

UNIVERSIDADE DE SÃO PAULO
FFCLRP - DEPARTAMENTO DE BIOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA COMPARADA

**Associations between *Hox* genes molecular evolution and the evolution of
morphological diversity in Squamata and Marsupialia**

**Associações entre a evolução molecular dos genes *Hox* e a evolução da
diversidade morfológica em Squamata e Marsupialia**

Sarah Ribeiro Milograna

Tese apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da USP, como parte das exigências para a obtenção do título de Doutora em Ciências, Área: Biologia Comparada.

Ribeirão Preto/SP

2015

UNIVERSIDADE DE SÃO PAULO
FFCLRP - DEPARTAMENTO DE BIOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA COMPARADA

**Associations between *Hox* genes molecular evolution and the evolution of
morphological diversity in Squamata and Marsupialia**

*Associações entre a evolução molecular dos genes *Hox* e a evolução da
diversidade morfológica em Squamata e Marsupialia*

Sarah Ribeiro Milograna

Orientadora: Prof^a. Dr^a. Tiana Kohlsdorf

Tese apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da USP, como parte das exigências para a obtenção do título de Doutora em Ciências, Área: Biologia Comparada.

VERSÃO CORRIGIDA

Ribeirão Preto/SP

2015

Ficha Catalográfica

Milograna, SR

Associations between *Hox* genes molecular evolution and the evolution of morphological diversity in Squamata and Marsupialia

202 p. : Il. ; 29,7 cm

Bibliografia: p. 183- 202

Tese de Doutorado apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto/USP – Área de Concentração: Biologia Comparada.

Orientadora: Kohldorf, Tiana.

1. *HOX* genes. 2. Evolution 3. Evolutionary developmental biology. 4. Bioinformatics 5. Snake 6. Amphisbaenia 7. Marsupial 8. *CsB*. 9. *Island I* 10. *CNS65* 11. Regulatory genes 12. Limbs 13. TFBS 14. Anterior-posterior body axis

To my parents and grandparents, my fundamental inspiration for all battles.

To Professor John McNamara, my scientific father.

To Jason Scealy, the co-author of my dreams and achievements.

**“It is not the strongest or the most intelligent who will survive but those who can best
manage change.”**

-Leon C. Megginson

SUMMARY OF CONTENTS

ACKNOWLEDGEMENTS	iv
ABSTRACT	vii
RESUMO	ix
GENERAL BACKGROUND	1
CHAPTER I - The independent evolution of snakelike squamate lineages registered both convergent and divergent regulatory signatures in the HoxD cluster enhancer <i>CsB</i>.	11
1.1 Abstract	12
1.2 Introduction	13
1.3 Methods	19
1.4 Results	24
1.5 Discussion	31
1.6 Conclusions	41
1.7 Appendix	42
CHAPTER II - Limb loss in snakes and degeneration of <i>CNS65</i> and <i>Island I</i> regulatory capacities: morphological evolution left footprints on <i>Hoxd</i> genes enhancers	55
2.1 Abstract	56
2.2 Introduction	57
2.3 Methods	62
2.4 Results	69
2.5 Discussion	78
2.6 Conclusions	87
2.7 Appendix	88
CHAPTER III - Evolution of limb development in marsupial mammals: noncoding RNAs associated with the regulation of <i>HOX</i> gene expression	97
3.1 Abstract	98
3.2 Introduction	99
3.3 Methods	106
3.4 Results	120
3.5 Discussion	136
3.6 Conclusions	144
3.7 Appendix	145
CONCLUDING REMARKS AND FUTURE PERSPECTIVES	178
LITERATURE CITED	183

ACKNOWLEDGEMENTS

I am very thankful to...

My supervisor Prof. Dr. Tiana Kohlsdorf, firstly, for her unconditional support regarding my carrier choices; she accepted my transference from another post-graduation program where I started my PhD and stimulated my international research internship. Thanks for having accepted and trusted me, and for always giving thoughtful advice in every step of this journey. Thanks for your strong example of hard, dedicated, passionate and tireless work, optimism and confidence. Thank you for your critical and insightful evaluation and editing in all reports, Thesis, manuscripts and presentations.

To the Post-graduation program in Comparative Biology, represented by the Coordinator Prof. Dr. Ricardo Macedo Corrêa and Castro and by Vice-Coordinator Prof. Dr. Tiana Kohlsdorf, for receiving me as a post-graduate student. To Prof. Márcia M. G. Bitondi for the careful revisions of my annual reports. To the secretary Vera Cassia Cicilini de Lucca, for all the kind and good-hearted support in administrative issues.

To the Biology Department of Faculdade de Filosofia, Ciências and Letras de Ribeirão Preto, USP, headed by Prof. Dr. Fernando Luis Medina Mantellato and by his surrogate Prof. Dr. Pietro Ciancaglini, and to all the helpful staff.

To Fundação de Amparo à Pesquisa do Estado de São Paulo, FAPESP and for the financial support, by providing me PhD regular (process #2012/13165-5) and research internship abroad (#2014/06503-7) scholarships.

To the Herpetological Collection from the Laboratory of Tetrapod Evolution, the Yale Peabody Museum of Natural History of Yale University (New Heaven, EUA), and the authors of Vargas *et al.* (2008), for gently providing genomic DNA samples. To the Herpetological Collections at *Universidade Federal do Mato Grosso* (UFMT, Cuiabá, Brazil), *Universidade Federal do Alagoas* (MUFAL, Maceió, Brazil), and *Universidade de Brasília* (CHUNB, Brazil); to Museum of Science and Technology at *Pontifícia Universidade Católica do Rio Grande do Sul* (MCP, Porto Alegre, Brazil); Zoology Museum at USP (MZUSP, São Paulo, Brazil); and to *Universidade do Estadual de São Paulo* (UNESP, Rio Claro, Brazil) for providing tissues from Squamata for DNA extraction. Genomic DNA and tissues were used in the research project of Chapter I.

To Prof. Dr. Pedro E. M. Guimarães for having performed the Bioinformatics analyses of Transcription Factor Binding Sites prediction from Chapter I, using a pipeline

developed by himself, and for his insights on interpreting data in the light of their biological meaning. To Prof. Igor Schneider and Acácio Freitas, for contributing with the gene report expression assays that supported our data from Chapter I.

To Prof. Igor Schneider, for his sapient contribution with ideas and guidance through the experiments performed in the research project correspondent to the Chapter II. For kindly receiving me in his laboratory at Universidade Federal do Pará, providing all expertise and material to construct the vectors used in transgenic assays. To himself and his adorable family for kindly welcoming me in their home and taking me to taste hundreds of ice cream flavors made from tropical amazonic fruits. To Gabriela Lima, who dedicated several weeks to accompany me during the experiments.

To Prof. Dr. Marilyn Renfree and Prof. Andrew Pask for having welcomed me in their research groups at the University of Melbourne-Australia, and guided me to accomplish the research corresponding to Chapter III. For having assisted me despite the many challenges that they had to face through the period I was there, and with all the paperwork required to make this possible.

To the Department of Biosciences from the University of Melbourne, for having received me there as a visitor researcher, providing every resource needed for my internship.

To Dr. Hongshi Yu, for technical supervising, his dedicated guidance through analyses and experiments. To Helen Clark and Elisabeth Pharo, for their support in and out of the laboratory. To my colleagues from the Walgroup and Pask-lab, for all the help with animal manipulation, experiments and ideas. Specially, to Jennifer Hetz, Yu Chen and Danielle Hickford for patiently taking me through the qPCR technique, and to Karen Hansen for sharing her hardly-learned experience with in situ hybridization. To Brandon and Helen, for their help in a hard moment. Thanks to all of them for welcoming me, for the friendship and emotional support.

To Prof. Dr. John C. McNamara, my former academic supervisor, for his unconditional support on my carrier choices and friendship, for never having stopped being my mentor through all this years, and for his tireless editing of our manuscripts. He never stopped contributing to my academic development.

To Tania Defina and Andrea C. Quiapim for sharing their expertise with sequencing protocols and for having prepared themselves many of the samples sequenced for Chapter I.

To the lovely technician in the Laboratory of Tetrapod Evolution (LET), Paula dos Santos, for her amazing support with laboratory logistic and maintenance, and for having saved me a few times with paperwork support, especially when I was abroad.

To Prof. Dr. Tiago C. Pereira, Prof. Dr. Milton Groppo, Prof. Dr. John C. McNamara, Prof. Dr. Maria Helena S. Goldman, Prof. Dr. Maria L. T. M. Polizeli, Prof. Dr. Zilá L. P. Simões, and their students for having always kindly shared thermo-cyclers and -80°C freezers when we faced technical issues with our equipment.

To all my colleagues and friends from LET whom with I have been sharing space, ideas, opinions, experiences, problems, and life through all these years. To Fernando Andrade for having pioneered the study of *CsB* element in our laboratory, and to him and Aline C. Dragalzew for guiding me on my first steps in the laboratory.

To Mariana B. Grizante and Marina E. Singarette, for so much...to start with our endless discussions about snakelike evolution mysteries, experimental and logistical support, co-authorships and friendship. Thanks to Mariana for helping me to organize myself when I left for my sandwich PhD in Australia, and for kindly X-raying the snake and the amphibaenian from Fig 1.1 in Chapter I. Thanks to Marina for helping me with Brazilian paperwork when I was abroad. For all the love and a lifetime's sisterhood!

To my parents, for having prepared me with education, support, stimulus and unconditional love that prepared me to handle undreamed challenges. For always having believed more than myself in my potential. To my sister, Julia, for always leading by example with her incredible strength. To all my family and friends for a lot of stimulus.

To my partner Jason Scealy, for all his patience, strength, support and love through these years and especially when he was my family in Australia. To all his family that welcomed, supported and made me feel at home there.

Finally, thanks to all the people that directly or indirectly contributed to this study. Without these people, this PhD Thesis could have never became a reality.

ABSTRACT

Hox genes pattern the vertebrate body during embryonic development, and understanding their evolution may unravel genetic mechanisms subjacent to morphological evolution. Molecular evolution of *Hox* genes entails signatures in regulatory regions that potentially affect gene expression, such as the *cis*-regulatory elements (CREs) that surround the *HoxD* cluster and its noncoding RNAs (ncRNAs). In this PhD Thesis, I have explored regulatory evolution of *HoxD* genes engaged in the development of appendicular and anterior-posterior body (AP) axes in amniotic lineages that exhibit homoplastic morphological peculiarities: snakelike squamates (Chapters I and II) and diprotodontid marsupials (Chapter III). In Chapter I, I investigated in snakes and amphisbaenians, whether equivalent regulatory signatures were registered in the Conserved Sequence B (*CsB*), a centromeric *Hoxd10-13* CRE, during evolution of snakelike morphologies. Using lizards and other tetrapods to represent the lacertiform morphology, conserved regions within *CsB* were sequenced from 38 squamate species, and transcription factor binding sites (TFBS) were predicted and compared among groups. Both snakelike lineages carry divergent and convergent regulatory signatures not identified in lizards; the convergence located in one *CsB* segment comprised loss of limb-related TFBS and gain of a binding site for PBX1. This convergent regulatory signature registered along two independent processes of snakelike evolution may relate to body elongation and limb loss, and evidences a role of *CsB* for AP axis development. In Chapter II, I investigated whether a telomeric (*CNS65*) and a centromeric (*Island I*) *Hoxd* enhancer that regulate gene expression respectively at proximal and distal regions of developing limbs retain their regulatory capacities in Serpentes. Gene reporter expression of these CREs from snakes were performed in transgenic mice and revealed that their regulatory activities were abrogated in limb buds. Comparison of predicted TFBS in these elements between snakes and limbed tetrapods revealed limb-related TFBS apparently lost in snakes, and pointed to one potential stilopodium/zeugopodium-specific element in *CNS65* and three likely autopodium-specific elements in *Island I*. Limb loss in snakes registered signatures in *Hoxd* CREs that possibly contributed for their functional impairment, putatively indicating limb-specific modules. Finally, in the chapter III, I studied ncRNAs from *HoxD* cluster in the context of hindlimb morphological evolution and heterochrony between fore and hindlimb development in the tammar wallaby *Macropus eugenii*. The ncRNAs mapped to *HoxD* cluster were selected from transcriptome performed using tammar embryo limbs at days 23 (d23) and 25 (d25) of

pregnancy, and their conservation, transcriptional profiles and expression patterns were explored. Comparison with orthologous sequences in other mammals revealed five ncRNAs conserved among mammals, and three transcripts apparently exclusive to marsupials. Transcriptional profiles of *HOXD10-13* and HoxD ncRNAs were mostly equivalent. *XLOC46* expression patterns resembled those of mouse and tammar terminal *HOXD* genes, whereas *XLOC52* and *XLOC53* showed identical expression patterns to those genes of tammar, except for *XLOC53* low expression at d25. The ncRNAs intergenic/intronic to *HOXD9-12* may regulate expression of terminal *HOXD* genes in mammals, and *XLOC52* and *XLOC53* are suitable for investigation regarding limb evolution in marsupial. This PhD Thesis demonstrates how studies of evolutionary footprints in regulatory elements of developmental genes contribute for elucidating specific processes during lineages divergence as well as functional aspects of these genes during development.

RESUMO

Os genes *Hox* padronizam o corpo dos vertebrados durante o desenvolvimento embrionário, e a compreensão de sua evolução pode elucidar mecanismos genéticos subjacentes à evolução morfológica. A evolução molecular dos genes *Hox* imprime assinaturas em regiões regulatórias, as quais potencialmente afetam sua expressão gênica, como os elementos *cis*-regulatórios (CREs) que ladeiam o cluster D de *Hox* e seus RNAs não-codificantes (ncRNAs). Essa Tese de Doutorado enfoca a evolução regulatória de genes *HoxD* envolvidos no estabelecimento dos eixos corpóreos axial ântero-posterior (AP) e apendiculares em linhagens de amniotas que exibem características morfológicas homoplásticas peculiares: os squamatas serpentiniformes (Capítulos I e II) e os marsupiais Diprotodontia (Capítulo III). No Capítulo I investigou-se, em serpentes e anfisbênias, se assinaturas regulatórias envolvidas no estabelecimento das morfologias serpentiniformes foram impressas na Sequência Conservada B (*Conserved Sequence B*, *CsB*), um CRE centromérico de *Hoxd10-13*. Usando lagartos e outros tetrápodes como referência para a morfologia serpentiniforme, regiões conservadas de *CsB* foram sequenciadas em 38 espécies de Squamata, cujos TFBS foram preditos e comparados. Ambas linhagens serpentiniformes exibem assinaturas regulatórias divergentes e convergentes ausentes em lagartos; a convergência localizou-se em um segmento de *CsB* que concentra perda nas linhagens serpentiniformes de diversos TFBS com funções no desenvolvimento de membros e a aquisição de um sítio de ligação para PBX1. Essa assinatura convergente impressa durante evoluções independentes da morfologia serpentiniforme pode estar relacionada à alongação corpórea e à perda dos membros, evidenciando um papel do *CsB* no desenvolvimento do eixo AP. No Capítulo II, foi investigado se um CRE telomérico (*CNS65*) e um centromérico (*Island I*) de *Hoxd*, os quais regulam respectivamente regiões proximais e distais dos membros tetrápodes em desenvolvimento, retêm suas capacidades regulatórias em Serpentes. Expressões de gene repórter desses CREs de serpentes foram realizadas em camundongo transgênico, revelando deficiência de suas atividades regulatórias nos brotos de membro. A comparação dos TFBS preditos nesses elementos entre serpentes e outros tetrápodes revelou que TFBS relacionados ao desenvolvimento dos membros foram perdidos nas sequências das serpentes. Ainda, essa comparação indicou um elemento em *CNS65* potencialmente envolvido especificamente na regulação da formação de estilopódio/zeugopódio, e três elementos na *Island I* exclusivamente reguladores do desenvolvimento autopodial. A perda de membros em

serpentes aparentemente imprimiu assinaturas nesses *CREs* de *Hoxd* que possivelmente contribuíram para sua degeneração funcional, putativamente indicando módulos específicos de regulação nos membros. No Capítulo III, ncRNAs do cluster D de Hox foram estudados no contexto da evolução morfológica do autopódio posterior e heterocronia entre o desenvolvimento de membros anteriores e posteriores em *Macropus eugenii*. Os ncRNAs mapeados sobre o cluster D de Hox foram selecionados a partir de transcrito de membros de embriões de *M. eugenii* nos dias 23 (d23) e 25 (d25) de gravidez, e sua conservação, perfis transcricionais e padrões de expressão foram explorados. A comparação com sequências ortólogas de outros mamíferos revelou cinco ncRNAs conservados em mamíferos, e três aparentemente exclusivos dos marsupiais. Os perfis transcricionais de genes *HOXD10-13* e dos ncRNAs do cluster D de Hox foram predominantemente equivalentes. Os padrões de expressão de *XLOC46* foi similar aos dos genes *HOXD* terminais de camundongo e *M. eugenii*, enquanto que *XLOC52* e *XLOC53* apresentaram expressão idêntica à desses genes em *M. eugenii*, exceto pela baixa expressão de *XLOC53* no d25. Os ncRNAs intergênicos/intrônicos aos genes *HOXD9-12* possivelmente regulam a expressão de genes *HOXD* terminais em mamíferos, enquanto que *XLOC52* e *XLOC53* constituem bons candidatos para investigação relacionada à evolução dos membros de marsupiais. Esta Tese demonstra como estudos de assinaturas regulatórias na evolução de genes do desenvolvimento contribuem para o entendimento das histórias evolutivas de divergência entre linhagens e dos aspectos funcionais desses genes do desenvolvimento.

GENERAL BACKGROUND

The present work corresponds to a PhD Thesis conceived in the Evolutionary Developmental Biology (EvoDevo) framework. The Thesis comprises three chapters focusing correlations between *HoxD* genes regulatory evolution and the evolution of morphological variation in tetrapods. In summary, I have explored whether and to what extent the evolution of snakelike and autopodial morphologies leaves footprints in diverse regulatory elements located externally and intergenically to *HoxD* cluster. The present section aims to provide a general theoretical background that introduces some concepts that permeate the three chapters, together with an explanation about the connections between *HoxD* gene expression during embryo development and the establishment of the adult phenotype. The section is divided in two items: first (named '*Hox genes roles on embryonic development*'), is a sub-section permeating *Hox* genes organization, functions and evolutionary mechanisms, where their roles and regulation on appendicular and axial axes are also introduced. Then, a second item (named '*Mechanism of morphological evolution associated with Hox regulatory systems*') is dedicated to detail the scenarios where associations of morphological and *Hox* molecular evolution were investigated. In the end of this theoretical background, the major goal shared by the three chapters of this Thesis was explained.

Hox genes roles on embryonic development

'*Hox*' is the name attributed to a group of several genes that comprise extremely conserved DNA-binding domains, termed homeobox, and operate as transcription factors regulating key developmental pathways widespread through vertebrate and invertebrate embryos (Lewis 1978; Carroll 1995; Spitz 2010), maintaining also several important functions in adults (Taylor et al. 1997; James & Kazenwadel 1991; Chen & Capecchi 1999; Takahashi

et al. 2004). In genomes of the prototypic Vertebrata and Cephalochordata, these genes were named after their relative positions in clusters, which usually reflect their patterns of spatiotemporal collinear expression in specific gradients of anterior-posterior axial (AP) and appendicular body axes during embryonic development (Carroll 1995; Spitz 2010). This collinear expression, initially described in model system organisms such as mice and drosophila, originated the concept of “Hox code” (Kessel & Gruss 1991), later improved after discovery of the “posterior prevalence” property (Duboule & Morata 1994; Duboule 2007), explaining how positional information is provided by *Hox* genes to axial and paraxial tissues, establishing the identity of several structures (Deschamps & Van Nes 2005; Spitz 2010). Along evolutionary history, *Hox* genes in ancestral vertebrates first encompassed one single cluster, which after two rounds of whole genome duplications resulted in four Hox clusters (HoxA to HoxD) that can be identified in most extant vertebrates (Panopoulou & Poustka 2005; Peer et al. 2009) except for bony fish, a lineage that experienced one additional genome duplication and exhibits eight Hox clusters (Crow & Wagner 2006). Hox clusters are spread among distinct chromosomes that vary according to the species, and yet, at least in most vertebrates, they conserve their typical arrangement inside each particular cluster, which in these animals are classified as ‘organized clusters’ or ‘type O clusters’ (Lemons & McGinnis 2006; Duboule 2007). *Hox* gene expression is fine-tuned by complex regulatory networks that favor their co-option in diverse contexts, and the similarity of developmental pathways by them weaved does not prevent highly specific functional outputs (Zakany & Duboule 2007). Such features characterize a robust but flexible genic system that allows evolutionary mechanisms to succeed in a way that little variation in *Hox* genes nucleotide sequences and on their interaction with downstream targets may elicit phenotypic changes (Carroll 2005; Montavon et al. 2011; Guerreiro et al. 2013). Consequently, *Hox* genes become very seductive for

scientists trying to elucidate evolution of phenotypic diversity and its relationships with embryonic development.

