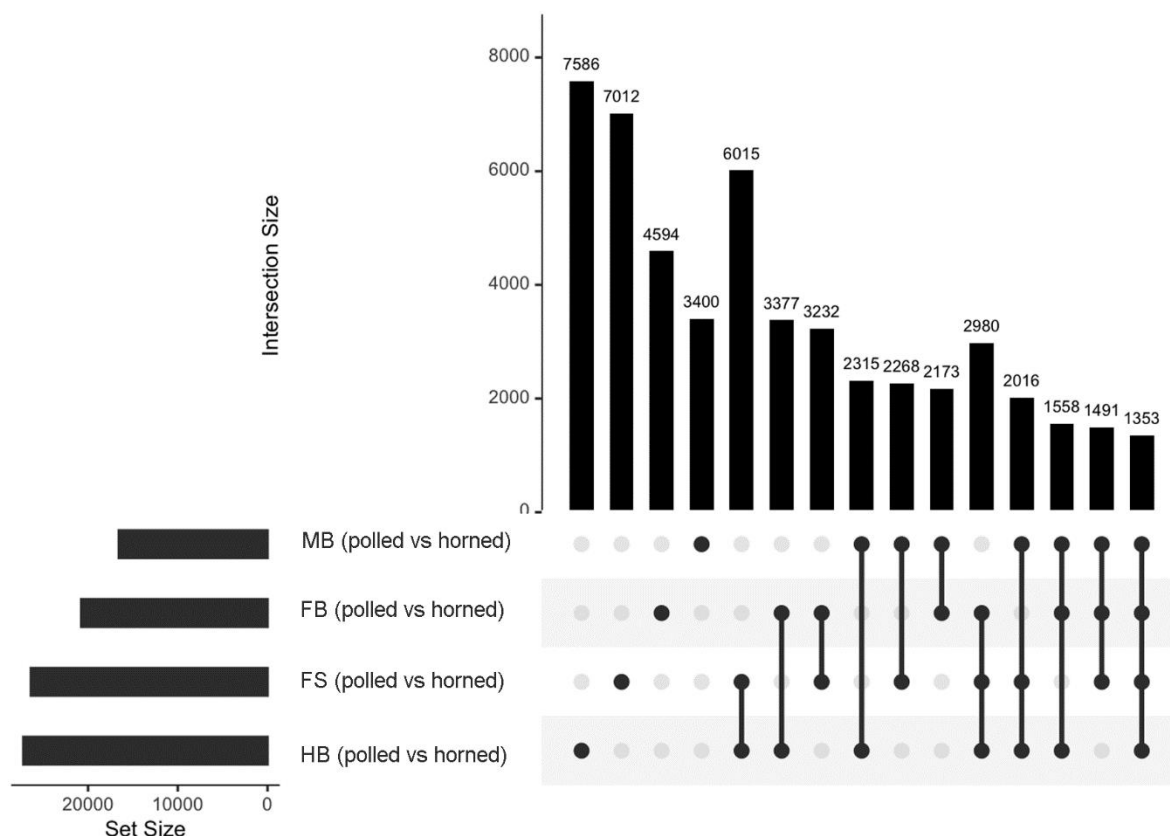
Figure 4.6: Overlapping DE genes between polled versus horned tissue comparisons. MB = midbrain skin; FB = forebrain skin; FS = frontal skin, HB = horn bud (horned) and horn bud region (polled).

4.3.2 Differential expression of genes in the POLLED region of BTA1

The genes (*IL10RB*, *IFNAR2*, *LOC526226*, *OLIG1*, *OLIG2,* and *C1H21orf62*), lincRNA (*LOC104970778*, *LOC112447121*, *LOC112447133* [LincRNA#1] and *LOC112447136* [LincRNA#2]) and ESTs (CB166156.1 and DN819280.1), which are located

in the POLLED region on cattle chromosome 1 (as defined in Chapter 2), were examined for differential expression. The POLLED region has four TADs encompassing the POLLED variants (Chapter 2). As the effect of the Celtic variant may not strictly operate within its own TAD, other nearby genes outside of this region (*DONSON, SON, GART, DNAJC28, TMEM50B, IFNGR2, IFNAR1, PAXBP1, SYNJ1, CFAP298* and *EVA1C*) and other features (*LOC104970777, LOC112447120*, Bta-miR-6501 and *RCAN1*) were also assessed. The expression differences were examined to identify the genes affected by genotype and tissue and the genes affected only by genotype.

In the horned fetuses, *C1H21orf62* had lower expression in the horn bud compared to forebrain skin and midbrain skin (i.e. horned HB ≠ FB or MB), although there was no difference in *C1H21orf62* expression between horn bud and frontal skin (i.e., horned HB = horned FS) (Table 4.4). There was no difference in *C1H21orf62* expression between the horn bud region and frontal skin, forebrain skin and midbrain skin in polled fetuses. Between horned and polled fetuses, *C1H21orf62* had lower expression in the horn bud region and frontal skin of horned fetuses.

Three other DE genes and lincRNA near the POLLED variants (*SON, EVA1C* and *LOC112447120*) were affected by genotype (Table 4.4). Between horned and polled fetuses, *SON* expression was higher in all tissues of horned fetuses. *EVA1C* expression was lower in the horn bud region of horned fetuses compared with the polled fetuses. *LOC112447120* expression was lower in the horn bud region, frontal skin and forebrain skin in horned fetuses compared to polled fetuses. *LOC112447120* also had lower expression in the horn bud compared to midbrain skin in horned fetuses.

Table 4.4: Differential expression of genes nearby the POLLED variants on BTA1* and candidate genes.

Genes	Horn Bud vs Frontal Skin		Horn Bud vs Forebrain Skin		Horn Bud vs Midbrain Skin		Horn vs Polled			
	HB vs FS	HB vs FS	HB vs FB	HB vs FB	HB vs MB	HB vs FB	HB vs HB	FS vs FS	FB vs FB	MB vs MB
Genes nearby POLLED variants*										
C1H21orf62	-	-	↓ 1.06 [a]	-	↓ 1.06	-	↓ 1.41	↓ 1.14	-	-
LOC112447120 (lincRNA)	-	-	-	-	↓ 1.38	-	↓ 1.82	↓ 1.76	↓ 1.21	-
SON	-	-	-	-	-	-	↑ 1.29	↑ 1.27	↑1.34	↑1.46
EVA1C	-	-	-	-	-	-	↓ 1.17	-	-	-
LOC526226 [b]	-	-	-	-	-	-	-	-	↑1.38	-
IFNGR2	-	-	-	-	↑ 1.23	-	-	-	-	-
Candidate genes										
RXFP2	↑ 2.29	-	↑ 2.15	-	↑ 3.23	-	↑ 2.38	-	-	-
TWIST2	-	-	↑ 1.22	-	↑ 1.36	-	↑ 1.43	↑ 1.40	-	-
ZEB2 (ENSBTAG00000048810)	-	-	↑ 1.06	-	↑ 2.14	↑ 1.54	↑ 1.67	↑ 1.55	↑ 1.26	↑ 1.06
ZEB2 [c] *(ENSBTAG00000012615)*	-	-	-	-	↑ 2.05	↑ 1.54	↑ 1.55	↑ 1.30	↑ 1.41	↑ 1.04
CDH2	-	-	↑ 1.06	-	↑ 2.14	↑ 1.54	-	↑ 1.55	-	-
KCNJ15	-	-	↑ 1.03	-	-	-	-	↓ 1.55	↓ 1.54	-
HOXD10	-	-	↓ 1.53	-	↓ 1.33	-	↓ 2.59	↓ 2.32	↓ 1.37	↓ 1.93
HOXD1	-	-	↓ 1.08	-	-	-	↓ 1.37	-	-	↓1.02
HOXD8	↓ 1.55	-	↓ 1.08	-	↓ 1.07	-	-	-	-	-
VIM	-	-	-	-	-	-	↑ 1.21	↑ 1.06	↑ 1.57	↑ 1.38
HAND1	-	-	-	-	-	-	↓ 1.17	↓ 1.93	-	↓ 1.15
FOXL2	-	-	↑ 2.97	↑ 3.08	↑ 1.93	↑ 1.73	-	-	-	-
TWIST1	-	-	-	-	-	-	-	-	↑ 1.43	-
FOXC2	-	-	↑ 1.20	-	-	-	-	-	-	-

* Genes within ~530 kb from the Celtic POLLED variant; HB = horn bud; FS = frontal skin; FB = forebrain skin; MB = midbrain skin; █ = homozygous horned; █ = homozygous polled;

↑ increased expression in horn bud or tissue from horned fetus when compared with polled fetus; ↓ decreased expression in horn bud or horned tissue for other comparisons; - not differentially expressed

[a] log fold change

[b] Histone H4/Osteogenic growth peptide

[c] There are two annotations for ZEB2 in Ensembl (release 108) which do not overlap. These were differentially expressed. There is only a single annotation in NCBI (Annotation release 106).

Genes and lincRNA nearby POLLED region analysed that were not differentially expressed: Protein-coding – *ITSN1, CRYZL1, DONSON, GART, DNAJC28, TMEM50B, IFNAR1, IL10RB, IFNAR2, OLIG1, OLIG2, PAXBP1, SYNJ1* and *CFAP298*; LincRNA - *LOC104970778, LOC112447121, LOC112447133* [*LincRNA#1*], and *LOC112447136* [*LincRNA#2*]

Candidate genes analysed but not differentially expressed: *MTX2, RCAN1, CDH1, OCLN* and *ERG*

4.3.3 Differential expression of candidate genes

Thirteen candidate genes, which are not located in the bovine chromosomal regions containing the polled variants, were differentially expressed (Table 4.4). *RXFP2, KCNJ15, HOXD1, HOXD8* and *HOXD10* expression patterns were influenced by both tissue and genotype. *RXFP2* expression was higher in the horn bud of horned fetuses compared with other horned tissues or the horn bud region of polled fetuses. *KCNJ15* had higher expression in the horn bud compared to forebrain skin, and had greater expression in horned fetuses than polled fetuses for frontal skin and forebrain skin. The *HOXD* genes generally had lower expression in horned tissue than polled tissues. *HOXD10* had lower expression in the horn bud compared to the forebrain skin and midbrain skin of horned fetuses and in all the tissues from horned fetuses compared to polled fetuses. *HOXD8* had lower expression in the horn bud compared to frontal skin, forebrain skin and midbrain skin in horned fetuses, but the expression was similar in all tissues of polled fetuses. *HOXD1* had lower expression in the horn bud compared to forebrain skin in horned fetuses and in the horn bud and midbrain of horned fetuses compared to polled fetuses. Notably, HOX cofactors, *PBX3* and *MEIS2*, had increased expression in the horn bud region of horned fetuses compared to polled fetuses (Appendix Table C7). MTX2 was not differentially expressed between any tissues.

Two EMT transcription factors and four EMT markers were differentially expressed (Table 4.4). *TWIST2* had greater expression in the horn bud compared to the forebrain and midbrain skin of horned fetuses and in the horn bud and frontal skin of horned fetuses compared to polled fetuses. These differences were not observed in the same tissue comparisons in polled tissues. There are two *ZEB2* genes annotated in the Ensembl annotation (version 108) for cattle that are in the same region of BTA2, separated by 3,787 bp (date accessed: Feb 2022; Cunningham *et al.* 2022), and both were differentially expressed. *ZEB2* expression was higher in the horned versus polled fetuses for all tissues. One *ZEB2* gene (ENSBTAG00000048810) had greater expression in the horn bud compared with forebrain skin in horned. The *ZEB2* genes

139

had greater expression in the horn bud compared to midbrain skin in both horned and polled fetuses. *CDH2* expression was higher in the horn bud compared with forebrain skin and midbrain skin in horned fetuses and in the horn bud region compared to midbrain of polled fetuses. *CDH2* also had higher expression in frontal skin of horned fetuses compared to the frontal skin of polled fetuses. *VIM* expression was higher in all tissues from horned fetuses compared to polled fetuses. *TWIST1* and *FOXC2* were only differentially expressed for one comparison and *CHD1* and *OCLN* were not differentially expressed.

4.3.4 Functional analysis of differentially expressed genes

Genes differentially expressed in eight out of ten comparisons were associated with major pathways by GO and KEGG analysis (Appendix Table C11-Table C26). Due to low numbers of differentially expressed genes, no GO or KEGG pathways were identified for horn bud versus frontal skin for either genotypes, and therefore, these pathway analyses were undertaken manually.

4.3.4.1 GO pathway analysis of differentially express genes

Significant GO pathways were identified in comparisons between tissues (Table 4.5; Appendix Figure C4-Figure C7) and genotypes (Table 4.6; Appendix Figure C8-Figure C10). The pathways extracellular region, extracellular space, intermediate filament, and intermediate filament cytoskeleton were common between horn bud versus forebrain skin and horn bud versus midbrain skin for horned fetuses (Table 4.5). Pathways involving perception of senses were also common for these comparisons. Sensory perception was significant for horn bud versus forebrain skin and visual perception was significant for horn bud versus midbrain skin. Only extracellular region pathways were significant in the polled fetuses for horn bud versus forebrain skin and for horn bud versus midbrain skin.

When the horn bud region was compared between horned and polled fetuses (Table 4.6), the top five pathways identified were extracellular region, visual perception, sensory perception

of light stimulus, sensory perception, and intermediate filament. The first four of these pathways and extracellular space were over-represented in the comparison of frontal skin between genotypes.

Table 4.5: Top five significant GO terms for comparisons between tissues within horned and polled genotypes.

Comparison	# Significant pathways	Pathway (GO term)	Ontology	Total # Genes	# DE Genes	P-value	P-value adjusted
Horned comparisons							
Horned HB vs Horned FS	0	No pathways					
Horned HB vs Horned FB	13	Extracellular region (GO:0005576)	CC	546	192	1.30E-16	1.84E-12
		Extracellular space (GO:0005615)	CC	243	92	1.45E-10	1.03E-06
		Intermediate filament (GO:0005882)	CC	45	28	9.77E-10	4.61E-06
		Intermediate filament cytoskeleton (GO:0045111)	CC	49	28	1.47E-08	5.18E-05
		Sensory perception (GO:0007600)	BP	114	47	2.43E-07	0.0005
Horned HB vs Horned MB	11	Extracellular region (GO:0005576)	CC	546	221	4.89E-17	6.92E-13
		Intermediate filament (GO:0005882)	CC	45	32	5.75E-11	4.07E-07
		Extracellular space (GO:0005615)	CC	243	105	1.01E-10	4.77E-07
		Intermediate filament cytoskeleton (GO:0045111)	CC	49	33	2.86E-10	1.01E-06
		Visual perception (GO:0007601)	BP	63	63	1.26E-07	0.0003
Polled comparisons							
Polled HBR vs Polled FS	0	No pathways					
Polled HBR vs Polled FB	13	Extracellular region (GO:0005576)	CC	546	16	1.08E-07	0.002
		Platelet-derived growth factor binding (GO:0048407)	MF	4	3	7.65E-07	0.005
		Ossification (GO:0001503)	BP	80	6	7.19E-06	0.03
		Growth factor binding (GO:0019838)	MF	22	4	7.34E-06	0.03
		System development (GO:0048731)	BP	788	16	1.28E-05	0.03
Polled HBR vs Polled MB	36	Ossification (GO:0001503)	BP	80	11	5.94E-10	2.80E-06
		Biomineral tissue development (GO:0031214)	BP	44	9	5.68E-10	2.80E-06
		Biomineralization (GO:0110148)	BP	44	9	5.68E-10	2.80E-06
		Extracellular region (GO:0005576)	CC	546	24	2.29E-09	8.11E-06
		Tissue development (GO:0009888)	BP	331	16	3.53E-07	0.001

CC = cell component; BP = biological process; MF = molecular function; HB = horn bud; HBR = horn bud region; FS = frontal skin; FB = forebrain skin; MB = midbrain skin; DE = differentially expressed; the p-value for each term was adjusted by the FDR method.

Table 4.6: Top five significant GO terms for comparisons between horned and polled genotypes.

Comparison	# Significant pathways	Pathway (GO term)	Ontology	Total # Genes	# DE Genes	P-value	P-value adjusted
Horned HB vs Polled HBR	20	extracellular region (GO:0005576)	CC	546	240	1.38E-12	1.96E-08
		visual perception (GO:0007601)	BP	63	44	7.02E-11	3.31E-07
		sensory perception of light stimulus (GO:0050953)	BP	63	44	7.02E-11	3.31E-07
		sensory perception (GO:0007600)	BP	114	65	1.53E-09	5.43E-06
		intermediate filament (GO:0005882)	CC	45	33	2.30E-09	6.50E-06
Horned FS vs Polled FS	16	extracellular region (GO:0005576)	CC	546	225	1.85E-11	2.62E-07
		sensory perception (GO:0007600)	BP	114	58	2.27E-07	0.001
		visual perception (GO:0007601)	BP	63	37	3.56E-07	0.001
		sensory perception of light stimulus (GO:0050953)	BP	63	37	3.56E-07	0.001
		extracellular space (GO:0005615)	CC	243	105	2.97E-07	0.001
Horned FB vs Polled FB	2	extracellular region (GO:0005576)	CC	546	158	1.17E-09	1.65E-05
		extracellular space (GO:0005615)	CC	243	74	4.08E-06	0.03
Horned MB vs Polled MB	17	extracellular region (GO:0005576)	CC	546	122	1.17E-07	0.002
		visual perception (GO:0007601)	BP	63	25	5.54E-07	0.003
		sensory perception of light stimulus (GO:0050953)	BP	63	25	5.54E-07	0.003
		supramolecular complex (GO:0099080)	CC	287	70	2.46E-06	0.009
		multicellular organismal process (GO:0032501)	BP	1326	245	3.82E-06	0.01

CC = cell component; BP = biological process; MF = molecular function; HB = horn bud; HBR = horn bud region; FS = frontal skin; FB = forebrain skin; MB = midbrain skin; DE = differentially expressed; the p-value for each term was adjusted by the FDR method.

4.3.4.2 KEGG pathway analysis of differentially expressed genes

KEGG pathway analysis was carried out for the differentially expressed genes to identify over-represented pathways between tissues (Table 4.7) and genotypes (Table 4.8). Pathways neuroactive ligand-receptor interaction, *Staphylococcus aureus* infection, cytokine-cytokine receptor interaction, and steroid hormone biosynthesis were common between forebrain skin and midbrain versus horn bud in horned fetuses (Table 4.7). These pathways were not over-represented for the tissue comparisons in polled fetuses (Table 4.7). The top five most over-represented pathways for the horn bud region compared between horned versus polled fetuses were neuroactive ligand-receptor interaction, *Staphylococcus aureus* infection, steroid hormone biosynthesis, protein digestion and absorption, cytokine-cytokine receptor interaction (Table 4.8). Except for the pathway protein digestion and absorption, they were also over-represented for horned frontal skin versus polled frontal skin.

Table 4.7: Top five significant KEGG terms for comparisons between tissues within horned and polled genotypes.

Comparison	# Significant pathways	Pathway (KEGG ID)	# Total genes	# DE genes	P-value	P-value adjusted
		Horned comparisons				
Horned HB vs Horn FS	0	No pathways				
Horned HB vs Horn FB	44	Neuroactive ligand-receptor interaction (path:bta04080)	345	144	3.50E-20	1.17E-17
		Staphylococcus aureus infection (path:bta05150)	101	59	3.87E-17	6.44E-15
		Cytokine-cytokine receptor interaction (path:bta04060)	296	120	6.79E-16	7.54E-14
		Steroid hormone biosynthesis (path:bta00140)	70	42	3.70E-13	3.08E-11
		Olfactory transduction (path:bta04740)	983	279	2.04E-10	1.36E-08
Horned HB vs Horn MB	51	Neuroactive ligand-receptor interaction (path:bta04080)	345	175	2.64E-26	8.79E-24
		Staphylococcus aureus infection (path:bta05150)	101	68	1.03E-19	1.71E-17
		Cytokine-cytokine receptor interaction (path:bta04060)	296	135	1.23E-15	1.37E-13
		Steroid hormone biosynthesis (path:bta00140)	70	47	5.49E-14	4.57E-12
		Protein digestion and absorption (path:bta04974)	113	60	5.48E-11	3.04E-09
		Polled comparisons				
Polled HBR vs Polled FS	0	No pathways				
Polled HBR vs Polled FB	5	Alcoholism (path:bta05034)	195	17	1.39E-15	4.61E-13
		Systemic lupus erythematosus (path:bta05322)	150	14	2.01E-13	3.35E-11
		Viral carcinogenesis (path:bta05203)	213	15	1.66E-12	1.84E-10
		ECM-receptor interaction (path:bta04512)	80	6	7.16E-06	0.0005
		Protein digestion and absorption (path:bta04974)	113	5	0.0005	0.03
Polled HBR vs Polled MB	6	ECM-receptor interaction (path:bta04512)	80	11	5.90E-10	1.96E-07
		Protein digestion and absorption (path:bta04974)	113	10	2.65E-07	4.41E-05
		Focal adhesion (path:bta04510)	192	11	5.06E-06	0.0006
		PI3K-Akt signaling pathway (path:bta04151)	354	13	8.13E-05	0.007
		Human papillomavirus infection (path:bta05165)	329	12	0.0002	0.01

HB = horn bud; HBR = horn bud region; FS = frontal skin; FB = forebrain skin; MB = midbrain skin; DE = differentially expressed; the p-value for each term was adjusted by the FDR method.

Table 4.8: Top five significant KEGG terms for comparisons between horned and polled genotypes.

Comparison	# significant pathways	Pathway (KEGG ID)	# Total genes	# DE genes	P-value	P-value adjusted
Horned HB vs Polled HBR	51	Neuroactive ligand-receptor interaction (path:bta04080)	345	192	9.67E-24	3.22E-21
		Staphylococcus aureus infection (path:bta05150)	101	76	5.62E-21	9.35E-19
		Steroid hormone biosynthesis (path:bta00140)	70	52	2.17E-14	1.93E-12
		Protein digestion and absorption (path:bta04974)	113	73	2.32E-14	1.93E-12
		Cytokine-cytokine receptor interaction (path:bta04060)	296	150	3.96E-14	2.64E-12
Horned FS vs Polled FS	72	Neuroactive ligand-receptor interaction (path:bta04080)	345	188	2.76E-25	9.19E-23
		Staphylococcus aureus infection (path:bta05150)	101	71	1.44E-18	2.40E-16
		Steroid hormone biosynthesis (path:bta00140)	70	52	1.33E-15	1.48E-13
		Cytokine-cytokine receptor interaction (path:bta04060)	296	145	1.51E-14	1.26E-12
		Linoleic acid metabolism (path:bta00591)	33	28	1.72E-11	1.14E-09
Horned FB vs Polled FB	27	Systemic lupus erythematosus (path:bta05322)	150	66	5.55E-13	1.85E-10
		Protein digestion and absorption (path:bta04974)	113	50	2.74E-10	4.56E-08
		Alcoholism (path:bta05034)	195	71	2.60E-09	2.88E-07
		Linoleic acid metabolism (path:bta00591)	33	20	1.03E-07	8.56E-06
		Arachidonic acid metabolism (path:bta00590)	78	34	2.95E-07	1.96E-05
Horned MB vs Polled MB	7	Spliceosome (path:bta03040)	128	41	1.95E-07	3.25E-05
		Protein digestion and absorption (path:bta04974)	113	38	1.31E-07	3.25E-05
		Cytokine-cytokine receptor interaction (path:bta04060)	296	69	1.46E-05	0.002
		Antigen processing and presentation (path:bta04612)	80	25	7.12E-05	0.005
		Fat digestion and absorption (path:bta04975)	49	18	7.02E-05	0.005

HB = horn bud; HBR = horn bud region; FS = frontal skin; FB = forebrain skin; MB = midbrain skin; DE = differentially expressed; the p-value for each term was adjusted by the FDR method.

4.3.4.3 Manual pathway analysis of differentially expressed genes between horn bud and frontal skin

Ninety-seven genes were differentially expressed between the horn bud and frontal skin of horned fetuses, fifty-four genes with increased expression in the horn bud and 43 with decreased expression. These genes were manually grouped based on their functions and two major groups appeared: 1) neural development and function and 2) cell structure and adhesion (Table 4.9). In general, the neural genes with increased expression in the horn bud compared to the frontal skin had a role in axon guidance and development, while genes that had reduced expression had a direct role in nerve function (Appendix Table C27). The structural and adhesion genes encoded keratins, desmocolins, desmoglein and cadherin. In contrast, only two genes (*MID1IP1* and *ENSBTAG00000048627*) were differentially expressed between the horn bud region and frontal skin in the polled fetuses, and both had higher expression in the horn bud region.

Table 4.9: Major functions of genes differentially expressed between horn bud and frontal skin in horned fetuses.

Function	Expression in horn bud	#	Differentially expressed Genes
Nerve development and function	↑	10	*BDNF, CDH7, CNTFR, EPHA3, FRMD7, NTNG1, OSTN, RORB, TAFA1, TMEM59L*
	↓	23	*ARPP21, ATP10A, CBLN4, CDH15, FABP7, GPM6A, GRIA2, HCN1, LHX8, LMO1, LMO3, LRRC7, NEFL, NETO2, NEUROG2, NRG1, SCN3A, SGIP1, SHROOM3, SLAIN1, SLCO1A2, ST8SIA2, SYBU*
Cytoskeleton structure and cell adhesion	↑	10	*KRT14, KRT1, KRT6B, KRT6A, KRT16, KRTDAP, DSG1, DSC1, DSC3, CDH7*
	↓	2	*CDH15, MFAP5*

4.3.5 WNT signaling in horned and polled bovine fetuses

WNT genes had lower expression in the horn bud compared to the forebrain skin and midbrain skin of horned fetuses (Table 4.10). These genes also had lower expression in all tissues of the horned fetuses compared to the same regions in polled fetuses. *FRZ8* had higher expression in the horned horn bud compared to other tissues of the horned fetuses and to the horn bud region of polled fetuses. *CTNNB1* had higher expression in the horn bud and midbrain of horned fetuses compared to the same regions in polled fetuses. *LEF1* had increased expression in the horn bud compared to forebrain skin and midbrain skin of horned fetuses, and higher expression in the horn bud region and frontal skin of horned fetuses compared to the same regions in polled fetuses. WNT signaling regulators, *SFRP2* and *AXIN2,* had increased expression in the horn bud of horned fetuses compared with the horn bud region of polled fetuses. Overall, the expression of genes that encode WNT signaling members was affected by genotype (Figure 4.7).

4.3.6 BMP signaling in horned and polled bovine fetuses

The expression of core BMP signaling genes, which included ligands, receptors, *SMADs,* and BMP antagonists (*CHRD* and *TWISG1*), was assessed. The *BMPs* generally had lower expression in the horn bud compared to forebrain skin and midbrain skin in horned fetuses and in horned tissues compared to polled tissues. The exceptions were *BMP1*, *BMP3* and *BMP5* which had higher expression in some horned tissues (Table 4.11). *SMAD1* had increased expression in the horn bud compared to midbrain skin in horned fetuses, and had increased expression the horn bud region and frontal skin of horned fetuses compared to polled fetuses.
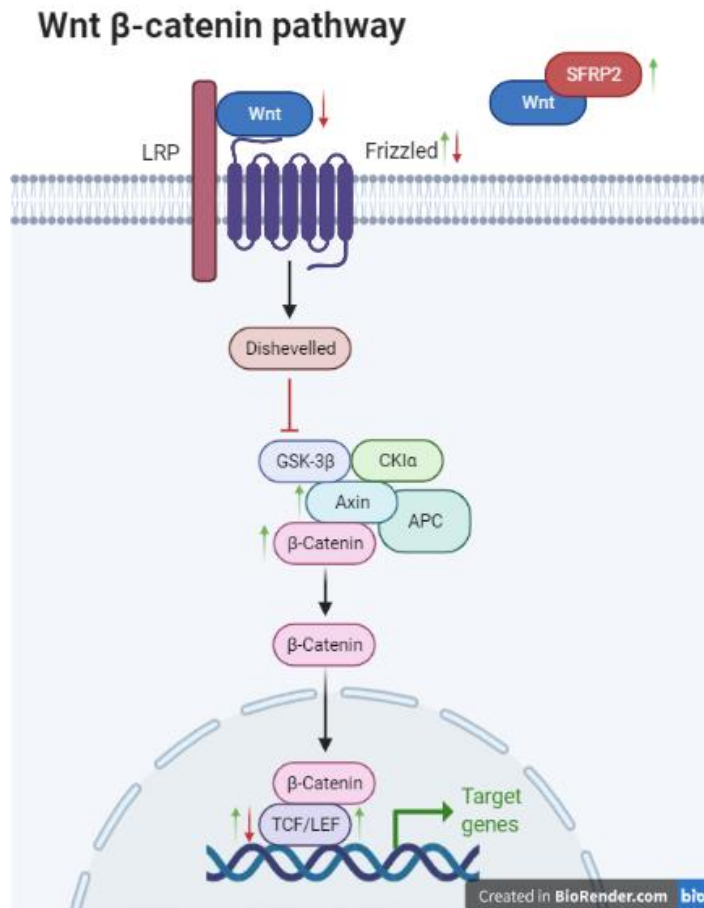
Figure 4.7: Differential expression of genes encoding members of WNT signaling in the horn bud of horned fetuses compared to the horn bud region of polled fetuses. The arrows indicate the direction of expression in the horn bud. Created in Biorender <https://biorender.com/>.

4.3.7 HGF/MET signaling in horned and polled bovine fetuses

Genes involved in HGF/MET signaling and downstream pathways were not differentially expressed nor were the genes involved in their downstream pathways (Appendix Table C28).

Table 4.10: Differential expression of WNT signaling genes in horned and polled bovine fetuses at 58 days of development.

WNT signaling genes	Horn Bud vs Frontal Skin		Horn Bud vs Forebrain Skin		Horn Bud vs Midbrain Skin		Horn vs Polled			
	HB vs FS	HB vs FS	HB vs FB	HB vs FB	HB vs MB	HB vs FB	HB vs HBR	FS vs FS	FB vs FB	MB vs MB
WNT	-	-	WNT2 ↓ WNT3A ↓ WNT8A ↓ WNT16 ↓	-	WNT2 ↓ WNT3A ↓ WNT7A ↓ WNT8A ↓ WNT16 ↓	-	WNT1 ↓ WNT2 ↓ WNT3 ↓ WNT3A ↓ WNT7A ↓ WNT7B ↓ WNT8A ↓ WNT9B ↓ WNT10A ↓ WNT10B ↓ WNT16 ↓	WNT1 ↓ WNT3 ↓ WNT3A ↓ WNT5B ↓ WNT6 ↓ WNT7A ↓ WNT7B ↓ WNT8A ↓ WNT9B ↓ WNT10A ↓ WNT10B ↓	WNT3 ↓ WNT5B ↓ WNT6 ↓ WNT8A ↓ WNT10B ↓ WNT16 ↑	WNT1 ↓ WNT7A ↓ WNT7B ↓ WNT8A ↓ WNT10A ↓ WNT10B ↓
FRZ	FZD8 ↑	-	FZD8 ↑	-	FZD5 ↑ FZD8 ↑	FZD5 ↑	FZD1 ↑ FZD8 ↑ FZD9 ↓ FZD10 ↑	-	FZD1 ↑	FZD1 ↑
CTNNB1	-	-	-	-	-	-	↑	-	-	↑
TCF	-	-	TCF7 ↑ TCF23 ↓	-	TCF7 ↑ TCF7L1 ↑ TCF21 ↓ TCF23 ↓	-	TCF4 ↑ TCF12 ↑ TCF15 ↓ TCF21 ↓ TCF23 ↓	TCF4 ↑ TCF7L1 ↑ TCF7L2 ↑ TCF12 ↑ TCF15 ↓ TCF21 ↓ TCF23 ↓	TCF12 ↑ TCF15 ↓	TCF12 ↑ TCF15 ↓ TCF23 ↓
LEF1	-	-	↑	-	↑	-	↑	↑	-	-
DVL	-	-	DVL2 ↑ DVL3 ↑	-	DVL3 ↑	-	-	-	-	-
AXIN2	-	-	↑	-	-	-	↑	-	-	-
SFRP2	↑	-	-	-	-	-	↑	↑	↑	↑

HB = horn bud; HBR = horn bud region; FS = frontal skin; FB = forebrain skin; MB = midbrain skin; ▮ = homozygous horned; ▮ = homozygous polled; ↑ increased expression in horn bud or horned tissue for other comparisons; ↓ decreased expression in horn bud or horned tissue for other comparisons; - not differentially expressed

150

Table 4.11: Differential expression of BMP signaling genes in horned and polled bovine fetuses at 58 days of development.

BMP signaling genes	Horn Bud vs Frontal Skin		Horn Bud vs Forebrain Skin		Horn Bud vs Midbrain Skin		Horned vs Polled			
	HB vs FS	HB vs FS	HB vs FB	HB vs FB	HB vs MB	HB vs FB	HB vs HB	FS vs FS	FB vs FB	MB vs MB
BMP	-	-	BMP1↑ BMP10↓ BMP15↓	-	BMP1↑ BMP3↓ BMP10↓ BMP15↓	-	BMP8A↓ BMP10↓ BMP15↓	BMP5↑ BMP8A↓ BMP10↓	BMP7↓ BMP8A↓ BMP10↓	BMP3↑ BMP8A↓
BMPR	-	-	-	-	-	-	-	BMPR2↑	BMPR2↑	BMPR2↑
ACVRL	-	-	-	-	-	-	-	-	-	-
R-SMADs (SMAD1/5/8)	-	-	-	-	SMAD1↑	-	SMAD1↑	SMAD1↑	-	SMAD5↑
Co-SMAD (SMAD4)	-	-	-	-	-	-	-	-	-	SMAD4↑
I-SMAD (SMAD6/7)	-	-	SMAD6↑ SMAD7↑	-	SMAD6↑	-	-	-	-	-
CHRD	-	-	-	-	-	-	-	-	-	-
TWSG1	-	-	-	-	-	-	-	-	↑	↑

HB = horn bud; FS = frontal skin; FB = forebrain; MB = midbrain; ▮ = homozygous horned; ▮ = homozygous polled; ↑ increased expression in horn bud or horned tissue for other comparisons; ↓ decreased expression in horn bud or horned tissue for other comparisons; - not differentially expressed

4.4 Discussion

The genetic pathways involved in bovine horn development are not known and the effect of POLLED variants on gene expression are not understood. This transcriptomic study of four horned and three polled fetuses tested whether the genes located within the POLLED region and candidate genes are involved in horn development at day 58.

Studies to date have examined horn buds from day 70 fetuses (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014), whereas the fetuses studied in this research were sampled at day 58. Day 58 was expected to be before the horn buds would be present. However, the horn bud was observed in the horned fetuses, and therefore, further studies on horn ontogeny should consider fetuses from 50 days to 55 days. Nevertheless, the analyses of the 58 day old fetuses contributed new information to our overall understanding of horn bud development (Chapter 5).

4.4.1 Genes in the POLLED region are differentially expressed

As the Celtic variant is intergenic, it may affect gene expression of nearby genes in the horn bud. Three nearby genes and one lincRNA (*C1H21orf62, SON, EVA1C* and *LOC112447120*) were differentially expressed in the horn bud between horned and polled fetuses, which indicates that their expression may be affected by the Celtic POLLED variant.

C1H21orf62 had decreased expression in the horn bud of horned fetuses compared to the horn bud region in polled fetuses at 58 days. In a previous RNAseq study, *C1H21orf62* was found to have lower expression in the horn bud of a horned fetus compared to a polled fetus at ~5 months of development although this was not validated by qPCR analysis (Wiedemar *et al.* 2014). However, in another qPCR study, the expression of *C1H21orf62* was the same in the horn bud region of horned and polled bovine fetuses at 90 days of development (Allais-Bonnet *et al.* 2013).

Also, in the present study, *C1H21orf62* expression in horned and polled fetuses was the same between the horn bud and frontal skin at 58 days of development. Similarly, Allais-Bonnet *et al.* (2013) did not find a difference in *C1H21orf62* expression between horn bud and frontal skin of horned or polled bovine fetuses at day 90 of development. However, Wiedemar *et al.* (2014) found through qPCR analysis that the expression of *C1H21orf62* had a trend towards being lower in the horn bud region compared to frontal skin in both horned and polled fetuses at ~70-175 days of development. The stage of development and tissues being compared (horned versus polled, horn bud versus frontal skin) may explain the differences in these results. A replicate study at similar and slightly different fetal ages is required to validate these results.

Unfortunately, the function of *C1H21orf62* is not known, although based on the predicted structure, it may be a secreted protein (further described in Appendix Section C5; Appendix; Figure C12A). The human orthologue is expressed in the brain and reproductive organs. As *C1H21orf62* expression is greater in polled horn bud region and frontal skin, the expression possibly has an inhibitory effect on horn bud formation.

SON had increased expression in the horn bud of horned fetuses compared to the horn bud region in polled fetuses. In horned fetuses, *SON* expression was the same between the horn bud and other tissues. *SON* has only been assessed in one other RNAseq study where it was not differentially expressed in the horn bud compared to frontal skin of sheep fetuses (Wang *et al.* 2019c), in agreement with the current study. SON is an mRNA splicing cofactor and is widely expressed in human tissues. SON has been shown to mediate splicing of genes critical for neuronal migration (Kim et al. 2016). SON also has a role in mitosis, as shown by RNA interference studies (Hickey *et al.* 2014). Therefore, *SON* expression in the horn bud may be required for increased or altered mRNA splicing during neural development or for mitosis, and reduced expression of *SON* in polled fetuses could halt development or migration of cells required for horn development.

EVA1C had decreased expression in the horn bud of horned fetuses compared to the horn bud region in polled fetuses. Differential expression of *EVA1C* has not been detected in previous gene expression studies (Wiedemar *et al.* 2014; Wang *et al.* 2019c). In mice, Eva1c has been shown to be involved is Slit-mediated axon guidance (James et al. 2013). This protein has been found in the axon shafts of neurons, including the dorsal root ganglia, which are a collection of sensory neuronal cell bodies and olfactory sensory axons (James et al. 2013). Furthermore, EVA-1 (EVA1C orthologue) in *C. elegans* had migration inhibitory activity in neuroblasts (Rella *et al.* 2021). The reduced expression of *EVA1C* may allow cells and axons to migrate freely into the horn bud. An increased expression of *EVA1C* in axons or migratory cells may inhibit outgrowth or migration towards the horn bud region of polled fetuses. The localisation of EVA1C and SLIT in horned and polled fetuses may elucidate how these proteins affect horn bud formation.

LOC112447120 had decreased expression in the horn bud of horned fetuses compared to the horn bud region in polled fetuses. Differential expression of *LOC112447120* has not been detected in previous gene expression studies (Wiedemar *et al.* 2014; Wang *et al.* 2019c). *LOC112447120* is a lincRNA located between *IFNGR2* and *IFNAR1* on BTA1. The sequence of *LOC112447120* is highly conserved in water buffalo, sheep and goat, but its function is unknown (Chapter 2). It can be only hypothesised that *LOC112447120* regulates gene expression.

These data suggest that the Celtic POLLED variant may increase the expression of *C1H21orf62*, *EVA1C* and *LOC112447120*, and decrease the expression of *SON*. The changes in gene expression near the Celtic POLLED variant could affect horn formation by inhibiting nerve development or migration of cells by increased expression of *EVA1C* or by inhibition of mitosis via down-regulation of *SON*. As *C1H21orf62*, *EVA1C* and *LOC112447120* are up-regulated in polled horn bud, they may act as inhibitors of horn development. Unfortunately,

the functions of *C1H21orf62* and *LOC112447120* are unknown so this hypothesis is difficult to test.

