Characterization of Transcription Factors and LncRNAs Involved in the Development of the Bat Wing

Zoe Gill

Department of Molecular and Cellular Biology
University of Cape Town
Rondebosch
7701

Supervisor:
Prof. Nicola Illing

February 2016

This dissertation was conducted in fulfilment of the requirements of a Masters of Science Degree in Molecular and Cellular Biology, MSc, at the University of Cape Town.

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

Declaration

I, Zoe Gill , know the meaning of plagiarism and declare that all of the work in the document, save for that which is properly acknowledged, is my own.

Signed by candidate

Date: 15th February 2016

Signature Removed

Abstract

Mammals have evolved a vast myriad of limb morphologies adapted for a wide range of activities. One of the most remarkable evolutionary adaptations of a mammalian limb is that of the forelimb wing of a bat used for powered flight. This capability evolved ~ 51 Mya from its arboreal ancestor without any fossil record of intermediate forms. To reconstruct how this transition occurred, an Evolutionary Developmental approach can be applied to investigate altered mechanisms present in bat limb development. Similar genes and signalling centres are present in both mice and bats, making mice a good model organism for comparison. This study used a pre-existing set of RNA-seq transcriptomes from three pivotal developmental stages (CS 15, CS 16 and CS 17) of bat development, to compare FL and HL gene expression. Of the list of differentially expressed genes, a subset was selected to characterise spatial expression patterns within the developing bat limb compared to mouse limbs by whole-mount *in situ* hybridisation. Five transcription factors: *Lef1*, *Lhx8*, *HoxA10*, *Mllt3* and *Tbx5*, as well as two Long non-coding RNAs: *Hottip* and *Tbx5-as1* were selected. Novel expression of *Mllt3* was detected in FL autopods at CS15, in a region slated to expand with digit elongation. *Lef1 in situ* signal was more robust in HL autopods of CS 15 embryos compared to FLs and equivalently staged mice. *Lhx8* displayed a strong signal in CS 16 and CS17 wrist tissue, as well as a faint signal in interdigital tissue in the FL autopods. The LncRNA *Hottip* displayed vastly different expression pattern between FL and HL, with staining being reduced in the digit and interdigital regions of the FL at CS 16L, whereas expression in the HL was robust in the digit, and even more so in the interdigital regions. The LncRNA *Tbx5-as1*, displayed a similar expression pattern to the known FL initiation transcription factor Tbx5 at late stages (CS 16L –CS 17L). Isoform characterization to validate the two LncRNAs, was performed on cDNA a CS 18L embryo. The cloned transcripts identified a new set of alternatively spliced isoforms for both LncRNAs. Unusual RNA-seq tracks in the *HoxA10* locus were investigated using qPCR. It was discovered this region is variable amongst biological samples; however there is a large reduction in expression in this region from CS 15 to CS 16.

Acknowledgements

I would like to thank my supervisor Professor Nicola Illing, who played a pivotal role in the concept and financing of this project. I would like to thank my lab-mates Stephen Schlebusch, Mandy Mason and Rafe Lyall, for being phenomenal scientific role models that have challenged me every day. I would like to thank the National Research Fund for awarding me the Innovation Scholarship, and the Functional Genomics grant. I would also like to thank Petra Muller for her eternal kindness and generosity with her equipment and time. I would like to thank my close friends for understanding the nature of my studies, and especially Robyn Waters and Cynthia Fan, for being a great support both in and out of the department. And lastly, but most importantly, I would like to thank my wonderful mother Eve Gill, for sticking out a job at the University of Cape Town for 22 years so that I could get the rebate on my fees. I would not have been able to complete this Masters without all this incredible support.

Table of Contents

Declaration..... 2

Acknowledgements..... 3

Abstract..... 4

Table of Contents..... 5

List of Illustrations..... 8

Glossary and Abbreviations 10

Chapter 1..... 15

1.1 Introduction 15

1.2 Vertebrate Limb Morphology 17

1.3 Evolutionary Developmental Biology..... 18

1.4 The Molecular Mechanisms that Shape Vertebrate Limbs 19

1.5 The Unique Morphology of the Bat Wing 22

1.6 The Molecular Mechanisms that Shape Bat Limbs..... 24

1.7 Strategies for Exploring Gene expression in Bat Limbs – RNA-seq..... 24

1.8 LncRNAs 26

1.9 Study Aims..... 28

Chapter 2 -Materials and Methods..... 29

2.1 Study Species: *Miniopterus natalensis* 29

2.1.2 Acquisition of Mouse Embryos 31

2.2 Whole-mount *in situ* Hybridization 32

2.3 Bioinformatics..... 43

2.3.1 Analysis of *M. schreibersii* RNA-seq Dataset..... 43

2.3.2 Bioinformatics Analysis of Cloned *in situ* Probes 44

2.3.3. Characterization of LncRNA Isoforms from RNA-seq Data..... 44

2.4 Characterizing *Hottip* and *Tbx5-as1* Transcripts..... 44

2.4.1 Primer Design..... 44

2.4.2 RNA Extraction ……..... 45
2.5 qPCR……..……..... 45
2.5.2 Primer Design……..... 45
2.5.3 MIQE Standards……..... 46
2.5.4 Sample Acquisition, Handling and Preparation ……..... 47
2.5.5 RNA Extractions……..... 47
2.5.6 cDNA Synthesis ……..... 48
2.5.7. λ gDNA Dilutions……..……..... 49
2.5.8. Experimental PCR Setup……..... 49
2.5.9 LRE Analysis……..... 49
Chapter 3 Transcription Factor Expression Analysis in Developing Mouse and Bat Limbs……..……..... 50
3.1 Introduction ……..... 50
3.1.1 *Lef1* ……..……..... 52
3.1.2 *Lhx8* ……..……..... 52
3.1.3 *Hox* Genes……..……... 53
3.1.4 *HoxA10*……..……..... 54
3.1.5 *Mllt3*……..……..... 55
3.2. Results……..……..... 58
3.2.1 Generation of In Situ Probes……..……..... 58
3.2.2 *M.natalensis* and *M. schreibersii* Dataset Comparison…….. 59
3.2.2.1 *Lef1* ……..……..... 61
3.2.2.2 *Lhx8*……..……..... 66
3.2.2.3 *HoxA10*……..……..... 69
3.2.2.4 *Mllt3*……..……..... 75
3.3 Discussion……..……..... 79
3.3.1 Novel region of *Lef1* expression in the HLs…….. 80
3.3.2 Novel zeugopod expression of *Lhx8*……..... 80
3.3.3 *HoxA10* ……..……..... 80
3.3.4. Novel expression domain of *Mllt3* in Bat FLs…….. 81

Chapter 4 LncRNA Expression Analysis and Isoform Characterisation in Developing Bat Limbs……..……..……..... 82

4.1 Introduction……..……..... 82

4.1.1 Identifying LncRNA in the *M. natalensis* dataset…… 82

4.1.2 *Tbx5-as1* ……..…..... 83

4.1.3 *Hottip* ……..……..... 84

4.1.4 Aims ……..……..... 87

4.2 Results……..……..... 88

4.2.1 Characterising Isoforms ……..……..... 88

4.2.2.Generatng *In situ* Probes ……..……..... 93

4.2.3 *M. natalensis* and *M. schreibersii* Dataset and WISH comparison …… 94

4.2.3.1 *Tbx5-as1* ……..……..……..... 94

4.2.3.2 *Hottip*……..……..……..... 98

4.3. Discussion ……..……..... 103

4.3.1 *Hottip* and *Tbx5-as1* Transcript Isoforms……..... 103

4.3.2 *Tbx5-as1* and *Tbx5* WISH……..……..... 104

Chapter 5 Conclusions and Recommendations……..……..... 105

5.1 The field of Evo Devo ……..... 105

5.2 Transcription Factors……..……..……..... 105

5.3 LncRNAs……..……..... 106

5.4 Limitations of This Study……..... 108

5.5 Concluding Comments……..……..……..... 108

List of illustrations

Fig.1.1 Variation in morphology of the three segments of mammalian limbs…….. 16

Fig.1.2 Human hand as an example of the basic plan of the tetrapod autopod……..... 17

Fig.1.3 Mouse limb-bud during early development with signalling centres domains 20

Fig.1.4 Anatomy of an adult *M. natalensis* with wing membrane and skeletal elements..... 23

Fig.1.5 Tissue sectioning used to generate the *M. schreibersii* RNA-seq dataset…… 26

Fig.1.6 The different types of LncRNAs…… 27

Fig.2.1 Serial dilution of Lambda gDNA as a LRE qPCR calibrator…… 48

Fig.2.2 Experimental Setup of the qPCR of two regions of the *HoxA10* gene locus…….. 49

Fig.3.1 The two states of Wnt/ β-catenin signalling……..…… 51

Fig.3.2 Non co-linear expression of HoxA genes……..…… 54

Fig.3.3 RNA-seq tracks of signal around the 5' HoxA gene locus …….. 55

Fig.3.4 Homozygous *Mllt3* mutant mouse and wild type mouse……. 56

Fig.3.5 Formaldehyde gel of In situ Probe Synthesis Steps ……. 58

Fig. 3.6 The dissection layout of the *M. schreibersii* RNA-seq dataset…….. 60

Fig.3.7 *M. schreibersii* and *M. natalensis Lef1* expression……. 61

Fig.3.8 *Lef1* expression patterns in the FL and HL autopods of *M. natalensis* embryos ……. 62

Fig.3.9 *M. natalensis* CS 16L head showing similar *Lef1 in situ* staining to E11.5 mouse…….. 63

Fig.3.10 Side-view photographs of *Lef1* WISH conducted on *M. natalensis* embryos……. 64

Fig.3.11 *M. schreibersii* and *M. natalensis Lhx8* expression…… 65

Fig.3.12 Side-view photographs of *Lhx8* WISH conducted on E10.5 and E13.0 mice…… 66

Fig.3.13 *Lhx8* expression patterns in the FL and HL autopods of *M. natalensis*……... 67

Fig.3.14 Side-view photographs of *Lhx8* expression in *M. natalensis* embryos……... 68

Fig.3.15 *M. schreibersii* and *M. natalensis HoxA10* expression…… 69

Fig.3.16 *HoxA10* WISH equivalently staged mouse and *M. natalensis* FL and HL autopods…… 71

Fig.3.17 Side-view of *HoxA10* WISH conducted on *M. natalensis* embryos……..... 72

Fig.3.19 qPCR of a region in close proximity to the 3' region of the *HoxA10* gene locus ……..... 74

Fig.3.20 *M. schreibersii* and *M. natalensis Mllt3* expression……. 75

Fig.3.21 *Mllt3* expression patterns in the FL and HL *M. natalensis* autopods…….. 77

Fig.3.22 Side-view of *Mllt3* WISH conducted on a CS 15 *M. natalensis* embryo…….. 78

Fig.4.1 *M. natalensis* genome browser displaying *Tbx5-as1* relative to *Tbx5*…… 83

Fig.4.2 *M. natalensis* genome browser displaying *Hottip* relative to *HoxA13*…….. 85

Fig.4.3 *In situ* showing expression domains of *Raldh2* in bat and mouse …… 86

Fig.4.4 Amplification and Colony PCR of cloned *Tbx5-as1* isoforms from CS 18L autopods 89

Fig.4.5 *Tbx5-as1* transcript variants identified by RNA-seq in *M. natalensis* …… 90

Fig.4.6 Amplification and Colony PCR of cloned Hottip isoforms from CS 18L autopods…..... 91

Fig.4.7 *Hottip* transcript variants identified by RNA-seq in *M. natalensis* autopods…….. 92

Fig.4.8 Steps involved in generating *Tbx5*, *Tbx5-as1* and *Hottip* probes by *in vitro* transcription…… 93

Fig.4.9 *Tbx5-as1* and *Tbx5* FPKMs in the *M. natalensis* and *M. schreibersii* datasets…….. 94

Fig.4.10 *Tbx5* and *Tbx5-as1* expression in the FL and HL autopods *M. natalensis* embryos…... 95

Fig.4.11 *Hottip* and *HoxA13* FPKMs in the *M. natalensis* and *M. schreibersii* datasets…….. 99

Fig.4.12 *Hottip* and *HoxA13* expression in the FL and HLs of *M. natalensis* embryos……. 100

Fig.4.13 Side-view of *Hottip* CS 15 and CS 16L embryo and *HoxA13* CS 17 embryo……..... 103

List of Tables

Table 2.1 Primer sets for *in situ* probes …… 31

Table 2.2 Primers for Transcript Characterisation of the LncRNA isoforms……. 44

Table 2.3 Primer sets for qPCR on *M. natalensis HoxA10*…….. 46

Table 2.4 Concentration of RNA from each extraction……. 47

Glossary and Abbreviations

AER	Apical ectodermal ridge
A-P	anterior-posterior
BCIP	5-bromo-4-chloro-3-idolyl phosphate
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
cDNA	Complementary DNA
Chiropatagium	The membrane structures of the bat wing that joins the digits together
ChIP-seq	Chromatin immunoprecipitation combined with massively paralleled DNA sequencing, a technique used to identify the binding of DNA-associated proteins.
Cis	A manner of a molecule acting in the same plane as another molecule
Cluster	A group of genes that are located in the same region, are co-expressed and have similar function
CS…	*Carollia stage* – a developmental staging system developed by Cretekos et. Al. (2005) on the short-tailed fruit bat, Carollia perspicillata.
DEPC	Diethylene pyrocarbonate
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
DBD	DNA binding domain
dpc	Days post coitus
DTT	1, 4-Dithiothreitol
D-V	Dorso-ventrally
E…	Embryonic day
EDTA	Ethylenediaminetetraacetic acid
Evo-Devo	Evolutionary Developmental Biology

…E	Early (an early developmental stage)
FC	Fold Change
Fgf	Fibroblast growth factor
FL	Fl
FPKM	Fragment Per Kilobase of exon per Million reads. A strategy of normalizing RNA-seq read counts to be able to compare gene expression between samples.
Fgf8	*Fibroblast growth factor 8*
Fz	Frizzled- transmembrane Wnt receptor
gDNA	Genomic DNA
Heterochrony	The change in timing of a developmental event, whether it be the onset, offset or tempo.
Heterotropy	The change in expression spatial domains of toolkit genes during development.
	Hl
Homeodomain	A domain in a protein that is encoded for by a Homeobox, that consists of about 60 amino acid residues which are usually similar from one such domain to another, and that recognizes and binds to specific DNA sequences in genes regulated by the homeotic gene
Hottip	A LncRNA known as: HOXA transcript at the distal tip
HoxA10	*Homeobox A10*
HoxA11	*Homeobox A11*
HoxA13	*Homeobox A13*
HP	Hand plate
IPA	Ingenuity pathway analysis
IVT	*In vitro* transcription
IPCD	Inter-digital Programmed Cell Death
IPTG	Isopropyl-beta-D-thiogalactopyranoside
KO	Knock-out
LAIX plates	Luria-Bertani agar plates supplemented with ampicillin (100µg/ml), IPTG (0.5mM) and X-Gal (80µg/ml)

LB	Luria-Bertani broth
Lef1	*Lymphoid enhancer factor-1*
Lhx8	*Lim Homeobox 8*
LincRNA	Long intergenic non-coding RNA
LncRNA	Long non-coding RNA
LOF	Loss-of-function study
LPM	Lateral plate mesoderm
Maternity roosts	Warm, humid roosts that are used by colonies of bats as shelter during their reproductive periods. Bats will travel long distances to inhabit these roosts.
Meis2	*Meis homeobox 2*
µg	Micrograms
MGE	Medial Ganglionic Eminence
MLL	Mixed Lineage Leukemia
Mllt3	*Mixed lineage leukemia translocated-to-3:* a transcription factor that is a common fusion partner of the mixed lineage leukemia complex (MLL)
Morphogen	Graded signalling molecules that emanate from a restricted area during development. The concentration of the molecule at a point gives the cell positional information and prefigures the developing structure.
mRNA	Messenger ribonucleic acid
NCBI	National Centre for Biotechnology Information
oligo (dT) primer	A single strand of poly T bases that contain a T7 promoter
oligonucleotides	Short sections of DNA with a sequence that corresponds to that of a specific gene
PBS	Phosphate buffered Saline
PBST	Phosphate buffered saline containing 1% Tween 20
PCR	Polymerase chain reaction
PCP	Wnt planar cell polarity pathway
P-D	Proximal-distal

PFA	Paraformaldehyde
Plagiopatagium	The largest membrane structure of the bat wing. It joins the hind and the forelimb and is extended along the outer edge of the fifth digit
Pleiotropy	When one gene influences two or more seemingly unrelated phenotypic traits. Consequently, a mutation in a pleiotropic gene may have an effect on some or all traits simultaneously
pmol/µl	Picomols per microlitre
polyA	Polyadenosine
Primordium	An organ or tissue in its earliest recognizable stage of development
Protopatagium	The small membrane structure of the bat wing that extends from the shoulder of the bat to its wrist
PZ	Progress Zone
RA	Retinoic acid
Raldh2	*Retinaldehyde Dehydrogenase two*
RNA	Ribonucleic acid
RNAP2	RNA Polymerase II
RNA-seq	Massively paralleled mRNA sequencing
RNP	Ribonucleoprotein complexes
rRNA	Ribosomal RNA
q-PCR	Real time polymerase chain reaction
SDS	Sodium dodecyl sulphate
Shh	*Sonic Hedgehog*
SSC	Sodium chloride sodium citrate
TBST	Tris-buffered Saline Solution plus Tween20
Tbx5	TATA-box 5
Tbx5-as1	A LncRNA in close proximity to *Tbx5*, that is transcribed in the antisense direction
Trx	*Trithorax*
TSS	Transcription Start Site
UCSC	A site developed and maintained by the University of California Santa Cruz, which contains the draft assemblies for a large collection of

	genomes.
UTR	Untranslated region
Uropatagium	The membrane structure of the bat that stretches between the tail and the hl skeletal elements
WDR5	An adaptor protein: WD repeat-containing protein 5
WISH	Wholemount *in situ* hybridisation
WT	Wild-type
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
ZPA	Zone of polarizing activity

Chapter 1

The Molecular Mechanisms of Bat Wing Development

1.1 Introduction

For 60 years, mammalian limb development has been a well-studied paradigm for understanding the fundamental processes that convert embryonic cells into the countless morphologies seen in nature (Zuniga 2015). Each homologous mammalian limb type is adapted for a unique set of functions. The adaptation to life on land known as the 'fin-to-limb' transition is one of the most remarkable evolutionary steps undertaken by mammals, requiring a whole new set of functions to adapt to the terrestrial habitat (Pyron 2011; Unsicker & Krieglstein 2006). This transition resulted in the mammalian hand/foot (autopod) becoming particularly divergent, with skin types and digit identities varying drastically allowing for specialisations such as gripping trees and running in fields, digging and flying. This divergence is depicted in Figure 1.1 with the first two sets of bones (white and yellow) being similar in shape and number, whereas the bones that make up the autopod (green) differ dramatically. These differences originate from the same base, as can be observed in the gross morphology in early mammalian development, with highly divergent autopod morphologies arising as the embryo approaches its adult form (Richardson 1999). Limb patterning during development is a rapid event caused by a common set of genes expressed in all mammals. Little is known about how the pattern of expression of these genes diverged to generate the observed morphological differences between mammalian limbs.

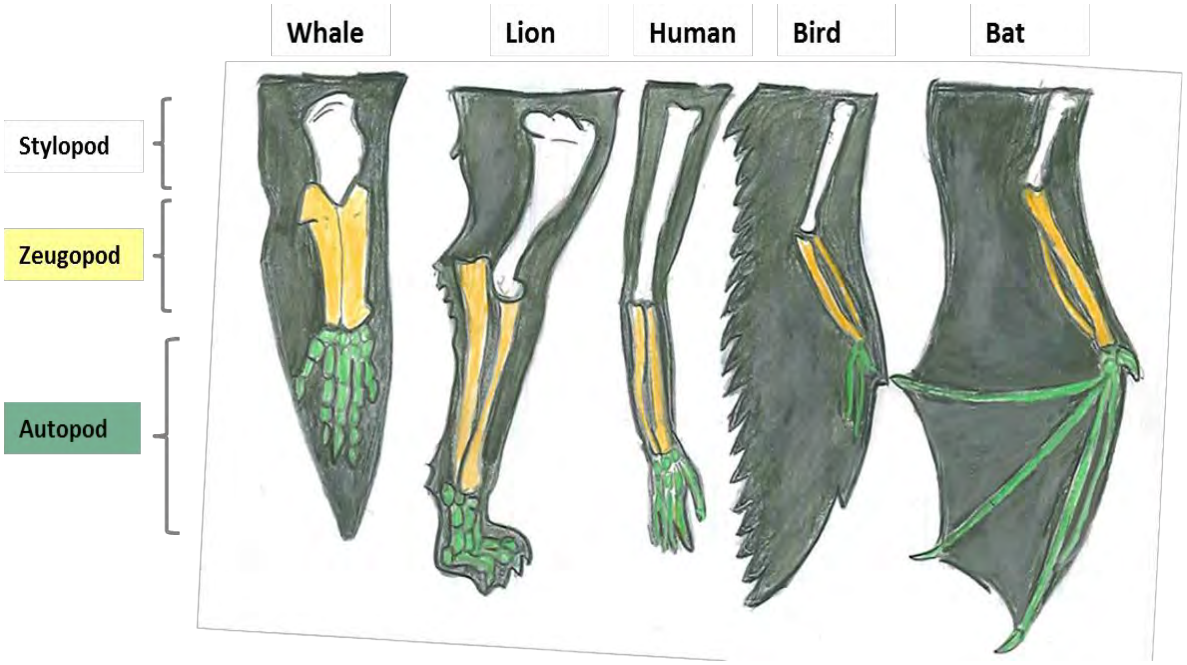


Figure 1.1 Variation in morphology of the three segments of mammalian limbs. The most proximal segment (shoulder and humerus) is known as the 'stylopod' (white), the middle segment (ulna and radius) is known as the 'zeugopod' (yellow) and the distal-most segment (hands and feet) is known as the 'autopod' (green).

An example of exaggerated mammalian autopod morphology is that of the wing of a bat. Bats are the only mammals capable of powered flight and underwent adaptive radiation approximately ~51 Mya to result in a highly exaggerated forelimb (FL) autopod. Several physiological and morphological changes were required for bats to acquire flight, such as dramatically elongated FL digits and retained interdigital tissue. The FLs are in stark contrast to the hindlimbs (HL) autopods made up of uniform, short and free digits (Liang et al. 2013). This contrast between the FL and HL morphology makes for a great internal control, and a great model for morphogenesis studies (Lancaster & Speakman 2001). Even more intriguing is that the same genes and genetic networks shape all mammalian autopods, therefore by understanding how bat wings develop, we can gain a greater understanding of mammalian hand formation (Gunnell & Simmons 2012).

1.2 Vertebrate Limb Morphology

Adult vertebrate limbs are made up of three basic parts: the most proximal segment known as the ‘stylopod’ (shoulder and humerus) (Figure 1.1; White segment); the middle segment known as the ‘zeugopod’ (ulna and radius) (Figure 1.1; Yellow segment); and the distal-most segment known as the autopod (Figure. 1.1; Green segment). Autopods are further divided into three sets of small bones; carpals (Figure1.2, A) metacarpals (Figure 1.2, B) and phalanges (Figure 1.2 C). Phalanges are typically constructed from three sets of small bones: Proximal (Figure 1.2. Orange); intermediate (Figure 1.2. Blue) and distal phalanges (Figure 1.2, Yellow) and are collectively referred to as ‘digits’.

Figure 1.2. Human hand as an example of the basic plan of the tetrapod autopod. A) Carpals (green) make up the wrist area and base of the palm and vary considerably between taxa B) Metacarpals (red) make up the palm of the hand. C) Digits are made up of three sets of phalanges: Proximal (orange), Intermediate (blue) and Distal (yellow).

Limbs develop in two stages; the first stage involves the outgrowth of small protrusions from the trunk known as ‘limb buds’. The second phase involves the patterning of mesenchymal progenitors into their future skeletal elements. Condensation and elongation of digit primordia is accompanied by interdigital programmed cell death (IPCD), whereby mesoderm between the condensing digits undergoes apoptosis (Towers & Tickle 2009; Kaufman, M. H., & Kaufman 1992).

1.3 Evolutionary Developmental Biology

Most knowledge about mammalian developmental comes from studying a single group of mammals; the mouse (*Mus Musculus*; order *Rodentia*). This is due to their short generation time, large number of progeny and a long history of genetics making them a good model organism for genetic studies. As a result, much is known about mouse development, but very little else is known about how the ~25 other mammalian orders develop. A relatively new field of study known as evolutionary developmental biology (informally: “Evo-Devo”) aims to study developmental processes of model organisms such as mice, in conjunction with non-model organisms, to compare differences during development (Heffer & Pick 2013). The field combines fossil record, embryology and genetic techniques, to study the molecular developmental processes that underpin innovation and diversification of morphology among taxa (Carroll 2008). The application of these techniques to a diverse set of taxa has shed light on the molecular events of mammalian body patterning, however; questions remain about the specific mechanisms that give rise to morphological novelty in evolution.

It was once thought that differences in amino acid sequences between homologous proteins drove the evolution of morphological novelty (King et al. 2007). This notion was challenged when homologous proteins in chimpanzees and humans were compared, to examine whether amino acid changes can account for distinct differences in anatomy. The findings were paradoxical, as there were only slight changes in protein sequence to result in the drastic phenotypic differences between chimps and humans. When homologous proteins from evolutionarily distant taxa were substituted *in vivo*, normal function remained. These data indicates that there is little change in the biochemical properties of homologous proteins and the ability to interact with other co-factors (Carroll 2008). Following these discoveries, it was proposed that instead of differences in protein sequence, differential expression of genes coding for these conserved proteins drives morphological variation (King et al. 2007). Despite mRNA being a precursor for protein, until recently researchers believed that mRNA transcript abundance was a poor predictor of protein levels. This was based on system-wide

studies that showed RNA translation rates to be the best predictor of protein levels, with transcription accounting for a mere 34% of protein levels (Greenbaum et al. 2003). A recent study described the four processes that typically affect protein levels namely, transcription, mRNA degradation, translation and protein degradation. These authors showed that large error due to bias in generating and analysing the mRNA and protein data was misleading and that transcription was by far the most substantial (73%) predictor of protein levels (Affairs et al. 2015).

There are a set of transcription factors and cell-signalling proteins known as 'toolkit' genes that regulate patterning and development across taxa. Coding sequence of toolkit genes is well conserved between taxa, as changes in protein sequence result is a multitude of negative and possibly fatal downstream affects (Carroll 2008). Regulation of toolkit genes is controlled by complex promoters with multiple transcription factor binding sites. Alterations to transcription of toolkit genes results in a phenomenon known as 'Pleiotropy', a biological phenomenon whereby a single gene locus affects two or more seemingly unrelated phenotypic traits (Stearns 2010; Hodgkin 1998). An example pleiotropy in nature would be the human disease Phenylketonuria, in which mutations to a single gene phenylalanine hydroxylase can cause mental retardation and reduced hair and skin pigmentation (Filiano, 2006). While the coding regions of toolkit genes are conserved, the regulatory regions are much more flexible and assimilate mutations that lead to novel patterns of expression. Modifications to the *spatial* regulation of toolkit genes is known as heterotropy (Hall 2003); while changes in the *timing* of expression of toolkit genes known as heterochrony (Carroll 2008). Alterations to timing and spatial expression of toolkit genes allow 'old genes to perform new tricks'.

1.4 The Molecular Mechanisms that Shape Vertebrate Limbs

Limb pattering occurs along three axes: the dorso-ventral (DV; top of the hand to palm), proximo-distal (PD; shoulder to fingers) and the anterior-posterior (AP; digit I to digit V). Limb development in these three axes is governed by molecular signals emitted from conserved signalling centres (Figure1.3).

Limbs begin as buds that arise as bulges of mesenchymal cells from the flank lateral plate mesoderm (LPM), encased in a layer of ectodermal cells known as the Apical ectodermal Ridge (AER) (Figure 1.3)(Rabinowitz & Vokes 2012). The AER acts as a signalling centre to initiate limb growth. The current model for limb proximo-distal outgrowth is known as the 'Two-signal model', based largely on chick studies.

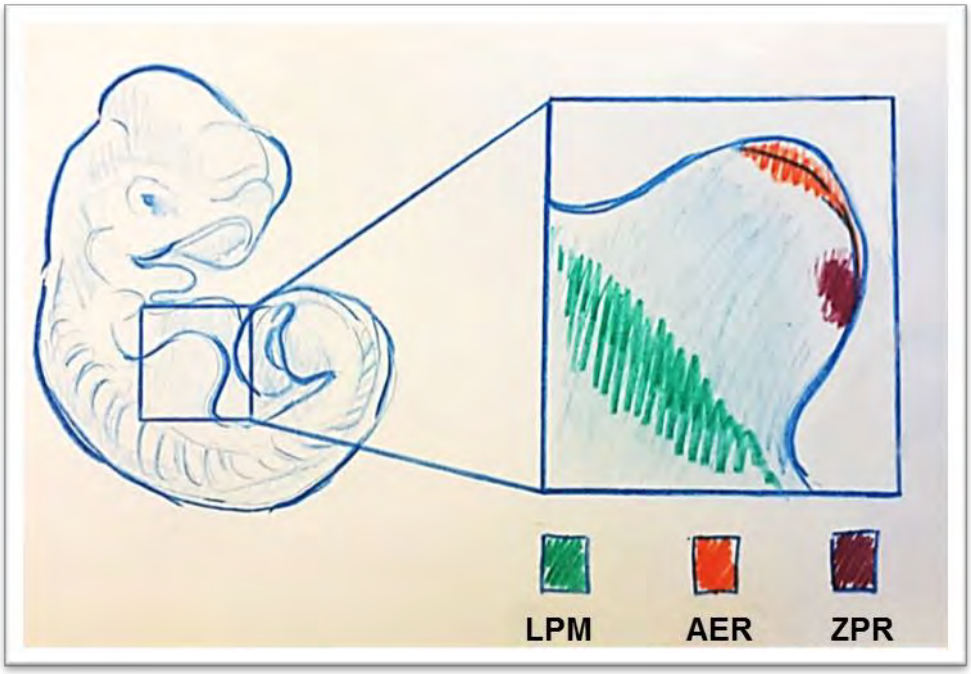


Figure 1.3 Mouse limb-bud during early development with signalling centres domains indicated. The Apical Ectodermal Ridge (AER) is the thin layer of ectodermal cells on the distal-most layer of the limb bud. The zone of polarizing activity (ZPR) is the region just below the AER. The Zone of Polarising Activity (ZPA) is in the posterior-most region of the limb bud.

1.4.2 Two-signal Model for PD Limb Patterning

Segmental patterning of the developing limb along the proximo-distal (PD) axis is guided by diffusible signalling molecules known as 'morphogens'. Morphogens give cells positional information within a limb-bud as they are secreted in a concentration gradient, and cells receiving the morphogen respond differently at different threshold concentrations. The two-signal model proposes that PD patterning occurs via opposing gradients of proximally secreted morphogen Retinoic Acid (RA) and distal ly secreted morphogen Fibroblast Growth Factor (FGF) (Cooper et al. 2011; Rosello-Diez et al.

2014). In the proximal region of the limb-bud LPM produces an enzyme, *retinaldehyde dehydrogenase 2* (*Raldh2*) that synthesizes RA from retinol (Dolle et al, 2012). In the distal end of the limb-bud FGF signalling activates *Cyp26b1* which oxidizes RA, causing its degradation and removal (Yashiro et al. 2004). The site-specific expression of RA synthesis and degradation creates a graded level of RA concentration with the highest level in the proximal region and lowest in the distal region (Tabin & Wolpert 2007). The positional information given by the levels of RA and FGF activates or represses expression of downstream homeobox transcription factors: *Meis1/2* (stylopod), *HoxA11* (zeugopod) and *HoxA13* (autopod)(Tabin & Wolpert 2007).

During the early stages of limb development RA activates *Meis1*/*2* expression throughout the bud, after which *Meis1/2* expression is successively downregulated in the distal end of the bud initiating *HoxA11* expression in the distal end of the limb bud. Expression of *HoxA13* is then initiated in a posterior distal region of the limb, whilst *HoxA11* expression is repressed in the distal end of the limb-bud (Mercader et al. 2009). Recent studies have questioned whether RA plays an instructive role in limb patterning and induction of *Meis1/2* expression in the proximal limb (Cunningham & Duester 2015; Zhao et al. 2009; Cunningham et al. 2011). *Meis2* is expressed in the autopod prior to interdigital tissue regression and emergence of the digits, and RA deficient embryos have retained interdigital webbing. However been previously shown that RA is responsible for limb induction, but loss-of-function studies show on patterning (Zhao et al. 2009).

1.4.3 Patterning along AP Axis

Patterning in the Anterior-Posterior (AP) axis is mediated by a group of cells located in the posterior and distal tip of the developing limb bud known as the zone of polarizing activity (ZPA) (Figure 1.3). The ZPA emits a concentration gradient of *Sonic Hedgehog* (*Shh*) along the AP axis, with the highest concentration in the ZPA and the lowest in the thumb region of the limb bud. Inactivation of *Shh* in mouse resulted in the loss of all digits except for the thumb (Harfe 2011). *Shh* causes the upregulation of mesenchymal genes *HoxA* and *HoxD* genes, which are responsible for specifying digit number and

identity (Sheth et. al, 2012). Although the separate axes were traditionally studied separately, recent studies have shown that they are tightly interconnected. Expression of the enzyme *Cyp26b1* is activated by the interplay between the AP and PD axes feedback loop: SHH-GREM1-AER-FGF (Probst et al. 2011).

1.4.3 Revised Model for Explaining Limb Development

The previous 'progress zone' model for limb patterning, postulated that the duration of exposure to the signalling molecule determine the cells fate. Recent studies have shown that neither models can fully explain the process (Rosello-Diez et al. 2014). Systems biology techniques were used to model limb bud outgrowth and patterning and found that both timing and strength of signal are essential in the patterning of the limb (Uzkudun et al. 2015). This revised model is known as the 'cross-over model', and calls for the inclusion of intrinsic timers, in addition to external signals (Zuniga 2015).

1.5 The Unique Morphology of the Bat Wing

Bats are part of the mammalian order Chiroptera, a word that directly translates to "hand-wing". Bats underwent vast anatomical diversification from their arboreal ancestors allowing them to invade a unique nocturnal niche (Miller-Butterworth et al. 2007). This ecological transition was possible with the innovation of flight, largely as a result of skeletal modification of its FL autopods (Gunnell & Simmons 2005). There is no fossil record documenting the transition from paw-to-wing (Gunnell & Simmons 2005; Lancaster & Speakman 2001), hence, to overcome this scarcity of data, scientists have studied the morphological and molecular events that occur as bats develop from embryos to adults (Eiting & Gunnell 2009; Weatherbee et al. 2006; Sears et al. 2006; Hockman et al. 2008; Mason et al. 2015; Cretekos et al. 2005; Chen et al. 2005).