In vertebrate embryos, *Hox* genes are essential for patterning structures across the AP body axis, which is composed by mesodermal modules, the somites, sequentially established in both sides of the neural tube during embryonic development in a process that constitutes the axial skeleton and its associated muscular apparatus, contributing also with cranial elements (Deschamps & van Nes 2005; Woltering 2012). Terminal *Hox* genes, which are also frequently named 'distal *Hox* genes', comprise those genes located closer to 5' positions in their respective clusters, which are transcription factors often crucial for patterning structures in posterior domains of embryo AP body axis (Roberts et al. 1995; Yokouchi et al. 1995; Warot et al. 1997). Examples of structures located in such posterior domains are the caudal somites, posterior regions of the notochord, neural tube and intestine, the cloaca and external genitalia (Griffith et al. 1992; Gajovic et al. 1993; Catala et al. 1995; Knezevic et al. 1998; Santa Barbara & Roberts 2002; Scott et al. 2005).

In addition to establishing the AP body axis, *Hox* genes contribute to the induction and patterning of limb buds, as well as other appendices in non-tetrapod vertebrates (Archambeault et al. 2014), which originate as a small protraction from the body wall (Capdevila & Belmonte 2001). During limb development, two main embryonic regions are established in this structure, the apical ectodermal ridge (AER), an inductor of proximal-distal growth, and the zone of polarizing activity (ZPA), which patterns the stylopodium (humerus and femur), the zeugopodium (radius/ulna and tibia/fibula) and the autopodium (carpi/tarsi and digits) (Fromental-Ramain et al. 1996), as illustrated in Figure 0.1. During development of the structures aforementioned, terminal *HoxA* and *HoxD* genes are collinearly expressed, assuming domains that are equivalent but not identical to the expression of these genes during trunk formation and extension (Zakany & Duboule 2007). Transcription of *HoxD* genes

during development of proximal and distal limb segments occurs in two independent phases (Nelson et al. 1996; Tarchini & Duboule 2006). The initial transcription phase determines stylopodium and zeugopodium patterning from sequential and homogeneous activation of *HoxD1-9* genes (Zakany & Duboule 2007; Montavon et al. 2011). Development of the most distal region in the limb is regulated by the expression of *HoxD10-13* genes, which occurs in inverted spatiotemporal collinearity up to the final stages of autopodial development, configuring the second *HoxD* transcriptional wave (Fromental-Ramain et al. 1996; Montavon et al. 2008; Montavon et al. 2011).

The two transcriptional phases described during limb development are regulated by several elements, which have not been completely elucidated yet, among which are *cis*-regulatory elements (CRE), noncoding RNAs (ncRNAs) (Lemons & McGinnis 2006; Soshnikova 2014) and IRES (viral internal ribosome entry sites)-like elements (Xue et al. 2015). Among these, some elements have an internal location (the intergenic *cis*-regulatory elements), whereas other ones are located externally (global *cis*-regulatory elements) to *Hox* clusters (Di-Poï et al. 2010; Montavon et al. 2011). The DNA segments engaged in several regulatory activities of long range, interacting with *Hox* genes from inside a remote regulatory complex, characterize the so-called global control regions (GRCs), which define chromosomal regulatory landscapes (Deschamps 2007; Spitz 2010). In these domains, genes share regulatory specificities (Spitz et al. 2003; Zuniga et al. 2004) that coordinate gene expression in different developing structures (Duboule 1998; Spitz et al. 2001; Spitz et al. 2005). The *HoxD* cluster encompasses GCRs which regulatory mechanism in traditional experimental organisms (e.g. mouse) are so far the best understood ones among *Hox* genes. The collinear proximal-distal transcription of *HoxD* genes during limb development experiences a transition between two opposed regulatory landscapes, one located centromeric and the other telomeric from the cluster (see Fig. 0.1), which partially overlap but correspond

to two independent topological domains (Dixon et al. 2012; Nora et al. 2012; Andrey et al. 2013). Each *HoxD* gene exhibits its specific tropism in the interaction with these domains, which depends of their relative genomic position in the cluster (Andrey et al. 2013). In the three chapters of this PhD Thesis I have investigated evolutionary aspects of the *HoxD* cluster regulatory components in the context of morphological evolution, as detailed in the following section.

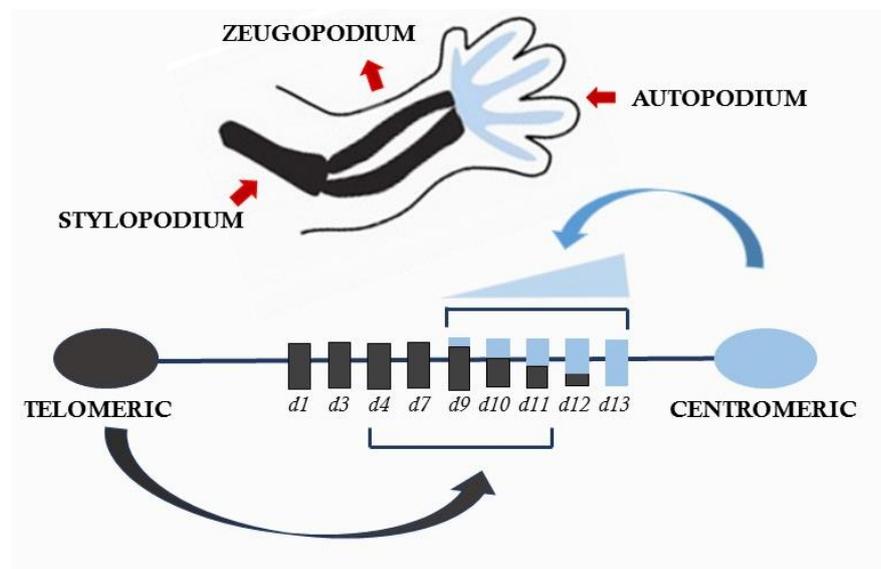


Figure 0.1. Mouse *Hoxd* genes are regulated by transcriptional enhancers located remotely from genes (in gray and blue), at opposed sides of the cluster. Enhancers that regulate initial transcription in proximal limb domains (gray) are located in the telomeric landscape of mouse chromosome 2, whereas enhancers regulating late expression at distal domains (blue) are located in a centromeric position. Adapted from Deschamps & van Nes (2005).

Mechanism of morphological evolution associated with Hox regulatory systems

The present PhD Thesis approaches regulation of *HoxD* expression in the context of morphological evolution using two chordate clades: Squamata (Sauria, Tetrapoda; chapters I and II) and Metatheria (Mammalia, Tetrapoda; chapter III). Squamata comprises lizards, snakes and amphisbaenians (Pough 1998), and represents a very intriguing biological system

for disclosing associations between molecular evolution and morphological divergence due to the recurrence in multiple lineages of remarkable morphological transitions involving limb reduction (and complete loss) associated with trunk elongation (Wiens & Slingluff 2001; Caldwell 2003). Several origins of snakelike morphologies occurred in this group, many of them associated with the occupancy of new environments imposing selective pressures that weaved or alternatively benefited from changes in developmental pathways (Coates & Ruta 2000; Wiens & Slingluff 2001; Woltering et al. 2009; Woltering 2012). However, *which* were these crucial changes in developmental pathways that produced new phenotypes, and *how* or *where* they have been settled in the genomes of these species, are issues that have sharpened the curiosity of scientists for decades, nurturing intense debate. In this context, it has been claimed that shifts in *Hox* gene expression patterns and associated downstream developmental networks might be strong candidates for connecting development with snakelike evolution, due to their consolidated roles in patterning both axial and appendicular axes (Cohn & Tickle 1999; Woltering et al. 2009; Di-Poï et al. 2010; Woltering 2012; Guerreiro et al. 2013), although evidences are still controversial (see Sanger & Gibson-Brown 2004; Head & Polly 2015). The anterior expansion of *HoxC6* and *HoxC8* gene expression domains along python paraxial and lateral plate mesoderm has been accepted for a long time as a developmental explanation for concomitant forelimb loss and trunk elongation (Cohn & Tickle 1999), although recent evidence suggests that functional changes in primaxial *Hox* genes did not cause axial de-regionalization in snakes (Head and Polly 2015). Other authors, conversely, suggest that alternative downstream interpretations of *Hox* codes in the pre-caudal region seems to determinate vertebral de-regionalization in caecilians and snakes (Woltering et al. 2009), which in snakes led to expansion of the rib cage (Guerreiro et al. 2013).

Beyond interest in the causal mechanisms that establish snakelike morphologies, questions concerning the evolution of genes essential for development of vertebrate structures

that no longer exist in snakelike lineages (i.e. the limbs) remain relegated in EvoDevo studies. The scarce information available for evolution of *Hox* genes in snakelike squamates is restricted to Serpentes and often focus on coding sequences or gene expression patterns (Cohn & Tickle 1999; Di-Poï et al. 2010; Woltering 2012; Head & Polly 2015; but see Singarete et al. 2015 for some exceptions), so comparisons of developmental processes among snakelike squamate lineages are practically inexistent, especially considering non-coding regions. In the first two chapters of this PhD Thesis (Chapters I and II), I have investigated molecular signatures of snakelike morphologies in *Hoxd* regulatory elements; I did not seek particularly for cause-effect correlations between molecular and morphological evolution. Because *Hox* genes are involved in multiple functions (Mallo et al. 2010; Montavon et al. 2011; Archambeault et al. 2014) and changes in their coding regions likely involve pleiotropy (Sivanantharajah & Percival-Smith 2015), variation in expression patterns associated with morphological evolution seem to concentrate on evolution of CREs (Warren et al. 1994; Averof & Akam 1995; Carroll 1995; Burke et al. 1995). Mechanisms of CRE evolution involve mutational loss and loss in the number, affinity and topology of transcription factor binding sites (TFBS; see Jeong et al. 2006; Wang & Chamberlin 2002; Gompel et al. 2005), besides the co-option of TFBS in pre-existing CREs (Wang & Chamberlin 2002; Gompel et al. 2005). Even single insertions or deletions among TFBS have the potential to affect function of regulatory elements and expression levels of genes regulated by them (Zinzen et al. 2006; Carroll 2008). Based on this scenario, in Chapter I, I have investigated whether independent evolution of snakelike morphologies in Serpentes and Amphisbaenia 1 mmolecular signatures in the Conserved Element B (*CsB*), a *HoxD* centromeric regulatory region. In Chapter II, I have evaluated whether elements that regulate *HoxD* gene expression [one centromeric (*Island D*) and one telomeric (*CNS65*) element] and are specially relevant for tetrapod limb

development retained in limbless organisms (snakes) their functional ability to regulate gene expression in limb buds.

The second taxonomic group used to evaluate regulation of *HoxD* expression in the context of morphological evolution was Metatheria. Macropodidae marsupials, such as the tammar wallaby *Macropus eugenii*, exhibit several peculiar limb characteristics in comparison with eutherian mammals. Their limb differences appear further beyond the adult morphology, being also registered in the heterochrony observable during development of forelimbs and hindlimbs. Forelimbs develop considerably early and allow the young newborn to climb its mother's body right after birth, encountering the pouch by himself while still blind and holding fetal hindlimbs (Sears 2009; Keyte & Smith 2010; Chew et al. 2014). Fascination does not end here because, after birth, hindlimb development surpasses forelimb growth, and they establish as very long and strong limbs specialized for hopping; yet, they present a digital formula that is unique to macropodids but includes an homoplastic component (Chew et al. 2012). Syndactylous, reduced and skin-bound digits II and III have evolved independently in two marsupial orders (Nilsson et al. 2004; Phillips et al. 2006), recapitulating a similar scenario to the aforementioned recurrent evolution of snakelike morphologies. However, information about the evolution of developmental genetic systems underlying the origin of such heterochronic processes and the establishment of these morphologies remains scarce for Macropodidae marsupials (Chew et al. 2012; Deakin 2012). Therefore, in the third chapter of this PhD Thesis (Chapter III) I have examined ncRNAs from the *HoxD* cluster present in *M. eugenii* developing limbs, aiming to disclose unknown regulatory elements that potentially emerged in the marsupial lineage that might have contributed to evolution of their peculiar limb morphological and developmental characteristics.

The core of this PhD Thesis settles on unraveling links between molecular evolution of *HoxD* cluster regulatory regions and the evolution of morphological diversity using two vertebrate groups as targets, squamates and marsupials.

CHAPTER I

The independent evolution of snakelike squamate lineages registered both convergent and divergent regulatory signatures in the HoxD cluster enhancer *CsB*.

1.1 ABSTRACT

Snakelike morphologies evolved independently several times within Squamata. Evolution of morphological diversity is often coupled with variation in developmental patterns of *Hox* gene expression, which may reside in nucleotide substitutions at regulatory regions. The genes *Hoxd10-13* are particularly relevant in this evolutionary context because they play essential roles during the development of structures remarkably modified in snakelike lineages: autopodia and axial anterior-posterior (AP) axis. We compared between Serpentes and Amphisbaenia the molecular signatures of a regulatory *Hoxd* element [Conserved Sequence B (*CsB*)], using lizards and other tetrapods as reference for the lacertiform morphology. Conserved regions within *CsB* (*CP1* [725 bp] and *CP2* [423 bp]) were sequenced from 38 squamate and 1 alligator species; our database was complemented by sequences from GenBank for chicken, turtle and mammals. Transcription factor binding sites (TFBS) were predicted by a pipeline that merges binding sites preferences to functional and evolutionary clustering. Variation within Squamata concentrated within *CP2*, whereas differences between squamates and non-squamates appeared also in *CP1*. Regulatory haplotypes of amphisbaenians seemed more similar to those of other lizards than to those of snakes. However, both snakelike lineages shared specific regulatory signatures not identified in lacertiform lizards, which were located in a short segment of *CP2* (initial positions 841-849 in our alignment) comprised by: 1) lack of TFBS for C-MAF, HIC1, NEUROD or TTF-1, AP-4 and XPF-1 or LBP-1, and 2) replacement at position 846 of a TFBS for HIC1, present in lizards, by one for PBX1. Molecular evolution of *CP2* in Squamata left convergent regulatory signatures during the independent origins of Serpentes and Amphisbaenia that may relate to body elongation and limb loss, and suggests a possible function of this element on AP axis in the developing vertebrate embryo, so far undescribed.

1.2 INTRODUCTION

Independent origins of equivalent morphologies are identified when a given phenotype is observed in relatively close taxonomic lineages but is absent in representatives of the group including their most recent common ancestral (Hall 2012). These origins may represent either parallel (based on the same developmental genetic mechanisms) or convergent (based on distinct developmental genetic mechanisms) processes (Wake et al. 2011). Patterns of homoplastic evolution are driven by specific mechanisms, such as natural selection, which depend on their respective evolutionary context (Grant 1999; Rundle et al. 2000). However, equivalence in a given morphological transition repeated along different lineages might have been settled through similar changes in developmental pathways (Sucena et al. 2003; Shapiro et al. 2006; Prud'homme et al. 2006; Abouheif 2008). In a scenario where the developmental bases of morphological similarity remain partially obscure, referring to such processes as 'independent origins of equivalent morphologies' is more accurate than classifying phenotypes as 'convergent' or 'parallel'. This consideration does not apply however to the genetic level, where nucleotide signatures that are common or exclusive to compared lineages may be respectively referred to as 'convergent' or 'divergent' molecular patterns. The conceptual counterpoint between these two biological levels is particularly relevant given that possible developmental bases of independent origins of equivalent morphologies have been poorly explored using empirical data, so the correspondence between phenotypic and genetic transitions during processes resulting in multiple independent origins of equivalent morphologies remains uncertain.

Among several examples of independent evolution of equivalent phenotypes, the multiple origins of snakelike morphologies within Squamata have received considerable attention in the past years. The snakelike body shape is characterized by an elongated body axis composed of 'de-regionalized' and numerous vertebrae (when compared to lacertiform

morphologies; see (Woltering 2012) and reduced or absent limbs (Wiens & Slingluff 2001; Caldwell 2003). In Squamata, snakelike morphologies evolved independently from lacertiform ancestors at least 26 times (Shapiro et al. 2003; Wiens et al. 2006), being observed in the families Amphisbaenidae, Anguillidae, Cordylidae, Dibamidae, Gerrhosauridae, Gymnophthalmidae, Pygopodidae, Scincidae and in the Serpentes clade (Wiens et al. 2006; Uetz 2007). In some lineages the evolution of this conspicuous morphology has occurred more than once (Pellegrino et al. 2001; Brandley et al. 2008), and it is widely accepted that the snakelike phenotype has evolved independently in snakes and amphisbaenians (Vidal & Hedges 2005; Vidal & Hedges 2009; Conrad 2008; Losos et al. 2012; Pyron et al. 2013), as illustrated in Figure 1.1. During these multiple transitions from lacertiform to snakelike morphologies, body elongation is recurrently associated with limb reduction, which suggests the presence of developmental constraints acting on limb development in the elongated taxa (Woltering 2012). Such constraints likely involve *Hox* genes expression because these pattern the anterior-posterior (AP) axis in the developing vertebrate embryo, and especially those from clusters A and D were apparently co-opted during the evolution of limb (Duboule 1992; Zákány & Duboule 1999; Tschopp & Duboule 2011) and genitalia (Dollé et al. 1989; Zákány & Duboule 1999; Tickle 2006; Spitz 2010; Tschopp et al. 2014) in tetrapods.

As mentioned in the General Background section, *Hoxd* genes are essential for development of both AP body axis (caudal regions of somites, neural tube, notochord, intestine, cloaca and external genitalia) (Griffith et al. 1992; Gajovic et al. 1993; Catala et al. 1995; Knezevic et al. 1998; Santa Barbara & Roberts 2002; Scott et al. 2005; Mallo et al. 2010) and limbs (Capdevila & Belmonte 2001; Zakany & Duboule 2007). Among these genes,

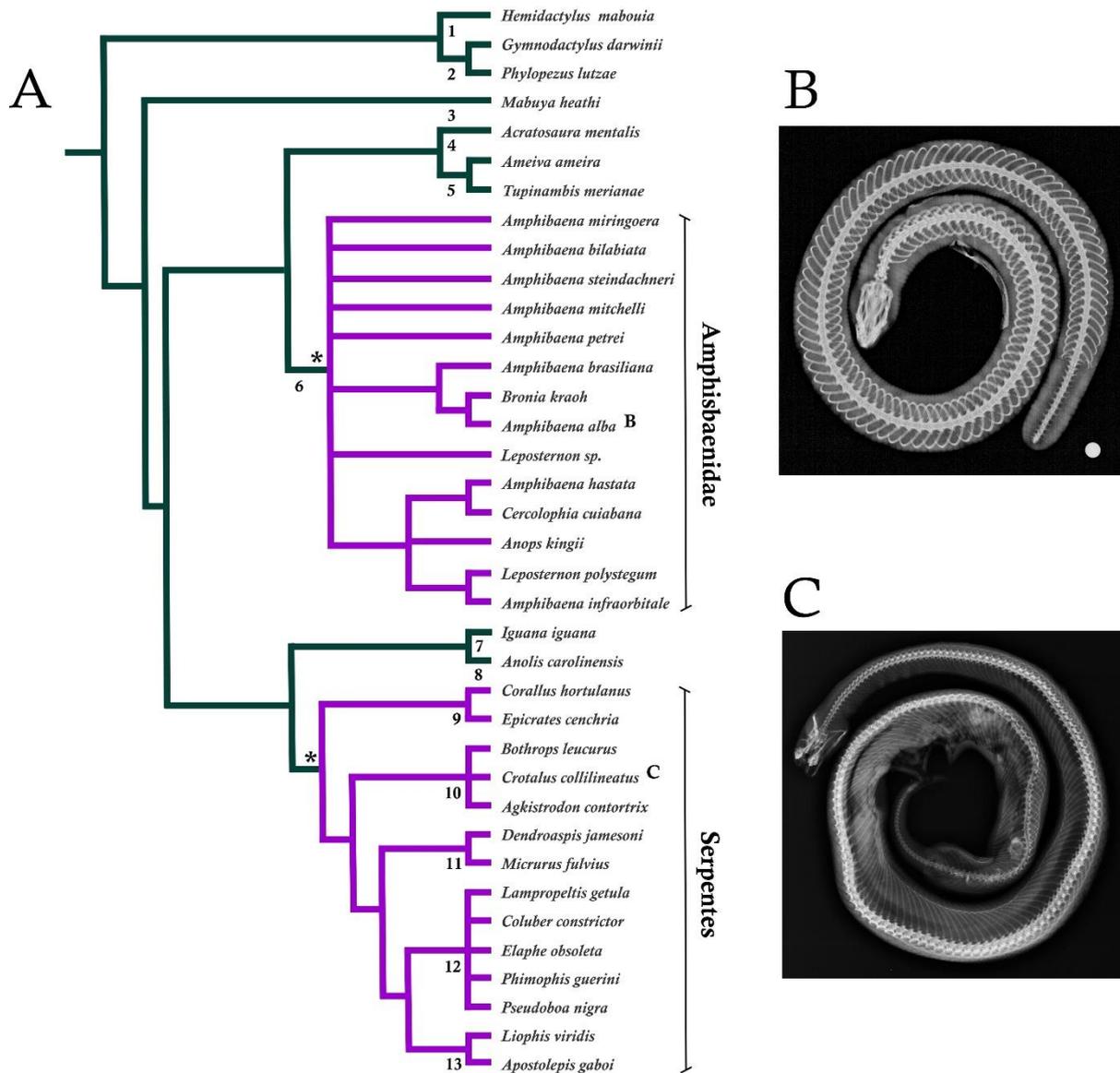


Figure 1.1. Squamata phylogenetic relationships and the snakelike morphology. (A) Topology illustrating phylogenetic relationships among the squamate taxa from which *CsB* fragments were sequenced, according to Pyron et al. (2013) and Mott & Vieites (2009), coupled with x-rays illustrating the similar morphologies of amphisbaenians (B), represented by *Amphisbaena alba*, and snakes (C), represented by *Crotalus collilineatus*. In (A) the asterisk highlights snakelike lineages, represented by purple branches. Green branches indicate lineages with a lacertiform morphology. Numbers correspond to squamate families as follows: 1. Gekkonidae, 2. Phyllodactylidae, 3. Scincidae, 4. Gymnophthalmidae, 5. Teiidae, 6. Amphisbaenidae, 7. Iguanidae, 8. Dactyloidae, 9. Boidae, 10. Viperidae, 11. Elapidae, 12. Colubridae, 13. Dipsadidae. X-rays gently given by Mariana Bortoletto Grizante.

terminal ones from the *HoxD* cluster represent good candidates for a genetic correspondence with phenotypic transitions during the independent origins of snakelike morphologies because they are involved in the establishment of both the AP axis and the limbs. During development of the axial skeleton, *Hoxd10* is expressed in mouse lumbar and sacral vertebrae and triggers interruption of ribs formation, inducing thoracic-lumbar transition, whilst *Hoxd11* is involved in the lumbar-sacral transition (Davis & Capecchi 1994; Burke et al. 1995; Wellik & Capecchi 2003; Carapuço et al. 2005). Two other genes, *Hoxd12* and *Hoxd13*, are expressed in mouse caudal vertebrae, the latter being involved in the mechanism that terminates somitogenesis, which modulates the final number of caudal vertebrae (Economides 2003; Kmita & Duboule 2003; Soshnikova & Duboule 2009). Limb development also involves *Hoxd* gene expression (Zakany & Duboule 2007; Nelson et al. 1996; Tarchini & Duboule 2006), and the same terminal genes aforementioned (*Hoxd10*, *Hoxd11*, *Hoxd12* and *Hoxd13*) are essential for canonical autopodium and zeugopodium formation (Tickle 2006; Montavon et al. 2008; Montavon et al. 2011).