These effects of the Celtic variant on gene expression could be a result of changing chromatin interactions, such as those between an enhancer and promotor. A gene editing study has shown that the 10 bp deletion without the 212 bp duplication in the Celtic variant does not cause individuals to be polled (Hennig *et al.* 2022a). However, the 212 bp duplication of a regulatory region in the Celtic variant may alter the expression of nearby genes. Mice with a duplicated enhancer near the transcription factor TBX15, which is important for limb development, has been associated with higher expression of this gene and polydactyly of the hind limbs (Flöttmann *et al.* 2018). Similarly, in humans, a ~5.5 kb duplication, including regulatory elements ~110 kb upstream of *BMP2,* was found to cause phalange malformations in a syndrome known as autosomal-dominant brachydactyly type A2 (Dathe *et al.* 2009). Dathe *et al.* (2009) did not measure *BMP2* expression, however, because conditional ablation of *BMP2* does not result in any limb phenotypes in mice, the authors concluded that the 5.5 kb duplication in humans most likely causes the malformations by increasing *BMP2* expression. Therefore, there is precedence for duplications of regulatory elements increasing gene expression and causing phenotypic changes, and the Celtic POLLED variant might function via this mechanism.

It is recognised that more samples per genotype ($n = 6$) would provide greater evidence for differential expression and may indicate which of these genes (*C1H21orf62*, *SON*, *EVA1C* and *LOC112447120*) is most important for horn development or polledness. However, the sample numbers herein are still improved compared to previous transcriptomic studies of horn development in cattle ($n = 1$) (Wiedemar *et al.* 2014) and sheep ($n = 2$) (Wang *et al.* 2019c). These small sample numbers reflect the difficulties in obtaining bovine fetal samples which are often opportunistic if collected from abattoirs or financially costly when generated as for the experiment herein.

4.4.2 Horn development candidate genes may be differentially expressed

RXFP2 had increased expression in the horn bud of horned fetuses compared to the horn bud region of polled fetuses at day 58. Previous gene expression studies on bovine fetuses also reported increased expression of *RXFP2* in the horn bud compared to the horn bud region between horned and polled fetuses (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014). In addition, *RXFP2* had increased expression in the horn bud compared to the frontal skin of horned fetuses at day 58. This result agrees with expression studies of bovine fetuses from ~70-175 days and 90 dpc ovine fetuses (Wiedemar *et al.* 2014; Wang *et al.* 2019c). However, in another study, *RXFP2* expression was not significantly decreased between in the frontal skin compared to the horn bud in horned bovine fetuses at 90 dpc (Allais-Bonnet *et al.* 2013). *RXFP2* can be up-regulated by SOX9 in mice (Feng *et al.* 2007; Feng *et al.* 2009). Expression of *SOX9* was also increased in the horn bud compared with frontal skin herein, suggesting that *SOX9* may affect *RXFP2* expression in the horn bud. RXFP2 is localised in the epithelium and nerves in the horn bud at 58 days of development (Chapter 3), and therefore, *RXFP2* is may be involved in development of these tissues in the horn bud. Nevertheless, the connection between the POLLED region and *RXFP2* is still not known.

The two bovine *ZEB2* genes had increased expression in all tissues from the horned fetuses compared to tissues from the polled fetuses. *ZEB2* was not identified as being differentially expressed in previous gene expression studies on bovine and ovine fetuses (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014; Wang *et al.* 2019c). Gene expression may differ depending on time-points of fetal sample collection. Two different variants, a deletion and a frameshift mutation, in *ZEB2* cause a polled phenotype in cattle. The *ZEB2* deletion also causes an abnormal skull shape, infertility in females and neurological disorders, among other symptoms (Capitan *et al.* 2012). The frameshift mutation in *ZEB2* causes abnormal skull shape, small body stature and sub-fertility in cattle (Gehrke *et al.* 2020b). In mice, the protein encoded by *Zeb2* (also referred to as *Zfhx1b*), Sip1, is expressed in pre-migratory and migrating cranial

neural crest cells (Van De Putte *et al.* 2003). Furthermore, it has been shown that cranial neural crest cells fail to migrate in mice when there is a homozygous deletion of exon 7 in *Zeb2* (Van De Putte *et al.* 2003). *Zeb2* has been also shown to promote immature Schwann cell differentiation and peripheral myelination in mice (Quintes *et al.* 2016). Therefore, ZEB2 may be required for cell migration and the development of Schwann cells and nerves in horned fetuses. Reduced expression in polled fetuses may indicate poor migration and arrested development of immature Schwann cells and nerves.

TWIST2 had higher expression in the horn bud and frontal skin of horned fetuses compared to polled fetuses at day 58. Differential expression of *TWIST2* has not been observed in association with horn development in previous gene expression studies (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014; Wang *et al.* 2019c). TWIST2 is a transcription factor that can activate and repress transcription (Franco *et al.* 2011). *TWIST2* was originally postulated as a candidate gene affecting horn development as mutations within *TWIST2* can cause distichiasis, a condition where eyelashes arise from the inner eyelid margin (Allais-Bonnet *et al.* 2013). A similar phenotype is found in polled cattle where additional eyelashes arise on the eyelid. Homozygous nonsense mutations in *TWIST2* are known to cause Setleis syndrome in humans (Franco *et al.* 2011). This syndrome is a form of focal facial dermal dysplasia characterized by bitemporal scar-like lesions with the absence of lower eyelashes and multiple rows of eyelashes on the upper eyelid, among other features (Tukel *et al.* 2010). In zebrafish, *twist2* is expressed by dermal fibroblasts during formation of scales, a type of skin appendage (Jacob *et al.* 2021). Within the promoter region of *twist2*, multiple LEF1 and TCF binding sites have been identified, which suggests that *twist2* expression, in zebrafish, can be regulated by WNT signaling (Jacob *et al.* 2021). If this genetic pathway is also involved in horn development, then WNT signaling may regulate *TWIST2* expression in the horn bud.

The expression of *HAND1* was decreased in horned fetuses compared with polled fetuses for the horn bud, frontal skin and midbrain skin at day 58. *HAND1* has not been detected

as differentially expressed in previous RNAseq studies of horn development (Wiedemar *et al.* 2014; Wang *et al.* 2019c). A putative HAND1 enhancer binding site has been predicted to overlap the 10-bp deletion site of the Celtic POLLED variant. Given that the Celtic POLLED variant leads to greater *HAND1* expression in polled fetuses, the significant involvement of HAND1 in horn development is unlikely. This is supported by a gene editing study that showed the horn bud still formed when the 10 bp deletion without the Celtic 212 bp duplication was edited into horned bovine fetuses (Hennig *et al.* 2022a). That is, the partial removal of the HAND1 predicted binding site did not cause polledness. Of note, *HAND1* encodes a bHLH transcription factor which is involved in the development of several systems during embryogenesis, including limb and osteogenic systems (Barnes *et al.* 2010; Funato *et al.* 2020). Therefore, *HAND1* may still have a role in the development of other tissues in the cranial regions, such as the cranial vault.

This study provides transcriptomic evidence that *RXFP2*, *ZEB2* and *TWIST2* are involved in early horn development. However, the exact function of these genes and their role in the horn developmental pathway are not understood. Other than *RXFP2*, it is not known which cells express these genes, and therefore, localisation of *ZEB2*, *TWIST2* and *HAND1* gene expression would contribute to understanding their role in horn development. *ZEB2* and *TWIST2* are involved in epithelial-mesenchymal transitions, which are also of interest (discussed in Section 4.4.7).

4.4.3 Neuronal, cytoskeletal and extracellular gene products affected by horn development

GO and KEGG pathway analyses of the differentially expressed genes and manual analysis of the 97 differentially expressed genes between the horn bud and frontal skin of horned fetuses identified neuronal, cell structure and adhesion functions as being in common between the analyses. The sensory nerve-related pathways were enriched for horned horn bud versus polled horn bud region and for horned frontal skin versus polled frontal skin. The

majority of the genes within these nerve-related pathways had lower expression in the horned tissues. This was unexpected as histological studies show that there is greater innervation at the horn bud of horned fetuses (Chapter 3; Wiener *et al.* 2015). An explanation may be that there are morphological differences between the innervation observed at the horn bud and frontal skin of horned fetuses compared to polled fetuses. This is supported by the histological analyses that showed that the developing nerves were larger in the horn bud of horned fetuses compared to the horn bud region and frontal skin of polled fetuses (Chapter 3). Nerve growth in the epithelium was also observed in the horn bud of horned fetuses and not in the polled tissues. The mesenchyme and epithelial cells at the horn bud may also express genes to direct nerve growth towards the epithelium, as has been observed in tooth development in the mouse (Luukko & Kettunen 2014).

Neural pathways were also enriched in analyses of headgear-specific genes, differentially expressed genes in sheep fetal horn buds, and differentially abundant horn bud proteins from yak fetuses (Li *et al.* 2018; Wang *et al.* 2019c). Headgear-specific genes were identified by analysing data from sheep, goats and deer and were highly expressed in horn or antler tissue (further described in Appendix Section C6) (Wang *et al.* 2019c). The pathway analyses of these headgear specific genes revealed enrichment of the GO pathways sensory perception, visual perception and sensory perception of light stimulus (Wang *et al.* 2019c). The KEGG pathway of neuroactive ligand-receptor interaction was not enriched (Wang *et al.* 2019c). Pathway analyses of differentially expressed genes between horn bud and frontal skin of the sheep at 90 dpc revealed enrichment of the KEGG pathway neuroactive ligand-receptor interaction, but none of the GO pathways were enriched (Wang *et al.* 2019c). Yak horn bud tissue from three horned and three polled fetuses at ~80-90 dpc was isolated and differentially abundant proteins were identified (Li *et al.* 2018). Highly abundant proteins were enriched for GO pathway neuron projection (GO term not provided), but no other pathways were identified.

The bovine horn is innervated by the cornual branch of the trigeminal nerve. The trigeminal nerve also branches to the inferior alveolar nerve which innervates the mouse molar (Luukko & Kettunen 2014). The regulation of nerve growth to the tooth bud occurs through signals in the mesenchyme and epithelium, and the signals to promote or dampen nerve growth occurs depending on the stage of development (Luukko & Kettunen 2014). A similar mechanism might be involved to promote nerve growth to the horn bud region.

Genes involved in axon guidance include neutropic factors (*NGF*, *GDNF*), semaphorins (*SEMA*)/plexin (*PLXN*), neurophilin, ephrin ligands (*EFN*)/Eph receptors (*EPH*), slit guidance ligand (SLIT)/roundabout guidance receptor (ROBO), and cell adhesion molecules, such as cadherins (*CDH*), netrins (*NTR*), and laminins (*LAMB*) (Thiede-Stan & Schwab 2015; Tong *et al.* 2019). Proteins from these gene families were differentially expressed in the horn bud in this study (Appendix Table C29). Furthermore, many of the axon guidance proteins are also involved in guiding cranial neural crest cell migration (Kulesa *et al.* 2010; Theveneau & Mayor 2012). Further scrutiny of the expression of the axon guidance genes may reveal their involvement in horn development.

Components of the cytoskeleton were enriched for horned horn bud versus polled horn bud region and for horned frontal skin versus polled frontal skin. This is possibly because the epidermis in the horn bud is differentiated (Chapter 3; Allais-Bonnet *et al.* 2013; Wiener *et al.* 2015) and has a different expression profile from the frontal skin tissue that has only one layer of epidermal cells (Chapter 3). The increased mechanical strength of the horn bud was evident in the histological studies (Chapter 3), where all six horn bud samples from horned fetuses were intact after histological processing, whereas the polled horn bud regions tended to fragment. Components of the cytoskeleton were enriched in the pathway analyses of headgear specific genes (Appendix Section C6), differentially expressed genes between horn bud and frontal skin of horned sheep fetuses, and differentially abundant proteins between horn bud of horned and polled yak fetuses described above (Li *et al.* 2018; Wang *et al.* 2019c). Headgear specific genes

and differentially expressed genes from sheep fetuses were enriched for the GO pathways intermediate filament (GO:0005882) and intermediate filament cytoskeleton (GO:0045111) (Wang *et al.* 2019c). The KEGG pathway *Staphylococcus aureus* infection (path:bta05150), which includes keratins, was enriched in the fetal ovine differentially expressed genes (Wang *et al.* 2019c). The differentially abundant proteins from yak fetuses were enriched in GO pathways that included intermediate filament, structural constituent of the cytoskeleton and focal adhesion, and the KEGG pathways tight junction and focal adhesion (Li *et al.* 2018).

Horn development also appears to be affected by the extracellular environment which was indicated by enrichment of GO terms extracellular region and extracellular space and KEGG term cytokine-cytokine receptor interaction herein. These GO pathways were enriched in head gear-specific genes and differentially expressed between horn bud and frontal skin from sheep fetuses (Wang et al. 2019b). The GO term extracellular matrix was enriched for differentially abundant proteins between the horn buds from horned and polled yak fetuses (Li et al. 2018). The KEGG pathway, cytokine-cytokine receptor interaction, was also enriched for differentially expressed genes between horn bud and frontal skin from sheep fetuses (Wang *et al.* 2019c). Proteins in the GO and KEGG pathways are part of the extracellular matrix, such as collagens and cadherin, or are signalling molecules (e.g. WNTs, BMPs, cytokines) and their receptors. It seems that the extracellular environment of the horn bud in horned fetuses is different from the environment of the horn bud region in polled fetuses.

4.4.4 WNT genes are differentially regulated in the horn bud

WNTs are secreted proteins that mediate cell-cell communication over a short distance, usually acting on nearby cells (Wiese et al. 2018). WNT signalling is involved in many processes, including cell fate specification, proliferation, 3D organization, migration, and differentiation (Wiese et al. 2018). There is evidence that WNT signalling can be activated by RXFP2 (Johnson *et al.* 2010) and the present study showed that expression of components of WNT signalling pathway are altered in the horn bud.

Some WNT signalling pathway genes (*FZD8, CTNNB1*, and *LEF1)* had increased expression in the horn bud of horned fetuses compared to the same region in polled fetuses. FZD8 has been shown to increase proliferation of human keratinocyte cells (Shen *et al.* 2017). *LEF* genes encode transcription factors that enhance transcription when β-catenin is located in the nucleus from the cytoplasm. *LEF1* is involved in ectodermal organ development (Van Genderen *et al.* 1994; Sasaki *et al.* 2005; Hermans *et al.* 2021). *Lef1*-knockout mice do not develop teeth, hair or mammary glands (Van Genderen *et al.* 1994). While *LEF1* had increased expression in the horn bud, its function depends on the nuclear location of β-catenin which is controlled by WNT signalling. In the absence of β-catenin in the nucleus to co-activate transcription, LEF transcription factors act as repressors (Kestler & Kuhl 2008). In mice, *WNT* genes are reported to be expressed in the epithelium and mesenchyme during normal early tooth development (Sarkar & Sharpe 1999; Hermans *et al.* 2021). *LEF1* is expressed in the dental lamina but later becomes more highly expressed in the mesenchyme during the cap and bell stages of tooth development (Figure 4.8) (Sasaki *et al.* 2005). However, in the present study, other genes that encode WNT signaling molecules had decreased expression in the horned fetuses compared to polled fetuses for all tissues.
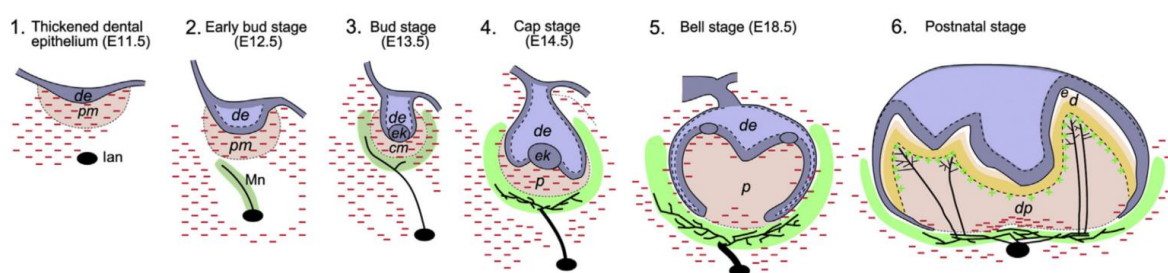

Figure 4.8: Stages of tooth development. Green colour indicates the mesenchymal dental axon pathway. Abbreviations: cm, condensed dental mesenchyme; d, dentin; de, dental epithelium; dp, dental pulp; e, enamel; ek, enamel knot; Ian, inferior alveolar nerve; p, dental papilla; Mn, molar nerve; pm, presumptive dental mesenchyme. Nerve fibers are indicated in black. Figure from Luukko and Kettunen (2014).

The changes in the expression of the WNT signaling genes suggest that regulation of the pathway may be different for the horn bud compared to surrounding tissues (Figure 4.7). The increased expression of some WNT pathway genes in the horn bud (*FRZ8, CTNNB1, LEF1,* and *AXIN2*) suggest the activation of WNT even if some *WNT* genes had reduced expression. Increased expression of *FRZ8* may confer greater receptivity of some cells to WNT signaling. WNT signaling activates expression of *AXIN2,* which is a negative regulator of WNT, and thus creates a negative feedback loop (Jho *et al.* 2002; Leung *et al.* 2002). *AXIN2* expression suggests that WNT signaling does occur in the horn bud. Together, these results show that the expression of WNT pathway genes is dependent on the tissue type and genotype. Further characterisation of the WNT pathway in the horn bud would contribute to the understanding of the role of WNT signaling in development.

4.4.5 HOXD cofactors are upregulated in the horn bud.

HOXD10 had lower expression in the horn bud of horned fetuses compared to forebrain skin and midbrain skin and had lower expression in all horned tissues compared to the polled tissues. Another HOXD gene, *HOXD1,* also had lower expression in the horn bud compared to forebrain skin in horned fetuses, and when the horn bud and midbrain skin were compared to the same tissues of polled fetuses. *HOXD8* expression was lower in the horn bud than frontal skin and forebrain skin in horned fetuses. *HOX* gene expression is tightly regulated by developmental stage and position of tissue on the anterior-posterior axis of the fetus (Figure 4.8) (Luo *et al.* 2019). However, the differences in expression cannot be explained by the general expression patterns of HOX genes. Considering that there was no differential expression of the HOXD genes among the polled tissues, the results suggest that *HOXD1, HOXD10* and *HOX8* expression is affected by the Celtic variant.

The HOX transcriptional activity is mediated by PBX and MEIS cofactors which provide temporal and spatial specificity (Mann *et al.* 2009). Between horned and polled fetuses,

PBX3 and *MEIS2* had increased expression in the horn bud of horned fetuses. In mice, Pbx/Meis cooperatively bind with Hox to initiate transcription (Mann *et al.* 2009). Interestingly, Meis2 is involved with limb development and double knockout of *Meis1* and *Meis2* results in down-regulation of *Lef1* expression (Delgado *et al.* 2021). A Meis2 binding site upstream of *Lef1* indicates that transcription of *Lef1* is directly regulated by Meis2 (Delgado *et al.* 2021). *PBX3* and *MEIS2* had higher expression in the horn bud of horned fetuses compared to polled fetuses. In the horn bud, the increased expression of *MEIS2* may explain the increased expression of *LEF1* that was also observed. As MEIS and PBX are cofactors for all *HOX* genes, the involvement of other HOX clusters (HOXA, HOXB and HOXC) in horn development cannot be ruled out.
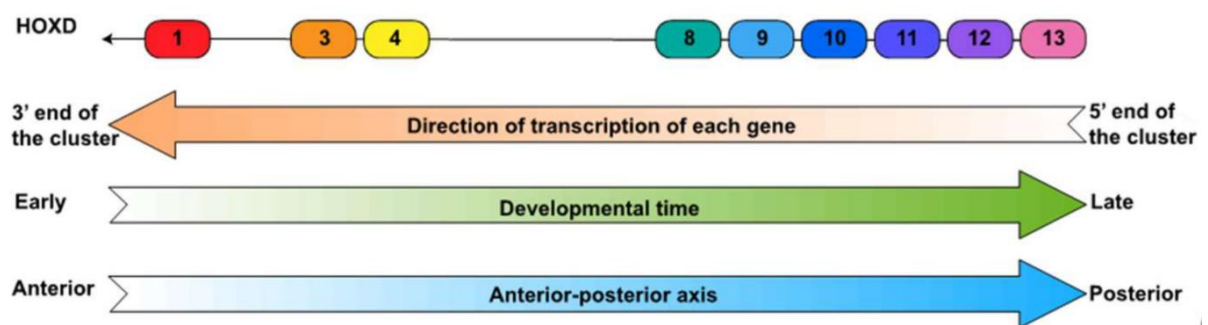

Figure 4.9: HOXD loci and developmental transcription. Figure modified from Luo et al. (2019).

It is important to note that HOX cofactors also act independently of *HOX* genes as not all phenotypes associated with HOX mutations and PBX/MEIS ortholog mutations in *Drosophila* and *Caenorhabditis elegans* overlap (Moens & Selleri 2006). PBX and MEIS can act as co-factors for other transcription factors, such as orphan HOX proteins and non-HOX homeodomain proteins (Moens & Selleri 2006). Therefore, the role of MEIS2 and PBX3 in the horn bud could be independent of the HOX genes.

4.4.6 BMP genes are differentially regulated in the horn bud

Genes encoding canonical BMP signalling pathway proteins had increased expression in the horn bud of horned fetuses. Genes that encode BMP ligands, receptor regulated SMADS (R-SMADS) and inhibitor SMADS (I-SMADS) were differentially expressed in the horn bud. *BMP1* had increased expression in the horn bud compared with forebrain skin and midbrain skin of horned fetuses. Despite the gene name, *BMP1* is not a BMP signaling molecule but a metalloproteinase that is involved in the proteolytic maturation of extracellular proteins such as collagens (Ge & Greenspan 2006). *R-SMAD* and *SMAD1* had higher expression in the horn bud of horned fetuses compared to the horn bud region of polled fetuses. In mice, conditional knock-outs of *SMAD1* in chondrocytes and osteocytes lead to reduced mineralization in the formation of the skull (Wang *et al.* 2011).

BMP signalling has a role in cranial development as it is required for cranial neural crest cell survival and migration to the facial primordia (Solloway & Robertson 1999; Nie *et al.* 2006). For example, BMP ligands are expressed in the brachial arches (Bennett *et al.* 1995; Solloway & Robertson 1999). The brachial arches are the regions where the cranial neural crest cells migrate to after the first wave of delamination. The brachial arches were under-developed due to cell death when *Bmp5/Bmp7* were double knocked-out in mice (Solloway & Robertson 1999). Given that BMP signalling is required for neural crest cell growth and craniofacial development, and can be activated by the INSL3/RXFP2 pathway, it is plausible that BMP signalling is also required for horn development. However, the nature of BMP signalling in the horn bud needs to be investigated further as the data are limited. Such studies should investigate the localization of the BMP signalling proteins as well as their expression.

4.4.7 EMT transcription factors and markers are upregulated in the horn bud at 58 days of development

TWIST1, *TWIST2*, *ZEB2* and *FOXC2* are transcription factors involved with epithelial-to-mesenchymal transitions (EMT). Other markers of EMT are E-cadherin (*CDH1*), N-cadherin (*CDH2*), occludin (*OCLN*) and vimentin (*VIM*). *TWIST2, ZEB2*, *CDH2* and *VIM* had increased expression in the horn bud, and their expression difference is consistent with the occurrence of EMT. However, *TWIST1*, *FOXC2*, *CDH1* and *OCLN* were not differentially expressed. ZEB proteins are repressors of the epithelial phenotype, whereas the TWIST proteins are inducers of the mesenchymal state (Migault *et al.* 2022). Together, the expression of these genes suggests that EMT may occur in the horn bud. However, this cannot be confirmed without identifying the cell types expressing *TWIST2, ZEB2*, *CDH2* and *VIM*.

EMT is a key developmental process that promotes the diversification of cells within an embryo (Kalcheim 2015). EMT allows neural crest cells to delaminate and migrate (Betancur *et al.* 2010). EMT is also required for the formation of mesenchymal sclerotomes from the ventral epithelium of somites (Kalcheim 2015). The sclerotome subsequently gives rise to the vertebral column and ribs of the skeleton (Christ *et al.* 2004; Tani *et al.* 2020). Delamination of neural crest cells and the formation of sclerotomes via EMT occurs earlier in development than observed in this study and in different regions. In the horn bud at day 58 of development, there is no clear role for EMT beyond diversification of the cells.

4.2.8 The frontal skin as a control tissue

There were 97 DE genes between horn bud and frontal skin suggesting the tissues are similar. In comparison, there were 5228 DE genes between the horn bud and forebrain skin, despite the forebrain skin being closer to the horn bud than frontal skin. Furthermore, when the tissues were compared between genotype, 75% of the DE genes for the horn bud region were in common with the frontal skin but only 45% DE genes were common with the forebrain skin.

If the horn bud and frontal skin have similar gene expression, then frontal skin may not be the best control tissue for horn development studies because genes important for horn development may have similar expression levels in the horn bud and frontal skin. However, there was a high degree of variation between the frontal skin samples from horned fetuses (Figure 4.4), which may lead to fewer differentially expressed genes being detected.

Gene expression in the horn bud and frontal skin may be similar because structures involved in horn development, such as nerves, are guided through the frontal skin towards to the horn bud. Neural crest cells, which may differentiate to progenitors of the horn bud, also presumably migrate through this region (Wu & Taneyhill 2019). Therefore, genes important for horn development may be expressed in both the frontal skin and horn bud. Future studies should include an additional control such as the forebrain skin to detect additional horn related genes and/or include polled tissues as controls.

4.4.9 Number of differentially expressed genes may reflect number of cells rather than regulation.

Expression of genes may be reduced because 1) a given gene is actively repressed, 2) expression of a given gene is not promoted, or 3) fewer cells in a population are expressing a given gene. Some genes with reduced expression may fall within the first two of these categories, that is, their gene expression is being regulated. However, some genes likely fall into the third group because the horn bud of horned fetuses has more cell types than the other tissues (Chapter 3). Some comparisons identified more differentially expressed genes with reduced expression rather than greater expression. For example, 86% of differentially expressed genes had lower expression when the horn bud of horned fetuses was compared to the forebrain skin. In contrast, only 2% of genes had decreased expression for the same comparison in polled fetuses. Additionally, only 16% of differentially expressed genes had increased expression compared to 84% which had reduced expression when the horn bud regions were compared between horned and polled fetuses. This affects the interpretation of the data because it cannot

be determined if the expression difference is important for development or if there are simply fewer cells expressing a given gene.

4.4.10 Samples did not cluster by tissue

Multi-dimensional scaling facilitates the comparison among samples, with those that are similar forming clusters. The samples within this study clustered by genotype and by fetus, with the exception of the samples from the horned fetus 736 which had the most variation amongst the samples (Figure 4.4B). Two samples from fetus 736, frontal skin and midbrain skin, were outliers. The polled samples clustered more tightly than the horned samples, which suggests there is greater variation between horned samples. However, the MDS plot indicated that the overall variation was not large as the range for the second dimension was only from 1.02 to -1.14 (Figure 4.4; y-axis).

Unexpectedly, the samples did not cluster by tissue. This may be because the tissues are histologically very similar, and therefore, the expression profiles may not segregate by sample region. For example, despite the histological differences between the horn bud and frontal skin samples (Chapter 3), most cell types are similar between the tissue regions, such as the mesenchymal and the epithelial cells outside of the horn bud. Other factors, such as sire, sex of the fetus, and slight differences in fetal age (due to minor variation in fertilisation, implantation and time of surgery) could also affect the clustering.

The greatest variation within tissue types was observed between the frontal skin samples and midbrain skin samples from the horned fetuses. These differences may have influenced the number for differentially expressed genes detected, particularly for samples from the horned fetuses. However, many DE genes were detected for most comparisons despite there being greater variation of horned samples. Nevertheless, the higher degree of variation in gene expression within the frontal skin from horned fetuses may have reduced the number of DE genes for the horn bud versus frontal skin comparison.

4.5 Conclusion

Transcriptomic analysis of tissues of horned and polled fetuses has provided a snapshot of gene expression related to horn ontogenesis at 58 days of development. The Celtic POLLED variant may directly affect *C1H21orf62*, *SON*, *EVA1C* and *LOC112447120* gene expression in the POLLED region. This effect, in turn, may then affect the expression of candidate genes, such as *RXFP2, TWIST2* and *ZEB2*, further downstream as the expression of these candidate genes also differed between horned and polled samples at 58 days. The Celtic variant also affected the expression of genes involved in the cytoskeleton, extracellular region, nerves, WNT signalling and BMP signalling. Transcription of *LEF1* of the WNT signalling pathway may be increased by HOX/MEIS2 expression in the horn bud. Further investigation is required to determine the importance of specific genes for horn development. In particular, localisation of proteins encoded by differentially expressed genes, such as *C1H21orf62, SON, EVA1C, LOC112447120, LEF1* and *FZD8*, in the tissue at different developmental time points would contribute to our understanding of horn ontogeny.

Chapter 5:

General Discussion

5.1 Discussion

The genetic basis of horn development is not understood. Hornless cattle (polled) occur naturally, which provides an opportunity to investigate horn development. Polledness in cattle is associated with four known intergenic DNA variants on bovine chromosome 1. Three of these variants overlap (Friesian, Mongolian and Guarani), whereas the Celtic variant is upstream of the others at a different site. Given the intergenic positions of these variants, it is not apparent how they cause the polled phenotype. One hypothesis tested in this work was that the POLLED variants affect the gene expression of neighbouring genes by altering the TAD structure of the chromosome. It was also hypothesised that horns are derived from neural crest cells and that the changes in gene expression caused by the POLLED variants affect the migration, condensation or differentiation of the neural crest cells. The potential effects of the POLLED variants on gene expression were explored bioinformatically, and the specific effects of the Celtic variant on horn bud development and gene expression were examined in the horn bud tissues of horned and polled fetuses at day 58 of development.

5.1.1 The TAD structure was the same for horn and polled sequence but the genes near the Celtic variant were differentially expressed

As altered TAD structure can cause phenotypic changes (Lupiáñez *et al.* 2015; Lupiáñez *et al.* 2016), the TAD structure of the POLLED genomic region was investigated. The TAD analysis refined the "POLLED region" and localised the four variants to two TADs within the range of chr1:2,240,000-2,759,999. The region was compared between wild-type and Celtic POLLED Hi-C sequences and it was found that the TAD structures were the same (Chapter 2). Therefore, the hypothesis that the Celtic POLLED variant changes gene expression by altering the TAD structure was not supported. As the sample number was limited though, this result needs to be validated, and the sequences of the other variants should be investigated as they may affect the TAD structure.

Although the TADs were the same between wild-type and Celtic POLLED sequences, gene expression of nearby genes were affected. Expression of *C1H21orf62, SON*, *EVA1C* and *LOC112447120* differed between horned and polled fetuses and therefore are candidate genes for horn ontogenesis identified from the present work which have not been proposed previously. *C1H21orf62* and *EVA1C* had decreased expression in the horn bud region of horned fetuses, while *SON* had increased expression (Chapter 4). A nearby lincRNA, *LOC112447120*, also had lower expression in the horn bud region of horned fetuses compared to polled fetuses (Chapter 4).

The Celtic variant consists of a 212 bp duplication and a 10 bp deletion. It has been shown by gene editing that the 10 bp deletion of the Celtic variant on its own does not cause the polled phenotype but the Celtic duplication has not been tested (Hennig *et al.* 2022a). Thus, the Celtic variant may cause the polled phenotype by the duplication of a regulatory region, such as an enhancer, and this could explain the effects on the expression of nearby genes.

5.1.2 Mesenchymal progenitor cells do not express neural crest markers at 58 days

Neural crest cells are involved in both bone and skin appendage development (Kaukua *et al.* 2014; Ishii *et al.* 2015) and therefore, may be involved in cranial headgear development. Condensed cells, thought to be neural crest cells, were discovered in the horn bud of horned fetuses at day 58. These cells were less evident or absent in the polled horn bud tissue sections. The lineage of these cells was tested by antibody staining of horn bud tissue with the neural crest cell markers SOX10 and NGFR, but the condensed cells were not positive for these markers (Chapter 3). This suggests that the condensed cells are not derived from neural crest cells. However, the neural crest cells required for horn development may have differentiated by day 58 to the point that they no longer express the neural crest markers. A neural crest cell tracking study of mouse dentition showed that mesenchymal neural crest cells near the developing tooth no longer expressed *Sox10* after migration at embryonic days 9-10 (Kaukua *et al.* 2014). SOX10 and NGFR positive cells have been reported in the horn bud epithelium

and dermal region of sheep at day 90 (Wang *et al.* 2019c). Therefore, cells derived from the neural crest cells may be found in the horn bud at a later developmental stage than examined here. A study using different markers and bovine fetuses at different stages of development is required to determine if mesenchymal neural crest cells are involved in horn bud development.

5.1.3 Nerves may be important for horn development

Nerves appear to be an underlying theme in horn development. Some of the genes and enhancers in the putative POLLED TAD, identified by comparative mapping in humans, have roles in neural development (Chapter 2, 4). Enhancers in the putative POLLED TAD from the RIKEN FANTOM5 database (Andersson *et al.* 2014; FANTOM Consortium *et al.* 2014) were investigated to identify tissues and cells in which they were active (Chapter 2). More than 80% of the enhancers were active in the brain tissues and the activity of ~15% were over-represented or active in an above average number of brain tissues (Chapter 2). The cell types in which enhancers were active included neuronal stem cells, neurons, and sensory epithelial cells (Chapter 2). Therefore, many of the enhancers in the POLLED region appear to affect gene expression in neural tissues. In humans at least, this chromosomal region is open for transcriptional activity in neuronal cell types. If this activity is conserved in cattle, enhancers within the POLLED region may be active in the same tissues and cell types, which is supported by the transcriptomic results from the present study (Chapter 4). Three out of four DE genes near the POLLED variants (*C1H21orf62, SON* and *EVA1C*) are expressed in neural tissue or involved with neural development (Chapter 4).

Nerve growth in the horn bud mesenchyme and epithelium was observed in the horned fetuses (Chapter 3). In polled fetuses, there were nerves in the horn bud region but they were only located in the mesenchyme. There was also less nerve tissue in the horn bud region and frontal skin of the polled fetuses than in the horn bud of horned fetuses, suggesting differences in nerve development within the horn bud region between the two genotypes (Chapter 3).

Whether this lack of nerves in horn bud region of the polled fetuses prevents horn bud development *or* whether the lack of a horn bud in the polled fetuses prevents nerve development is not known. In tooth development, the dental bud forms, *then* nerves grow and project towards the bud (Luukko & Kettunen 2014). The nerves are attracted to the tooth bud by signals from the epithelium and mesenchyme (Figure 4.8; Luukko & Kettunen 2014). A similar mechanism may be involved in horn development. Nerves are present in the mesenchyme of both horned and polled tissues, suggesting that the presence of nerves does not lead to horn bud formation. However, nerve growth into the epithelium was only observed in the horn bud of the horned fetuses but not in the horn bud region of the polled fetuses. As the horn bud does not form in polled fetuses, it is likely that there are no signals attracting nerve growth into the epithelium in the horn bud region. Therefore, the differences in nerve-related gene expression and nerve growth between horn and polled fetuses may be a downstream consequence of the POLLED variants rather than a direct effect of them.

5.1.4 Nerves may deliver Schwann cell precursors to the horn bud

The horn bud nerves may have a "non-neural" role in addition to its expected neural role in horn development. Peripheral nerves can deliver neural crest-derived Schwann cell precursors to peripheral tissues, which allows these multipotent cells to migrate to emerging structures at later fetal developmental stages, when migration through the mesenchyme is inefficient (Kaukua *et al.* 2014; Furlan & Adameyko 2018; Solovieva & Bronner 2021). Despite their name, Schwann cell precursors are not restricted to a Schwann cell fate (Figure 5.1) (Kaukua *et al.* 2014; Furlan & Adameyko 2018; Solovieva & Bronner 2021). For example, in mouse incisor tooth development, neural crest cells from both the mesenchyme *and* Schwann cell precursors contribute to the dental pulp and odontoblast layer of the tooth (Kaukua *et al.* 2014).
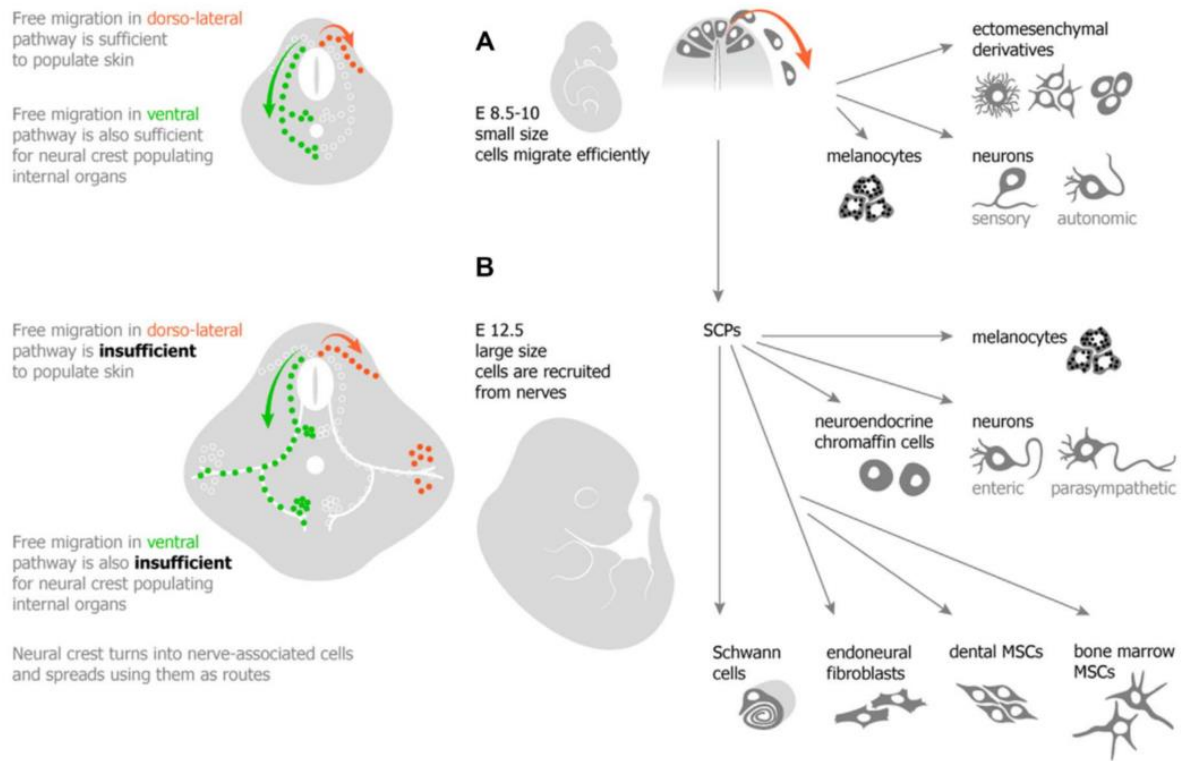
Figure 5.1: Neural crest cell migration pathways by A) free migration and B) becoming nerve-associated cells. Neural crest cell and Schwann cell precursor fates are also described. SCP = Schwann cell precursor, MSC = mesenchymal stem cells. Figure from Furlan and Adameyko (2018).

Glial cells in the horn bud region were identified as pro-myelinating Schwann cells, based on their position in relation to nerves and the localisation of SOX10 (Chapter 3), which is a marker for neural crest cells and cells from the Schwann cell lineage. If the nerves deliver Schwann cell precursors to the horn bud, then two populations of neural crest cells might contribute to horn development: the head mesenchyme comprising the condensed cells and Schwann cell precursors identified as SOX10 positive nerve-associated cells (Chapter 3). However, the neural crest origin of the condensed cells needs to be confirmed. The presence of SOX10 positive cells in the epithelial and dermal regions has been also reported in the sheep horn bud at day 90 of development (Wang *et al.* 2019c). Day 90 of development in sheep fetuses is much later in development than the equivalent stage in cattle, given the shorter gestation

period of sheep. These SOX10 positive cells in sheep may be derived from Schwann cell precursors, as only glial cells and nerves express SOX10 in bovine fetuses at day 58. However, the glial cells detected at 58 days in the bovine fetuses seemed to have differentiated beyond Schwann cell precursors to pro-myelinating Schwann cells. Therefore, additional investigation is required to determine the true identity of these cells and their potential contribution to horn growth.