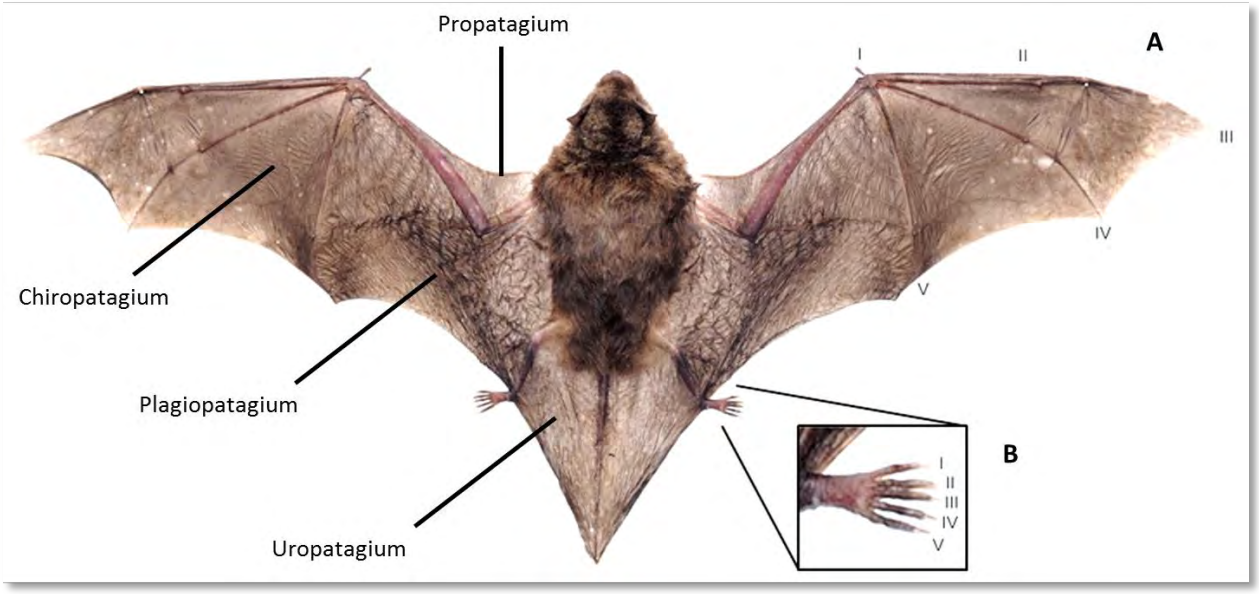


Figure 1.4 Anatomy of an adult bat *Miniopterus natalensis*. A) An outstretched left FL autopod, with a truncated digit I, and dramatically elongated digits II –V. B) A close-up of the HL autopod with its uniform digits indicated. Forelimb digit I and HL digits are similar in morphology and size.

Bat limbs begin much like mice, as buds that progress into paddle-shaped appendages. It is only in their late embryonic phase that their morphology diverges from mouse, and later becomes dramatically divergent from other mammals. The digits of adult bat FL autopods are dramatically elongated, with digits II-V being considerably longer than digit I (Figure 1.4 A). Despite the elongation of the digits, there is a reduction in the number of phalanges in each digit, with digit II having a single phalange, digit III having three phalanges and digits IV and V having two phalanges (Figure 1.4 A). The zeugopod has a low-bone-mass phenotype, created by enlarged cavities in the bone and reduced cortical bone width (Cooper et al. 2012). Across the elongated FL digits stretches an expanded and thinned membrane (Chiropatagium) that extends from the FL to the HLs (Plagiopatagium) and from the HL zeugopod to the tail (Uropatagium) (Figure 1.4). This membrane is collectively called a ‘patagia’ and is highly adapted for aerial manoeuvres with an extensive network of collagen fibres, microscopic muscles, and an abundance of cutaneous receptors (Sterbing-D’Angelo et al. 2011). Bat HL autopods have uniform AP patterning, and are made up of equally sized, short and free digits (Figure 1.4, B)(Chen et al, 2005). This adaptation aids in a firm grip while perching upside down.

1.6 Alterations to AP patterning that Shape Bat Limbs

A recent finding in bat autopod development, is an expanded distal region of *Meis2* expression in the developing bat FL compared to mice (Mason et al. 2015). *Meis2* is a marker for the stylopod and is responsible for specifying proximal identity of the limb. In chick studies, ectopic *Meis1/2* expression in distal regions of the limb bud, result in an expansion of proximal identity, such as a reduction in distal digit length and the persistence of interdigital webbing (Mercader et al. 1999). Overexpression of *Meis1/2* also caused a delayed expression of *HoxA13*, and a prolonged expression of *HoxA11* (Rosello-Diez et al. 2014). From mouse studies, it was discovered that interdigital *Meis2* expression is RA-independent, and thus, retention of interdigital webbing in bat wings is not due to the suppression of RA-induced cell death. It is possible that RA signalling plays a role in the thinning of the interdigital tissue in bat FL.

Although certain candidate genes have been explored in bat development, the full complexity of the mechanism of how this dramatic adaptation evolved remains to be shown. Seeing as gene expression has been shown to be a good proxy for protein levels (Affairs et al. 2015), a global expression study can provide key insight to the mechanisms at play.

1.7 Strategies for Exploring Gene expression in Bat Limbs

There are two general strategies for exploring gene expression in limb development (Mallarino & Abzhanov 2012). A bottom-up or 'Candidate Gene' approach which uses pre-existing knowledge of genes known to play a role in the system of interest (limb development)(Chen et al., 2005; Cretekos et al., 2008;Hockman et al., 2008). The second strategy is a top-down or 'Genome-Wide' approach that employs global gene expression patterns in the tissue of interest to inform gene selection (Hockman et al. 2009). A 'Genome-wide' approach was applied to this project using massively parallel RNA sequencing (RNA-seq) which allowed for the analysis of entire transcriptomes by generating a snapshot of transcript levels in limb tissue. This was done by high through-put sequencing of cDNA of the natal long-fingered bat *Miniopterus natalensis* (Eckalbar et. al., manuscript under review).

The specific developmental stages (Carollia Staging: CS) for the study were selected at the point of development when the bat FL diverges from the HL, and the FL of equivalently staged mice. Previous work described by Hockman et al. (2009) identified CS 15 as the stage when the FL autopod enlarges distal ly in the AP axis, becoming asymmetrical compared to the HL, which remains symmetrical and rounded. At the subsequent stage CS 16 the FL diverges noticeably from the HL, and at stage CS 17 the morphology of the FL is dramatically altered from the HL and mouse autopods at equivalent stages. These rapid changes in FL shape can be attributed to digit elongation, establishing the hand-wing prior to digit ossification. In the membrane, there is a novel interplay between apoptosis that is thinning the membrane while preventing IPCD and retraction of the tissue (as it does in the HL and mouse autopods) (Mason et al. 2015).

Six transcriptomes (CS 15 FL and HL; CS 16 FL and HL; CS 17 FL and HL) were annotated using an assembled *M. natalensis* genome (Stephen Schlebush). The gene counts for each transcriptome were normalised to fragment per kilobase of exon per million reads (FPKM), allowing for comparison between datasets. Expression between FL and HL was compared using DE-seq, which identified 2952 genes as being differentially expressed (adjusted p-value ≤ 0.01). This discovery cannot give insight as to whether the differentially expressed genes are involved in digit elongation or reduced IPCD. In order for this to be elucidated, the spatial expression of individual genes must be explored. However; this number of genes is too large for each gene to be individually investigated; thus, the list was narrowed down by literature, to select genes that are known to regulate development (Eckalbar, manuscript in preparation).

An RNA-seq dataset was recently published, from a closely related bat species *Miniopterus schreibersii* (Wang et al. 2014). The sampling of this transcriptome was complementary to the *M. natalensis* transcriptome, as the same developmental stages were selected, however the tissue was divided up into spatial regions and CS 15- CS17 tissue was pooled (Figure 1.5). These data provided a unique opportunity to validate our RNA-seq data, as well as gain insight into the spatial expression of candidate genes.

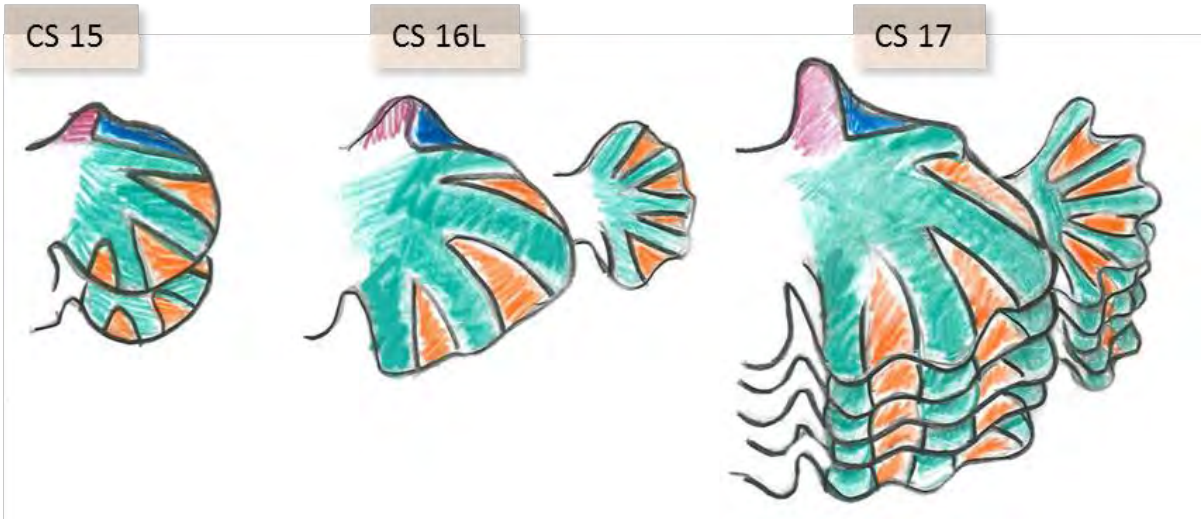


Figure 1.5 Tissue sectioning used to generate the *M. schreibersii* RNA-seq dataset. Schematic displaying the size and relative contribution of each tissue division from each developmental stage that was pooled. The tissue bias in the RNA-seq sample as CS17 makes up the majority of the tissue. The different colours represent the different domains that each autopod was divided into: maroon: FL digit I to represent the truncated region,. Navy: inter-digital tissue between digit I and II. Green: digit tissue from FL digits II-V and HL digits. Orange: inter-digital tissue between digits II and V.

1.8 LncRNAs

A particular class of non-coding RNAs known as long non-coding RNAs (LncRNAs) biochemically resemble mRNA with polyadenylated tails, and introns, but lack intact stop/start codons (Hu et al. 2015). This type of non-coding RNA has been found to be involved in regulating gene expression during development, controlling factors such as temporal and spatial expression of genes (Knosp et al, 2004).

The various types of LncRNAs that have been discovered have been defined by their position and transcriptional orientation relative to their neighbouring protein-coding genes (Figure 1.6) (Cabili et al. 2011). 'Antisense' LncRNAs initiate 3' or 5' of, or within a protein coding gene, are transcribed in the opposite direction to the gene, and overlap at least one exon of the gene (Figure 1.6 A). The *M. natalensis* RNA-seq dataset was generated using stranded reads in order to identify transcripts being transcribed in the antisense direction, and thus identify antisense LncRNAs (Eckalbar et. al., manuscript under review).

Another means of classifying LncRNAs, is by the manner in which they influence the regulation of genes. LncRNAs that regulate neighbouring genes are known as '*cis*-acting' LncRNAs, whereas LncRNAs that regulate distant genes on other chromosomes are known as '*trans*-acting' LncRNAs (Rinn & Chang 2012). A well-defined antisense *cis*-acting LncRNAs *Hottip* (HoxA transcribed at distal tip) as well as the less well-defined antisense cis-acting '*Tbx5-as1' (*transcribed from the 3' end of *Tbx5),* were one of many LncRNAs identified by RNA-seq to be differentially expressed between FL and HL. They were selected for further investigation by *in situ* hybridisation to uncover their spatial expression.

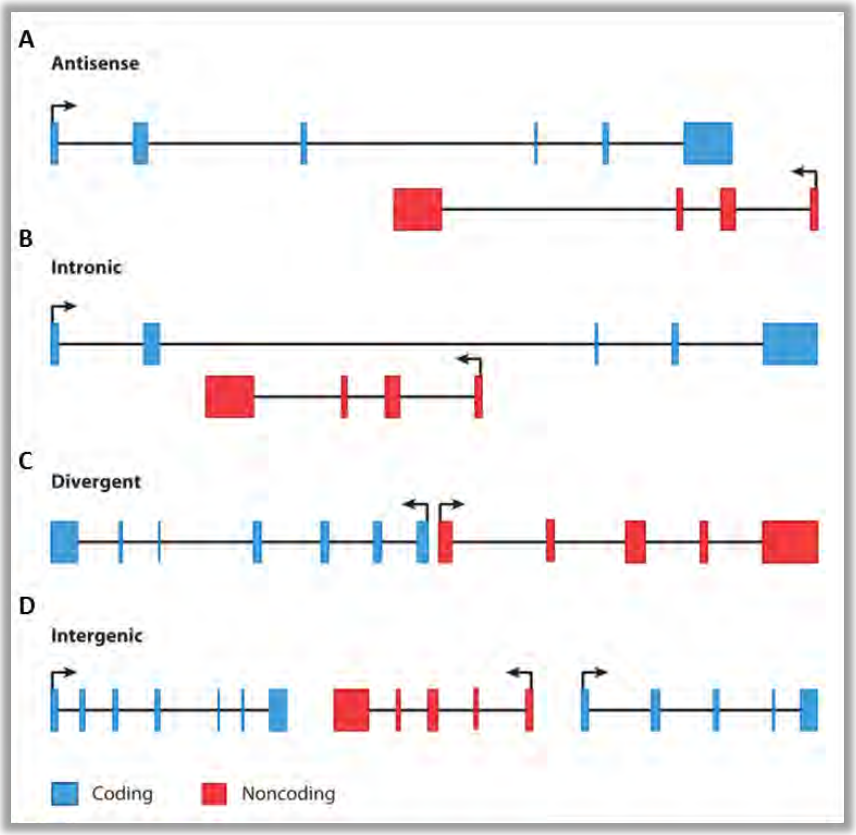


Figure 1.6 The different types of LncRNAs. A) **'**Antisense' LncRNAs are transcribed in the antisense direction, on the the 3' end of their neighbouring protein-coding genes. B) 'Intronic' LncRNAs initiate and terminate in the introns of protein coding genes, and can be transcribed in either direction. C) 'Divergent' LncRNAs initiate from the promoter of a protein coding gene. D) 'Intergenic' LncRNAs (LincRNAs) are LncRNAs with independent transcriptional units to their neighbouring protein-coding genes (image from Rinn & Chang 2012).

1.9 Study Aims

The broad aim of this study was to test whether differences in the expression of key development genes have played an important role in generating the morphological diversity of mammalian limbs. The first aim of the project was to validate the RNA–seq results from the *M. natalensis* transcriptomes (Eckalbar et. al., manuscript under review) using the *M. schreibersii* dataset (Wang et al. 2014). This was done by extracting and comparing the read counts between our dataset and the published data. Another element to the RNA-seq validation was to investigate a rapidly fluctuating RNA-seq signal in the 3' *HoxA10* gene locus using qPCR.

The aim of the second part of this study was to characterise the spatial and temporal expression patterns of selected transcription factors and LncRNAs. This was done by conducting whole-mount *in situ* hybridisation (WISH), with antisense DIG-labelled probes, of five transcription factors: *Lef1*, *Lhx8*, *HoxA10*, *Mllt3* and *Tbx5,* and two LncRNAs *Hottip* and *Tbx5-as1.* WISH was conducted on bats within the range of the RNA-seq dataset. Due to limited numbers of embryos that fit precisely into a developmental stage, some embryos were either older or younger than the RNA-seq stages.

The last element of this project was to validate the LncRNAs that were selected for WISH, to ensure that they are not artefacts of the RNA-seq. There is a scarcity of literature on LncRNAs, however; it has been established that LncRNAs can arise as spurious transcripts in cDNA syntheses and sequencing. For this reason, a subset of isoforms of the LncRNAs were cloned from CS 18L cDNA, and sequenced to confirm their identity against the RNA-seq transcript isoforms.

Chapter 2

Materials and Methods

2.1 Study Species: *Miniopterus natalensis*

A small insectivorous bat, *M. natalensis,* was selected as the study species for this project The genus *Miniopterus* (Bonaparte,1937) are often referred to as 'bent-wing' or 'long-fingered' bats as they have highly elongated third digit that folds back on itself when the wings are closed (Peter John Taylor 2000). The population was sampled under Western Cape Nature Conservation Board permit number: AAA007000410056. The pregnant bats were caught as they emerged from a maternity roosting cave known as the Guano Cave (34°26'S; 20°25'E) that hosts a large population of bats from different areas of South Africa during the breeding season (McDonald et al., 1990b). Ethical approval by the University of Cape Town for the use of these embryos was granted by the UCT Animal Ethics Approval committee, code: 2012V39NI *Evolution of the development of the bat wing.*

Bats were captured during the evening emergence from the maternal nesting cave using a Harp Trap (Austbat 3-bank, Faunatech, Mount Taylor, Victoria, Australia). The trap was placed near but out of the way of the main stream of bats, as not to hinder emergence. Captured bats were studied and only pregnant *M. natalensis* bats were selected and each bat was placed in its own black cloth bag for transport. The bats were transported in silence to the research house where they were hung up in a ventilated but dark cupboard for the night. The selected bats were euthanized with halothane – (Safe Line Pharmaceuticals Pty. Ltd.) which was administered to individual bats in a small sealed container and absorbed by inhalation. The bats were prevented from coming into direct contact with the halothane as it was dripped into cotton wool surrounded by tin foil. After 5 minutes, the bats were removed from the container and underwent a cervical dislocation, as an added measure, to ensure there were dead before dissection.

2.2 *M. natalensis* Dissection and Embryonic Staging

Dissection was carried out according to procedures outlined by Mandy Mason (pers. Comm.). Following cervical dislocation, the adult mother bat wings were pinned open in three places. The hair on the abdomen was wet with ethanol (70%) and smoothed over to prevent hair from interfering and attaching to the dissecting implements. Abdomen hair and flesh was lifted with forceps and an incision of 1cm was made above the pubic bone using a pair of medium dissecting scissors. Once there was a clear incision through the abdomen flesh and peritoneum, an upside down Y shaped incision was made from this point down to the pubic bone exposing the amniotic sack containing the embryo. Fresh sterile microdissection scissors and forceps were used to cut through the connective tissue binding the sack to the cervix, pulling the sack freely from the abdomen. Care was taken to not collect hair with the sack as hair contains RNAses that can degrade the RNA in the embryo once it has been removed from the amniotic sack. This sack was placed in a petri dish filled with chilled PBS.

Developmental stages were determined by a species specific staging system developed by Hockman et al (2009) in accordance with a previous bat and mouse equivalent staging system by Cretekos for *C. perspicillata* (Cretekos et al. 2005). Embryos were staged using ear, eye and limb developmental progression to classify the embryos into each stage. Embryos were staged in PBS using a stereomicroscope and those embryos that were to be used for RNA extraction were beheaded, eviscerated and placed in RNAlater® according to the manufacturer's instructions (QIAGEN). Embryos in RNAlater® were stored on ice for 24 hours and then stored at -20˚C. The embryos used for WISH were fixed in 4% Paraformaldehyde (PFA) for 24 hours at 4˚C, and then taken through a methanol dehydration series of washes (25%, 50%, 75% and 100% methanol). Once in 100% methanol the embryos were placed on ice for the trip back to Cape Town (4 hours) and were then stored at 20˚C.

2.1.2 Acquisition of Mouse Embryos

Mouse embryos were obtained from pregnant C57BL/6 mice from the Animal Unit at the University of Cape Town Medical School. Ethical approval was granted to the supervisor of this project, Professor Nicola Illing by the Faculty of Health Sciences Animal Research Committee (HSC 012/052 and HSC 014/07). Noon on the day a vaginal plug was observed, was taken to be E0.5. Mice were sacrificed at midday to obtain embryos E10.5, E11.5, E12.5 and E13.5. Pregnant mice were euthanized by intraperitoneal injection of sodium pentobarbitone (400mg/kg (diluted 1:2 in sterile saline) followed by cervical dislocation. The reproductive tract was dissected out via a midline laparotomy, and uteri were placed in 4˚C phosphate buffered solution (PBS) in a 50ml Sterilin tube (Thermo Scientific). Embryos were removed from the amniotic sacks in pre-chilled 4˚C PBS in a sterile petri dish.

Embryonic stages were determined by comparing the progress of limb development to the staging systems provided by Martin (1990) and Kaufman (1992). Embryos used for RNA extractions were placed in RNAlater® and stored at -20˚C, while embryos for WISH were placed in chilled 4% PFA overnight at 4˚C. Following fixation, embryos were taken through a methanol series (25%; 50%; 75%; 100%) and stored at -20˚C.

2.2 Whole-mount *in situ* Hybridization

2.2.1 Primer Design

In situ probes were designed to the *M. natalensis* assembled transcriptome (Eckalbar et. al., manuscript under review), and were informed by the RNA-seq read counts and exon information. Where possible, probes were designed over regions with the highest reads aligning in the RNA-seq dataset, and areas with splice variants were avoided. Primers (Table 2.1) were designed using Integrated DNA Technology website (IDT) and were analysed using the OligoAnalyzer® Tool (SciTools®) to determine the primer set with the best homodimer, heterodimer and hairpin values, as well as Tms closest to 60°C. The optimal size of a probe ranged from 500 to 1000bp as *in situ* probes need to be long enough for Digoxigenin (DIG) labelling to intercalate between the mRNA to yield sufficient signal, but not too long as they will not penetrate the flesh efficiently. Exceptions were made to this rule in certain cases, as the transcript variants did not have sufficient overlap to allow for such a large probe. Each probe came with its own set of specifications and rationale detailed below:

Transcription Factor Primers

Bat HoxA10- primers were designed over the 3' UTR, as all *Hox* genes contain a highly conserved region known as a Homeodomain encoded for by a homeobox consisting of 60 amino acid residues which are similar from one domain to another. This domain was avoided to prevent non-specific binding of the *in situ* probe to any of the 39 *Hox* genes.

Bat Lhx8 - The *Lhx8* probe was designed to avoid the conserved LIM domain. Primers were designed in order to bind to both bat and mouse regions of the gene.

Bat Mllt3- Primers were designed that matched both mouse and bat sequence and were designed over an area with very little transcript variance in *M. natalensis*.

Bat Lef1 – The *M. natalensis* gene had 27 transcript variants annotated in the transcriptome. Primers were designed over the 5' region of the gene that

included all transcript variants in order to gain the greatest mRNA *in situ* signal. The bat primers matched the mouse sequence, and were used to amplify the mouse probe template.

Bat Tbx5 - A short probe was designed over the 3' region of the *M. natalensis* sequence**,** over a region of high signal and overlap.

LncRNAs Primers

Bat *Tbx5-as1* - Human, bat and mouse *Tbx5-as1* transcripts were mapped and an area of overlap between human and the 3' region of the M. natalensis Tbx5-as1 was chosen to design primers. This region overlapped 10 of the 13 transcript variants as well as the human *Tbx5-as1* (Figure 2.1 'IS probe').

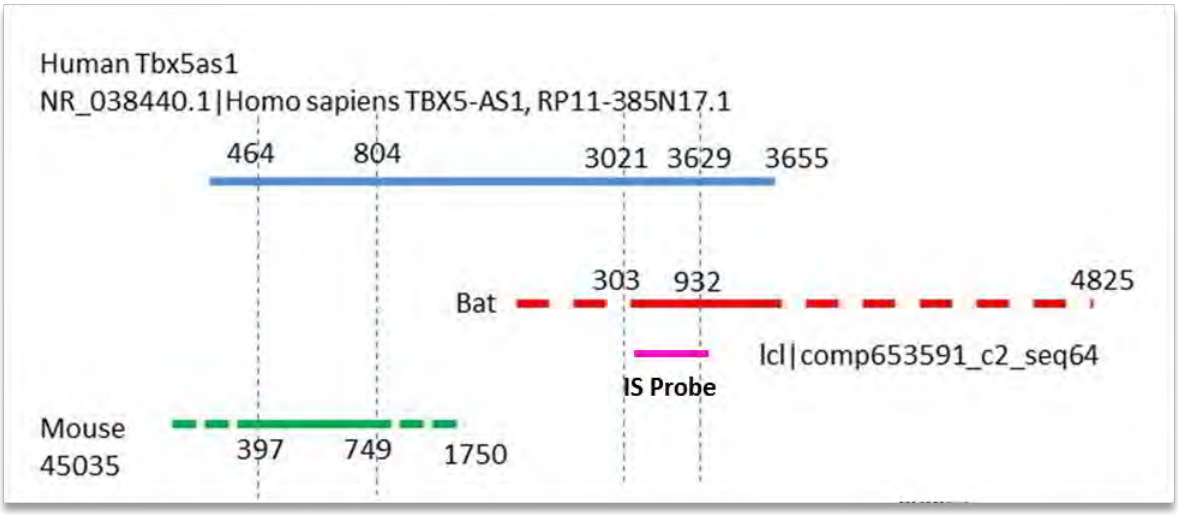


Figure 2.1 Position of the *M. natalensis* transcript 2 (red) relative to the mouse (green) and human (blue) *Tbx5-as1.* The *in situ* probe sits on the 3' end of the *M. natalensis Tbx5-as1* transcript 2, in a region of overlap with the human *Tbx5-as1*.

Mouse Tbx5-as1 - Several transcript variants were identified that mapped in close proximity to the *Tbx5* locus. All of these reads had low associated FPKM counts with the majority of the coverage coming from the 1st exon. Probes were designed to this region on the minus strand.

Bat *Hottip*- probes were designed over the most conserved region which lies in the 5' region of the gene and is present in 7 of the 8 isoforms.

Mouse *Hottip* - Primers were designed to mouse *Hottip* extracted from Ensembl (ENSMUSG00000055408). Probes were first designed over the 5' region of the gene. After testing the first set, a second set of primers were designed over the same region and were tested. When this second set of primers failed to amplify Hottip, primer sequences were obtained from Wang (2011).

2.2.2 RNA Extraction and First Strand cDNA Synthesis

Total RNA was extracted from head, FL and HL tissue from a CS17 embryo that had been stored in RNAlater® (2014). Mouse RNA was extracted from E13.5 mouse head, FLs and HLs that has been collected in RNAlater® and stored at -20°C for 5 days. RNA was extracted using a RNeasy® Mini Lipid Tissue Kit (QIAGEN, Germany) according to manufacturer's protocol. The concentration and quality of the RNA was determined using a Nanodrop 3000 by measuring the absorbance at 260nm and 280nm. RNA integrity was determined by gel electrophoresis on a 1% formaldehyde denaturing gel. The RNA sample was prepared for gel electrophoresis as follows: 1µg of RNA in 2x volume RNA loading buffer (10 ml deionized formamide, 3.5ml 37% formaldehyde, 2ml 5x MOPS and 1% ethidium bromide added to the working stock) was heated at 70°C for 5 minutes and snap-cooled for 1 minute on ice. The RNA was DNase1 treated.

200µg of RNA was converted to cDNA in a 20µl cDNA synthesis reaction prepared as follows: Of the DNase treated RNA, 5µL (0.75µg) was transferred to a nuclease free 200µl PCR tube, where 0.5 mM dNTPs (a mixture of dATP, dGTP, dCTP, dTTP at equal concentrations), 50 µM Oligo dT primers and nuclease-free water were added. The reaction was spun briefly in a microcentrifuge and heated to 70°C for 10 minutes using a GeneAmp Thermocycler, and then snap-cooled on ice for 1 minute. Following denaturation, the sample was collected at the bottom of the tube by microcentrifugation. To each tube, 1x First Strand buffer (Invitrogen), 2.5mM 0.1M DTT, 40U RNase inhibitor (Roche) and 200U Superscript™ III Reverse Transcriptase (Invitrogen) was added. Samples were then placed in the thermocycler and heated to 25°C for 15 minutes,

followed by 50˚C for 60 minutes, followed by reaction termination at 85˚C for 5 minutes. cDNA was stored at 4˚C for no longer than 1 week.

Bat Primers		Primer Sequence		Tm (˚C)	Amplicon (bp)
BHottip_F	5’	TTCCTTACACGCACACTTAGG	3’	62	548
BHottip_R	5’	CTCTTCCAGCACCGACATC	3’	62	
BTbx5_F	5’	CGACGACCACAGATACAAGTT	3’	62	306
BTbx5_R	5’	TAAACGCCGTCTCAGGAAAG	3’	62	
BTbx5as1_F	5’	GGACTGGGTTTCGTGGTC	3’	62	499
BTbx5as1_R	5’	TCCCTTCTGGAGAGGCTATT	3’	62	
BLef1_F	5’	AACAAGGGACCCTCCTACT	3’	62	573
BLef1_R	5’	CTCCTGCTCCTTTCTCTGTTC	3’	62	
BMllt3_F	5’	aTGCAGCAGATCGTGAACC	3’	55.5	931
BMllt3_R	5’	GTCCTACAGGTTTATTGACTAAGGC	3’	55.1	
BLhx8_F	5’	GGCAAGTGTGTGTGCAGC	3’	57.7	828
BLhx8_R	5’	ATGGGGTAACAAGGGCTGG	3’	57.4	
BHoxA10_F	5’	aTCTCCTCCCTCTGTCTTCGG	3’	58.1	463
BHoxA10_R	5’	AGCCCTTCTCTTTGCCAAGG	3’	59.6	
BHoxA13_F	5’	AAGGGTGGGCAGATGTTTAC	3’	55.1	579
BHoxA13_R	5’	CAGAACAGGAGGGTCAAAGAG	3’	55	

Mouse Primers	Sequence			Tm (˚C)	
MHottip_F1	5’	ACTCACCAGCAAGGAGAAAG	3’	62	455
MHottip_R1	5’	GGATCTAGTCCCACAGACAAAG	3’	62	
MLef1_F	5’	AACAAGGGACCCTCCTACT	3’	62	484
MLef1_R	5’	CTCCTGCTCCTTTCTCTGTTC	3’	62	
MMllt3_F	5’	aTGCAGCAGATCGTGAACC	3’	55.5	898
MMllt3_R	5’	GTCCTACAGGTTTATTGACTAAGGC	3’	55.1	
MLhx8_F	5’	GGCAAGTGTGTGTGCAGC	3’	57.7	823
MLhx8_R	5’	ATGGGGTAACAAGGGCTGG	3’	57.4	
MHoxA10_F	5’	aCGCGCAGAACATCAAAGAGG	3’	58.1	432
MHoxA10_R	5’	aCTTGCTGCTCTCGGAGGG	3’	59.6	

Table 2.1 Primer sets for amplification of *in situ* probes designed to *M. natalensis* and mouse RNA.

2.2.3 PCR Amplification and A-Tailing

The DNA fragments were amplified using CloneAmp^{TM} HiFi PCR Premix reactions (Clonetech, California) in the following 50µl reaction: 2x ClobeAmp HiFi PCR Premix, 0.3µM Primers, approximately 200ng Template DNA. With the following cycling parameters: 98˚C for 2 minutes (1 cycle), 98˚C for 10 seconds, 59˚C for 15 seconds, 72 for 20 seconds (30 cycles), and 72˚C for 1 minute final extension. PCR products were resolved on a 1% low melting point agarose gel. The PCR reaction was purified using the Wizard® SV Gel and PCR Clean-Up system (Promega) according to the manufacturer's protocol and eluted into 30µl nuclease-free water and 2µl of the purified DNA fragment was run on a 1% agarose checking gel to check whether there was a single band of the expected size. Concentration of the purified sample was determined using the Nanodrop ND1000 (Thermoscientific, USA).

Amplification by the CloneAmp^{TM} HiFi kit results in blunt-ended PCR products. Following purification, the blunt-ended DNA fragments were A-tailed so as to be used in the TA vector pGEM-T Easy®. A-tailing involved adding a maximum of 6µl of template, 1µl Super-Therm Taq Polymerase (Southern Cross Biotechnology), 25mM $MgCl_2$ (25mM), 10x Buffer and PCR-grade H_2O to the total volume of 10µl. The reaction was placed in the GeneAmp ThermoCylcer PCR System (Applied Biosystems) for 30 minutes at 70˚C.

2.2.4 Cloning and Transformation

pGEM-T Easy® has both a T7 and Sp6 site on either side of the multiple cloning site, making it a good vector for *in vitro* transcription as you can transcribe the probe irrespective of the orientation. A-tailed DNA fragments were ligated into pGEMT-Easy® (Promega) in the following 10µl reaction: 3µl DNA, 5µl 2X Rapid Ligation Buffer, and 1µl T4 DNA Ligase. This reaction was left to ligate overnight at 4˚C. 5µl of the ligation mix was transformed into 100µl competent *Escherichia coli* (E. coli) strain XL-1 Blue cells by heat shock transformation. Positive transformants (white colonies) were identified selected using β-galactosidase insertional inactivation and ampicillin resistance

selection on Luria-Bertani plates containing 100µg/ml Ampicillin, 0.5mM IPTG and 80µg/ml X-Gal (LAIX).

2.2.5 Screening Transformants and Sequencing

White colonies were randomly picked and screened for inserts by colony PCR using either gene specific primers or M13 primers in the following 20µl PCR reaction: 2x Reaction Buffer, 1.5 mM MgCl2, 0.2 mM dNTPs, 0.4 µM of each primer, and 1U Super-Therm Taq Polymerase (Southern Cross Biotechnology). The PCR was placed in a Thermo-cycler with the following cycling conditions: 98˚C for 5 minutes (1 cycle), 98˚C Colonies that for 30 seconds, 55˚C for 30 seconds, 72˚C for 1 minute (30 cycles) and 72˚C for 5 minutes (1 cycle). PCR products were resolved on a 1% agarose gel through gel electrophoresis. Colonies that contained the insert were grown overnight in a 2ml Luria Bertani media containing 100µg/ml ampicillin, shaking at 37°C. Plasmids were extracted from the overnight culture using Promega PureYield™ Plasmid Miniprep System (Promega, Fitchberg, WI, US) according to the manufacturer's protocol. Concentration of the purified plasmid was determined by Nanodrop 1000ND (Thermoscientific, USA). The purified probes were sent for sequencing to The Central Analytical Facilities at the University of Stellenbosch, where Sanger sequencing (Sanger et al. 1977) was used to confirm the identity and orientation of insertion of the gene product into the plasmid.