Expression of *Hox* genes in vertebrates is mainly regulated by *cis-trans* transcriptional mechanisms, although a translational control is also involved (Kondrashov et al. 2011; Montavon & Soshnikova 2014). The regulatory machinery modulating expression patterns of terminal *Hoxd* genes during limb development has been very well characterized in the mouse: a landscape located telomeric to *HoxD* cluster regulates gene expression during stylopodium and zeugopodium formation (Montavon et al. 2011; Andrey et al. 2013), whereas autopodium development relies on a second transcriptional wave mastered by a centromeric regulatory landscape that coordinates *Hoxd13* to *Hoxd10* inverted collinear expression (Tickle 2006; Montavon et al. 2008; Montavon et al. 2011). This centromeric regulatory landscape comprises several enhancers which complex interaction modulates expression of terminal *Hoxd* genes in the embryo in the so-called ‘distal phase’ (DP) expression pattern. This pattern

has been mostly established based on developmental processes of digit formation (Tickle 2006; Montavon et al. 2008; Montavon et al. 2011), but it was recently shown to regulate several other distally elongated structures in vertebrates, representing an ancestral feature co-opted to pattern diverse body plan features (Archambeault et al. 2014). The Conserved Sequence B (*CsB*), in particular, belongs to a global control region (GCR; see (Spitz et al. 2003; Gonzalez et al. 2007) that regulates *Hoxd10-13* genes expression during development of autopodia, external genitalia and the dorsal neural tube (Spitz et al. 2003; Gonzalez et al. 2007; Schneider & Shubin 2013). Among vertebrates, three domains within *CsB* (hereafter described as conserved peaks *CP1*, *CP2* and *CP3*) are extremely conserved (Gonzalez et al. 2007; Schneider & Shubin 2013), and the first two encompass tetrapod-specific elements (nominated *B1* and *B2* by (Gonzalez et al. 2007).

Comparative analyses focusing on the *CsB* element seem especially promising for evaluating the likelihood of molecular convergence coupled to the independent origins of snakelike morphologies because the relevance of this region for *Hoxd* gene expression has been already well characterized, and there is indication of variation in the expression patterns of *Hox* genes during snake development (e.g. Di-Poï et al. 2010; Woltering 2012). For example, *Hoxd10* and *Hoxd11* are expressed more posteriorly in the AP body axis of snakes than in mouse and lizard, a pattern likely related to the elongated trunks lacking the transition from rib-less to rib-bearing dorsal vertebrae regions (equivalent to thoracic-lumbar transition in mammals; see (Dequéant & Pourquié 2008; Woltering et al. 2009; Di-Poï et al. 2010; Woltering 2012). Moreover, *Hoxd12* gene is not found in the genome of snakes whereas *Hoxd13* expression is weak in the tail region, opposed to their strong expression in the lizard tail (Di-Poï et al. 2010); such differences apparently contribute to a delay in somitogenesis termination and likely contribute to tail elongation in Serpentes (Burke et al. 1995; Di-Poï et al. 2009; Woltering 2012). Despite recent discussions regarding developmental mechanisms

underlying the evolution of a snakelike body shape (e.g. (Cohn & Tickle 1999; Kohlsdorf & Wagner 2006; Di-Poï et al. 2009; Woltering 2012; Head & Polly 2015), to our knowledge the topic remains unexplored under a comparative phylogenetic framework where regulatory footprints are addressed in the scenario of independent origins of snakelike morphologies in Squamata [but see Singarete et al. (2015) for a recent study on protein evolution]. The present study aims to establish this connection between genotypic and phenotypic dimensions in the context of multiple origins of snakelike morphologies. Specifically, I have performed comparative analyses of a *CsB* fragment comprising the first two conservation peaks (*CP1* and *CP2*), and focused on regulatory haplotypes that are shared by all species of Serpentes or Amphisbaenia evaluated and *exclusive* to these respective groups (hereafter referred to as ‘divergent signatures’) and those that are present in all species of both lineages but absent in lacertiform species (hereafter referred to as ‘convergent signatures’). I tested the following hypothesis: 1) some regulatory signatures in the *CsB* fragment regarding predicted transcription factor binding sites (TFBS) would be exclusive to Serpentes or to Amphisbaenia, a pattern that likely reflects the independent evolutionary histories of these two lineages; 2) concomitantly, it would be possible to identify equivalent regulatory signatures in the *CsB* that seem coupled to the independent evolution of snakelike morphologies in these two lineages, providing evidence for possible developmental constraints acting during the different origins of such phenotypes. Both divergent and convergent regulatory signatures within *CP2* fragment in snakes and amphisbaenians were found, suggesting that they may have been registered by axial elongation and limblessness evolution. The results obtained here were complemented with a functional dissection of *CsB* through mice transgenic assays, performed by Prof. Dr. Igor Schneider, to explore whether the fragment analyzed here would alone cover the regulatory capacity held by the entire *CsB* element in limbs and AP axis.

1.3 METHODS

Sequencing CsB fragments

Fragments of *CsB* were amplified and sequenced from 14 species of Serpentes, 14 species of Amphisbaenia, 9 lizard species and one alligator (see Fig. 1.1). Voucher numbers of tissue samples and corresponding GenBank accession numbers for the sequences are provided in the supplementary material (see appendix for Table A1.1). Genomic DNA for sequencing *CsB* fragments was obtained from: 1) aliquots given by Yale Peabody Museum of Natural History from Yale University (New Heaven, EUA) and the authors of Vargas et al. (2008) (*A. mississippiensis*) 2) tissue samples available at Herpetological Collection from Laboratory of Tetrapod Evolution (*Laboratório de Evolução de Tetrápodes*, LET) at *Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto* (FCLRP-USP), Ribeirão Preto, Brazil; *Conselho de Gestão do Patrimônio Genético* (CGen) registry n° 02000.002450/2011-23), and 3) tissue samples gently given by Herpetological Collections at *Universidade Federal do Mato Grosso* (UFMT, Cuiabá, Brazil), from *Universidade Federal do Alagoas* (MUFAL, Maceió, Brazil), and from *Universidade de Brasília* (CHUNB, Brazil); from Museum of Science and Technology at *Pontifícia Universidade Católica do Rio Grande do Sul* (MCP, Porto Alegre, Brazil); from Zoology Museum at USP (MZUSP, São Paulo, Brazil); and from *Universidade do Estadual de São Paulo* (UNESP, Rio Claro, Brazil).

I extracted DNA from samples previously preserved in 95% ethanol using the DNAEasy Tissue kit (Qiagen), following protocol provided by the manufacturer. Regions *CP1* and *CP2* were amplified by polymerase chain reaction (PCR) using the following primers: region *CP1* (approximately 652-699 bp) - forward 5' GTCCTTCTGTCTAACTAATAATTGC and reverse 5' GCTTCTATTATGATCTCTTG;

region CP2 (approximately 403-407 bp) - forward 5' CAGATCAAGAGATCATAATAG and reverse 5' TTGTAATGCTAACAGGCAAT. Annealing temperatures used in the PCR ranged between 50 and 54°C (see appendix for Table A1.2). PCR efficacy was verified by 1% agarose gel electrophoresis, subsequently repeated with doubled volume (50 µl) and ran in 2% agarose gel. Bands of interest, presenting size correspondent to fragments amplification, were purified using *Gel Purification* kit (*Qiagen*), according to manufacturer's protocol. Purified PCR products were ligated to pGEM-t vector (Promega) using T4 ligase enzyme, incubated overnight at 16°C, and used in heat-shock transformation of thermo-competent *E. coli* (DH5α or Top10 lineages). Transformed cells were plated in solid agar containing ampicillin, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and Isopropyl β-D-1-thiogalactopyranoside (IPTG), to ensure double selection of clones containing insert. Six white colonies from each Petri dish, corresponding to clones expected to encompass the gene of interest, were incubated in Lysogeny broth (LB) for 14 hours at 37°C. Concomitantly, PCRs using colonies smear as templates and T7/SP6 primers (given at the pGEM-t vector/Promega protocol) or respective specific primers were performed to ensure that the colonies selected contained inserts of expected sizes. Cultures containing the insert were purified using the *PureYield™ Plasmid Miniprep System Protocol* kit (Promega), following the protocol provided by the manufacturer, but one fraction of each clone (~500 µl) was stored in 50% glycerol. Multiple clones (from 3 to 5 for each species) were sequenced in both directions (forward and reverse, using respectively the primers T7 and SP6. This stage consisted in performing PCR reactions containing 350 µg of plasmid DNA, primers T7 or SP6 (5 pM), *Byg Dye Terminator* v.3 and 5X *Big Dye* buffer. PCR products were precipitated using isopropanol (65%) and ethanol (60%) before plates were sent to sequencing.

A consensus sequence for each species was obtained from the sequences confirmed in BLASTn (NCBI) as corresponding to the *CsB* element. For each one of the species

aforementioned, a consensus was first obtained for each cloned sequenced bidirectionally using T7 and SP6 primers, and then a consensus of at least three clones was produced. Consensus sequences obtained for *CP1* and *CP2* regions were concatenated into a single *CsB* fragment. Additional *CsB* sequences were retrieved from the *UCSC Genome Browser* database using the BLAST tool searching genomes of *Gallus gallus*, *Pelodiscus sinensis*, *Anolis carolinensis*, *Mus musculus* and *Homo sapiens* (see appendix for Table A1.1). All sequences were aligned using the Clustal-V algorithm (Higgins et al. 1992) implemented in the software BioEdit, and the resulting alignment was used in the subsequent analyses, which aimed to find regulatory signatures registered in the snakelike lineages compared (see below). Analyses were performed having a major focus on comparisons between snakes and amphisbaenians, and can be grouped in two categories: 1) cluster analyses, and 2) estimation of TFBS.

Cluster analyses

Cluster analyses based on Maximum Likelihood were implemented in JModelTest 2 (Darriba et al. 2012), a program that also calculates the most suitable model of evolution prior to tree estimation. The base of tree search chosen for computing likelihood was Best, which automatically selects the most appropriate parameter between NNI and SPR for each analysis. The model of evolution HKY+I+G was estimated using BIC (Bayesian Information Criterion). Heuristic search estimation was set for a maximum of 5,000 trees. In order to reassure our results, we also performed cluster analyses using Parsimony, implemented in PHYLIP 3.695 (Felsenstein 2005), and 1,000 bootstrap replicates were applied to confer statistical confidence to nodes (Felsenstein 1985; Hillis & Bull 1993). Heuristic searches performed in both analyses were unrooted.

Prediction of Transcription Factor Binding Sites

Prediction of TFBS was performed individually for each sequence of the alignment by an in-house developed pipeline that uses the Match algorithm (Kel 2003) and TRANSFAC vertebrate databases, in collaboration with Prof. Dr. Pedro E. M. Guimarães from *Universidade Federal de Uberlândia*, Brazil. In order to avoid the prone to predict false positive TFBS with little or none biological value attached, very common in sequence analysis, we implemented a series of steps to improve the reliability of our data. After comparing results using different minimum detection thresholds, we settled the value for this study as 0.8 for both the matrix core and the complete matrix profiles. Lower values were too permissive and increased noise in the analyses, while higher values were too strict and resulted in loss of biological relevant binding sites (data not shown). TFBS matrices were clustered into families of functional and evolutionary related transcription factors, based on the hierarchical classification provided by TRANSFAC database. The family clustering procedure accommodates the expected TFBS sequence variation associated with the evolutionary history of each genome evaluated, avoids the by-product sequence changes deriving from relaxation of selective constraints and reduces the prediction of overlapping binding sites (Balhoff & Wray 2005; Akhtar & Veenstra 2011). A total of 1106 vertebrate matrices were clustered into 72 TFBS families of functional and evolutionary related transcription factors. Consequently, the possibility of analyzing regulatory sequences across species was expanded and comparisons of genomes phylogenetically as distant as human and the coelacanth were enabled. The predicted binding sites were transformed into strings of TFBS families, respecting the prediction order and recording the relative position of the regulatory elements. The relative position was used to identify TFBS signatures constrained during evolution, an approach recognized as *phylogenetic footprinting*, based on the principle that functional regulatory sequences show lower mutation rates than non-functional. Using the prediction

pipeline associated with the control steps described above we searched for TFBS that are conserved in all sequences from predefined sets of species, at the same relative position in the alignment, but are not predicted in the same position in the other groups compared. Estimated TFBS often extend for 5-8 bp, so we settled the initial relative position in the alignment of a given predicted TFBS as the reference point for comparing positional differences among groups. Aware of the restrictions of predictions based algorithms and the necessity of experimentally validate the data, we used the above mentioned approaches to identify gains and losses of reliable TFBS during the evolution of *CsB* across tetrapods.

CHAPTER II

Limb loss in snakes and degeneration of *CNS65* and *Island I* regulatory capacities: morphological evolution left footprints on *Hoxd* genes enhancers.

2.1 ABSTRACT

The Serpentes clade comprises elongated squamates either entirely limbless or exhibiting vestigial hindlimbs. Current literature recently incorporated evidence for associations between the origin of snakelike morphologies and evolution of *Hox* genes, but information about *Hox* enhancers during Serpentes evolution remains scarce. Regulation of *Hoxd* gene expression is of particular interest due to their relevance for limb development. Here I combine transgenic assays with bioinformatics to explore footprints of snake limb loss in *Hoxd* enhancers essential for developing the tetrapod limb. Reporter expression of telomeric [*CNS65*] and centromeric [*Island I*] *Hoxd* enhancers from snakes in transgenic mice were performed to investigate if these retained their regulatory capacities. Activities of both enhancers from snakes were abrogated in the limbs of most mouse transgenic embryos, which contrasts with the strong regulatory signals reported in the literature for assays performed with both enhancers from mouse, and also fish for *Island I*. Regulatory signatures in *CNS65* and *Island I* were evaluated through prediction of transcription factor binding sites (TFBS). Comparisons between snakes and ‘limbed reptiles’ (lizard, alligator, turtles and birds) revealed that 43-52% of TFBS predicted in all ‘limbed reptiles’ except snakes were limb-related; when snake enhancers were compared with those from other limbed amniotes (reptiles and mammals), 44-70% among the predicted TFBS diverging between the two groups were limb-related. I identified one segment that potentially represents a stilopodium/zeugopodium-specific element in *CNS65* and three possible autopodium-specific elements in *Island I*, comprised by several binding sites for limb-related transcription factors estimated in all limbed amniotes that seem lost by snakes. In conclusion, loss of *CNS65* and *Island I* regulatory capacities in snakes suggests existence of regulatory segments specifically dedicated to limb development in these enhancers that apparently degenerated after limb loss. Regulatory signatures suggest that snakes lost at least four segments that in limbed amniotes are likely enriched by binding sites for transcription factors essential for limb development. Our results suggest that these specific regulatory signatures could represent the major sites which degeneration caused by relaxation during evolution of Serpentes impaired ability of these enhancers to regulate *Hoxd* expression on developing limbs, providing clues on the specific mechanisms of *Hox* genes molecular evolution linked to the origin of snakes. Additionally, they reveal segments within *CNS65* and *Island I* that possibly belong to regulatory modules strictly associated with vertebrate limb development.

2.2 INTRODUCTION

The snakelike morphology characteristic of *Serpentes* remarkably diverges from standard vertebrate body shapes due to the lack of limbs and extreme elongation of the trunk; this phenotype evolved from lacertoid lizard ancestors exhibiting limbs (Coates & Ruta 2000; Tchernov et al. 2000; Bejder & Hall 2002) not only in this clade but also in many other squamate lineages (see Sanger & Gibson-Brown 2004; Wiens et al. 2006; Brandley et al. 2008; Woltering et al. 2009; Kohlsdorf et al. 2010). All snakes lack forelimbs and scapular girdles (Cohn & Tickle 1999; Woltering 2012). Colubroids, commonly referred as “advanced” snakes, are entirely limbless and have axial skeletons composed by numerous rib-bearing vertebrae; these trunks seem more homogeneous than those of lineages having earlier origins (scolecophidians and booids), which retained rudimentary pelvic girdles and a vestigial femur (Bellairs & Underwood 1950; Cohn & Tickle 1999). Evolution of a phenotype so remarkably divergent from its ancestor both regarding structures gain (e.g. number of vertebrae) and traits loss (i.e. limbs) has intrigued scientists for decades, being subject of intense debate focusing not only on the historical circumstances of snake origins (see Caldwell & Lee 1997; Coates & Ruta 2000; Tchernov et al. 2000; Rieppel et al. 2003; Vidal & Hedges 2004; Lee 2005; Apesteguía & Zaher 2006), but also on the developmental variations that resulted in such elongated limbless morphologies (see Cohn & Tickle 1999; Wiens & Slingluff 2001; Woltering et al. 2009; Woltering 2012).

Developmental changes identified as possible candidates for explaining the origin of *Serpentes* have placed *Hox* genes at the focal point of this discussion. As detailed in the General Background section, current literature sustains that shifts in *Hox* gene expression patterns and downstream developmental pathways are the likely connection between morphological transitions and developmental pathways in the context of snake origins, as

these genes play essential roles in patterning both axial and appendicular axes (Cohn & Tickle 1999; Woltering et al. 2009; Di-Poï et al. 2010; Woltering 2012; Guerreiro et al. 2013, but see Head and Polly 2015; Sanger and Gibson-Brown 2004 for controversial opinions). Snake evolution has registered specific signatures in the sequence of *Hoxa13* first exon (Kohlsdorf et al. 2008), a pattern apparently absent in the sequence of other snakelike lineages, as amphisbaenians and caecilians (Singarette et al. 2015). Together, these examples reinforce the idea that the way *Hox* genes are expressed and how they interact with other molecules ought to have been modified during the evolution of snakes. Paradoxically, our knowledge about *Hox* enhancers in snakes remains scarce. The present study aims to fill this gap by testing whether *Hox* enhancers involved in the development of tetrapod limbs remain regulatory functional after ~100 million years of limb loss in Serpentes (Brandley et al. 2008; Vidal et al. 2010; Pyron et al. 2013; Reynolds et al. 2014; Head 2015). This question is particularly intriguing given that *Hox* genes are often involved in multiple functions during embryo development (e.g. Fromental-Ramain et al. 1996; Deschamps & van Nes 2005; Scott et al. 2005; Zakany & Duboule 2007; Archambeault et al. 2014), and therefore sequence changes in these genes likely involve considerable pleiotropy (Sivanantharajah & Percival-Smith 2015).

In the context of limbless morphologies evolution, regulation of *Hoxd* genes expression is of particular interest due to their relevance for development of the autopodium, a feature absent in Serpentes. One can now study in limbless organisms some of the key regulatory elements involved in tetrapod limb development because the regulatory dynamics of *Hoxd* gene expression has been recently very well characterized in mice limb buds (Gonzalez et al. 2007; Rinn et al. 2007; Montavon et al. 2008; Montavon et al. 2011; Tschopp & Duboule 2011; Tschopp et al. 2011; Delpretti et al. 2013; Soshnikova 2014). During limb development (as well as when other vertebrate appendages

are developing; see Archambeault et al., 2014), *Hoxd* gene expression occurs in two independent phases that will establish proximal and distal segments (Nelson et al. 1996; Tarchini & Duboule 2006). The initial transcription phase is a homogeneous and sequential expression of *Hoxd1-11*, which together with other genes coordinate patterning of stilopodium and zeugopodium (Zakany & Duboule 2007; Montavon et al. 2011). The second transcriptional phase defines autopodium formation and involves expression of *Hoxd9-13* genes in an inverted spatio-temporal collinear mode (Tickle 2006; Montavon et al. 2008; Montavon et al. 2011). These transcriptional phases are regulated by several overlapping elements including noncoding RNAs (Soshnikova 2014) and cis-regulatory elements (CREs), among which are intergenic and global enhancers (Di-Poï et al. 2009; Montavon et al. 2011; Andrey et al. 2013). Regulatory landscapes comprise many CREs that are located far away from the genes they regulate (Deschamps 2007; Spitz 2010; Montavon et al. 2011; Andrey et al. 2013): regulation of the first transcriptional wave during stylopodium/zeugopodium development engages elements from the telomeric landscape (downstream of *HoxD* genes; see Tarchini & Duboule 2006; Andrey et al. 2013; Woltering et al. 2014), while autopodium development involves a shift between regulatory domains and is controlled by a centromeric region located upstream of *Hoxd13* (Dixon et al. 2012; Nora et al. 2012; Andrey et al. 2013).