5.1.5 *Nerve guidance and cell migration genes have a potential role in horn development*

GO pathway analysis showed that genes involved with sensory perception were differentially expressed between the horned and polled fetuses in the horn bud region and frontal skin (Chapter 4). KEGG pathway analysis showed that genes involved with neuroactive ligand-receptor interaction were differentially expressed between the horned and polled fetuses in the horn bud and frontal skin (Chapter 4). When the DE genes from horn bud and frontal skin were manually analysed for pathways, many genes relating to neural function had lower expression in the horn bud of horned fetuses than in the frontal skin (Chapter 4: Section 4.3.4.3). These results were surprising because the nerves in the horn bud of horned fetuses appeared to be more developed than in the frontal skin or than in the polled samples (Chapter 3). However, the genes that had higher expression in the horn bud of the horned fetuses were involved with axon growth and guidance, whereas the genes that had higher expression in the frontal skin were involved in nerve function (Chapter 4: Section 4.3.4.3).

Nerve guidance and neural crest cell migration have overlapping systems of attraction and repulsion. The repulsion cues are of particular interest because they channel axons and neural crest cells to the correct location. The main components in nerve and neural crest repulsion are the semaphorins (SEMA), ephrins/ephrin receptors (EFN/EPH), and slit guidance ligand/roundabout guidance receptors (SLIT/ROBO) (Theveneau & Mayor 2012). For example, when an axon has ROBO incorporated in its cell membrane, ROBO interacts with

SLIT from the extracellular region and the axon is repelled from that region (Tong *et al.* 2019). Genes within these families were differentially expressed between the horn bud regions of the horned and polled fetuses (Chapter 4).

One such gene, *EVA1C,* had increased expression in the horn bud region of polled fetuses compared to horned fetuses. In mice, Eva1c may have a role in axon repulsion (James *et al.* 2013). Therefore, *EVA1C* expression by neurons may cause pioneering axons to grow slower or be repelled from the horn bud region of polled fetuses. *EVA1C* is located ~530 kb downstream from the Celtic *POLLED* variant and is outside both the POLLED region and the putative POLLED TAD, albeit only ~30 kb from the putative TAD boundary. As this gene is outside the POLLED region, it is possible that the increase in its expression in the polled fetuses is an indirect effect, rather than a direct effect, of the Celtic variant.

Interestingly, there is evidence that EVA1C may also affect cell migration. A study of neuroblast migration in *Caenorhabditis elegans* showed that EVA-1 (EVA1C orthologue), which is upregulated by WNT signalling, inhibited the migration of the neuroblasts (Rella *et al.* 2021). If EVA1C also has a role in neural crest cell migration, then it and other regulators of migration may prevent the neural crest cells from reaching the horn bud region, ultimately preventing horns from developing. In the horned fetuses, the reduced *EVA1C* expression by migrating cells may mean that the cells do not respond to repulsion cues in the horn bud (such as SLIT) and continue migration into the horn bud region. In polled fetuses, the increased expression of *EVA1C* may mean that cells are sensitive to repulsion cues and cannot enter the horn bud region.

This hypothesis is supported by the histological data. The histological data showed that the condensed cells are significantly reduced in the mesenchyme, initiation of horn bud formation does not occur and innervation is reduced in polled fetuses. To better understand which cells are over-expressing *EVA1C* in the polled fetuses and whether *EVA1C* does have a

role in stopping the migration of neural crest cells, further localization or spatial transcriptomic studies are required.

The product of another gene in the POLLED region, SON, mediates splicing of genes critical for neuronal migration in humans (Kim et al. 2016). *SON* had higher expression in the horn bud of horned fetuses compared to polled fetuses and is located ~500kb upstream of the Celtic POLLED variant (Chapter 4). Interestingly, SON has been shown to splice the pre-mRNA of *FLNA* (Kim et al. 2016), which also had increased expression in the horn bud of horned fetuses compared to the horn bud region in polled fetuses (Chapter 4). FLNA (filamin A) is a protein that interacts with F-actin and, among other functions, is involved brain neurogenesis and neuronal migration (Fox *et al.* 1998; Zhou *et al.* 2010). Therefore, altered expression of *SON* may also affect cell migration.

5.1.6 The "lost cells" hypothesis

There may be an association between headgear and the phenotype of other neural crest derived structures. Horned cattle have a single row of eyelashes, whereas polled cattle have 'bushy' eyelashes arising from additional eyelashes on the eyelid. These two traits are most likely connected (Allais-Bonnet *et al.* 2013). Similarly, musk deer (*Moschidae* species) and Chinese water deer (*Hydropotes inermis*) no longer have antlers through independent convergent evolution (Figure 5.2; Wang *et al.* 2019c), yet both species have elongated upper canine teeth. This suggests a link between antler and canine tooth development. The size and complexity of antlers of a given deer species may be associated with the reduction or loss of upper canine teeth, and there seems to be a trade-off between the two traits (Heckeberg 2017). Eyelashes, dental lamina and dental papilla in humans develop from a combination of neural crest cells and head mesoderm (Kao *et al.* 2007; Edgar *et al.* 2013). Therefore, neural crest cell may be the connection between eyelashes or canine teeth and headgear. These structures are also connected by the trigeminal nerve.

Given that cell migration-related genes are differentially expressed between horned and polled fetuses (Chapter 4), it is possible that the migration path of the neural crest cells required for horn development is disrupted. The mis-directed cells may converge at the eyelid in polled cattle, which would explain why they have more eyelashes, and by analogy, these cells may result in elongated canine teeth in musk deer and Chinese water deer. In cattle, the changes in migration could arise from the altered expression of *SON* or *EVA1C* caused by the Celtic POLLED variant. The contribution of neural crest cells to the eyelid and eyelashes should be tested by comparing horned and polled fetuses at different fetal developmental stages and with different cell markers as described above.
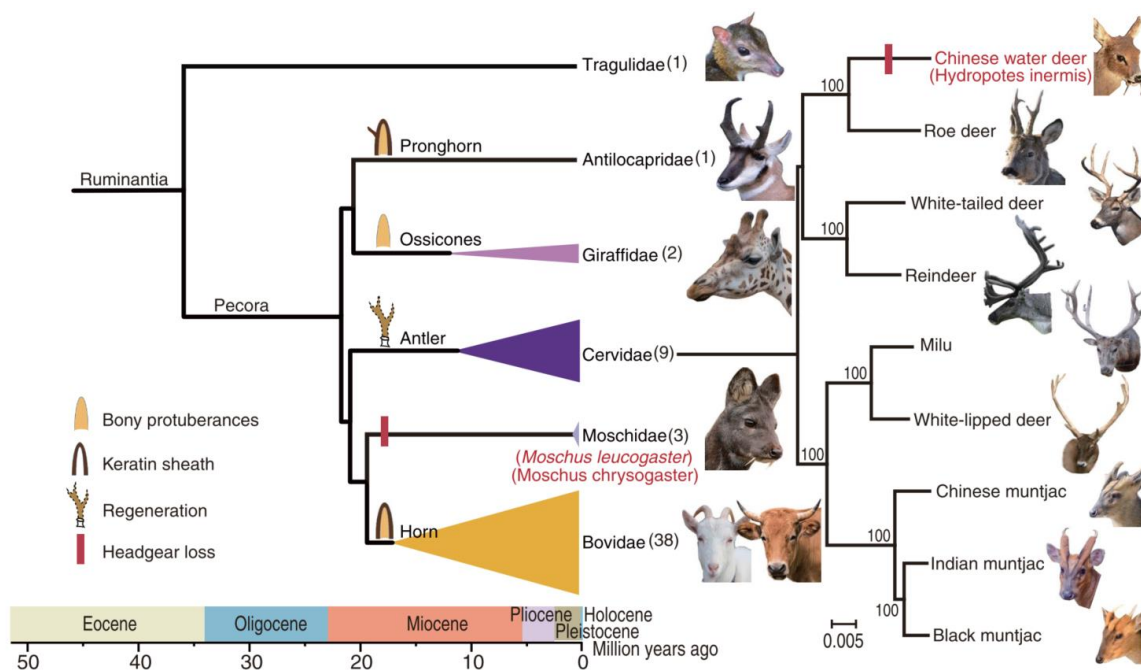
Figure 5.2: Phylogeny of ruminant species with and without headgear. Moschidae and Hydropotes independently lost antlers and have large canine teeth. Figure from Wang *et al.* (2019c).

5.1.7 Early horn bud and tooth bud formation are similar

The horn bud was apparent on the head of the homozygous horned fetuses at day 58 of development as a ring of depressed skin with a raised centre (Chapter 3). This depression was absent in the homozygous polled fetuses. The main histological differences between horned and polled fetuses at 58 days of development were the epithelium and the condensed cells in the horn bud of the horned fetuses (Chapter 3). The horn bud had a stratified epithelium and a group of condensed cells in the mesenchyme layer. Polled samples of the same region had a single layer of epithelium and the condensed cells were less evident or absent from the tissue sections.

This horn bud development shares similarities with the early stages of tooth development (Luukko & Kettunen 2014). Both the horn bud and the tooth bud begin with the stratification of epithelium and condensation of the mesenchyme, and both are innervated by branches of the trigeminal nerve. However, the position of the condensed cells differs: the mesenchymal cells condense directly below the epithelium when the tooth bud forms (Luukko & Kettunen 2014; Svandova *et al.* 2020), whereas in horn development, there is a region of apparently normal mesenchyme separating the epithelium and condensed cells in the horn bud. This suggests a difference between ectodermal appendage development (teeth) and headgear development (horns). However, the difference observed may be because the horn bud at 58 days is not at its earliest stage of formation. To determine if the cells condense just below the horn bud when the horn bud first forms, fetuses will need to be sampled prior to 58 days of development.

Given that cell condensation is vital for appendage development (Fuchs 2007; Li *et al.* 2016; Puthiyaveetil *et al.* 2016; Salhotra *et al.* 2020), altered proliferation of these condensed cells could affect horn development. SON has a role in mitosis, as the RNA interference of *SON* mRNA results in mitotic arrest (Hickey *et al.* 2014). *SON* is located near the Celtic variant and

has higher expression in the horn bud of horned fetuses compared to polled fetuses. Therefore, mitotic arrest of the precursor cells may prevent horn development in polled fetuses.

5.1.8 Other elements near the POLLED variants may be important for horn development

C1H21orf62 expression was affected by the Celtic variant, with lower expression in the horn bud of horned fetuses compared to the horn bud region of polled fetuses (Chapter 4). The function of *C1H21orf62* is not known, therefore its role in horn development cannot be predicted. *C1H21orf62* is downstream of the Guarani, Friesian and Mongolian variants and is located at the border of the POLLED TAD 4 (Chapter 2: Section 2.3.8). Intriguingly, although the Celtic variant is located in TAD 1, the Celtic variant is predicted to interact with a region 20 kb upstream of *C1H21orf61* (Chapter 2: Section 2.3.9). The protein structure of C1H21orf62 suggests that it is a secreted protein and its human orthologue is expressed in the brain (Uhlén *et al.* 2015). As *C1H21orf62* has higher expression in the horn bud region of polled fetuses than in the horn bud, it may negatively affect horn development. It is possible that chromatin changes caused by the Guarani, Friesian and Mongolian variants could increase the expression *C1H21orf62*, though this hypothesis needs to be tested. Therefore, the function of *C1H21orf62* and its involvement in horn development warrant further investigation.

LOC112447120 is a long intergenic non-coding RNA located ~256 kb upstream of the Celtic POLLED variant between *IFNGR2* and *IFNAR1*. It is outside of the POLLED region defined herein but is within the putative POLLED TAD (Chapter 2). Although the function of *LOC112447120* is unknown, it is one of three lincRNA in the putative POLLED TAD that is highly conserved in horned species (water buffalo, sheep and goat) but not in hornless species (pig, horse, dog and human) (Chapter 2). This suggests that it may have a bovine or ruminant specific function such as horn development.

The POLLED region also contains a protein-coding gene *LOC526226* and two ESTs (CB166156.1 and DN819280.1) (Chapter 2) that are only found in cattle. *LOC526226* was differentially expressed between the forebrain skin of horned and polled fetuses. The ESTs were

not differentially expressed between tissues or genotypes (Chapter 3). Despite the lack of differential expression between the genotypes, *LOC526226* and the ESTs should not be completely disregarded as elements affected by the POLLED variants as their potential differential expression has been only measured at one fetal developmental time point, day 58.

Previous studies have proposed other genes within the POLLED region as candidates involved with horn development, namely *OLIG1*, *OLIG2*, *LincRNA#1* (*LOC100848368*) and *LincRNA#2* (*LOC112447133*), as they were shown to be differentially expressed between horned and polled fetuses or between the horn bud and frontal skin from 70 days of development and later (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014). There was no evidence of differential expression of these genes in the horn bud region between the horned and polled fetuses at day 58 for any of these genes though.

5.1.9 Expression of candidate genes, WNT signalling genes and HOX pathway at 58 days of development differ between genotypes

Candidate genes outside the POLLED region (*RXFP2, ZEB2* and *TWIST2*) and genes from the WNT and HOX pathways were differentially expressed in the horn bud region of the horned and polled fetuses at day 58. In particular, *RXFP2* had increased expression in the horn bud of horned fetuses compared with the other tissues (Chapter 4), which is consistent with previous transcriptomic studies of horned and polled bovine fetuses at 70 and 90 days (Allais-Bonnet *et al.* 2013; Wiedemar *et al.* 2014). *RXFP2* also had increased expression in the horn bud compared to frontal skin of horned ovine fetuses at 90 days (Wang *et al.* 2019c). The RXFP2 protein was localised to the nerves and epithelium in the horn bud of the horned fetuses in the histology studies (Chapter 3). However, the RXFP2 protein was also present in the nerves within the horn bud region of the polled fetuses. Therefore, RXFP2 may be involved with nerve development and epithelium differentiation in the horn bud region (Chapter 3).

There is evidence that RXFP2 plays a role in the horn development in cattle and sheep. *RXFP2* has been associated with scurs in cattle (Wang & Gill 2021). Scurs are horn-like

appendages that range from scabs in the horn region to large outgrowths resembling true horns. In addition, various studies have identified variants in *RXFP2* that are associated with horn status in sheep (Johnston *et al.* 2011; Wang *et al.* 2014b; Wiedemar & Drögemüller 2015). Single-nucleotide polymorphisms in *RXFP2* have been associated with horns, scurs and polledness in Tan sheep (Wang *et al.* 2014b) and in Soay sheep (Johnston *et al.* 2011). In semi-feral populations of Prairie Tibetan sheep and Oula sheep, a specific *RXFP2* haplotype is associated with horn size (Pan *et al.* 2018). Loss-of-function mutations in *RXFP2* may be the cause of the evolutionary loss of antlers in musk deer and Chinese water deer (Wang *et al.* 2019c).

The role of RXFP2 in tissue development is not well understood, although the presence of RXFP2 has been implicated in the growth and development of organs. Its expression in the horn bud and presence in epithelial cells and nerves may enhance the growth of these tissues in the horn bud. *RXFP2* is expressed in the epithelium and mesenchyme of developing teeth (Duarte *et al.* 2014). Additionally, RXFP2 has been shown to activate expression of axon guidance genes plus WNT, HGF and BMP signalling in rat gubernacular cells (Johnson *et al.* 2010). The connection between RXFP2 and downstream signalling pathways, such as WNT/BMP signalling, needs to be explored further to understand the importance of RXFP2 in horn development.

Components of WNT and BMP signalling were among the genes differentially expressed between horned and polled fetuses (Chapter 4). WNT ligands had lower expression in the horn bud of horned fetuses compared with other tissues and the polled fetuses, but genes encoding a WNT receptor (*FZD8*), β-catenin (*CTNNB1*), and a transcription factor (*LEF1*) had higher expression in the horn bud of horned fetuses. Activation of frizzled receptors in WNT signalling leads to the translocation of β-catenin from the cytoplasm to the nucleus (Wiese *et al.* 2018). β-catenin then promotes transcription in cooperation with TCF/LEF transcription factors (Wiese *et al.* 2018). SMAD1 of the BMP signalling also had increased expression in the

horn bud of horned fetuses compared to polled fetuses. Thus, *FZD8, CTNNB1, LEF1* and *SMAD1* are good candidates for genes involved in horn development.

While there is little evidence of WNT or BMP signalling is involved in headgear development, both the WNT and BMP signalling pathways are essential for limb development (Dathe *et al.* 2009; Lin & Zhang 2020) and skin appendage development (Van Genderen *et al.* 1994; Sarkar & Sharpe 1999; Sasaki *et al.* 2005; Lan *et al.* 2014; Hermans *et al.* 2021). In particular, the knockout of *Lef1* results in mice that do not form ectodermal appendages, such as hair, teeth and mammary glands, nor develop the trigeminal nerve (Van Genderen *et al.* 1994). In cattle, the trigeminal nerve innervates teeth, upper eyelid and horns among other regions.

HOX, MEIS2 and PBX3 may promote *LEF1* transcription in the horn bud. *HOXD* genes may be specifically involved in headgear development (Wang *et al.* 2019c). A highly conserved pecoran-specific element is located near the *HOXD* gene cluster (Wang *et al.* 2019c). However, *HOXD* genes either had lower expression in the horn bud or similar levels of expression among tissues. In contrast, *MEIS2* and *PBX3* had higher expression in the horn bud of horned fetuses compared to the same region in polled fetuses. MEIS2 and PBX3 are transcription cofactors that bind cooperatively to HOX transcription factors to initiate transcription of genes. These cofactors are responsible for the temporal and spatial specificity of HOX-related transcription. A study of limb development in mice showed that double knockout of *Meis1*/2 reduces the expression of *Lef1* (Delgado *et al.* 2021). Meis DNA binding sites have been identified upstream of the *Lef1* gene (Delgado *et al.* 2021). The HOX transcription factors may be activated in the cranial region by HOX cofactors, MEIS2 and PBX3, may enable transcription of horn-specific genes including *LEF1*.

ZEB2 and *TWIST2,* candidate genes previously linked to horn development, also had increased expression in the horn bud of horned fetuses compared to polled fetuses. This is the first study to observe the differential expression of *ZEB2* and *TWIST2* and suggests that their

encoded proteins are required for horn development, although their exact role in horn development is not known. Deletion and frameshift mutations of *ZEB2* in cattle cause the polled phenotype, as well as other changes including poor growth and infertility in females (Capitan *et al.* 2009; Capitan *et al.* 2012; Gehrke *et al.* 2020b). In humans, nonsense mutations in *TWIST2* cause Setleis syndrome, which is characterised by scar-like lesions on the temporal lobe and the absence of eyelashes (Tukel *et al.* 2010). This is interesting because the phenotypes of the Setleis syndrome and the POLLED variants affect both the temple and eyelashes. Thus, TWIST2 is likely to affect similar pathways in cattle.

5.2 Future directions

The research in this thesis has led to new hypotheses regarding horn ontogenesis in ruminants, however, the work was limited by the number of samples, the cell and tissue types, and the fetal developmental stage. Some of these limitations can be addressed in future studies using the new tools that are now available.

5.2.1 Gene knockout studies to study gene effect on development

Gene knockout studies can be used to assess the effect of a given gene on a phenotype. Gene editing has been used to understand the effect of the Celtic POLLED variant. For example, the effect of the Celtic variant on horn development was confirmed by producing polled calves from edited wild-type somatic cells and somatic cell nuclear transfer (Carlson *et al.* 2016). Furthermore, the Crisper-Cas9 editing system was used test the effect of the 10 bp deletion on horn development without the 212 bp duplication. A 133 bp region including the Celtic polled 10 bp deletion site in cattle was deleted, however, the fetuses remained horned (Hennig *et al.* 2022a). The same research group used CRISPER-Cas9 to knock-out *LincRNA#1* in heterozygous fetuses (P$_c$/p). *LincRNA#1* was shown to be up-regulated in the horn bud region of polled fetuses in an earlier study (Allais-Bonnet *et al.* 2013). If *LincRNA#1* was knocked-out in heterozygous cattle, it was reasoned that horns would then develop (Hennig *et al.* 2022b).

However, the *LincRNA#1* knockout bovine fetuses remained polled, suggesting that up-regulation of *LincRNA#1* is not the causative factor for polledness.

Future studies could use a similar approach to investigate the involvement of genes in horn development. For example, the genes near the POLLED variants that had increased expression in the horn bud of horned fetuses (e.g. *SON* and *LOC112447120*) could be knocked out by gene editing genetically horned fetuses. If these genes are vital to horn development, lack of expression should create a polled phenotype. As *C1H21orf62* and *EVA1C* expression is increased in polled fetuses, duplication of these genes in horned fetuses could determine whether their increased expression causes polledness. Another approach could be to investigate whether knocking out these genes in genetically polled fetuses achieves a horned phenotype.

5.2.2 Spatial transcriptomic study of structure and function in horn bud development

The research described in this thesis investigated how the tissue physiology or 'function' relates to tissue organisation or 'structure' in horn development. The approach was to analyse genomic sequence, histological and gene expression data to understand the formation of the horn bud structure. Immunohistochemistry provided a limited link between function and structure. Uncovering the basis of horn development is difficult as several cell types are involved, and gene expression analysis provides an average of the cells that were sampled. Due to time constraints, the location of only one protein from a differentially expressed gene, *RXFP2*, was determined. While many genes putatively involved in horn development have been suggested, identifying the location of their protein products has yet to be addressed. Protein localisation in the horn bud seems to lag despite being important for understanding function.

Spatial transcriptomics could be used to localise the expression of specific genes. Spatial gene expression is the current 'gold-standard' for studying the intersection of physiology and structure by capturing mRNA distribution *in situ* to developing a transcriptomic 'atlas' (Palla *et al.* 2022). The technique can provide gene expression information at tissue and cellular resolution (Waylen *et al.* 2020; Palla *et al.* 2022), unlike RNAseq, where gene expression data

is the average of the whole sample. Furthermore, as the location of expression is known, cells with similar expression can be clustered and differential gene expression between clusters analysed. Lastly, predictions on the role of differentially expressed genes can be formed because the cells expressing the genes are known.

High definition spatial transcriptomics (Vickovic *et al.* 2019) and Slide-seq (Rodriques *et al.* 2019) are two such methods that could be used to study fetal horn development and define groups of cells in the horn bud based on gene expression, which then can be compared. For example, the condensed cells and 'normal' mesenchymal cells within in the horn bud may cluster separately, and their expression patterns explored. Furthermore, the expression levels of specific genes may be mapped so the cells and regions expressing candidate genes can be detected. The expression of genes that modulate cell migration and axon growth could be mapped to reveal migration and growth paths, respectively, using common repulsion cues and their receptors, such as SLIT/EVA1C and SLIT/ROBO. The spatial transcriptomics data collected from horned fetuses could be compared to polled fetuses to more accurately determine the effect of POLLED variants on horn development.

5.2.3 Hi-C analysis, isoform sequencing, and cap analysis of gene expression

Data from technologies, such as Hi-C analysis, Cap Analysis of Gene Expression (CAGE), and long read isoform sequencing (e.g. Iso-seq) are critical for understanding the effect of the POLLED variants. As mentioned in Chapter 2, a TAD analysis of individuals with the Friesian and Guarani variants would be interesting. These variants are larger duplications, and thus, they are more likely to change the chromatin interactions than the smaller Celtic variation studied here. CAGE facilitates the identification of transcription start sites, and can be used to analyse RNAs transcribed by active enhancers (Morioka *et al.* 2020), and therefore, could be used to examine enhancers in the POLLED region and test if the POLLED variants interrupt any of those enhancers. Iso-seq, may be used to study the expressed isoforms of genes within the POLLED region. This would be particularly useful for investigating if *LOC526226*

is involved with horn development, as both histone H4 and osteogenic growth peptide may be transcribed from this gene. A recent study used Iso-seq to study the isoforms of known fertility-related genes in Brahman cattle and showed that using more than one type of sequencing is important for a fuller characterisation of the transcriptome landscape (Ross *et al.* 2022).

5.2.4 Fetal developmental time-points and sampling

Fetal developmental time points and the sampling methods need to be considered for future histology experiments. The initial horn bud development may not have been captured at 58 days when the horn bud was visible. Ideally, horn bud formation prior to this time point should be assessed. However, sampling will be difficult. The biopsy punches used in the histology study sampled precise locations because the horn bud was just visible. However, the frontal skin and polled tissues were very fragile which meant that many sections were poor quality. Cryopreservation and sectioning might better preserve tissue structure, but detecting the horn bud would be more difficult as it would have to be identified post-staining.

Lastly, although the number of samples for the histology and RNA sequencing was more than used in previous studies, additional samples would be valuable to validate the results obtained. Unfortunately, obtaining bovine fetal samples at specific developmental stages is both difficult and expensive, making extensive experiments challenging.

5.3 Conclusion

This thesis describes the POLLED genomic region, horn bud morphology and gene expression of horned and polled fetuses at 58 days of development. The horn bud regions of horned and polled fetuses were morphologically differentiated at this time-point and the horn bud did not form in polled fetuses. In the horn bud, a population of aggregated cells observed in the mesenchyme may represent horn progenitor cells. These aggregated cells did not express neural crest markers, suggesting that they are not derived from the cranial neural crest or that they have already differentiated and no longer express these markers. Schwann cell precursors

may contribute to horn development. The POLLED variants may alter the genomic interactions in the region and the Celtic variant was shown to affect the expression of three protein-coding genes (*C1H21orf62*, *EVA1C* and *SON*) and one lincRNA (*LOC112447120*). The altered expression of these genes may affect cell migration, cell aggregation and/or nerve growth required for horn development.

5.4 Final remarks

This work contributes to the understanding of horn development in cattle and other species with headgear. The knowledge is particularly important for horned species, including cattle, sheep and goats, where horns are often a commercially undesirable trait (Simon *et al.* 2022). Horned animals are often disbudded or dehorned through painful procedures which impacts animal well-being. Knowledge from this research may contribute to the development of new procedures with improved welfare outcomes. An example would be the development of an injectable agent that inhibits post-natal horn development by affecting horn developing cells without impacting other tissues.

Appendices

Appendix A: Supplementary material for Chapter 2

Table A1: Assemblies used for conservation analyses of the POLLED region. Ten horned, one antlered and 13 species without headgear were analysed.

Species	Common name	Headgear phenotype	Assembly (NCBI accessed 2/7/20)
Bos taurus	Cattle	Horns	ARS-UCD1.2
Bos indicus	Indicine cattle	Horns	UOA_Brahman_1
Bos grunniens	Domestic yak	Horns	BosGru3.0
Bos mutus	Wild yak	Horns	BosGru_v2.0
Bubalus bubalis	Water buffalo	Horns	UOA_WB_1
Bison bison	Bison	Horns	Bison_UMD1.0
Ovis aries	Sheep	Horns	Oar_rambouillet_v1.0
Capra hircus	Goat	Horns	ARS1
Capra aegagrus	Wild goat	Horns	CapAeg_1.0
Pantholops hodgsonii	Tibetan antelope	Horns	PHO1.0
Cervus elaphus hippelaphus	Red deer	Antlers	CerEla1.0
Moschus moschiferus	Siberian musk deer	No headgear	MosMos_v2_BIUU_UCD
Camelus bactrianus	Bactrian camel	No headgear	Ca_bactrianus_MBC_1.0
Camelus dromedarius	Arabian camel	No headgear	CamDro3
Camelus ferus	Wild Bactrian camel	No headgear	BCGSAC_Cfer_1.0
Sus scrofa	Pig	No headgear	Sscrofa11.1
Equus caballus	Horse	No headgear	EquCab3.0
Canis lupus familiaris	Dog	No headgear	canfam4
Homo sapiens	Human	No headgear	GRCh38.p13
Loxodonta africana	African elephant	No headgear	Loxafr3.0
Monodelphis domestica	Opossum	No headgear	MonDom5
Ornithorhynchus anatinus	Platypus	No headgear	mOrnAna1.pri.v3
Choloepus hoffmanni	Sloth	No headgear	C_hoffmanni-2.0.1
Dasypus novemcinctus	Armadillo	No headgear	Dasnov3.0

Table A2: Assemblies used for gene synteny and lincRNA analyses of the POLLED region. A proxy POLLED region from DONSON to SYNJ1 was used because the 1-kb sequence from the cattle were not conserved across all species, and therefore, could not be defined.

Species	Headgear phenotype	Assembly	Source	Region
Cow	Horned	ARS-UCD1.2	Ensembl	Chr1:1,946,384 – 2,921,213
Water buffalo	Horned	UOA_WB_1	NCBI	Chr1:46,565,723 – 47,541,217
Sheep	Horned	Oar_rambouillet_v1.0	NCBI	Chr1: 130,024,872 – 131,033,373
Goat	Horned	ARS1	Ensembl	Chr1: 369,751 – 1,374,417
Pig	No headgear	Sscrofa11.1	Ensembl	Chr13: 196,230,089 – 197,193,656
Horse	No headgear	EquCab3.0	Ensembl	Chr26: 30,595,136 – 31,450,213
Dog	No headgear	CanFam3.1	Ensembl	Chr31: 27,509,522 – 28,117,154
Human	No headgear	GRCh38.p12	GENCODE	Chr21: 32,628,759 – 33,588,684

Table A3: Query coverage and identity of Human enhancers from the RIKEN FANTOM 5 database obtained when aligned to the Bos Taurus reference genome (ARSUCD1.2). Only matches (8/34) to the putative POLLED TAD are shown (Chr1:1,946,384 – 2,921,213).

	RIKEN FANTOM5 enhancer coordinates (Chr21)		Enhancer coordinates (Chr21) (GRCh38.p13)		BLAST alignment to Bos taurus (Chr1)			
Enhancer	Start	End	Start	End	Start	End	Query coverage (%)	Identity (%)
E3	34,350,736	34,350,972	32,978,364	32,978,764	2,588,333	2,588,724	99	77
E5	34,379,947	34,380,464	33,007,727	33,008,127	2,566,794	2,567,110	79	74
E8	34,407,597	34,407,715	33,035,139	33,035,539	2,539,031	2,539,445	99	66
E12	34,451,403	34,451,550	33,078,994	33,079,394	2,489,957	2,490,222	68	81
E13	34,524,822	34,524,923	33,152,373	33,152,773	2,432,700	2,433,095	97	84
E14	34,571,192	34,571,326	33,198,718	33,199,118	2,381,477	2,381,517	10	85
E26	34,746,472	34,746,562	33,374,022	33,374,422	2,173,228	2,173,449	57	75
E27	34,746,676	34,747,085	33,374,433	33,374,833	2,172,722	2,173,072	81	69

Table A4: All gene-lincRNA-gene triplets from cattle that were conserved in other species. The cattle triplets were compared to water buffalo, sheep, goat, pig, horse, dog and human. A bovine lincRNA was considered conserved in another species based on conservation of synteny, i.e, when a lincRNA in another species was placed between the same genes and was transcribed in the same direction.

Bovine lincRNA	Species	LincRNA	Gene1	Gene2
LOC100848368	Water buffalo	*LOC102394079*	*OLIG1*	*OLIG2*
	Horse	*ENSECAG00000030232*	*OLIG1*	*OLIG2*
	Dog	*ENSCAFG00000037762*	*OLIG1*	*OLIG2*
		LOC111093515	*OLIG1*	*OLIG2*
		ENSCAFG00000044452	*OLIG1*	*OLIG2*
	Human	*LINC00945*	*OLIG1*	*OLIG2*
LOC112447133	Water buffalo	*LOC112583890*	*OLIG2*	*C1H21orf62*
	Goat	*ENSCHIG00000005617*	*OLIG2*	*C1H21orf62*
	Dog	*ENSCAFG00000048727*	*OLIG2*	*C1H21orf62*
	Human	*ENSG00000232360.1*	*OLIG2*	*C1H21orf62*
LOC112447120	Pig	*ENSSSCG00000051450*	*IFNGR2*	*IFNAR1*
LOC104970778	Sheep	*LOC114113258*	*IFNAR1*	*IL10RB*
	Pig	*ENSSSCG00000049143*	*IFNAR1*	*IL10RB*

Table A5: Number of total lincRNA annotated for the putative POLLED TAD and comparative regions in water buffalo, sheep, goats, pigs, horses, dogs and humans. The number of lincRNA was obtained from Ensembl and NCBI.

Species	Number of lincRNA
Cow	5
Water buffalo	5
Sheep	1
Goat	2
Pig	6
Horse	7
Dog	13
Human	7

Table A6: Statistics from the alignment of haplotype-resolved Hi-C sequences to Angus (UOA_Angus_1), Brahman (UOA_Brahman_1) and the reference genome (ARS-UCD1.2). Hi-C reads were obtained from the lung tissue of an Angus-Brahman F1 hybrid fetus at 90 days of development. Hi-C reads were separated into Angus- and Brahman-specific reads using a k-mer approach. The data were produced by Low *et al.* (2020) and downloaded from NCBI SRR6691720 of PRJNA432857.

Sub-species	Angus		Brahman	
Genome	**UOA_Angus_1**	**ARS-UCD1.2**	**UOA_Brahman_1**	**ARS-UCD1.2**
Sequenced Read Pairs	144,922,493	144,922,493	151,034,776	151,034,776
Normal Reads	107,395,215 (74.11%,R1) - 97,296,013 (67.14%,R2)	101,980,331 (70.37%,R1) - 91,695,666 (63.27%,R2)	113,194,805 (74.95%,R1) - 102,035,917 (67.56%,R2)	106,114,951 (70.26%,R1) - 95,437,555 (63.19%,R2)
Chimeric Unambiguous	18,381,090 (12.68%,R1) - 16,950,535 (11.7%,R2)	18,449,480 (12.73%,R1) - 17,007,333 (11.74%,R2)	19,106,953 (12.65%,R1) - 17,629,456 (11.67%,R2)	19,170,699 (12.69%,R1) - 17,682,388 (11.71%,R2)
Ligations	12.94% (R1) - 11.91% (R2)	12.94% (R1) - 11.91% (R2)	12.9% (R1) - 11.88% (R2)	12.9% (R1) - 11.88% (R2)
Unmapped	7,941,350 (5.48%)	11,514,666 (7.945%)	6,871,116 (4.549%)	11,924,754 (7.895%)
Low Mapping Qual	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Unique Aligned Pairs	60,432,110 (41.7% / 100%)	61,266,575 (42.275% / 100%)	64,770,331 (42.884% / 100%)	63,944,423 (42.338% / 100%)
Valid Contacts	31,825,876 (21.96% / 52.66%)	32,278,892 (22.27% / 52.69%)	34,297,288 (22.71% / 52.95%)	33,749,034 (22.35% / 52.78%)
Duplicate Contacts	1,222,607 (0.84% / 2.02%)	1,243,451 (0.86% / 2.03%)	1,322,184 (0.88% / 2.04%)	1,298,895 (0.86% / 2.03%)

Table A6: Hi-C alignment statistics (Continued).

Intra Fragment	15,779,662 (10.89% / 26.11%)	15,714,954 (10.84% / 25.65%)	16,838,357 (11.15% / 26.0%)	16,362,345 (10.83% / 25.59%)
Inter Chromosomal	14,650,349 (10.11% / 24.24%)	14,929,118 (10.3% / 24.37%)	15,730,318 (10.42% / 24.29%)	15,644,631 (10.36% / 24.47%)
Intra Chromosomal	15,952,920 (11.01% / 26.4%)	16,106,323 (11.11% / 26.29%)	17,244,786 (11.42% / 26.62%)	16,805,508 (11.13% / 26.28%)
Intra Short Range (< 20kb)	2,481,636 (1.71% / 4.11%)	2,463,119 (1.7% / 4.02%)	2,644,372 (1.75% / 4.08%)	2,560,636 (1.7% / 4.0%)
Intra Long Range (> 20kb)	13,471,284 (9.3% / 22.29%)	13,643,204 (9.41% / 22.27%)	14,600,414 (9.67% / 22.54%)	14,244,872 (9.43% / 22.28%)
Read Pair Type (L-I-O-R)	24.65%-26.15%-24.59%-24.61%	24.65%-26.11%-24.61%-24.62%	24.65%-26.14%-24.59%-24.62%	24.65%-26.1%-24.62%-24.63%

Table A7: TAD coordinates within the putative POLLED TAD (Chr1:1,946,384 – 2,921,213; ARS-UCD1.2) (Wang *et al.* 2018) for Angus (polled) and Brahman (horned). Hi-C sequences were mapped to the bovine reference genome ARSUCD1.2. Hi-C reads were from the lung tissue of an Angus-Brahman F1 hybrid fetus at 90 days of development. Hi-C reads were separated into Angus- and Brahman-specific reads using a k-mer approach. The data were produced by Low *et al.* (2020) and downloaded from NCBI SRR6691720 of PRJNA432857.

Species	Chromosome	Start	End
Angus (ARSUCD1.2 genome)	chr1	1960000	2079999
		2080000	2239999
		2240000	2439999
		2440000	2519999
		2520000	2599999
		2600000	2759999
		2760000	2959999
Brahman (ARSUCD1.2 genome)	chr1	1960000	2079999
		2080000	2239999
		2240000	2439999
		2440000	2519999
		2520000	2599999
		2600000	2759999
		2760000	2959999

Table A8: CTCF binding site coordinates (ARSUCD1.2) predicted by Find Individual Motif Occurrences (version 4.12.0) from the MEME suite (Bailey *et al.* 2009) based on the CTCF position weight matrix from the JASPAR database (motif id: MA0139.1).

Chromosome	Start	End	Strand
chr1	1944233	1944251	+
chr1	2129358	2129376	+
chr1	2202953	2202971	+
chr1	2257839	2257857	+
chr1	2296078	2296096	+
chr1	2373568	2373586	-
chr1	2384654	2384672	+
chr1	2445776	2445794	+
chr1	2523866	2523884	-
chr1	2549857	2549875	-
chr1	2634288	2634306	-
chr1	2648138	2648156	-
chr1	2733863	2733881	+
chr1	2824524	2824542	-
chr1	2893391	2893409	-
chr1	2936456	2936474	+
chr1	2963068	2963086	+

Table A9: Significant chromatin interactions between the Celtic region (chr1:2,420,000-2,440,000) and nearby regions analysed from Angus- and Brahman-specific Hi-C sequences. Coordinates unique to Angus (polled) or Brahman (horned) sequence are highlighted in grey.