2.2.6 DIG-Labelled RNA Probe Synthesis

Thirteen DIG-labelled RNA probes were synthesized: *BMllt3, BLef1, BTbx5, BLhx8, BHoxA10, BHottip1, BHottip2, BTbx5-as1* and five mouse probes: *MMllt3, MLef1, MLhx8, MHoxA10.* 100 ml cultures were prepared using insert-positive plasmids in LB medium containing 100µg/ml ampicillin and left shaking at 37˚C for 16 hours. Plasmids were purified from the overnight cultures using PureYield™ plasmid Midiprep System (Promega) according to the manufacturer's protocol for large cultures and centrifugation. Purified plasmids were checked for inserts by a restriction endonuclease (RE) digest for 1 hour in a 37˚C water bath in the following 20µl reaction: 500ng plasmid

DNA, 1x RE buffer and 2 U Restriction Enzyme (Fermentas). The results were resolved on a 1% low melt agarose gel.

The insert positive plasmid DNA was linearized in the following 50µl reaction: 10µg plasmid DNA, 1x RE Buffer and 5 U RE. Reactions were left to digest overnight at 37˚C in a water bath. Complete digestion was visualized by running 1µl of the reaction on a 1% agarose gel stained with ethidium bromide (10mg/mL). Plasmid linearization reactions were purified using the Wizard® SV Gel and PCR Clean-Up system (Promega) according to the manufacturer's protocol and eluted into 30µl Tris-EDTA Buffer (pH 8.0). Concentration was quantified by measuring absorbance at 260nm with a NanoDrop 1000 (Thermo Fisher Scientific, Inc).

The purified linearized plasmid was used in the following 20µl *in vitro* transcription RNA labelling reaction: 1µg linearized plasmid, 6µl of RNase-free water, 10x Transcription buffer, 2µ 0.1M DTT, 2µ DIG RNA labelling mix (Roche, contains Digoxigenin labelled UTP), 1µl RNase inhibitor (Roche), Reactions were incubated at 37˚C in the Gene Amp Thermocycler for 3 hours. Genomic contamination was removed by treating the plasmid template with 5U of the endonuclease DNase I (RNase-free) (Ambion®, USA) at 37°C for 15 minutes. The RNA probe solution was purified using Sigma Spin™ Post Reaction Clean-up column (Sigma) according to the manufacturer's protocol. The purified probes were run on a 1% agarose denaturing formaldehyde gel to check for RNA degradation. Concentration was quantified by measuring absorbance at 260nm. Probes were placed directly into hybridisation solution (50% formamide, 5x standard saline citrate (SSC), 2% Blocking Reagent (Roche), 0.1 % Tween 20, 0.05µg/ml Heparin (Sigma) and 0.05µg/ml yeast tRNA (Sigma) and stored at -20°C.

2.2.7 Pre-Treatment of Embryos

The working Protocol for WISH used in this project originates from a protocol developed by D. Wilkinson, as modified and taught to A.C. Burke by M. Hollyday, in the McMahon laboratory at Roche (1993) and was further modified in the Tabin laboratory. All washes in were performed at in volumes of 5-10mls, at room temperature on a shaker, unless

otherwise noted. All solutions were diluted with 1xDEPC Phosphate-buffered solution containing 1% Tween 20 (PBST) and vacuum filter sterilized using a 0.22µm pore Stericup® Filter Unit (SigmaAldrich).

Due to the limited number of correctly staged bat embryos obtained in 2014, embryos were halved along their rostral-caudal axis to conserve samples. As a further measure to conserve bat samples, the *M. natalensis* probes were first tested in mouse embryos. Ideally each gene of interest would have a CS 15, 16 and 17 embryo for WISH for a direct comparison to be drawn between the RNA-seq and the WISH; however, due to the limited number of embryo's and exact stages of embryos, each gene was allocated a minimum of two halves of two differentially staged bisected embryos (an early and a late embryo). Where possible, there were three to four stages per gene. All embryos had been previously fixed overnight in 4% PFA (Sigma, St. Louis, MO, USA), followed by dehydration by a methanol series (25%, 50%, 75%, 2x 100% methanol) and stored at -20^{o}C (See 2.1). Embryos were rehydrated through a decreasing methanol series (75%, 50%, 25% methanol in DEPC PBS for 10 minutes at each step). They were then bisected and immersed in 6% Hydrogen Peroxide (H_2O_2,) for an hour. After three PBT washes to remove the H_2O_2, embryos were treated with 10µg/ml Proteinase K (Fermentas) in PBST for the following time periods: mouse embryos: E10.5: 10 minutes, E11.0: 10 minutes, E11.5: 12 minutes, E12.0: 15 minutes, E12.5: 20 minutes, E13.0: 22 minutes, E13.5: 25 minutes, E14.0: 30 minutes, E14.5: 35 minutes; Bat Embryos: CS 13L: 10 minutes, CS 14: 10 minutes, CS 14L: 12 minutes, CS 15: 14 minutes, CS 15L: 16 minutes, CS 16: 20 minutes, CS16L: 22 minutes, CS 17: 30 minutes, CS 17L 35 minutes. Following Proteinase K treatment, embryos were washed with freshly made and filtered 2mg/ml glycine (Mercke), for 10 minutes (to stop the Proteinase K reaction) and rinsed twice with PBST. The embryos were then post-fixed in freshly prepared 4% PFA (Sigma) and 0.2% EM grade glutaraldehyde (SigmaAldrich) for 20 minutes.

2.2.8 Hybridisation of the DIG-labelled RNA Probe

Embryos were prepared for hybridization of the probes by emersion in 10mls of hybridization solution (previously described), and incubation in a pre-warmed

hybridization oven for an hour at 70°C shaking. The DIG-labelled probe was added to 10ml of hybridisation solution, and incubated with the embryos overnight at 70°C. The probe solution was decanted into a 15ml Sterilin tube (Thermo Scientific) and stored at -20^{0}C for use in consecutive *in situ* runs. The initial post-hybridisation washes were conducted in a pre-heated solution 1 (50% Formamide, 5X SSC, pH4.5, 1% SDS) for three, hour-long washes in a 70°C hybridization oven. These washes were followed by lower stringency washes with Solution3 (50% Formamide, 2x SSC) for three, 30 minute long washes at 65°C, shaking in the hybridisation oven.

2.2.9 Blocking of the Embryos and the Addition of the anti-DIG Antibody

Following the post-hybridisation washes, embryos were washed 3 times for 10 minutes in 1x Tris-buffered Saline and Tween20 (TBST) in the following solution: 0.14M NaCl, 2.7 mM KCl, 25mM Tris HCL, pH 7.5, 1% Tween-20, 2 mM Levimasole .The embryos were pre-blocked for 2.5 hours at room temperature in 10% sheep serum (Sigma, heat inactivated at 56°C for 50 minutes) diluted in TBST. The embryos were washed at 4°C overnight in 1% sheep serum diluted in TBST, and Roche Anti-Digoxigenin- AP Fab fragments (150U/200µl) at a concentration of 1:50 000. The following day, embryos were washed three times in TBST for 5 minutes each, followed by 5, hour-long washes in TBST, and were left washing at 4°C in TBST overnight. Unbound antibody was removed by washing embryos for three 10 minute washes in PBST, followed by five hour-long washes at room temperature. An additional overnight wash was carried out at 4°C.

2.2.10 Signal Visualisation by Alkaline Phosphatase

The embryos were washed three times for 10 minutes in freshly made and filtered alkaline phosphatase buffer (NTMT) (100Mm NaCl, 100mM TrisHCl, pH. 9.5, 50 mM $MgCl_2$, 0.1% Tween-20, 2mM Levimasole). Levimasole acts to block endogenous enzyme activity that can produce large amounts of unspecific background. The colour reaction was made fresh using NTMT and NBT made fresh from powder (0.075g/ 1ml 70% dimethyl formamide) and BCIP (0.050g/1ml H_2O). The embryos were added to 10 mls of colour reaction and covered with a box that omitted all light, as light causes

background staining. The embryos were left in the dark, and the reaction progress was checked every 20 minutes. The CS17-late embryos required a second wash with the colour reaction, as well as an extra hour of staining, to obtain good signal.

Embryos were washed in NTMT before the colour reaction was complete, as the colour reaction continues to react until the embryos are washed in pH 5.5 PBST. The embryos were then post-fixed in 4% PFA and 0.1% Gluturaldehyde for an hour. The embryos were then taken through a methanol series. Dehydrating the embryo flushes out any non-specific, residual colour reaction that has built up in the tissue. Another benefit of Methanol is that it turns the BCIP-NBT stain dark blue, and the embryo flesh white, increasing the clarity of the stain. The embryos are kept in the dark on a shaker at 4˚C overnight in methanol.

2.2.11 Embryo Preparation and Photography

Due to the nature of BCIP NBT colour reaction causing background staining with time, all photographs were taken the day following the colour reaction. The embryos were first rehydrated in a series of reduced methanol PBS washes for 15 minutes each on a shaker (80%, 70%, 50%, 25%). The embryos were then taken through a glycerol series (25%, 50%, 70%, 80%) which required longer shaking durations of 30 minutes each, as glycerol does not easily infiltrate the embryo. Glycerol causes the embryos to become translucent when photographed under the light of the microscope, allowing for better stain resolution.

Photographs were taken using a Nikon Stereoscopic Zoom Microscope SMZ1500 equipped with a Nikon Digital Sight Camera Control Unit (DS-U2) and a DS-5M Camera head (Nikon Instruments Inc., Melville. NY, USA). NIS-Elements digital 3D imaging software (Nikon Instruments Inc. Melville. NY, USA) was used to capture, format and add scale-bars to the images. Embryos were photographed on a petri dish containing 50mls of 2% agarose (to remove any glare and scratches from the microscope stage) and remained submerged in 80% glycerol PBS during photography. Side-view images were taken to be able to directly compare FL vs HL staining as well as over-staining of

the rest of the embryo. Floating tissue was removed from the backgrounds and images were compiled using Adobe® Photoshop® CS4 11.0.1.

2.3 Bioinformatics

2.3.1 Analysis of *M. schreibersii* RNA-seq Dataset

Wang et al. (2014) published a RNA-seq dataset (Accession GSE50699) of a closely related species (Stoffberg et al. 2004) of bent-wing bat *M. schreibersii* (Kuhl, 1819). This transcriptome was used as a comparison to the *M. natalensis* dataset (Eckalbar et. al manuscript under review). The *M. schreibersii* dataset was designed in a complementary manner to our dataset, as tissue from the same developmental stages (CS 15-17) was pooled but divided into digit and interdigital tissue as well as FL digit 1 and the interdigital tissue between digit I and digit II (Figure 1.5). Out of the 8 individuals, there were 2 CS15 FLs, no HLs (the digits weren't obvious in the HL), 1 CS16L individual and 5 CS17 individuals. Due to the number and size of CS 17 FLs compared to the other stages, the data is biased towards representing CS 17 gene expression (Figure 1.5). The dataset was downloaded from the Gene Expression Omnibus (GEO). The transcriptome is not annotated by gene name and the genes were fragmented due to the transcriptomes being assembled *de novo* (without a reference genome) resulting multiple "genes' aligning to *Tbx5-as1*. To extract the gene count information for our genes of interest, a Blastn was performed on the *M. schreibersii* transcriptome using transcripts from the *M. natalensis* assembly, this was conducted in BioLinux using the Blastn function, and the M. natalensis transcriptome as the Blast database. *M. schreibersii* sequences with high similarity were considered to represent the candidate gene, and the associated normalized gene counts were extracted.

2.3.2 Bioinformatics Analysis of Cloned *in situ* Probes

Returned *in situ* probe sequences were aligned to genomic scaffold sequence to determine the orientation of insertion into the vector. A consensus sequence for each insert was created, and remaining vector sequence was removed using BioEdit Sequence Alignment Editor Version 7.2.5 (T. Hall, 1999). Sequences were positively identified as the target gene from bat using the Basic Local Alignment Search Tool (BLAST)(Altschul et al. 1997) provided by the National Centre for Biotechnology Information (NCBI). The LncRNA sequences were Blasted against the 'Reference RNA

sequence (refseq_RNA)' database. The appearance of gene-specific sequences within the top five Blastn results indicated that the insert contained the correct sequence. A BLAST-like alignment tool (BLAT) was conducted to against a microbat (*Miotis lucifugus*) genome in order to visualize the sequence position relative to the *M. lucifugus* gene homologue. Sequences were exported into RestrictionMapper Version 3 and Serial Cloner Version 2.5 (Serial Basics, France) to determine suitable restriction enzymes (RE) for linearizing the plasmids containing the probes. Double-cutter REs, as well as REs that generate a 3' overhangs were avoided as 3' overhangs reduce the efficiency of *in vitro* transcription reaction.

2.3.3. Characterization of LncRNA Isoforms from RNA-seq Data

Different variants were identified and manually aligned with one another and compared to human, mouse and *M. schreibersii* sequence. Alignment and editing of the sequences was done using BioEdit, Sequence Alignment Editor Version 7.2.5 (Hall 1999b). Vector sequences were removed, and the transcripts were manually aligned, as CLustalW could not create large enough gaps to account for the large exons excluded in some splice variants.

2.4 Characterizing *Hottip* and *Tbx5-as1* Transcripts

2.4.1 Primer Design

Primers were designed against the sequence variants generated by the RNA-seq dataset, using Integrated DNA Technology website (IDT). One set of primers (F1 and R1) were designed for *Hottip*, and two sets (F1 and R1; F2 and R1) were designed for *Tbx5-as1*. The Primer sets were analysed using the OligoAnalyzer® Tool (SciTools®).

	Sequence			**Tm (°C)**
BHottip_TC_F1	5'	CAGGTGTCTACTCGTCTTTCTG	3'	60.6
BHottip_TC_R1	5'	AGGGACGGTCCTGAAGAT	3'	60.5
BTbx5as1_TC_F1	5'	TGGTTCCTCTCTGTCTACACC	3'	61.4
BTbx5as1_TC_F2	5'	TGCAACTGACAGAGGTTCAG	3'	60.7
BTbx5as1_TC_R1	5'	CAGGTTCATGCGGCTAGTG	3'	61.3

Table 2.2 Primers for Transcript Characterisation of the LncRNA isoforms.

2.4.2 RNA Extraction and Cloning

RNA was extracted from the head, FL and HLs of a CS 18L embryo, as previously described. The embryo had been stored at -20˚C for 8 months. The *LncRNA t*ranscripts were amplified as previously described using primer sets in listed in table 2.2. A positive control: TATA box binding protein-like 1 (*Tbpl1*) validated as a reference gene by Mason (2012) on CS 16 and CS 17, FL and HLs was included in the PCR run. The PCR reaction was visualised on a 1% agarose gel by gel electrophoresis.

2.5 qPCR

2.5.1 Linear Regression Efficiency Method

LRE method was used to compare the expression profiles of three regions in and around the *HoxA10* gene. Linear Regression Efficiency (LRE) yields absolute values for a targeted region, without the need for a standard curve but requires the efficiency of the reaction to be as close to 100% as possible for the prediction to be accurate (Ruijter et al. 2013). LRE utilizes the curve of λ gDNA with primers for an amplicon CAL1 (151bp) and CAL2 (113bp). The kinetics of this plot can be studied by plotting cycle efficiency (y axis) by fluorescence (x axis).

2.5.2 Primer Design

Primers for *HoxA10* qPCR were designed over three different regions in and just beyond the *HoxA10* gene. Primers were designed using Integrated DNA Technology website (IDT) and were analysed using the OligoAnalyzer® Tool (SciTools®). The first primer set (F1 and R1) was designed over a region just past the 3' untranslated region (UTR) that gave high and inconsistent RNA-seq signal. The second primer set (F2 and R2) were designed over the 3' end of the gene. The third set (F3 and R3) were designed over the *in situ* probe region (Table 2.3).

	Sequence			Tm (°C)	GC%	Size (bp)
BHoxA10_F1	5’	AGAATTGTGGTGTGCTTGT	3’	60.1	42.1	168
BHoxA10_R1	5’	CTATTTCCCTTTCCTCTGAATCT	3’	59.9	39.1	
BHoxA10_F2	5’	CATGGGCATACAGGTTGG	3’	60.0	55.5	91
BHoxA10_R2	5’	CATCTTTCTGACTCTCTCCATTT	3’	60.3	39.1	
BHoxA10_F3	5’	GTTTGTCGTCTCCGTGTC	3’	59.8	55.5	118
BHoxA10_R3	5’	GCCTTCCGCTAGGTTTG	3’	59.7	58.8	

Table 2.3 Primer sets for qPCR on *M. natalensis HoxA10*, F1 and R1 amplify a region just past the 3’ UTR. F2 and R2 amplify a region within the 3’ UTR. F3 and R3 amplify a region within the in situ probe as a direct comparison for what is seen in the *in situ*.

2.5.3 MIQE Standards

The Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines were followed closely to ensure that the results are publication-worthy (Bustin et al. 2009). Bat embryos were placed directly in RNAlater® following excision from the maternal uterus. The sample was then kept on ice for 12 hours and stored at -20°C for 1 month prior to RNA extraction.

RNA was extracted as previously described, with the amendment of increasing the tissue homogenization step to 10 minutes from the previously stated 5 minutes. RNA was eluted into ultrapure H_2O and DNase treated as previously described. DNase treated RNA was aliquoted into 2µl Eppendorf® DNA LoBind tubes (Sigma-Aldrich) and stored at -70˚C. Samples were converted to cDNA within one month to prevent degradation. RNA was extracted from 3 CS 15 FL and HLs, and 3 CS 16 FLs and HLs. RNA was run on a denaturing formaldehyde gel to check for RNA quality (data not shown). Concentration of RNA was determined by Nanodrop (listed in Table 2.4)

Carollia Stage	ID Code	FL (ng/µl)	HL (ng/µl)
CS 15	Mn15_45	159.8	118.1
	Mn15_36	132.8	75.5
	Mn15_22	156.4	64.8
CS 16	Mn15_38	237.1	146.9
	Mn15_42	196.8	116.9
	Mn15_43	224.5	175.2

Table 2.4 Concentration of RNA from each extraction. In order to have 500ng of RNA in a 20µl qPCR reaction, concentrations needed to be 65 ng/µl and above.

2.5.4 cDNA Synthesis

All 18 samples (se experimental setup) of cDNA were generated at the same time. Of the DNase treated RNA, 5µL (0.75µg) was transferred to a nuclease free 200µl PCR tube, where 1µl of 10mM dNTPs, 0.5M Nonomers (Sigma-Aldrich) and 6µl of nuclease-free water were added, the reaction was spun briefly in a microcentrifuge and heated to 70°C for 10 minutes using a GeneAmp Thermocycler. Samples were then cooled on ice and centrifuged for 30 seconds. To each tube, 4µl of 5x First Strand buffer, 1µl 0.1M DTT, 1µl 40U/µl RNase inhibitor and 200U Superscript™ III Reverse Transcriptase (Invitron) was added. Samples were then placed in the thermocycler and heated to 25°C for 15 minutes, followed by 50°C for 60 minutes, and incubated at 70°C for 15 minutes to inactivate the reaction. The cDNA was diluted with 60µl nuclease-free water and 4µl was used as the template for each qPCR reaction. The cDNA was diluted to increase the accuracy and precision of the QIAgen QIAgility liquid Handling workstation for PCR setup.

2.5.5 λ gDNA Dilutions

Λ gDNA was heated at 60°C on a heating block for 2 minutes prior to dilution in Tris (10mM). The 1ng-100fg λ gDNA concentration range was elected as it yielded the least amount of variance in curve efficiency. In order to successfully achieve this low concentration dilution with accuracy, a serial dilution was conducted as described in

Figure 2.1. The concentration of each dilution from 100ng to 1ng was measured three times by the Nanodrop, in order to generate a stable average and be able to aliquot the correct volume into each dilution. The 1ng measurement needed to be as close to 1ng as possible, as the 10pg-100fg could not be measured prior to the qPCR. Dilutions were performed in the lamina flow hood to prevent contamination.

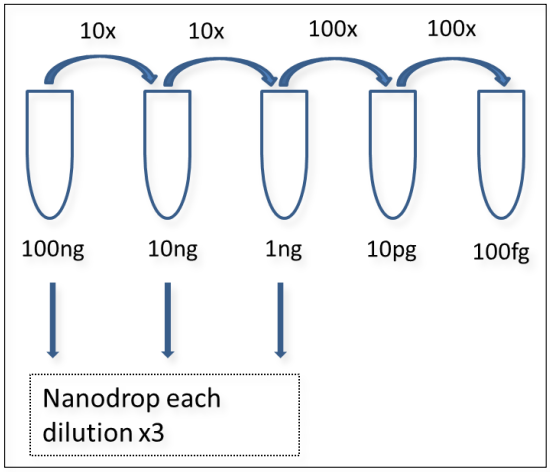


Figure 2.1 Serial dilution of Lambda gDNA as a LRE qPCR calibrator.

2.5.6 Experimental PCR Setup

The setup described in Figure 2.2 was programmed using the QIAgility Sotware (Version 4.15.1). Reactions were aliquoted using the QIAgen QIAgility liquid Handling workstation and were setup as follows: 1.5µl (0.375Mm) forward primer, 1.5µl (0.375Mm) reverse primer, 10µl 2x SensiMix™ SYBR No ROX (Bioline), 6µl cDNA (1µg), 1 µl Ultrapure H_2O. Reactions were set up in a pre-cooled block, in a ring of 100 tubes, and sealed with Rotor-Disc® heat sealer prior to qPCR in the Rotor-Gene 6000 (Corbett Research Pty Ltd). Standard SYBR green settings were used for a two-step qPCR with a melt step. Runs were analysed using the Rotor-Gene software (ScreenClust HRM software).

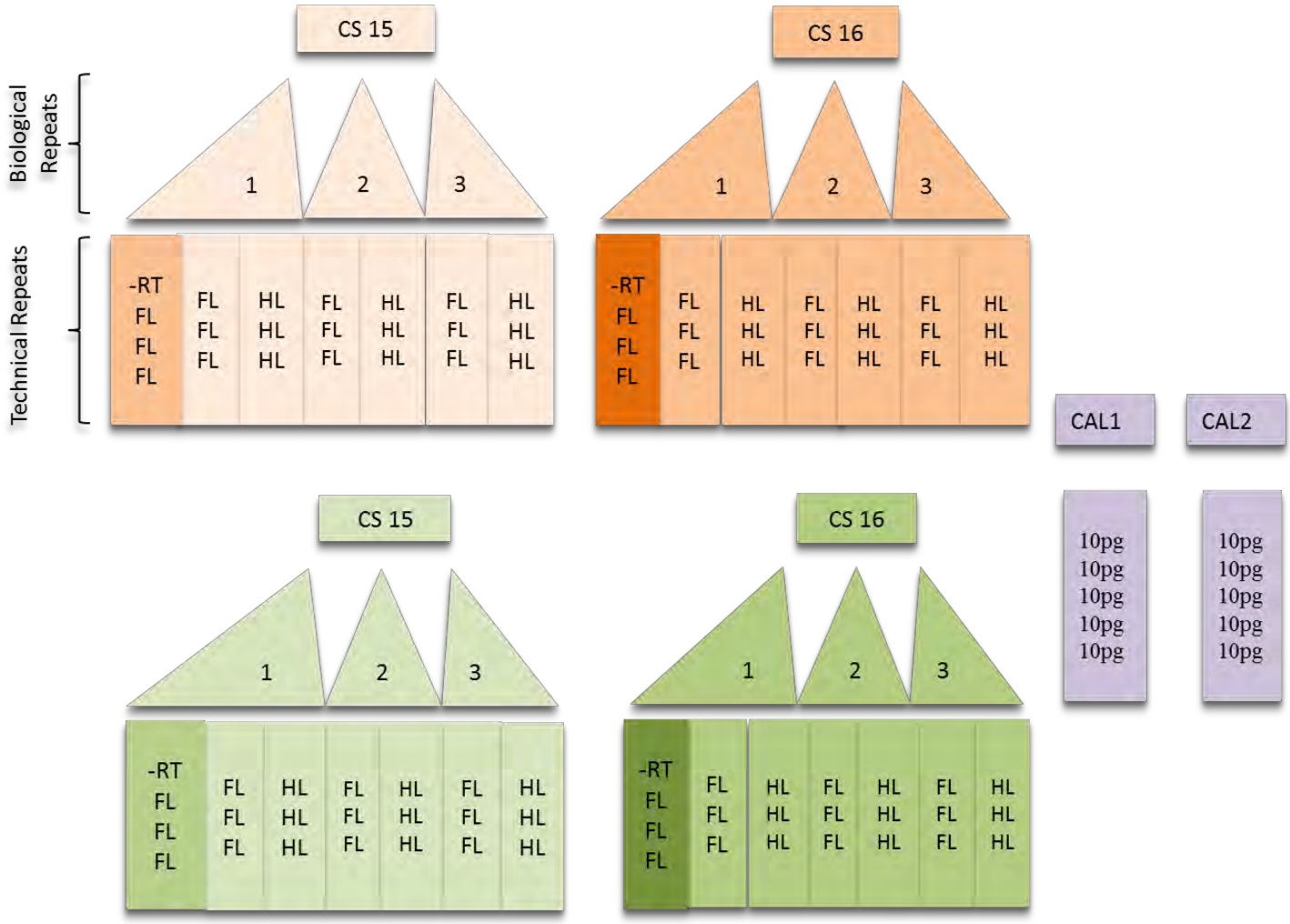


Figure 2.2 Experimental Setup of the qPCR of two regions of the *HoxA10* gene locus. Orange represents primer set 1 (designed over the region just past the 3' UTR of the *HoxA10* locus). Green represents primer set 3 (designed over the *in situ* probe). The –RT samples were conducted in triplicate on one of the biological samples. CAL1 is the calibrator primer set 1 of the lambda DNA, CAL2 is the calibrator primer set 2 of the lambda DNA.

2.5.7 LRE Analysis

qPCR data was analysed using the LRE Analyzer (Version 9.10) (Rutledge 2011). The LRE analyser is capable of using the λ gDNA sigmoidal curve as a calibrator to determine the OCF, which is then used to calibrate the analyser to generate absolute values for concentration of starting material. Efficiencies of the reactions needed to reach a window of between 95%-105% efficiency for the prediction to be accurate.

Chapter 3

Transcription Factor Expression Analysis in Developing Mouse and Bat Limbs

3.1 Introduction

Transcription factors play a crucial role in the development of morphological variation by regulating spatial expression (heterotopy) and timing or rate of expression (heterochrony) of toolkit genes during development. Transcription factors can have a multitude of pleiotropic effects during development by binding to stretches of DNA known as enhancers. Enhancers recruiting co-activators or repressors to either stimulate or inhibit the binding of RNA polymerase II (RNAP2) (Spitz & Furlong 2012).

Five transcription factors: *Lef1*, *Lhx8*, *HoxA10*, *Mllt3* and *Tbx5* were selected from the RNA-seq dataset for further spatial expression characterization. These selected transcription factors were associated with two known developmental regulators: Wnt /β-catenin signalling and regulation by *Hox* genes. In this section, the two developmental regulators will be expanded upon, followed by an introduction to each of the selected transcription factors.

WNT/β-catenin pathway

Wnt/β-catenin signal transduction is a pivotal developmental signalling pathway mediated by a set of powerful wingless-type (Wnt) signalling proteins (Tamamura et al. 2005; Church & Francis-West 2002). Wnt/β-catenin signalling is involved in bone formation by endochondral ossification, a process whereby mesenchymal cells condense to create a template for cartilage to form, that later matures into bone (Tamamura et al. 2005; Hartmann 2006). The ‘ON’ and ‘OFF’ states of Wnt/β-catenin signalling have been clearly define. The ‘OFF’ state of Wnt/β-catenin, in the absence of Wnt ligands, a protein complex known as the Axin complex (consisting of: adenomatous polyposis coli gene product (APC), casein kinase 1 (CK1), and glycogen synthase

kinase 3 (GSK3)) causes the phosphorylation of a transcription co-activator known as β-catenin. Once phosphorylated, β-catenin gets ubiquinated and degraded by the ubiquitin-proteosome pathway (Figure 3.1, A). This process prevents β-catenin from building-up in the cytoplasm and reaching the nucleus, thus preventing the transcription of Wnt responsive genes.

The 'ON' state of Wnt/β-catenin involves the binding of Wnt ligands to the transmembrane cell receptors Frizzled (Fz) and its co-receptor low-density lipoprotein receptor-related protein 6 (LRP6)(Figure 3.1, B). Once the Wnt-Fz-LRP6 complex is formed, the destructive complex diffuses, inhibiting the degradation of β-catenin. This results in a build-up of β-catenin in the cytoplasm and translocation through the nuclear pore to bind to TCF/LEF transcription factors and activate the transcription of Wnt responsive genes (Figure 3.1, B) (MacDonald et al. 2009).

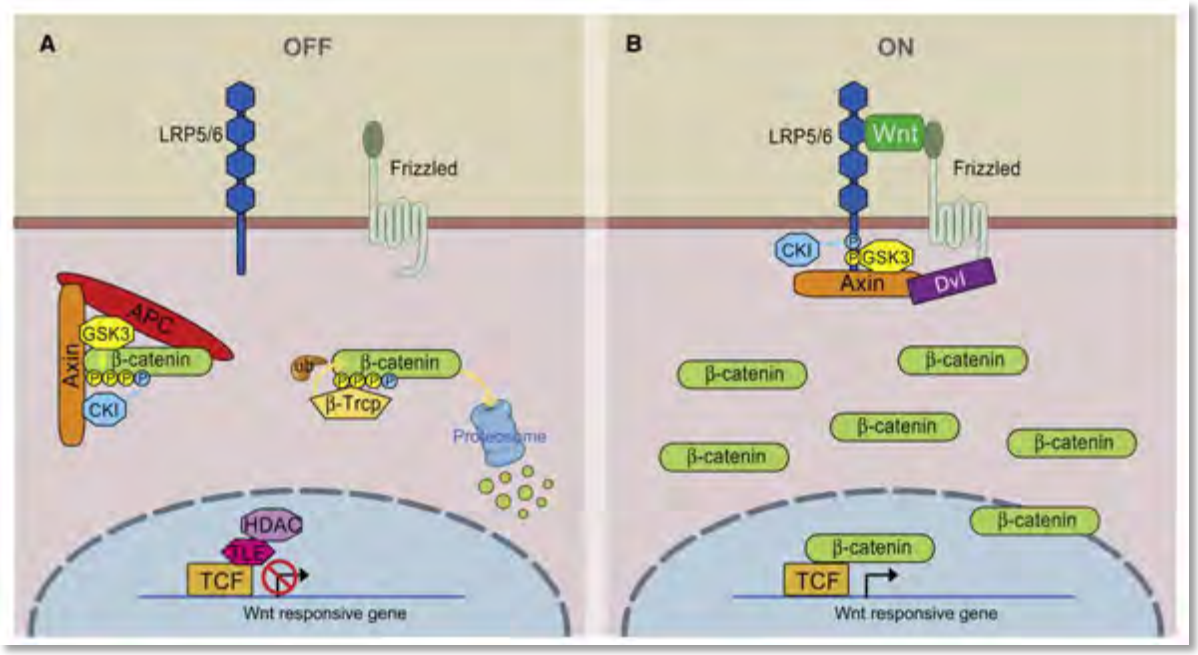


Figure 3.1 The two states of Wnt/ β-catenin signalling. A) In the absence of Wnt signalling, β-catenin is phosphorylated by the destructive complex (Axin, CKI, GSK3, APC). B) The active state of Wnt signalling. When a Wnt signal binds the Frizzled and LRP5/6 receptors, this causes the destructive complex t be attracted to the transmembrane complex, thus inhibiting the breakdown of β-catenin. B-catenin builds up in the cytoplasm and translocates through the nuclear pore where it binds to TCF/LEF transcription factors and activates the transcription of Wnt responsive genes (Image from MacDonald et al. 2009).

3.1.1 *Lef1*

Lymphoid enhancer factor 1 (Lef1) is a member of the TCG family of transcription factor genes of known effectors of Wnt/β-catenin signalling. There are four TCF genes in mammals: *Tcf1*; *Lef1*; *Tcf3* and *Tcf4* (Hoppler & Kavanagh 2007). Disruption of *Lef1* has been shown to have a number of pleiotropic effects, from a disorganised growth plate to inhibition of chondrocyte and osteoblast maturation (Tamamura et al. 2005). In the absence of Wnt signalling *Lef1* can repress gene transcription with the co-repressor *Groucho*, while *Lef1* can activate transcription of Wnt responsive genes in the presence of Wnt signalling (Figure 3.1, B). In mice WISH studies, *Lef1* has been shown to be expressed in early development (E9.5) in the mesenchyme of the FL bud, with expression being excluded from the AER (Galceran et al. 1999). Double *Tcf-1* and *Lef1* null mutant mice were found to have a deficiency in the formation of paraxial mesoderm that later develops into muscles and dermis (Galceran et al. 1999). A dominant-negative form of *Lef1* interfered with the function of *Wnt3a* in inducing the expression of *Bmp2*, *Fgf4*, and *Fgf8* in the AER (Kengaku et al. 1998).

In the *M. natalensis* dataset *Lef1* was significantly higher in the HLs compared to the FLs at CS 15 (p-value 0.043), suggesting an up-regulation of Wnt/ β-catenin signalling in the HLs (Eckalbar et. al., manuscript under review).

3.1.2 *Lhx8*

Lim Homeobox 8 (Lhx8) is one of multiple *Lhx* genes containing LIM homeodomains that serve a variety of regulatory functions during development. In early mouse development, *Lhx8* is expressed in the bony ridge of the upper jaw known as the maxillary arch and is regulated by the WNT/β-catenin pathway, as a specific enhancer *'Lhx8_enh1'*, was found to bind a myriad of signalling molecules, with LEF/TCF family members being some of them (Landin et al. 2014). The exact target of *Lhx8* is not as clearly defined, however; *Lhx6* and *Lhx8* have been found to stimulate *Shh* expression in the Medial Ganglionic Eminence (MGE) (Flandin et al. 2011). *Lhx8* has also been found to regulate cholinergic neuron development (Zhao et al. 2003)(Tomioka et al. 2014). This pathway is particularly interesting as bat wing membranes have a highly

adapted set of nerves to perform complex aerial manoeuvres (Sterbing-D'Angelo et al. 2011)

Lhx8 was selected as a candidate for spatial expression characterisation as *Lhx8* has not been recorded in mouse or chick limb development, however; *Lhx8* displayed a highly differential expression at later stages (CS 16 and CS17), with a $\log_2$ Fold Change of 3.86 at CS 17 in the *M. natalensis* dataset (Eckalbar et. al., manuscript under review).

3.1.3 *Hox* Genes

39 *Hox* genes encode transcription that exist as multi-gene 'clusters', a term given to groups of genes that are located in the same region, are co-expressed and have similar function (Tschopp & Duboule 2011). These clusters have undergone four-fold duplication in modern vertebrates to result in four clusters termed *HoxA*, *B, C* and *D*, and are spread across 4 chromosomal loci, with each cluster contains 9-11 genes that can be aligned to 13 paralogous groups (Knosp et al. 2004).