The essential roles for limb development performed by some CREs located in these regulatory landscapes may guide our view towards regulatory elements likely modified during the evolution of limbless phenotypes. For example, *Hoxd* gene expression along the anterior-posterior limb bud axis is modulated by the *early limb control region* (ELCR) located at the *HoxD* telomeric regulatory landscape (Zákány et al. 2004; Deschamps & van Nes 2005; Tarchini & Duboule 2006; Schneider et al. 2011), which comprises several conserved noncoding sequences (CNS) assigning enhancers for limb development (Andrey et al.

2013). Among these, one fragment named *CNS65* induces the initial transcription of *Hox* genes in a broad proximal domain of developing limb buds and, when deleted, reduces in about 30-40% the physiological levels of *Hoxd8-11 mRNAs* at the most proximal portion of mice limb buds (Andrey et al. 2013). The centromeric landscape also comprises several regulatory elements involved in the development of limb distal portions and the central neural system (Spitz et al. 2003; Gonzalez et al. 2007; Montavon et al. 2011), and a 600 Kb region so-called genic desert located between *Lnp* and *Atp5g3* contains a 'regulatory archipelago', comprised by *Island I* to *V*, essential for autopodium development (Montavon et al. 2011). The complete transcriptional activation of terminal *Hoxd* genes requires interactions among islands I and IV, and other enhancers such as the ones in the global control region (*GRC*) and *Prox*; mouse *Island I* reporter expression shows that its activity occurs from the proximal extremity of the bud to the top of the paddle across the region of digits III and IV in E12.5 mouse embryos (Montavon et al. 2011). Enhancers located in these two regulatory regions, such as the *CNS65* (ELCR at the HoxD telomeric regulatory landscape) and *Island I* (genic desert at the HoxD centromeric regulatory landscape), represent therefore proper candidates for testing if the regulatory machinery involved in tetrapod limb development has degenerated in Serpentes. Differences in *Hox* gene expression associated to morphological evolution likely entail variation in CREs (Warren et al. 1994; Averof & Akam 1995; Carroll 1995; Burke et al. 1995) because these genes are often involved in multiple functions (Sivanantharajah & Percival-Smith 2015) and a complex, fine-tuned association of several Hox regulatory elements confers both specificity for gene expression in particular structures and evolutionary flexibility (Duboule 2007; Gonzalez et al. 2007; Deschamps 2007; Montavon et al. 2011). Consequently, sequence variation affecting, for example, transcription factor binding sites (TFBS) of a given enhancer may impose a functional loss (Zinzen et al. 2006) that in *Hox* genes could affect

only one developmental function or involve considerable pleiotropy, depending on the specificity of the regulatory module (see Tschopp et al. 2011). This matter as approached in the context of Serpentes evolution by focusing on one element from HoxD telomeric (*CNS65*) and another one from the centromeric (*Island I*) regulatory landscapes to test if regulatory capacities of limb-specific enhancers degenerated in limbless lineages. Specifically, I combined bioinformatics predicting TFBS with reporter expression of *CNS65* and *Island I* from snakes in transgenic mouse. I retrieved genome sequences for *Python bivittatus* (Pythonidae, Boidae) (Fig. 2.1A), a snake presenting femoral vestiges, and *Ophiophagus hannah* (Elapidae, Colubroidae) (Fig. 2.1B), a completely limbless snake, to construct plasmids for transgenic assays; the sequences were also compared with limbed amniotes regarding predicted TFBS. I predict that evolution of Serpentes involved functional impairment of *CNS65* and *Island I* regulatory capacities in developing limbs, a degeneration of these sequences characterized by the loss of TFBS involved in limb developmental pathways.

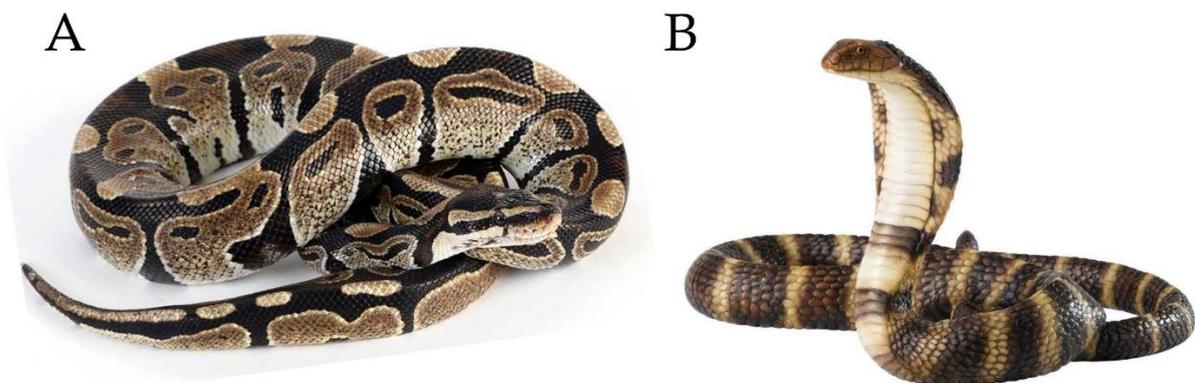


Figure 2.1. Images of representatives from *Python molurus bivittatus* (Pythonidae, Boidae), bearer of femoral vestige (A), and *Ophiophagus hannah* (Elapidae, Colubroidae), completely limbless (B).

2.3 METHODS

Gene reporter assays: plasmid generation and mouse transgenesis

The fragments studied were retrieved from *Whole-genome shotgun contigs (wgc)* deposits on Genbank by their similarity with mouse sequences gently given by Prof. Igor Schneider, using BLASTn tool: *CNS65* from *Phyton mulurus bivittatus* (1,168 bp; AEQU02037710.1) and *Ophiophagus hannah* (1,178 bp; AZIM01000055.1); *Island I* from *P. bivittatus* (1,060 bp; AEQU02066876.1). The sequences were synthesized and inserted in pUC57 vector by FASTBIO (Ribeirão Preto-BR). The primers pUC77-F, 5'-GTAAAACGACGGCCAGTG-3'; pUC77-F, 5'-GGAAACAGCTATGACCATG-3', were used to amplify those sequences from snakes by PCR, using conditions of 48-50°C annealing temperature, 1mM MgCl₂, and extension time varied from 1 minute for *CNS65* and 1.5 min for *Island I*. When amplification was successful, PCR were repeated using doubled volume (50 µl), which were run on 2% agarose gel electrophoresis and purified using *QIAquick Gel Extraction Kit (Qiagen)*. Products were ligated to PCR8-Topo TA (Invitrogen) and cloned into *Top-10* or *DH5α* thermo-competent bacteria according to manufacturers' instructions. Plasmids were extracted using *PureYield™ Plasmid Miniprep System (Promega)* and sequenced to confirm insert. PCR8-Topo TA was used because it can be easily ligated to DNA fragments by TOPOR cloning system and contains attL sites from lambda bacteriophage, specific for rapid and highly efficient recombination with any vector containing attR sites. PCR8-Topo TA plasmid DNA carrying sequences of *CNS65* from *P. bivittatus* and *O. hannah*, and of *Island I* from *P. bivittatus*, were cleaved using the enzyme EcoRI as a second confirmation of inserts. *CNS65* clones 1 and 4 from *P. bivittatus* and clone 3 from *O. hannah* (Fig.2.2), were picked for LR recombination reaction. All of the *Island I* clones from *P. bivittatus* contained insert (Fig. 2.3).

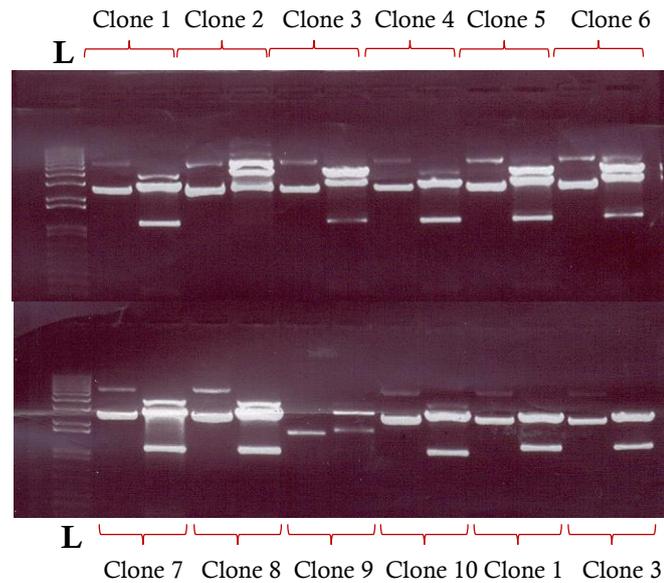


Figure 2.2. Agaroses gel for diagnosing clones containing insert of expected size cloned into vector PCR8. Clones 1 to 6 (top) and 7 to 10 (bottom) indicated the presence of *P. bivittatus* (Pmb) *CNS65* insert. Clones 1 to 3 (bottom) correspond to *O. hannah* (Oh) *CNS65*. First well in each clone corresponds to intact plasmid DNA, and second well corresponds to plasmid cleaved with *EcoRI* enzyme. Clones 1 and 4 of Pmb *CNS65* and clone 3 of Oh *CNS65* were picked for LR reactions. L=ladder.

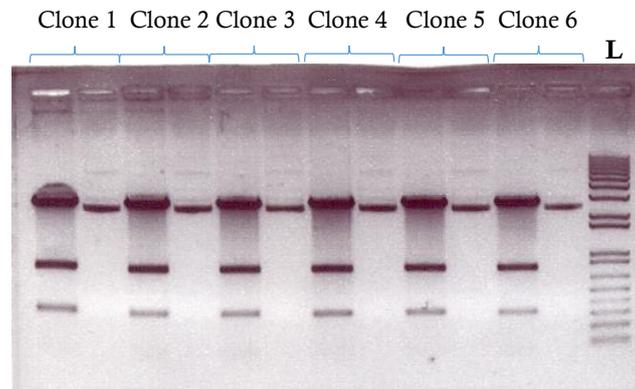


Figure 2.3. Agaroses gel for diagnosing presence of *P. bivittatus* *Island I* insert into PCR8. Each clone cleaved with *EcoRI* enzyme, and their intact DNA, respectively, were run in adjacent wells for each clone (1 to 6). All clones apparently contained expected insert. L=ladder.

LR recombination reaction from Gateway System (Invitrogen) was subsequently performed in order to transfer inserts into the Gateway-Hsp68-LacZ vector, according to

manufacturer's protocol. 50-150 ng (1-7 μ l) of PCR8 plasmid DNA containing respective inserts were added to 1 μ l of Hsp68-LacZ vector, topping with TE buffer, pH 8.0, up to 8 μ l. Every sample mix prepared was homogenized, briefly centrifuged, and added to 2 μ l LR Clonase™ II enzymatic mix previously left resting on ice for 2 minutes and vortexed twice for 2 seconds each. Final mixtures were again homogenized, vortexed twice and quickly centrifuged. Reactions were incubated at 25°C for 3-4 hours. Subsequently, 1 μ l of proteinase K solution was added to each sample to stop enzymatic reaction, following up with vortexing and incubating at 37°C for 10 minutes. Products of LR recombination reaction were cloned in thermo-competent *Top-10* bacteria, and 6-8 colonies were inoculated in liquid LB overnight at 37°C. Plasmids were extracted using QIAprep Spin Miniprep Kit (Qiagen) and subjected to enzymatic cleavages and sequencing to confirm inserts. Band patterns expected after enzymatic cleavages were predicted using the software *NEBcutter V2.0* (www.neb.com/external-links/nebcutter), which performs *in silico* cleavage of sequences, in case they present binding sites for the chosen enzymes. For instance, the *Island I* fragment studied has a binding site for EcoRI enzyme (Fig. 2.4).

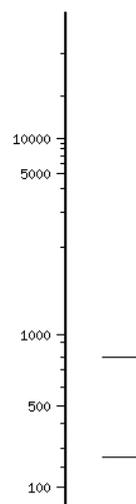


Figure 2.4. Pattern of bands expected after cleaving *Island I* from *P. biviwtatus* with EcoRI enzyme, estimated by NEBcutter V2.0 software.

Besides sequences themselves, HSP68-lacZ vector also comprises binding sites for EcoRI and EcoRV enzymes. The patterns revealed in the agarose gel were compared with the patterns estimated for each sequence. For example, combining information obtained from enzymatic digestions with both enzymes (EcoRV, Figs 2.5 and 2.7; EcoRI, Figs 2.6 and 2.8), clones 1, 2, 3, 5, 6, 7, 8 of *P. bivittatus* CNS65, and clones 1, 2, 4 and 6 of *O. hahhah* CNS65 apparently contained respective inserts of interest. Estimation of HSP68-lacZ clones containing *P. bivittatus* Island I as inserts were confirmed using the same method. Finally, the selected clones were sequenced for ensuring that they were definitively suitable for transgenic assays. Most of the aforementioned experiments were accomplished in Prof. Igor Schneider's laboratory at *Universidade Federal do Pará*.

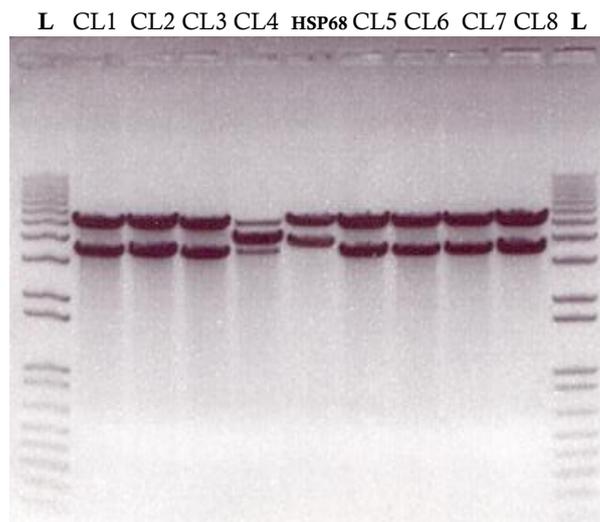


Figure 2.5. Cleavage with EcoRV for diagnosing *P. bivittatus* CNS65 enhancer cloned into HSP68-lacZ vector. Apparently, clone 4 was the only one lacking insert. L=ladder.

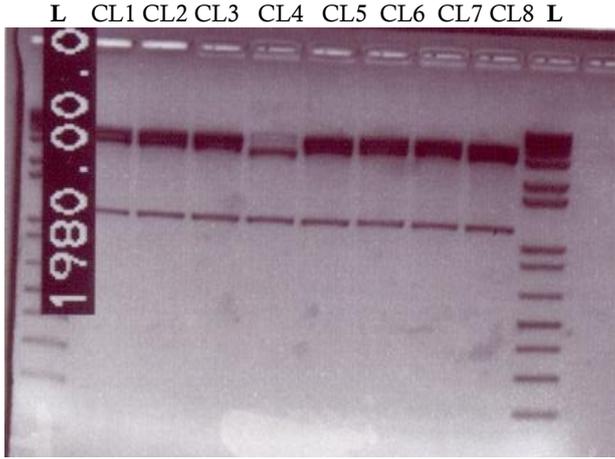


Figure 2.6. Cleavage using EcoRI for diagnosing *P. bivittatus* CNS65 insert cloned into HSP68-lacZ vector. This cleavage confirms information obtained by cleavage with EcoRV, showing that clone 4 lacks insert. L=ladder.

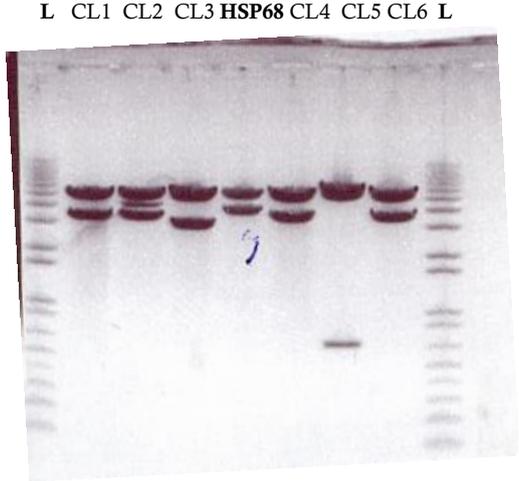


Figure 2.7. Cleavage using EcoRV for diagnosing *O. hannah* CNS65 insert cloned into HSP68-lacZ vector. Apparently, clones 1, 2, 4 and 6 contained expected inserts. L=ladder.

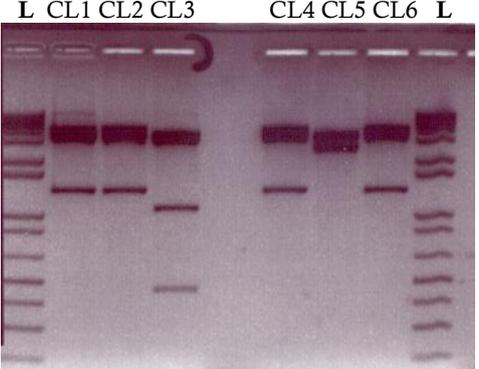


Figure 2.8. Cleavage with EcoRI enzyme for diagnosing *O. hannah* CNS65 insert cloned into HSP68-lacZ vector. Apparently, clones 1, 2, 4 and 6 contained expected inserts. L=ladder.

The vector backbone was excised by digesting 50 µg of each plasmid with SallI enzyme, purified using *QIAquick Gel Extraction Kit* (Qiagen), and the DNA to be injected was further purified using a standard ethanol precipitation and diluted to 2 ng/µL, which was performed at the University of Chicago, by Prof. Igor Schneider's group. Pronuclear injections of CD1 mouse embryos, followed by embryos harvesting, staining and fixing, were performed by *Cyagen* (Chicago-EUA). Each construct resulted in at least 3 LacZ-positive embryos displaying similar reporter activity.

Estimation of Transcription Factor Binding Sites

Orthologous sequences of *CNS65* and *Island I* from nine limbed amniotes were retrieved from Genbank for comparisons with snake sequences (accession numbers are provided in Table A2.1): the lizard *Anolis carolinensis*, the alligator *Alligator mississippiensis*, the avian *Gallus gallus* and *Columba livia*, the turtles *Pelodiscus scinensis* and *Chrysemis picta*, and the mammals *Mus musculus*, *Rattus norvegicus* e *Homo sapiens*. Alignments were produced using ClustalW algorithm (Thompson et al. 1994) implemented in the software BioEdit sequence alignment editor, for adjusting 3'-5' orientation and assuring equivalence of fragments compared. Transcription Factor Binding Sites (TFBS) were predicted individually for each sequence of the alignment and also for different groups of species, using MultiTF/Mulan online software (Ovcharenko et al. 2005). I performed multiple TFBS comparisons by contrasting different groups of species, defined as follows: 1) snakes (*P. bivittatus* and *O. hannah*); 2) squamates (snakes and the lizard *Anolis carolinensis*); 3) snakes and the alligator *Alligator mississippiensis*; 4) 'reptiles' (snakes, lizard, alligator, the turtles *Pelodiscus scinensis* and *Chysemys picta*, the birds *Gallus gallus* and *Columba livia*); 5) 'reptiles' and mammals (all species listed in 4 and *Homo sapiens*, *Mus musculus* and *Rattus norvegicus*); 6) 'reptiles' excluding the snakes; 7) mammals and 'reptiles' excluding snakes.

Results obtained using different minimum detection thresholds were compared and we settled the value for complete matrix profiles as 0.85, which allowed prediction of a considerable amount of TFBS while avoiding noise in the analyses resulting from excessive permissiveness. I also performed an 'optimized for function' prediction in order to narrow down comparisons and access information on a smaller group of likely more relevant candidate TFBS that were lost or acquired in Serpentes. MultiTF online software does not process pre-aligned sequences nor sequences containing gaps, performing instead inside alignments prior to TFBS prediction. Therefore, it was not possible to have the same interval positions for equivalent TFBS in all groups compared in every comparison. In groups including snakes, I used *P. bivittatus* TFBS positions as a reference for data description. For groups excluding snakes, the reference used was *A. mississippiensis*. The TFBS predicted were plotted in Microsoft Excel, where I searched for TFBS conserved in sequences from predefined sets of species that were not predicted in equivalent positions in snake sequences. I also annotated TFBS predicted in snake sequences that were not identified in the regulatory elements from the limbed groups. This approach enabled identification of TFBS gains and losses during evolution of *CNS65* and *Island I* in Serpentes, although our discussion focuses on TFBS losses possibly related to limb loss.