Breed	Chromosome	Start	End
Angus	chr1	970000	980000
		1770000	1780000
		2720000	2730000
		2830000	2840000
		2910000	2920000
		10880000	10890000
Brahman	chr1	1770000	1780000
		2720000	2730000
		2830000	2840000
		2910000	2920000
		3090000	3100000
		10880000	10890000
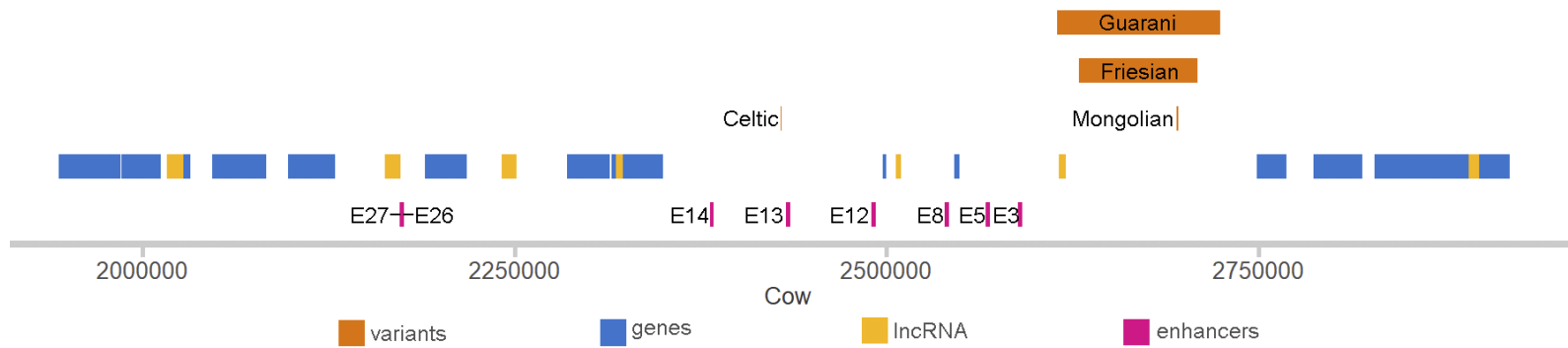
Figure A1: RIKEN FANTOM5 enhancers aligned to the *Bos taurus* reference genome (ARSUCD1.2). Only eight of 34 human enhancers from the comparative

POLLED region in humans aligned to ARS-UCD1.2.
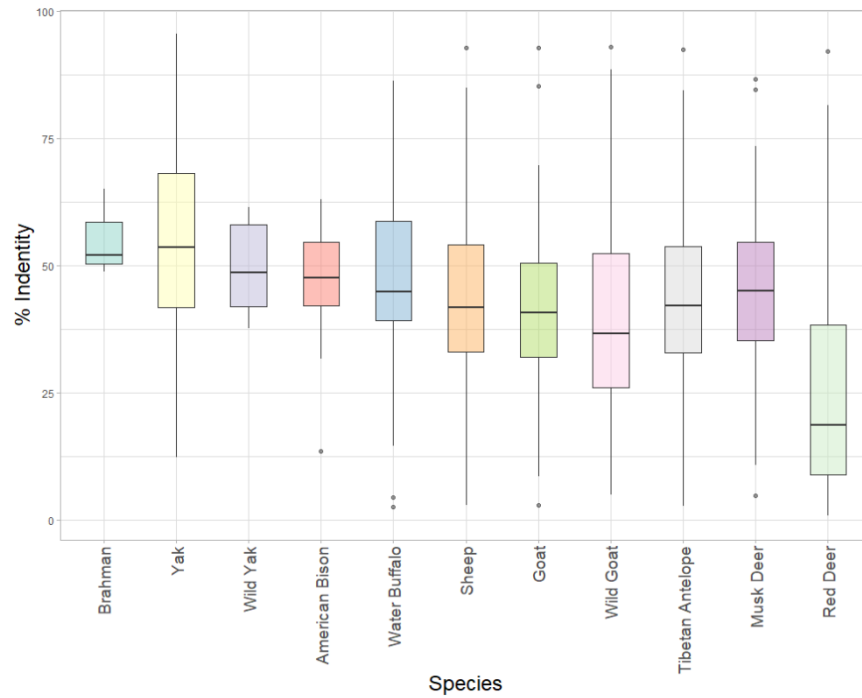
Figure A2: Boxplot of base pair identity (%) of conserved segments when the POLLED region is aligned to Pecora species via LASTZ. All species, except musk deer and red deer, have horns. Red deer have antlers and musk deer do not have head gear.
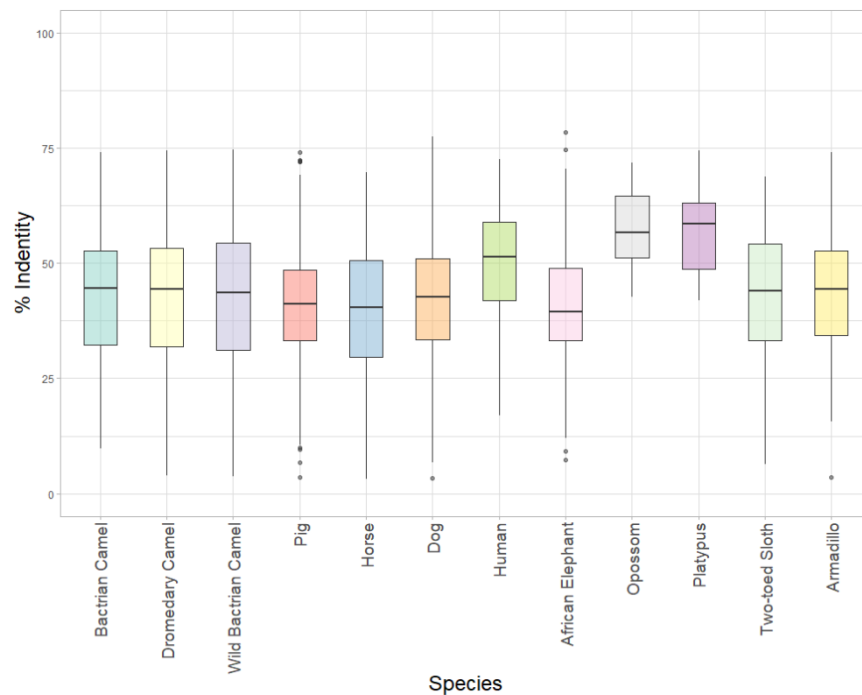
Figure A3: Boxplot of base pair identity (%) of conserved segments when the POLLED region is aligned to species without headgear via LASTZ. Note: this excluded regions where there were no matches.
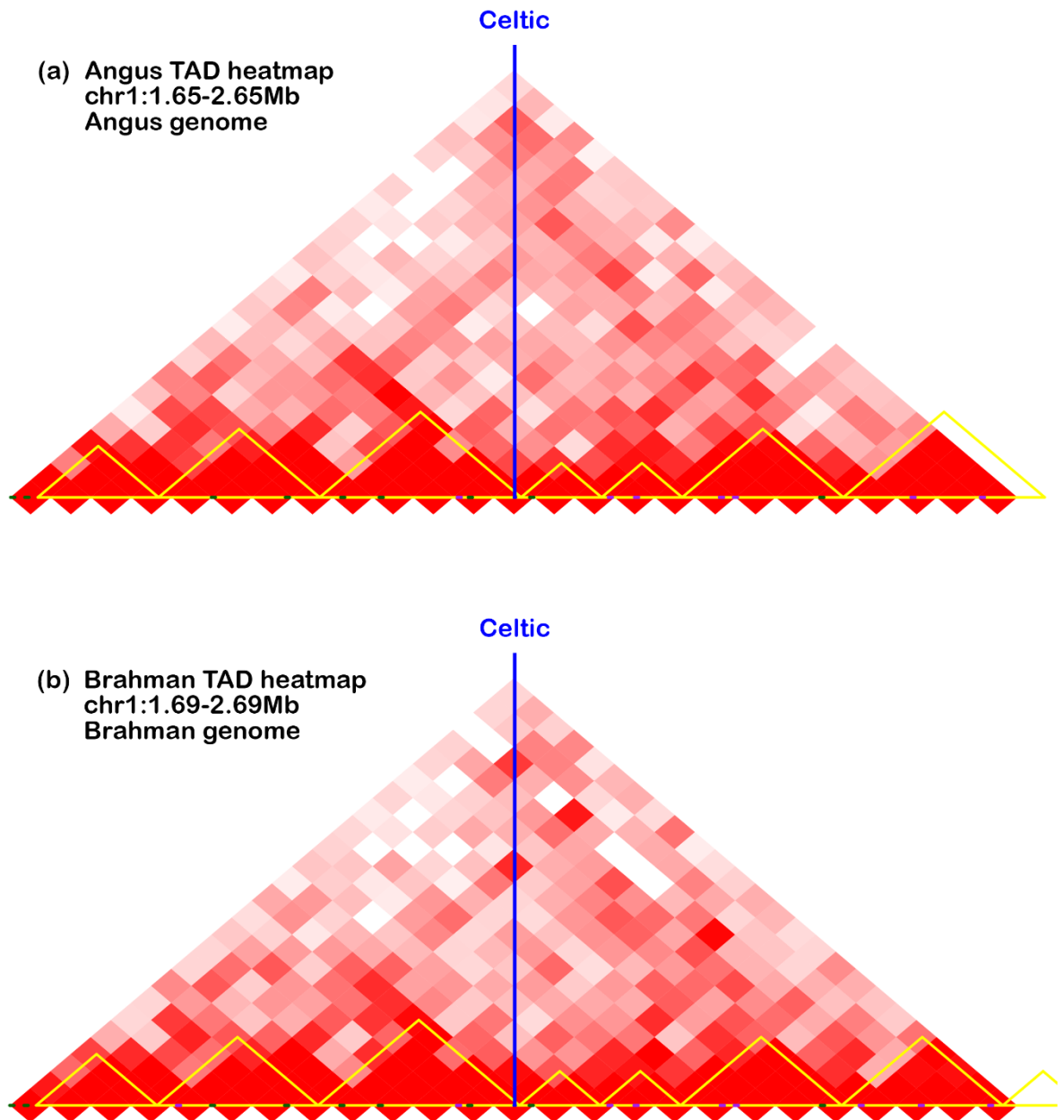
Figure A4: TAD structure in the putative POLLED TAD analysed from (a) Angus Hi-C reads aligned to UOA_Angus_1 and (b) Brahman Hi-C reads aligned UOA_Brahman_1. Hi-C reads were from the lung tissue of an Angus-Brahman F1 hybrid fetus at 90 days of development. Hi-C reads were separated into Angus- and Brahman-specific reads using a k-mer approach. The data were produced by Low *et al.* (2020) and downloaded from NCBI SRR6691720 of PRJNA432857.

Appendix B: Supplementary material for Chapter 3

Table B1: Haematoxylin and eosin staining protocol as conducted by the Histology Services at the Adelaide Health and Medical Sciences.

Reagent	Time (minutes)
Xylene	3:00
Xylene	3:00
Absolute ethanol	1:00
Absolute ethanol	1:00
Absolute ethanol	1:00
70% ethanol	1:00
Tap water	1:00
Haematoxylin (CS709)	1:00
Tap water	1:00
Blue (CS702)	1:00
Tap water	1:00
70% ethanol	1:00
Eosin (CS710)	4:30
Absolute ethanol	0:30
Absolute ethanol	1:00
Absolute ethanol	1:00
Pre-exit Xylene	1:00
Histoclear	1:00

Table B2: The number of slides of moderate to good quality obtained from each tissue from the bovine fetuses. Horn bud (HB; horned)/horn bud region (HBR; polled) and frontal skin (FS) tissues were collected at 58 days of fetal development. Morphometric measurements were taken from the sections on these slides.

Fetus	Age	Genotype	Number of slide with moderate – good quality	
			HB/HBR	FS
618	58	Horned	80	25
546	58	Horned	108	0
736	58	Horned	58	0
532	58	Horned	79	0
668	58	Horned	62	0
581	58	Horned	44	0
667	58	Polled	57	0
701	58	Polled	0	0
689	58	Polled	0	0
709	58	Polled	61	0
694	60	Polled	0	38

Table B3: Shapiro-Wilk normality test conducted on tissue layer measurements from horn bud and frontal skin from horned and polled fetuses.

Variable	P-value
Total depth	1.819e-05
Epithelium depth (as proportion of total depth)	3.83e-16
Mesenchyme depth	4.789e-09
Condensed cell depth	1.838e-10

Table B4: Shapiro-Wilk normality test conducted on image segmentation data for horn bud and frontal skin sections from horned and polled fetuses stained with SOX10, NGFR and RXFP2 antibodies.

Antibody	Variable	P-value
SOX10	SOX10 nuclei	0.24
	Background	0.45
	SOX10 nerve	0.13
	Negative nuclei	0.72
NGFR	Background	0.65
	NGFR nerve	0.22
	Negative nuclei	0.50
RXFP2	Background	0.68
	RXFP2 nerve	0.10
	Negative nuclei	0.27
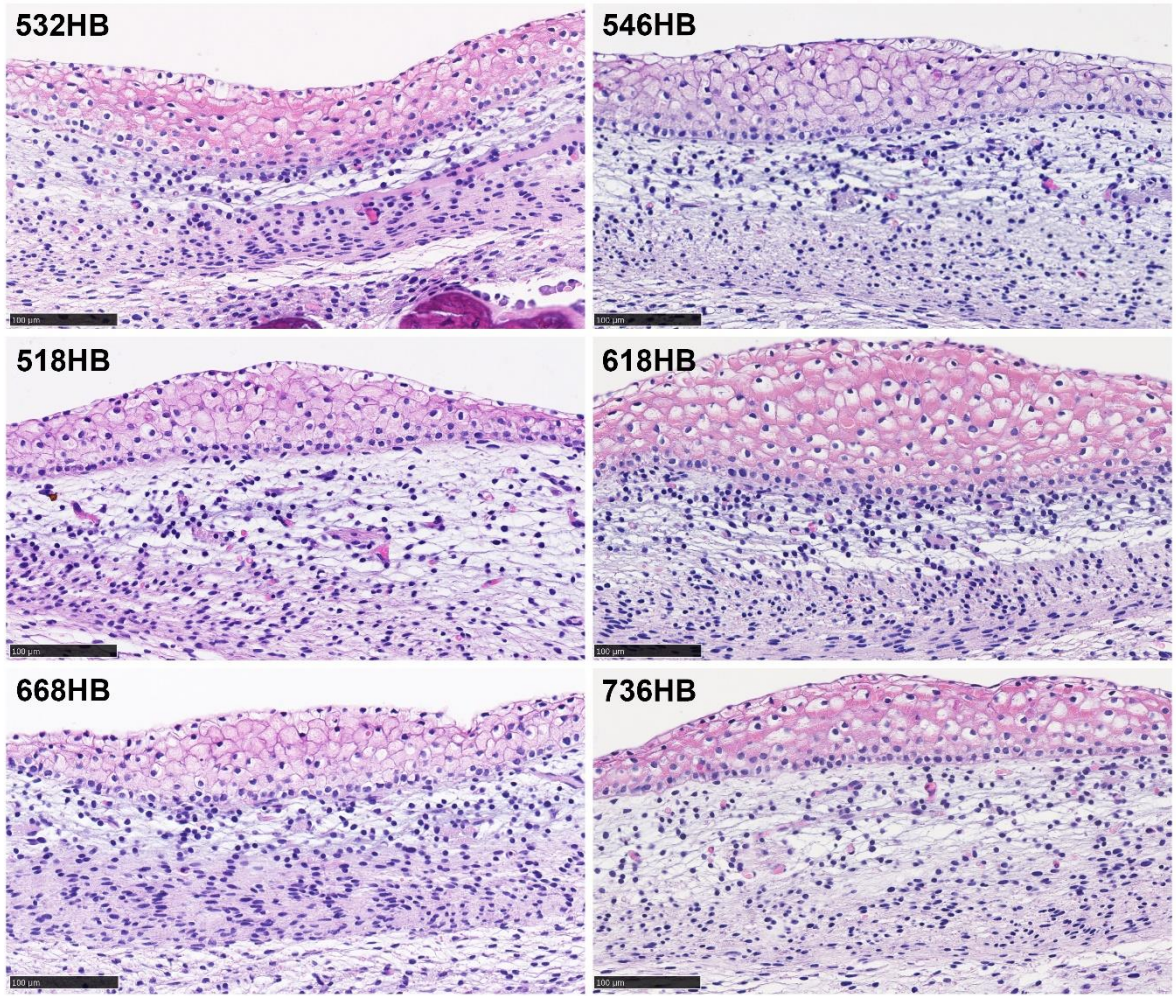
Figure B1: Haematoxylin and eosin stained sections of the horn bud centre, where the epithelial cell depth is the thickest, collected from six horned fetuses.

532HB

546HB

581HB

618HB

668HB

736HB
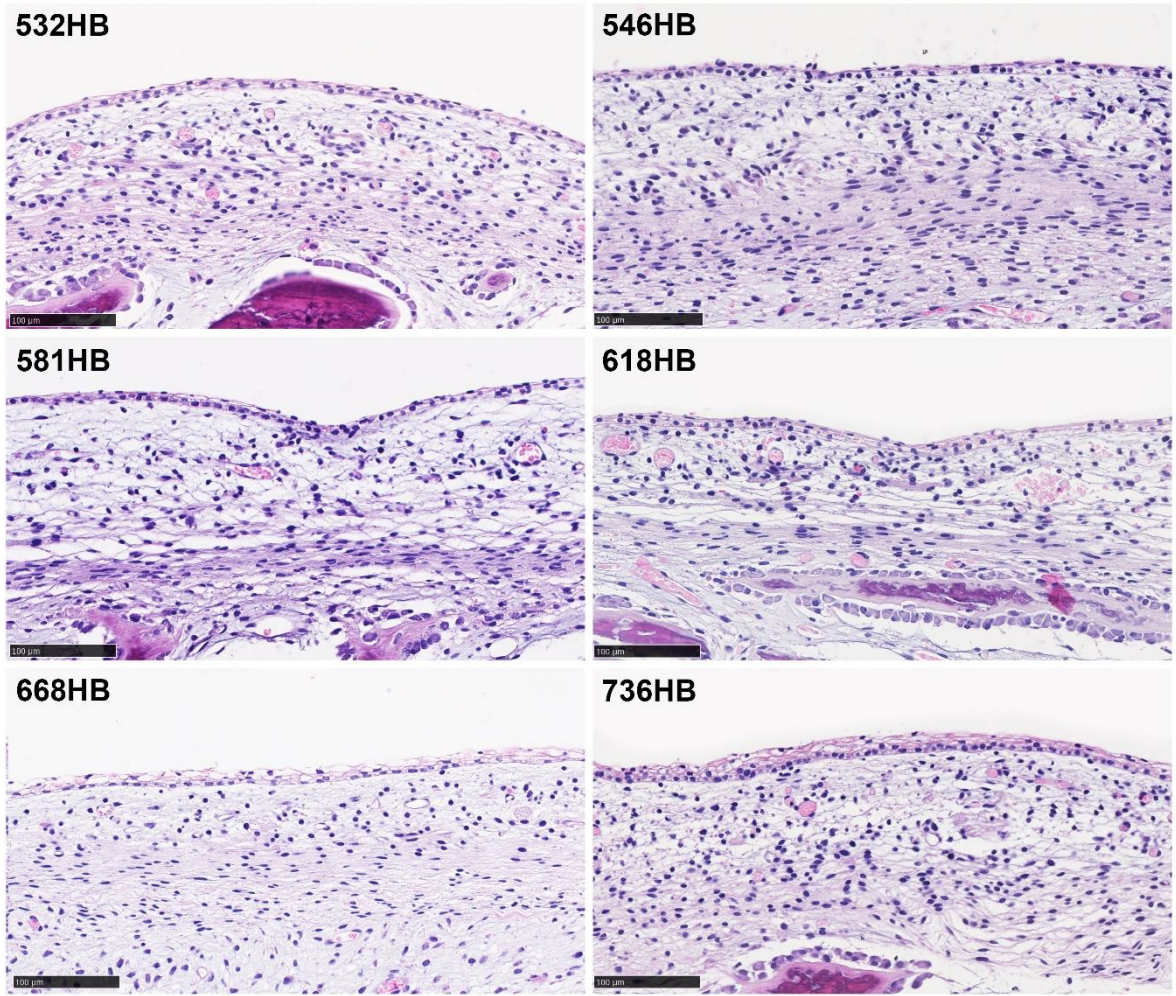
Figure B2: Haematoxylin and eosin stained sections of tissue immediately outside of the horn bud, where there is only a single epithelial layer of cells, collected from six horned fetuses.
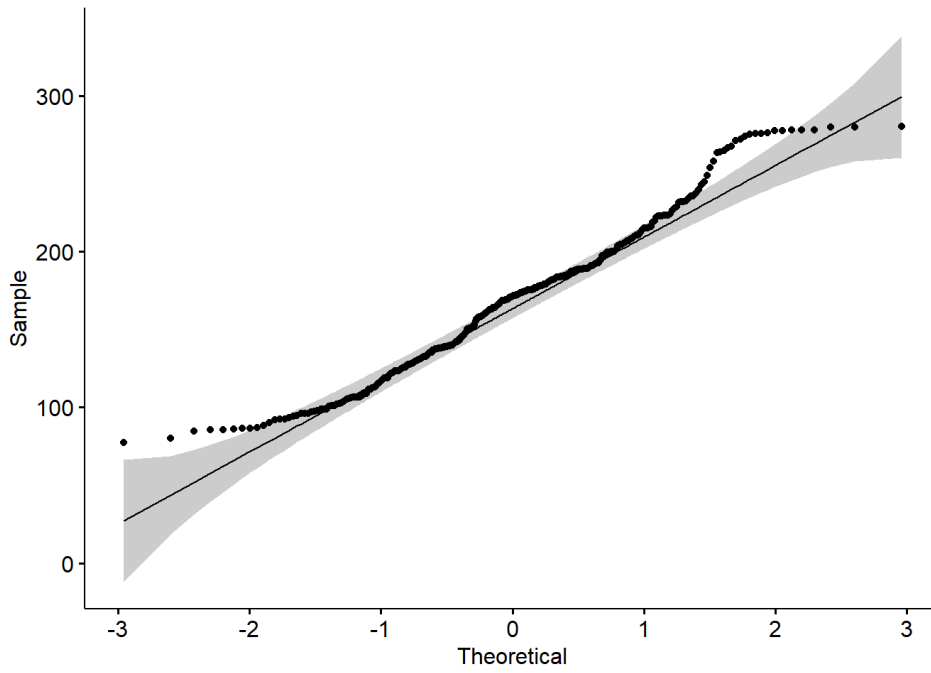
Figure B3: Q-Q plot for histomorphometric measurement of total depth from tissues collected from horned and polled fetuses. Total depth was measured from the epidermis to the end of the mesenchyme.
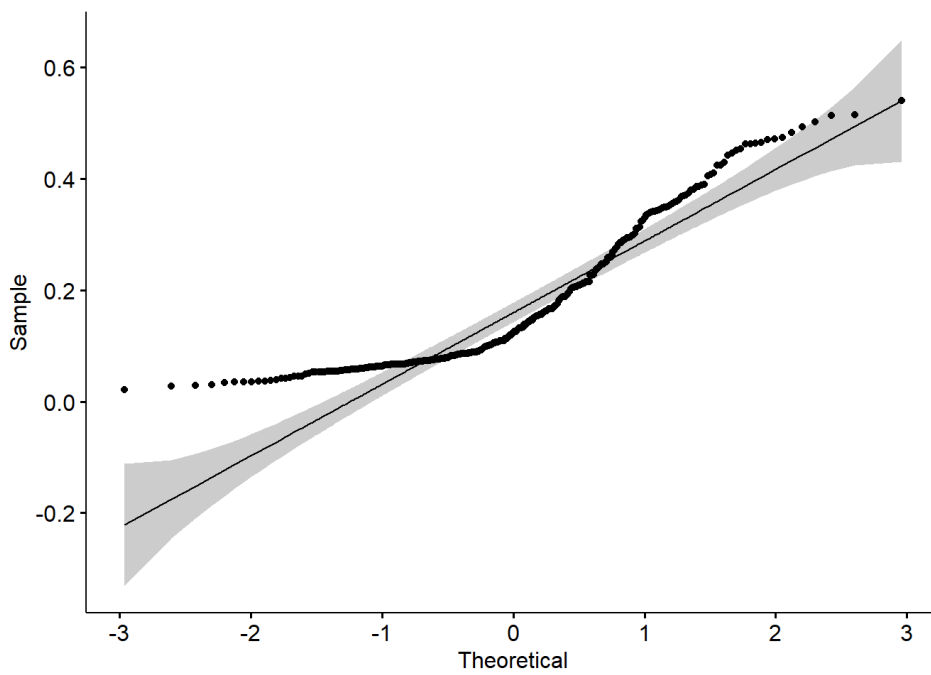
Figure B4: Q-Q plot for histomorphometric measurement of epithelium depth (as a proportion of total depth) from tissues collected from horned and polled fetuses.
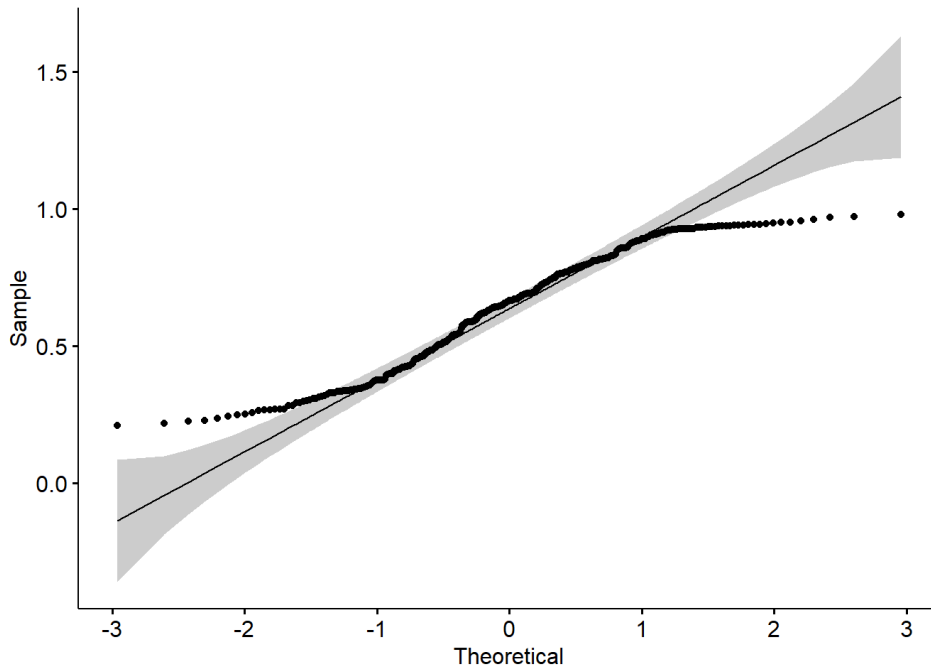
Figure B5: Q-Q plot for histomorphometric measurement of mesenchyme depth (as a proportion of total depth) from tissues collected from horned and polled fetuses. Mesenchyme depth was the depth of loose cells between the epidermis and condensed cells.
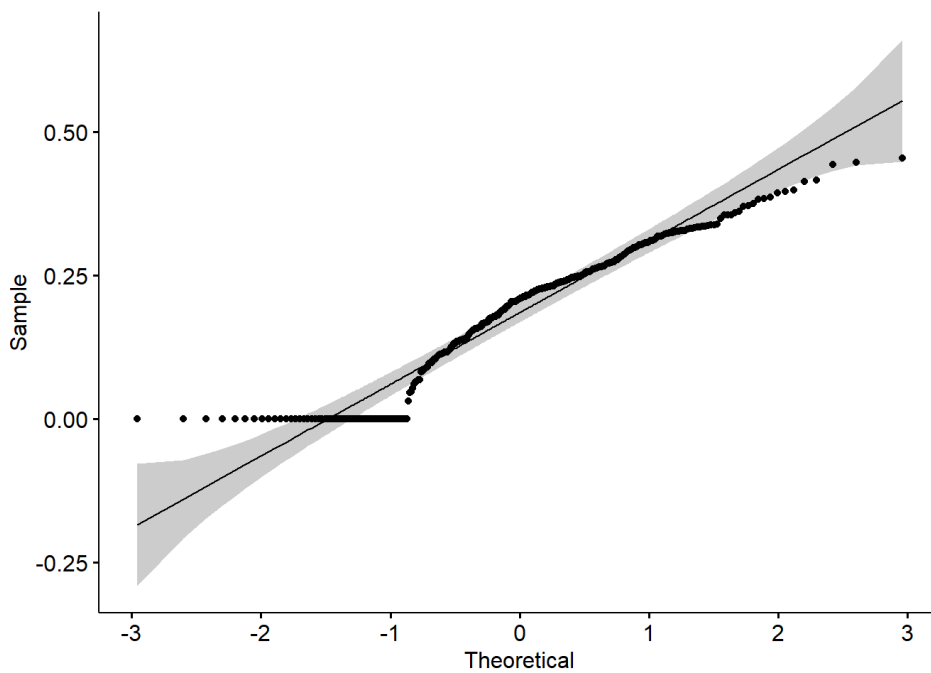
Figure B6: Q-Q plot for histomorphometric measurement of condensed cell depth (as a proportion of total depth) from tissues collected from horned and polled fetuses. Condensed cell depth was the depth of tightly compacted cells within the mesenchyme.
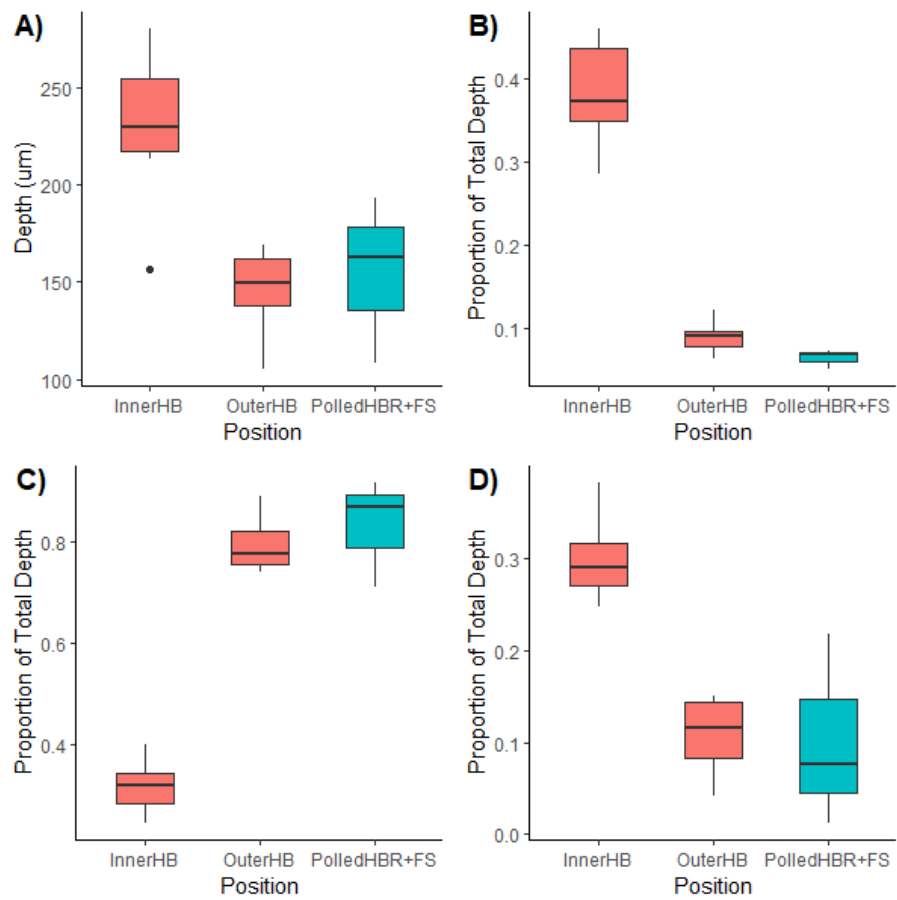
Figure B7: Boxplots displaying the range of data points for A) the total depth, B) epithelium depth, C) mesenchyme depth and D) condensed cell depth for the inner horn bud (InnerHB), outer horn bud (OuterHB) and tissue from the polled horn bud region and frontal skin (PolledHBR+FS). Tissue was collected from bovine fetuses at 58 days of development. B-D) Tissue depths are reported as proportions of the total depth. Box represents the interquartile range and the whiskers represent the smallest and largest values within 1.5 times the range of the 25th and 75th percentile, respectively.
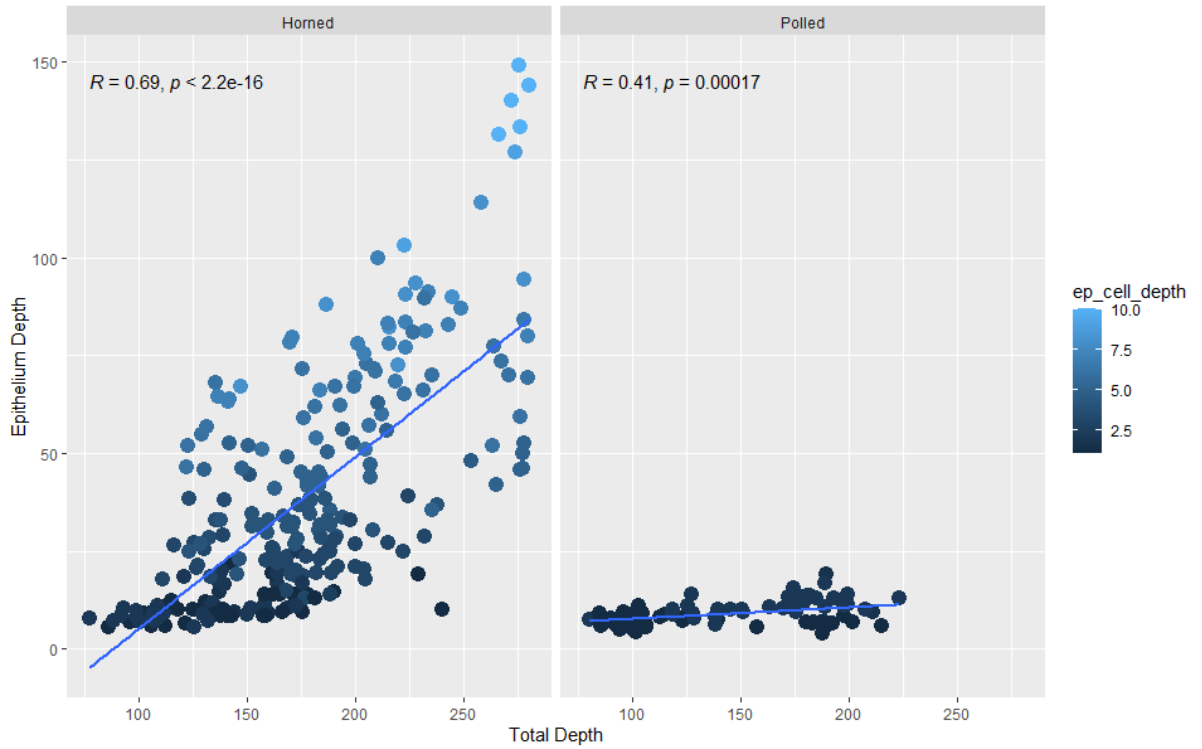
Figure B8: Scatterplot of total depth and epithelium depth of tissue from the horn bud of horned and horn bud region of polled fetuses at 58 days of development. The total tissue depth included the epithelium, mesenchyme and condensed cells.
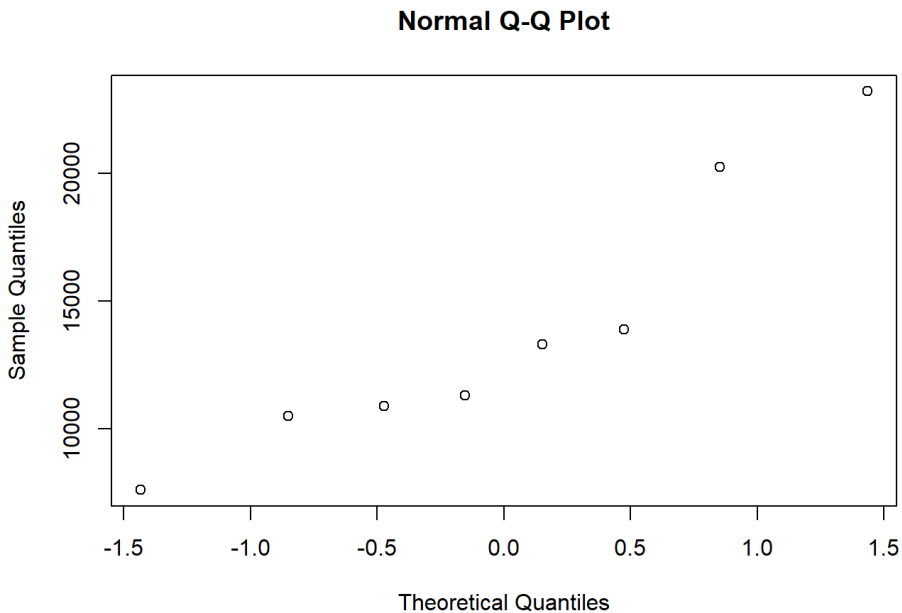
Figure B9: Q-Q plot for area of positive nuclei stained with SOX10 antibody in the horn bud and frontal skin samples from horned and polled fetuses at 58 days of development.
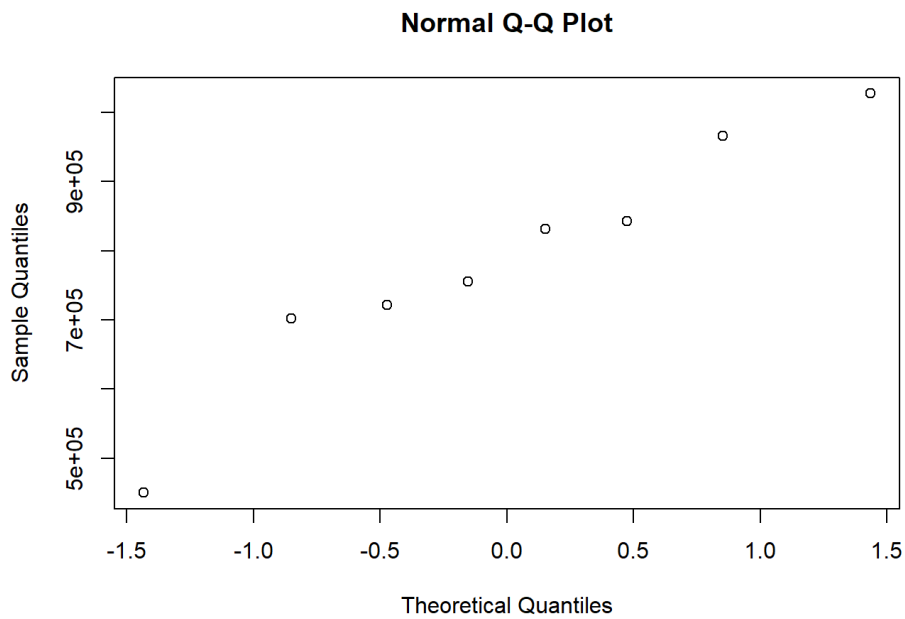
Normal Q-Q Plot
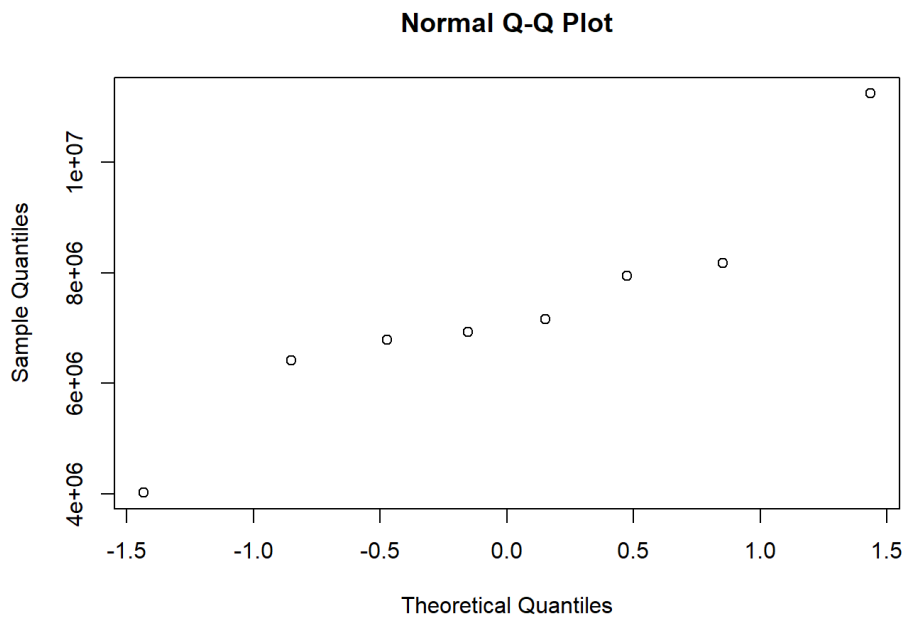
Figure B10: Q-Q plot for area of background from sections stained with SOX10 antibody in the horn bud and frontal skin samples from horned and polled fetuses at 58 days of development.
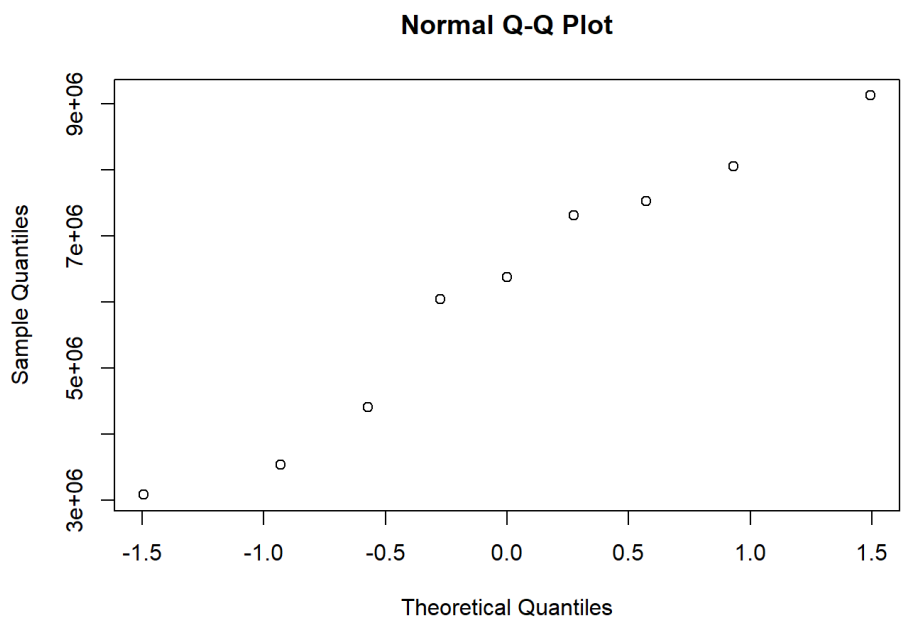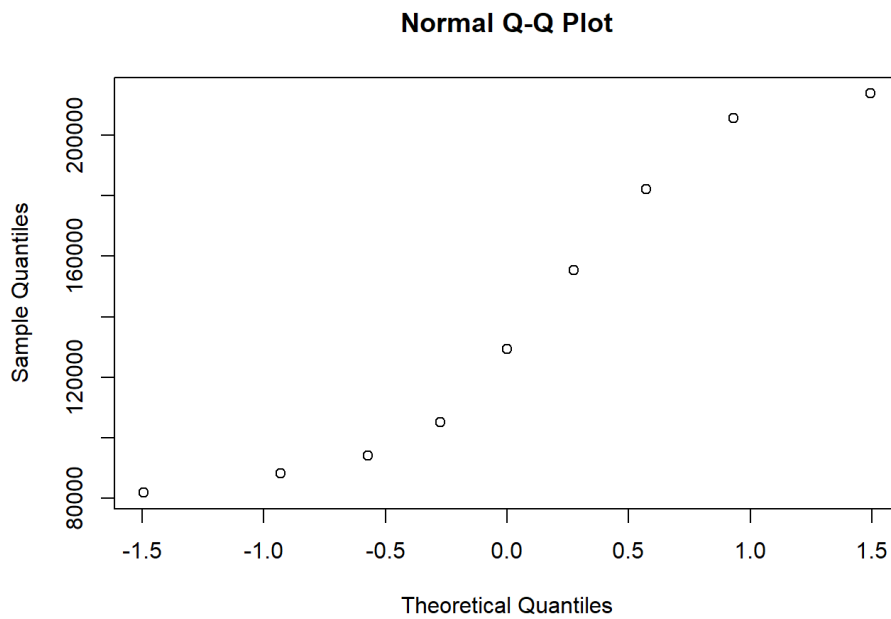