During development, *Hox* gene expression controls the identity of body regions according to the rules of spatiotemporal collinearity, whereby genomic order of genes within the cluster determines its domain of influence (Zakany & Duboule 2007). Therefore, genes existing on the 3' end of the cluster governing the patterning of anterior structures, while genes located on the 5' end of the cluster govern patterning on the posterior end of the embryo (Slack et al. 1993). This applies to timing of expression as well, with genes located toward the 3' end of the *Hox* cluster being activated earlier than genes that are located on the 5' end of the *Hox* cluster (Zakany & Duboule 2007).

The particular *Hox* genes involved with regulating limb patterning are genes that reside on the 5' end of the *HoxA* and *D* cluster, namely genes *Hox9-13* (Davis 2013). This was discovered in mutant studies where loss of *HoxA* and *HoxD* clusters causes truncation of limbs to approximately a third of wildtype limbs (Zakany & Duboule 2007).

During development there are two major phases of *HoxA* and *D* gene expression that pattern the limb: the 'Early Phase' and 'Late phase'. The early phase initiates expression of *Hox* genes in the collinear pattern, while the late phase activates expression of *Shh* in the posterior region of the limb bud (Suzuki 2013). The 'Late Phase' of *Hox* gene expression initiates reverse collinear expression across the A-P axis of the developing hand. Genes that are more 3' are expressed in the posterior region, whereas more 5' *Hox* genes are expressed in the anterior region of the limb bud (Suzuki, 2013).

3.1.4 *HoxA10*

In the late phase of expression of the 5' *HoxA* genes, *HoxA10* is expressed in the entire autopod (Figure 3.2 B), whereas *HoxA11* is restricted to the presumptive zeugopod region (Figure 3.2). Seeing as Hox genes function in overlapping regions, and *in situ's* of *HoxA11* and *HoxA13* have already been conducted on *M. natalensis* embryos (Eckalbar et al., manuscript under review) spatial characterisation of *HoxA10* would complete the picture of the 5' *HoxA* genes expression in the developing bat wing.

Figure 3.2 Non co-linear expression of HoxA genes. *HoxA10* genes are responsible for patterning the stylopod, *HoxA11* genes the zeugopod and *HoxA13* the autopods (Berlivet et al. n.d.).

Unusual *HoxA10* signal in the 3' region of the gene

An inconsistency was discovered in the RNA-seq tracks in a region 3' of the *HoxA10* gene locus, where there is a high signal in FL (CS 15), a high signal in the HL (CS 16) and a high signal in the FLs (CS 17) (Figure 3.3, red box) (Eckalbar, manuscript under review). In an attempt to decipher whether this signal is real or an artefact, qPCR was

conducted on this region as well as on a region within the *in situ* probe to use as a comparison (Figure 3.3).

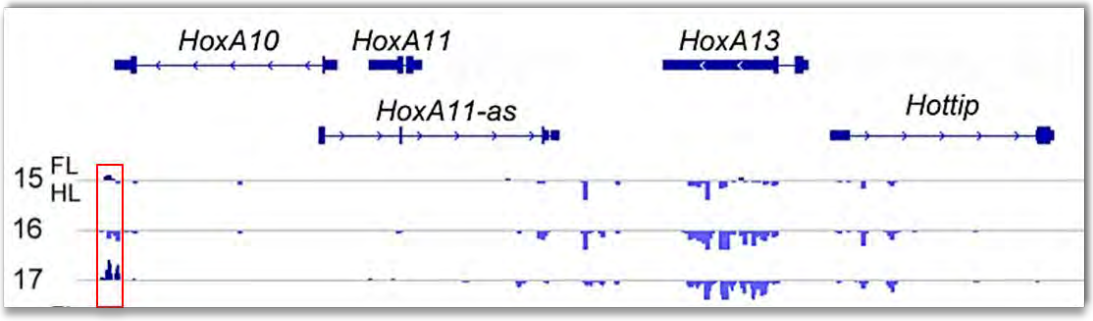


Figure 3.3 RNA-seq tracks of signal around the 5' HoxA gene locus. The box in red demarcates a region of inconsistency in the results as there is a region of high expression of this region in the FL at CS 15, an increase in expression in HLs at CS 16 and increase in expression in FLs at CS 17 (Eckalbar, manuscript under review).

3.1.5 *Mllt3*

Mllt3 or *mixed-lineage leukaemia translocated to 3,* is a common fusion partner of the mixed lineage leukaemia (MLL) gene product. MLL is frequently rearranged by reciprocal chromosomal translocations resulting in the expression of chimeric transcripts that are fused in-frame between the MLL gene and various partner genes (Vogel & Gruss 2009). The *Mllt3*/MLL fusion in particular, has been associated with acute myeloid leukemia and acute lymphocytic leukemia. The *Mll* gene encodes a large (431kDa) protein with a DNA-binding motif in its N-terminus, as well as a DNA methyltransferase domain that potentially regulates transcription of its target genes (Pramparo et al. 2005). MLL knockouts have a similar phenotype to *Mllt3* knockouts, suggesting that the fusion partners act together to affect skeletal patterning.

Mllt3 is one of 20 MLL fusion partners and contains a transcriptional activation site which is maintained in the MLL-T3 fusion protein. The function of the *Mllt3* gene is largely unknown, however; *Mllt3* null mutations in mice cause defects in axial skeletal segmentation, while heterozygous *Mllt3*$^{-/+}$ mutations have little phenotypic effect. Homozygous mutants display transformation of thoracic and cervical vertebrae to a more anterior identity (Figure 3.4 B)(Collins et al. 2002).

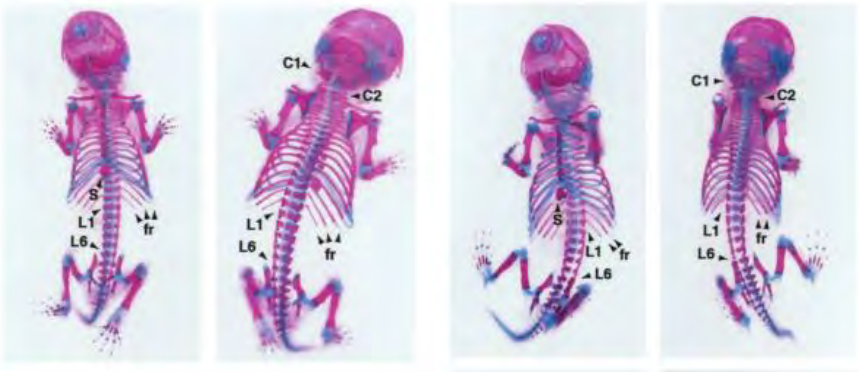


Figure 3.4 Ventral and dorsal views of A) a wild type mouse, and B) a homozygous *Mllt3* mutant mouse. B) lacks a pair of floating ribs (Fr) as well as abnormalities in the development of the first and second cervical vertebrae (C1 and C2)(Collins et al. 2002).

Mllt3 and *Hox* Genes

The MLL protein contains homologous domains to the Trithorax (Trx) protein in Drosophila such as a SET protein binding domain and three PDH zinc finger domains (Kawagoe et al. 2001). MLLT3 chimeric protein positively regulates expression of *Hox genes as Mllt3* null mutants are similar to Hox4 gene (*HoxA4*, *HoxB4* and *HoxD4*) null mutants (Kawagoe et al. 2001). *Mllt3* perturbation causes an anterior extension of HoxD4 expression, however; there is no evidence for its effect on any other *Hox* genes. Despite the lack of evidence in mice, the association of the C-terminal portion of the MLLT3 protein with N-terminal MLL may target the fusion protein sites not normally affected by MLLT3 (through the DNA-binding properties of MLL) creating a novel *Hox* gene regulator. It has also been revealed that *Mllt3* recruits HPC3, a member of the human Polycomb group of proteins known to regulate *Hox* gene expression and axial skeleton development.

Mllt3 was selected from the *M. natalensis* dataset as it displayed strong FL expression at CS15 and CS16 (Eckalbar, et. al., manuscript under review), as well as being a known *Hox* gene regulator *Mllt3*-null mouse mutants exhibiting axial defects (Collins et al. 2002).

3.1.6 Aims

In this chapter, spatial expression domains of 4 transcription factors: *Lef1*, *Lhx8*, *HoxA10* and *Mllt3* were investigated using WISH on *M. natalensis* and equivalently staged mouse embryos. WISH required the preparation of DIG-labelled RNA probes that are complementary to the mRNA of the transcription factors. *In situ* expression domains (i.e. whether expression was in digital or interdigital tissue) were compared with RNA-seq data from digit and interdigital tissue from *M. schreiberseii* FL and HL autopods.

3.2. Results

3.2.1 Generation of Probes for *In Situ* Experiments

The cloned gene sequences were used as templates for RNA probe synthesis by *in vitro* transcription. The multiple steps involved in generating a RNA probe can be seen in Figure 3.7, with all the synthesized and purified probes having a single clear band of the expected size.

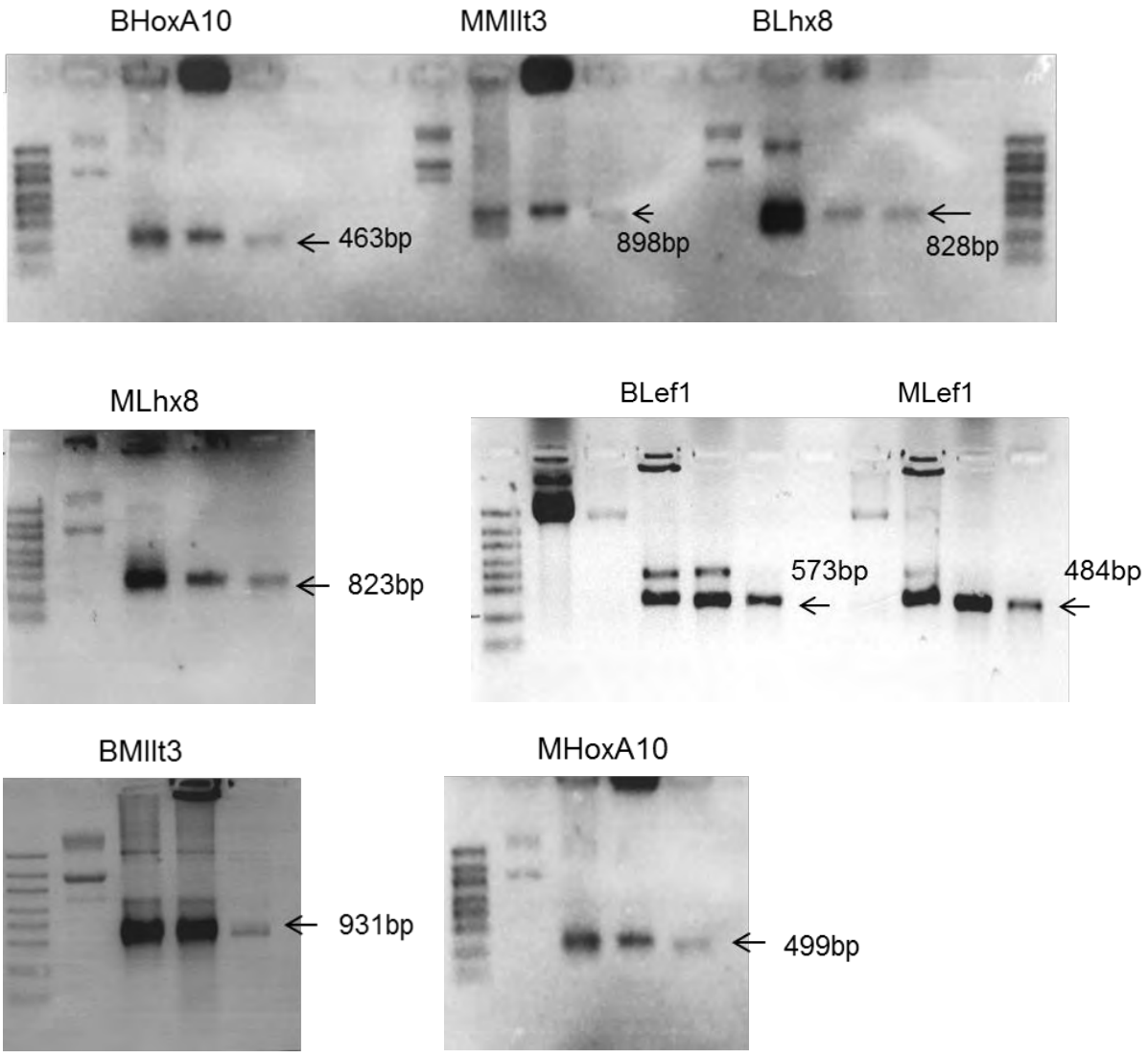


Figure 3.5 Formaldehyde Probe Synthesis Steps. Step 1) Linearized plasmid Step 2) In vitro transcribed probe. Step 3) DNase treated Probe. Step 4) Purified probe.

3.2.2 *M.natalensis* and *M. schreibersii* Dataset Comparison

Initially the *M. schreibersii* dataset (Wang et al. 2014) appeared to be a promising complementary comparison for the *M. natalensis* dataset (Eckalbar et. Al., manuscript under review), as the species are closely related and sample the same developmental stages. After personal communication about sampling, and closer inspection of the *M. schreibersii* data three issues arose: 1) dissection of CS 15 and CS 16 limb buds into the various tissue domains (Figure 1.6) is problematic as carry-over is easy. 2) There is a bias of the tissue towards CS 17 in sample number and tissue mass (as illustrated in Figure 1.5). 3) Lastly, the *M. schreibersii* transcriptomes were assembled *de novo*, meaning it was assembled without a reference genome, making gene annotations problematic. The *M. schreibersii* RNA-seq dataset was assembled using Trinity package to yield 60 000 'gene' fragments with associated gene counts to be listed in the transcriptome, compared to our ~15000 expressed genes. Nonetheless, the data can be used to validate expression between FL and HL for CS 17, as well as earlier expression patterns in CS 14 that precede our RNA-seq dataset.

A benefit of having the *M. schreibersii* dataset is the ability to determine whether the selected genes are localised into digit or interdigital tissue, or both. Another benefit is being able to determine this pattern in the stage that precedes and succeeds the *M. natalensis* RNA-seq dataset.

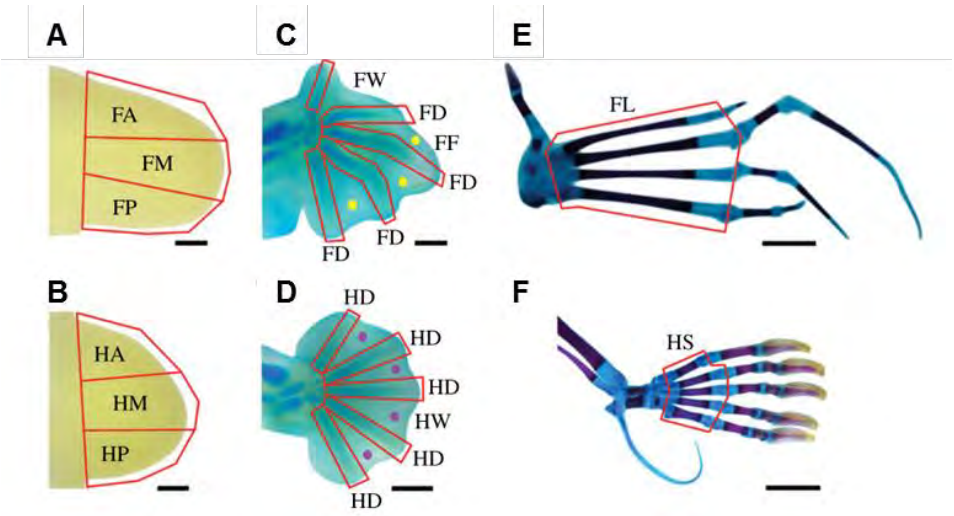
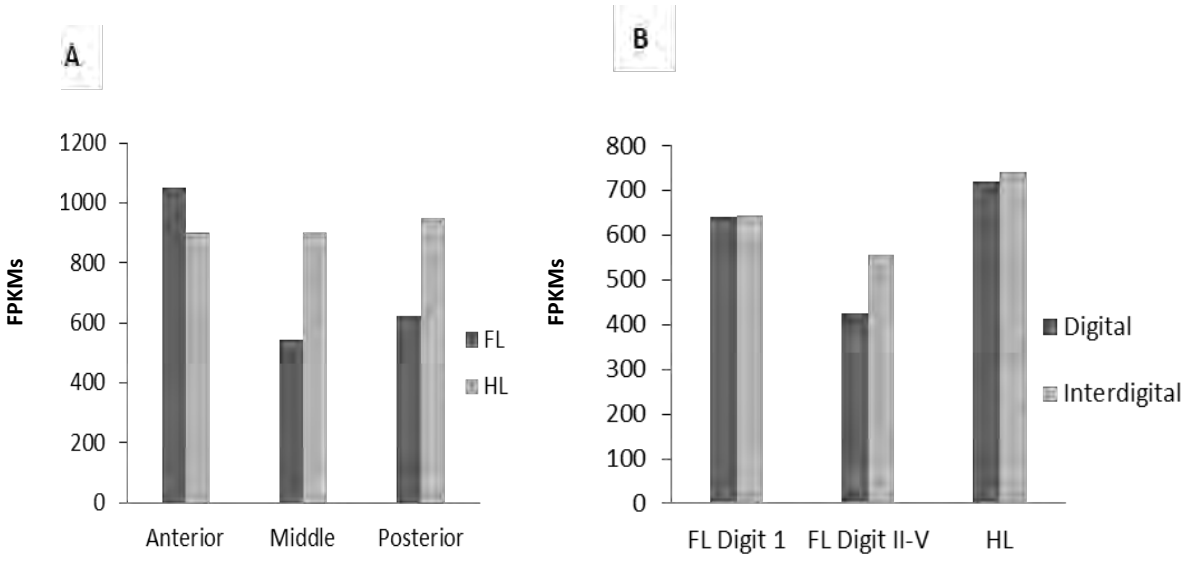


Figure 3.6 Dissection layout of the *M. schreibersii* RNA-seq dataset. A and B) CS 14 autopods; FA: FL anterior domain, FM: FL middle domain, FP: FL posterior domain. HA: HL anterior domain; HM: HL middle domain; HP: HL posterior domain. C and D) CS 15-CS 17 autopods; FI: digit 1 tissue; FW: Interdigital tissue between digit I and II; FD: digit II-V tissue; FF (yellow dots) interdigital tissue between digits II and V; HD: hindlimb digit tissue; HW: (purple dots) interdigital tissue. E and F) The red box outlines the region of sampling. FL: elongating metacarpal tissue from digit II-V; HS: Elongating digit tissue from metatarsal tissue (Image taken from Wang et al. 2014).

3.2.2.1 Lef1

At CS 14 expression of *Lef1* is similar in the anterior region of the FL and HL, with expression being higher in the HL in the middle and posterior region. At the patterning phase, expression is virtually identical in the digit I and interdigital tissue between digit I and digit II. Expression decreases in the FL digit and interdigital region between digit II-V, with signal being stronger in the interdigital tissue than digit tissue. Consistent with the CS 14 pattern of expression, *Lef1* expression is higher in the HL in both the digit and interdigital tissue relative to the FL. This pattern reverses in the elongating digit tissue by the middle fetal phase, with expression being 3 fold greater in the FL than the HL (Figure 3.7. C). In the *M. natalensis* dataset, expression is consistently higher in the HL relative to the FL at all 3 stages, with $\log_2$ fold changes of -0.43 (CS 15), -0.40 (CS 16), -0.49 (CS 17).



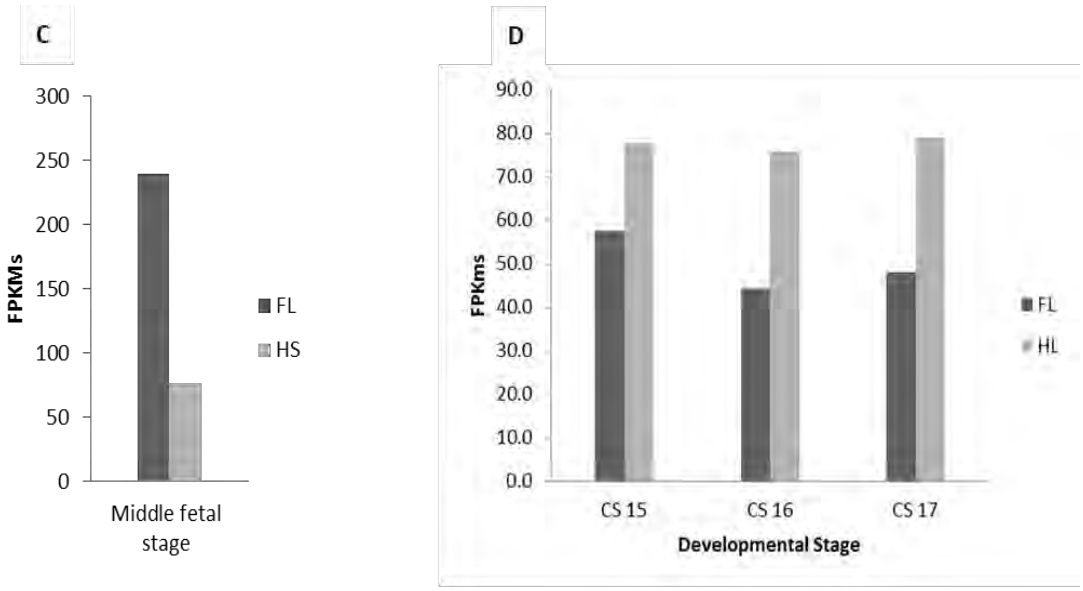


Figure 3.7 A) *M. schreibersii* CS 14 *Lef1* expression B) *M. schreibersii* pooled CS15-17 expression. C) *M. schreibersii* middle fetal stage expression patterns from metacarpal tissue (FL) and metatarsal tissue (HS). D) *M. natalensis* RNA-seq FPKMs for *Lef1* in FLs and HLs at CS 15, CS 16 and CS 17.

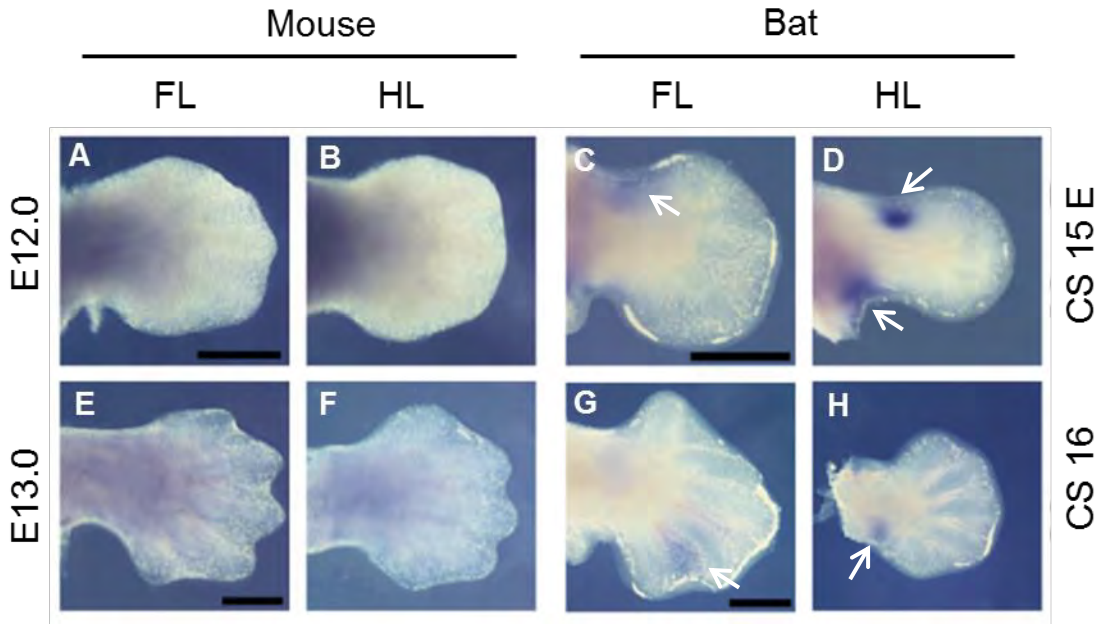


Figure 3.8 *Lef1* expression patterns in the FL and HL autopods of *M. natalensis* embryos from CS 15E and CS 16, and mouse embryos from equivalently staged E12.0 and E13.0 At CS 15E there is a comet of bright staining in the proximal and anterior–most region of the HL autopod (D). No expression is detected in mice at equivalent stages. Scale bars are standardised to 0.5mm.

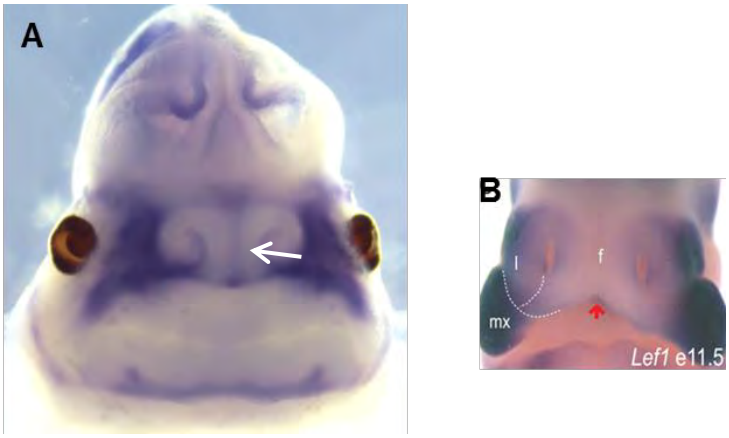


Figure 3.9 Photograph of a A) *M. natalensis* CS 16L face showing similar *in situ* staining to *Lef1 B)* E11.5 mouse. A) There is clear exclusion of *in situ* signal in the frontonasal prominence (white arrow), surrounded by dark *in situ* signal in the maxillary. B) Mouse facial region, where a similar *Lef1* pattern of expression can be observed with dark staining in the maxillary (mx) and exclusion of staining in the frontonasal region (mouse *in situ* taken from Brugmann et al. 2007)**.**

Lef1 Expression Pattern in Developing Mouse Limbs

Despite the successful cloning of *Lef1* from mouse head RNA, and generation of a mouse *Lef1* probe, no gene expression was detected in mouse limbs at either of the observed stages E12.0 and E13.0.

Lef1 Expression Pattern in Developing Bat Limbs

At CS 15E there are two clear regions of expression in the FLs: in the arch of the anterior zeugopod in the developing propatagium, and a small region in the join of the posterior region of the autopod and the plagiopatagium (Figure 3.8 C). The HLs display a distinctly stronger signal in the anterior zeugopod region as well as a region in the posterior zeugopod that will later develop into the calcar (Figure 3.8, D). At CS 16 there is light expression visible in the interdigital tissue between digit IV and V in the FL (Figure 3.8, G), the presumptive chiropatagium, and in a comet in the forming Calcar of the HL (Figure 3.8, H). There is faint staining in the plagiopatagium in a region above the HL (Figure 3.10, B), however; overall expression is faint in both FLs and HLs. Seeing as *Lef1* had been shown to display distinctive facial staining, the head of the CS 16 was used as a positive control. The facial region displayed a robust signal in the maxillary, as well as exclusion from the frontonasal protrusion (Figure 3.9, A arrow).

This is the expected and recorded pattern of *Lef1* expression in mice (Figure 3.9, B) (mouse *in situ* taken from Brugmann et al. 2007)**.**

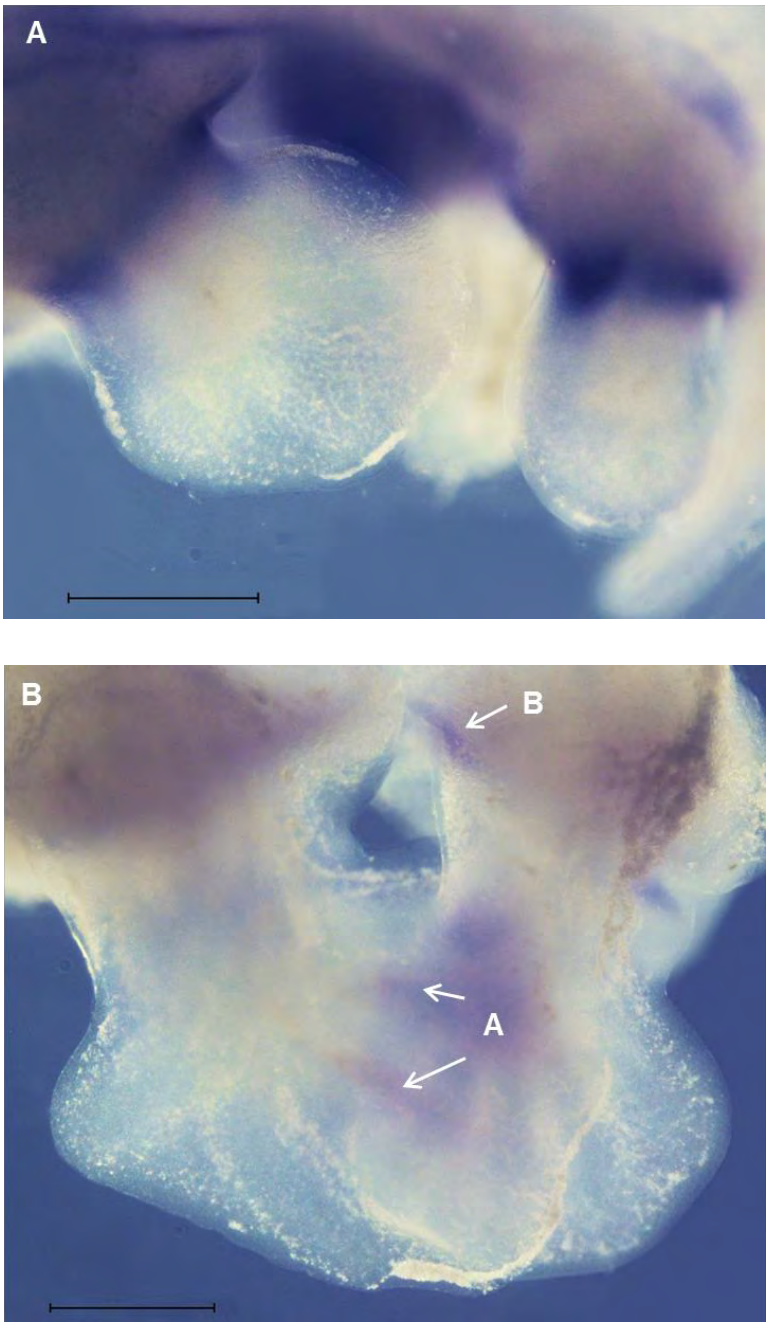
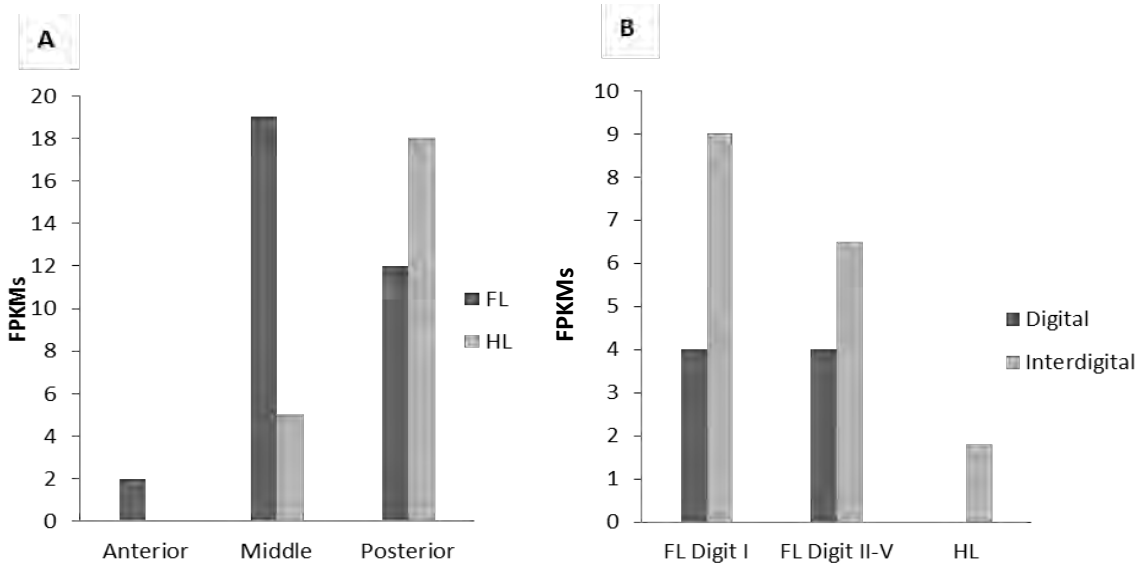


Figure 3.10 Side-view photographs of *Lef1* WISH conducted on *M. natalensis* embryos. These side-view images are from the same embryo as the montages and directly compare staining between the FL and HL. A) Side-view of a CS 15E embryo B) A side-view of a CS 16 embryo with the limbs with arrows indicating the regions of *in situ* signal. Scale bars are standardised to 0.5mm.

3.2.2.2. Lhx8

In the *M. schreibersii* CS 14 dataset there is very low expression in the FL anterior region, with stronger expression in the middle segment. In the HL there is no expression the anterior domain, with expression increasing from the middle segment to the posterior region. In the *M. schreibersii* CS 15-17 dataset there is higher expression in the interdigital tissues in the FL compared to the digit tissue, with expression being the strongest in the interdigital tissue of digit I. Expression is non-existent in the HL digit tissue, and very low in the digit tissue. Expression in the middle fetal stage is not shown as it was very low (below 2 FPKMs). In the *M. natalensis* dataset Lhx8 is expressed at much higher levels in the FL compared to the HL at all 3 stages, with expression increasing in the FL from CS 15-17 (Figure 3.11, C). Although overall expression of *Lhx8* is low, the fold change between FL and HL is one of the greatest recorded in the *M. natalensis* dataset (Eckalbar et al., manuscript under review).



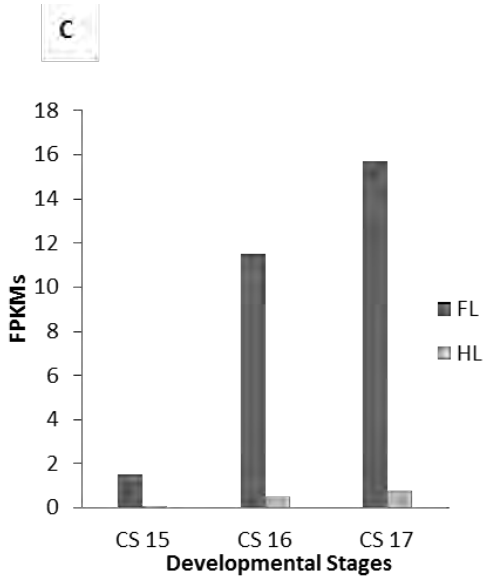


Figure 3.11 A) *M. schreibersii* CS 14 *Lhx8* expression patterns B) *M. schreibersii* pooled CS15-17 expression patterns. C) *M. natalensis* RNA-seq FPKMs for *Lhx8* in FLs and HLs at CS 15, CS 16 and CS 17.