CHAPTER III

Evolution of limb development in marsupial mammals: noncoding RNAs
associated with the regulation of *HOX* gene expression

3.1 ABSTRACT

Limb development in the tammar wallaby *Macropus eugenii* is characterized by accelerated forelimb development before birth, which allows the newborn to reach the pouch, while hindlimbs initially remain at a very early stage of differentiation and are effectively nonfunctional. Additionally, kangaroos' feet lack digit I, develop syndactyl digits II and III and an elongated digit IV. *HOXD* genes are essential for limb patterning during tetrapod development, and its complex regulatory system might contribute to the origin of phenotypic novelties. Recent literature recognizes noncoding RNAs (ncRNAs) as a crucial parcel of this regulatory system, and investigation of their conservation and function can provide insights into mechanisms of tetrapod limb evolution. Here, I have identified and compared ncRNAs transcribed in tammar limbs at days 23 (d23) and d25 (d25) of pregnancy, using RNA-seq analyses, and real-time quantitative PCR (qPCR) to validate comparisons; performed conservation analyses using orthologue genomic sequences from opossum, Tasmanian devil, mouse and human; and characterized their spatiotemporal expression patterns in the developing autopodia by whole mount *in situ* hybridisation (WISH). I have also, for the first time, shown the expression patterns of *HOXD10* at d25, and *HOXD11-D12* at d23 and d25 in tammar wallaby autopodia, using WISH, and compared Hox coding and ncRNA expression patterns. I identified five lncRNAs (long noncoding RNAs) located in intergenic and intronic positions among terminal *HOXD* coding genes that are conserved among mammals, and three that are located downstream to *HOXD9* that are exclusive to marsupials. Transcriptional profiles in the limbs of the lncRNAs *XLOC46-49* and the putative pre-miRNA *XLOC53* partially resemble those of *HOXD10-13* genes. The lncRNA *XLOC46* shows spatial expression patterns in the autopodia that resemble those of mouse and tammar *HOXD11-12* genes, whereas *XLOC52* and *XLOC53* are expressed in patterns identical to those in the tammar, except for low *XLOC53* expression in the hind autopodium at d25. These findings suggest that the lncRNAs located intergenically and intronically to the terminal *HOXD* genes (*XLOC46-49*) may regulate expression of these genes during autopodium development in mammals. *XLOC52* and *XLOC53* likely regulate *HOXD* expression in marsupials, and may have contributed to the evolution of specific morphological traits in this lineage. Finally, *XLOC53* may represent a stage-specific pre-miRNA, contributing to the rapid development of the forelimbs compared to that of the hindlimbs in kangaroos. Knowledge on lncRNAs evolution and functions may reveal elements that contributed to morphological evolution in marsupials.

3.2 INTRODUCTION

Macropodid marsupials, as the tammar wallaby *Macropus eugenii* (Fig. 3.1), exhibit highly modified hindlimbs in comparison with other marsupials and with eutherians, and heterochrony between development of fore- and hindlimbs (Chew et al. 2012; Chew et al. 2014). When the tammar mother delivers her altricial young, it possesses relatively well-developed forelimbs, used to climb independently, using swimming motions, from the birth canal opening to one of the four teats in the pouch (Shaw & Renfree 2006; Schneider et al. 2009). At this stage, hindlimbs are still fetal and cannot help the young during its journey to the pouch, but after birth the hindlimbs develop very fast and become bigger than the forelimbs, allowing the animal to hop, an efficient way of locomotion for wallabies and kangaroos (Sears 2009; Keyte & Smith 2010; Chew et al. 2014). This specialized appendage is highly modified in comparison with the typical vertebrate autopodium that very often have five separate digits (Weisbecker & Nilsson 2008; Chew et al. 2012). In the tammar foot, digit I is absent even in the fetus, digits II and III are syndactyl (reduced, fused and bound by skin), and digit IV is elongated (Fig. 3.2; Chew et al., 2012; Jones, 1925; Weisbecker and Nilsson, 2008).



Figure 3.1. Tammar wallaby *Macropus eugenii* at the marsupial breeding colony from the University of Melbourne. Image credit: Prof. Geoff Shaw.

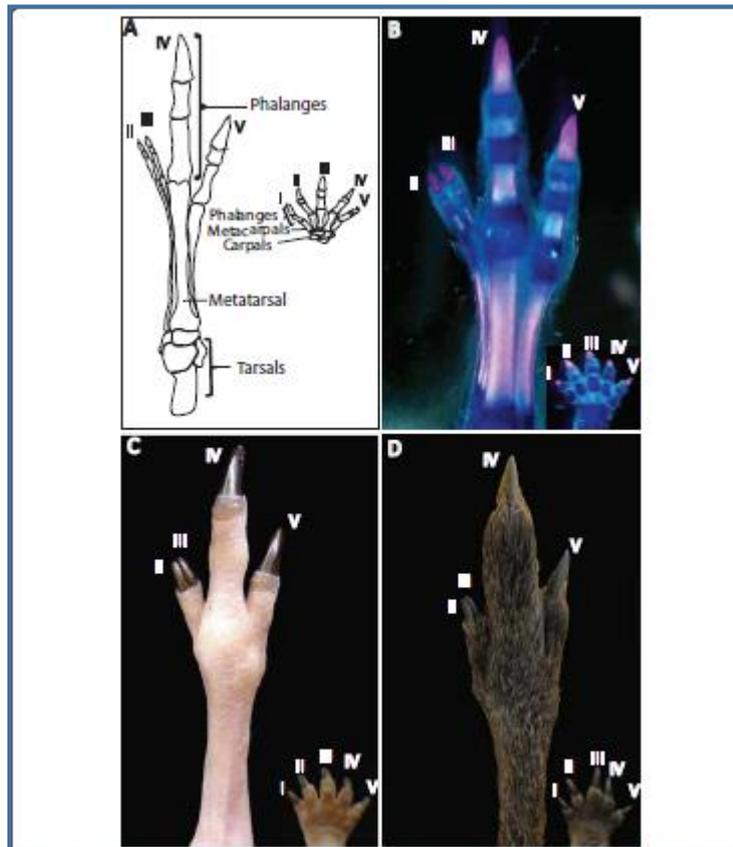


Figure 3.2. Morphology of the tammar wallaby autopodium. (A) Diagram of the macropodid fore- and hind autopodial bones. (B) Fore autopodium at day 150 post-partum (pp) and hind autopodium stained for bone and cartilage with alcian blue and counter-stained with alizarin red. (C) External morphology of fore- and hind autopodia at day 120 pp. (D) Adult tammar fore- and hind autopodia. In the fore autopodium, all five digits are present, whereas hind autopodium lacks digit I. Digits are numbered with Roman numerals. (Images copied from Chew et al., 2012).

The syndactyl morphology is a classic example of homoplasy, as it is identified in the foot of all species from the marsupial Orders Diprotodontia (koalas, wombats, possums, kangaroos, and allies) and Peramelemorpha (bilbies and bandicoots); such phenotype is characterized by ossification similarity and strong integration between the syndactyl pedal digits (Fig. 3.3; Weisbecker & Nilsson 2008). Because of the unlikelihood of a complex trait such as syndactyly evolving multiple times, this character was first used to classify marsupials in a "former" clade, the Syndactyla (Hall 1987; Lockett 1994; Marshall 1973; Szalay 1994), but such theory was refuted by a new classification based on dentition characteristics and

molecular systematics (Amrine-Madsen et al. 2003; Kirsch 1977; Nilsson et al. 2004; Phillips et al. 2006). Nowadays, Diprotodontia and Peramelemorpha are widely accepted as non-sister groups [(Kirsch 1977; Nilsson et al. 2004; Phillips et al. 2006), but see (Szalay 1994; Baker et al. 2004)]. The marsupial syndactyly common to both orders is considered non-adaptative because it apparently does not favor locomotion, so the morphological similarities identified argue for a parallelism settled on equivalent developmental pathways (Weisbecker & Nilsson 2008). A malformation, which was not deleterious because the marsupial syndactyl digits remain functional (differently from human syndactyly), apparently explains the origin of this morphology. It seems to have been fixed when the ancestral population went through a bottleneck event (Hall 1987), culminating in an ontogenetic constraint (Weisbecker & Nilsson 2008). The processes regulating morphogenesis of such differentiated appendages and the developmental heterochrony between the macropodid fore- and hindlimb remain however obscure (Chew et al. 2012; Deakin 2012).

Investigation of developmental processes underlying evolution of morphological diversity in marsupial limbs recalls *Hox* genes as main target candidates (Zákány & Duboule 1999; Zakany & Duboule 2007). Among these, *HoxD* genes are particularly relevant because they are intimately connected with limb development in all tetrapods, and their specific expression patterns are frequently associated with the origin of new limb morphologies (Schneider & Shubin 2013). *HoxD* genes are pleiotropic, as their expression is also essential to pattern several other body structures during development, such as the anterior-posterior body axis (Deschamps & van Nes 2005; Mallo et al. 2010). The pleiotropic nature of *HoxD* genes can be illustrated by *HOXD13* mutation in human causing limb and genital malformations (Bruneau et al. 2001; Amiel et al. 2004; Sivanantharajah & Percival-Smith 2015). Phenotypic evolution involving changes in gene expression domains, particularly the pleiotropic ones, is often associated with the evolution of regulatory noncoding sequences (Spitz & Duboule 2008;

Spitz 2010b; Montavon et al. 2011). Despite the prominent morphological and developmental peculiarities of the tammar limbs and the fact that *HOXD13* gene expression patterns in the tammar limb is detected at earlier stages of embryonic development in the forelimb than in the hindlimb (which differs from its mouse orthologue), the *HoxD13* coding sequence is highly conserved between mouse and tammar (Chew et al. 2012). In scenarios where coding regions are too conserved, morphological differences may be explained by variation in the way that genes are expressed, which is modulated by regulatory elements. One important parcel of *HOX* gene expression regulatory elements is comprised by noncoding RNAs (ncRNAs) (Petruk et al. 2006; Kumar et al. 2015), including the micro- (miRNAs) (Yekta et al. 2008) and long noncoding RNAs (lncRNAs) (Rinn et al. 2007; Chew et al. 2012). These elements provide great specificity in gene expression during digit development (Montavon et al. 2011), which have been thoroughly investigated in traditional model systems.

Noncoding RNAs (ncRNAs) comprise the vast majority of genome transcripts (Ponting et al. 2009; Djebali et al. 2012), and are likely involved in evolution of specific phenotypes. Transcription of hundreds of ncRNAs from Hox loci have been observed in fly, mouse and human (Bae et al. 2002; Noordermeer & Duboule 2013; Dasen 2013). Amongst those are the miRNAs, comprising a class of short noncoding RNA molecules that regulate expression of their mRNA targets (Mansfield et al. 2004), some of which have been shown to regulate *HoxD* genes during limb development. In mice, limb bud development relies on three miRNAs, *miR-let-7c* and *miR-let-7e* and *miR-10b*, all involved in cell fate determination (Mansfield et al. 2004). In the tammar wallaby, both conserved (*mir-10a*, *mir-10b*, *mir-196a*, *mir-196a2*, and *mir-196b*) and one novel miRNA located in Hox clusters have been identified (Yu et al. 2012).

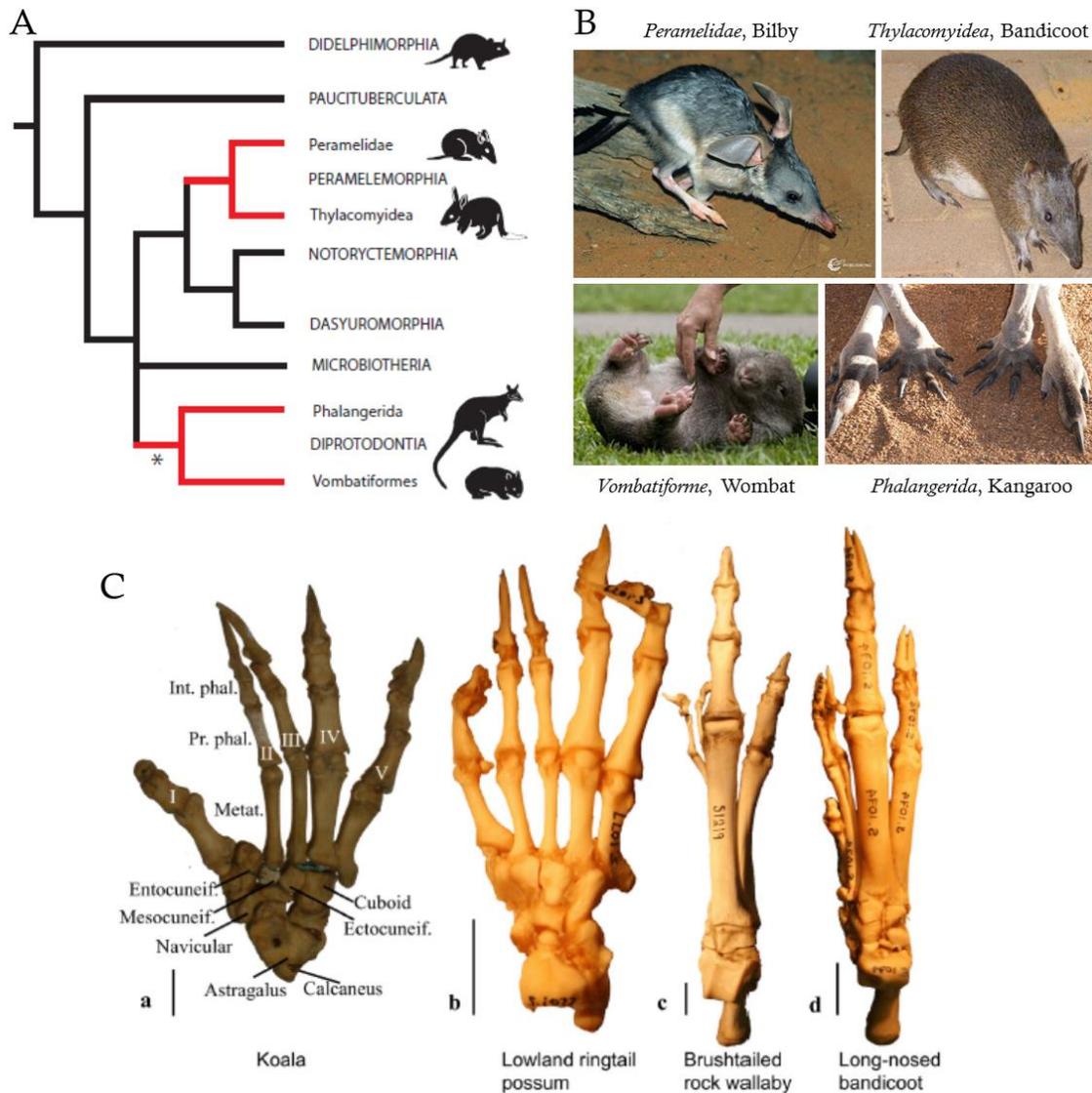


Figure 3.3. Syndactyl feet occur in both Peramelemorpha (bilbies and bandicoots, top B) and Diprotodontia (wombats and kangaroos, bottom B), which are not sister groups according to the phylogenies proposed by Weisbecker and Nilsson (2008) and Amrine-Madsen et al., (2003) (A). C) Right pedal skeleton of four syndactylous marsupials, showing tarsal bones and separated digits of a koala (a), a lowland ringtail possum *Pseudochirulus canescens* (b), a brush-tail rock wallaby (c), and a long-nosed bandicoot *Perameles nasuta* (d). Figures a-c are diprotodontian marsupials, while d is a peramelemorph. Int. phal., Intermediate phalanx; Metat., metatarsal; Prox. Phal, proximal phalanx. Roman numerals correspond to digit numbers. Vertical scale bar = 1 cm. Adapted from Figs. 1 and 3 from Weisbecker and Nilsson (2008).

Long non-coding RNAs, on the other hand, constitute an extremely variable class of transcripts that are longer than 200 base pairs (Saxena & Carninci 2011), which participate in

target gene silencing or activation, working as key *cis*- or *trans*- acting regulators of transcriptional and translational output (Ponting et al. 2009; Saxena & Carninci 2011; Dinger et al. 2008). These transcripts are fascinating for engaging several different genetic and epigenetic mechanisms involved in regulation of gene expression in specific cellular contexts and stages of differentiation (Rinn et al. 2007; Ponting et al. 2009; Blignaut et al. 2012; Dasen 2013), reason for these elements to be defined as ‘fine-tuners’ of cell fate (Fatica & Bozzoni 2014). The Hox lncRNAs *HOXA11AS*, *HOTAIRM1*, and *HOTAIR* regulate site-specific *Hox* gene expression in eutherians that have also been identified in the tammar wallaby (Yu et al. 2012). Among these, *HOTAIR* is known to play a role in the regulation of *HoxD* genes during limb development, although it originates from the HoxC cluster located on another chromosome, representing the first *trans*-acting lncRNAs ever described (Rinn et al. 2007; Saxena & Carninci 2011). In the tammar, it is expressed at the early head-fold stage of the embryo before limb buds develop (Chew et al. 2012; Yu et al. 2012). In human fibroblasts, *HOTAIR* knockdown decreases activity of *HOXD* genes repressor complexes, increasing their expression across 40 kilobases (Rinn et al. 2007); in mouse, it causes homeotic transformation of spinal vertebrae and limb defects (Dasen 2013; Fatica & Bozzoni 2014). Another lncRNA that influences limb development in the mouse, is “*HOXA* Transcript at the distal TIP” (*HOTTIP*), which is transcribed from 5’ end of the human *HOXA locus* upstream of *HOXA13* and coordinates the activation of several *HOXA* genes in vivo. *HOTTIP* knockdown in mouse limb buds decreases the expression of *HOXA13*, *HOXA11* and *HOXA10* and affects limb morphology in chick and mouse embryos, resembling some of the defects observed in mice lacking *HOXA11* and *HOXA13* (Wang et al. 2011; Fromental-Ramain et al. 1996; Small & Potter 1993; Davis et al. 1995; Fatica & Bozzoni 2014). Specific functional roles of these conserved lncRNAs were not yet investigated in marsupials.

NcRNAs are recently-discovered regulators, important for addressing questions about the evolution of lineage-specific transcriptional regulation (Yu et al. 2012). Investigating the lncRNAs of tammar wallaby HoxD cluster that are transcribed in developing limbs is particularly interesting due to its unique characteristics such as their digital formula, locomotion by hopping, and limb developmental heterochrony (Lindsay et al. 2012). The divergence time between eutherian and marsupials, estimated for about 160 Mya (Luo 2007), is a sufficiently long period for allowing regulatory complexes to evolve concomitantly with new traits, which turns mammals an interesting target for investigating ncRNA evolution (Yu et al. 2012). In this context, I have investigated conservation and spatiotemporal expression of *HOXD* ncRNAs transcribed in *M. eugenii* developing autopodia, in order to identify regulatory elements that may have contributed to the evolution of marsupial developmental characteristics and unique digit phenotypes. I predict that some of those ncRNAs will be conserved amongst mammals, whereas other ones might have evolved exclusively in marsupials' genomes, representing important candidates for modulating morphological evolution. I also hypothesize that some of the ncRNAs exhibit transcription profiles and expression patterns in the fore- and hind autopodia at distinct developmental stages that resemble those of *HOXD* coding genes, indicating possible regulatory roles of these ncRNAs during autopodial development.

3.3 METHODS

Tissue assessment: collecting tammar wallaby and programming gestation

All sampling techniques and tissue collection procedures conformed to the Australian National Health and Medical Research Council guidelines (2013) and were approved by the University of Melbourne Animal Experimentation and Ethics Committees. The tammar wallaby *Macropus eugenii* were regularly collected at the Kangaroo Island, SA, and maintained in the marsupial colony of Prof. Marilyn Renfree's group in Melbourne, VIC. The expeditions were composed by four people who travelled to the Kangaroo Island and spent 7 days performing nocturnal collection of tammar individuals (approximately one hundred females and a couple of males). Researchers chased animals belonging to an overpopulation that inhabits local farm's paddocks. Individuals collected over the night were screened the next day for reproductive stage based on pouch examination, and classified as: 1) pouch empty, dirty and no plug (absence of a young tammar in the pouch, which was dirty because the female was not close to give birth, no lactating mammary glands, and no evidence of sperm in the genital, which would indicate recent copulation); 2) pouch empty, dirty and plug; 3) pouch empty and clean (indicating proximity of birth); 4) pouch young (presence of a newborn fetus or furry young in the pouch; see Fig. 3.4); 5) pouch empty with lactating gland (indicating that the female was breast-feeding). Animals were also checked for injuries, treated and ear-tagged, receiving a number that identifies each individual for future screening and experimental use. Collected animals were transported to the marsupial colony and distributed in several yards where they fed on green grass and received lucerne cubes (Fig. 3.5) as a supplement three times per week. In the colony, they were regularly screened for reproductive stage and rotated between yards during reproduction management (see Fig. 3.6 for collecting method).



Figure 3.4. Examining tammar wallaby pouches. The females are carrying pouch youngs, at an early stage on the left, and a late stage on the right.



Figure 3.5. Tammar wallaby feeding on lucern cubes at the marsupial colony. Image credit: Prof. Geoff Shaw.



Figure 3.6. Students catching a wallaby specimen with the help of a sack and a net, at the marsupial colony.