Normal Q-Q Plot
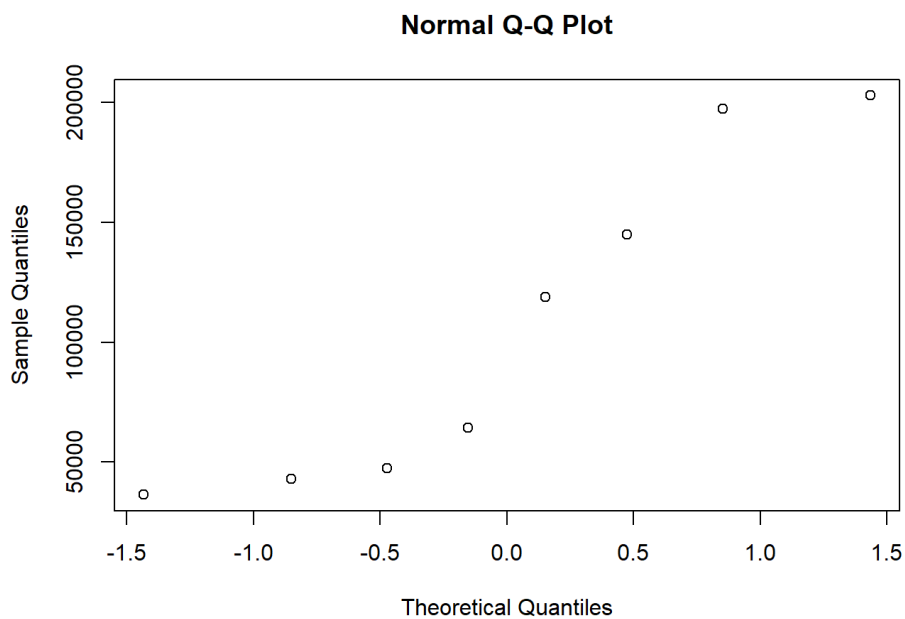
Figure B11: Q-Q plot for area of positive nerve regions stained with SOX10 antibody in the horn bud and frontal skin samples from horned and polled fetuses at 58 days of development.
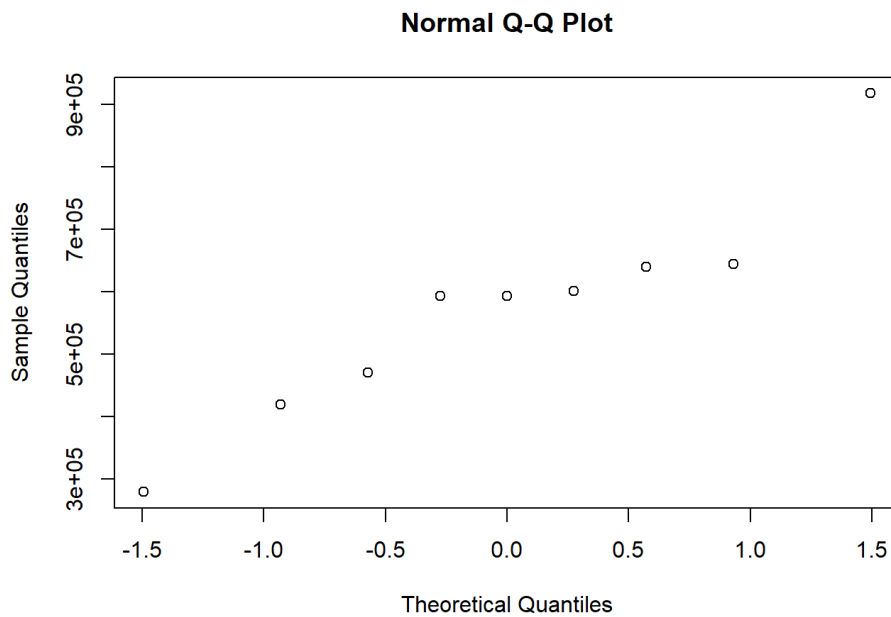
Normal Q-Q Plot

Figure B12: Q-Q plot for area of negative nuclei from sections stained with SOX10 antibody in the horn bud and frontal skin samples from horned and polled fetuses at 58 days of development.
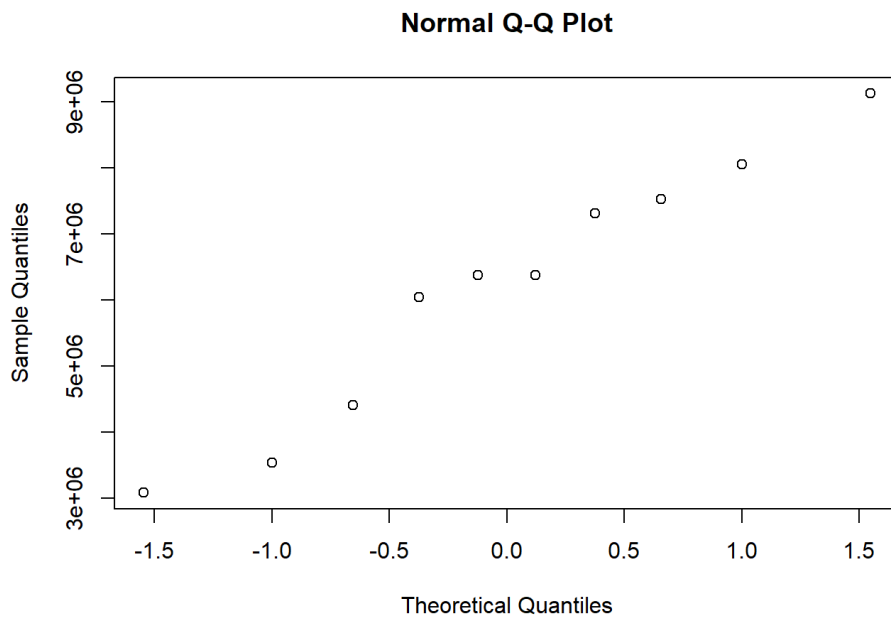
Normal Q-Q Plot

Figure B13: Q-Q plot for area of background from sections stained with NGFR antibody in the horn bud and frontal skin samples from horned and polled fetuses at 58 days of development.

Normal Q-Q Plot

Figure B14: Q-Q plot for area of positive nerve regions stained with NGFR antibody in the horn bud and frontal skin samples from horned and polled fetuses at 58 days of development.

Normal Q-Q Plot

Figure B15: Q-Q plot for area of negative nuclei from sections stained with NGFR antibody in the horn bud and frontal skin samples from horned and polled fetuses at 58 days of development.

Normal Q-Q Plot

Figure B16: Q-Q plot for area of background from sections stained with RXFP2 antibody in the horn bud and frontal skin samples from horned and polled fetuses at 58 days of development.
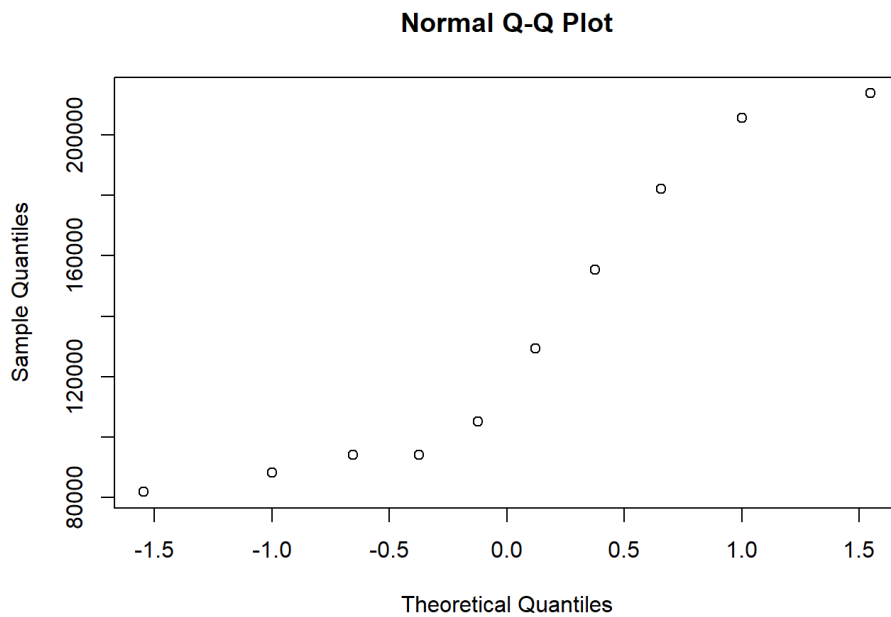
Normal Q-Q Plot

Figure B17: Q-Q plot for area of positive nerve regions stained with RXFP2 antibody in the horn bud and frontal skin samples from horned and polled fetuses at 58 days of development.
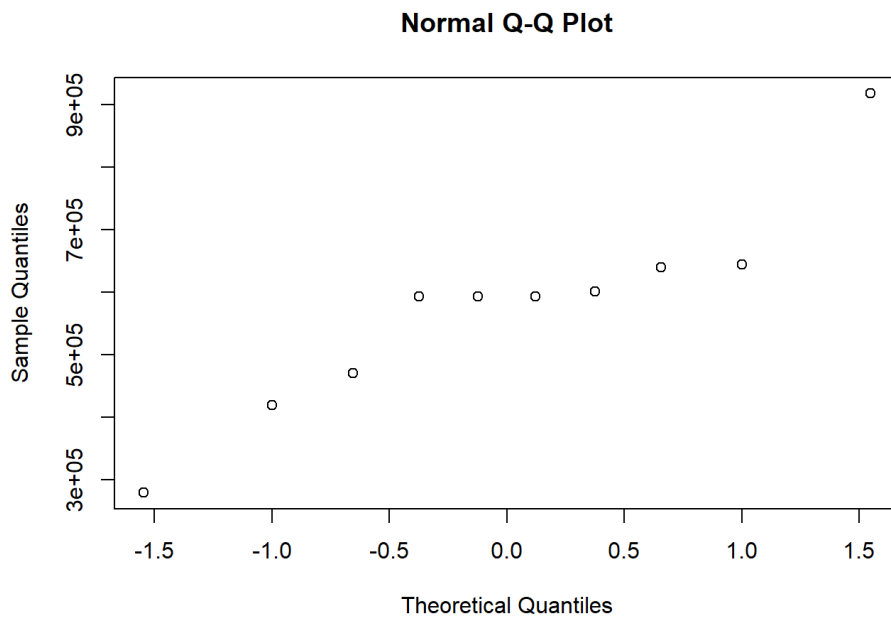
Normal Q-Q Plot

Figure B18: Q-Q plot for area of negative nuclei from sections stained with RXFP2 antibody in the horn bud and frontal skin samples from horned and polled fetuses at 58 days of development.
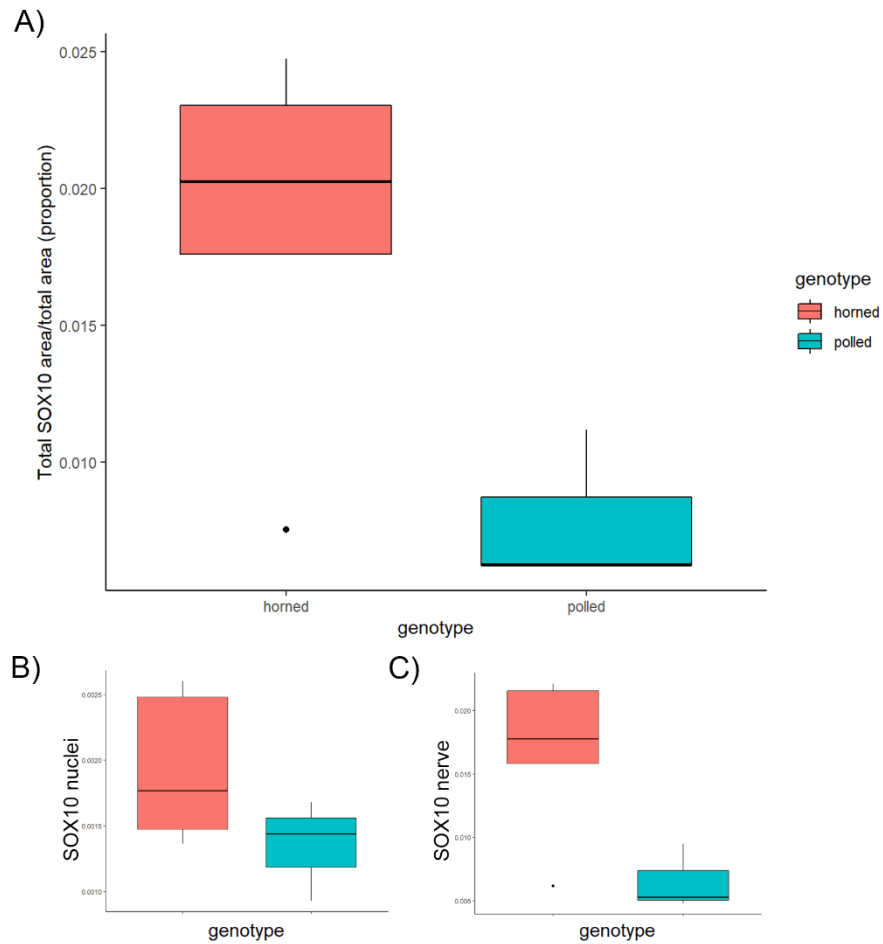
Figure B19: Boxplots displaying range of area as a proportion of total area for A) positive nuclei and nerves, B) positive nuclei and C) positive nerves from SOX10 antibody stained sections of horn bud from horned (n = 6) and horn bud region and frontal skin of polled fetuses (n = 3) at 58 days of development. Note: one sample from the polled group was collected at 60 days of development. Areas are reported as proportion of the total area. Box represents the interquartile range and the whiskers represent the smallest and largest values within 1.5 times the range of the 25th and 75th percentile, respectively.
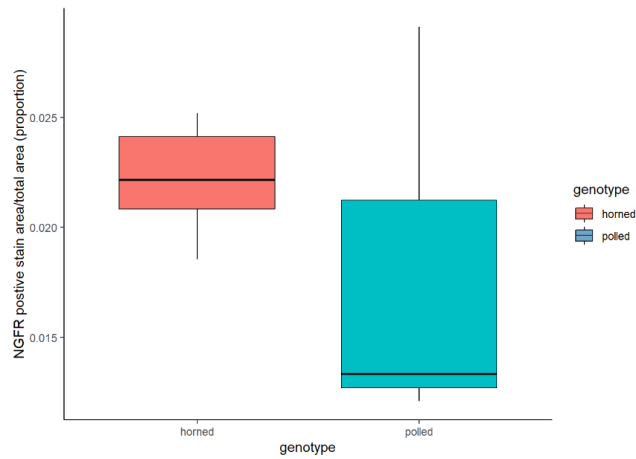
Figure B20: Boxplots displaying range of data for area of positive nerves from NGFR antibody stained sections of horn bud from horned (n = 6) and horn bud region and frontal skin of polled fetuses (n = 3) at 58 days of development. Note: one sample from the polled group was collected at 60 days of development. The area is reported as proportion of the total area. Box represents the interquartile range and the whiskers represent the smallest and largest values within 1.5 times the range of the 25th and 75th percentile, respectively.

Figure B21: Boxplots displaying range of data for area of positive nerves from RXFP2 antibody stained sections of horn bud from horned (n = 6) and horn bud region and frontal skin of polled fetuses (n = 3) at 58 days of development. Note: one sample from the polled group was collected at 60 days of development. The area is reported as proportion of the total area. Box represents the interquartile range and the whiskers represent the smallest and largest values within 1.5 times the range of the 25th and 75th percentile, respectively.
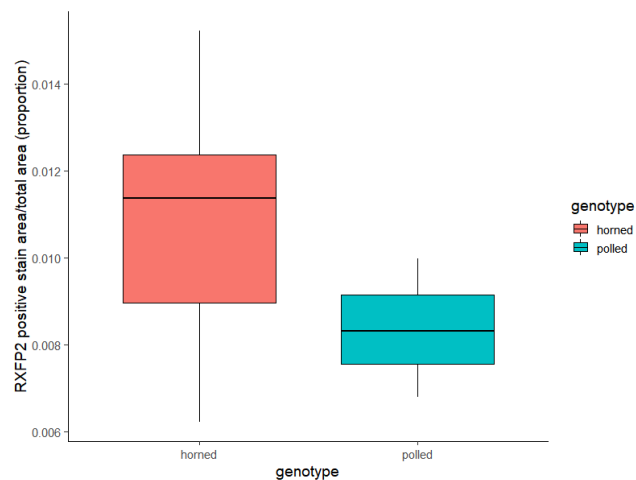
Appendix C: Supplementary material for Chapter 4

Section C1: Description of WNT signalling

Genes involved in WNT signalling were upregulated in rat gubernacular cells after being exposed to the RXFP2 ligand, INSL3 (Johnson et al. 2010). Canonical WNT signalling can be activated by WNT3A, WNT4, WNT6 WNT7B, WNT8B, WNT9A and WNT10B (Benhaj *et al.* 2006). When WNT is not present, β-catenin is bound to a destruction complex consisting of several proteins (axis inhibition protein [AXIN], adenomatous polyposis coli [APC] tumour suppressor, casein kinase 1 [CK1] and glycogen synthase kinase 3β [GSK3β]) (Wiese et al. 2018). In the canonical WNT signalling pathway, when WNT binds to Frizzled receptors (FZD), β-catenin dissociates from the destruction complex, allowing β-catenin to accumulate in the cytosol and nucleus. In the nucleus, β-catenin activates transcription by binding to TCF/LEF (Wiese et al. 2018). These transcription factors act as repressors in the absence of β-catenin.

SFRP2 and AXIN2 regulate WNT signaling. At high concentrations, SFRP2 act as an antagonist to WNT signaling in mouse mesenchymal stem cells (Alfaro *et al.* 2010), but at low levels, SFRP2 functions as an agonist promoting WNT signaling in various tissue (Van Loon *et al.* 2021). AXIN2 regulates WNT signaling through a negative feedback loop (Jho *et al.* 2002; Leung *et al.* 2002).

Section C2: Description of BMP signaling

Expression of the BMP signaling pathway genes were increased by INSL3/RXFP2 in rat gubernacular cells (Johnson et al. 2010. In the canonical pathway, BMP binds to its receptor, which is a heterotetrameric complex consisting of a dimer of type I (bone morphogenetic protein receptor type 1B [BMPR1B] and activin A receptor like type 1 [ACVRL1]) and a dimer of type II (bone morphogenetic protein receptor type 2 [BMPR2] and activin A receptor type 2B [ACVR2B]) serine/threonine kinase receptors (Wang et al. 2014a; Montanari et al. 2021). This binding leads to the phosphorylation of receptor-regulated SMADs, namely SMAD family members 1/5/8 (SMAD1/5/8) (Wang et al. 2014a; Montanari et al. 2021). The receptor-regulated SMADs associate with the co-mediator, SMAD4, and then can act as nuclear transcription factors (Wang et al. 2014a; Montanari et al. 2021). The pathway is regulated by SMAD6 and SMAD7, which act as inhibitors of chordin (CHRD) and twisted gastrulation BMP signaling modulator 1 (TWSG1) (Montanari et al. 2021).

Section C3: Description of HGF/MET signaling

HGF/MET Proto-Oncogene, Receptor Tyrosine Kinase (MET) signaling is also upregulated in response to INSL3/RXFP2 in rat gubernacular cells (Johnson et al. 2010). HGF/MET interactions contribute to morphogenesis, cell survival, proliferation and motility. HGF/MET is also involved in various aspects of neural development, including neural induction (Bronner-Fraser 1995), neuron survival, guidance and axon outgrowth (Thompson et al. 2004), and acts as a mitogen for rat Schwann cells (Krasnoselsky et al. 1994). Furthermore, HFG/MET induces morphogenic changes in epithelial cells in a tissue-specific manner, such as the enterocyte-like brush border and junction complexes in SW 1222 cells (human colon) (Brinkmann et al. 1995). The various roles of HGF/MET overlap with changes observed in the horn bud and occur through the activation of downstream pathways, such as STAT3, PI3K/AKT, MEK/ERK and JNK signaling (Liu et al. 2020b).

Section C4: RNA quality check

To assess RNA quality, the RNA concentration, OD 260/280 ratio and RNA integrity number (RIN) were considered. The concentration for some tissues were below 10 ng/µl (Appendix Table C30). In addition, the RNA integrity number (RIN value) of these samples was also low (< 3), indicating highly degraded material (Appendix Figure C10). The Aligent Bioanalyser manual states that below 25 ng/µl, no accurate RIN may be obtained (Mueller *et al.* 2016). The ratio of absorbance at 260 nm and 280 nm is used to assess the purity of DNA and RNA. A ratio of ~2.0 is generally accepted as "pure" for RNA. In many cases, the 260/280 ratio detected was < 2.0, suggesting that there are potential impurities such as proteins or phenols However, the samples with a low RIN value also had low RNA concentrations. The library preparation kit, SMART-Seq Stranded kit (Takara Bio USA Inc.), was chosen because of the low concentration (recommended input is 10 pg–10 ng) and quality RNA. This meant that good quality sequencing data could be generated from the RNA.

Section C5: Functional investigation of *C1H21orf62*

The Human Protein atlas reports that the human ortholog, *C21orf62*, is expressed in the brain (particularly the choroid plexus), epididymis in males and ovaries in females (Uhlén et al. 2015). *C21orf62* may be similar to *WNT7B*, *PHC3* and *MTHFR* according to GenesLikeMe (Stelzer et al. 2016), however, the scores are low (0.71/8, 0.67/8 and 0.65/8). *C21orf62*, *WNT7B* and *MTHFR* are part of the skeleton phenotype term in the Mammalian Phenotype Browser (Smith & Eppig 2009) therefore, *C21orf62* could also have a function in bones. The protein structure of *C21orf62* (Appendix Figure C12A), includes an N-terminal signal peptide (amino acids 1-19) suggesting this protein is secreted (Bateman et al. 2021). The human and bovine amino acid sequence has 78.5% identity, and some amino acid changes are in the signal peptide. These amino acid changes may alter translocation efficiency (Owji et al. 2018).

Section C6: Headgear-specific genes

The headgear-specific genes are 'recruited' from other pathways (such as bone development) and were recruited for horn and antler development (Wang et al. 2019b). The horn-specific genes have been identified in transcriptome analyses of sheep and goats, which included data from seven goat and three sheep horn sprouts (Wang et al. 2019b). Horn-specific genes had a t index > 0.8 and were most highly or second most highly expressed in horn tissue. Antler-specific genes were identified from transcriptomes from 20 roe deer and 20 sika deer, and included data from neonatal antlers (Wang et al. 2019b). Antler-specific genes had a t index > 0.8 and were most highly or second most highly expressed in antler tissue. Genes common between the horn-specific gene list (n = 624) and antler-specific gene list (n = 761) were labeled as headgear-specific genes (n = 201) (Wang et al. 2019b). GO and KEGG pathway analyses were conducted on the headgear-specific genes.

Table C1-Table C10 are available at https://figshare.com/: <DOI: 10.25909/19319966 >

Table C1: Differentially expressed genes (p<0.05, logFC>1) between horn bud (HB) and frontal skin (FS) of horned fetuses (pp; n = 4) (reference = FS) at 58 days of development.

Table C2: Differentially expressed genes (p<0.05, logFC>1) between forebrain skin (FB) and horn bud (HB) of horned fetuses (pp; n = 4) (reference = HB) at 58 days of development.

Table C3: Differentially expressed genes (p<0.05, logFC>1) between midbrain skin (MB) and horn bud (HB) of horned fetuses (pp; n = 4) (reference = HB) at 58 days of development.

Table C4: Differentially expressed genes (p<0.05, logFC>1) between horn bud (HB) and frontal skin (FS) of polled (PP; n = 3) fetuses (reference = FS) at 58 days of development.

Table C5: Differentially expressed genes (p<0.05, logFC>1) between forebrain skin (FB) and horn bud region (HBR) of polled fetuses (PP; n = 3) (reference = HB) at 58 days of development.

Table C6: Differentially expressed genes (p<0.05, logFC>1) between midbrain skin (MB) and horn bud region (HBR) of polled fetuses (PP; n = 3) (reference = HB) at 58 days of development.

Table C7: Differentially expressed genes (p<0.05, logFC>1) between horn bud (HB) of horned fetuses (pp; n = 4) and horn bud region (HBR) of polled fetuses (PP; n = 3) (reference = ppHB) at 58 days of development.

Table C8: Differentially expressed genes (p<0.05, logFC>1) between frontal skin (FS) of horned fetuses (pp; n = 4) and frontal skin (FS) of polled fetuses (PP; n = 3) (reference = ppFS) at 58 days of development.

Table C9: Differentially expressed genes (p<0.05, logFC>1) between forebrain skin (FB) of horned fetuses (pp; n = 4) and forebrain skin (FB) of polled fetuses (PP; n = 3) (reference = ppFB) at 58 days of development.

Table C10: Differentially expressed genes (p<0.05, logFC>1) between midbrain skin (MB) of horned fetuses (pp; n = 4) and midbrain skin (MB) of polled fetuses (PP; n = 3) (reference = ppMB) at 58 days of development.

Table C11-Table C26 are available at https://figshare.com/: <DOI: 10.25909/19335482>

Table C11: Enriched GO pathways for DE genes between forebrain skin (FB) and horn bud (HB) of horned fetuses (pp). DE genes were analysed from RNAseq data collected from horned fetuses (n = 4) at 58 days of development.

Table C12: Enriched KEGG pathways for DE genes between forebrain skin (FB) and horn bud (HB) of horned fetuses (pp). DE genes were analysed from RNAseq data collected from horned fetuses (n = 4) at 58 days of development.

Table C13: Enriched GO for DE genes between midbrain skin (MB) and horn bud (HB) of horned fetuses (pp) (reference = HB). DE genes were analysed from RNAseq data collected from horned fetuses (n = 4) at 58 days of development.

Table C14: Enriched KEGG pathways for DE genes between midbrain skin (MB) and horn bud (HB) of horned fetuses (pp) (reference = HB). DE genes were analysed from RNAseq data collected from horned fetuses (n = 4) at 58 days of development.

Table C15: Enriched GO pathways for DE genes between forebrain skin (FB) and horn bud (HB) of polled fetuses (PP) (reference = HB). DE genes were analysed from RNAseq data collected from polled fetuses (n = 3) at 58 days of development.

Table C16: Enriched KEGG pathways for DE genes between forebrain skin (FB) and horn bud (HB) of polled fetuses (PP) (reference = HB). DE genes were analysed from RNAseq data collected from polled fetuses (n = 3) at 58 days of development.

Table C17: Enriched GO pathways for DE genes between midbrain skin (MB) and horn bud (HB) of polled fetuses (PP) (reference = HB). DE genes were analysed from RNAseq data collected from polled fetuses (n = 3) at 58 days of development.

Table C18: Enriched KEGG pathways for DE genes between midbrain skin (MB) and horn bud (HB) of polled fetuses (PP) (reference = HB). DE genes were analysed from RNAseq data collected from polled fetuses (n = 3) at 58 days of development.

Table C19: Enriched GO pathways for DE genes between horn bud (HB) of horned fetuses (pp) and horn bud (HB) of polled fetuses (PP) (reference = ppHB). DE genes were analysed from RNAseq data collected from polled fetuses (n = 3) at 58 days of development.

Table C20: Enriched KEGG pathways for DE genes between horn bud (HB) of horned fetuses (pp) and horn bud (HB) of polled fetuses (PP) (reference = ppHB). DE genes were analysed from RNAseq data collected from horned (n = 4) and polled fetuses (n = 3) at 58 days of development.

Table C21: Enriched GO pathways for DE genes between frontal skin (FS) of horned fetuses (pp) and frontal skin (FS) of polled fetuses (PP) (reference = ppFS). DE genes were analysed from RNAseq data collected from horned (n = 4) and polled fetuses (n = 3) at 58 days of development.

Table C22: Enriched KEGG pathways for DE genes between frontal skin (FS) of horned fetuses (pp) and frontal skin (FS) of polled fetuses (PP) (reference = ppFS). DE genes were analysed from RNAseq data collected from horned (n = 4) and polled fetuses (n = 3) at 58 days of development.

Table C23: Enriched GO pathways for DE genes between forebrain skin (FB) of horned fetuses (pp) and forebrain skin (FB) of polled fetuses (PP) (reference = ppFB). DE genes were analysed from RNAseq data collected from horned (n = 4) and polled fetuses (n = 3) at 58 days of development.

Table C24: Enriched KEGG pathways for DE genes between forebrain skin (FB) of horned fetuses (pp) and forebrain skin (FB) of polled fetuses (PP) (reference = ppFB). DE genes were analysed from RNAseq data collected from horned (n = 4) and polled fetuses (n = 3) at 58 days of development.

Table C25: Enriched GO pathways for DE genes between midbrain skin (MB) of horned fetuses (pp) and midbrain skin (MB) of polled fetuses (PP) (reference = ppMB).

Table C26: Enriched KEGG pathways for DE genes between midbrain skin (MB) of horned fetuses (pp) and midbrain skin (MB) of polled fetuses (PP) (reference = ppMB). DE genes were analysed from RNAseq data collected from horned (n = 4) and polled fetuses (n = 3) at 58 days of development.

Table C27: Functions of neuronal genes differentially expressed between the horn bud and frontal skin of horned fetuses at 58 days of development (n = 4).

Gene	DE[a]	Functions	Citation
BDNF	↑	Modulates neuronal activity, neurite outgrowth and differentiation	Zagrebelsky and Korte (2014)
CDH7	↑	Expressed in cranial neural crest cells and involved with trigeminal nerve development	Wu and Taneyhill (2019)
CNTFR	↑	Neurotropic factor which plays a role in neuronal cell survival, and differentiation	Vlotides *et al.* (2004)
EPHA3	↑	Regulate axon guidance of neuronal growth cones, and selective bundling and dispersal of axons	Klein (2004).
FRMD7	↑	Neurite outgrowth, survival, synapse formation and function, and determination and differentiation	Watkins *et al.* (2012)
NTNG1	↑	Axonal guidance and N-methyl-D-aspartate receptor signaling in neurons	Lin *et al.* (2003); Archer *et al.* (2006)
RORB	↑	Differentiation of cerebral cortex and the retina	Liu *et al.* (2017)
TAFA1	↑	Promotes neuronal cell differentiation (also known as FAM19A1)	Sarver *et al.* (2021)
TMEM59L	↑	Mediates neuronal cell death	Zheng *et al.* (2017)
ARPP21	↓	Regulates dendritic branching	Rehfeld *et al.* (2018)
ATP10A	↓	May be involved with myelination	Uhlén *et al.* (2015)
CBLN4	↓	Synapse formation	Yuzaki (2008)
CDH15	↓	Cell adhesion molecule expressed in the cerebellum and pons	Uhlén *et al.* (2015)
FABP7	↓	Fatty acid binding protein expressed in the central nervous system	Yun *et al.* (2012); Killoy *et al.* (2020)
GPM6A	↓	Synapse formation	León *et al.* (2021)
GRIA2	↓	Receptor for excitatory neurotransmitters and expressed in most brain regions	Uhlén *et al.* (2015).
HCN1	↓	Sensory transduction	Barravecchia and Demontis (2021)

Table C27: (Continued)

Gene	DE[a]	Functions	Citation
LHX8	↓	Regulator of cholinergic neuronal function	Tomioka *et al.* (2014)
LMO1	↓	Hindbrain patterning and adult central nervous system function	Hinks *et al.* (1997); Matis *et al.* (2007)
LMO3	↓	Amygdala development and function	Reisinger *et al.* (2020)
LRRC7	↓	Component of the post-synaptic density of excitatory synapses in central nervous system	Liu *et al.* (2013); Kim *et al.* (2020)
NEFL	↓	Neurofilament, cytoskeleton protein	Wang *et al.* (2019b)
NETO2	↓	Interacts with kainite receptor subunits at the post synaptic density	Mennesson *et al.* (2019)
NEUROG2	↓	Differentiation of neural crest cells to neurons rather than glial cells in dorsal root ganglia	(Liu *et al.* 2020a)
NRG1	↓	Neuronal function (fear, memory) Regulates the behavior of myelinating Schwann cells in peripheral nervous system	Mei and Nave (2014); Chen *et al.* (2021)
SCN3A	↓	Voltage-gated ion channel responsible for generating action potentials in neurons	Uhlén *et al.* (2015).
SGIP1	↓	Synaptic protein that acts as an selective endocytic adapter	Lee *et al.* (2019)
SHROOM3	↓	Inhibition of axon outgrowth	Taylor *et al.* (2008)
SLAIN1	↓	Involved in microtubule elongation and expressed in the developing nervous system	Hirst *et al.* (2010); Van Der Vaart *et al.* (2011)
SLCO1A2	↓	A sodium-independent transporter of organic ions, steroid hormones and thyroid hormones in the brain	Zhou *et al.* (2015)
ST8SIA2	↓	Weakens cell-cell contacts by decreasing the adhesive properties of neural cell adhesion molecule 1 (NCAM1) in neural crest cells.	Szewczyk *et al.* (2017)
SYBU	↓	Synaptic function	Xiong *et al.* (2021)

[a]DE = differential expression in horn bud compared to frontal skin

Table C28: Differential expression of HGF/MET signalling genes between tissues collected from horned (n = 4) and polled fetuses (n = 3) at 58 days of development.

HGF/MET signaling genes	Horn Bud vs Frontal Skin		Horn Bud vs Forebrain Skin		Horn Bud vs Midbrain Skin		Horn vs Polled			
	HB vs FS	HB vs FS	HB vs FB	HB vs FB	HB vs MB	HB vs FB	HB vs HB	FS vs FS	FB vs FB	MB vs MB
HGFAC	-	-	-	-	HGFAC ↑	-	-	-	-	HGFAC ↓
HGF	-	-	-	-	-	-	-	-	-	
MET	-	-	-	-	-	-	-	-	-	-
STAT3	-	-	-	-	-	-	-	-	-	-
PI3K	-	-	-	-	-	-	-	-	-	-
AKT	-	-	-	-	-	-	AKT1 ↑	AKT3 ↑	-	-
mTOR	-	-	-	-	-	-	-	-	-	-
S6	-	-	-	-	-	-	-	-	-	-
MEK	-	-	-	-	-	-	-	-	-	-
ERK	-	-	-	-	-	-	-	-	-	-

HB = horn bud; FS = frontal skin; FB = forebrain; MB = midbrain; ▮ = homozygous horned; ▮ = homozygous polled; ↑ increased expression in horn bud or horned tissue for other comparisons; ↓ decreased expression in horn bud or horned tissue for other comparisons; - not differentially expressed

Table C29: Axon growth regulatory genes differentially expressed in the horn bud of horned fetuses (n = 4) compared to horn bud region of polled fetuses (n = 3).

Family	DE in horn bud of horned fetuses	Genes
SEMA/PLXN/NRP	↑	SEMA6A, SEMA3C, NRP1, PLXNB2
	↓	SEMA4A, SEMA7A, SEMA6B
EFN/EPH	↑	EFNA5, EFNB2, EPHA3, EPHA4, EPHA7
	↓	EFNA3, EPHA8, EPHA10
CDH	↑	CDH11, PCDH7, PCDH18
	↓	CDH4, CDH9, CDH10, CDH12, CDH15, CDH16, CDH17, CDH20, CDH22, CDH23, CDH26, CDHR1, CDHR2, CDHR3, CDHR4, CDHR5, PCDH8, PCDHA2, PCDHB1,
LAMB	↑	LAMB1
	↓	N/A
SLIT/ROBO	↑	SLITRK6, ROBO1,
	↓	SLIT1, SLITRK3, ROBO3

Table C30: RNA sample concentration and quality collected for differential expression analysis of horned and polled day 58 fetuses.

Fetus #	Genotype	Sex	Tissue	Concentration[a] (ng/ul)	260/280	RIN (pico)	RIN (nano)
546	pp	F	HB	120	2.01	-	8.2
			FS	72	1.98	-	7.5
			FB	53	1.97	-	7.5
			MB	58	2.00	-	6.4
618	pp	F	HB	33	2.00	-	7.6
			FS	4	1.96	-	5.7
			FB	3	1.97	-	2.8
			MB	4	1.89	-	3.5
736	pp	F	HB	44	1.90	-	3.9
			FS	35	1.91	-	6.8
			FB	27	1.72	-	4.8
			MB	33	1.97	-	7.5
668	pp	F	HB	63	1.83	-	3.7
			FS	23	1.86	-	3.9
			FB	24	1.78	-	5.9
			MB	14	1.79	-	2.1
709	PP	M	HB	Low	1.58	-	0
			FS	Low	1.63	-	0
			FB	0.47	1.55	5.6	-
			MB	2	1.58	3.3	-
667	PP	F	HB	31	1.73	-	4.1
			FS	28	1.86	-	1.9
			FB	18	1.87	-	2.2
			MB	16	1.68	-	2
689	PP	M	HB	Low	1.62	-	0
			FS	Low	1.55	-	4.2
			FB	Low	1.61	-	2.9
			MB	Low	1.57	-	1

HB = horn bud; FS = frontal skin; FB = forebrain skin; MB = midbrain skin, pp = homozygous horned, PP = homozygous polled.