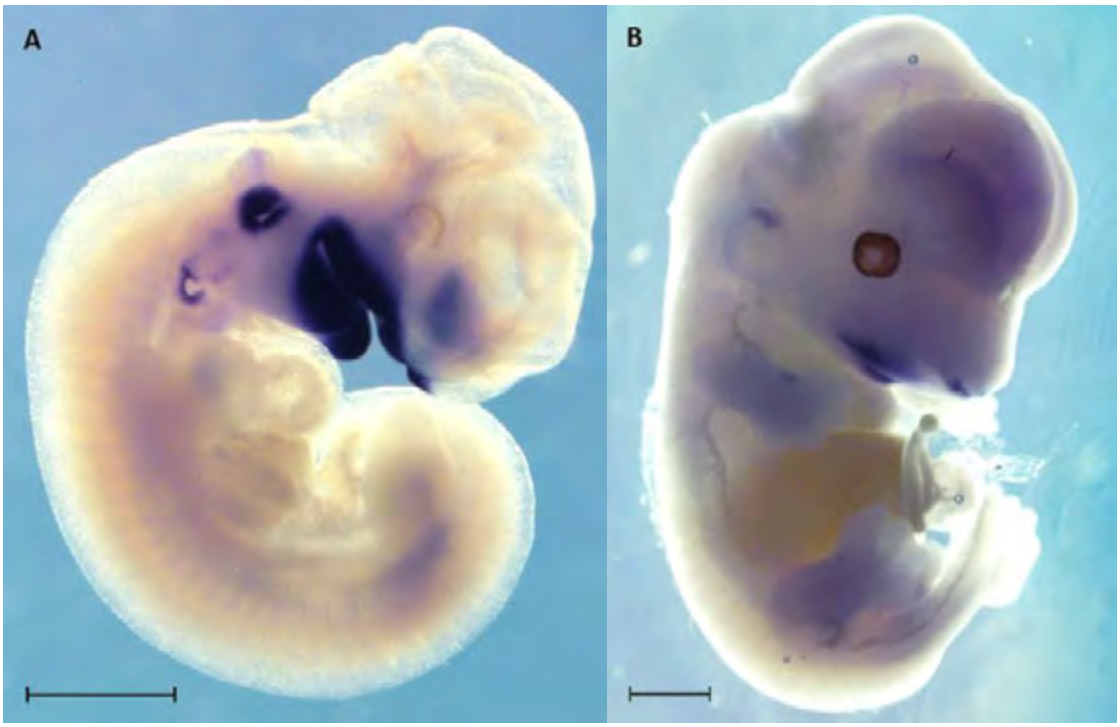


Figure 3.12 Side-view photographs of *Lhx8* WISH conducted on E10.5 and E13.0 mice, these side-view images are from the same embryo as the montages and directly compare staining between the FL and HL A) Embryo showing the efficacy of the mouse *BLhx8 in situ* probe in binding the expected region of the maxillary arch. B)

MLhx8 probe binding to the maxillary arch and nasal cavity, but expression is lacking in the limbs. Scale bars are standardised to 0.5mm.

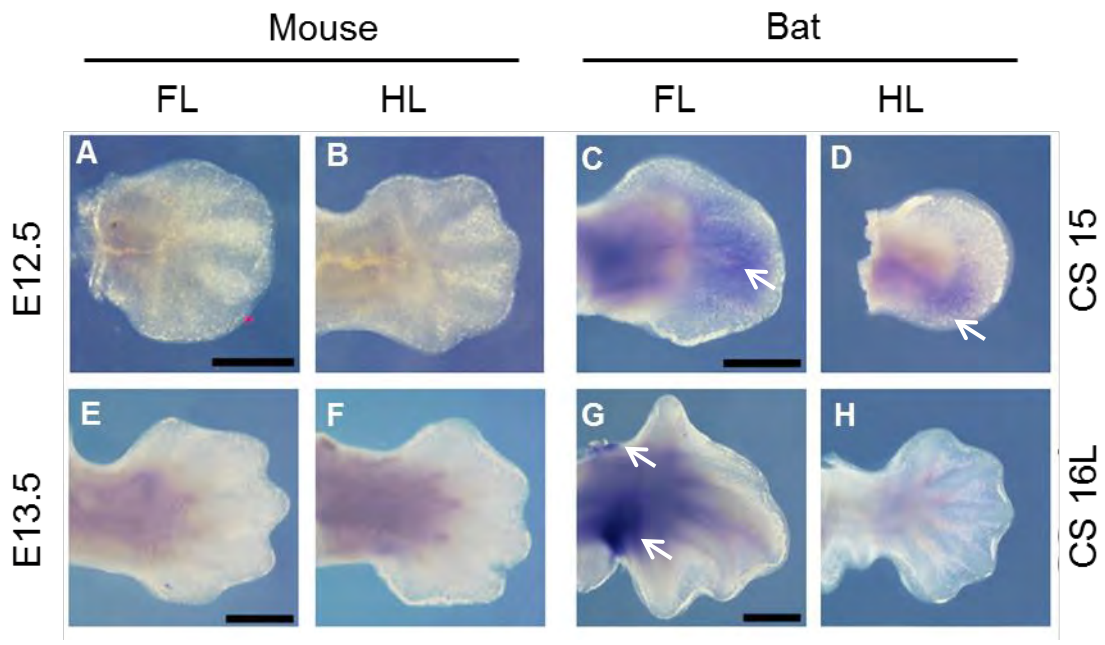


Figure 3.13 *Lhx8* expression patterns in the FL and HL autopods of *M. natalensis* embryos from CS 15 and CS 16L, and mouse embryos from equivalently staged E12.5 and E13.5 Scale bars are standardised to 0.5mm.

Lhx8 Expression in Developing Mouse Limbs

Lhx8 has been reported to be important for craniofacial development (Landin Malt et al. 2014b), the *M. natalensis* probe was used as a positive control on E10.5 and E13.0 embryos. The *M. natalensis Lhx8* probe displayed a robust signal in the presumptive mandibular arch region as reported in the literature (Figure 3.12 A and B); confirming the identity and ability of the probe to bind *Lhx8* mRNA, as well as the ability of the probe to penetrate the mouse tissue. When the BLhx8 probe (data not shown), as well as a mouse specific MLhx8 probe was tested on mice of later developmental stages there was no expression of *Lhx8* in the limbs (Figure 3.12 A-B; E-F).

Lhx8 Expression in Developing Bat Limbs

At the earlier stage (CS 15) there is clear, light staining in the zeugopod of the FL and in a triangle of expression from the beginning of the metacarpals covering presumptive digits II to IV (Figure 3.13 C). The HL shows expression in the posterior region of the zeugopod and autopod (Figure 3.13 D). By CS 16L there is abundant expression in the

region between the zeugopod and the propatagium, as well as a region the autopod and plagiopatagium. There is faint staining in the interdigital regions between digit III and IV, and IV and V (Figure 3.13, demarcated by white arrows), while expression is absent from the HLs (Figure 3.13 H).

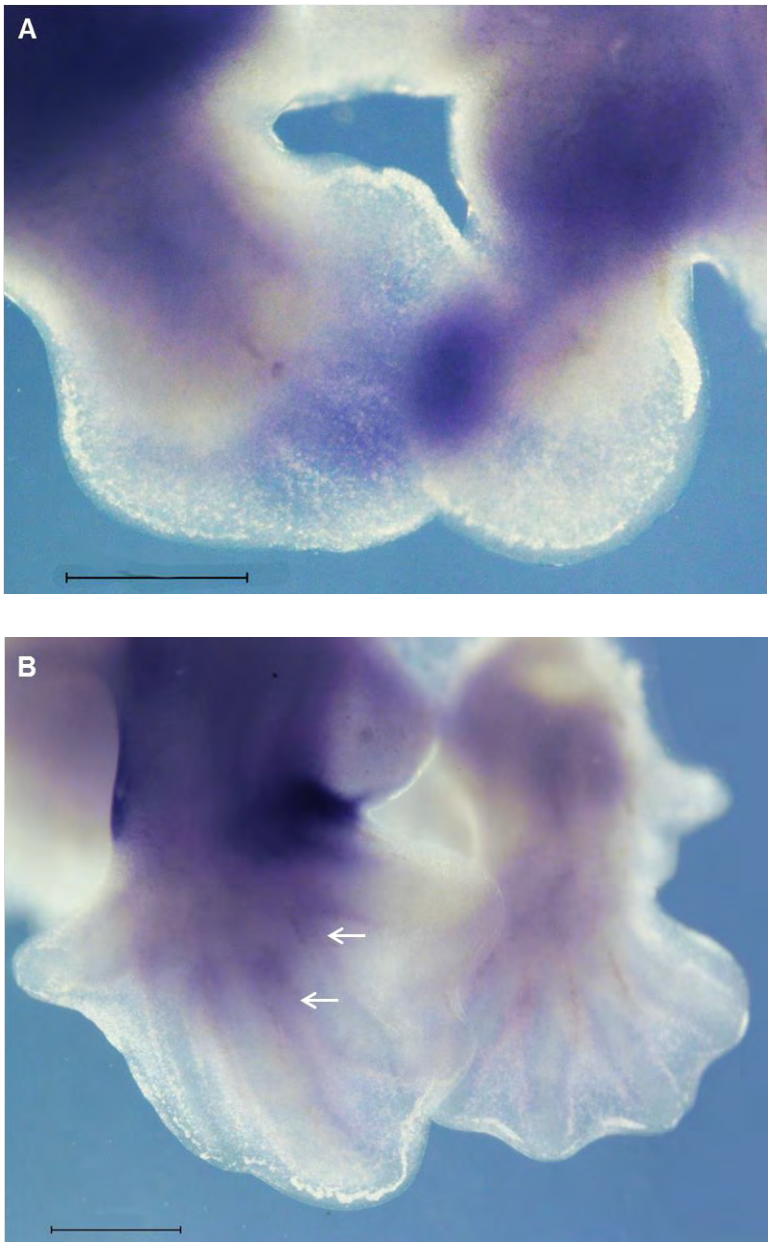
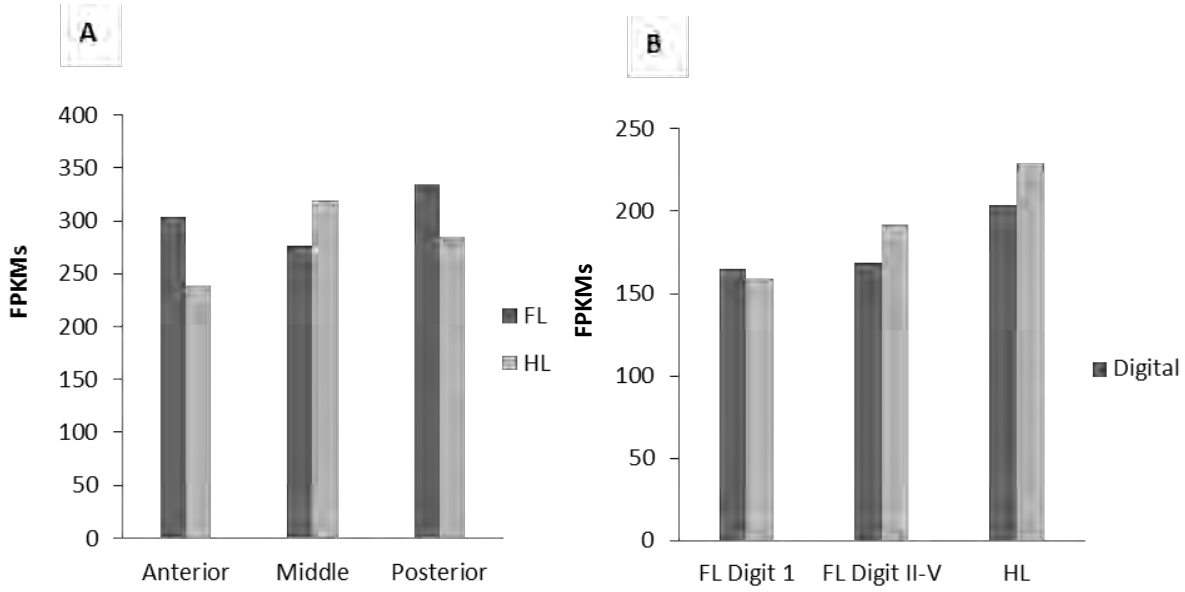


Figure 3.14 Side-view photographs of *Lhx8* WISH conducted on A) CS 15 B) CS 16L *M. natalensis* embryos. These side-view images are from the same embryo as the montages and directly compare staining between the FL and HL A) Side-view of a bat CS 15 bisected embryo showing the concentration of the *Lhx8 in situ* probe. B) Side-view of a bat CS 16L bisected embryo showing the concentration of the *Lhx8 in*

situ probe in the join between the FL autopod and the Plagiopatagium. Scale bars are standardised to 0.5mm.

3.2.2.3. HoxA10

At CS 14 in the *M. schreibersii* dataset (Wang et al. 2014), expression is comparable between FL and HL, with slightly elevated signal in the anterior and posterior regions of the FL compared to the middle domain of the FL, while expression is strongest in the middle domain of the HL (Figure 3.15 A). At CS 15-17 in the *M. schreibersii* dataset, FPKMs are half as abundant as FPKMs at CS 14 (Figure 3.15 A vs B). At this patterning phase expression of *HoxA10* is similar between the digit and interdigital tissue in the FL and HL, with expression being marginally higher in the HLs compared to FLs (digits i-v) (Figure 3.15 B) At the middle fetal stage, expression is higher in the FL metacarpals compared to the HL metatarsals. In the *M. natalensis* dataset, expression is higher in the HL compared to the FL (Figure 3.15, D), however; this relationship is not reflected in the fold change: CS 15: -0.29; CS 16: -0.27; CS 17: 0.4 and p-values: CS 15: 0.47; CS 16: 0.37 and CS 17: 0.49.



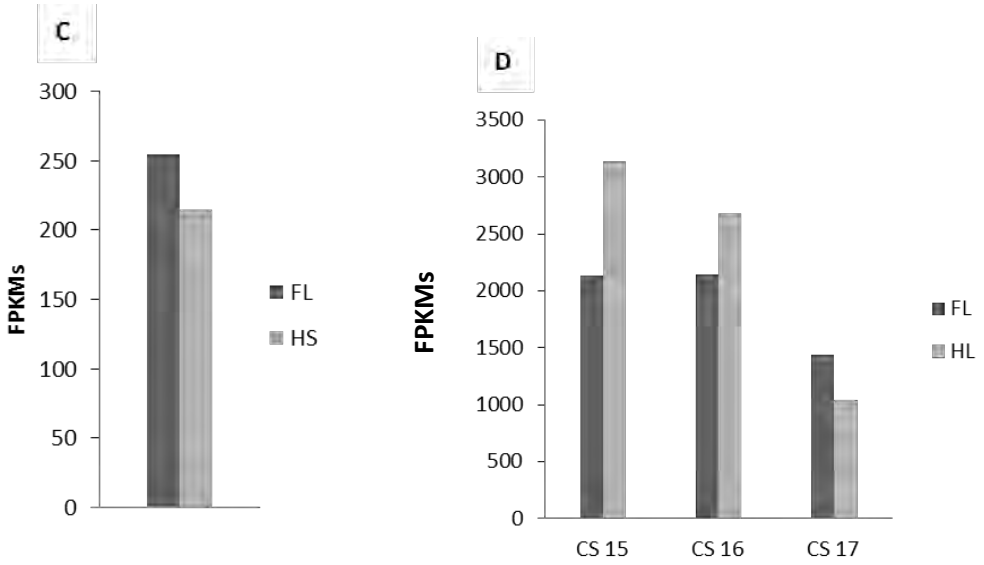


Figure 3.15 A) *M. schreibersii* CS 14 *Lef1* expression patterns B) *M. schreibersii* pooled CS15-17 expression patterns. C) *M. schreibersii* middle fetal stage expression patters from metacarpal tissue (FL) and metatarsal tissue (HS). D) *M. natalensis* RNA-seq FPKMs for *Mllt3* in FLs and HLs at CS 15, CS 16 and CS 17.

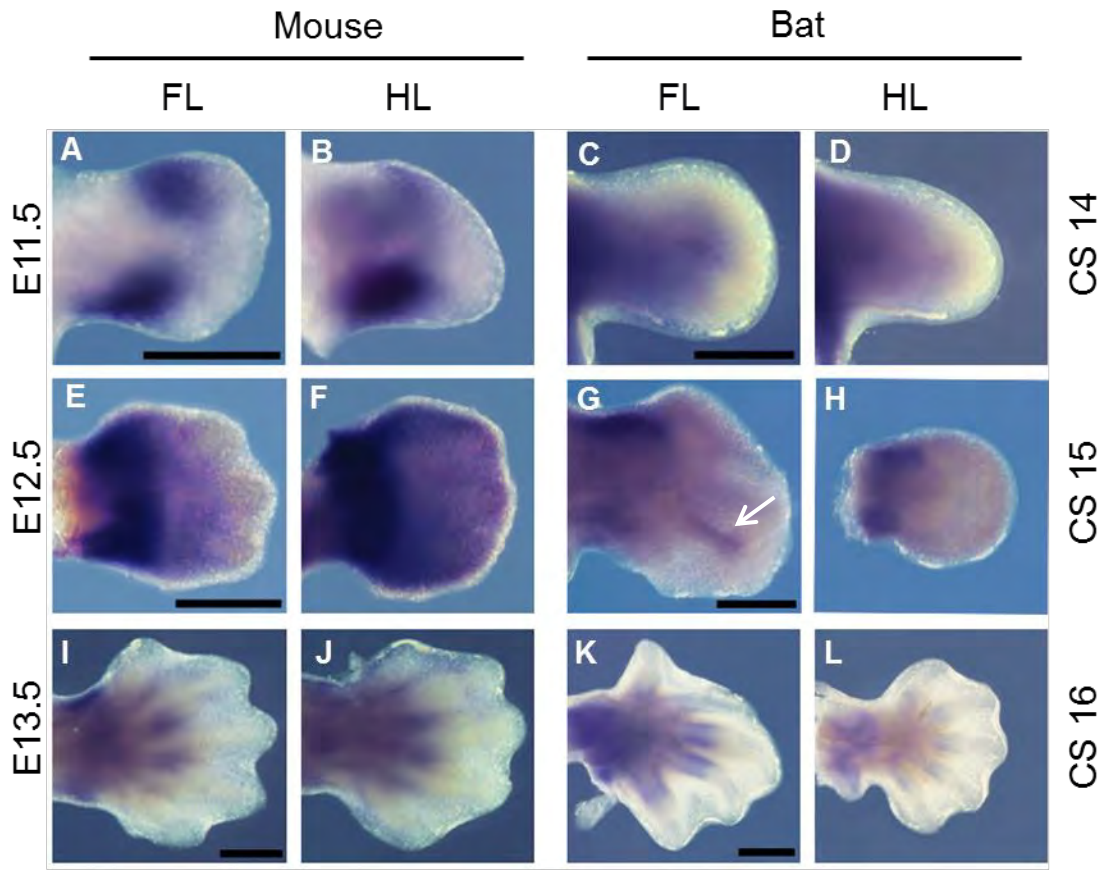


Figure 3.16 *HoxA10* expression patterns in the FL and HL autopods of equivalently staged mouse and *M. natalensis* embryos. At E11.5 there is clear staining in the FL and HL zeugopod, with very little staining in the autopod region. The HL has stronger signal in the posterior region of the zeugopod compared to the anterior region. Scale bars are standardised to 0.5mm.

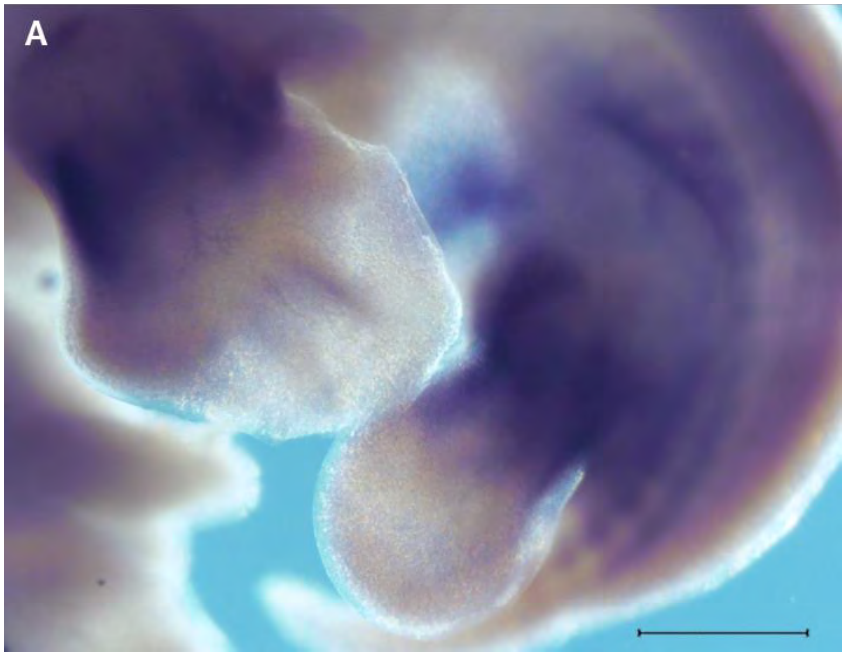
HoxA10 Expression in Developing Mouse Limbs

At E11.5 there is distinct staining in the zeugopod regions in both FL and HLs, with signal being particularly robust in the proximal zeugopod region of the HL. Expression is absent from the distal autopod region in both FL and HLs (Figure 3.14, A and B). E12.5 displays similar staining pattern to that reported (data not shown), with expression being stronger in the HL autopod relative to the FL autopod, with a stronger signal concentrated in the zeugopod region (Figure 3.14, E and F). At E13.5 staining is comparable between FL and HL, with signal having dissipated in the distal autopod in FL and HLs. There is also clear staining restricted to the proximal interdigital regions,

between the base of the presumptive metacarpals and metatarsals (Figure 3.14, I and J).

HoxA10 Expression in Developing Bat Limbs

At CS 14 there is signal in the proximal region of the limb bud with expression being excluded from the distal region. At CS 15 expression is restricted to the zeugopod region in both FL and HL with a distinctive area of expression in the IV digit ray. There is evenly spread low expression in the FL and HL autopods, aside for a strip of higher expression in digit ray IV. In the side-view image it is evident that there is comparable expression between the FL and the HL (Figure 3.15, A). By CS 16 there is clear expression in the zeugopod region of the FL and HL, with clear expression in the mesenchymal condensation where the metacarpals of digits II- V will form. Expression is lacking in the HL autopod, but is evident in the HL zeugopod.



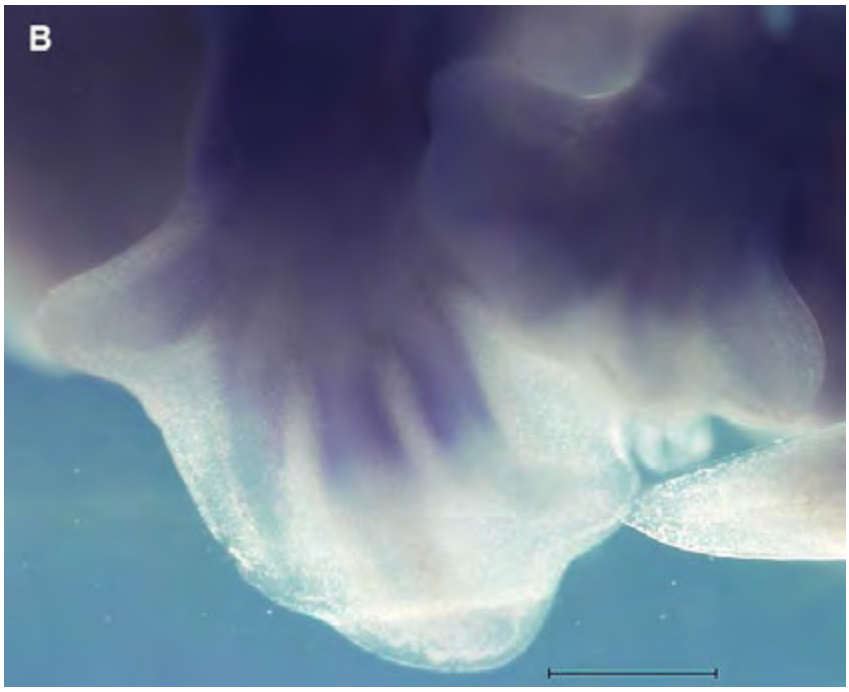


Figure 3.17 Side-view photographs of *HoxA10* WISH conducted on *M. natalensis* embryos. These side-view images are from the same embryo as the montages and directly compare staining between the FL and HL. A) A CS 15 embryo B) A CS 16 FL and HL that has been over-stained, this is apparent from the purple flesh of the trunk.

qPCR of the *HoxA10* gene locus

An inconsistency was discovered in the RNA-seq tracks in a region just past the 3' region of the *HoxA10* gene locus. This region displayed a higher signal in CS 15 FL, a higher signal in the CS 16 HL and a high signal in the CS 17 FL (Figure 3.3) (Eckalbar, manuscript under review). To investigate this rapidly fluctuating signal, absolute qPCR using the LRE method, was conducted on this region (primer set 1) as well as on a region within the *in situ* probe (primer set 3) to use as a comparison. Only two stages were selected for comparison to conserve precious *M. natalensis* RNA, and is sufficient to compare the outlier which is CS 16 HL expression. MIQE standards were adhered to, to ensure consistency of the qPCR methods by describing the minimum information necessary for evaluating qPCR experiments (Bustin et al. 2009).

Various verification steps were taken when conducting qPCR. Following the first run, the 20µl reaction was run on an agarose gel to ensure that the primers were amplifying

amplicon of the expected size (data not shown). Of the three primer sets designed, only two sets were selected for qPCR, as the third was not able to yield curves that reached 100% efficiency necessary for LRE absolute quantification (Rutledge 2011). Primer concentration, cDNA concentration and λ DNA curves were all optimized in separate runs. The final run contained both primer sets, thus the melt curves were observed for contamination prior to the experimental run (Appendix 1.3).

LRE Analysis

In order for the LRE to accurately predict the absolute values reliably, the efficiency of each reaction has to be as close to 100% as possible (with a 5% margin on either side). With optimization of the reactions, each biological repeat managed to reach this level of efficiency. There was a great amount of variability amongst biological samples in the 3' region of the *HoxA10* gene locus that can be seen with the disproportionately large standard deviation between HL biological samples (Figure 3.19, A). This large standard deviation is not a matter of technical repeat variation, or variation in cDNA starting concentration. A possible reason for this variable expression is that this region of UTR is highly variable amongst individuals, as there is less standard deviation among the biological samples in the *in situ* region amplified by primer set 3 (Figure 3.19, B).

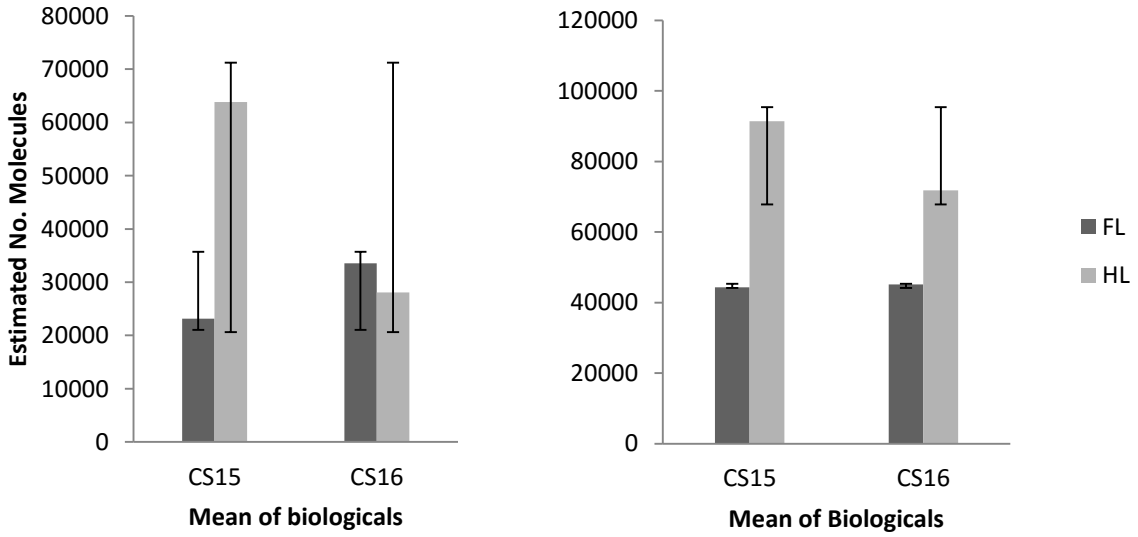
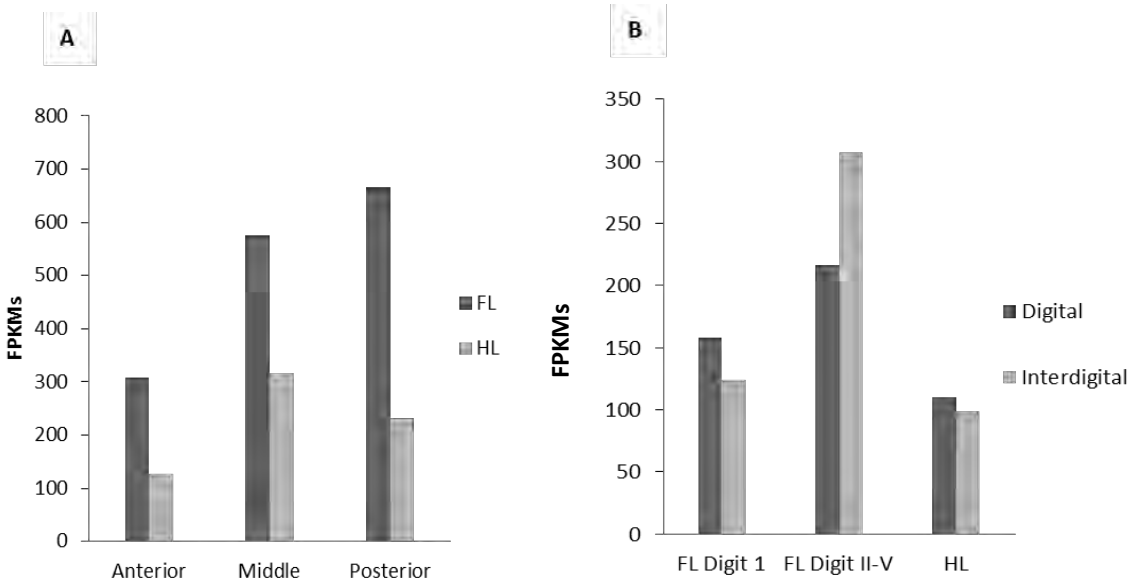
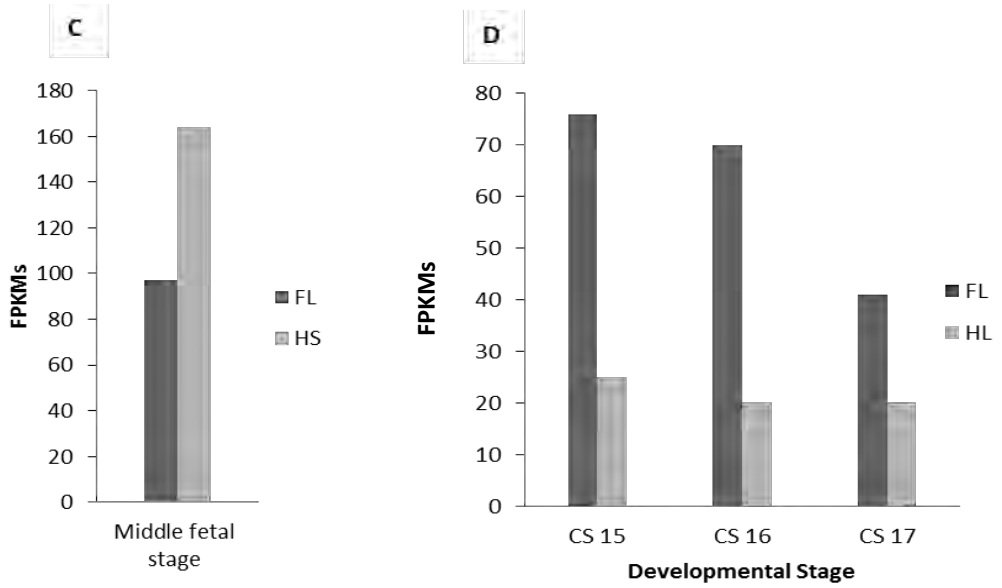


Figure 3.19 LRE qPCR absolute prediction of molecules of A) a region 3' of the *HoxA10* gene locus B) A region within the *HoxA10 in situ* probe.

3.2.2.4. *Mllt3*

The *M. schreibersii* ‘gene’ (comp5651) aligned to the *M. natalensis Mllt3* gene. At CS 14 expression is consistently higher in the FL compared to the HL, with expression increasing from the anterior domain to the posterior domain. In the HL, expression is highest in the middle domain. Overall FPKMs decrease from CS 14 to the patterning phase with FL expression being higher than the HL, and expression being considerably higher in the digit II-V region than digit I. The interdigital tissue of the FL displays higher expression than the digit II-V tissue. In the middle fetal phase (FL and HS) that has sampled metacarpal and metatarsal tissue, the pattern switches and is considerably higher in the HL compared to the FL, this could be due to an exclusion of Mllt3 in the FL digits at this elongation stage. In the *M. natalensis* dataset, the fold change between FL and HL is highest at CS 15, with the difference decreasing with each stage. It is possible that expression of *Mllt3* rapidly increases in the FL compared to the HL at CS 15, and this is not reflected in the *M. schreibersii* dataset as the sampling for the patterning phase is skewed towards CS 17.





Graphs 3.20 A) *M. schreibersii* CS 14 *Mllt3* expression patterns B) *M. schreibersii* pooled CS15-17 expression patterns. C) *M. schreibersii* middle fetal stage expression patters from metacarpal tissue (FL) and metatarsal tissue (HS). D) *M. natalensis* RNA-seq FPKMs for *Mllt3* in FLs and HLs at CS 15, CS 16 and CS 17.