Most embryos used in this project were collected in January of 2014 and 2015 at days 23 (d23), 24 (d24) and 25 (d25) of pregnancy, as the tammar wallaby reproduction is seasonal, occurring from January to June. The tammar wallaby mate at a post-partum oestrus, but the blastocyst entries diapause controlled by both lactation and photoperiod (Renfree & Tyndale-Biscoe 1973; Browner et al. 2004), synchronizing births to occur around end of January and the first half of February (Andrewartha & Barker 1969; Berger 1966; Renfree & Tyndale-Biscoe 1973; Sadleir & Tyndale-Biscoe 1977). These reproductive characteristics allow programming females to give birth by removing the young from the mother's pouch and injecting the female with bromocriptine, a dopamine agonist that suppresses the peripartum pulse of plasma prolactin (Fletcher et al. 1990), inducing immediate blastocyst reactivation and allowing control of embryonic developmental stage. Five females were programmed and sacrificed on the d23 of pregnancy for obtaining additional embryos. Embryos' whole bodies or individualized limbs destined to RNA extraction were immediately frozen using liquid nitrogen and maintained at -80°C . The tissues used for whole-mount *in situ* hybridisation were dissected and immediately fixed at 4% paraformaldehyde (PFA)/PBS overnight at 4°C , then washed next day with PBS and 25, 50, 75 and 100% methanol (MeOH)/PBS for 20 min each step on ice, then stored at -20°C .

Using RNA-seq data to access transcripts expressed in tammar wallaby developing limbs

Total RNA extracted from forelimbs and hindlimbs from d23 and d25 (N=3, pooled together) was used to perform RNA-seq, in collaboration with Professor Asao Fujiyama and his team at National Institute of Genetics (Japan). The RNA-seq was analyzed through VLSCI Supercomputer Facility (<https://www.vlsci.org.au>) implemented to allow data processing. The protocol described in Trapnell et al. (2012) was used to identify transcripts of interest and to compare their expression levels between tissues and embryonic stages.

Transcription reads were mapped to *Macropus eugenii* genome (Renfree et al. 2011) using TopHat (Trapnell et al. 2009), a tool that also infers read spans when segments align to the genome far apart from one another, and estimates location of junction's splice sites, working independently from gene or splice site annotations (Trapnell et al. 2012). The resulting alignment files were inserted on to Cufflinks package (Trapnell et al. 2010) in order to generate a transcriptome assembly for each specific tissue and developmental stage. Using Cufflinks, the splicing structure of each gene was inferred, the expression level of each transfrag in the sample was quantified, and background or artefactual transfrags were filtered out. Subsequently, Cuffmerge was implemented to merge these assemblies together, generating a uniform database for calculating gene and transcript expression in each condition. Once each sample was assembled and all samples were merged, genes and transcripts that are differentially expressed or regulated between samples were identified from the final assembly. Both reads and merged assembly were inserted in Cuffdiff to calculate transcriptional levels and to test for statistical significance of observed changes. Text files reporting results were displayed in the software CummeRbund (<http://compbio.mit.edu/cummeRbund/>), which was used to create expression plots by converting Cufflinks output files into statistical objects that were then analyzed in R package version 2.2 and accessed through the Bioconductor website (<http://www.bioconductor.org/>).

Reads aligned to HoxA and HoxD clusters, in intergenic and intronic positions, were selected as target transcripts for investigation. As transcripts identified outside annotated genes can correspond either to new protein-coding genes or to noncoding RNA, I predicted coding potential of the sequences using online software ORF Finder (Rombel et al. 2002) and Coding Potential Assessment Tool (Wang et al., 2013) prior to investigating transcripts expression profiles and associated roles.

Investigating conservation of HoxD cluster transcripts among mammals

The transcripts from HoxD cluster identified by RNA-seq analyses were mapped into genomes of the marsupials opossum (*Monodelphis domestica*) and Tasmanian devil (*Sarcophilus harrisii*), and the eutherians mouse (*Mus musculus*) and human (*Homo sapiens*) using the BLAT tool on Ensemble Genome Browser database. Matches were individually evaluated, and those located in the HoxD cluster of the referred species were used to extract conserved regions. The genomic sequences corresponding to transcriptional fragments were used in the online software mVista to perform custom alignments. Annotation information, where present, was also valuable to infer known functions of these transcripts in other organisms.

Confirming ncRNAs transcription by reverse transcriptase polymerase chain reaction, cloning and sequencing

The construction of transcriptome assemblies allow for inferences of novel genes and transcripts, which must be validated by traditional cloning and PCR-based techniques (Trapnell et al. 2012). In order to validate transcription of *HOXA* and *HOXD* intergenic noncoding RNAs identified in the transcriptome, I first performed reverse transcriptase polymerase chain reaction (PCR). Total RNA was extracted from fore and hind autopodia dissected at d23 and d25 that have been instantly frozen, using the GenElute™ Mammalian Total RNA Miniprep Kit (Sigma-Aldrich, AU) and following the manufacturer's protocol. To assure that there was no genomic DNA in the samples, RNA was decontaminated using DNA-free™ kit (Ambion), following the manufacturer's protocol. DNA-decontaminated RNA was used as a template for a PCR using tammar *HOX* intronic primers (kind gift of Dr. Hongshi Yu) in order to verify that no amplification occurred, which would indicate genomic contamination. In case there was a smear band, decontamination reaction and PCR were repeated. RNA was finally stored at -80 °C.

Subsequently, cDNA was synthesized from RNA from fore- and hindlimb of d23 and d25 using SuperScript™ III First Strand Synthesis System for RT-PCR (Invitrogen), according to the manufacturer's protocol. The cDNA synthesis product was stored at -20°C and then used as template for PCR. For confirming transcription of intergenic noncoding RNAs identified by the transcriptome in the HoxA and HoxD clusters, primers were designed (Tables 3.1 and 3.2) using Primer 3 online software (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/>, Untergasser et al. (2007)).

Table 3.1. Primers used for amplifying transcripts mapped to the tammar wallaby HoxD cluster; sequences are presented together with melting temperatures used in the PCRs.

Transcript name/loci	Forward (F) and Reverse (R) Primers (3'-5')	PCR melting temperature
XLOC_033946 (<i>XLOC46</i>)	F:CCCTGTCACAAACCCTCACT R:ATTCGACAAGAGCGAGCAAT	58°C
XLOC_033947 (<i>XLOC47</i>)	F:AGCTCCCTGAATTATCCTTCAAC R:AGGAGTTGCCCTTAAACAAGAAC	60°C
XLOC_033948 (<i>XLOC48</i>)	F:CTTTGTTTGTGACTTGGGAAAAG R:ATGAGAAACTGTTGTGGCAACTT	60°C
XLOC_033949 (<i>XLOC49</i>)	F:TGGTTAGTAGCGAGGGGAAA R:CCCTGTGTGGTTTCATAGAGAA	55°C
XLOC_033950 (<i>XLOC50</i>)	F:CCGGTGCGGTAAATGTAAAC R:GCGGGTTCACAAAGTAGGAG	58°C
XLOC_033951 (<i>XLOC51</i>)	F:TTTTCCCTTCCTTGCCTTTT R:ATGAGGGAAACACCTATATTGC	52°C
XLOC_033952 (<i>XLOC52</i>)	F:TGTTAAAGTCCTTGCAAATTATGG R:AGCGGCTGTTACACAAGCAT	55°C
XLOC_033953 (<i>XLOC53</i>)	F:CAGGTGCTAGGTGGCATCTTA R:GGGTTGAAGGAGGAGTGGTC	60°C
XLOC_033954 (<i>XLOC54</i>)	F:CTACCCTCATCCCCACTTT R:AGGGCATTCTCCCACTG	60°C
XLOC_033955 (<i>XLOC55</i>)	F:TTTGTTCCTTGGGGTGA R:GGGGAAGCTCTGTAGCTG	52°C

Table 3.2. List of primers used for amplifying transcripts mapped to the tammar wallaby HoxA cluster; sequences are presented together with melting temperatures used in the PCRs.

Transcript name/loci	Forward (F) and Reverse (R) Primers (3'-5')	PCR melting temperature
XLOC_053797 (XLOC97)-isoform 1	F: CTCTGGTGAATGGGACCTGT R: AGCAAGCCAGACTTCAAAGC	62°C
XLOC_053797 (XLOC97)-isoform 2	F: GACCGCAGTGGTTTTTCTGT R: AGCAAGCCAGACTTCAAAGC	62°C
XLOC_053798 (XLOC98)	F: CTTGCCGAAGTCTCCTTGAC R: CAGCCTTCCTCTCAGAATGG	55°C
(XLOC_053799) (XLOC99)-isoform 1, 2 and 3	F: GTGCTCAGCAGTGAGATGGA R: CCATGTTCTTGCCAGTCCTT	55°C
XLOC_053799 (XLOC99)-isoform 2	F: GGGAACAACCCCTGTTCATA R: CCACTCTCTGCCACTTCTCC	-
XLOC_053799 (XLOC99)-isoform 4	F: GGTGAAGGTGCAGTCAGGAT R: GATGCCTGTATGTGGCCTCT	-
XLOC_053800 (XLOC800)- isoform 1	F: CCAACCCAGCAAACCTTTCAG R: AGGTGCCAACAGAAGGAAAA	55°C
XLOC_053800 (XLOC800)- isoform 2	F: GAGAGTCCCAGCTCTGGTCA R: AGGTGCCAACAGAAGGAAAA	-
XLOC_053800 (XLOC800)- isoform 3	F: GCCTGCTCAGTCACAATGAA R: CACACACCAGGGAAACTCCT	55°C

The efficiency of PCRs (d25 µl) were verified by electrophoresis on 1% agarose gel. After amplification conditions were optimized, reactions were repeated using double volume (50 µl), and ran at 2% agarose gel, where DNA from the bands having suitable sizes were purified using *QIAquick Gel Extraction Kit* (Qiagen), following the manufacturer's protocol. The purified DNA was ligated to pGEM-t easy vector (*Promega*), according to the manufacturer's protocol. Ligation reactions (5 µL) were transformed into thermo-competent bacteria (JM109, *Promega*) through thermic shock at 42°C for 45 seconds. Bacteria were grown for one hour in 800 µL LB broth, centrifuged and plated on solid agar containing x-gal and ampicillin for double selection, incubated overnight at 37°C. Three white colonies of each plate were scrubbed using a sterile pipette tip into a PCR tube reaction containing specific primers of the putative insert; the pipette tips were incubated overnight in LB broth at 37°C. Clones exhibiting band having the expected sizes after PCR were purified using Wizard® Plus SV

Minipreps DNA Purification (Promega), following the manufacturer's protocol. Before purifying, 500 μ L of bacteria in LB broth were added to 500 μ L 50% glycerol and stored at -80°C for stock. Plasmid DNA samples were sent to sequencing for insert confirmation.

Quantifying transcription using real time quantitative polymerase chain reaction (qPCR)

Primers for qPCR were designed following the protocol described in Thornton and Basu (2011), and the oligonucleotides used are provided in Tables 3.3 and 3.4. The cDNAs were designed from RNA extracted of a minimum of 6 individuals for each condition compared (fore and hind autopodia of embryos with 23 and 25 days of development). The RNA extracted as previously described was quantified using Qubit equipment, and the same amount of RNA was used to synthesize every cDNA sample used. Standard curves were performed for every gene using cDNA serial dilutions. Pure DNA was diluted in 1 volume of nuclease free water, corresponding to dilution 1. Subsequent dilutions were 5, 25, 125, 625, 1265 or 10, 100, 1000, 2000, 10000. Optimal dilutions fell within the same range, and all cDNA pools were ultimately diluted 50 times for every differential expression experiment. Reactions were set up according the protocol from SYBR Green PCR Master mix (Thermo Fisher Scientific). Standard curves were managed on Max Pro software, where threshold values were manually adjusted in order to obtain cycle threshold (Ct) values based on primer efficiency (PE) and Pearson Correlation Coefficient (RSq) as closest as possible to 100% and 1, respectively (see individual values on Table 3.3).

Table 3.3. List of primers used to quantify tammar wallaby ncRNA from HoxD cluster by RT-qPCR. TN (transcript name); PC (primer concentration); PCR MT (melting temperature); TV (threshold value); RSq (R squared, Pearson Correlation Coefficient); PE (primer efficiency).

TN	Forward (qF) and Reverse (qR) Primers (3'-5')	PCR MT (°C)	PC(M)	TV	RSq	PE (%)
TBP	qF: GGACAAACTGAAGCAAAGGGACC qR: AGGGCATCATTGGGCTAAAGATAG	58-60	0.5	1000	0.997	96.6
GAPDH	qF: TCCCAATGTATCTGTTGTGGATCTG qR: AACCATACTCATTGTCATACCAAGAAAT	60	0.4	1000	0.999	91
<i>XLOC46</i>	qF: AATCCCACAAATCCCTGACCC qR: ACACAAAGAAAGAAACCCAAACC A	64	1	1362.7	0.995	96.2
<i>XLOC47</i>	qF: ATCGTTTGTGTCTTTCTTGCC qR: GCGTCCTGTCTCTGCC	66	1	6565.8	0.998	100
<i>XLOC48</i>	qF: GACAGAGGAGATTTATTGGGAAG GG qR: CTGGACTTGGATTTCAGGCGG	60	1	1000	0.898	98.7
<i>XLOC49</i>	qF: AAATATCTTCCCGTGTGTGGC qR: AAATTCTCATCTACATTCTTGGGC	55	0.5	546.08	0.997	99.8
<i>XLOC53</i>	*F: CAGGTGCTAGGTGGCATCTTA *R: GGGTTGAAGGAGGAGTGGTC	60	0.4	1000	0.992	97.2
<i>XLOC54</i>	qF: AAAGTGAAGGTTTGATTCTATGCC qR: CTCTCCTACATTCCATAAGCCC	60	0.5	1368.5	0.99	105.5
<i>HOXD10</i>	qF: CTACTTGCTCCTTCACTGCC qR: CGGAAGAAATTAGTTTGGTTCGC	60	0.5	8353.3	0.999	101.9
<i>HOXD11</i>	qF: TGAATAATAGTTTCCGAGACTGGG qR: ACTCAGACTTATGCCAATCCG	60	0.5	3763.4	0.998	100
<i>HOXD12</i>	qF: TACAAGGAGAAGAAGAAGTTAGC G qR: ATTCTAATTGAGGTGGTAGGAGG G	60	0.2	527.01 9	0.995	99.5
<i>HOXD13</i>	qF: ATCTTCCTTCCAGGGGATG qR: ATGGCGTATTCGTTCTCCAG	60	1	1528	0.998	100.7

*Primers designed following qPCR specific protocol did not successfully amplify this gene, so primers designed for cloning were used.

Differential expression experiments were set up according to Bustin et al. (2009), and all samples were compared in the same plate for each gene of interest. The plate model used was exemplified in Fig. 3.7. Real time-qPCR results were adjusted on Max pro software

according to individual threshold values, and analyzed both on REST (Pfaffl et al. 2002) statistical software and manually on Microsoft Excel using a normalization formula.

RT-qPCR plate

Date: _____

Experiment: **Hoxd13 differential expression** _____

	1	2	3	4	5	6	7	8	9	10	11	12
A	Negative Control-Hoxd13	Negative Control-Hoxd13	Negative Control-Hoxd13	Negative Control-Hoxd13	Negative Control-Hoxd13	Negative Control-Hoxd13	Negative Control-TBP	Negative Control-TBP	Negative Control-TBP	Negative Control-GAPDH	Negative Control-GAPDH	Negative Control-GAPDH
B	Calibration-template (1:125 dilution)	Calibration-template (1:125 dilution)	Calibration-template (1:125 dilution)	Calibration TBP	Calibration TBP	Calibration TBP	Calibration GAPDH	Calibration GAPDH	Calibration GAPDH	Empty well	Empty well	Empty well
C	FL-d23 Replicate 1	FL-d23 Replicate 1	FL-d23 Replicate 1	HL-d23 Replicate 1	HL-d23 Replicate 1	HL-d23 Replicate 1	FL-d25 Replicate 1	FL-d25 Replicate 1	FL-d25 Replicate 1	HL-d25 Replicate 1	HL-d25 Replicate 1	HL-d25 Replicate 1
D	FL-d23 Replicate 2	FL-d23 Replicate 2	FL-d23 Replicate 2	HL-d23 Replicate 2	HL-d23 Replicate 2	HL-d23 Replicate 2	FL-d25 Replicate 2	FL-d25 Replicate 2	FL-d25 Replicate 2	HL-d25 Replicate 2	HL-d25 Replicate 2	HL-d25 Replicate 2
E	FL-d23 Replicate 3	FL-d23 Replicate 3	FL-d23 Replicate 3	HL-d23 Replicate 3	HL-d23 Replicate 3	HL-d23 Replicate 3	FL-d25 Replicate 3	FL-d25 Replicate 3	FL-d25 Replicate 3	HL-d25 Replicate 3	HL-d25 Replicate 3	HL-d25 Replicate 3
F	FL-d23 Replicate 4	FL-d23 Replicate 4	FL-d23 Replicate 4	HL-d23 Replicate 4	HL-d23 Replicate 4	HL-d23 Replicate 4	FL-d25 Replicate 4	FL-d25 Replicate 4	FL-d25 Replicate 4	HL-d25 Replicate 4	HL-d25 Replicate 4	HL-d25 Replicate 4
G	FL-d23 Replicate 5	FL-d23 Replicate 5	FL-d23 Replicate 5	HL-d23 Replicate 5	HL-d23 Replicate 5	HL-d23 Replicate 5	FL-d25 Replicate 5	FL-d25 Replicate 5	FL-d25 Replicate 5	HL-d25 Replicate 5	HL-d25 Replicate 5	HL-d25 Replicate 5
H	FL-d23 Replicate 6	FL-d23 Replicate 6	FL-d23 Replicate 6	HL-d23 Replicate 6	HL-d23 Replicate 6	HL-d23 Replicate 6	FL-d25 Replicate 6	FL-d25 Replicate 6	FL-d25 Replicate 6	HL-d25 Replicate 6	HL-d25 Replicate 6	HL-d25 Replicate 6

Figure 3.7. Example of plate set up applied to all genes differential expression qPCR experiments performed to validate RNA-seq data.

Spatial characterization of noncoding RNA expression profiles using whole-mount in situ hybridisation (WISH)

Based on the expression profiles acquired by RNA-seq and qPCR data, I chose the lncRNAs *XLOC46*, *XLOC52* and *XLOC53* as target genes to perform expression spatial characterization using WISH. These genes have been chosen because they had higher transcriptional levels and were adequate representatives of distinct transcriptional profiles (see Results section). The coding genes *HOXD10*, *HOXD11* and *HOXD12* were chosen as reference to interpret lncRNAs expression profiles. I used lncRNAs plasmid DNA produced as previously described to design probes, and used the aforementioned methods to clone *HOX* genes coding sequences (Table 3.4).

Table 3.4. List of primers used to amplify tammar wallaby *HOXD* genes, together with melting temperatures used in the PCR reactions.

Transcript name/loci	Forward (F) and Reverse (R) Primers (3'-5')	PCR melting temperature
<i>HOXD9</i>	F: GCTGTCCCTACACAAAATACCAG R: CTGGCACCTAAAGCAAATAAGA	55°C
<i>HOXD10</i>	F: CCACAATTACACAGGGAATGTTT R: TGAGGGACTTTAGCTGAGTCTTG	58°C
<i>HOXD11</i>	F: TTGTCTGAGTTCCATACCGAGAT R: TAGGGGAAAGTCATTTGACAAGA	55°C
<i>HOXD12</i>	F: GCCTTTGAGTTCCCTTGTTACTT R: ACCTACTGGGAGAGAAGTCATCC	60°C

In order to design RNA probes (riboprobes) necessary for performing WISH, plasmid DNA was first linearized and the orientation of inserts in the vectors was identified by PCRs. The following set of primers was used in four distinct reactions *per* target gene: 1) Gene specific forward primer + M13 forward primer (see pGEM-t easy vector protocol); 2) Gene specific reverse primer + M13 reverse primer (see pGEM-t easy protocol); 3) Gene specific forward primer + M13 reverse primer; 4) Gene specific reverse primer + M13 forward primer. As a result of these four reactions, I obtained one pair of equal bands presenting the approximate insert size, representing the predominant orientation of respective plasmid inserts. After the correct pair of primers was identified, PCRs using 50-100 μ L were repeated in order to obtain a high number of insert copies. Samples were run in 2% agarose gel, and DNA from gel bands was purified as previously described. Purified DNA was used to perform *in vitro* transcription reactions according to DIG RNA Labeling Mix (Roche) according to the manufacturer's protocol.

The technique WISH was performed on limb buds preserved in 100% methanol and stored at -20°C as previously described. On the first day of WISH protocol application, pre-hybridisation was performed starting with rehydrating limbs through MeOH/PBTX series in reverse (75%, 50%, 25% and 0% MeOH) on ice. Limbs were washed three times with PBTX

for 5 minutes each time at room temperature on a rocker; then, they were bleached with 6% hydrogen peroxide in PBTX for 15 minutes at room temperature, washed three times in PBTX for 5 minutes each, and finally treated with 10 μ g/ml Proteinase K in PBTX at room temperature. Time for the Proteinase K treatment ranged between 1 and 2 hours, depending on tissue and developmental stage, representing the most critical step of this protocol.