[a] Concentration was based on the measure obtained from the Bioanalyser. (-) not measured
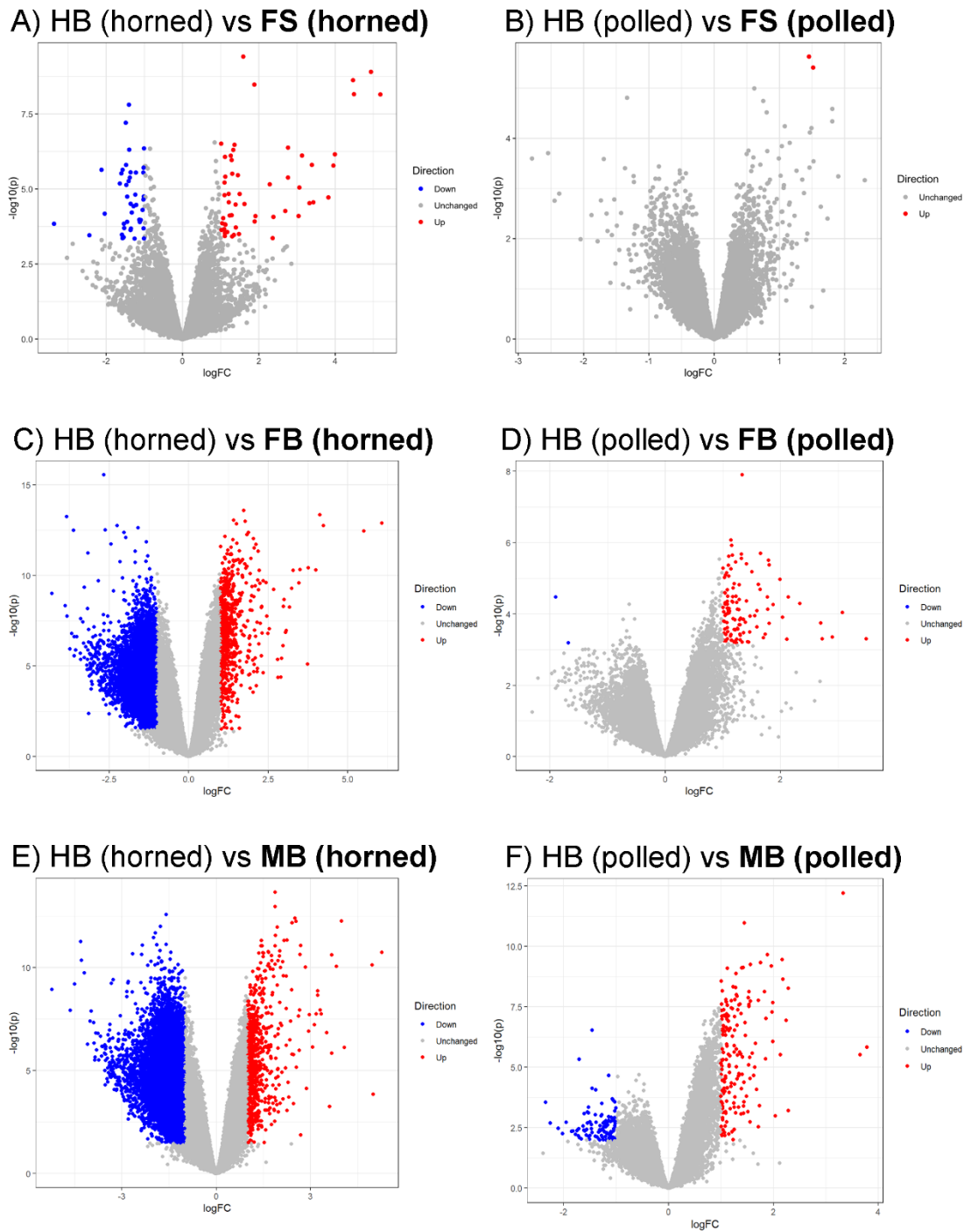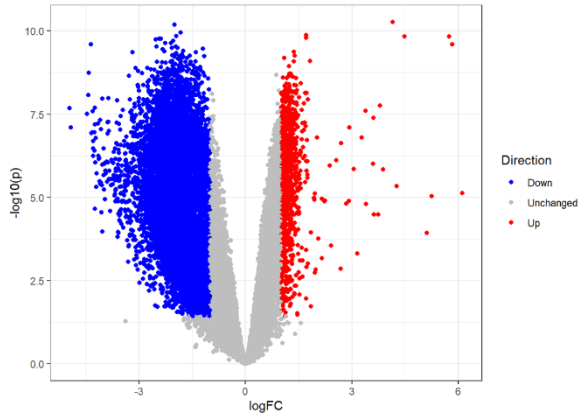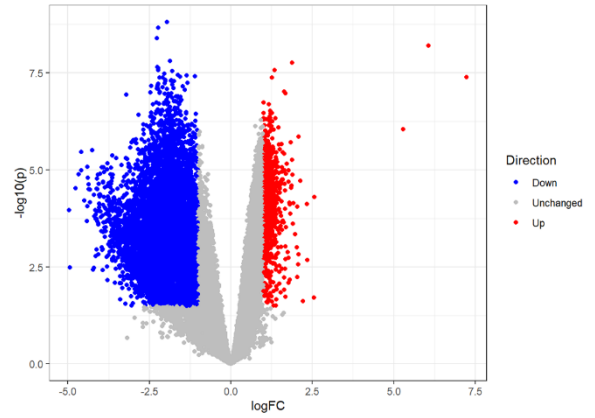
Figure C2: Volcano plots of differentially expressed genes between tissue of horned fetuses (n = 4) and polled fetus (n = 3) samples at 58 days of development. A-B) horn bud (HB) vs frontal skin (FS), C-D) horn bud vs forebrain skin (FB), and E-D) horn bud vs midbrain skin (MB). More genes are differentially expressed for the comparisons of horned tissues (A, C and E) than in polled tissues (B, D and F). Reference tissues are **boldface.**
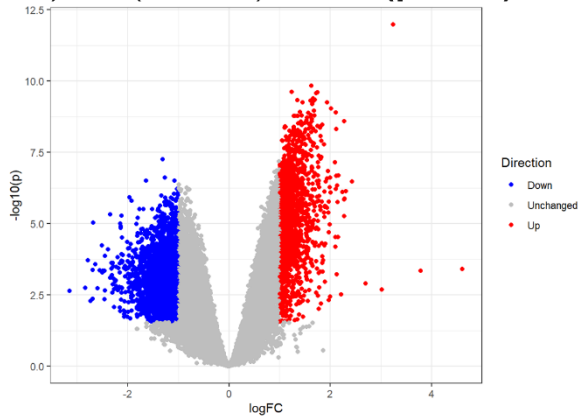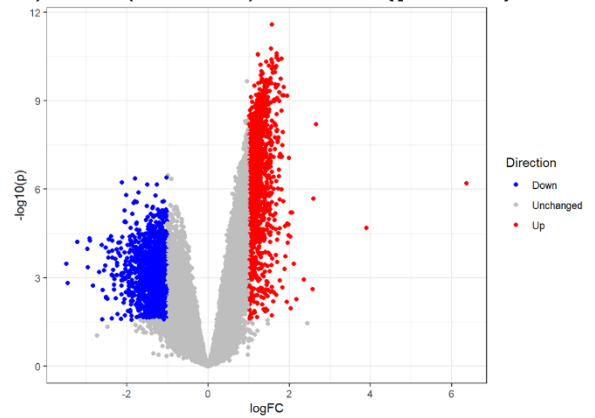
232

Figure C3: Volcano plots of differentially expressed genes between horned (n = 4) and polled fetus (n = 3) samples at 58 days of development. Reference tissues are **boldface.** HB = horn bud, FS = frontal skin, FB = forebrain skin, MB = midbrain skin.
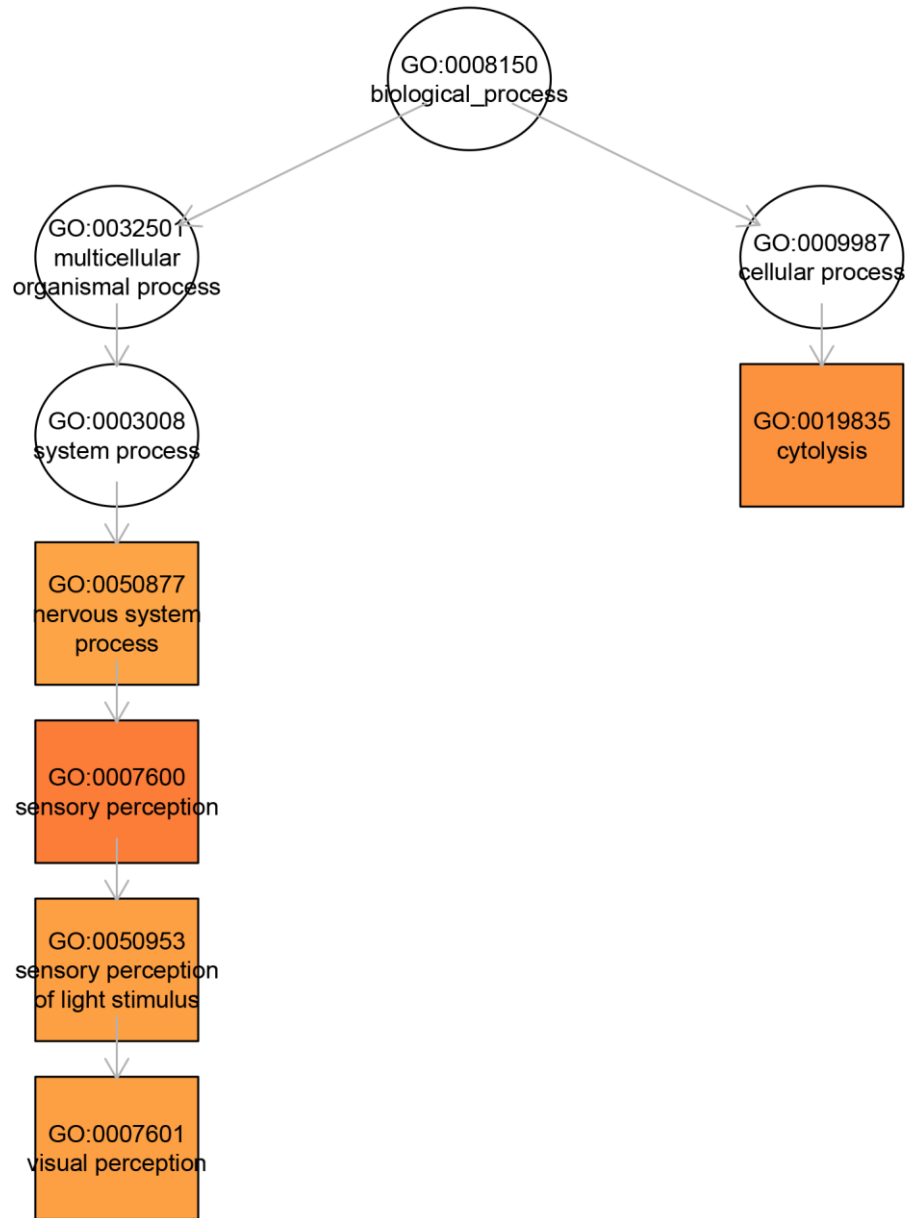
Figure C4: Enriched GO pathways for horn bud versus forebrain skin comparison in horned fetuses at 58 days of development (n = 4). High resolution image available at https://figshare.com/: <DOI: 10.25909/19341869>.

Figure C5: Enriched GO pathways for horn bud versus midbrain skin comparison in horned fetuses at 58 days of development (n = 4). High resolution image available at https://figshare.com/: <DOI: 10.25909/19341869>.

Figure C6: Enriched GO pathways for horn bud versus forebrain skin comparison in polled fetuses (n = 3) at 58 days of development. High resolution image available at https://figshare.com/: <DOI: 10.25909/19341869>.
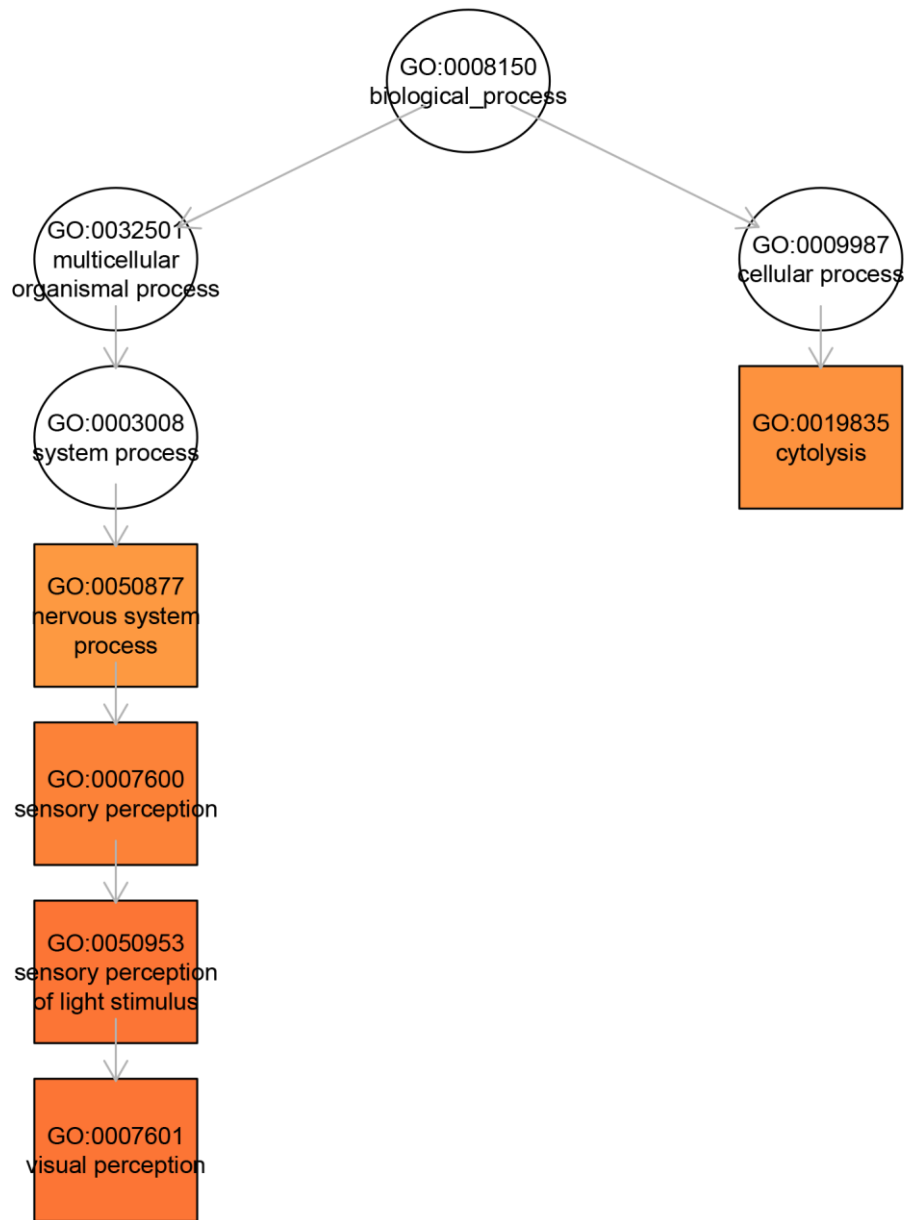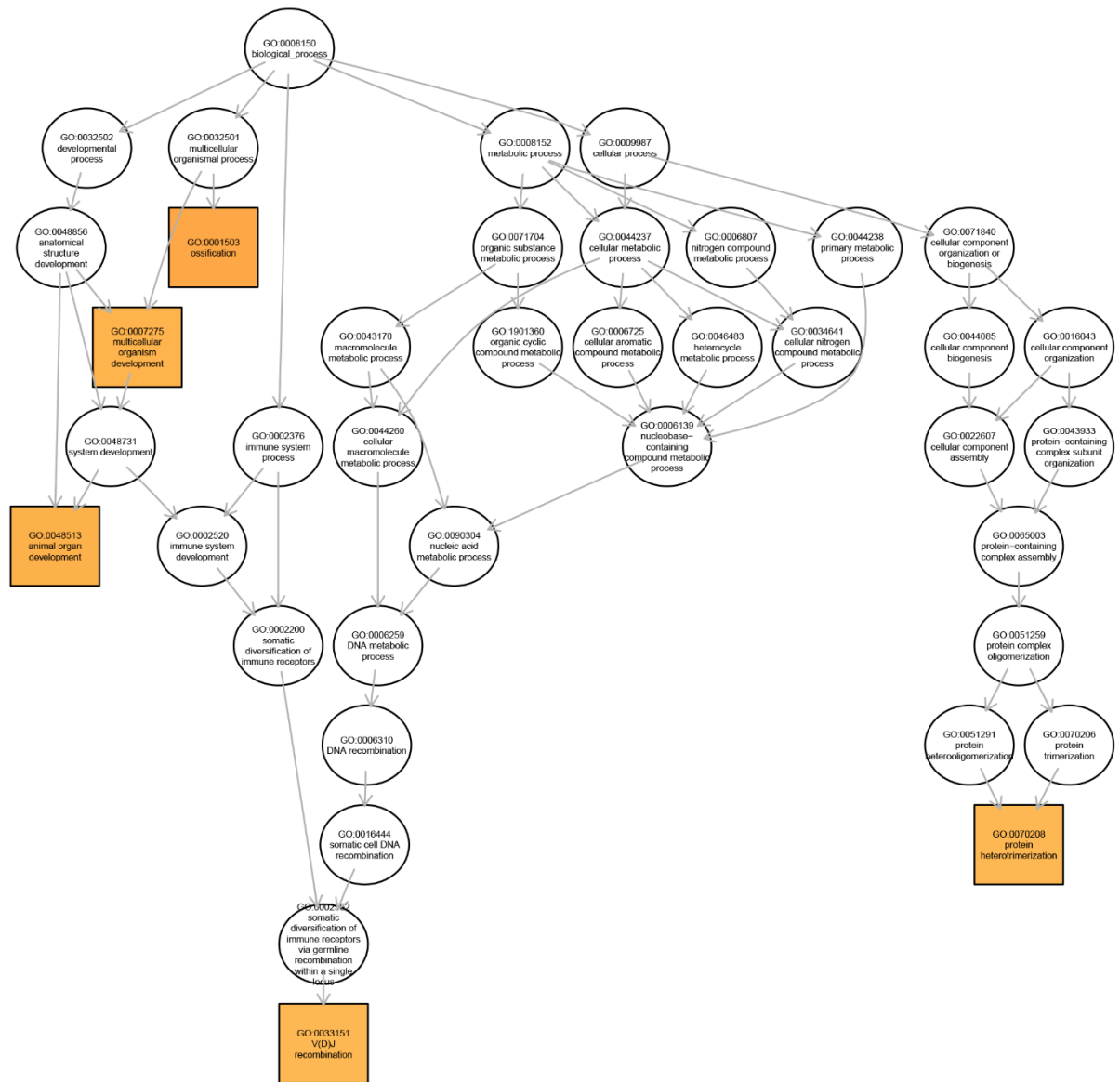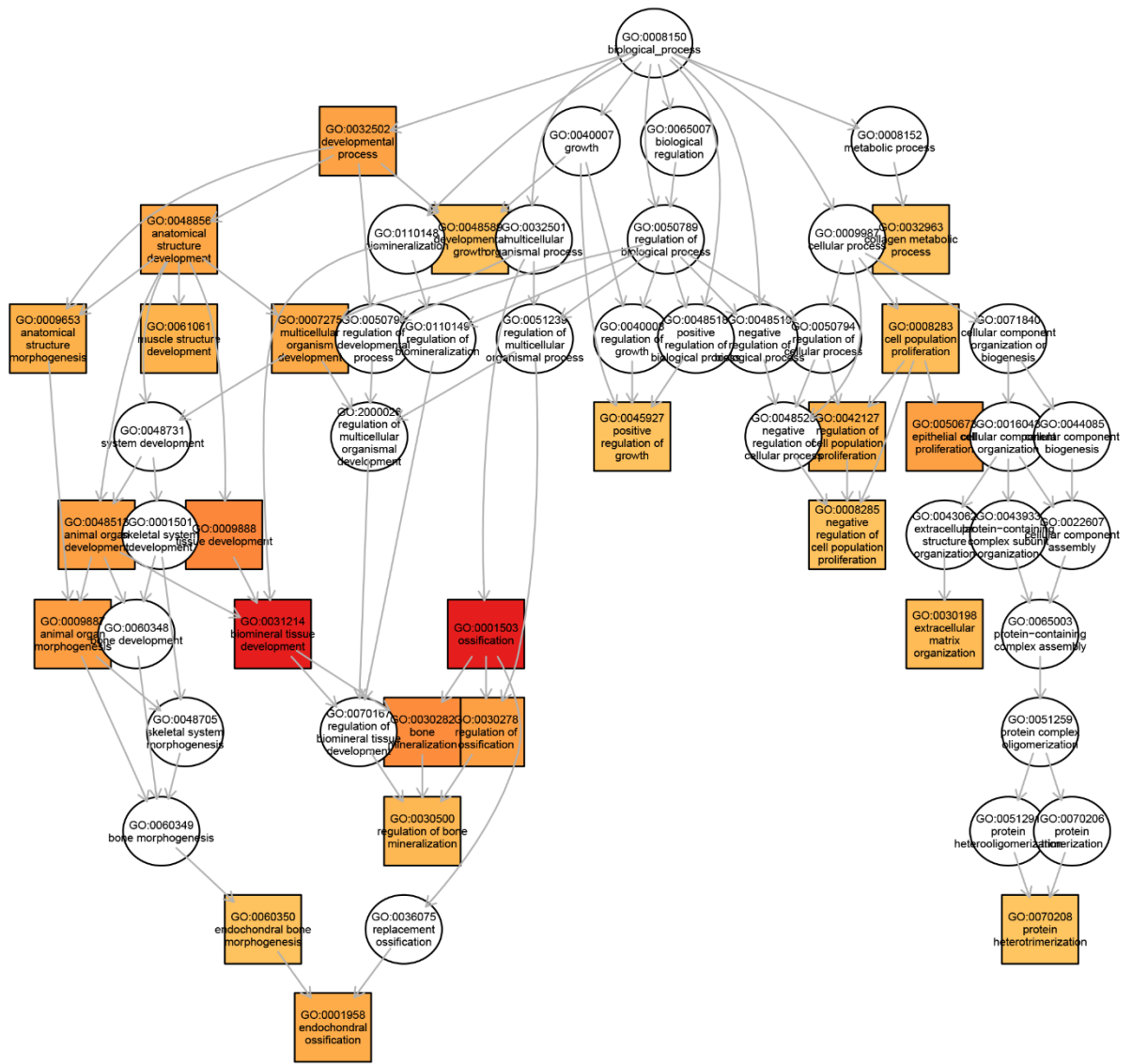
Figure C7: Enriched GO pathways for horn bud versus midbrain skin comparison in polled fetuses (n = 3) at 58 days of development. High resolution image available at https://figshare.com/ <DOI: 10.25909/19341869>.

Figure C8: Enriched GO pathways for horn bud (horned; n = 4) versus horn bud region (polled; n = 3) differentially expressed genes at 58 days of development. High resolution image available at https://figshare.com/: <DOI: 10.25909/19341869>.
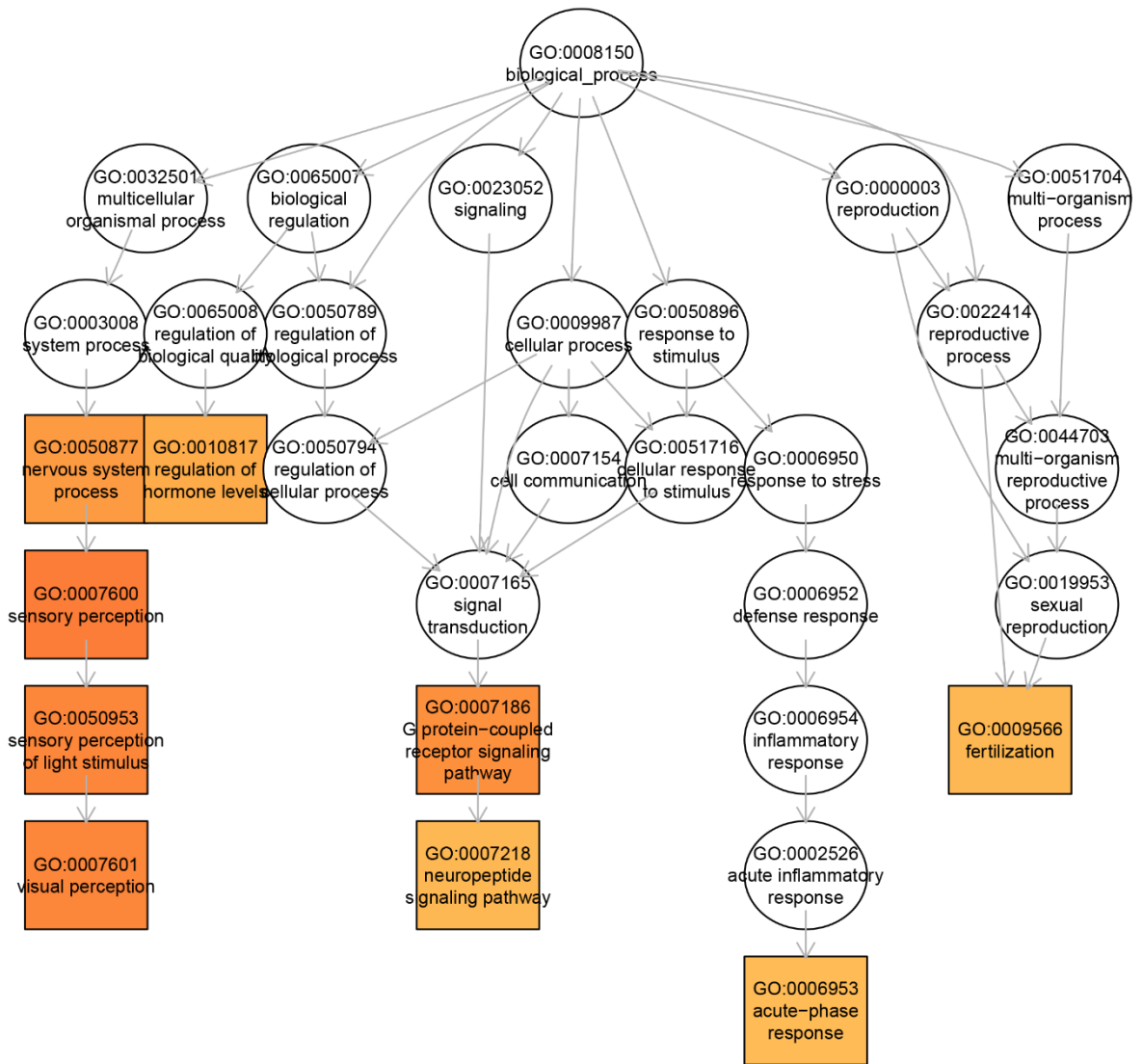
Figure C9: Enriched GO pathways for frontal skin (horned; n = 4) versus frontal skin (polled; n = 3) differentially expressed genes at 58 days of development. High resolution image available at https://figshare.com/: <DOI: 10.25909/19341869>.

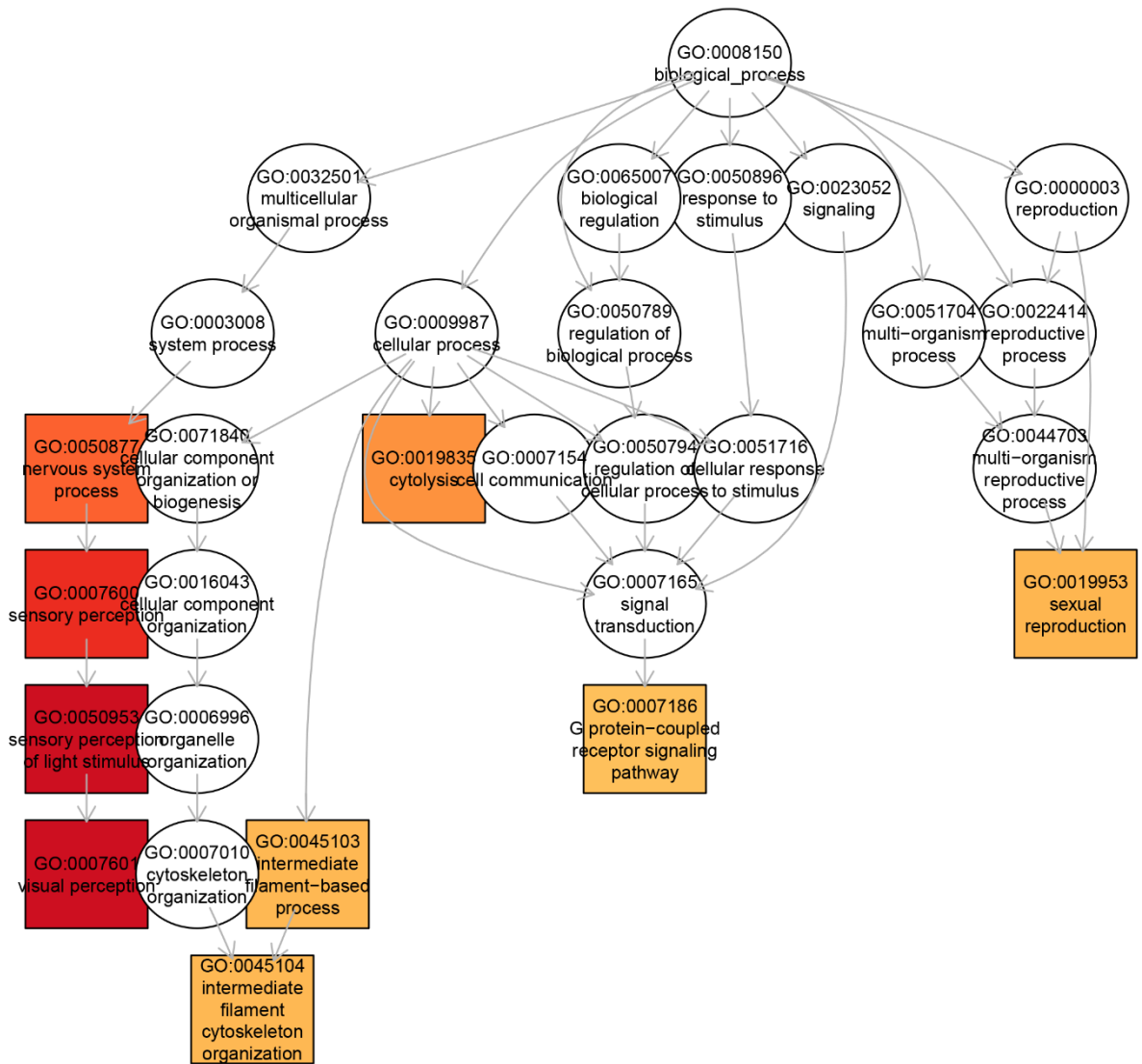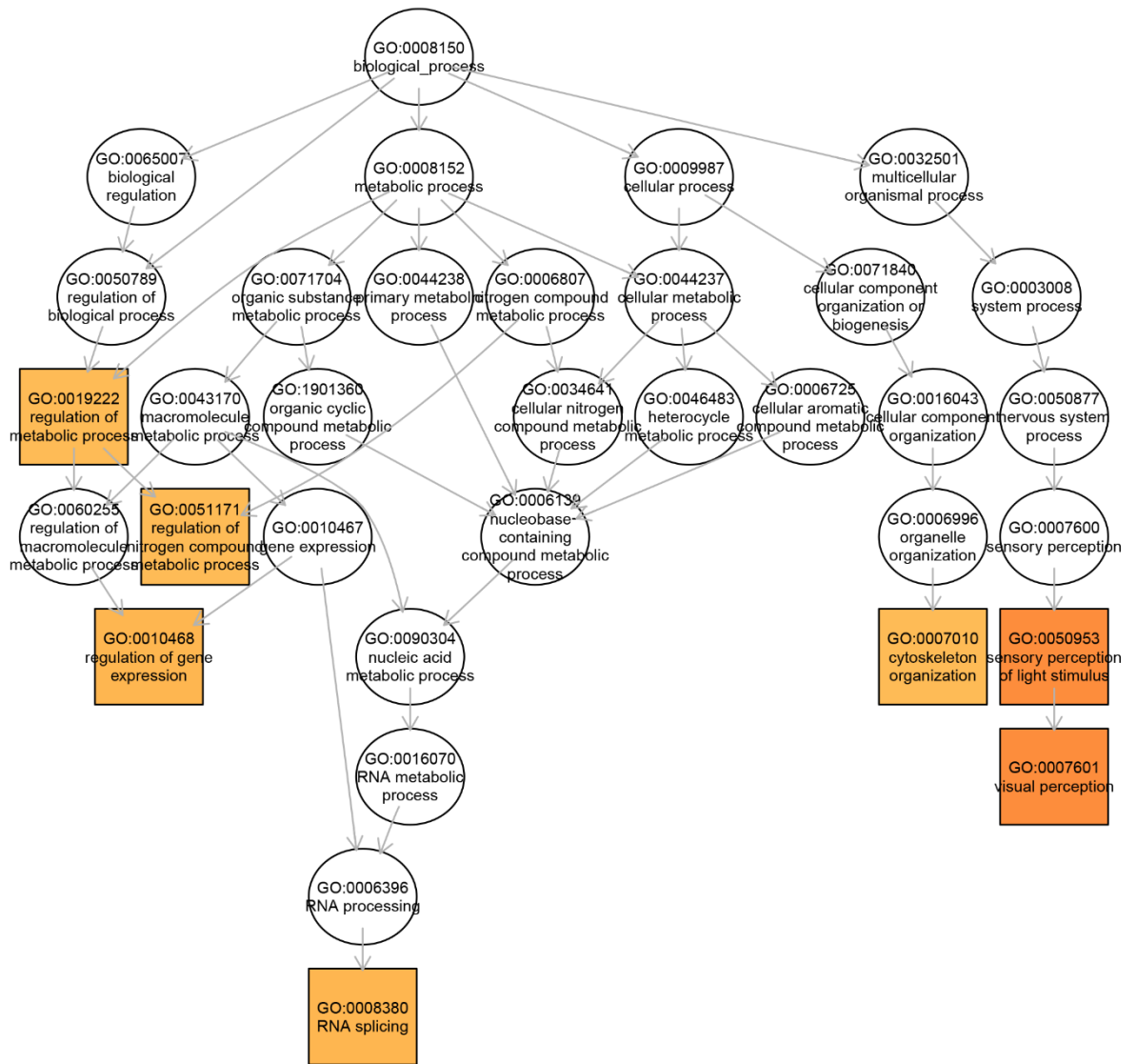Figure C10: Enriched GO pathways for midbrain skin (horned; n = 4) versus midbrain skin (polled; n = 3) differentially expressed genes at 58 days of development. High resolution image available at https://figshare.com/: <DOI: 10.25909/19341869>.
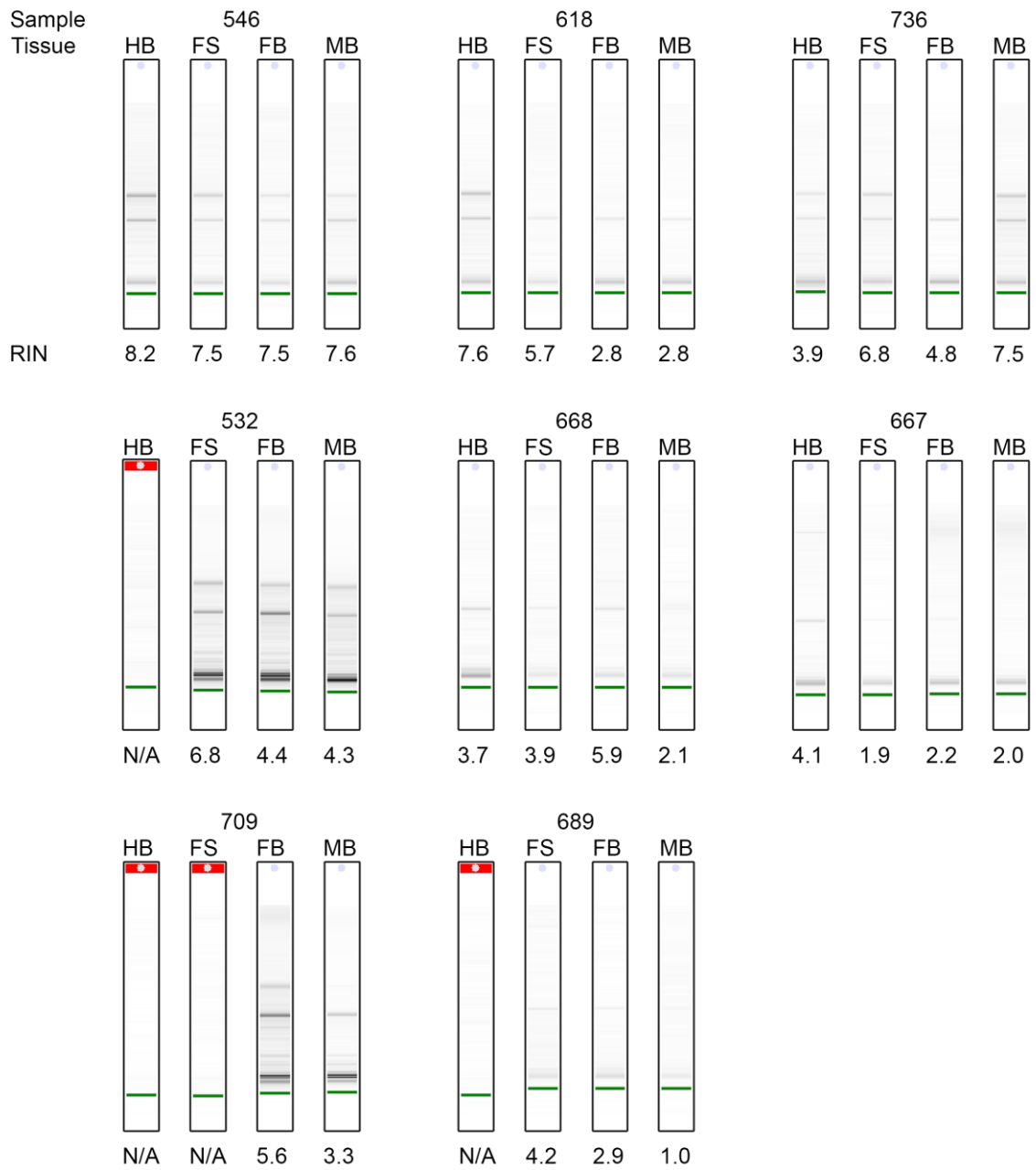
Figure C11: RNA integrity numbers for RNA samples extracted for RNAseq.
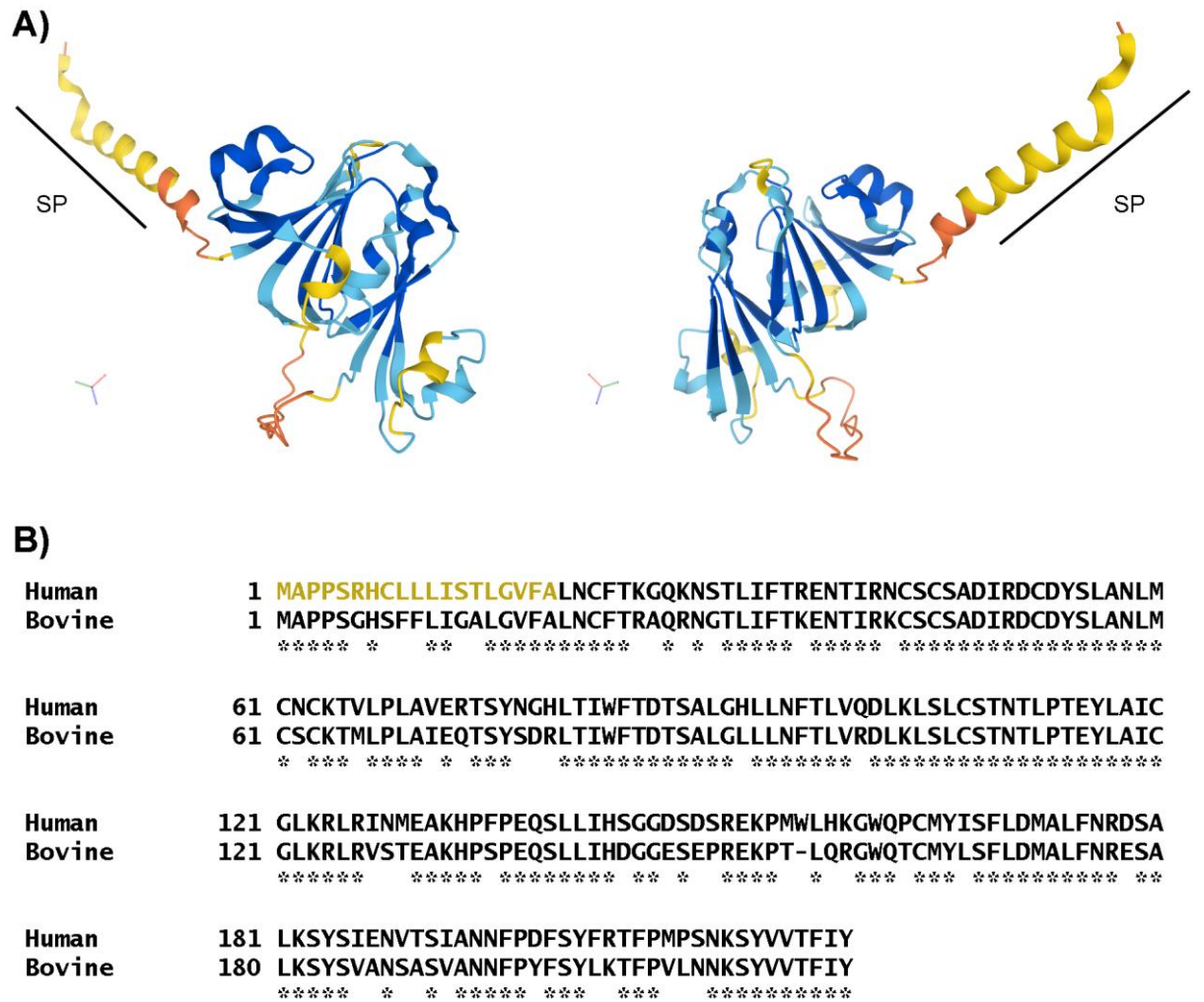
Figure C12: Structure and amino acid sequence of C21orf62. A) Predicted structure of C21orf62 (ortholog of bovine C1H21orf62) using Alphafold (Jumper *et al.* 2021; Varadi *et al.* 2021). Per-residue confidence score (pLDDT) between 0 and 100 indicates confidence of structure prediction. B) Amino acid sequence of C21orf62 (human) and C1H21orf62 (bovine). The signal peptide (SP) is highlighted as the first 19 amino acids. Model Confidence: ■ Very high (pLDDT > 90) ■ Confident (90 > pLDDT > 70) ■ Low (70 > pLDDT > 50) ■ Very low (pLDDT < 50).