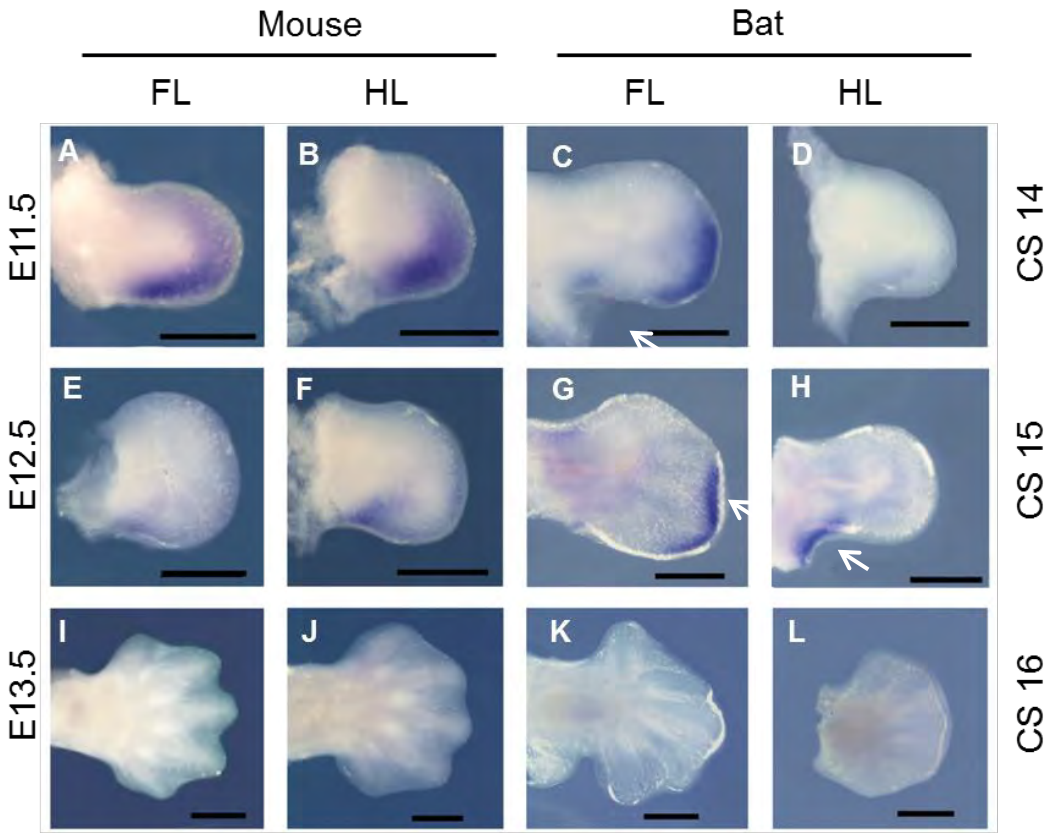


Figure 3.21 *Mllt3* expression patterns in the FL and HL autopods of *M. natalensis* embryos from CS 14, CS 15 and CS 16, and mouse embryos from equivalently staged E11.5, E12.5 and E13.5. There is a novel *Mllt3* expression domain present in CS 14 and CS 15 FL autopods that is not present in HLs or in mice. This domain of expression subsides by CS 16. Scale bars are standardised to 0.5mm.

Mllt3 Expression in Developing Mouse Embryos

Expression at E11.5 is highly similar between FL and HLs with expression following the outer edge of the limb buds and being concentrated in the posterior region (Figure 3.21 A and B). By E12.5 expression is restricted to a slight smear in the proximal posterior-most region of both FL and HL (Figure 3.21 E and F). By E13.5 expression of *Mllt3* is absent from the mouse autopods.

Mllt3 Expression in Developing Bat Embryos

At CS 14 there is a region of strong expression in the posterior-distal domain of the FL in the region where there will be expansion, as well as in a small region where the zeugopod meets the presumptive plagiopatagium (Figure 3.21 B) while expression is

barely visible in the HL (Figure 3.21 C). At CS 15 there is robust expression restricted to the posterior distal mesenchyme running from presumptive digit II to V (Figure 3.21 G; and Figure 3.22). Expression in the HLs is restricted to the most proximal and posterior region of the autopod, where the calcar will later develop (Figure 3.21 H). At CS 16, expression is faint in a region between digit IV and V (Figure 3.21 K), and is absent from the HLs.

Figure 3.22 Side-view photograph of *Mllt3* WISH conducted on a CS 15 *M. natalensis* embryo, the same embryo that was used in the in situ montage.There is clear staining in the differentiating somites and in the distal tip of the FL (digit III- V) and presumptive calcar of the HL**.** Scale bars are standardised to 0.5mm.

3.3 Discussion

3.3.1 Novel region of *Lef1* expression in the HLs

At CS 15 RNA–seq results show differential expression between the FL and the HL, with expression being higher in the HL (p-value 0.043). This differential expression decreases by CS 16 where there is no longer a significant difference between the FL and the HL (p-value: 0.12), very little differential expression at CS 17 (p-value: 0.37). The expression pattern revealed by the RNA-seq is consistent with the *in situ* results, with CS 15E displaying a strong *Lef1* signal in a 'comet' in the proximal anterior region of the HL (Figure 3.8 D). Wnt ligands are secreted from the limb bud ectoderm and block cartilage formation in the periphery of the limb bud via the β-catenin pathway. *Lef1* is a known Wnt/ β-catenin receptor (Kozhemyakina, 2015), and its higher expression level in the HL relative to the FL could suggest suppression of the canonical Wnt/ β-catenin in the FLs relative to the HLs, thus cartilage formation is permitted in the FL, accounting for the elongated digits. This connection may be a stretch as the region of *Lef1* expression in the HL is not over the digit region, thus suppression of *Lef1* expression in the proximal anterior region of the FL would not affect digit elongation. *Lef1* expression has been detected in mice FLs at early stages of development (E9.5), it would be potentially interesting to see what the expression pattern is in bat limb buds at the equivalent stage.

By CS 16 the *Lef1* signal is low and visibly similar between FL and HL (Figure 3.8 G and H; Figure 3.10 B). This faint staining could be viewed as a technical issue considering that the RNA-seq displays a similar signal in CS 15 and 16 HLs, with a larger difference between the FL and the HLs at CS 16. This was not the case as the head of the CS 16 sample that was used as a positive control received the same storage, fixation and *in situ* treatment as the limbs, and displayed strong signal and a clear resemblance to mouse *Lef1* expression in the head. This suggests that the probe had efficiently penetrated the tissue (Figure 3.9, A) (Brugmann et al. 2007).

3.3.2 Novel Expression of *Lhx8*

Lhx8, a known regulator of neuronal development (Zhao et al. 2003). WISH of CS 16L FLs showed localized *Lhx8* expression in the posterior portion of the wrist region, specifically in the junction between the base of digit V and the plagiopatagium, while no expression was detected in HLs and mouse limbs.

The RNA-seq shows no significant difference in expression between FL and HL at CS 15 (Figure 3.11, C), while there is a rapid increase in expression in the FL at CS 16, with significantly higher expression in the FL autopod compared to the HL (p-value 0.0035). This pattern continues as the $\log_2$ fold change increases from 2.34 (CS 16) to 3.86 (CS 17). The *in situ* results do not reveal a novel domain of strong expression in the autopod to account for the high fold change at CS 16 and 17. Expression of *Lhx8* is strongest in the most proximal region of the autopod (wrist) (Figure 3.11, G) especially along the anterior and posterior edges of the limb.

The elevated FL autopod signal could be attributed to inclusion of zeugopod tissue in the RNA-seq sample when the autopods were being dissected from the zeugopod. Dissection of the autopod in the field is a challenging task. Removal of all of the zeugopod tissue is not always possible given the equipment and time constraints of gathering many different sample types in a short space of time without degradation of the samples. Despite the probable contamination of autopod tissue with zeugopod tissue, the expression domains present are completely absent from mice limbs.

Lhx8 is a known regulator of neuronal development (Flandin et al. 2011). Bats evolved a unique set of nerves in their patagia as a specialisation to handle complex aerial manouvers it is possible that *Lhx8* could be involved in this evolution (Flandin et al. 2011). The complete absense of *Lhx8* expression from mice limbs versus the strong signal in the bat implies a unique adaptation for this gene. The stage of the expression implies that that *Lhx8* could be responsible for nerve development but this research would benefit from sectioning to determine the expression localization in the tissue.

3.3.3 *HoxA10* expression in the metacarpals of CS 16 embryos

There is no significant differential expression between FLs and HLs at any stage. These data was available prior to selection of *in situ* candidates, however; it was selected on the basis of the importance of completing the spatial characterization of the 5' HoxA genes. The *in situs* for *M. natalensis* displayed similar patterns of expression in the FL and HL at all three stages. However; at the later stage of CS 16 there is staining in the metacarpals of digit II-V. Over-staining of the CS 16 embryo is evident in the side-view (Figure 3.17 B), the apparent dark staining evident for CS 16 FL and HLs in Figure 3.17, may be less pronounced had the embryo not been overstained, yet this pattern cannot solely be attributed to background staining.

3.3.4. Novel expression domain of *Mllt3* in Bat FLs

A novel *Mllt3* expression domain was identified in *M. natalensis* embryos at early stages CS 14 and CS 15 that is not present in bat HLs, or mice autopods of the equivalent stage (Figure 3.21 C and G). This novel expression domain exists in the distal region of the CS 15 limb bud slated to develop into digit II-V that will later undergo rapid elongation. This pattern of expression complies with the expression profiles of both the *M. natalensis* and *M. schreibersii* datasets, which exhibit significantly higher expression in the FL compared to HLs at CS 15 (*M. natalensis*: p-value 0.00053). The *in situ* results for CS 16 do not fit the RNA-seq data as there is faint *in situ* signal at CS 16, compared to CS15. For this reason, the *in situ* was repeated on another CS 16 (more recently acquired and fixed embryo, 2012), using the probe that was shown to work on the CS 15, however; the staining remained the same (data not shown).

There is a scarcity of literature about the role of *Mllt3* during limb development; which lead our collaborators at UCSF to generate transgenic homozygous *Mllt3*$^{-/-}$ mutant mice to investigate the effect of reduced *Mllt3* expression on axial patterning, as *Mllt3* null mutant mice exhibited axial defects (Figure 3.4)(Collins et al. 2002). Unfortunately bat transgenic studies are not possible at this time, so mice were used as a model. The mouse skeletal elements were stained with Alizarin red and cartilage with Alcian blue to better observe any differences in skeletal and cartilage formation. The homozygous

Mllt3 $^{-/-}$ displayed no gross skeletal abnormalities differences compared to the wildtype littermate (data not shown). To further this investigation, it would be important to explore the effect of ectopic expression of *Mllt3* on limb development, as the bat displays upregulated expression in the distal FLs.

Chapter 4

LncRNA Expression Analysis and Isoform Characterisation in Developing Bat Limbs

4.1 Introduction

With the advent of next generation genome sequencing came the discovery that mRNA transcripts that do not code for proteins (non-coding) make up a far greater contribution to the total number of genes, than previously thought (Rinn & Chang 2012). The current definition of LncRNA is a RNA gene that is 200 bp or longer (up to 10kb), that does not encode a protein. This size cut-off is somewhat arbitrary, but allows for the exclusion of smaller RNAs (Rinn & Chang 2012). The function of LncRNAs is largely unknown, however; some have been implicated in regulating development by forming ribonucleoprotein complexes (RNP), which in turn aid in regulating gene expression (Mattick & Rinn 2015).

There is low sequence similarity among LncRNAs between species suggesting a novel mechanism for these rapidly changing sequences to drive morphological variation amongst taxa (Gaiti et al. 2015). The antisense LncRNAs were more surefire candidates for real LncRNAs as they are transcribed in the opposite direction and are transcribed in close proximity to known autopod regulators.

4.1.1 Identifying LncRNAs

Obtaining artefacts when locating LncRNAs is possible; to avoid this, a number of criteria were considered before selecting LncRNA for further investigation. Identification of LncRNAs was performed by Stephen Schlebusch. The first step he took was to annotate transcripts that showed no similarity to known protein-coding genes. This was done by removing all protein-coding genes from the *M. natalensis* dataset by a one-to-one Blastp search of all transcripts to the Mammalian Unitprot database. This identified 227 potential LncRNAs, 188 of which exhibited partial sequence conservation across mammals. This list was further narrowed-down by searching LncRNA database

LncRNAdb (v2.027) which resulted in 12 matches that exhibited sequence similarity to characterized LncRNAs. Within this dataset, 8 known LncRNAs showed differential expression between FLs and HLs. The two LncRNAs: *Tbx5-as1* and *Hottip* were selected for further investigation based on two factors: 1) Both LncRNAs had a sizable fold change between FL and HL expression. 2) They both are transcribed in close proximity to known developmental regulators genes (An in-depth explanation of the identification process can be found in Appendix 1.2).

4.1.2 *Tbx5-as1*

Tbx5-as1 was discovered in the *M. natalensis* RNA-seq dataset that mapped in close proximity and in the antisense direction to T-box transcription factor (*Tbx5)* (Figure 4.1 A), a known regulator of FL bud outgrowth (Hasson et al. 2007). *Tbx5-as1* is the most differentially expressed LncRNA in the RNA-seq dataset (at CS 16), with elevated expression in the FL relative to HL across all stages. The function of *Tbx5-as1* is unknown, but given the name of antisense LncRNAs it is possible that it is involved in regulating its neighbouring gene *Tbx5* (Figure 4.1 A)

Figure 4.1 Screen-shot of the *M. natalensis* genome browser displaying the position of A) *Tbx5* transcripts relative *to* B) *Tbx5-as1* transcripts. The red arrows display the direction of transcription.

Tbx5 is involved in limb bud initiation

Tbx5 knockout mouse disrupts the development of all limb elements (Rallis et al. 2003). Prior and during limb bud initiation, *Tbx5* forms a positive feedback loop with *Fgf10* in the lateral plate mesoderm (LPM) (Agarwal et al. 2003; Rallis et al. 2003). Mutations to the human TBX5 gene are associated with a disorder known as 'Holt-Oram Syndrome' which results in FL and heart deformities (Basson et al., 1997). With a single functional copy of the TBX5 gene, the disease is fully penetrant; however, the severity of the phenotypes can vary significantly to include extended phalangeal elements of the thumb, to more severe deformities such as deletion of entire skeletal elements (Hasson et al. 2007). Despite *Tbx5* involvement in limb bud initiation, which occurs at CS 12 in *M. natalensis* embryos, the dataset displayed late *Tbx5* expression increase from CS15 to CS 17 (Figure 4.10 E and F).

4.1.3 *Hottip*

Hottip is transcribed on the 5' end of the HoxA gene cluster, and is 919bp upstream of *HoxA13* in *M. natalensis*. *Hottip* has been found to act in *cis* with Trithorax protein WDR5 (WD repeat-containing protein 5), a cofactor of the histone methyltransferase MLL1 that driving histone H3 lysine 4 trimethylation, a marker of unwound and transcriptionally active chromatin associated with gene activation (Mercer & Mattick 2013). Through chromosomal looping the Hottip-WD5-MLL complex is brought into close proximity to activate the transcription of the neighbouring 5' *HoxA* genes (Wang et al. 2011).

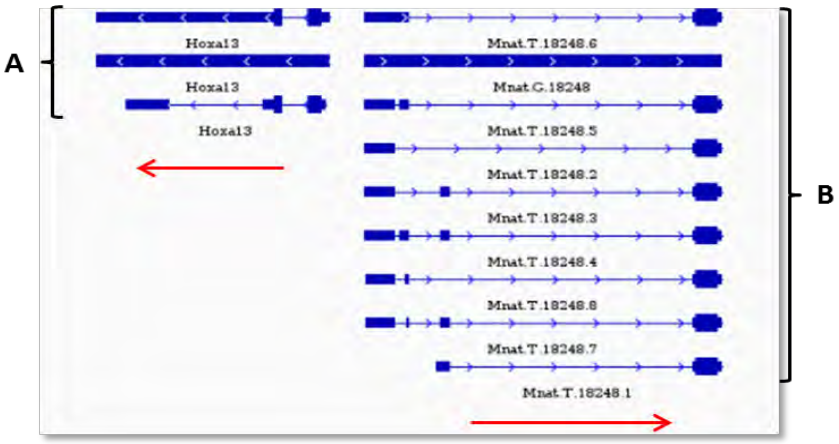


Figure 4.2 Screen-shot of the *M. natalensis* genome browser displaying the position of A) *HoxA13* transcripts relative to B) *Hottip* transcripts. Red arrows indicate transcriptional orientation.

In humans, transcription of the *Hottip* gene yields a 3764 nucleotide product that is both spliced and polyadenylated. Its exact location of initiation is 330 bases upstream of the *HoxA13* gene (Wang et al. 2011). *Hottip* knockdown by siRNA in foreskin tissue had the greatest effect on *HoxA11* and *HoxA13*, with progressively less effect on *HoxA10* and *HoxA9* (Wang et al. 2011). Functional studies were conducted on chicks to see the effect of interfering with *Hottip in vivo* (RNAi). It was discovered that the interference of *Hottip* had the greatest effect on *HoxA13* expression in the disitl tips of the limb buds.

HoxA13 and Inter-digital Programmed Cell Death

In both mice and humans, disruption to wildtype *HoxA13* causes defects in development of distal digit and interdigital tissue of the autopods (Knosp et al. 2004). *HoxA13*, like all *Hox* genes, regulates a number of processes in the genetic cascade of limb development, including chrondrogenesis, mesenchymal cell adhesion, skeletal formation, and IPCD (Knosp et al. 2004; Stadler et al. 2001; Knosp et al. 2007; Perez et al. 2010). *HoxA13* directly regulates *Bmp2 and Bmp7,* genes responsible for maintaining RA signalling in the distal autopod mesenchyme by binding to, and activating transcription of retinaldehyde dehydrogenase two (*Raldh2)* (Shou et al. 2013; Knosp et al. 2004). *Raldh2* converts retinaldehyde to retinoic acid (RA), a potent morphogen involved in inter-digital programmed cell death (IPCD) (Shou et al. 2013). In mice, heterozygous mutants (*HoxA13*$^{-/+}$) displayed reduced apoptosis between digits II

and III (Stadler et al. 2001), and homozygous mutants (*HoxA13*$^{-/-}$) display delayed and reduced *Raldh2* expression and RA signalling.

In *M. natalensis* embryos, *Raldh2* has been shown to be reduced in the interdigital tissue between digit I-III (Figure 4.3, F). In bat FL autopods, *HoxA13* has been shown to be downregulated in the FL relative to the HL (Eckalbar, manuscript under review). This downregulation could be deregulating the *Raldh2* pathway, and thus be a contributor to the retained inter-digital webbing. Reduced *HoxA13* expression in bat FL autopods could be directly regulated by a reduction in *Hottip* in the FL autopod. By uncovering the spatial expression patterns of *Hottip* and comparing these to *HoxA13*, a region of overlapping expression can be elucidated.

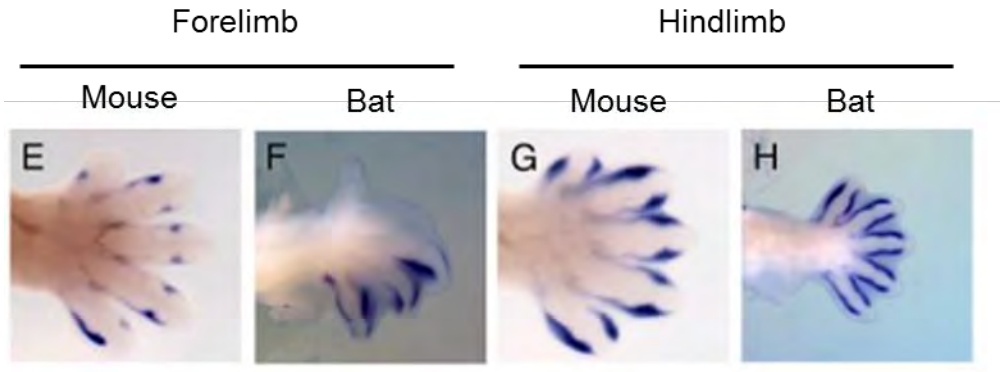


Figure 4.3 *In situ* showing expression domains of *Raldh2* in bat and mouse FL compared to the HL at CS17. *Raldh2* expression in the bat FL autopod is restricted to the tissue surrounding digits III-V, whereas *Raldh2* is expressed in the tissue surrounding all digits in the mouse FL and HL.

4.1.4 Aims

In this chapter, transcript variants for both *Hottip* and *Tbx5-as1* were characterized to ensure that the LncRNAs are not technical artefacts. This was verified by cloning and sequencing the transcripts from RNA extracted at a stage outside of the RNA-seq dataset: CS 18L.

The second element was to characterise the spatial expression domains in *M. natalensis* embryos of the LncRNAs: *Tbx5-as1* and *Hottip* by WISH, and compare this to the expression domains of the neighbouring transcription factors: *Tbx5* and *HoxA13* respectively.

4.2 Results

4.2.1 Characterising Isoforms

Tbx5-as1

Two Forward primers (F1 and F2) were design with one corresponding reverse primer (R1). F1 was designed over the 100 bp region, to amplify transcript variants 2, 3, 4, 6, 9 and 11. F2 was designed to amplify transcripts 5 and 13 (Figure 4.5). Due to the very repetitive and GC heavy sequence of *Tbx5-as1,* especially in the region over which the F1 was designed*,* the F1and R1 primer set had a very bad self-dimer ΔG score, and were not successful in amplifying isoforms despite PCR optimization (Figure 4.4 F1). Primer set F2 and R1 were successful in amplifying transcripts of 4 different sizes. Sequencing and manual alignment to the RNA-seq transcripts revealed that these transcript variants were unlike the variants predicted by the RNA-seq conducted on the early fetal stages, but aligned to the genomic sequence (Figure 4.5). It is possible that there has been alternative splicing at this later developmental stage (CS 18L) compared to the post patterning phase (CS15- CS17). The returned sequences were blasted against the *M. natalensis* genome to see whether they bind to any other region on the genome. All sequences mapped to the predicted region. The sequences were searched 20-50nt upstream of the acceptor site for branch sequences to denote splice sites, and the branch code ‘CTGAC’ was present in 4 of the returned sequences (colony E, F, H and I).

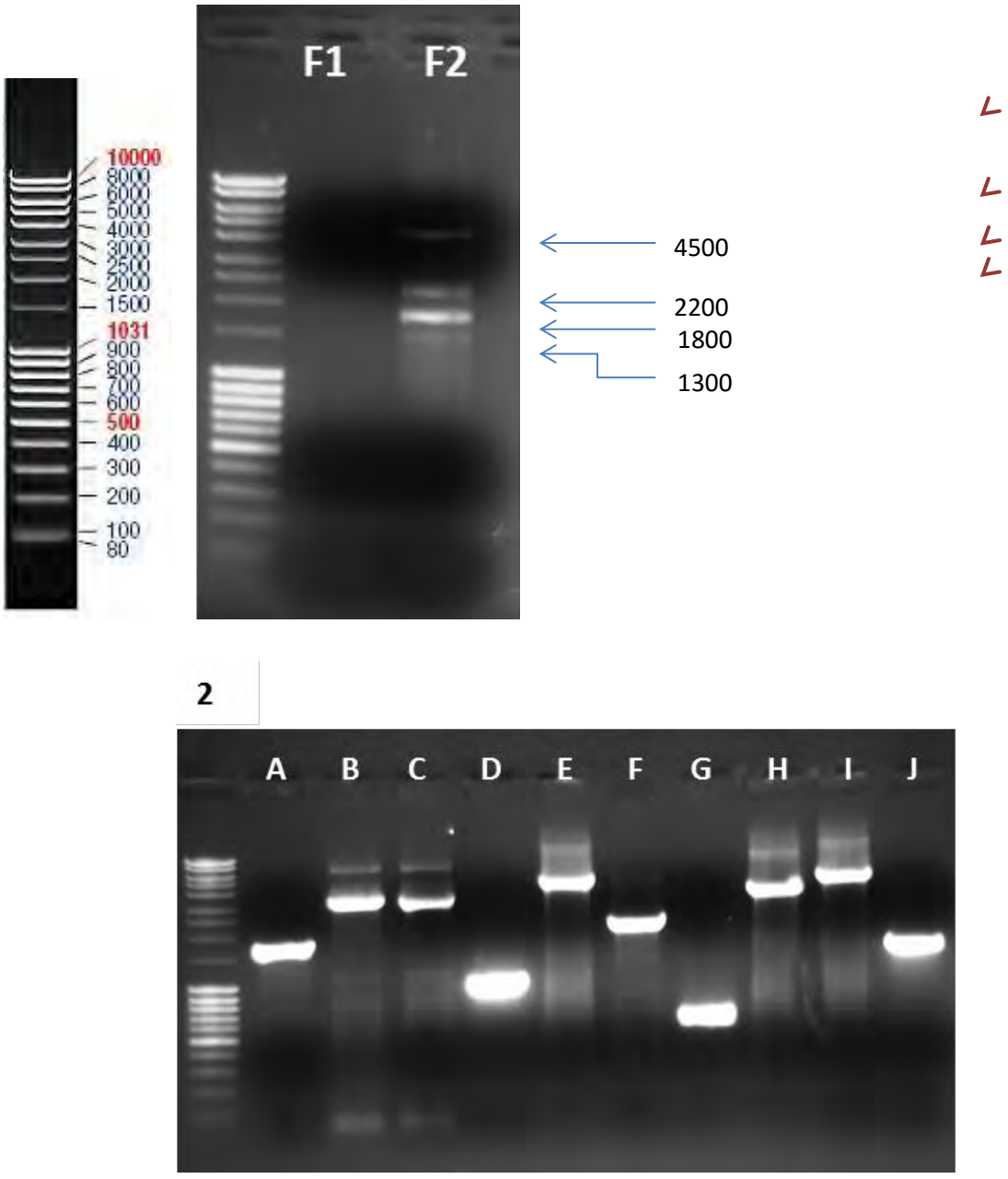


Figure 4.4 1) *Tbx5-as1* amplification from CS19 *M. natalensis* HL cDNA. 2) Colony PCR of cloned *Tbx5-as1* variants.

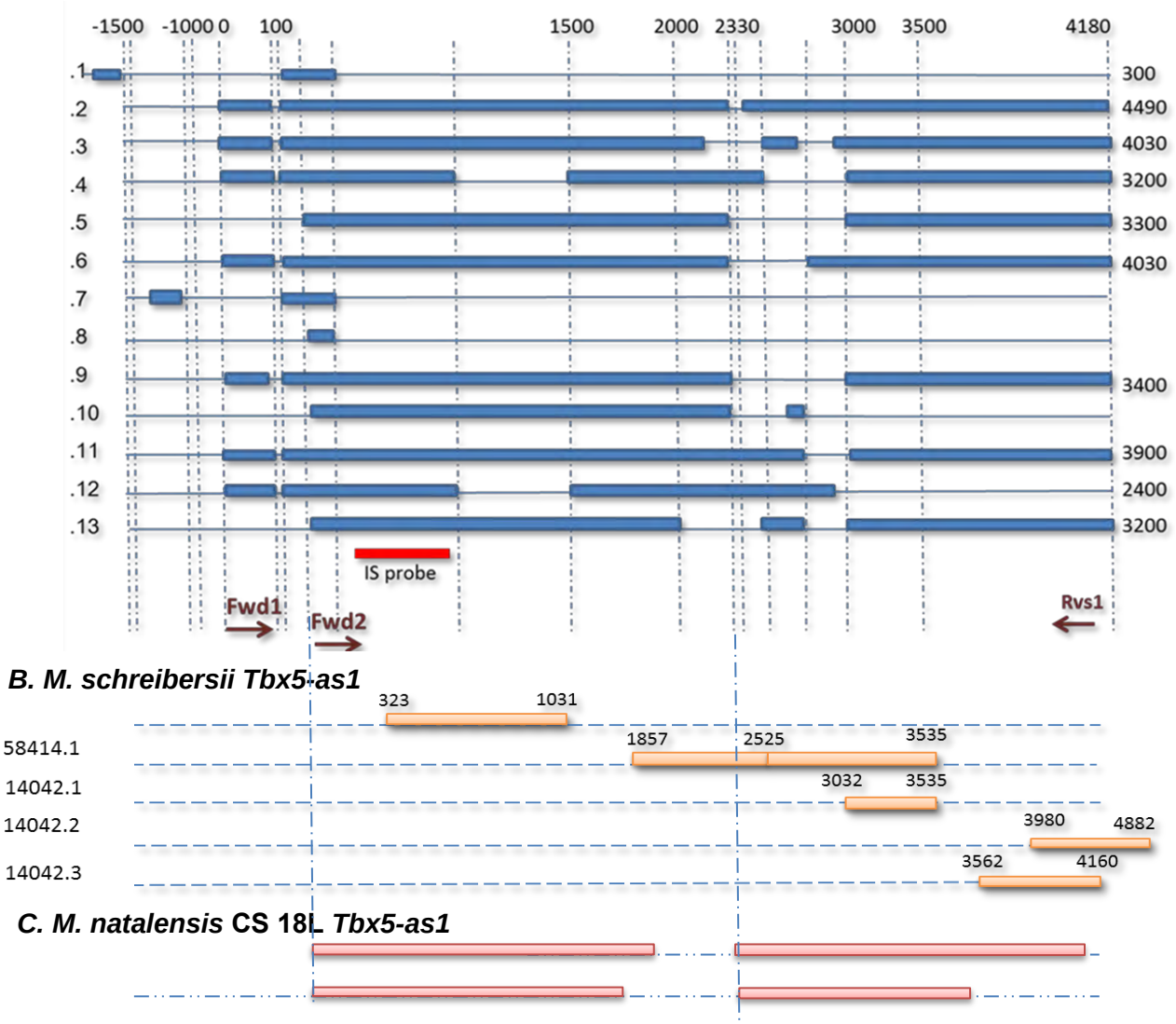


Figure 4.5 A) Aligned *M. natalensis and M. schreibersii Tbx5-as1* transcript variants from CS 15, CS 16 and CS 17 tissue, identified by RNA-seq.C) M. natalensis transcript isoforms cloned from CS 18L FL RNA A) Numbers down the left-hand side are the numbered isoforms. Numbers down the right-hand side are the lengths (bp) of each transcript variant. The red bar represents (IS probe) displays the position of the *in situ* probe. Fwd1: Forward primer 1; Fwd2: forward primer 2; Rvs1: reverse primer that matches Fwd1 and Fwd2.

Hottip

The *Hottip* primers were designed over the furthest ends of the 3' and the 5' exons of 7 of the 8 isoforms identified by RNA-seq. These isoforms were amplified and produced unexpected sizes from the CS 18L cDNA (Figure 4.6, 1). Following cloning, colony produced a number of similarly sized bands (Figure 4.6 2). This region included 6

different isoforms, some of which were difficult to identify by size by gel electrophoresis as they were very close in size (1966bp, 1979bp and 1943bp). The roughly 1700bp band appeared to be most abundant as it was the brightest band (Figure 4.6). All 14 colonies were sent for sequencing despite appearing a similar size (Figure 4.6 A-N). The colonies were sequenced and the returned sequences mapped to the genome. The transcripts displayed a unique two *Hottip* isoforms (Figure 4.7). It is possible that following the patterning phase, and in the fetal phase (CS 18L) the LncRNA has different transcript variants to perform different regulatory roles.

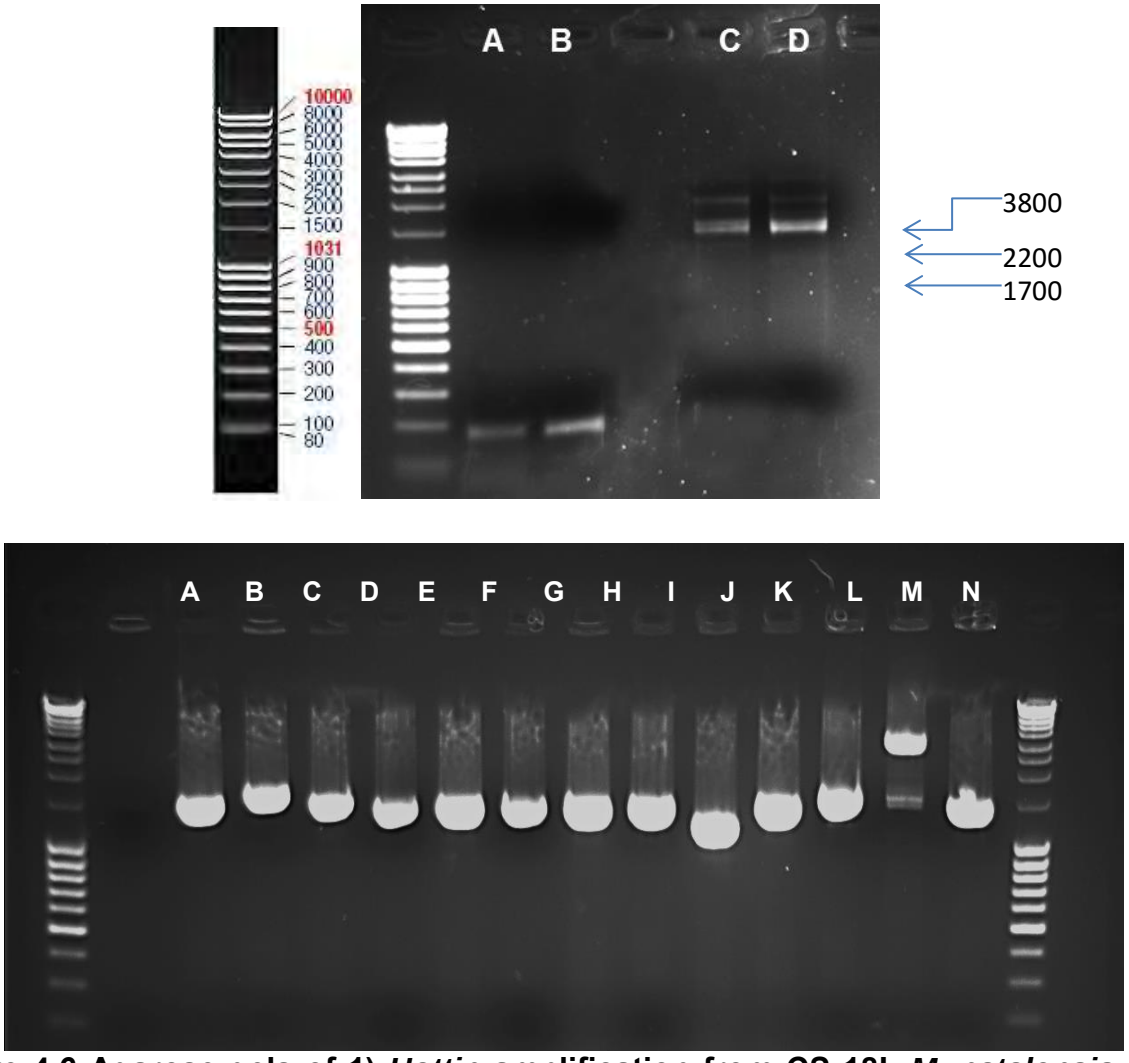


Figure 4.6 Agarose gels of 1) *Hottip* amplification from CS 18L *M. natalensis* HL cDNA, 2) Colony PCR of the cloned *Hottip* isoforms. Gel 1) A and B) Positive controls (*Tbpl1*) using two different cDNA synthesis kits. C) *Hottip* isoforms from first cDNA synthesis. D) *Hottip* from second cDNA synthesis.Gel2) Agarose gel Image of a

colony PCR used to screen for *Hottip* transcripts variants of CS 16 HL cDNA. Colonies A - L, and N were sent for sequencing to identify the isoform identity.