Limbs were washed twice in PBTX for 5 minutes, re-fixed carefully with 0.2% glutaraldehyde/4% PFA in PBTX for 20 minutes at room temperature on rocker, washed twice in PBTX for 10 minutes, immersed in 500 μ l of pre-hybridisation solution (Table 3.5) and then incubated overnight at 65°C. After this step, embryos can be stored for several days at -20°C before following up with the experiment. In the second day, limbs were hybridised by adding 15-20 μ l of probes *per* sample pre-warmed in 500 μ l pre-hybridisation solution, and were incubate overnight at 65°C. On the third day, pre-hybridisation solution was removed from limbs and stored at -20 °C, because they can be used again. Samples were washed for 5 mins with solutions pre-warmed in heating block to 65°C, as follows: 1) 100% Solution 1 (Table 3.5); 2) 75% Solution 1/25% 2xSSC; 3) 50% Solution 1/50% 2xSSC and 4) 25% Solution 1/75% 2xSSC. Then, they were washed with 2xSSC, 0.1% CHAPS (Table 3.5) and 0.2xSSC, 0.1% CHAPS once for 30 mins each, at 65°C on rocker. Subsequently, they were washed twice with TBTX 1X (Table 3.5) for 10 mins at room temperature rocker and preblocked with 10% sheep serum, 2% BSA in TBTX for 2-3 hours at room temperature rocker. Limbs were then removed, immersed in fresh pre-block solution with DIG antibody (1:1000) and rocked overnight at 4°C. In the fourth day, limbs were washed five times with 1xTBTX containing 0.1% BSA for 30-60 minute. Either in the same or the next day, color reaction was started by washing embryos twice with 1xTBTX for 30 mins each on rocker at room temperature, then washed three times with freshly made NTMT (Table 3.5) for 10 mins each on rocker, at room temperature. They were incubated with 20 μ L NBT/BCIP (Roche)

per 1 ml of NTMT, rocking for 20 min in the dark. Then, incubating samples were kept on the bench, protected from light, and checked periodically until color was developed (usually taking 2-6 hours). All reaction pairs (sense and antisense probes) were stopped at the same time, for allowing valid comparisons. To continue color reaction on the next day, limbs were washed in NTMT for 10mins, then TBTX for 10min and left in TBTX overnight at 4°C in foil. To stop color reaction, they were washed in NTMT for 10mins and in PBTX for at least 20 min (but they were washed for several hours if background was intense), then washed in Ethanol/PBTX series of (25%, 50%, 75%, 100%, 75%, 50%, 25%) for 20-30 min each, and PBTX for 20-30 min. Finally, limbs were fixed in 4%PFA/PBS for 20 min and stored in 50% glycerol at 4°C. Photographs were taken under stereomicroscope coupled to a camera, with stained limb buds positioned in plastic plates filled with 50% glycerol, sometimes over an agarose substrate, under several dark and light backgrounds, in order to adjust images individually.

Table 3.5. Whole-mount *in situ* hybridisation reagents.

Solution	Reagents
5% CHAPS	10g CHAPS; 200ml MQH ₂ O
20X SSC	350.6g NaCl; 176.4g Sodium citrate; 2lMQH ₂ O (pH = 7.0)
10x TBTX	125ml 1M Tris.Cl pH7.5; 75ml 5M NaCl; 25ml 10% Triton X-100; 25ml MQH ₂ O
PBTX	1L PBS; 1ml Triton X-100
Hybridisation Solution	25ml100% Formamide; 12.5 ml 20x SSC; 25 µl 100mg/ml Heparin; 500 µl 10% Triton X-100; 5ml 5% CHAPS; 50 mg Yeast RNA; 500 µl 500mM EDTA; 1g blocking powder; MQH ₂ O to 50 ml
Solution 1	25ml 100% Formamide; 12.5 ml 20XSSC; 500 µl 10% Triton-X; 5 ml CHAPS MQH ₂ O to 50 ml
NMTM	1 ml NaCl; 5 ml Tris pH9.5; 2.5 ml MgCl ₂ ; 50 µl Twen 20; MQHO to 50 ml

Because ncRNAs both sense and antisense probes induced strong staining, I checked transcription orientation to investigate if they were uni or bidirectionally transcribed. cDNA was synthesized from RNA extracted of limbs using the protocol previously described, except that primers specific to each ncRNA studied were used instead of oligo(DT) primers, using either forward or reverse primers in distinct reactions. cDNA from each reaction was used as template for PCRs using primers specifically designed (Table 3.6) to amplify a region internal to the first pair of primers.

Table 3.6. Relation of primers designed for amplifying tammar wallaby transcripts in order to check transcription orientation from cDNA synthesized using unidirectional specific primers (Table 3.1).

Primers name	Forward (F) and Reverse (R) Primers (3'-5')	PCR melting temperature
oriXLOC46	F: GGCACAAAATACCGCATCT R: CTCAGGTCCTCACAGCAACA	55-60°C
oriXLOC52	F: TATTCACGCCGAGTCAGTTG R: TACCAGCCAACACAAGTGGA	55-60°C
qXLOC53	F: GCTGGAGCTATGCAGCG R: AAGGAGGAGTGGTCGGG	60°C

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

The idea of morphological evolution *happening together with* the evolution of regulatory elements of developmental genes is the framework connecting the three chapters of this PhD Thesis. By saying “*happening together with*”, I try to group different scenarios, such as ‘being driven by’, ‘leaving footprints on’ or ‘having its paths opened by’, permeating causal and consequential correlations of evolution both at phenotypic and genetic levels. The regulatory network focused was the HoxD cluster because it constitutes an intriguing and complex system involved in the development of several body structures. This study brought important clues into developmental aspects of scarcely explored evolutionary contexts, which may guide future work into unravelling evolutionary developmental mechanisms involved in morphological divergence among related lineages.

Chapter I explored whether independent evolution of snakelike morphologies in Serpentes and Amphisbaenia registered regulatory signatures that might argue for common developmental mechanisms underlying recurrent origin of such phenotypes among Squamata. One footprint identified in *CsB*, a *Hoxd* centromeric regulatory region, seems remarkable because it was predicted in all of the 14 species analyzed from each snakelike group, encompassing loss of TFBS involved in limb developmental pathways and gain of a TFBS essential for AP axis development in tetrapods. This signature, together with a functional dissection of *CsB*, provides evidence for a possible ancient role of *CR2-CsB* fragment during axial elongation. This study showed how strong and consistent predictions based solely on bioinformatics can be, as long as a series of steps is embraced to ascertain accuracy of predictions derived from robust sampling. This approach is important because validation of the predicted convergent signature would represent a dual groundbreaking insight to the field of Evolutionary Developmental Biology. The confirmation of these prediction would reinforce the still relegated role of *CsB* regulatory element for regulating *Hoxd* genes during

development of vertebrates AP axis. Moreover, it would consent a scenario of parallel snakelike evolution in snakes and amphisbaenians, involving similar changes in nucleotide sequence of the same regulatory element that likely account for establishment of limbless-elongated phenotypes. This study nurtures very interesting questions for future projects, especially after availability of snake genomes: 1) Are there additional signatures of independent snakelike evolution in the third conserved peak of *CsB*, which is displaced about 2 kb downstream in comparison with other vertebrates? 2) Does the complete snake *CsB* retain its regulatory potential on limb buds? 3) If we replace the snake sequence fragment where the snakelike regulatory signature was identified by its mouse orthologous region, does the *CsB* display its full regulatory capacity?

Chapter II described that snake evolution impaired regulatory ability of two limb-related regulatory elements, *CNS65* (*HoxD* telomeric landscape) and *Island I* (*HoxD* centromeric landscape). Functional evidence of their regulatory capacities degeneration was complemented with comparative analyses of predicted TFBS, which identified one putative stilopod/zeugopod-specific segment in *CNS65* and three autopod-specific candidate segments in *Island I*. Such regions are likely comprised in limb-specific regulatory modules within these enhancers that potentially evolved under relaxed selection together with or after limb loss. This chapter inspires future studies towards answering whether substitutions of the regions where the signatures were predicted in the sequences of snakes by their mouse orthologue segments would recover their regulatory activities in transgenic mice limb buds. Moreover, one might ask if mouse sequences of these elements would still elicit a similar regulatory activity if the same fragments of snakes occupied the regions of their mouse orthologues? Both complementary approaches would evaluate if the loss of regulatory function in the snake enhancers resides in degeneration of the candidate segments proposed here, also testing the relevance of these motifs for limb development. This chapter demonstrates that studies on the

molecular evolution of regulatory sequences contribute for identifying specific elements dedicated to unique developmental processes affected by particular evolutionary events. Specifically, I showed that by focusing on functional loss of regulatory sequences it is possible to select fragments within sequences that are possibly module-specific, using simple TFBS prediction (not as strict as predictions performed in Chapter I, because they are based on functional evidence), an approach that paves the way for discovering relevant elements within CREs that are dedicated to development of specific morphological structures in vertebrates.

Chapter III explored transcripts of noncoding sequences in the *HoxD* cluster of a marsupial family that exhibits several limb peculiarities when compared to other limbed vertebrates. This unity is connected with Chapter I through the conceptual framework focusing on recurrent evolution of specific phenotypes: evolution of particular traits in the tammar wallaby (i.e. the syndactylous feet) is also observed in a distantly related marsupial group (Peramelemorpha). The scenario is slightly different, however, as marsupial syndactyly is interpreted as a non-adaptive trait because its relevance for locomotion remains unknown; Diprotodontia (the family to which the kangaroos belong), for instance, is the most diversified marsupial order regarding locomotion and ecology (Weisbecker & Nilsson 2008). The most accepted current hypothesis states that marsupial syndactyly represents a case of parallelism established through an ontogenetic constraint (Weisbecker & Nilsson 2008). Although the topic has not been directly approached in my Thesis, this study represents an important step towards understanding the enigmatic genetic system that builds autopodial morphology in marsupials, as ncRNAs have been disclosed in the past decade as important regulatory elements. I explored transcription profiles and expression patterns of *HoxD* ncRNAs identified in embryo limbs, among which five lncRNAs were conserved in mammals and three of them might represent marsupial elements. These putative marsupial elements may constitute important regulatory elements involved in morphological evolution, therefore

representing interesting targets for functional investigation. As most ncRNAs examined showed similar transcriptional profiles and overlapping expression domains with their putative target *HOXD* coding genes, they may regulate *HOXD* expression during autopodial development. Because expression of the putative pre-miRNA *XLOC53* decreased in forelimb at the day before birth, it may also be involved in defining heterochrony between forelimbs and hindlimbs, a peculiarity of marsupials. This work provides the first foundation for exploring function of ncRNAs from cluster HoxD in the scenario of limb development and the evolutionary mechanisms linked to autopodial diversification in marsupials. Next steps of investigation may involve performing functional assays to access specific ncRNA roles during development, such as interference RNA and overexpression assays using tammar wallaby cell cultures where both coding and noncoding HoxD regions here studied are transcribed. As future questions addressed, I exemplify: 1) does inhibition or overexpression of these ncRNAs affect expression profiles of their putative target terminal *HOXD* genes? 2) do these ncRNAs act in *cis* or *trans*-regulatory mechanisms?

Overall, the present PhD Thesis validates the general hypothesis that regulatory elements of HoxD cluster carry histories of morphological divergence both in mammals and reptiles, placing the foundations for understanding developmental mechanisms that originated the diversity of extant vertebrate forms standing in our Planet.

LITERATURE CITED

- Abouheif, E., 2008. Parallelism as the Pattern and Process of Mesodermevolution. *Evolution and Development*, 10(1), pp.3–5.
- 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(3), pp.623–633.
- Akhtar, W. & Veenstra, G.J.C., 2011. TBP-related factors: a paradigm of diversity in transcription initiation. *Cell & bioscience*, 1(1), p.23.
- Akker, E. Van Den & Forlani, S., 2002. Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation. *Development*, 129, pp.2181–2193.
- Alexandrovich, A. et al., 2008. A role for GATA-6 in vertebrate chondrogenesis. *Developmental Biology*, 314(2), pp.457–470.
- Amemiya, C.T. et al., 2013. The African coelacanth genome provides insights into tetrapod evolution. *Nature*, 496(7445), pp.311–6.
- Amiel, J. et al., 2004. Polyalanine expansions in human. *Human Molecular Genetics*, 13(2), pp.235–243.
- Amrine-Madsen, H. et al., 2003. Nuclear gene sequences provide evidence for the monophyly of australidelphian marsupials. *Molecular phylogenetics and evolution*, 28, pp.186–196.
- Andrewartha, H. & Barker, S., 1969. Introduction to a Study of the Ecology of the Kangaroo Island Wallaby *Protemnodon-eugenii* within Flinders Chase Kangaroo Island South-Australia. *Transactions of the Royal Society of South Australia*, 93, pp.127–132.
- Andrey, G. et al., 2013. A Switch Between Topological Domains Underlies HoxD Genes Collinearity in Mouse Limbs. *Science*, 340, p.1195.
- Apesteguía, S. & Zaher, H., 2006. A Cretaceous terrestrial snake with robust hindlimbs and a sacrum. *Nature*, 440, pp.1037–1040.
- Archambeault, S., Taylor, J.A. & Crow, K.D., 2014. HoxA and HoxD expression in a variety of vertebrate body plan features reveals an ancient origin for the distal Hox program. *EvoDevo*, 5(44), pp.1–10.
- Aulehla, A. et al., 2003. Wnt3a plays a major role in the segmentation clock controlling somitogenesis. *Developmental Cell*, 4, pp.395–406.

- Averof, M. & Akam, M., 1995. Hox genes and the diversification of insect and crustacean body plans. *Nature*, 376, pp.420–423.
- Bae, E. et al., 2002. Characterization of the intergenic RNA profile at abdominal-A and Abdominal-B in the *Drosophila bithorax* complex. *Proc Natl Acad Sci U S A*, 99(26), pp.16847–16852.
- Baker, M. et al., 2004. Relationships among the families and orders of marsupials and the major mammalian lineages based on recombination activating gene-1. *Journal of Mammalian Evolution*, 11, pp.11–16.
- Balhoff, J.P. & Wray, G.A., 2005. Evolutionary analysis of the well characterized endo16 promoter reveals substantial variation within functional sites. *Proceedings of the National Academy of Sciences of the United States of America*, 102(24), pp.8591–6.
- Bejder, L. & Hall, B.K., 2002. Limbs in whales and limblessness in other vertebrates: Mechanisms of evolutionary and developmental transformation and loss. *Evolution and Development*, 4(6), pp.445–458.
- Bellairs, B.Y.A.D.A. & Underwood, G., 1950. The origin of snakes.
- Bellon, E., Luyten, F.P. & Tylzanowski, P., 2009. Delta-EF1 is a negative regulator of *Ihh* in the developing growth plate. *Journal of Cell Biology*, 187(5), pp.685–699.
- Bel-Vialar, S., Itasaki, N. & Krumlauf, R., 2002. Initiating Hox gene expression: in the early chick neural tube differential sensitivity to FGF and RA signaling subdivides the HoxB genes in two distinct groups. *Development*, 129(22), pp.5103–5115.
- Benton, M.J., 1990. Phylogeny of the major tetrapod groups: morphological data and divergence dates. *Journal of molecular evolution*, 30(5), pp.409–24.
- Berge, D. ten et al., 1998. Prx1 and Prx2 in skeletogenesis: roles in the craniofacial region, inner ear and limbs. *Development (Cambridge, England)*, 125(19), pp.3831–3842.
- Berger, P., 1966. No TitleEleven-month “Embryonic Diapause” Marsupial. *Letters to Nature*, 211, pp.435–436.
- Blignaut, M., Morris, K. V & Scripps, T., 2012. regulation of gene expression A book edited by Kevin Morris. , 7(6), pp.664–666.
- Bonnelye, E. et al., 1997. The ERR-1 orphan receptor is a transcriptional activator expressed during bone development. *Molecular endocrinology*, 11(7), pp.905–916.
- Brandley, M.C., Huelsenbeck, J.P. & Wiens, J.J., 2008. Rates and patterns in the evolution of snake-like body form in squamate reptiles: evidence for repeated re-evolution of lost digits and long-term persistence of intermediate body forms. *Evolution*, 62(8), pp.2042–64.

- Brand-Saberi, B., 2005. Genetic and epigenetic control of skeletal muscle development. *Annals of Anatomy*, 187(3), pp.199–207.
- Browner, N.C., Sellak, H. & Lincoln, T.M., 2004. Downregulation of cGMP-dependent protein kinase expression by inflammatory cytokines in vascular smooth muscle cells. *American journal of physiology. Cell physiology*, 287(1), pp.C88–96.
- Bruneau, S. et al., 2001. The Mouse Hoxd13spdh Mutation, a Polyalanine Expansion Similar to Human Type II Synpolydactyly (SPD), Disrupts the Function but Not the Expression of Other Hoxd Genes. *Developmental Biology*, 237(2), pp.345–353.
- Burel, A. et al., 2006. Role of HOXA7 to HOXA13 and PBX1 genes in various forms of MRKH syndrome (congenital absence of uterus and vagina). *Journal of negative results in biomedicine*, 5, p.4.
- Burke, A. et al., 1995. Hox genes and the evolution of vertebrate axial morphology. *Development*, 121(2), pp.333–346.
- Büscher, D. et al., 1997. Evidence for genetic control of Sonic hedgehog by Gli3 in mouse limb development. *Mechanisms of development*, 62(2), pp.175–82.
- 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.
- Caldwell, M., 2003. “Without a leg to stand on”: on the evolution and development of axial elongation and limblessness in tetrapods. *Canadian Journal of Earth Sciences*, 40, pp.573–588.
- Caldwell, M.W. & Lee, M.S.Y., 1997. A snake with legs from the marine Cretaceous of the Middle East. *Letters to Nature*, 386, pp.705–709.
- Capdevila, J. & Belmonte, J., 2001. Patterning mechanisms controlling vertebrate limb development. *Annual review of cell and developmental biology*, 17, pp.87–132.
- Capdevila, J. & Izpisua-Belmonte, J.-C., 2001. Patterning Mechanisms Controlling Vertebrate Limb Development. *Annual review of cell and developmental biology*, 17, pp.87–132.
- Capellini, T.D., Handschuh, K., et al., 2011. Control of pelvic girdle development by genes of the Pbx family and Emx2. *Developmental dynamics : an official publication of the American Association of Anatomists*, 240(5), pp.1173–89.
- Capellini, T.D. et al., 2008. Pbx1/Pbx2 govern axial skeletal development by controlling Polycomb and Hox in mesoderm and Pax1/Pax9 in sclerotome. *Developmental biology*, 321(2), pp.500–14.
- Capellini, T.D. et al., 2006. Pbx1/Pbx2 requirement for distal limb patterning is mediated by the hierarchical control of Hox gene spatial distribution and Shh expression. *Development (Cambridge, England)*, 133(11), pp.2263–73.

- Capellini, T.D. et al., 2010. Scapula development is governed by genetic interactions of PBX1 with its family members and with Emx2 via their cooperative control of Alx1. *Development*, 137, pp.2559–2569.
- Capellini, T.D., Zappavigna, V. & Selleri, L., 2011. Pbx homeodomain proteins: TALEnted regulators of limb patterning and outgrowth. *Developmental dynamics: an official publication of the American Association of Anatomists*, 240(5), pp.1063–86.
- Carapuço, M. et al., 2005. Hox genes specify vertebral types in the presomitic mesoderm. *Genes & development*, 19(18), pp.2116–21.
- Carroll, S.B., 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell*, 134(1), pp.25–36.
- Carroll, S.B., 2005. Evolution at two levels: on genes and form. *PLoS biology*, 3(7), pp.e245–e245.
- Carroll, S.B., 1995. Homeotic genes and the evolution of arthropods and chordates. *Nature*, 376, pp.479–485.
- Catala, M., Teillet, M.A. & Le Douarin, N.M., 1995. Organization and development of the tail bud analyzed with the quail-chick chimaera system. *Mechanisms of development*, 51(1), pp.51–65.
- Chen, A., Ginty, D. & Fan, C.-M., 2005. Protein kinase A signalling via CREB controls myogenesis induced by Wnt proteins. *Nature*, 433(7023), pp.317–322.
- Chen, C. et al., 2005. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic acids research*, 33(20), p.e179.
- Chen, F. & Capecchi, M.R., 1999. Paralogous mouse Hox genes, Hoxa9, Hoxb9, and Hoxd9, function together to control development of the mammary gland in response to pregnancy. *Proceedings of the National Academy of Sciences of the United States of America*, 96, pp.541–546.
- Chew, K.Y. et al., 2014. Heterochrony in the regulation of the developing marsupial limb. *Developmental dynamics : an official publication of the American Association of Anatomists*, 243(2), pp.324–338.
- Chew, K.Y. et al., 2012. HOXA13 and HOXD13 expression during development of the syndactylous digits in the marsupial *Macropus eugenii*. *BMC developmental biology*, 12(2), pp.1–15.
- Chimal-Monroy, J. et al., 2003. Analysis of the molecular cascade responsible for mesodermal limb chondrogenesis: Sox genes and BMP signaling. *Developmental Biology*, 257(2), pp.292–301.
- Cho, E. & Dressler, G.R., 1998. TCF-4 binds β -catenin and is expressed in distinct regions of the embryonic brain and limbs. *Mechanisms of Development*, 77(1), pp.9–18.