References

Abbate F., Germana G.P., De Carlos F., Montalbano G., Laura R., Levanti M.B. & Germana A. (2006) The oral cavity of the adult zebrafish (Danio rerio). *Anatomia, Histologia, Embryologia: Journal of Veterinary Medicine Series C* **35**, 299-304.

Agarwala R., Barrett T., Beck J., Benson D.A., Bollin C., Bolton E., Bourexis D., Brister J.R., Bryant S.H., Canese K., Cavanaugh M., Charowhas C., Clark K., Dondoshansky I., Feolo M., Fitzpatrick L., Funk K., Geer L.Y., Gorelenkov V., Graeff A., Hlavina W., Holmes B., Johnson M., Kattman B., Khotomlianski V., Kimchi A., Kimelman M., Kimura M., Kitts P., Klimke W., Kotliarov A., Krasnov S., Kuznetsov A., Landrum M.J., Landsman D., Lathrop S., Lee J.M., Leubsdorf C., Lu Z., Madden T.L., Marchler-Bauer A., Malheiro A., Meric P., Karsch-Mizrachi I., Mnev A., Murphy T., Orris R., Ostell J., O'Sullivan C., Palanigobu V., Panchenko A.R., Phan L., Pierov B., Pruitt K.D., Rodarmer K., Sayers E.W., Schneider V., Schoch C.L., Schuler G.D., Sherry S.T., Siyan K., Soboleva A., Soussov V., Starchenko G., Tatusova T.A., Thibaud-Nissen F., Todorov K., Trawick B.W., Vakatov D., Ward M., Yaschenko E., Zasypkin A. & Zbicz K. (2018) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research* **46**, D8-D13.

Aldersey J.E., Sonstegard T.S., Williams J.L. & Bottema C.D.K. (2020) Understanding the effects of the bovine POLLED variants. *Animal Genetics* **51**, 166-76.

Alfaro M.P., Vincent A., Saraswati S., Thorne C.A., Hong C.C., Lee E. & Young P.P. (2010) sFRP2 Suppression of Bone Morphogenic Protein (BMP) and Wnt Signaling Mediates Mesenchymal Stem Cell (MSC) Self-renewal Promoting Engraftment and Myocardial Repair. *Journal of Biological Chemistry* **285**, 35645-53.

Allais-Bonnet A., Grohs C., Medugorac I., Krebs S., Djari A., Graf A., Fritz S., Seichter D., Baur A., Russ I., Bouet S., Rothammer S., Wahlberg P., Esquerre D., Hoze C., Boussaha

M., Weiss B., Thepot D., Fouilloux M.-N., Rossignol M.-N., van Marle-Koster E., Hreidarsdottir G.E., Barbey S., Dozias D., Cobo E., Reverse P., Catros O., Marchand J.-L., Soulas P., Roy P., Marquant-Leguienne B., Le Bourhis D., Clement L., Salas-Cortes L., Venot E., Pannetier M., Phocas F., Klopp C., Rocha D., Fouchet M., Journaux L., Bernard-Capel C., Ponsart C., Eggen A., Blum H., Gallard Y., Boichard D., Pailhoux E. & Capitan A. (2013) Novel insights into the bovine polled phenotype and horn ontogenesis in bovidae. *PLOS One* **8**, e63512.

Andersson R., Gebhard C., Miguel-Escalada I., Hoof I., Bornholdt J., Boyd M., Chen Y., Zhao X., Schmidl C., Suzuki T., Ntini E., Arner E., Valen E., Li K., Schwarzfischer L., Glatz D., Raithel J., Lilje B., Rapin N., Bagger F.O., Jørgensen M., Andersen P.R., Bertin N., Rackham O., Burroughs A.M., Baillie J.K., Ishizu Y., Shimizu Y., Furuhata E., Maeda S., Negishi Y., Mungall C.J., Meehan T.F., Lassmann T., Itoh M., Kawaji H., Kondo N., Kawai J., Lennartsson A., Daub C.O., Heutink P., Hume D.A., Jensen T.H., Suzuki H., Hayashizaki Y., Müller F., Forrest A.R.R., Carninci P., Rehli M. & Sandelin A. (2014) An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455-61.

Andrews S. (2010) FASTQC. A quality control tool for high throughput sequence data.

Andrews S.J. & Rothnagel J.A. (2014) Emerging evidence for functional peptides encoded by short open reading frames. *Nature Reviews Genetics* **15**, 193-204.

Archer H.L., Evans J.C., Millar D.S., Thompson P.W., Kerr A.M., Leonard H., Christodoulou J., Ravine D., Lazarou L., Grove L., Verity C., Whatley S.D., Pilz D.T., Sampson J.R. & Clarke A.J. (2006) NTNG1mutations are a rare cause of Rett syndrome. *American Journal of Medical Genetics Part A* **140A**, 691-4.

Arévalo J.C. & Wu S.H. (2006) Neurotrophin signaling: many exciting surprises! *Cellular and Molecular Life Sciences* **63**, 1523-37.

Bab I., Gavish H., Namdar-Attar M., Greenberg Z., Chen Y., Mansur N., Muhlrad A., Shteyer A. & Chorev M. (1999a) Isolation of mitogenically active C-terminal truncated pentapeptide of osteogenic growth peptide from human plasma and culture medium of murine osteoblastic cells. *The Journal of Peptide Research* **54**, 408-14.

Bab I., Smith E., Gavish H., Attar-Namdar M., Chorev M., Chen Y.-C., Muhlrad A., Birnbaum M.J., Stein G. & Frenkel B. (1999b) Biosynthesis of osteogenic growth peptide via alternative translational initiation at AUG85 of histone H4 mRNA. *Journal of Biological Chemistry* **274**, 14474-81.

Bailey T.L., Boden M., Buske F.A., Frith M., Grant C.E., Clementi L., Ren J., Li W.W. & Noble W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* **37**, W202-W8.

Bancroft J.D. & Layton C. (2019) The hematoxylins and eosin. In: *Bancroft's theory and practice of histological techniques* (pp. 126-38. Elsevier.

Barbosa A.C., Funato N., Chapman S., Mckee M.D., Richardson J.A., Olson E.N. & Yanagisawa H. (2007) Hand transcription factors cooperatively regulate development of the distal midline mesenchyme. *Developmental Biology* **310**, 154-68.

Barnes R.M., Firulli B.A., Conway S.J., Vincentz J.W. & Firulli A.B. (2010) Analysis of the Hand1 cell lineage reveals novel contributions to cardiovascular, neural crest, extra-embryonic, and lateral mesoderm derivatives. *Developmental Dynamics* **239**, 3086-97.

Barravecchia I. & Demontis G.C. (2021) HCN1 channels: A versatile tool for signal processing by primary sensory neurons. *Progress in Biophysics and Molecular Biology* **166**, 133-46.

Benhaj K., Akcali K. & Ozturk M. (2006) Redundant expression of canonical Wnt ligands in human breast cancer cell lines. *Oncology Reports*.

Bennett J.H., Hunt P. & Thorogood P. (1995) Bone morphogenetic protein-2 and -4 expression during murine orofacial development. *Archives of Oral Biology* **40**, 847-54.

Betancur P., Bronner-Fraser M. & Sauka-Spengler T. (2010) Assembling neural crest regulatory circuits into a gene regulatory network. *Annual Review of Cell and Developmental Biology* **26**, 581-603.

Bonilla-Claudio M., Wang J., Bai Y., Klysik E., Selever J. & Martin J.F. (2012) Bmp signaling regulates a dose-dependent transcriptional program to control facial skeletal development. *Development* **139**, 709-19.

Bourdon C., Bardou P., Aujean E., Le Guillou S., Tosser-Klopp G. & Le Provost F. (2019) RumimiR: a detailed microRNA database focused on ruminant species. *Database* **2019**.

Brinkmann V., Foroutan H., Sachs M., Weidner K.M. & Birchmeier W. (1995) Hepatocyte growth factor/scatter factor induces a variety of tissue-specific morphogenic programs in epithelial cells. *Journal of Cell Biology* **131**, 1573-86.

Britsch S., Goerich D.E., Riethmacher D., Peirano R.I., Rossner M., Nave K.-A., Birchmeier C. & Wegner M. (2001) The transcription factor Sox10 is a key regulator of peripheral glial development. *Genes & Development* **15**, 66-78.

Brown T.A. (2002) Understanding a genome sequence. In: *Genomes* (ed. by Brown TA). Wiley-Liss, Oxford.

Buda S., Bragulla H. & Budras K.-D. (2011) Central nervous system and cranial nerves. In: *Bovine anatomy* (ed. by Budras K-D), pp. 50-5. Schlutersche, Hannover, Germany.

Bush S.J., Muriuki C., McCulloch M.E.B., Farquhar I.L., Clark E.L. & Hume D.A. (2018) Cross-species inference of long non-coding RNAs greatly expands the ruminant transcriptome. *Genetics Selection Evolution* **50**, 1-17.

Capitan A., Allais-Bonnet A., Pinton A., Marquant-Le Guienne B., Le Bourhis D., Grohs C., Bouet S., Clement L., Salas-Cortes L., Venot E., Chaffaux S., Weiss B., Delpeuch A.,

Noe G., Rossignol M.N., Barbey S., Dozias D., Cobo E., Barasc H., Auguste A., Pannetier M., Deloche M.C., Lhuilier E., Bouchez O., Esquerre D., Salin G., Klopp C., Donnadieu C., Chantry-Darmon C., Hayes H., Gallard Y., Ponsart C., Boichard D. & Pailhoux E. (2012) A 3.7 Mb deletion encompassing ZEB2 causes a novel polled and multisystemic syndrome in the progeny of a somatic mosaic bull. *PLOS One* **7**, e49084.

Capitan A., Grohs C., Gautier M. & Eggen A. (2009) The scurs inheritance: new insights from the French Charolais breed. *BMC Genetics* **10**.

Capitan A., Grohs C., Weiss B., Rossignol M.N., Reverse P. & Eggen A. (2011) Newly described bovine type 2 scurs syndrome segregates with a frame-shift mutation in TWIST1. *PLOS One* **6**, e22242.

Carlson D.F., Lancto C.A., Zang B., Kim E.-S., Walton M., Oldeschulte D., Seabury C., Sonstegard T.S. & Fahrenkrug S.C. (2016) Production of hornless dairy cattle from genome-edited cell lines. *Nature Biotechnology* **34**, 479-81.

Cervantes-Barragan D.E., Villarroel C.E., Medrano-Hernandez A., Duran-McKinster C., Bosch-Canto V., Del-Castillo V., Nazarenko I., Yang A. & Desnick R.J. (2011) Setleis syndrome in Mexican-Nahua sibs due to a homozygous TWIST2 frameshift mutation and partial expression in heterozygotes: review of the focal facial dermal dysplasias and subtype reclassification. *Journal of Medical Genetics* **48**, 716-20.

Chen F., Li Z. & Zhou H. (2019) Identification of prognostic miRNA biomarkers for predicting overall survival of colon adenocarcinoma and bioinformatics analysis: A study based on The Cancer Genome Atlas database. *Journal of Cellular Biochemistry* **120**, 9839-49.

Chen M., Li Y., Liu Y., Xu H. & Bi L.-L. (2021) Neuregulin-1-dependent control of amygdala microcircuits is critical for fear extinction. *Neuropharmacology* **201**, 108842.

Christ B., Huang R. & Scaal M. (2004) Formation and differentiation of the avian sclerotome. *Anatomy and Embryology* **208**, 333-50.

Cuevas-Diaz Duran R., Wei H., Kim D.H. & Wu J.Q. (2019) Invited review: Long non-coding RNAs: Important regulators in the development, function and disorders of the central nervous system. *Neuropathology and Applied Neurobiology* **45**, 538-56.

Cunningham F., Allen J.E., Allen J., Alvarez-Jarreta J., Irina, Austine-Orimoloye O., Andrey, Barnes I., Bennett R., Berry A., Bhai J., Bignell A., Billis K., Boddu S., Brooks L., Charkhchi M., Cummins C., Luca, Davidson C., Dodiya K., Donaldson S., Bilal, Tamara, Fatima R., Giron C.G., Genez T., Jose, Guijarro-Clarke C., Gymer A., Hardy M., Hollis Z., Hourlier T., Hunt T., Juettemann T., Kaikala V., Kay M., Lavidas I., Le T., Lemos D., Marugán J.C., Mohanan S., Mushtaq A., Naven M., Denye, Parker A., Parton A., Perry M., Piližota I., Prosovetskaia I., Manoj, Ahamed, Bianca, Schuilenburg H., Sheppard D., José, Stark W., Steed E., Sutinen K., Sukumaran R., Sumathipala D., Suner M.-M., Szpak M., Thormann A., Tricomi F.F., Urbina-Gómez D., Veidenberg A., Thomas, Walts B., Willhoft N., Winterbottom A., Wass E., Chakiachvili M., Flint B., Frankish A., Giorgetti S., Haggerty L., Sarah, Garth, Jane, Fergal, Moore B., Jonathan, Muffato M., Perry E., Ruffier M., Tate J., Thybert D., Stephen, Dyer S., Peter, Kevin, Andrew, Daniel & Flicek P. (2022) Ensembl 2022. *Nucleic Acids Research* **50**, D988-D95.

Darbellay F., Bochaton C., Lopez-Delisle L., Mascrez B., Tschopp P., Delpretti S., Zakany J. & Duboule D. (2019) The constrained architecture of mammalian Hox gene clusters. *Proceedings of the National Academy of Sciences* **116**, 13424-33.

Dathe K., Kjaer K.W., Brehm A., Meinecke P., Nürnberg P., Neto J.C., Brunoni D., Tommerup N., Ott C.E., Klopocki E., Seemann P. & Mundlos S. (2009) Duplications involving a conserved regulatory element downstream of BMP2 are associated with brachydactyly type A2. *The American Journal of Human Genetics* **84**, 483-92.

Dávila López M., Martínez Guerra J.J. & Samuelsson T. (2010) Analysis of gene order conservation in eukaryotes identifies transcriptionally and functionally linked genes. *PLOS One* **5**, e10654.

De Vienne D.M. (2016) Lifemap: exploring the entire tree of life. *PLOS Biology* **14**, e2001624.

Delgado I., Giovinazzo G., Temiño S., Gauthier Y., Balsalobre A., Drouin J. & Torres M. (2021) Control of mouse limb initiation and antero-posterior patterning by Meis transcription factors. *Nature Communications* **12**.

Divisato G., Passaro F., Russo T. & Parisi S. (2020) The key role of microRNAs in self-renewal and differentiation of embryonic stem cells. *International Journal of Molecular Sciences* **21**, 6285.

Dixon J.R., Gorkin D.U. & Ren B. (2016) Chromatin Domains: The Unit of Chromosome Organization. *Molecular Cell* **62**, 668-80.

Dixon J.R., Selvaraj S., Yue F., Kim A., Li Y., Shen Y., Hu M., Liu J.S. & Ren B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-80.

Donizetti A., Fiengo M., Del Gaudio R., Iazzetti G., Pariante P., Minucci S. & Aniello F. (2015) Expression pattern of zebrafish rxfp2 homologue genes during embryonic development. *Journal of Experimental Zoology* **324**, 605-13.

Duarte C., Kobayashi Y., Kawamoto T. & Moriyama K. (2014) Relaxin receptors 1 and 2 and nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor) mRNAs are expressed in oral components of developing mice. *Archives of Oral Biology* **59**, 111-8.

Duijvesteijn N., Bolormaa S., Daetwyler H.D. & Van Der Werf J.H.J. (2018) Genomic prediction of the polled and horned phenotypes in Merino sheep. *Genetics Selection Evolution* **50**, 1-11.

Edgar R., Mazor Y., Rinon A., Blumenthal J., Golan Y., Buzhor E., Livnat I., Ben-Ari S., Lieder I., Shitrit A., Gilboa Y., Ben-Yehudah A., Edri O., Shraga N., Bogoch Y., Leshansky L., Aharoni S., West M.D., Warshawsky D. & Shtrichman R. (2013) LifeMap Discovery™: The embryonic development, stem cells, and regenerative medicine research portal. *PLOS One* **8**, e66629.

Eliveld J., Daalen S.K.M., Winter-Korver C.M., Veen F., Repping S., Teerds K. & Pelt A.M.M. (2020) A comparative analysis of human adult testicular cells expressing stem Leydig cell markers in the interstitium, vasculature, and peritubular layer. *Andrology* **8**, 1265-76.

Engstrom P.G., Ho Sui S.J., Drivenes O., Becker T.S. & Lenhard B. (2007) Genomic regulatory blocks underlie extensive microsynteny conservation in insects. *Genome Research* **17**, 1898-908.

Esteban-Lopez M. & Agoulnik A.I. (2020) Diverse functions of insulin-like 3 peptide. *Journal of Endocrinology* **247**, R1-R12.

Evans E.H. & Sack W.O. (1973) Prenatal development of domestic and laboratory mammals: growth curves, external features and selected references. *Anatomia, Hostologia, Embryologia* **2**, 11-45.

Falomir-Lockhart A.H., Ortega Masague M.F., Rudd Garces G., Zappa M.E., Peral García P., Morales H.F., Holgado F.D., Rogberg Muñoz A. & Giovambattista G. (2019) Polledness in Argentinean Creole cattle, five centuries surviving. *Animal Genetics* **50**, 381-5.

FANTOM Consortium, RIKEN PMI & RIKEN CLST (DGT) (2014) A promoter-level mammalian expression atlas. *Nature* **507**, 462-70.

Farré M., Kim J., Proskuryakova A.A., Zhang Y., Kulemzina A.I., Li Q., Zhou Y., Xiong Y., Johnson J.L., Perelman P.L., Johnson W.E., Warren W.C., Kukekova A.V., Zhang G.,

O'Brien S.J., Ryder O.A., Graphodatsky A.S., Ma J., Lewin H.A. & Larkin D.M. (2019) Evolution of gene regulation in ruminants differs between evolutionary breakpoint regions and homologous synteny blocks. *Genome Research* **29**, 576-89.

Feng S., Bogatcheva N.V., Truong A., Korchin B., Bishop C.E., Klonisch T., Agoulnik I.U. & Agoulnik A.I. (2007) Developmental expression and gene regulation of insulin-like 3 receptor RXFP2 in mouse male reproductive organs. *Biology of Reproduction* **77**, 671-80.

Feng S., Ferlin A., Truong A., Bathgate R., Wade J.D., Corbett S., Han S., Tannour-Louet M., Lamb D.J., Foresta C. & Agoulnik A.I. (2009) INSL3/RXFP2 signaling in testicular descent. *Annals of the New York Academy of Sciences* **1160**, 197-204.

Ferlin A., De Toni L., Agoulnik A.I., Lunardon G., Armani A., Bortolanza S., Blaauw B., Sandri M. & Foresta C. (2018) Protective role of testicular hormone INSL3 from atrophy and weakness in skeletal muscle. *Frontiers in Endocrinology* **9**.

Ferlin A., De Toni L., Sandri M. & Foresta C. (2017) Relaxin and insulin-like peptide 3 in the musculoskeletal system: from bench to bedside. *British Journal of Pharmacology* **174**, 1015-24.

Ferlin A., Perilli L., Gianesello L., Taglialavoro G. & Foresta C. (2011) Profiling insulin like factor 3 (INSL3) signaling in human osteoblasts. *PLOS One* **6**, e29733.

Filippova D., Patro R., Duggal G. & Kingsford C. (2014) Identification of alternative topological domains in chromatin. *Algorithms for Molecular Biology* **9**, 1-11.

Finzsch M., Schreiner S., Kichko T., Reeh P., Tamm E.R., Bösl M.R., Meijer D. & Wegner M. (2010) Sox10 is required for Schwann cell identity and progression beyond the immature Schwann cell stage. *Journal of Cell Biology* **189**, 701-12.

Fiore R., Siegel G. & Schratt G. (2008) MicroRNA function in neuronal development, plasticity and disease. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1779**, 471-8.

Firulli B.A., Fuchs R.K., Vincentz J.W., Clouthier D.E. & Firulli A.B. (2014) Hand1 phosphoregulation within the distal arch neural crest is essential for craniofacial morphogenesis. *Development* **141**, 3050-61.

Firulli B.A., Milliar H., Toolan K.P., Harkin J., Fuchs R.K., Robling A.G. & Firulli A.B. (2017) Defective Hand1 phosphoregulation uncovers essential roles for Hand1 in limb morphogenesis. *Development* **144**, 2480-9.

Flöttmann R., Kragesteen B.K., Geuer S., Socha M., Allou L., Sowińska-Seidler A., Bosquillon De Jarcy L., Wagner J., Jamsheer A., Oehl-Jaschkowitz B., Wittler L., De Silva D., Kurth I., Maya I., Santos-Simarro F., Hülsemann W., Klopocki E., Mountford R., Fryer A., Borck G., Horn D., Lapunzina P., Wilson M., Mascrez B., Duboule D., Mundlos S. & Spielmann M. (2018) Noncoding copy-number variations are associated with congenital limb malformation. *Genetics in Medicine* **20**, 599-607.

Foissac S., Djebali S., Munyard K., Vialaneix N., Rau A., Muret K., Esquerré D., Zytnicki M., Derrien T., Bardou P., Blanc F., Cabau C., Crisci E., Dhorne-Pollet S., Drouet F., Faraut T., Gonzalez I., Goubil A., Lacroix-Lamandé S., Laurent F., Marthey S., Marti-Marimon M., Momal-Leisenring R., Mompart F., Quéré P., Robelin D., Cristobal M.S., Tosser-Klopp G., Vincent-Naulleau S., Fabre S., Der Laan M.-H.P.-V., Klopp C., Tixier-Boichard M., Acloque H., Lagarrigue S. & Giuffra E. (2019) Multi-species annotation of transcriptome and chromatin structure in domesticated animals. *BMC Biology* **17**, 1-25.

Fornes O., Castro-Mondragon J.A., Khan A., Robin, Zhang X., Richmond P.A., Modi B.P., Correard S., Gheorghe M., Baranašić D., Santana-Garcia W., Tan G., Chèneby J., Ballester B., Parcy F., Sandelin A., Lenhard B., Wasserman W.W. & Mathelier A.

(2019) JASPAR 2020: Update of the open-access database of transcription factor binding profiles. *Nucleic Acids Research* **48**, D87–D92.

Fox J.W., Lamperti E.D., Ekşioğlu Y.Z., Hong S.E., Feng Y., Graham D.A., Scheffer I.E., Dobyns W.B., Hirsch B.A., Radtke R.A., Berkovic S.F., Huttenlocher P.R. & Walsh C.A. (1998) Mutations in filamin 1 prevent migration of cerebral cortical neurons in human periventricular heterotopia. *Neuron* **21**, 1315-25.

Franco H.L., Casasnovas J., Rodriguez-Medina J.R. & Cadilla C.L. (2011) Redundant or separate entities? Roles of Twist1 and Twist2 as molecular switches during gene transcription. *Nucleic Acids Research* **39**, 1177-86.

Fuchs E. (2007) Scratching the surface of skin development. *Nature* **445**, 834-42.

Fudenberg G., Abdennur N., Imakaev M., Goloborodko A. & Mirny L.A. (2017) Emerging evidence of chromosome folding by loop extrusion. *Cold Spring Harbor Symposia on Quantitative Biology* **82**, 45-55.

Funato N., Taga Y., Laurie L.E., Tometsuka C., Kusubata M. & Ogawa-Goto K. (2020) The transcription factor HAND1 is involved in cortical bone mass through the regulation of collagen expression. *International Journal of Molecular Sciences* **21**, 8638.

Furlan A. & Adameyko I. (2018) Schwann cell precursor: A neural crest cell in disguise? *Developmental Biology* **444**, S25-S35.

Furlong E.E.M. & Levine M. (2018) Developmental enhancers and chromosome topology. *Science* **361**, 1341-5.

Gabarin N., Gavish H., Muhlrad A., Chen Y.-C., Namdar-Attar M., Nissenson R.A., Chorev M. & Bab I. (2001) Mitogenic Gi protein-MAP kinase signaling cascade in MC3T3-E1 osteogenic cells: Activation by C-terminal pentapeptide of osteogenic growth peptide [OGP(10-14)] and attenuation of activation by cAMP. *Journal of Cellular Biochemistry* **81**, 594-603.

Ge G. & Greenspan D.S. (2006) Developmental roles of the BMP1/TLD metalloproteinases. *Birth Defects Research Part C: Embryo Today: Reviews* **78**, 47-68.

Gehrke L.J., Capitan A., Scheper C., König S., Upadhyay M., Heidrich K., Russ I., Seichter D., Tetens J., Medugorac I. & Thaller G. (2020a) Are scurs in heterozygous polled (Pp) cattle a complex quantitative trait? *Genetics Selection Evolution* **52**, 1-13.

Gehrke L.J., Upadhyay M., Heidrich K., Kunz E., Klaus-Halla D., Weber F., Zerbe H., Seichter D., Graf A., Krebs S., Blum H., Capitan A., Thaller G. & Medugorac I. (2020b) A de novo frameshift mutation in ZEB2 causes polledness, abnormal skull shape, small body stature and subfertility in Fleckvieh cattle. *Scientific Reports* **10**, 1-14.

Godinho H.P. (1968) A comparative anatomical study of the cranial nerves in goat, sheep and bovine (Capra hircus, Ovis aries and Bos taurus): Their distribution and related autonomic components. p. 390. Iowa State University, Retrospective Theses and Dissertations.

Goncharuk S.A., Artemieva L.E., Nadezhdin K.D., Arseniev A.S. & Mineev K.S. (2020) Revising the mechanism of p75NTR activation: Intrinsically monomeric state of death domains invokes the "helper" hypothesis. *Scientific Reports* **10**, 1-15.

Gong Y., Lazaris C., Sakellaropoulos T., Lozano A., Kambadur P., Ntziachristos P., Aifantis I. & Tsirigos A. (2018) Stratification of TAD boundaries reveals preferential insulation of super-enhancers by strong boundaries. *Nature Communications* **9**, 1-12.

Graf D., Malik Z., Hayano S. & Mishina Y. (2016) Common mechanisms in development and disease: BMP signaling in craniofacial development. *Cytokine & Growth Factor Reviews* **27**, 129-39.

Grobler R., Visser C., Capitan A. & Van Marle-Köster E. (2018) Validation of the POLLED Celtic variant in South African Bonsmara and Drakensberger beef cattle breeds. *Livestock Science* **217**, 136-9.

Guo J., Jiang R., Mao A., Liu G.E., Zhan S., Li L., Zhong T., Wang L., Cao J., Chen Y., Zhang G. & Zhang H. (2021) Genome-wide association study reveals 14 new SNPs and confirms two structural variants highly associated with the horned/polled phenotype in goats. *BMC Genomics* **22**, 1-10.

Halls M.L., Bathgate R.A.D. & Summers R.J. (2006) Relaxin family peptide receptors RXFP1 and RXFP2 modulate cAMP signaling by distinct mechanisms. *Molecular Pharmacology* **70**, 214-26.

Halls M.L., Bathgate R.A.D. & Summers R.J. (2007) Comparison of signaling pathways activated by the relaxin family peptide receptors, RXFP1 and RXFP2, using reporter genes. *Journal of Pharmacology and Experimental Therapeutics* **320**, 281-90.

Hansen K. D., Gentry J., Long L., Gentleman R., Falcon S., Hahne F. & D. S. (2021) Rgraphviz: Provides plotting capabilities for R graph objects. version 2.36.0., R package version 2.36.0.

Harris R.S. (2007) Improved pairwise alignment of genomic DNA. In: *Computer Science and Engineering*. The Pennsylvania State University, Pennsylvania.

He S., Di J., Han B., Chen L., Liu M. & Li W. (2020) Genome-wide scan for runs of homozygosity identifies candidate genes related to economically important traits in Chinese merino. *Animals* **10**, 1-13.

He X., Zhou Z., Pu Y., Chen X., Ma Y. & Jiang L. (2016) Mapping the four-horned locus and testing the polled locus in three Chinese sheep breeds. *Animal Genetics* **47**, 623-7.

Heckeberg N.S. (2017) A comprehensive approach towards the phylogeny and evolution of Cervidae. In: *Fakultat fur Geowissenschaften der Ludwig–Maximilians*. Universitat München, München.

Hennig S.L., Owen J.R., Lin J.C., McNabb B.R., Van Eenennaam A.L. & Murray J.D. (2022a) A deletion at the polled PC locus alone is not sufficient to cause a polled phenotype in cattle. *Scientific Reports* **12**, 1-12.

Hennig S.L., Owen J.R., Lin J.C., McNabb B.R., Van Eenennaam A.L. & Murray J.D. (2022b) Investigating the casual mechanism behind hornless cattle. University of California – Davis, Poster presented at: Plant and Animal Genome conference.

Hermans F., Hemeryck L., Lambrichts I., Bronckaers A. & Vankelecom H. (2021) Intertwined signaling pathways governing tooth development: A give-and-take between canonical Wnt and Shh. *Frontiers in Cell and Developmental Biology* **9**, 1-22.

Hickey C.J., Kim J.-H. & Ahn E.-Y.E. (2014) New discoveries of old SON: A link between RNA splicing and cancer. *Journal of Cellular Biochemistry* **115**, 224-31.

Hinks G.L., Shah B., French S.J., Campos L.S., Staley K., Hughes J. & Sofroniew M.V. (1997) Expression of LIM protein genes Lmo1, Lmo2, and Lmo3 in adult mouse hippocampus and other forebrain regions: Differential regulation by seizure activity. *The Journal of Neuroscience* **17**, 5549-59.

Hirst C.E., Lim S.-M., Pereira L.A., Mayberry R.A., Stanley E.G. & Elefanty A.G. (2010) Expression from a betageo gene trap in the Slain1 gene locus is predominantly associated with the developing nervous system. *The International Journal of Developmental Biology* **54**, 1383-8.

Horikiri T., Ohi H., Shibata M., Ikeya M., Ueno M., Sotozono C., Kinoshita S. & Sato T. (2017) SOX10-nano-lantern reporter human iPS cells; A versatile tool for neural crest research. *PLOS One* **12**, e0170342.

Huntzinger E. & Izaurralde E. (2011) Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nature Reviews Genetics* **12**, 99-110.

Ienasescu H., Li K., Andersson R., Vitezic M., Rennie S., Chen Y., Vitting-Seerup K., Lagoni E., Boyd M., Bornholdt J., De Hoon M.J.L., Kawaji H., Lassmann T., Hayashizaki Y., Forrest A.R.R., Carninci P. & Sandelin A. (2016) On-the-fly selection of cell-specific enhancers, genes, miRNAs and proteins across the human body using SlideBase. *Database* **2016**, 1-10.

Irimia M., Tena J.J., Alexis M.S., Fernandez-Minan A., Maeso I., Bogdanovic O., De La Calle-Mustienes E., Roy S.W., Gomez-Skarmeta J.L. & Fraser H.B. (2012) Extensive conservation of ancient microsynteny across metazoans due to cis-regulatory constraints. *Genome Research* **22**, 2356-67.

Ishii M., Sun J.J., Ting M.C. & Maxson R.E. (2015) The development of the calvarial bones and sutures and the pathophysiology of craniosynostosis. In: *Craniofacial Development* (ed. by Chai Y), pp. 131-56. Elsevier Inc., Waltham, USA.

Ivell R., Alhujaili W., Kohsaka T. & Anand-Ivell R. (2020) Physiology and evolution of the INSL3/RXFP2 hormone/receptor system in higher vertebrates. *General and Comparative Endocrinology*, 1-11.

Jacob T., Chakravarty A., Panchal A., Patil M., Ghodadra G., Sudhakaran J. & Nuesslein-Volhard C. (2021) Zebrafish twist2/dermo1 regulates scale shape and scale organization during skin development and regeneration. *Cells & Development* **166**, 203684.

James G., Foster S.R., Key B. & Beverdam A. (2013) The expression pattern of EVA1C, a novel slit receptor, is consistent with an axon guidance role in the mouse nervous system. *PLOS One* **8**, e74115.

Jho E.-h., Zhang T., Domon C., Joo C.-K., Freund J.-N. & Costantini F. (2002) Wnt/β-Catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. *Molecular and Cellular Biology* **22**, 1172-83.

Johnson K.J., Robbins A.K., Wang Y., McCahan S.M., Chacko J.K. & Barthold J.S. (2010) Insulin-like 3 exposure of the fetal rat gubernaculum modulates expression of genes involved in neural pathways. *Biology of Reproduction* **83**, 774-82.

Johnston S.E., McEwan J.C., Pickering N.K., Kijas J.W., Beraldi D., Pilkington J.G., Pemberton J.M. & Slate J. (2011) Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. *Molecular Ecology* **20**, 2555-66.

Jumper J., Evans R., Pritzel A., Green T., Figurnov M., Ronneberger O., Tunyasuvunakool K., Bates R., Žídek A., Potapenko A., Bridgland A., Meyer C., Kohl S.A.A., Ballard A.J., Cowie A., Romera-Paredes B., Nikolov S., Jain R., Adler J., Back T., Petersen S., Reiman D., Clancy E., Zielinski M., Steinegger M., Pacholska M., Berghammer T., Bodenstein S., Silver D., Vinyals O., Senior A.W., Kavukcuoglu K., Kohli P. & Hassabis D. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-9.

Kalcheim C. (2015) Epithelial–mesenchymal transitions during neural crest and somite development. *Journal of Clinical Medicine* **5**, 1-15.

Kao W.W., Jester J.V., Liu C.-Y., Hayashi Y., Weng D. & Carlson E. (2007) Neural Crest Derived Periocular Mesenchymal Cell Lineages in Ocular Surface Tissues. *Investigative Ophthalmology & Visual Science* **48**, 4926.

Kardos M., Luikart G., Bunch R., Dewey S., Edwards W., McWilliam S., Stephenson J., Allendorf F.W., Hogg J.T. & Kijas J. (2015) Whole-genome resequencing uncovers molecular signatures of natural and sexual selection in wild bighorn sheep. *Molecular Ecology* **24**, 5616-32.

Kaukua N., Shahidi M.K., Konstantinidou C., Dyachuk V., Kaucka M., Furlan A., An Z., Wang L., Hultman I., Ährlund-Richter L., Blom H., Brismar H., Lopes N.A., Pachnis V., Suter

U., Clevers H., Thesleff I., Sharpe P., Ernfors P., Fried K. & Adameyko I. (2014) Glial origin of mesenchymal stem cells in a tooth model system. *Nature* **513**, 551-4.

Kaul A., Bhattacharyya S. & Ay F. (2020) Identifying statistically significant chromatin contacts from Hi-C data with FitHiC2. *Nature protocols* **15**, 991-1012.

Kawamura K., Kumagai J., Sudo S., Chun S.Y., Pisarska M., Morita H., Toppari J., Fu P., Wade J.D., Bathgate R.A.D. & Hsueh A.J.W. (2004) Paracrine regulation of mammalian oocyte maturation and male germ cell survival. *Proceedings of the National Academy of Sciences* **101**, 7323-8.

Kawashima T. (2019) Comparative and Evolutionary Genomics. In: *Encyclopedia of Bioinformatics and Computational Biology* (eds. by Ranganathan S, Gribskov M, Nakai K & Schönbach C), pp. 257-67. Elsevier, Cambridge MA.

Kestler H.A. & Kuhl M. (2008) From individual Wnt pathways towards a Wnt signalling network. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**, 1333-47.

Kijas J.W., Hadfield T., Naval Sanchez M. & Cockett N. (2016) Genome-wide association reveals the locus responsible for four-horned ruminant. *Animal Genetics* **47**, 258-62.

Kikuta H., Laplante M., Navratilova P., Komisarczuk A.Z., Engstrom P.G., Fredman D., Akalin A., Caccamo M., Sealy I., Howe K., Ghislain J., Pezeron G., Mourrain P., Ellingsen S., Oates A.C., Thisse C., Thisse B., Foucher I., Adolf B., Geling A., Lenhard B. & Becker T.S. (2007) Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. *Genome Research* **17**, 545-55.

Killoy K.M., Harlan B.A., Pehar M. & Vargas M.R. (2020) FABP7 upregulation induces a neurotoxic phenotype in astrocytes. *Glia* **68**, 2693-704.

Kim C.H., Kim S., Kim S.H., Roh J., Jin H. & Song B. (2020) Role of densin-180 in mouse ventral hippocampal neurons in 24-hr retention of contextual fear conditioning. *Brain and Behavior* **10**, 1-19.

Kim D., Langmead B. & Salzberg S.L. (2015) HISAT: a fast spliced aligner with low memory requirements. *Nature Methods* **12**, 357-60.

Kim J., Lo L., Dormand E. & Anderson D.J. (2003) SOX10 maintains multipotency and inhibits neuronal differentiation of neural crest stem cells. *Neuron* **38**, 17-31.

Klein R. (2004) Eph/ephrin signaling in morphogenesis, neural development and plasticity. *Current Opinion in Cell Biology* **16**, 580-9.

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

Kozomara A., Birgaoanu M. & Griffiths-Jones S. (2019) miRBase: from microRNA sequences to function. *Nucleic Acids Research* **47**, D155-D62.

Krasnoselsky A., Massay M., Defrances M., Michalopoulos G., Zarnegar R. & Ratner N. (1994) Hepatocyte growth factor is a mitogen for Schwann cells and is present in neurofibromas. *The Journal of Neuroscience* **14**, 7284-90.