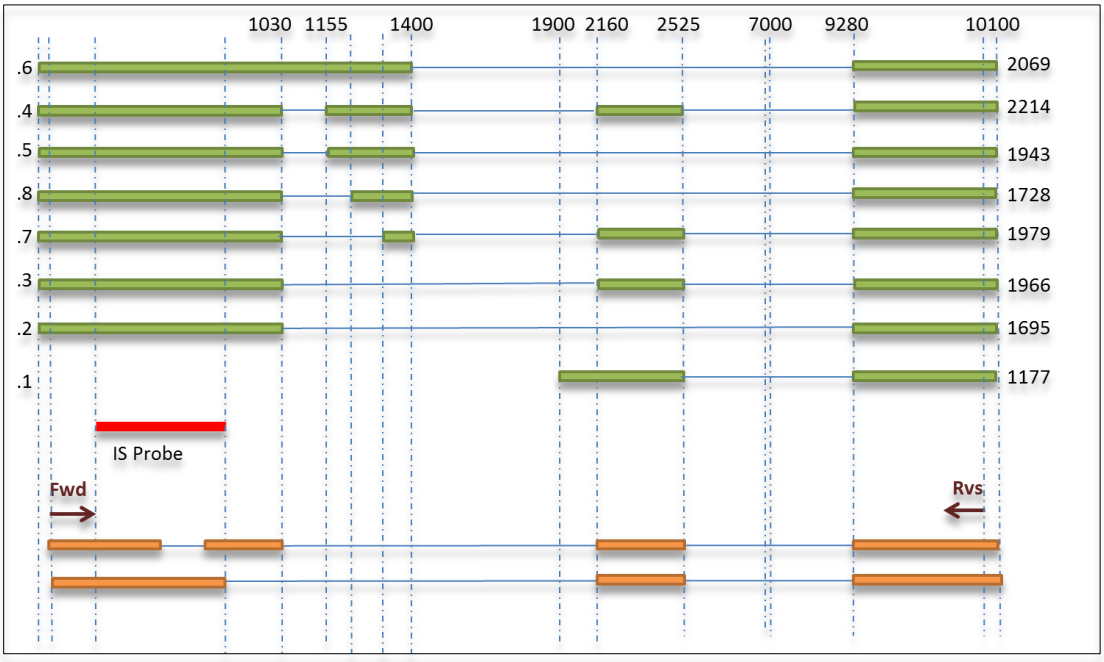


Figure 4.7 Schematic displaying the *Hottip* transcript variants identified by RNA-seq in *M. natalensis* limb autopods at the developmental stages CS 15, CS16 and CS17. The numbers down the left-hand side are the various transcript numbers of the read Mnat.T.18248. The numbers down the right-hand side are the lengths of each transcript variant. The red bar represents the position of the *in situ* probe.

4.2.2 Generating *In situ* Probe Synthesis

Tbx5-as1

Generation of a *Tbx5-as1* mouse probe was not successful as expression was too low in the FLs. The low expression is supported by the low FPKMs from mouse RNA-seq dataset (FPKMs below 5 in the FL, no expression in the HLs). The *M. natalensis* probe generation was successful as can be seen in Figure 4.9 with the final cleaned up product being a single band of the expected size (499bp).

Hottip

Mouse: Neither the *Hottip* primers designed over sequence extracted from Ensembl (ENSMUSG00000055408), nor the primers used in Wang paper (obtained through personal communication) (Wang et al. 2011), were able to amplify product a product from E13.5 cDNA. The published data on E13.5 mice was used as a comparison for spatial expression patterns (equivalent developmental stage as CS 16) (Wang et al. 2011). *M. natalensis*: probe synthesis was successful as can be seen in Figure 4.8 with the final cleaned-up product being a single band of the expected size (548bp).

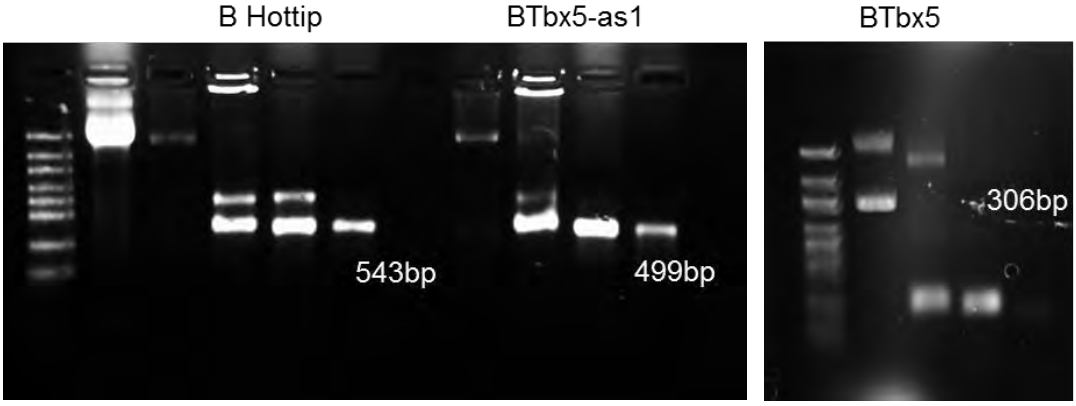


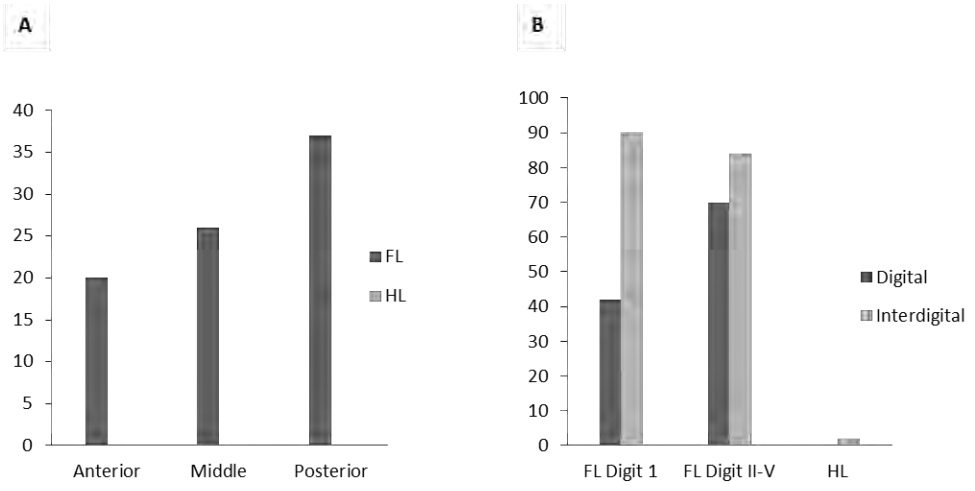
Figure 4.8 Steps involved in generating the *Tbx5-as1* and *Hottip* probes by *in vitro* transcription. Step 1) Linearized plasmid Step 2) In vitro transcribed probe. Step 3) DNase treated Probe. Step 4) Purified probe

4.2.3. *M. natalensis* and *M. schreibersii* Dataset and WISH comparison

4.2.3.1 Tbx5-as1

There were four gene 'fragments' in the *M. schreibersii* that aligned to the *M. natalensis* transcript isoforms annotated as *Tbx5-as1*. The fragment in the *M. schreibersii* labelled as 'comp14042' displayed the most sequence similarity and %coverage to the *M. natalensis Tbx5-as1* of all the gene fragments. The gene counts for comp14042 at the early stage (CS 14) reveal no expression in the HLs, while the FLs display higher expression in the middle and posterior domains than the anterior region. The middle and posterior region of the FL autopod will later expand with digit elongation (Figure 4.10, A). In the *M. schreibersii* dataset at CS 17 expression is still absent in the HLs (bar for a signal of 1 in the HL interdigital tissue), and expression in the FLs is higher in the interdigital tissue then its corresponding digits, with expression being reduced in digit I compared to the rest of the autopod, and expression being the highest in the interdigital tissue between digit I and II (Figure 4.10 B).

In the *M. schreibersii* dataset, the gene fragment that aligned to the *M. natalensis Tbx5* contained variants that extended far beyond the annotated *Tbx5* in the *M. natalensis* dataset. This incorrect inclusion of regions that are not *Tbx5* could explain the presence of expression in the HL at both CS 14 and CS 17 in the *M. schreibersii* dataset (Figure 4.10, C and D).



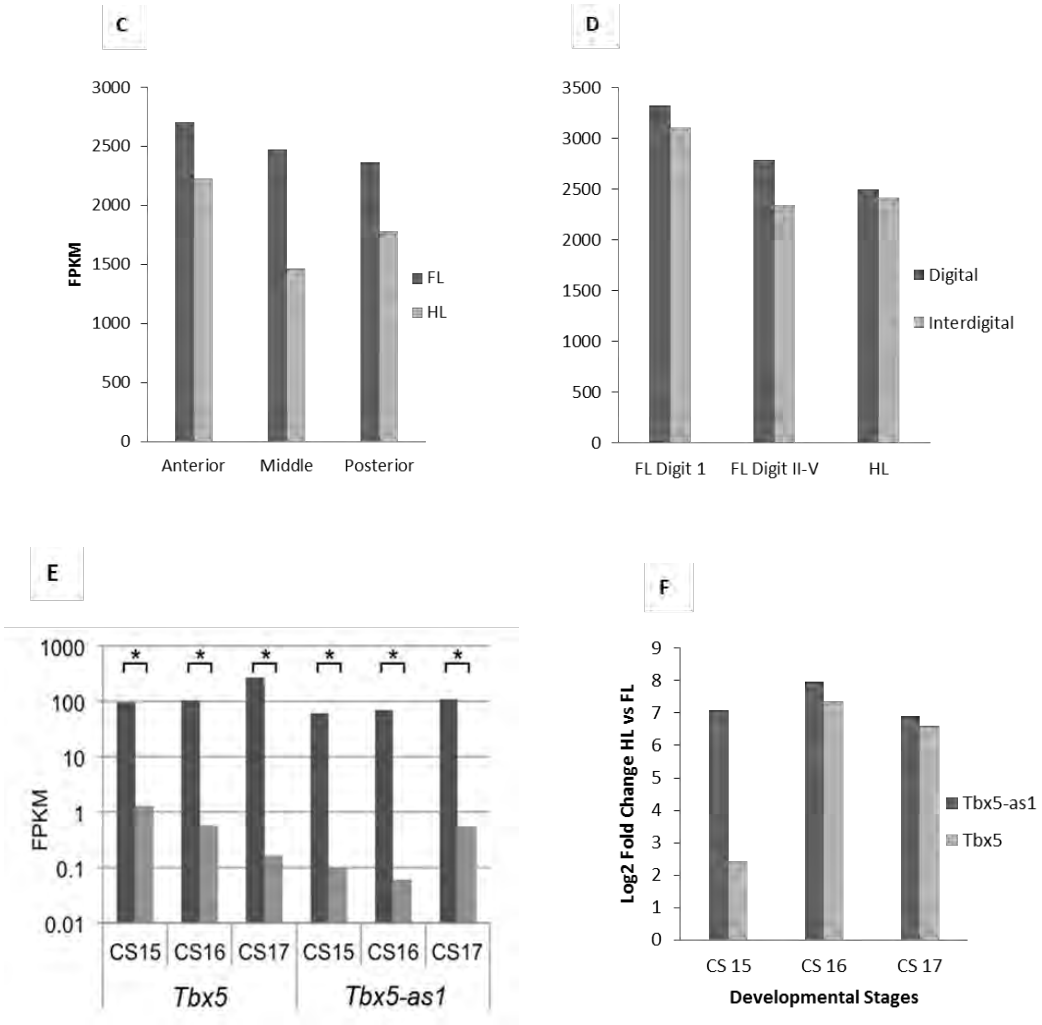


Figure 4.9 *Tbx5-as1* and *Tbx5* FPKMs in the *M. natalensis* and *M. schreibersii* datasets A) *M. schreibersii* CS 14 *Tbx5-as1* expression patterns B) *M. schreibersii* pooled CS15-17 *Tbx5-as1* expression patterns. C) *M. schreibersii CS 14 Tbx5* expression patterns D) *M. schreibersii* pooled CS15-17 *Tbx5* expression patterns. E) *M. natalensis* expression for *Tbx5* and *Tbx5-as1* asterics display significant FL versus HL expression changes by stage. F) *M. natalensis* Log_2 Fold Change HL vs FL for *Tbx5-as1* and *Tbx5*.

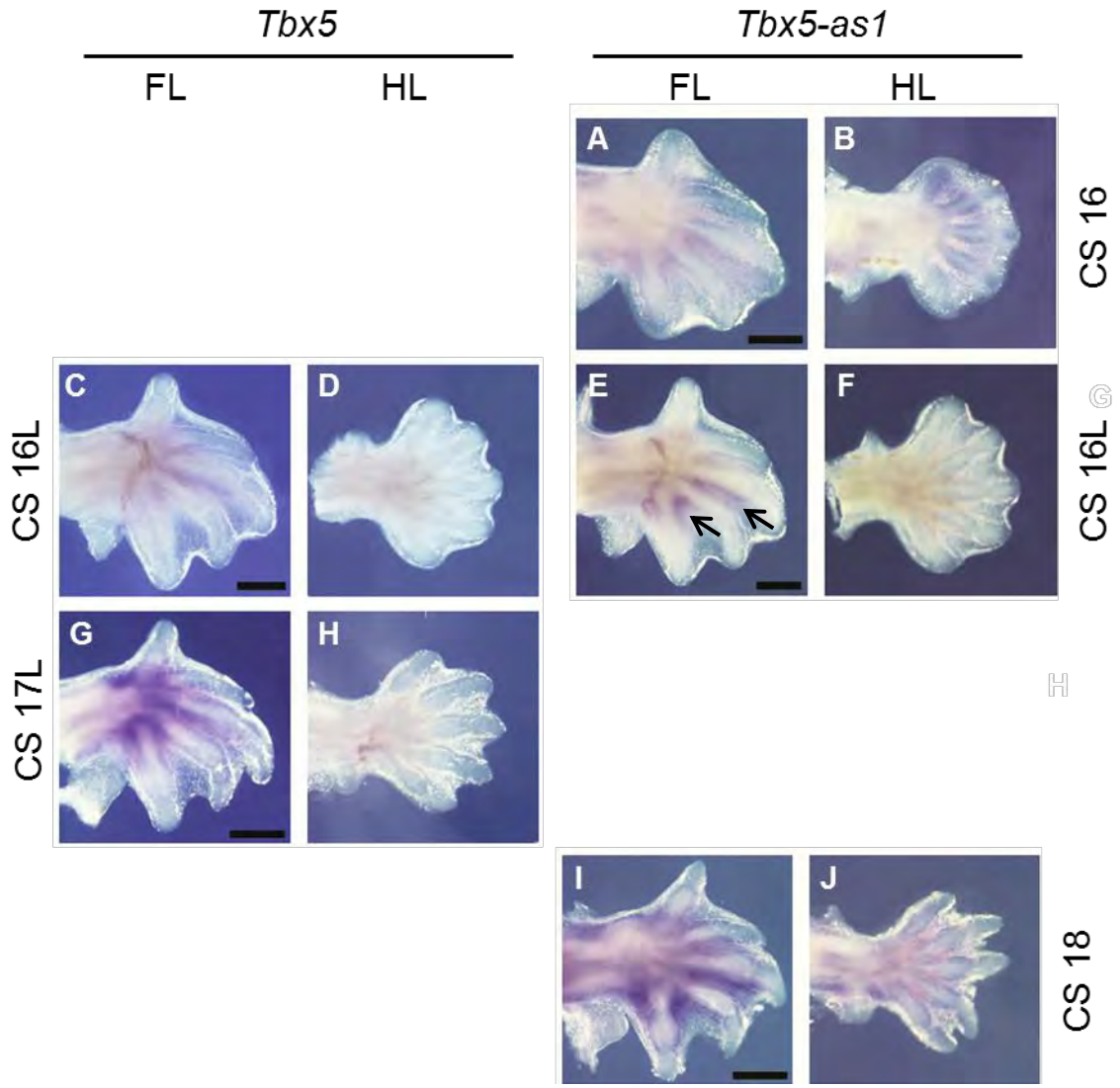


Figure 4.10 *Tbx5* and *Tbx5-as1* expression in the FL and HL of *M. natalensis* embryos. A and B; E and F; I and J) CS 16 and CS 16L and CS 18 *Tbx5-as1* expression. C and D; G and H) CS 16 and CS 17L *Tbx5* expression.

Tbx5 Expression in Developing Bat Embryos

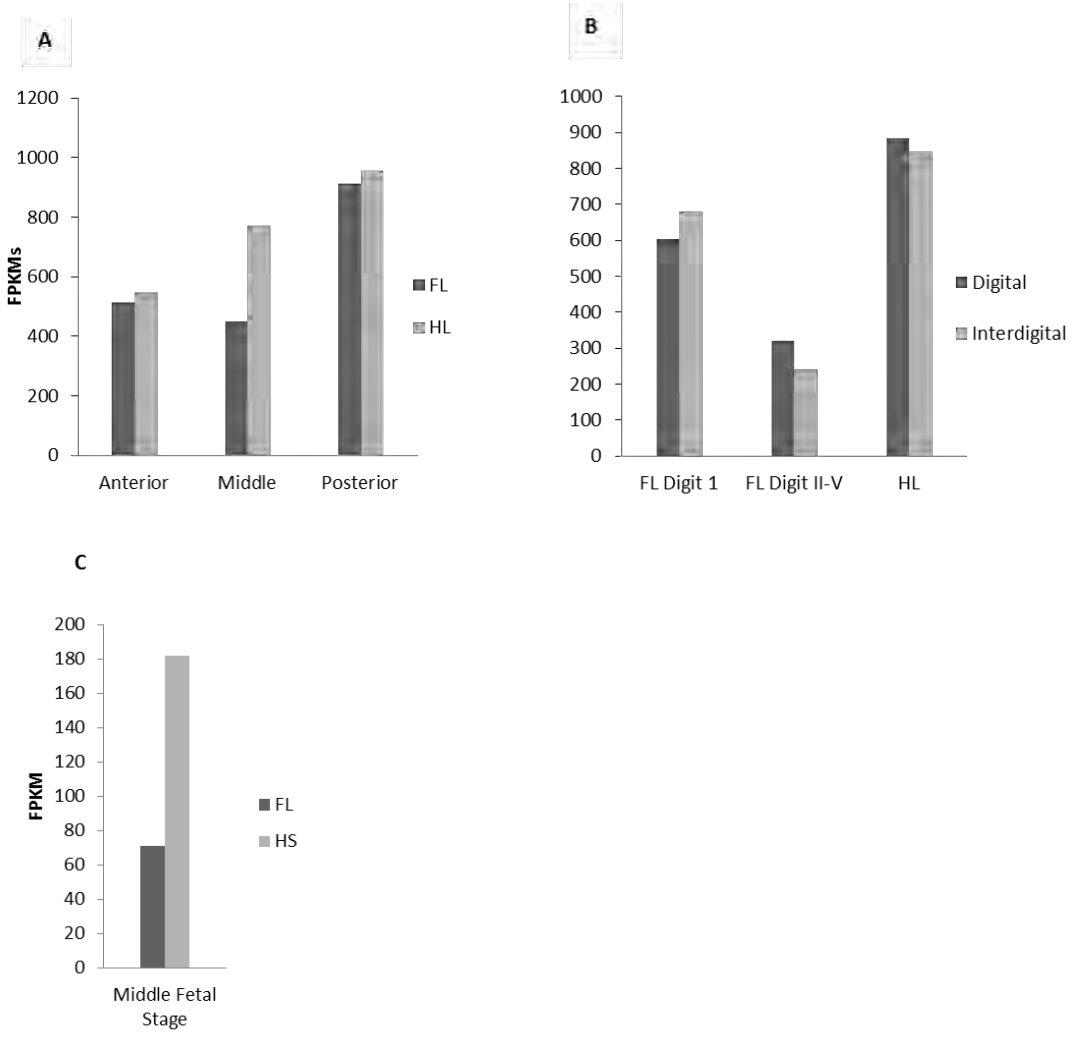
Tbx5 shows faint staining at CS 16L in the interdigital tissue in the centre of the FL, in the presumptive metacarpal domain, while expression is absent from the HLs (Figure 4.1, C and D). At CS 17L staining is much stronger at the base of the metacarpals and the interdigital tissue of digit III-V, and is absent from the HLs (Figure 4.11, G and H).

Tbx5-as1 Expression in Developing Bat Limbs

Overall expression is faint at CS 16L, with weak expression in the FL proximal interdigital tissue between digit III and V (Figure 4.11, A), while the HLs show minor signal dispersed throughout the limb bud (Figure 4.11, B). At CS 16L, the *in situ* signal is more prominent and concentrated in the proximal interdigital tissue at the base of the metacarpals (Figure 4.11, E), while expression is completely absent from the HLs (Figure 4.11, F). The CS 18 embryo was overstained, so a comparison of signal with the other stages won't be accurate; however, the pattern of expression in the FLs remains similar to CS 16L, with expression being robust in the proximal interdigital tissue (Figure 4.11, I), while expression is lacking in the HLs, as the colour present is due to over-staining (Figure 4.11 J).

4.2.3.2 *Hottip*

There was a single *M. schreibersii* gene 'comp7364' consisting of 6 transcripts that aligned to the 8 *M. natalensis Hottip* transcripts. At CS 14 in *M. schreibersii*, *Hottip* expression is comparable between FL and HL anterior and posterior regions, with expression being up-regulated in the middle domain of the HL compared to the FL (Figure 4.11 A). By CS 15- 17, expression in the FL digit ii-v is roughly half of the expression in digit I and its interdigital tissue, and a third of expression of the HL. At the middle fetal stage, expression is more than double in the HL metatarsal tissue compared to the FL metacarpal tissue, suggesting that the pattern of expression present during the patterning phase extends into the middle fetal stage. In the *M. natalensis* dataset, overall expression of *Hottip* is less than *HoxA13* (Figure 4.11 G).



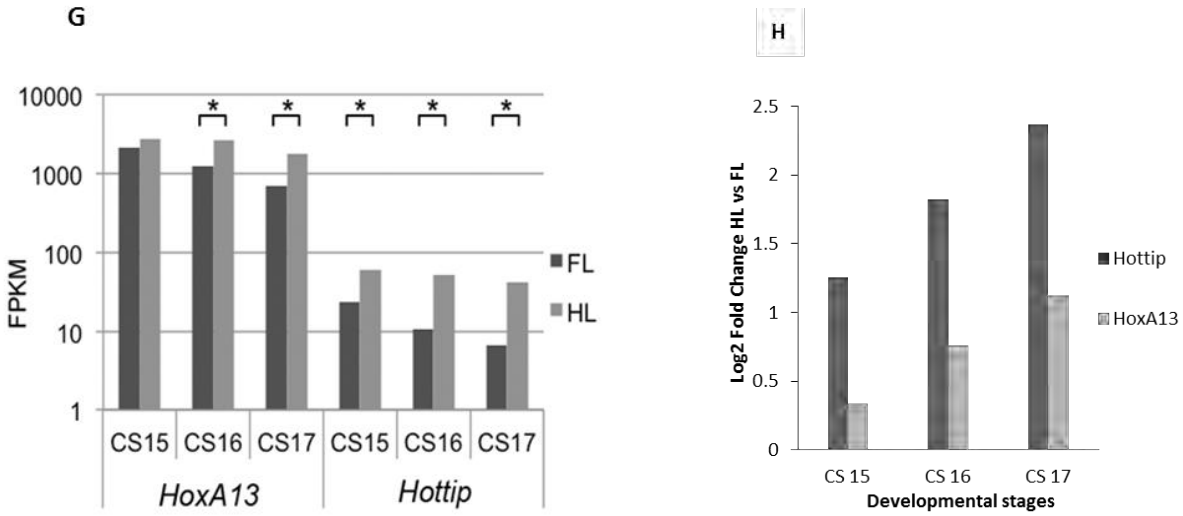


Figure 4.11 *Hottip* and *HoxA13* FPKMs in the *M. natalensis* and *M. schreibersii* datasets. A) *M. schreibersii* CS 14 *Hottip* expression patterns B) *M. schreibersii* pooled CS15-17 *Hottip* expression patterns. D) *M. schreibersii CS 14 HoxA13* expression patterns E) *M. schreibersii* pooled CS15-17 *HoxA13* expression patterns. F) *M. schreibersii* middle fetal FL: metacarpal and HS: metatarsal *HoxA13* expression G) *M. natalensis* expression for *HoxA13* and *Hottip* asterics display significant FL versus HL expression changes by stage. H) Log_2 fold-change between FL and HL in *M. natalensis.*

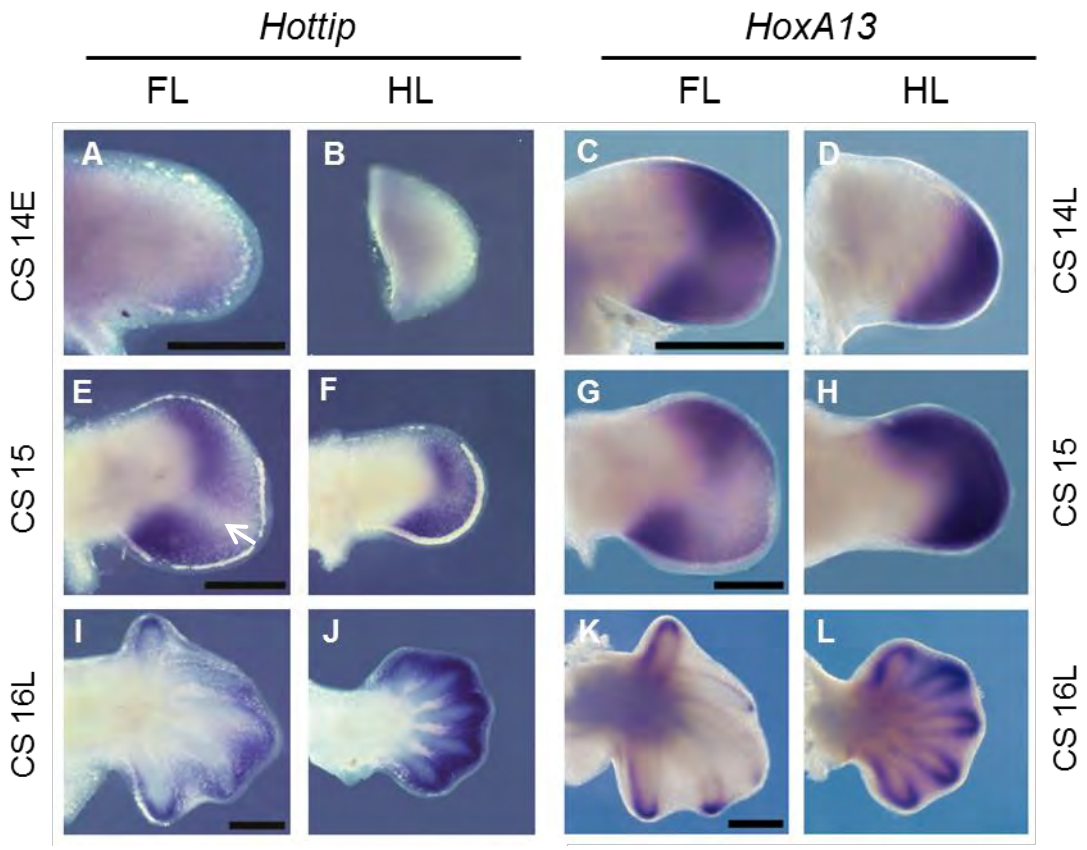


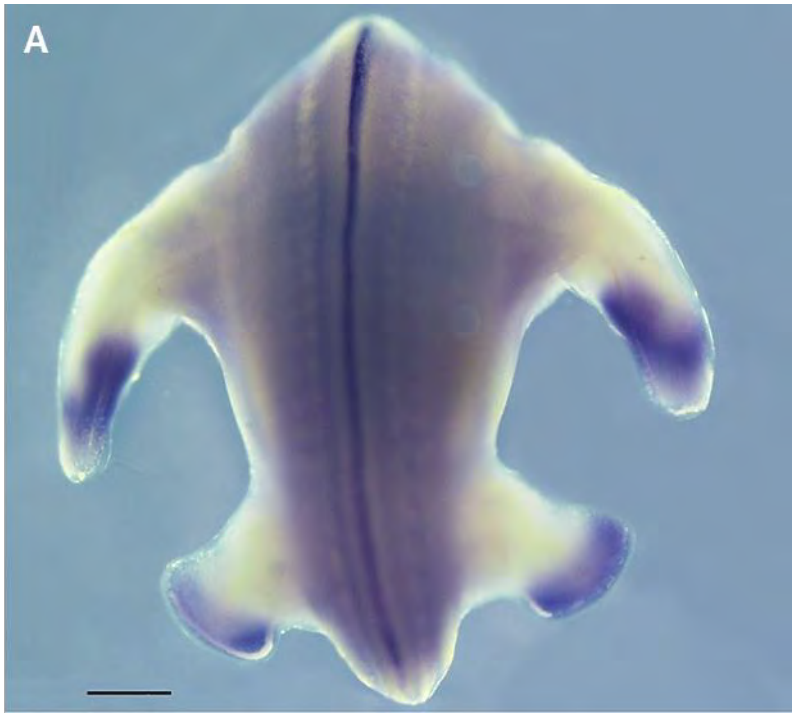
Figure 4.12 *Hottip* and *HoxA13* expression patterns in the FL and HLs of *M. natalensis* embryos. *HoxA13*: The flesh and stain of these embryos is slightly redder than the rest of the embryos as they were not taken through a methanol series. (Eckalbar et. Al. manuscript under review).

Hottip Expression in Developing Bat Limbs

At CS 14E there is faint staining in the distal-most posterior region of the FL bud (Figure 4.12 A), with expression being absent from the HL (Figure 4.12 B). At CS 15 expression is concentrated in the presumptive metacarpal region of the FL bud, with expression being the strongest in a proximal posterior region, with exclusion of expression in the region that will later develop into digit IV (Figure 4.12, E). Expression in the HL follows a similar pattern as the FL, with a bold band of expression over the proximal-middle domain of the autopod (Figure 4.12 F). At CS 16L this pattern alters considerably, with FL expression being completely absent from the metacarpal region, and is faint and restricted to distal interdigital tissue (Figure 4.12 I). In the HLs, *Hottip* is expressed at a much higher level compared to the FL in the interdigital tissue, with less robust expression in the digit primordia (Figure 4.12 J).

HoxA13 Expression in Developing Bat Limbs

At CS 14L expression is similar between FL and HL, with slight exclusion of staining in the FL in a triangle from the centre of the limb bud to the middle of the distal limb bud (Figure 4.12 C). At CS 15 this triangle becomes more pronounced with *HoxA13* expression being saturated in the presumptive digit I domain and the most proximal and posterior region of the autopod. In the HLs there is robust staining, with signal becoming more concentrated in the posterior region of the limb bud (Figure 4.12 H). By CS 16L expression is dramatically reduced in the FL compared to the HL, with expression absent from the interdigital tissue becoming restricted to the digit tips as well as the mesenchyme lining the anterior edge of digit I and the posterior edge of digit V (Figure 4.12 K). There is expression throughout the HLs, with robust signal lining all five digits (Figure 4.12 L).



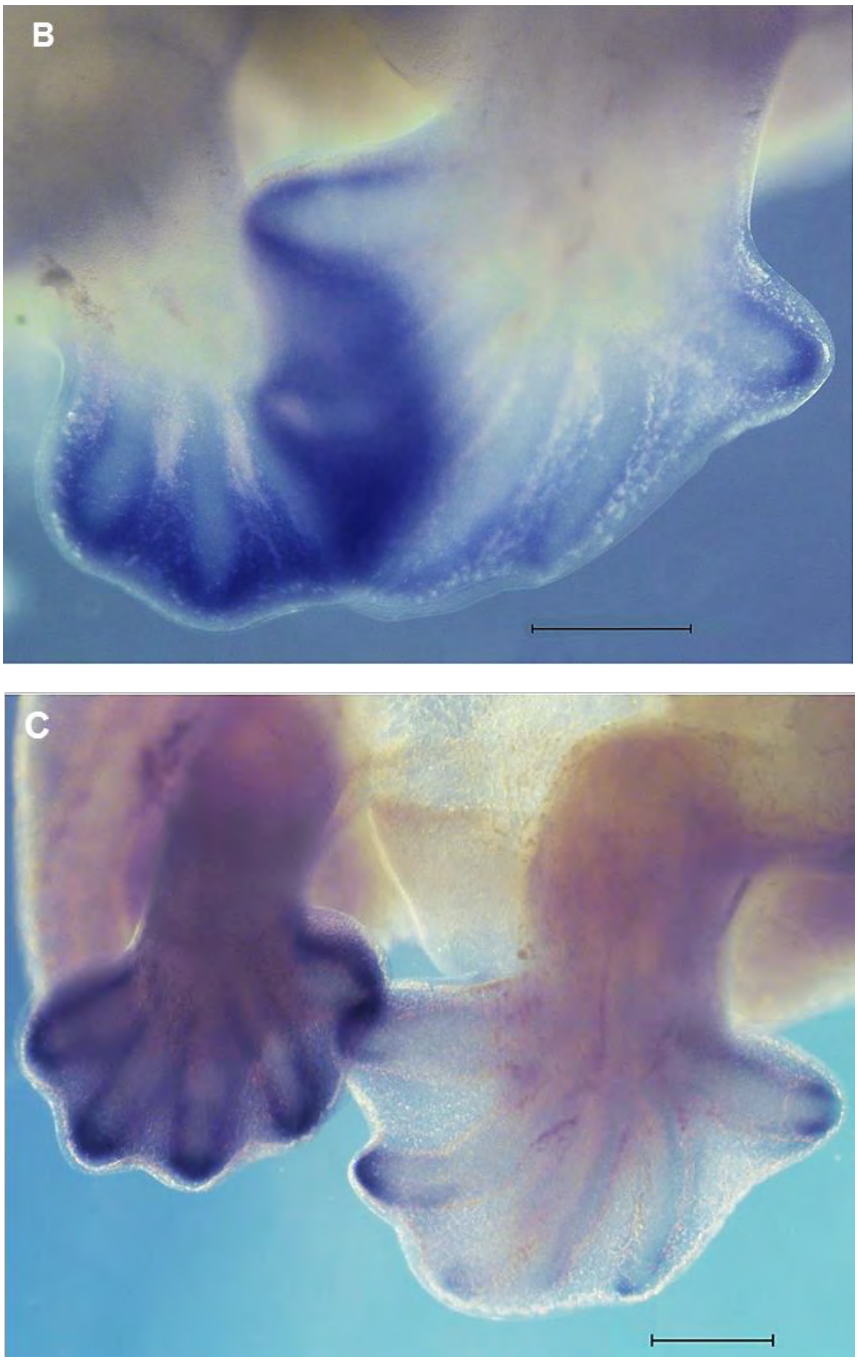


Figure 4.13 Side-view images from the same embryo as the montages, taken to directly compare staining between the FL and HL A) Dorsal view of CS 15 embryo with *Hottip* WISH. B) Side-view of *Hottip* CS 16L embryo and C) Side-view of a *HoxA13* CS 17 embryo. A) *Hottip* expression is comparable between FL and HL at this stage**.** B) The FL is overlaps the HL, but the translucent nature of the embryo in glycerol, the HL staining is visible through the FL. Expression is in distal interdigital regions of the FL, and is reduced in the FL relative to the HL. C) *HoxA13* expression is reduced in the FL relative to the HL, and is restricted to the tips of digits I, IV and V.