Kressler D., Hurt E. & Baßler J. (2017) A puzzle of life: Crafting ribosomal subunits. *Trends in Biochemical Sciences* **42**, 640-54.

Krietenstein N., Abraham S., Venev S.V., Abdennur N., Gibcus J., Hsieh T.-H.S., Parsi K.M., Yang L., Maehr R., Mirny L.A., Dekker J. & Rando O.J. (2020) Ultrastructural Details of Mammalian Chromosome Architecture. *Molecular Cell* **78**, 554-65.e7.

Krueger F. (2015) TrimGalore. The Babraham Institute.

Kulesa P.M., Bailey C.M., Kasemeier-Kulesa J.C. & McLennan R. (2010) Cranial neural crest migration: New rules for an old road. *Developmental Biology* **344**, 543-54.

Kumagai J., Hsu S.Y., Matsumi H., Roh J.-S., Fu P., Wade J.D., Bathgate R.A.D. & Hsueh A.J.W. (2002) INSL3/Leydig Insulin-like peptide activates the LGR8 receptor important in testis descent. *Journal of Biological Chemistry* **277**, 31283-6.

Lan Y., Jia S. & Jiang R. (2014) Molecular patterning of the mammalian dentition. *Seminars in Cell and Developmental Biology* **25-26**, 61-70.

Laurie L.E., Kokubo H., Nakamura M., Saga Y. & Funato N. (2016) The Transcription Factor Hand1 Is Involved In Runx2-Ihh-Regulated Endochondral Ossification. *PLOS One* **11**, e0150263.

Lee A.P., Koh E.G.L., Tay A., Brenner S. & Venkatesh B. (2006) Highly conserved syntenic blocks at the vertebrate Hox loci and conserved regulatory elements within and outside Hox gene clusters. *Proceedings of the National Academy of Sciences* **103**, 6994-9.

Lee S.-E., Jeong S., Lee U. & Chang S. (2019) SGIP1α functions as a selective endocytic adaptor for the internalization of synaptotagmin 1 at synapses. *Molecular Brain* **12**, 1-11.

León A., Aparicio G.I. & Scorticati C. (2021) Neuronal glycoprotein M6a: An emerging molecule in chemical synapse formation and dysfunction. *Frontiers in Synaptic Neuroscience* **13**, 1-12.

Leung J.Y., Kolligs F.T., Wu R., Zhai Y., Kuick R., Hanash S., Cho K.R. & Fearon E.R. (2002) Activation of AXIN2 expression by β-Catenin-T cell factor. *Journal of Biological Chemistry* **277**, 21657-65.

Lewis M.W., Li S. & Franco H.L. (2019) Transcriptional control by enhancers and enhancer RNAs. *Transcription* **10**, 171-86.

Li C., Sheard P.W., Corson I.D. & Suttie J.M. (1993) Pedicle and antler development following sectioning of the sensory nerves to the antlerogenic region of red deer (Cervus elaphus). *Journal of Experimental Zoology* **267**, 188-97.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G. & Durbin R. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-9.

Li H., Lu Y., Smith H.K. & Richardson W.D. (2007) Olig1 and Sox10 interact synergistically to drive myelin basic protein transcription in oligodendrocytes. *Journal of Neuroscience* **27**, 14375-82.

Li J., Chatzeli L., Panousopoulou E., Tucker A.S. & Green J.B.A. (2016) Epithelial stratification and placode invagination are separable functions in early morphogenesis of the molar tooth. *Development* **143**, 670-81.

Li M.N., Wu X.Y., Guo X., Bao P.J., Ding X.Z., Chu M., Liang C.N. & Yan P. (2018) Comparative iTRAQ proteomics revealed proteins associated with horn development in yak. *Proteome Science* **16**, 1-11.

Li R., Beaudoin F., Ammah A.A., Bissonnette N., Benchaar C., Zhao X., Lei C. & Ibeagha-Awemu E.M. (2015) Deep sequencing shows microRNA involvement in bovine mammary gland adaptation to diets supplemented with linseed oil or safflower oil. *BMC Genomics* **16**, 1-16.

Li Y., Sheftic S.R., Grigoriu S., Schwieters C.D., Page R. & Peti W. (2020) The structure of the RCAN1:CN complex explains the inhibition of and substrate recruitment by calcineurin. *Science Advances* **6**, eaba3681.

Liao Y., Smyth G.K. & Shi W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-30.

Lin G.-H. & Zhang L. (2020) Apical ectodermal ridge regulates three principal axes of the developing limb. *Journal of Zhejiang University-SCIENCE B* **21**, 757-66.

Lin J.C., Ho W.-H., Gurney A. & Rosenthal A. (2003) The netrin-G1 ligand NGL-1 promotes the outgrowth of thalamocortical axons. *Nature Neuroscience* **6**, 1270-6.

Liu D.-C., Jow G.-M., Chuang C.-C., Peng Y.-J., Hsu P.-H. & Tang C.-Y. (2013) Densin-180 is not a transmembrane protein. *Cell Biochemistry and Biophysics* **67**, 773-83.

Liu H., Aramaki M., Fu Y. & Forrest D. (2017) Retinoid-Related Orphan Receptor β and Transcriptional Control of Neuronal Differentiation. In: *Current Topics in Developmental Biology* (eds. by Forrest D & Tsai S), pp. 227-55. Elsevier.

Liu J.A., Tai A., Hong J., Cheung M.P.L., Sham M.H., Cheah K.S.E., Cheung C.W. & Cheung M. (2020a) Fbxo9 functions downstream of Sox10 to determine neuron-glial fate choice in the dorsal root ganglia through Neurog2 destabilization. *Proceedings of the National Academy of Sciences* **117**, 4199-210.

Liu R., Holik A.Z., Su S., Jansz N., Chen K., Leong H.S., Blewitt M.E., Asselin-Labat M.-L., Smyth G.K. & Ritchie M.E. (2015) Why weight? Modelling sample and observational level variability improves power in RNA-seq analyses. *Nucleic Acids Research* **43**, e97.

Liu S., Gao Y., Canela-Xandri O., Wang S., Yu Y., Cai W., Li B., Pairo-Castineira E., D'Mellow K., Rawlik K., Xia C., Yao Y., Li X., Yan Z., Li C., Rosen B.D., Van Tassell C.P., Vanraden P.M., Zhang S., Ma L., Cole J.B., Liu G.E., Tenesa A. & Fang L. (2020b) A comprehensive catalogue of regulatory variants in the cattle transcriptome. Cold Spring Harbor Laboratory.

Liu X., Sun R., Chen J., Liu L., Cui X., Shen S., Cui G., Ren Z. & Yu Z. (2020c) Crosstalk mechanisms between HGF/c-Met axis and ncRNAs in malignancy. *Frontiers in Cell and Developmental Biology* **8**, 1-18.

Low W.Y., Tearle R., Liu R., Koren S., Rhie A., Bickhart D.M., Rosen B.D., Kronenberg Z.N., Kingan S.B., Tseng E., Thibaud-Nissen F., Martin F.J., Billis K., Ghurye J., Hastie A.R., Lee J., Pang A.W.C., Heaton M.P., Phillippy A.M., Hiendleder S., Smith T.P.L. & Williams J.L. (2020) Haplotype-resolved genomes provide insights into structural

variation and gene content in Angus and Brahman cattle. *Nature Communications* **11**, 1-14.

Lu B., Pang P.T. & Woo N.H. (2005) The yin and yang of neurotrophin action. *Nature Reviews Neuroscience* **6**, 603-14.

Lühken G G., Krebs S., Rothammer S., Kupper J., Mioc B., Russ I. & Medugorac I. (2016) The 1.78-kb insertion in the 3'-untranslated region of RXFP2 does not segregate with horn status in sheep breeds with variable horn status. *Genetics Selection Evolution* **48**, 1-14.

Luo Z., Rhie S. & Farnham P. (2019) The enigmatic HOX genes: Can we crack their code? *Cancers* **11**, 1-14.

Lupiáñez D.G., Kraft K., Heinrich V., Krawitz P., Brancati F., Klopocki E., Horn D., Kayserili H., Opitz J.M., Laxova R., Santos-Simarro F., Gilbert-Dussardier B., Wittler L., Borschiwer M., Haas S.A., Osterwalder M., Franke M., Timmermann B., Hecht J., Spielmann M., Visel A. & Mundlos S. (2015) Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* **161**, 1012-25.

Lupiáñez D.G., Spielmann M. & Mundlos S. (2016) Breaking TADs: how alterations of chromatin domains result in disease. *Trends in Genetics* **32**, 225-37.

Luppino J.M. & Joyce E.F. (2020) Single cell analysis pushes the boundaries of TAD formation and function. *Current Opinion in Genetics & Development* **61**, 25-31.

Luukko K. & Kettunen P. (2014) Coordination of tooth morphogenesis and neuronal development through tissue interactions: Lessons from mouse models. *Experimental Cell Research* **325**, 72-7.

Lyons R. & Randhawa I. (2020) Improving the Australian poll gene marker test. Meat and Livestock Australia, Sydney, Australia.

Madekurozwa M.-C.N. (1996) Studies on the development and innervation of the bovine and caprine horn. In: *Veterinary Medicine*, p. 339. University of Glasglow, Glasgow.

Maina F., Hilton M.C., Ponzetto C., Davies A.M. & Klein R. (1997) Met receptor signaling is required for sensory nerve development and HGF promotes axonal growth and survival of sensory neurons. *Genes & Development* **11**, 3341-50.

Mann R.S., Lelli K.M. & Joshi R. (2009) Hox specificity: Unique roles for cofactors and collaborators. *Current Topics in Developmental Biology* **88**, 63-101.

Matis C., Oury F., Remacle S., Lampe X., Gofflot F., Picard J.J., Rijli F.M. & Rezsohazy R. (2007) Identification of Lmo1 as part of a Hox-dependent regulatory network for hindbrain patterning. *Developmental Dynamics* **236**, 2675-84.

Medugorac I., Graf A., Grohs C., Rothammer S., Zagdsuren Y., Gladyr E., Zinovieva N., Barbieri J., Seichter D., Russ I., Eggen A., Hellenthal G., Brem G., Blum H., Krebs S. & Capitan A. (2017) Whole-genome analysis of introgressive hybridization and characterization of the bovine legacy of Mongolian yaks. *Nature Genetics* **49**, 470-5.

Medugorac I., Seichter D., Graf A., Russ I., Blum H., Goepel K.H., Rothammer S., Foerster M. & Krebs S. (2012) Bovine polledness - an autosomal dominant trait with allelic heterogeneity. *PLOS One* **7**, e39477.

Mei L. & Nave K.-A. (2014) Neuregulin-ERBB signaling in the nervous system and neuropsychiatric diseases. *Neuron* **83**, 27-49.

Mennesson M., Rydgren E., Lipina T., Sokolowska E., Kulesskaya N., Morello F., Ivakine E., Voikar V., Risbrough V., Partanen J. & Hovatta I. (2019) Kainate receptor auxiliary subunit NETO2 is required for normal fear expression and extinction. *Neuropsychopharmacology* **44**, 1855-66.

Migault M., Sapkota S. & Bracken C.P. (2022) Transcriptional and post-transcriptional control of epithelial-mesenchymal plasticity: Why so many regulators? *Cellular and Molecular Life Sciences* **79**, 1-19.

Mildner M., Eckhart L., Lengauer B. & Tschachler E. (2002) Hepatocyte growth factor/scatter factor inhibits UVB-induced apoptosis of human keratinocytes but not of keratinocyte-derived cell lines via the Phosphatidylinositol 3-Kinase/AKT Pathway. *Journal of Biological Chemistry* **277**, 14146-52.

Moens C.B. & Selleri L. (2006) Hox cofactors in vertebrate development. *Developmental Biology* **291**, 193-206.

Montanari M.P., Tran N.V. & Shimmi O. (2021) Regulation of spatial distribution of BMP ligands for pattern formation. *Developmental Dynamics* **251**, 178–92.

Morgan J.M., Navabi H., Schmid K.W. & Jasani B. (1994) Possible role of tissue-bound calcium ions in citrate-mediated high-temperature antigen retrieval. *The Journal of Pathology* **174**, 301-7.

Morioka M.S., Kawaji H., Nishiyori-Sueki H., Murata M., Kojima-Ishiyama M., Carninci P. & Itoh M. (2020) Cap analysis of gene expression (CAGE): A quantitative and genome-wide assay of transcription start sites. In: *Bioinformatics for Cancer Immunotherapy: Methods and Protocols* (ed. by Boegel S), pp. 277-301. Springer US, New York, NY.

Mueller O., Lightfoot S. & Schroeder A. (2016) RNA integrity number (RIN) – Standardization of RNA quality control. Agilent Technologies, Waldbronn, Germany.

Murillo-Maldonado J.M. & Riesgo-Escovar J.R. (2019) The various and shared roles of lncRNAs during development. *Developmental Dynamics* **248**, 1059-69.

National Centre for Biotechnology Information (2020) Bos taurus ESTs. https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&PROG_DEF=blastn&BLAST_PROG_DEF=megaBlast&BLAST_SPEC=OGP__9913__10708.

Nesbit M., Mamo J.C., Majimbi M., Lam V. & Takechi R. (2021) Automated quantitative analysis of ex vivo blood-brain barrier permeability using Intellesis machine-learning. *Frontiers in Neuroscience* **15**, 1-8.

Nguyen Q.H., Tellam R.L., Naval-Sanchez M., Porto-Neto L.R., Barendse W., Reverter A., Hayes B., Kijas J. & Dalrymple B.P. (2018) Mammalian genomic regulatory regions predicted by utilizing human genomics, transcriptomics, and epigenetics data. *Gigascience* **7**, 1-17.

Nie X., Luukko K. & Kettunen P. (2006) BMP signalling in craniofacial development. *The International Journal of Developmental Biology* **50**, 511-21.

O'Brien J., Hayder H., Zayed Y. & Peng C. (2018) Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Frontiers in Endocrinology* **9**, 1-12.

Pailhoux E., Vigier B., Chaffaux S., Servel N., Taourit S., Furet J.-P., Fellous M., Grosclaude F., Cribiu E.P., Cotinot C. & Vaiman D. (2001) A 11.7-kb deletion triggers intersexuality and polledness in goats. *Nature Genetics* **29**, 453-8.

Palla G., Fischer D.S., Regev A. & Theis F.J. (2022) Spatial components of molecular tissue biology. *Nature Biotechnology*.

Pan Z., Li S., Liu Q., Wang Z., Zhou Z., Di R., Miao B., Hu W., Wang X., Hu X., Xu Z., Wei D., He X., Yuan L., Guo X., Liang B., Wang R., Li X., Cao X., Dong X., Xia Q., Shi H., Hao G., Yang J., Luosang C., Zhao Y., Jin M., Zhang Y., Lv S., Li F., Ding G., Chu M. & Li Y. (2018) Whole-genome sequences of 89 Chinese sheep suggest role of RXFP2 in the development of unique horn phenotype as response to semi-feralization. *Gigascience* **7**, 1–15.

Park S.-M., Park S.-J., Kim H.-J., Kwon O.-H., Kang T.-W., Sohn H.-A., Kim S.-K., Moo Noh S., Song K.-S., Jang S.-J., Sung Kim Y. & Kim S.-Y. (2013) A known expressed sequence tag, BM742401, is a potent lincRNA inhibiting cancer metastasis. *Experimental & Molecular Medicine* **45**, e31.

Parkinson J. & Blaxter M. (2009) Expressed sequence tags: An overview. In: *Expressed Sequence Tags (ESTs)* (ed. by Parkinson J), pp. 1-12. Humana Press.

Patel A., Yamashita N., Ascano M., Bodmer D., Boehm E., Bodkin-Clarke C., Ryu Y.K. & Kuruvilla R. (2015) RCAN1 links impaired neurotrophin trafficking to aberrant development of the sympathetic nervous system in Down syndrome. *Nature Communications* **6**, 1-17.

Pathirana I.N., Kawate N., Büllesbach E.E., Takahashi M., Hatoya S., Inaba T. & Tamada H. (2012) Insulin-like peptide 3 stimulates testosterone secretion in mouse Leydig cells via cAMP pathway. *Regulatory Peptides* **178**, 102-6.

Peng Q., Pevzner P.A. & Tesler G. (2006) The fragile breakage versus random breakage models of chromosome evolution. *PLOS Computational Biology* **2**, e14.

Petrie E.J., Lagaida S., Sethi A., Bathgate R.A.D. & Gooley P.R. (2015) In a class of their own – RXFP1 and RXFP2 are unique members of the LGR family. *Frontiers in Endocrinology* **6**, 1-9.

Pevzner P. & Tesler G. (2003) Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution. *Proceedings of the National Academy of Sciences* **100**, 7672-7.

Pigossi S., Medeiros M., Saska S., Cirelli J. & Scarel-Caminaga R. (2016) Role of osteogenic growth peptide (OGP) and OGP(10–14) in bone regeneration: A review. *International Journal of Molecular Sciences* **17**, 1-15.

Pitia A.M., Minagawa I., Uera N., Hamano K.-I., Sugawara Y., Nagura Y., Hasegawa Y., Oyamada T., Sasada H. & Kohsaka T. (2015) Expression of insulin-like factor 3 hormone-receptor system in the reproductive organs of male goats. *Cell and Tissue Research* **362**, 407-20.

Pitia A.M., Uchiyama K., Sano H., Kinukawa M., Minato Y., Sasada H. & Kohsaka T. (2017) Functional insulin-like factor 3 (INSL3) hormone-receptor system in the testes and

spermatozoa of domestic ruminants and its potential as a predictor of sire fertility. *Animal Science Journal* **88**, 678–90.

Polychronopoulos D., James, Nash A.J., Tan G. & Lenhard B. (2017) Conserved non-coding elements: developmental gene regulation meets genome organization. *Nucleic Acids Research* **45**, 12611-24.

Puthiyaveetil J.S.V., Kota K., Chakkarayan R., Chakkarayan J. & Thodiyil A.K.P. (2016) Epithelial - mesenchymal interactions in tooth development and the significant role of growth factors and genes with emphasis on mesenchyme - a review. *Journal of clinical and diagnostic research* **10**, ZE05-ZE9.

Quintes S., Brinkmann B.G., Ebert M., Fröb F., Kungl T., Arlt F.A., Tarabykin V., Huylebroeck D., Meijer D., Suter U., Wegner M., Sereda M.W. & Nave K.-A. (2016) Zeb2 is essential for Schwann cell differentiation, myelination and nerve repair. *Nature Neuroscience* **19**, 1050-9.

Rapizzi E., Benvenuti S., Deledda C., Martinelli S., Sarchielli E., Fibbi B., Luciani P., Mazzanti B., Pantaleo M., Marroncini G., Vannelli G.B., Maggi M., Mannelli M., Luconi M. & Peri A. (2020) A unique neuroendocrine cell model derived from the human foetal neural crest. *Journal of Endocrinological Investigation* **43**, 1259-69.

Rehfeld F., Maticzka D., Grosser S., Knauff P., Eravci M., Vida I., Backofen R. & Wulczyn F.G. (2018) The RNA-binding protein ARPP21 controls dendritic branching by functionally opposing the miRNA it hosts. *Nature Communications* **9**, 1-13.

Reisinger S.N., Bilban M., Stojanovic T., Derdak S., Yang J., Cicvaric A., Horvath O., Sideromenos S., Zambon A., Monje F.J., Boehm S. & Pollak D.D. (2020) Lmo3 deficiency in the mouse is associated with alterations in mood-related behaviors and a depression-biased amygdala transcriptome. *Psychoneuroendocrinology* **111**, 1-9.

Rella L., Fernandes Póvoa E.E., Mars J., Ebbing A.L.P., Schoppink L., Betist M.C. & Korswagen H.C. (2021) A switch from noncanonical to canonical Wnt signaling stops neuroblast migration through a Slt–Robo and RGA-9b/ARHGAP–dependent mechanism. *Proceedings of the National Academy of Sciences* **118**, e2013239118.

Remsburg C., Konrad K., Sampilo N.F. & Song J.L. (2019) Analysis of microRNA functions. In: *Methods in Cell Biology* (eds. by Hamdoun A & Foltz KR), pp. 323-34. Elsevier.

Rodriques S.G., Stickels R.R., Goeva A., Martin C.A., Murray E., Vanderburg C.R., Welch J., Chen L.M., Chen F. & Macosko E.Z. (2019) Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463-7.

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

Ross E.M., Sanjana H., Nguyen L.T., Cheng Y., Moore S.S. & Hayes B.J. (2022) Extensive variation in gene expression is revealed in 13 fertility-related genes using RNA-Seq, ISO-Seq, and CAGE-Seq from Brahman cattle. *Frontiers in Genetics* **13**.

Salhotra A., Shah H.N., Levi B. & Longaker M.T. (2020) Mechanisms of bone development and repair. *Nature Reviews Molecular Cell Biology* **21**, 696-711.

Sanborn A.L., Rao S.S.P., Huang S.-C., Durand N.C., Huntley M.H., Jewett A.I., Bochkov I.D., Chinnappan D., Cutkosky A., Li J., Geeting K.P., Gnirke A., Melnikov A., McKenna D., Stamenova E.K., Lander E.S. & Aiden E.L. (2015) Chromatin extrusion explains

key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences* **112**, E6456-E65.

Sanges R., Hadzhiev Y., Gueroult-Bellone M., Roure A., Ferg M., Meola N., Amore G., Basu S., Brown E.R., De Simone M., Petrera F., Licastro D., Strähle U., Banfi S., Lemaire P., Birney E., Müller F. & Stupka E. (2013) Highly conserved elements discovered in vertebrates are present in non-syntenic loci of tunicates, act as enhancers and can be transcribed during development. *Nucleic Acids Research* **41**, 3600-18.

Sarkar L. & Sharpe P.T. (1999) Expression of Wnt signalling pathway genes during tooth development. *Mechanisms of Development* **85**, 197-200.

Sarver D.C., Lei X. & Wong G.W. (2021) FAM19A (TAFA): An emerging family of neurokines with diverse functions in the central and peripheral nervous system. *ACS Chemical Neuroscience* **12**, 945-58.

Sasaki T., Ito Y., Xu X., Han J., Bringas P., Maeda T., Slavkin H.C., Grosschedl R. & Chai Y. (2005) LEF1 is a critical epithelial survival factor during tooth morphogenesis. *Developmental Biology* **278**, 130-43.

Schindelin J., Arganda-Carreras I., Frise E., Kaynig V., Longair M., Pietzsch T., Preibisch S., Rueden C., Saalfeld S., Schmid B., Tinevez J.-Y., White D.J., Hartenstein V., Eliceiri K., Tomancak P. & Cardona A. (2012) Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**, 676-82.

Schubert M., Lindgreen S. & Orlando L. (2016) AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC research notes* **9**, 1-7.

Schuster F., Aldag P., Frenzel A., Hadeler K.-G., Lucas-Hahn A., Niemann H. & Petersen B. (2020) CRISPR/Cas12a mediated knock-in of the Polled Celtic variant to produce a polled genotype in dairy cattle. *Scientific Reports* **10**, 1-9.

Sedaghat K., Shen P.J., Finkelstein D.I., Henderson J.M. & Gundlach A.L. (2008) Leucine-rich repeat-containing G-protein-coupled receptor 8 in the rat brain: Enrichment in thalamic neurons and their efferent projections. *Neuroscience* **156**, 319-33.

Seo J.-Y., Jung Y., Kim D.-Y., Ryu H.G., Lee J., Kim S.W. & Kim K.-T. (2019) DAP5 increases axonal outgrowth of hippocampal neurons by enhancing the cap-independent translation of DSCR1.4 mRNA. *Cell Death & Disease* **10**, 1-15.

Serpico D. (2020) Beyond quantitative and qualitative traits: three telling cases in the life sciences. *Biology & Philosophy* **35**, 1-26.

Servant N., Varoquaux N., Lajoie B.R., Viara E., Chen C.-J., Vert J.-P., Heard E., Dekker J. & Barillot E. (2015) HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biology* **16**, 1-11.

Shen H., Tian Y., Yao X., Liu W., Zhang Y. & Yang Z. (2017) MiR-99a inhibits keratinocyte proliferation by targeting Frizzled-5 (FZD5) / FZD8 through β-catenin signaling in psoriasis. *Pharmazie* **72**, 461-7.

Shruti K., Shrey K. & Vibha R. (2011) Micro RNAs: Tiny sequences with enormous potential. *Biochemical and Biophysical Research Communications* **407**, 445-9.

Shyer A.E., Rodrigues A.R., Schroeder G.G., Kassianidou E., Kumar S. & Harland R.M. (2017) Emergent cellular self-organization and mechanosensation initiate follicle pattern in the avian skin. *Science* **357**, 811-5.

Simon R., Drögemüller C. & Lühken G. (2022) The complex and diverse genetic architecture of the absence of horns (polledness) in domestic ruminants, including goats and sheep. *Genes* **13**, 832.

Solloway M.J. & Robertson E.J. (1999) Early embryonic lethality in Bmp5;Bmp7 double mutant mice suggests functional redundancy within the 60A subgroup. *Development* **126**, 1753-68.

Solovieva T. & Bronner M. (2021) Schwann cell precursors: Where they come from and where they go. *Cells & Development* **166**, 1-7.

Srinivasan R., Sun G., Keles S., Jones E.A., Jang S.-W., Krueger C., Moran J.J. & Svaren J. (2012) Genome-wide analysis of EGR2/SOX10 binding in myelinating peripheral nerve. *Nucleic Acids Research* **40**, 6449-60.

Stelzer G., Rosen N., Plaschkes I., Zimmerman S., Twik M., Fishilevich S., Stein T.I., Nudel R., Lieder I., Mazor Y., Kaplan S., Dahary D., Warshawsky D., Guan-Golan Y., Kohn A., Rappaport N., Safran M. & Lancet D. (2016) The GeneCards suite: From gene data mining to disease genome sequence analyses. *Current Protocols in Bioinformatics* **54**, 1-33.

Svandova E., Peterkova R., Matalova E. & Lesot H. (2020) Formation and developmental specification of the odontogenic and osteogenic mesenchymes. *Frontiers in Cell and Developmental Biology* **8**, 1-16.

Szewczyk L.M., Brozko N., Nagalski A., Röckle I., Werneburg S., Hildebrandt H., Wisniewska M.B. & Kuznicki J. (2017) ST8SIA2 promotes oligodendrocyte differentiation and the integrity of myelin and axons. *Glia* **65**, 34-49.

Tani S., Chung U.-I., Ohba S. & Hojo H. (2020) Understanding paraxial mesoderm development and sclerotome specification for skeletal repair. *Experimental & Molecular Medicine* **52**, 1166-77.

Taylor J., Chung K.-H., Figueroa C., Zurawski J., Dickson H.M., Brace E.J., Avery A.W., Turner D.L. & Vojtek A.B. (2008) The scaffold protein POSH regulates axon outgrowth. *Molecular Biology of the Cell* **19**, 5181-92.

Theveneau E. & Mayor R. (2012) Neural crest migration: interplay between chemorepellents, chemoattractants, contact inhibition, epithelial-mesenchymal transition, and collective cell migration. *Wiley Interdisciplinary Reviews: Developmental Biology* **1**, 435-45.

Thiede-Stan N.K. & Schwab M.E. (2015) Attractive and repulsive factors act through multi-subunit receptor complexes to regulate nerve fiber growth. *Journal of Cell Science* **128**, 2403-14.

Ting M.-C., Wu N.L., Roybal P.G., Sun J., Liu L., Yen Y. & Maxson R.E. (2009) EphA4 as an effector of Twist1 in the guidance of osteogenic precursor cells during calvarial bone growth and in craniosynostosis. *Development* **136**, 855-64.

Tomioka T., Shimazaki T., Yamauchi T., Oki T., Ohgoh M. & Okano H. (2014) LIM homeobox 8 (Lhx8) is a key regulator of the cholinergic neuronal function via a tropomyosin receptor kinase A (TrkA)-mediated positive feedback loop. *Journal of Biological Chemistry* **289**, 1000-10.

Tong M., Jun T., Nie Y., Hao J. & Fan D. (2019) The role of the Slit/Robo signaling pathway. *Journal of Cancer* **10**, 2694-705.

Tukel T., Šošić D., Al-Gazali L.I., Erazo M., Casasnovas J., Franco H.L., Richardson J.A., Olson E.N., Cadilla C.L. & Desnick R.J. (2010) Homozygous nonsense mutations in TWIST2 cause setleis syndrome. *The American Journal of Human Genetics* **87**, 289-96.

Uhlén M., Fagerberg L., Hallström B.M., Lindskog C., Oksvold P., Mardinoglu A., Sivertsson Å., Kampf C., Sjöstedt E., Asplund A., Olsson I., Edlund K., Lundberg E., Navani S., Szigyarto C.A.-K., Odeberg J., Djureinovic D., Takanen J.O., Hober S., Alm T., Edqvist P.-H., Berling H., Tegel H., Mulder J., Rockberg J., Nilsson P., Schwenk J.M., Hamsten M., Feilitzen K.v., Forsberg M., Persson L., Johansson F., Zwahlen M., Heijne G.v., Nielsen J. & Pontén F. (2015) Tissue-based map of the human proteome. *Science* **347**, 1-9.

Utsunomiya Y.T., Torrecilha R.B.P., Milanesi M., Paulan S.D.C., Utsunomiya A.T.H. & Garcia J.F. (2019) Hornless Nellore cattle (Bos indicus) carrying a novel 110 kbp duplication variant of the polled locus. *Animal Genetics*, 1-2.

Van De Putte T., Maruhashi M., Francis A., Nelles L., Kondoh H., Huylebroeck D. & Higashi Y. (2003) Mice lacking Zfhx1b, the gene that codes for smad-interacting protein-1, reveal a role for multiple neural crest cell defects in the etiology of Hirschsprung disease–mental retardation syndrome. *The American Journal of Human Genetics* **72**, 465-70.

Van Der Vaart B., Manatschal C., Grigoriev I., Olieric V., Gouveia S.M., Bjelić S., Demmers J., Vorobjev I., Hoogenraad C.C., Steinmetz M.O. & Akhmanova A. (2011) SLAIN2 links microtubule plus end–tracking proteins and controls microtubule growth in interphase. *Journal of Cell Biology* **193**, 1083-99.

Van Genderen C., Okamura R.M., Fariñas I., Quo R.G., Parslow T.G., Bruhn L. & Grosschedl R. (1994) Development of several organs that require inductive epithelial-mesenchymal interactions is impaired in LEF-1-deficient mice. *Genes & Development* **8**, 2691-703.

Van Loon K., Huijbers E.J.M. & Griffioen A.W. (2021) Secreted frizzled-related protein 2: A key player in noncanonical Wnt signaling and tumor angiogenesis. *Cancer and Metastasis Reviews* **40**, 191-203.

Varadi M., Anyango S., Deshpande M., Nair S., Natassia C., Yordanova G., Yuan D., Stroe O., Wood G., Laydon A., Žídek A., Green T., Tunyasuvunakool K., Petersen S., Jumper J., Clancy E., Green R., Vora A., Lutfi M., Figurnov M., Cowie A., Hobbs N., Kohli P., Kleywegt G., Birney E., Hassabis D. & Velankar S. (2021) AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research* **50**, D439-D44.

Vickovic S., Eraslan G., Salmén F., Klughammer J., Stenbeck L., Schapiro D., Äijö T., Bonneau R., Bergenstråhle L., Navarro J.F., Gould J., Griffin G.K., Borg Å., Ronaghi M., Frisén J., Lundeberg J., Regev A. & Ståhl P.L. (2019) High-definition spatial transcriptomics for in situ tissue profiling. *Nature Methods* **16**, 987-90.

Vlotides G., Zitzmann K., Stalla G.K. & Auernhammer C.J. (2004) Novel neurotrophin-1/B cell-stimulating factor-3 (NNT-1/BSF-3) / cardiotrophin-like cytokine (CLC)—a novel gp130 cytokine with pleiotropic functions. *Cytokine & Growth Factor Reviews* **15**, 325-36.

Wang D., Berg D., Ba H., Sun H., Wang Z. & Li C. (2019a) Deer antler stem cells are a novel type of cells that sustain full regeneration of a mammalian organ—deer antler. *Cell Death & Disease* **10**, 1-13.

Wang G. & Gill C.A. (2021) Dissection of the scurs phenotype to refine the mapping of scurs. Texas Agricultural and Mechanical University, Poster presented at: International Society for Animal Genetics conference, July 26-30, 2021, [Online].

Wang M., Hancock T.P., Chamberlain A.J., Jagt C.J.V., Pryce J.E., Cocks B.G., Goddard M.E. & Hayes B.J. (2018) Putative bovine topological association domains and CTCF binding motifs can reduce the search space for causative regulatory variants of complex traits. *BMC Genomics* **19**, 1-17.

Wang M., Hancock T.P., Macleod I.M., Pryce J.E., Cocks B.G. & Hayes B.J. (2017) Putative enhancer sites in the bovine genome are enriched with variants affecting complex traits. *Genetics Selection Evolution* **49**, 1-16.

Wang M., Jin H., Tang D., Huang S., Zuscik M.J. & Chen D. (2011) Smad1 plays an essential role in bone development and postnatal bone formation. *Osteoarthritis and Cartilage* **19**, 751-62.

Wang R.N., Green J., Wang Z., Deng Y., Qiao M., Peabody M., Zhang Q., Ye J., Yan Z., Denduluri S., Idowu O., Li M., Shen C., Hu A., Haydon R.C., Kang R., Mok J., Lee M.J., Luu H.L. & Shi L.L. (2014a) Bone morphogenetic protein (BMP) signaling in development and human diseases. *Genes & Diseases* **1**, 87-105.

Wang S., Ji D., Yang Q., Li M., Ma Z., Zhang S. & Li H. (2019b) NEFLb impairs early nervous system development via regulation of neuron apoptosis in zebrafish. *Journal of Cellular Physiology* **234**, 11208-18.

Wang W., Rai A., Hur E.-M., Smilansky Z., Chang K.T. & Min K.-T. (2016) DSCR1 is required for both axonal growth cone extension and steering. *Journal of Cell Biology* **213**, 451-62.

Wang X.L., Zhou G.X., Li Q., Zhao D.F. & Chen Y.L. (2014b) Discovery of SNPs in RXFP2 related to horn types in sheep. *Small Ruminant Research* **116**, 133-6.

Wang Y., Zhang C.Z., Wang N.N., Li Z.P., Heller R.A.M., Liu R., Zhao Y., Han J.G., Pan X.Y., Zheng Z.Q., Dai X.Q., Chen C.S., Dou M.L., Peng S.J., Chen X.Q., Liu J., Li M., Wang K., Liu C., Lin Z.S., Chen L., Hao F., Zhu W.B., Song C.C., Zhao C., Zheng C.L., Wang J.M., Hu S.W., Li C.Y., Yang H., Jiang L., Li G.Y., Liu M.J., Sonstegard T.S., Zhang G.J., Jiang Y., Wang W. & Qiu Q. (2019c) Genetic basis of ruminant headgear and rapid antler regeneration. *Science* **364**, 1153-60.

Watkins R.J., Thomas M.G., Talbot C.J., Gottlob I. & Shackleton S. (2012) The role of FRMD7 in idiopathic infantile nystagmus. *Journal of Ophthalmology* **2012**, 1-7.

Waylen L.N., Nim H.T., Martelotto L.G. & Ramialison M. (2020) From whole-mount to single-cell spatial assessment of gene expression in 3D. *Communications Biology* **3**, 1-11.

Wheeler D.L., Church D.M., Federhen S., Lash A.E., Madden T.L., Pontius J.U., Schuler G.D., Schriml L.M., Sequeira E., Tatusova T.A. & Wagner L. (2003) Database resources of the National Center for Biotechnology. *Nucleic Acids Res* **31**, 28-33.

Wiedemar N. & Drögemüller C. (2015) A 1.8-kb insertion in the 3-UTR of RXFP2 is associated with polledness in sheep. *Animal Genetics* **46**, 457-61.

Wiedemar N., Tetens J., Jagannathan V., Menoud A., Neuenschwander S., Bruggmann R., Thaller G. & Drögemüller C. (2014) Independent POLLED mutations leading to complex gene expression differences in cattle. *PLOS One* **9**, e93435.

Wiener D.J., Wiedemar N., Welle M.M. & Drögemüller C. (2015) Novel features of the prenatal horn bud development in cattle (bos taurus). *PLOS One* **10**, e0127691.

Wiese K.E., Nusse R. & Van Amerongen R. (2018) Wnt signalling: conquering complexity. *Development* **145**, 1-9.

Wilkinson L. (2011) ggplot2: Elegant graphics for data analysis by Wickham, H. *Biometrics* **67**, 678-9.

Wong E.S., Zheng D., Tan S.Z., Bower N.I., Garside V., Vanwalleghem G., Gaiti F., Scott E., Hogan B.M., Kikuchi K., McGlinn E., Francois M. & Degnan B.M. (2020) Deep conservation of the enhancer regulatory code in animals. *Science* **370**, 1-8.

Wu C.Y. & Taneyhill L.A. (2019) Cadherin-7 mediates proper neural crest cell–placodal neuron interactions during trigeminal ganglion assembly. *genesis* **57**, e23264.

Xiong G.-J., Cheng X.-T., Sun T., Xie Y., Huang N., Li S., Lin M.-Y. & Sheng Z.-H. (2021) Defects in syntabulin-mediated synaptic cargo transport associate with autism-like synaptic dysfunction and social behavioral traits. *Molecular Psychiatry* **26**, 1472-90.

Yamaguchi N., Osaki M., Onuma K., Yumioka T., Iwamoto H., Sejima T., Kugoh H., Takenaka A. & Okada F. (2017) Identification of MicroRNAs involved in resistance to sunitinib in renal cell carcinoma cells. *Anticancer Research* **37**, 2985-92.

Yan Y. & Wang Q. (2021) BMP signaling: Lighting up the way for embryonic dorsoventral patterning. *Frontiers in Cell and Developmental Biology* **9**, 1-16.

Yu M. & Ren B. (2017) The three-dimensional organization of mammalian genomes. *Annual Review of Cell and Developmental Biology* **33**, 265-89.

Yun S.W., Leong C., Zhai D., Tan Y.L., Lim L., Bi X., Lee J.J., Kim H.J., Kang N.Y., Ng S.H., Stanton L.W. & Chang Y.T. (2012) Neural stem cell specific fluorescent chemical probe binding to FABP7. *Proceedings of the National Academy of Sciences* **109**, 10214-7.

Yuzaki M. (2008) Cbln and C1q family proteins – New transneuronal cytokines. *Cellular and Molecular Life Sciences* **65**, 1698-705.

Zagrebelsky M. & Korte M. (2014) Form follows function: BDNF and its involvement in sculpting the function and structure of synapses. *Neuropharmacology* **76**, 628-38.

Zhang M., Wang J., Deng C., Jiang M.H., Feng X., Xia K., Li W., Lai X., Xiao H., Ge R.-S., Gao Y. & Xiang A.P. (2017) Transplanted human p75-positive stem Leydig cells replace disrupted Leydig cells for testosterone production. *Cell Death & Disease* **8**, 1-10.

Zheng Q., Zheng X., Zhang L., Luo H., Qian L., Fu X., Liu Y., Gao Y., Niu M., Meng J., Zhang M., Bu G., Xu H. & Zhang Y.-W. (2017) The neuron-specific protein TMEM59L mediates oxidative stress-induced cell death. *Molecular Neurobiology* **54**, 4189-200.

Zhou A.X., Hartwig J.H. & Akyürek L.M. (2010) Filamins in cell signaling, transcription and organ development. *Trends Cell Biol* **20**, 113-23.

Zhou B.-R. & Bai Y. (2019) Chromatin structures condensed by linker histones. *Essays in Biochemistry* **63**, 75-87.

Zhou Y., Yuan J., Li Z., Wang Z., Cheng D., Du Y., Li W., Kan Q. & Zhang W. (2015) Genetic polymorphisms and function of the organic anion-transporting polypeptide 1A2 and its clinical relevance in drug disposition. *Pharmacology* **95**, 201-8.