HoxA13 is expressed is throughout the HL and is strongest in the tips and lining of the digit ray progenitors.

4.3. Discussion

4.3.1 *Hottip* and *Tbx5-as1* Transcript Isoforms

The transcripts for both *Hottip* and *Tbx5-as1* from the CS 18L RNA did not have the same splice variants abundant in the RNA-seq for the patterning phase. The isoforms that were cloned and sequenced matched *Hottip* and *Tbx5-as1* transcripts, but with different exon combinations. It is possible that these differential splice variants are due to different regulation requirements at this stage of development where the interdigital tissue is already retained and thinned. It is also the stage that precedes ossification of autopod elements (Hockman et al. 2008). The presence of these transcripts validates that these LncRNAs are not artefacts.

4.3.2 *Tbx5-as1* and *Tbx5*

Tbx5 is a known FL bud initiation factor which mediates initiation at E9.5 in mice (Agarwal et al. 2003) and CS 12L in bats (Hockman et al. 2009). Seeing as *Tbx5* is expressed at high levels in the FL at initiation (Rallis et al. 2003), one would expect the *M. natalensis* RNA-seq to display higher *Tbx5* expression in the FL compared to the HL at the earliest sampled developmental stage CS 15. However; *Tbx5* expression increases from CS 15-17 (Figure 4.10 E and F), with fold-changes increasing from 2.43 (CS 15) to 7.36 and 6.59 (CS 16 and CS 17). For this reason, older stages of development were selected for spatial expression analysis by WISH.

The CS 16 embryo used in this sample was obtained in 2008, so it is likely that the mRNA had degraded, resulting in a faint signal. The *in situ* was repeated on a CS 16 L embryo that was bisected and divided between *Tbx5* and *Tbx5-as1* probes, to control for staging and tissue treatment making the comparison between the results more accurate. In support of a coupled activity for these transcripts, the expression patterns of *Tbx5* and *Tbx5-as1* were similar, being restricted to the base of digits I to V at CS 16L and CS 17, with clear expression in the proximal inter-digital tissue. It is likely that

these overlapping regions of expression could predict a region of late expression of *Tbx5* in the FL.

4.3.3 Hottip and *HoxA13*

In situ data reveals early expression of *HoxA13* and *Hottip* in FL autopods compared to HL autopods (Figure 4.12 A and B; C and D). This pattern is consistent with a role that *HoxA13* plays in driving early digit mesenchymal condensation (Stadler et al., 2001). Mesenchymal condensations occur earlier in bat FL autopods (Hockman et al., 2009) compared to bat HL autopods, and the elevated expression of markers of early mesenchymal condensations (*Rac1*, *Sox6* and *N-cadherin;* data not shown) in CS15 FL autopods is consistent with early bat FL digit chondrogenesis.

Bats have reduced IPCD in their FLs to result in the retained interdigital tissue making up the Chiropatagium, whereas the HLs have normal IPCD to result in free digits. A possible mechanism that belies this adaptation is the *Hottip* regulated reduction of *HoxA13* in bat FLs. An indicator of regulation of *HoxA13* by *Hottip* would be overlapping domains of expression. *Hottip* is significantly higher in the HL compared to the FL at CS 15 (p-value: 2.47E-06), whereas expression of *HoxA13* is not differentially expressed between FL and HL at CS 15 (P-value: 0.30). At later stages the genes follow a similar pattern of fold change between FL and HL (Figure 4.12 G and H), suggesting *HoxA13* regulation by *Hottip*, as the pattern of reduced FL *Hottip* expression precedes reduced *HoxA13* FL expression. The *in situ* results revealed that there is considerably reduced *Hottip* expression in the digit and interdigital tissue of CS 16 embryos. Further investigation is needed to substantiate and understand the regulation of *HoxA13 by Hottip.*

Chapter 5

Conclusions and Recommendations

5.1 The field of Evo Devo

The field of Evo Devo aims to study developmental processes of model organisms such as mice in conjunction with non-model organisms to compare differences during development (Heffer & Pick 2013). The central hypothesis of the field is that differences in morphology are brought about by alterations to regulation of key developmental toolkit genes rather than changes in protein sequences (Carroll 2008).

An RNA-seq dataset of embryos from three key developmental stages allowed for the identification of genes that were differentially expressed between FL and HL (Eckabar et. al., manuscript under review). Five transcription factors *Lef1*, *Lhx8*, *HoxA10*, *Mllt3* and *Tbx5*, as well as two LncRNAs *Hottip* and *Tbx5-as1* were selected for further investigation in order to characterise their spatial expression and reveal their domain of influence in autopod development. WISH was conducted on two or more stages of bat, with equivalently staged mouse embryos.

5.2 Transcription Factors

A known Wnt/β catenin effector *Lef1* was shown to be upregulated in CS 15 HLS compared to FLs. WISH validated this expression, with *Lef1* displaying a strong signal in a region in the anterior proximal boundary of the HL, as well as a less robust signal in the proximal posterior region of the HL that will later develop into the calcar (Figure 3.8), while expression was weak in this region of the FL. The region in the junction between the patagia and the anterior boundary of autopod cannot be explained by Wnt signalling without any more information, however; the position of the posterior expression in the HL could suggest elevated mesenchymal condensations as it is the region where the cartilaginous calcar will develop at CS 19 (Hockman et al. 2009).

A known regulator of neuronal and maxillary arch development: *Lhx8*, has not been previously identified in limb development, was shown to have a high fold change in FLs compared to HLs in the *M. natalensis* dataset (Eckalbar et. al. manuscript under review). Spatial characterisation by WISH revealed a unique *Lhx8* domain of expression in the FL zeugopod at CS 16, that is absent in bat HLs or mice limbs. There is light signal in the FL autopods in at CS 15, and the interdigital tissue between digits III-V at CS 16, however this region is not strong enough to account for the high fold-change between FL and HL. This signal could be due to inclusion of FL zeugopod tissue in the samples as the *Lhx8* signal is incredibly strong in the junction between the wrist and autopod. The link between *Lhx8* and neuronal development is potentially very interesting as bats have highly developed microscopic muscles, and an abundance of cutaneous receptors adapted for complex aerial manoeuvres (Sterbing-D'Angelo et al. 2011). It is possible that this novel expression of *Lhx8* in the presumptive patagia between bat limbs is due to the bat-specific complex nerve-system that will develop in the adult wing membrane.

HoxA10 expression was not significantly differentially expressed between FL and HL in the *M. natalensis* dataset at any stage (Eckalbar et al. manuscript under review). The *in situs* validate this expression as the signal is comparable between FL and HL at all three stages. However; there is a definite signal in the metacarpals at CS 16L which could suggest an involvement of *HoxA10* in chondrogenesis. However the qPCR of the in situ probe displays higher expression in the HL at this stage, so it is possible that this metacarpal staining is exaggerated due to overall overstaining of this embryo (Figure 3.17 B).

A novel expression domain was identified for *Mllt3* in the distal edge of bat FLs that is not present in the HLs or in mouse limbs of equivalent stages. *Mllt3* has been shown to be involved in axial patterning by regulating *Hox* genes, however; deletions to *Mllt3* in transgenic mice show no differences to wt mice (data not shown) (Eckalbar, manuscript under review). Additional functional studies conducted on bat tissue would need to be conducted to elucidate the exact mechanism of action of this upregulated *Mllt3*, and the target genes involved.

5.3 LncRNAs

Tbx5-as1 and *Hottip* were validated by the successful cloning of transcript variants from CS 18L FL and HL respectively. The transcript variants that were most abundant at this stage of development displayed an alternate slicing pattern to those identified in the RNA-seq dataset from stages CS 15-CS 17. The idea that alternatively splice variants of LncRNAs govern different regulatory roles at different developmental points is potentially very interesting. To fully characterise the effect of LncRNAs on regulation, functional studies would need to be conducted.

The spatial characterisation of the LncRNAs in *M. natalensis* was successful. *Hottip* has shown to regulate expression of distal *HoxA* genes, with the greatest effect on the autopod patterning gene *HoxA13* (Wang et al. 2014). WISH on CS 14E (*Hottip*) and CS 14L (*HoxA13*) displayed comparable expression between FL and HL. This expression is consistent with the involvement of *HoxA13* in early mesenchymal condensation. The later mechanism of IPCD is also regulated by *HoxA13*. There is a similar pattern of expression at CS 16L, with an absence of expression of both genes in the digit and interdigital regions and restriction of expression to the digit tips in CS 16L *HoxA13* embryos compared to HLs.

The other LncRNA selected for further investigation: *Tbx5-as1*, is not characterised in literature, but exists in close proximity (Figure 4.1) to the FL initiation regulator *Tbx5*. Older embryos were selected for *Tbx5* and the *Tbx5* antisense LncRNA *Tbx5-as1* WISH as both *M. natalensis* and *M. schreibersii* datasets displayed an increase in expression of both genes in the later stages. The *M. schreibersii* dataset displayed high expression of *Tbx5-as1* at CS 14 (Figure 4.10, A), which precedes the high fold change of *Tbx5* later in development (Figure 4.10 F). The specific role that *Tbx5* would play in the FL at this late stage of development is unclear, and would require more functional work. Taken together, the spatial characterisation of both LncRNAs follows a similar pattern of expression thus validating the *M. natalensis* RNA-seq dataset (Eckalbar et. al., manuscript under review).

5.4 Limitations of This Study

The restricted sample size of bats is a great limitation of the study as it results in reduced biological repeats, thus reducing the temporal description of the spatial expression of each selected gene. The reason for the small sample size is due to the difficulties associated with obtaining animals from the wild, especially nocturnal animals such as bats. Seasonal differences in climate in 2014 caused the gestation period to shift forward from the previous trips, resulting in an earlier gestation period and embryos that were too old.

An unavoidable technical issue was gross sampling of the tissue used for RNA-seq. From observations on genes expressed at high levels in the wrist (*HoxA11* and *Lhx8*) it is apparent that the high RNA-seq signal may be due to wrist tissue being included in the 'autopod sample'. This is likely as tissue sampling was done in the field, and wrist tissue is difficult to execute with such small limbs and indistinguishable boundaries between autopod and zeugopod at stages such as CS 15.

5.5 Conclusion

This study has provided further insight into a few known and a few previously unknown key transcription factors and LncRNAs involved in shaping the bat wing. The knowledge gleaned from this work can aid in informing future studies of similar anatomical anomalies evident amongst mammals. Bats are not currently not an option for transgenic experimentation, future functional characterization could be carried out using technologies such as chromatin capture, sectioning and transgenic experiments performed in the mouse. Enhancer-regulation studies would greatly benefit the understanding of regulation of key genes involved in development. The ever-growing datasets generated from 'omic' studies, a greater understanding of mathematical modelling can aid in understanding the complex interplay between all mechanisms involved in limb development.

Although limb development has been extensively studied as a paradigm for morphogenesis and organogenesis, questions still remain. Further Evo Devo studies

can help deepen our understanding of the complex processes involved in shaping mammalian limbs. Despite the caveats of this work, the spatial and temporal characterisation of selected genes supports the RNA-seq analyses (Eckalbar et. al. manuscript under review) and highlight genes that have previously uncharacterized roles in limb development.

References

Affairs, R., Angeles, L. & Berkeley, L., 2015. Statistics requantitates. , 2018(2013), pp.2013–2015.

Agarwal, P. et al., 2003. Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. *Development (Cambridge, England)*, 130, pp.623–633.

Altschul, S. et al., 1997. Gapped BLAST and PSI- BLAST: a new generation of protein database search programs. *Nucleic acids Res*, 25(17), pp.3389–3402. Available at: http://nar.oxfordjournals.org/content/25/17/3389.short.

Basson1, C.T. et al., 1997. Mutations in human cause limb and cardiac malformation in Holt-Oram syndrome. *Nat Genet*, 15(1), pp.30–35. Available at: http://dx.doi.org/10.1038/ng0197-30.

Berlivet, S. et al., Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing limbs. , pp.1–60.

Brugmann, S. a et al., 2007. Wnt signaling mediates regional specification in the vertebrate face. *Development (Cambridge, England)*, 134(18), pp.3283–3295.

Bustin, S. a. et al., 2009. The MIQE guidelines:Minimum Information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, 55(4), pp.611–622.

Cabili, M. et al., 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes and Development*, 25(18), pp.1915–1927.

Carroll, S.B., 2008. Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. *Cell*, 134(1), pp.25–36.

Chen, C.-H. et al., 2005. Hoxd13 expression in the developing limbs of the short-tailed fruit bat, Carollia perspicillata. *Evolution & development*, 7(2), pp.130–41. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15733311.

Church, V.L. & Francis-West, P., 2002. Wnt signalling during limb development. *The International journal of developmental biology*, 46(7), pp.927–936.

Collins, E.C. et al., 2002. Mouse Af9 Is a Controller of Embryo Patterning , Like Mll , Whose Human Homologue Fuses with AF9 after Chromosomal Translocation in Leukemia. , 22(20), pp.7313–7324.

Cooper, K.L. et al., 2011. Initiation of proximal-distal patterning in the vertebrate limb by signals and growth. *Science*, 332(6033), pp.1083–1086. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3258580&tool=pmcentrez&rendertype=abstract.

Cooper, L.N., Cretekos, C.J. & Sears, K.E., 2012. The evolution and development of mammalian flight. *Wiley Interdisciplinary Reviews: Developmental Biology*, 1(5), pp.773–779. Available at: http://doi.wiley.com/10.1002/wdev.50.

Cretekos, C.J. et al., 2005. Embryonic staging system for the short-tailed fruit bat, Carollia perspicillata, a model organism for the mammalian order Chiroptera, based upon timed pregnancies in captive-bred animals. *Developmental Dynamics*, 233(3), pp.721–738.

Cretekos, C.J. et al., 2008. Regulatory divergence modifies limb length between mammals. *Genes and Development*, 22(2), pp.141–151.

Cunningham, T.J. et al., 2011. Rdh10 mutants deficient in limb field retinoic acid signaling exhibit normal limb patterning but display interdigital webbing. *Developmental Dynamics*, 240(5), pp.1142–1150.

Cunningham, T.J. & Duester, G., 2015. Mechanisms of retinoic acid signalling and its roles in organ and limb development. *Nat Rev Mol Cell Biol*, 16(2), pp.110–123. Available at: http://dx.doi.org/10.1038/nrm3932.

Davis, M.C., 2013. The Deep Homology of the Autopod: Insights from Hox Gene Regulation. *Integrative and Comparative Biology*, pp.1–9. Available at: http://icb.oxfordjournals.org/cgi/doi/10.1093/icb/ict029 [Accessed April 30, 2013].

Eiting, T.P. & Gunnell, G.F., 2009. Global Completeness of the Bat Fossil Record. *Journal of Mammalian Evolution*, 16(3), pp.151–173. Available at: http://link.springer.com/10.1007/s10914-009-9118-x.

Flandin, P. et al., 2011. Lhx6 and Lhx8 coordinately induce neuronal expression of Shh that controls the generation of interneuron progenitors. *Neuron*, 70(5), pp.939–50. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3153409&tool=pmcentrez&rendertype=abstract [Accessed May 24, 2014].

Gaiti, F. et al., 2015. Dynamic and Widespread lncRNA Expression in a Sponge and the Origin of Animal Complexity. *Molecular Biology and Evolution*, 32(9), pp.2367–2382. Available at: http://mbe.oxfordjournals.org/cgi/doi/10.1093/molbev/msv117.

Galceran, J. et al., 1999. deficiency in Lef1 − / − Tcf1 − / − mice. , pp.709–717.

Greenbaum, D. et al., 2003. Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome biology*, 4(9), p.117.

Gunnell, G.F. & Simmons, N.B., 2012. Evolutionary history of bats: fossils, molecules and morphology. *Cambridge studies in morphology and molecules: new paradigms in evolutionary biology*, 94(2), pp.i–xii, 1–560.

Gunnell, G.F. & Simmons, N.B., 2005. Fossil Evidence and the Origin of Bats. *Journal of Mammalian Evolution*, 12(1-2), pp.209–246. Available at: http://link.springer.com/10.1007/s10914-005-6945-2.

Hall, B.K., 2003. Evo-Devo: evolutionary developmental mechanisms. *Int J Dev Biol*, 47(7-8), pp.491–495. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14756324.

Hall, T., 1999a. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, pp.95–98.

Hall, T., 1999b. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, pp.95–98. Available at: http://jwbrown.mbio.ncsu.edu/JWB/papers/1999Hall1.pdf.

Harfe, B.D., 2011. Keeping up with the zone of polarizing activity: New roles for an old signaling center. *Developmental dynamics : an official publication of the American Association of Anatomists*, 240(5), pp.915–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21360795.

Hartmann, C., 2006. A Wnt canon orchestrating osteoblastogenesis. *Trends in Cell Biology*, 16(3), pp.151–158.

Hasson, P., Del Buono, J. & Logan, M.P.O., 2007. Tbx5 is dispensable for forelimb outgrowth. *Development (Cambridge, England)*, 134(1), pp.85–92.

Heffer, A. & Pick, L., 2013. Conservation and Variation in *Hox* Genes: How Insect Models Pioneered the Evo-Devo Field. *Annual Review of Entomology*, 58(1), pp.161–179. Available at: http://www.annualreviews.org/doi/abs/10.1146/annurev-ento-120811-153601.

Hockman, D. et al., 2008. A second wave of Sonic hedgehog expression during the development of the bat limb. *Proceedings of the National Academy of Sciences of the United States of America*, 105(44), pp.16982–16987.

Hockman, D. et al., 2009. The role of early development in mammalian limb diversification: a descriptive comparison of early limb development between the Natal long-fingered bat (Miniopterus natalensis) and the mouse (Mus musculus). *Developmental dynamics : an official publication of the American Association of Anatomists*, 238(4), pp.965–79. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19253395 [Accessed August 21, 2014].

Hodgkin, J., 1998. Seven types of pleiotropy. *International Journal of Developmental Biology*, 42(3), pp.501–505.

Hoppler, S. & Kavanagh, C.L., 2007. Wnt signalling: variety at the core. *Journal of cell science*, 120, pp.385–393.

Hu, L. et al., 2015. A common set of distinct features that characterize noncoding RNAs across multiple species. *Nucleic Acids Research*, 43(1), pp.104–114. Available at: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gku1316.

Kaufman, M. H., & Kaufman, M.H. (1992). T. atlas of mouse development (Vol. 428). L.A. press., 1992. *The atlas of mouse development* Vol. 428., London: Academic press.

Kawagoe, H., Kawagoe, R. & Sano, K., 2001. Targeted down-regulation of MLL-AF9 with antisense oligodeoxyribonucleotide reduces the expression of the HOXA7 and -A10 genes and induces apoptosis in a human leukemia cell line, THP-1. *Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K*, 15(11), pp.1743–1749.

Kengaku, M. et al., 1998. Distinct WNT pathways regulating AER formation and dorsoventral polarity in the chick limb bud. *Science (New York, N.Y.)*, 280(5367), pp.1274–1277.

King, M.-C., Wilson, a. C. & Url, S., 2007. Evolution at Two Levels in Humans and Chimpanzees Mary-Claire King; A. C. Wilson. *Science*, 188(4184), pp.107–16.

Knosp, W.M. et al., 2007. Elucidation, quantitative refinement, and in vivo utilization of the HOXA13 DNA binding site. *Journal of Biological Chemistry*, 282(9), pp.6843–6853.

Knosp, W.M. et al., 2004. HOXA13 regulates the expression of bone morphogenetic proteins 2 and 7 to control distal limb morphogenesis. *Development (Cambridge, England)*, 131(18), pp.4581–92. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15342482 [Accessed March 23, 2013].

Lancaster, W.C. & Speakman, J.R., 2001. Variations in respiratory muscle activity during echolocation when stationary in three species of bat (Microchiroptera: Vespertilionidae). *The Journal of experimental biology*, 204(Pt 24), pp.4185–4197.

Landin Malt, A. et al., 2014a. Identification of a Face Enhancer Reveals Direct Regulation of LIM Homeobox 8 (Lhx8) by Wingless-Int (WNT)/β-Catenin Signaling. *Journal of Biological Chemistry*, 289, pp.30289–30301. Available at: http://www.jbc.org/lookup/doi/10.1074/jbc.M114.592014.

Landin Malt, A. et al., 2014b. Identification of a Face Enhancer Reveals Direct Regulation of LIM Homeobox 8 (Lhx8) by Wingless-Int (WNT)/β-Catenin Signaling. *Journal of Biological Chemistry*, 289(44), pp.30289–30301. Available at: http://www.jbc.org/lookup/doi/10.1074/jbc.M114.592014.

Liang, L. et al., 2013. Adaptive evolution of the hox gene family for development in bats and dolphins. *PloS one*, 8(6), p.e65944. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3692524&tool=pmcentrez&rendertype=abstract [Accessed July 8, 2013].

MacDonald, B.T., Tamai, K. & He, X., 2009. Wnt/??-Catenin Signaling: Components, Mechanisms, and Diseases. *Developmental Cell*, 17(1), pp.9–26. Available at: http://dx.doi.org/10.1016/j.devcel.2009.06.016.

Mallarino, R. & Abzhanov, A., 2012. Paths Less Traveled: Evo-Devo Approaches to Investigating Animal Morphological Evolution. *Annual Review of Cell and Developmental Biology*, 28(1), pp.743–763.

Martin, P., 1990. Tissue patterning in the developing mouse limb. *International Journal of Developmental Biology*, 34(3), pp.323–336.

Mason, M.K. et al., 2015. Retinoic acid-independent expression of Meis2 during autopod patterning in the developing bat and mouse limb. *EvoDevo*, 6(1). Available at: http://www.evodevojournal.com/content/6/1/6.

Mattick, J.S. & Rinn, J.L., 2015. Discovery and annotation of long noncoding RNAs.

Nature Publishing Group, 22(1), pp.5–7. Available at: http://dx.doi.org/10.1038/nsmb.2942.

Mercader, N. et al., 1999. Conserved regulation of proximodistal limb axis development by Meis1/Hth. *Nature*, 402(1994), pp.425–429.

Mercader, N. et al., 2009. Ectopic Meis1 expression in the mouse limb bud alters P-D patterning in a Pbx1-independent manner. *International Journal of Developmental Biology*, 53(8-10), pp.1483–1494.

Mercer, T.R. & Mattick, J.S., 2013. Structure and function of long noncoding RNAs in epigenetic regulation. *Nature structural & molecular biology*, 20(3), pp.300–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23463315.

Miller-Butterworth, C.M. et al., 2007. A family matter: Conclusive resolution of the taxonomic position of the long-fingered bats, Miniopterus. *Molecular Biology and Evolution*, 24(7), pp.1553–1561.

Perez, W.D. et al., 2010. Survival of Hoxa13 homozygous mutants reveals a novel role in digit patterning and appendicular skeletal development. *Developmental Dynamics*, 239(2), pp.446–457.

Peter John Taylor, 2000. *Bats of Southern Africa*, University of Natal Press Pietermaritzburg.

Pramparo, T. et al., 2005. Loss-of-function mutation of the AF9/MLLT3 gene in a girl with neuromotor development delay, cerebellar ataxia, and epilepsy. *Human genetics*, 118(1), pp.76–81.

Probst, S. et al., 2011. SHH propagates distal limb bud development by enhancing CYP26B1-mediated retinoic acid clearance via AER-FGF signalling. *Development (Cambridge, England)*, 138(10), pp.1913–23. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3082298&tool=pmcentrez&rendertype=abstract.

Pyron, R.A., 2011. Divergence time estimation using fossils as terminal taxa and the origins of lissamphibia. *Systematic Biology*, 60(4), pp.466–481.

Rabinowitz, A.H. & Vokes, S. a, 2012. Integration of the transcriptional networks regulating limb morphogenesis. *Developmental biology*, 368(2), pp.165–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22683377 [Accessed May 7, 2013].

Rallis, C. et al., 2003. Tbx5 is required for forelimb bud formation and continued outgrowth. *Development (Cambridge, England)*, 130(12), pp.2741–2751.

Rhinn, M. & Dolle, P., 2012. Retinoic acid signalling during development. *Development*, 139(5), pp.843–858. Available at: http://dev.biologists.org/cgi/doi/10.1242/dev.065938.

Richardson, M.K., 1999. Vertebrate evolution: the developmental origins of adult variation. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 21, pp.604–613.

Rinn, J.L. & Chang, H.Y., 2012. Genome regulation by long noncoding RNAs. *Annual review of biochemistry*, 81, pp.145–66. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3858397&tool=pmcentrez&rendertype=abstract [Accessed July 9, 2014].

Rosello-Diez, A. et al., 2014. Diffusible signals and epigenetic timing cooperate in late proximo-distal limb patterning. *Development*, 141(7), pp.1534–1543. Available at: http://dev.biologists.org/cgi/doi/10.1242/dev.106831.

Ruijter, J.M. et al., 2013. Evaluation of qPCR curve analysis methods for reliable biomarker discovery: Bias, resolution, precision, and implications. *Methods*, 59(1), pp.32–46. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1046202312002290.

Rushikesh Sheth, 1*† Luciano Marcon, 2, 3* M. Félix Bastida, 1, 4 Marisa Junco, 1 Laura Quintana, 2, 3 Randall Dahn, 5 Marie Kmita, 6‡ James Sharpe, 2, 3, 7‡

Maria A. Ros1, 4‡, 2012. Hox Genes Regulate Digit Patterning by Controlling the Wavelength of a Turing-Type Mechanism. *Science*, 338(December), pp.1476–1480.

Rutledge, R.G., 2011. A Java Program for LRE-Based Real-Time qPCR that Enables Large-Scale Absolute Quantification. *PLoS ONE*, 6(3), p.e17636. Available at: http://dx.plos.org/10.1371/journal.pone.0017636.

Sanger, F., Nicklen, S. & Coulson, a R., 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), pp.5463–5467.

Sears, K.E. et al., 2006. Development of bat flight: morphologic and molecular evolution of bat wing digits. *Proceedings of the National Academy of Sciences of the United States of America*, 103(17), pp.6581–6586.

Shou, S. et al., 2013. HOXA13 regulates Aldh1a2 expression in the autopod to facilitate interdigital programmed cell death. *Developmental dynamics : an official publication of the American Association of Anatomists*, pp.1–38. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23553814 [Accessed April 29, 2013].

Slack, J.M., Holland, P.W. & Graham, C.F., 1993. The zootype and the phylotypic stage. *Nature*, 361(6412), pp.490–492.

Spitz, F. & Furlong, E.E.M., 2012. Transcription factors: from enhancer binding to developmental control. *Nature Reviews Genetics*, 13(9), pp.613–626.

Stadler, H.S., Higgins, K.M. & Capecchi, M.R., 2001. Loss of Eph-receptor expression correlates with loss of cell adhesion and chondrogenic capacity in Hoxa13 mutant limbs. *Development (Cambridge, England)*, 128(21), pp.4177–88. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11684655.

Stearns, F.W., 2010. One hundred years of pleiotropy: A retrospective. *Genetics*, 186(3), pp.767–773.

Sterbing-D'Angelo, S. et al., 2011. Bat wing sensors support flight control. *Proceedings of the National Academy of Sciences of the United States of America*, 108(27), pp.11291–11296.

Stoffberg, S., Jacobs, D.S. & Miller-Butterworth, C.M., 2004. Field identification of two morphologically similar bats, Miniopterus schreibersii natalensis and Miniopterus fraterculus (Chiroptera: Vespertilionidae). *African Zoology*, 39(1), pp.47–53.

Suzuki, T., 2013. How is digit identity determined during limb development? *Development, growth & differentiation*, 55(1), pp.130–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23230964 [Accessed March 15, 2013].

Tabin, C. & Wolpert, L., 2007. Rethinking the proximodistal axis of the vertebrate limb in the molecular era. *Genes and Development*, 21(12), pp.1433–1442.

Tamamura, Y. et al., 2005. Developmental regulation of Wnt/??-catenin signals is required for growth plate assembly, cartilage integrity, and endochondral ossification. *Journal of Biological Chemistry*, 280(19), pp.19185–19195.

Tomioka, T. et al., 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 *. , 8.

Towers, M. & Tickle, C., 2009. Growing models of vertebrate limb development. *Development (Cambridge, England)*, 136(2), pp.179–190.

Tschopp, P. & Duboule, D., 2011. A Genetic Approach to the Transcriptional Regulation of Hox Gene Clusters.

Unsicker, K. & Krieglstein, K., 2006. Cell Signaling and Growth Factors in Development: From Molecules to Organogenesis. , 37(12), p.21702.

Uzkudun, M., Marcon, L. & Sharpe, J., 2015. Data-driven modelling of a gene regulatory network for cell fate decisions in the growing limb bud. *Molecular systems biology*, 8(815), pp.1–15.

Vogel, T. & Gruss, P., 2009. Expression of Leukaemia associated transcription factor Af9/Mllt3 in the cerebral cortex of the mouse. *Gene expression patterns : GEP*, 9(2), pp.83–93. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19000783 [Accessed June 19, 2014].

Wang, K.C. et al., 2011. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*, 472(7341), pp.120–4. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3670758&tool=pmcentrez&rendertype=abstract [Accessed July 10, 2014].

Wang, Z. et al., 2014. Unique expression patterns of multiple key genes associated with the evolution of mammalian flight Unique expression patterns of multiple key genes associated with the evolution of mammalian flight.

Weatherbee, S.D. et al., 2006. Interdigital webbing retention in bat wings illustrates genetic changes underlying amniote limb diversification. *Proceedings of the National Academy of Sciences of the United States of America*, 103(41), pp.15103–15107.

Yashiro, K. et al., 2004. Regulation of Retinoic Acid Distribution Is Required for Proximodistal Patterning and Outgrowth of the Developing Mouse Limb. *Developmental Cell*, 6(3), pp.411–422. Available at: http://www.cell.com/article/S1534580704000620/fulltext.

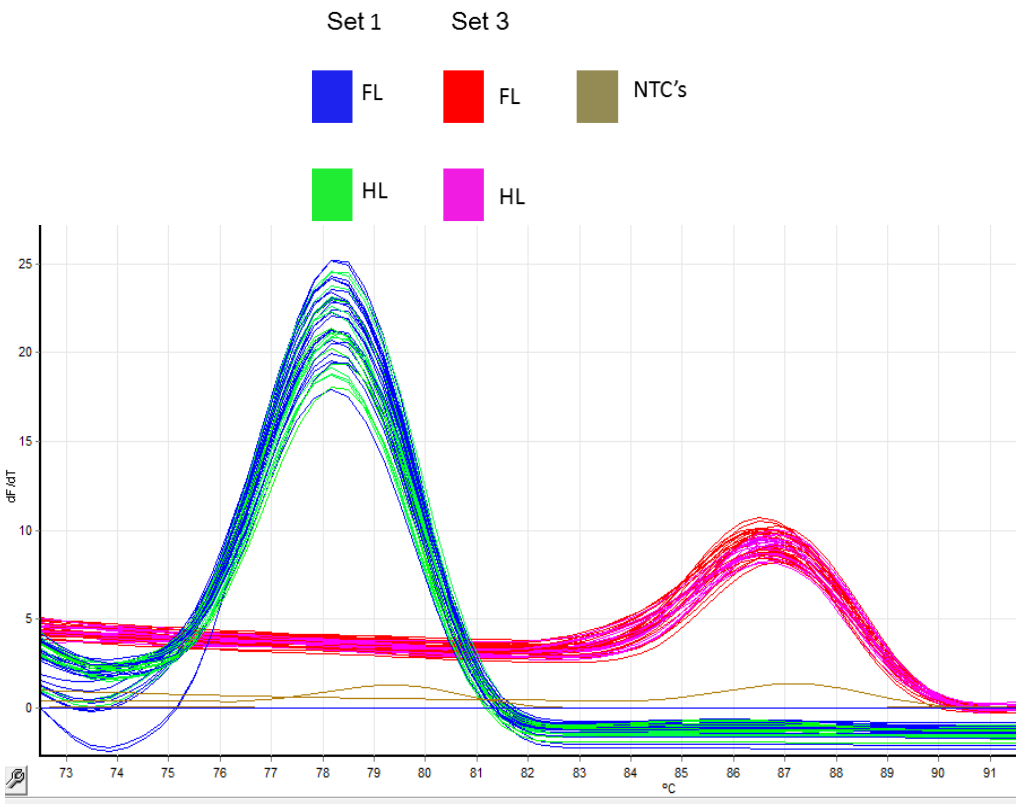
Zakany, J. & Duboule, D., 2007. The role of Hox genes during vertebrate limb development. *Current opinion in genetics & development*, 17(4), pp.359–66. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17644373 [Accessed March 14, 2013].

Zhao, X. et al., 2009. Retinoic acid promotes limb induction through effects on body axis extension but is unnecessary for limb patterning. *Current biology : CB*, 19(12), pp.1050–7. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2701469&tool=pmcentrez&rendertype=abstract.

Zhao, Y. et al., 2003. The LIM-homeobox gene Lhx8 is required for the development of many cholinergic neurons in the mouse forebrain. *Proceedings of the National Academy of Sciences of the United States of America*, 100(15), pp.9005–9010.

Zuniga, A., 2015. Next generation limb development and evolution: old questions, new perspectives. *Development*, 142(22), pp.3810–3820. Available at: http://dev.biologists.org/cgi/doi/10.1242/dev.125757.

Appendix



Appendix 1.2 qPCR Melt curves of the two *HoxA10* regions. The first peak represents Primer set 1, while the second peak represents primer set 3.

1.3 Discovering LncRNAs – In Detail

Protein-coding genes were removed from the assembled RNA-seq transcriptome by a pairwise blastp of all transcripts to the Mammalian Uniprot database. The uniprot database was utilised as it is a well-curated database that contains the least amount of spurious putative proteins which arise from automated gene annotation predictions. This blastp resulted in a set of 436 transcripts corresponding to 227 gene loci that did not encode any known proteins. The sequences corresponding to the 436 transcripts were blasted against other bat genomes, mouse, human, dog, cow, and cat genomes. Seeing as Transcripts from 142 genes showed some conservation across mammals, 12 of which have homologues in the LncRNA databases.

Transcripts from 63 genes showed conservation across one or more bat species, but not other mammals. Transcripts from 34 genes did not have a match elsewhere (ie specific to *M. natalensis*). 130 putative lncRNAs did not have homologues in existing lncRNA databases, but did share similarity when blasted against other mammalian genomes. 11 of these were differentially expressed with FPKM values > 10 Table. There were two groups of bat *Tbx5-as1* transcripts: lcl|comp653591_c2_seq64; lcl|comp653591_c2_seq39 and Mnat.G.8148.