- Chotteau-Lelievre, A. et al., 2001. Expression patterns of the Ets transcription factors from the PEA3 group during early stages of mouse development. *Mechanisms of Development*, 108(1-2), pp.191–195.
- Coates, M. & Ruta, M., 2000. Nice snake, shame about the legs. *Trends in Ecology & Evolution*, 15(12), pp.503–507.
- Cohn, M. & Tickle, C., 1999. Developmental basis of limblessness and axial patterning in snakes. *Nature*, 399, pp.474–479.
- Colosimo, P.F., 2005. Widespread Parallel Evolution in Sticklebacks by Repeated Fixation of Ectodysplasin Alleles. *Science*, 307(5717), pp.1928–1933.
- Conrad, J.L., 2008. Phylogeny And Systematics Of Squamata (Reptilia) Based On Morphology. *Bulletin of the American Museum of Natural History*, 310(310), pp.1–182.
- Crow, K.D. & Wagner, G.P., 2006. Proceedings of the SMCBE Tri-National Young Investigators' Workshop 2005. What is the role of genome duplication in the evolution of complexity and diversity? *Molecular biology and evolution*, 23(5), pp.887–92.
- Darriba, D. et al., 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature methods*, 9(8), p.772.
- Dasen, J.S., 2013. Long noncoding RNAs in development: solidifying the Lncs to Hox gene regulation. *Cell reports*, 5(1), pp.1–2.
- Davis, a P. et al., 1995. Absence of radius and ulna in mice lacking *hoxa-11* and *hoxd-11*. *Nature*, 375(6534), pp.791–795.
- Davis, A. & Capecchi, M., 1994. Axial homeosis and appendicular skeleton defects in mice with a targeted disruption of *hoxd-11*. *Development (Cambridge, England)*, 120(8), pp.2187–2198.
- Deakin, J.E., 2012. Marsupial genome sequences: providing insight into evolution and disease. *Scientifica*, 2012, pp.1–22.
- Delpretti, S. et al., 2013. Multiple Enhancers Regulate *Hoxd* Genes and the Hotdog LncRNA during Cecum Budding. *Cell Reports*, 5(1), pp.137–150.
- Denans, N., Iimura, T. & Pourquié, O., 2015. Hox genes control vertebrate body elongation by collinear Wnt repression. *eLife*, 4, pp.1–33.
- Dequéant, M.-L. & Pourquié, O., 2008. Segmental patterning of the vertebrate embryonic axis. *Nature reviews. Genetics*, 9(5), pp.370–82.
- Deschamps, J., 2007. Ancestral and recently recruited global control of the Hox genes in development. *Current opinion in genetics & development*, 17(5), pp.422–7.

- Deschamps, J. & van Nes, J., 2005. Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development*, 132(13), pp.2931–42.
- Deutsch, U., Dressier, R. & Gruss, P., 1988. Pax 7, A Member of a Paired Box Homologous Murine Gene Family, Is Expressed in Segmented Structures during Development. *Cell*, 53, pp.617–625.
- Dinger, M.E. et al., 2008. Differentiating protein-coding and noncoding RNA: challenges and ambiguities. *PLoS computational biology*, 4(11), p.e1000176.
- Di-Poï, N. et al., 2010. Changes in Hox genes' structure and function during the evolution of the squamate body plan. *Nature*, 464(7285), pp.99–103.
- Di-Poï, N., Montoya-Burgos, J.I. & Duboule, D., 2009. Atypical relaxation of structural constraints in Hox gene clusters of the green anole lizard. *Genome research*, 19(4), pp.602–10.
- Dixon, J. et al., 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485, pp.376–380.
- Djebali, S. et al., 2012. Landscape of transcription in human cells. *Nature*, 489(7414), pp.101–8.
- Dollé, P. et al., 1989. Coordinate expression of the murine Hox-5 complex homeobox-containing genes during limb pattern formation. *Nature*, 342, pp.767–72.
- Duboule, D., 2007. The rise and fall of Hox gene clusters. *Development*, 134(14), pp.2549–2560.
- Duboule, D., 1992. The vertebrate limb: a model system to study the Hox/HOM gene network during development and evolution. *BioEssays*, 14(6), pp.375–384.
- Duboule, D., 1998. Vertebrate hox gene regulation: clustering and/or colinearity? *Current opinion in genetics & development*, 8(5), pp.514–518.
- Duboule, D. & Morata, G., 1994. Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends in Genetics*, 10(10), pp.358–364.
- Duprez, D. et al., 1996. Overexpression of BMP-2 and BMP-4 alters the size and shape of developing skeletal elements in the chick limb. *Mechanisms of Development*, 57(2), pp.145–157.
- Economides, K.D., 2003. Hoxb13 is required for normal differentiation and secretory function of the ventral prostate. *Development*, 130(10), pp.2061–2069.
- Fatica, A. & Bozzoni, I., 2014. Long non-coding RNAs: new players in cell differentiation and development. *Nature reviews. Genetics*, 15(1), pp.7–21.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, pp.783–791.

- Felsenstein, J., 2005. Phylip (phylogeny inference package) version 3.6 (2004)
Distributed by the author.
- Fisher, S. et al., 2006. Conservation of RET Regulatory Function from Human to Zebrafish Without Sequence Similarity. *Science*, 312(5771), pp.276–279.
- Fletcher, T., Shaw, G. & Renfree, M., 1990. Effects of bromocriptine at parturition in the tammar wallaby, *Macropus eugenii*. *Reproduction, Fertility and Development*, 2(1), pp.79–88.
- Fromental-Ramain, C. et al., 1996. Hoxa-13 and Hoxd-13 play a crucial role in the patterning of the limb autopod. *Development*, 122(10), pp.2997–3011.
- Fukuda, K. & Kikuchi, Y., 2005. Endoderm development in vertebrates: fate mapping, induction and regional specification. *Development, growth & differentiation*, 47, pp.343–55.
- Gajovic, S., Kostovic-Knezevic, L. & Svajger, A., 1993. Morphological evidence for secondary formation of the tail gut in the rat embryo. *Anatomy and Embriology*, 187(3), pp.291–297.
- Galceran, J. et al., 1999. Wnt-/- like phenotype and limb deficiency in Lef1-/-Tcf1-/- mice. *Genes*, 13, pp.709–717.
- Gehrke, A.R. et al., 2014. Deep conservation of wrist and digit enhancers in fish. *Proceedings of the National Academy of Sciences*, 112(3), pp.803–808.
- Gompel, N. et al., 2005. Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature*, 433, pp.481–487.
- Gonzalez, F., Duboule, D. & Spitz, F., 2007. Transgenic analysis of Hoxd gene regulation during digit development. *Developmental biology*, 306(2), pp.847–859.
- Goulding, M.D. et al., 1991. Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *The EMBO journal*, 10(5), pp.1135–47.
- Grant, P.R., 1999. *Ecology and evolution of Darwin's finches*, Princeton, NJ.: Princeton Univ. Press.
- Griffith, C., Wiley, M. & Sanders, E., 1992. The vertebrate tail bud: three germ layers from one tissue. *Anatomy and Embriology*, 185, pp.101–113.
- Guerreiro, I. et al., 2013. Role of a polymorphism in a Hox/Pax-responsive enhancer in the evolution of the vertebrate spine. *Proceedings of the National Academy of Sciences of the United States of America*, 110(26), pp.10682–6.
- Gustavsson, P. et al., 2007. Increased expression of Grainyhead-like-3 rescues spina bifida in a folate-resistant mouse model. *Human molecular genetics*, 16(21), pp.2640–6.

- Hall, B.K., 2012. Parallelism, deep homology, and evo-devo. *Evolution & development*, 14(1), pp.29–33.
- Hall, L., 1987. Syndactyly in marsupials-problems and prophecies. In *Possums and Opossums*. pp. 245–255.
- Head, J.J., 2015. Fossil calibration dates for molecular phylogenetic analysis of snakes 1 : *Paleontologica Electronica*, pp.1–17.
- Head, J.J. & Polly, P.D., 2015. Evolution of the snake body form reveals homoplasy in amniote Hox gene function. *Nature*, 520, pp.86–89.
- Higgins, D.G., Bleasby, A.J. & Fuchs, R., 1992. CLUSTAL V: improved software for multiple sequence alignment. *Computer applications in the biosciences*, 8(189-191).
- Hillis, D. & Bull, J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic biology*, 42(2), pp.182–192.
- Hirata, M. et al., 2009. C/EBP β promotes transition from proliferation to hypertrophic differentiation of chondrocytes through transactivation of p57Kip2. *PLoS ONE*, 4(2), p.e4543.
- Hogan, P.G. et al., 2003. Transcriptional regulation by calcium, calcineurin. and NFAT. *Genes Dev.*, 17(18), pp.2205–2232.
- Hosaka, T. et al., 2004. Disruption of forkhead transcription factor (FOXO) family members in mice reveals their functional diversification. *Proceedings of the National Academy of Sciences of the United States of America*, 101(9), pp.2975–2980.
- James, R. & Kazenwadel, J., 1991. Homeobox gene expression in the intestinal epithelium of adult mice. *The Journal of biological chemistry*, 266(5), pp.3246–51.
- Jeong, S., Rokas, A. & Carroll, S.B., 2006. Regulation of body pigmentation by the Abdominal-B Hox protein and its gain and loss in Drosophila evolution. *Cell*, 125(7), pp.1387–99.
- Jones, W.F., 1925. *The mammals of south australia*, Adelaide, Australia: Government Printer.
- Kanagae, Y. et al., 1998. Role of Rel/NF-kB transcription factors during the outgrowth of the vertebrate limb. *Nature*, 392, pp.611–614.
- Karamboulas, K., Dranse, H.J. & Underhill, T.M., 2010. Regulation of BMP-dependent chondrogenesis in early limb mesenchyme by TGF β signals. *Journal of cell science*, 123, pp.2068–2076.
- Kardon, G., Harfe, B.D. & Tabin, C.J., 2003. A Tcf4-positive mesodermal population provides a prepattern for vertebrate limb muscle patterning. *Developmental Cell*, 5(6), pp.937–944.

- Kawakami, Y. et al., 2006. Wnt/beta-catenin signaling regulates vertebrate limb regeneration. *Genes & Development*, (858), pp.3232–3237.
- Kel, A.E., 2003. MATCHTM: a tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Research*, 31(13), pp.3576–3579.
- Kessel, M. & Gruss, P., 1991. Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell*, 67(1), pp.89–104.
- Keyte, A. & Smith, K.K., 2012. Heterochrony in somitogenesis rate in a model marsupial, *Monodelphis domestica*. *Evolution & development*, 14(1), pp.93–103.
- Keyte, A.L. & Smith, K.K., 2010. Developmental origins of precocial forelimbs in marsupial neonates. *Development (Cambridge, England)*, 137(24), pp.4283–4294.
- Kirsch, J., 1977. The comparative serology of Marsupialia, and a classification of marsupial. *Australian Journal of Zoology Supplementary Series*, 25(52), pp.1–152.
- Kmita, M. et al., 2005. Early developmental arrest of mammalian limbs lacking HoxA/HoxD gene function. *Nature*, 435(7045), pp.1113–6.
- Kmita, M. & Duboule, D., 2003. Organizing axes in time and space; 25 years of colinear tinkering. *Science (New York, N.Y.)*, 301(5631), pp.331–333.
- Knezevic, V., De Santo, R. & Mackem, S., 1998. Continuing organizer function during chick tail development. *Development*, 125(10), pp.1791–801.
- Kohlsdorf, T. et al., 2010. Data and Data Interpretation in the Study of Limb Evolution: a Reply To Galis Et Al. on the Reevolution of Digits in the Lizard Genus *Bachia*. *Evolution*, 64(8), pp.2477–2485.
- Kohlsdorf, T. & Wagner, G., 2006. Evidence for the reversibility of digit loss: a phylogenetic study of limb evolution in *Bachia* (Gymnophthalmidae: Squamata). *Evolution*, 60(9), pp.1896–1912.
- Kondrashov, N. et al., 2011. Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell*, 145(3), pp.383–97.
- Kos, L., Chiang, C. & Mahon, K.A., 1998. Mediolateral patterning of somites: multiple axial signals, including Sonic hedgehog, regulate Nkx-3.1 expression. *Mechanisms of development*, 70(1-2), pp.25–34.
- Kozhemyakina, E., Ionescu, A. & Lassar, A.B., 2014. GATA6 Is a Crucial Regulator of Shh in the Limb Bud. *PLoS Genetics*, 10(1), p.e1004072.
- Krawchuk, D. et al., 2010. Twist1 activity thresholds define multiple functions in limb development. *Developmental biology*, 347(1), pp.133–46.
- Kumar, B. et al., 2015. Analysis of dynamic changes in retinoid induced transcription and epigenetic profiles of murine Hox clusters in ES cells. *Genome research*, in press.

- Lee, M.S.Y., 2005. Molecular evidence and marine snake origins. *Biology Letters*, 1(2), pp.227–230.
- Lemons, D. & McGinnis, W., 2006. Genomic Evolution of Hox Gene Clusters. *Science*, 313(5795), pp.1918–1922.
- Lemos, B. et al., 2004. Phylogenetic footprinting reveals extensive conservation of Sonic Hedgehog (SHH) regulatory elements. *Genomics*, 84(3), pp.511–523.
- Lewis, E.B., 1978. A gene complex controlling segmentation in *Drosophila*. *Nature*, 276(5688), pp.565–570.
- Li, L. et al., 2013. Targeted disruption of *Hotair* leads to homeotic transformation and gene derepression. *Cell reports*, 5(1), pp.3–12.
- Li, L. & Chang, H.Y., 2014. Physiological roles of long noncoding RNAs: insight from knockout mice. *Trends in Cell Biology*, 24(10), pp.594–602.
- Lindsay, J. et al., 2012. Unique small RNA signatures uncovered in the tammar wallaby genome. *BMC genomics*, 13, p.559.
- Long, F. et al., 2001. The CREB family of activators is required for endochondral bone development. *Development*, 128(4), pp.541–550.
- Lorda-Diez, C.I. et al., 2009. Transforming growth factors β coordinate cartilage and tendon differentiation in the developing limb mesenchyme. *Journal of Biological Chemistry*, 284(43), pp.29988–29996.
- Losos, J.B., Hillis, D.M. & Greene, H.W., 2012. Who speaks with a forked tongue. *Science*, 338, pp.1428–1429.
- Luckett, W., 1994. Suprafamilial relationships within Marsupialia: resolution and discordance from multidisciplinary data. *Journal of Mammalian Evolution*, 2, pp.225–283.
- Luo, Z.-X., 2007. Transformation and diversification in early mammal evolution. *Nature*, 450, pp.1011–1019.
- Mallo, M., Wellik, D.M. & Deschamps, J., 2010. Hox genes and regional patterning of the vertebrate body plan. *Developmental Biology*, 344(1), pp.7–15.
- Mansfield, J.H. et al., 2004. MicroRNA-responsive “sensor” transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. *Nature genetics*, 36(10), pp.1079–1083.
- Marigo, V. & Tabin, C.J., 1996. Regulation of patched by sonic hedgehog in the developing neural tube. *Proceedings of the National Academy of Sciences of the United States of America*, 93(18), pp.9346–51.

- Marshall, C.R., Raff, E.C. & Raff, R.A., 1994. Dollo's law and the death and resurrection of genes. *Proceedings of the National Academy of Sciences of the United States of America*, 91(25), pp.12283–12287.
- Marshall, L., 1973. Evolution of the peramelid tarsus. *Proceedings of the Royal Society of Victoria*, 85, pp.51–60.
- McMahon, J.A. et al., 1998. Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. *Genes & Development*, 12, pp.1438–1452.
- Mercader, N. et al., 1999. Conserved regulation of proximodistal limb axis development by Meis1/Hth. *Nature*, 402(6760), pp.425–429.
- Mercader, N., Tanaka, E.M. & Torres, M., 2005. Proximodistal identity during vertebrate limb regeneration is regulated by Meis homeodomain proteins. *Development*, 132(18), pp.4131–4142.
- Moens, C.B. & Selleri, L., 2006. Hox cofactors in vertebrate development. *Developmental biology*, 291(2), pp.193–206.
- Montavon, T. et al., 2011. A regulatory archipelago controls Hox genes transcription in digits. *Cell*, 147(5), pp.1132–1145.
- Montavon, T. et al., 2008. Modeling Hox gene regulation in digits: reverse collinearity and the molecular origin of thumbness. *Genes & development*, 22(3), pp.346–359.
- Montavon, T. & Soshnikova, N., 2014. Hox gene regulation and timing in embryogenesis. *Seminars in cell & developmental biology*, 34C, pp.76–84.
- Mott, T. & Vieites, D.R., 2009. Molecular phylogenetics reveals extreme morphological homoplasy in Brazilian worm lizards challenging current taxonomy. *Molecular Phylogenetics and Evolution*, 51(2), pp.190–200.
- Münsterberg, A. & Lassar, A., 1995. Combinatorial signals from the neural tube, floor plate and notochord induce myogenic bHLH gene expression in the somite. *Development*, 121, pp.651–660.
- Nelson, C.E. et al., 1996. Analysis of Hox gene expression in the chick limb bud. *Development*, 122(5), pp.1449–1466.
- Niederreither, K. et al., 2002. Embryonic retinoic acid synthesis is required for forelimb growth and anteroposterior patterning in the mouse. *Development*, 129(15), pp.3563–3574.
- Nilsson, M. et al., 2004. Marsupial relationships and a timeline for marsupial radiation in South Gondwana. *Gene*, 340(2), pp.189–196.
- Noordermeer, D. & Duboule, D., 2013. Chromatin Architectures and Hox Gene Collinearity. In *Current Topics in Developmental Biology*. pp. 113–148.

- Nora, E. et al., 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, 485(7398), pp.381–385.
- O'Rourke, M.P. et al., 2002. Twist Plays an Essential Role in FGF and SHH Signal Transduction during Mouse Limb Development. *Developmental Biology*, 248(1), pp.143–156.
- Ovcharenko, I. et al., 2005. Mulan: multiple-sequence local alignment and visualization for studying function and evolution. *Genome research*, 15(1), pp.184–194.
- Packard, D.S.J. & Jacobson, A.G., 1976. The influence of axial structures on chick somite formation. *Developmental Biology*, 53(1), pp.36–48.
- Palci, A. & Caldwell, M.W., 2007. Vestigial forelimbs and axial elongation in a 95 million-year-old non-snake squamate. *Journal of Vertebrate Paleontology*, 27(1), pp.1–7.
- Pang, K.C., Frith, M.C. & Mattick, J.S., 2006. Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends in genetics*, 22(1), pp.1–5.
- Panopoulou, G. & Poustka, a, 2005. Timing and mechanism of ancient vertebrate genome duplications – the adventure of a hypothesis. *Trends in Genetics*, 21(10), pp.559–567.
- Park, H.L. et al., 2000. Mouse Gli1 mutants are viable but have defects in SHH signaling in combination with a Gli2 mutation. *Development*, 127(8), pp.1593–1605.
- Parker, H.J. et al., 2011. Ancient Pbx-Hox signatures define hundreds of vertebrate developmental enhancers. *BMC Genomics*, 12, p.637.
- Parr, B.A. & McMahon, A.P., 1995. Dorsalizing signal Wnt-7a required for normal polarity of D–V and A–P axes of mouse limb. *Nature*, 374, pp.350–353.
- Peer, Y. Van de, Maere, S. & Meyer, A., 2009. The evolutionary significance of ancient genome duplications. *Nature Reviews Genetics*, 10, pp.725–732.
- Pellegrino, K. et al., 2001. A molecular perspective on the evolution of microteiid lizards (Squamata, Gymnophthalmidae), and a new classification for the family. *Biological Journal of the Linnean Society*, 74(3), pp.315–338.
- Petruk, S. et al., 2006. Transcription of bxd Noncoding RNAs Promoted by Trithorax Represses Ubx in cis by Transcriptional Interference. *Cell*, 127(6), pp.1209–1221.
- Pfaffl, M.W., Horgan, G.W. & Dempfle, L., 2002. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic acids research*, 30(9), p.e36.
- Phillips, M.J. et al., 2006. Combined mitochondrial and nuclear DNA sequences resolve the interrelations of the major Australasian marsupial radiations. *Systematic biology*, 55(1), pp.122–137.

- Ponting, C.P., Oliver, P.L. & Reik, W., 2009. Evolution and functions of long noncoding RNAs. *Cell*, 136(4), pp.629–641.
- Pough, H., 1998. *Herpetology* 1st ed., Prentice Hall.
- Price, M. et al., 1992. Regional expression of the homeobox gene Nkx-2.2 in the developing mammalian forebrain. *Neuron*, 8(2), pp.241–55.
- Provot, S. & Schipani, E., 2005. Molecular mechanisms of endochondral bone development. *Biochemical and Biophysical Research Communications*, 328(3), pp.658–665.
- Prud'homme, B. et al., 2006. Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature*, 440(7087), pp.1050–3.
- Pyron, R.A., Burbrink, F.T. & Wiens, J.J., 2013. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC evolutionary biology*, 13, p.93.
- Rallis, C. et al., 2003. Tbx5 is required for forelimb bud formation and continued outgrowth. *Development*, 130(12), pp.2741–2751.
- Ralston, G., 2013. Article HOTAIR : Can we make sense of this anti-sense RNA ? *Chemistry in New Zealand*, (October), pp.116–119.
- Ranger, A.M. et al., 2000. The nuclear factor of activated T cells (NFAT) transcription factor NFATp (NFATc2) is a repressor of chondrogenesis. *The Journal of experimental medicine*, 191(1), pp.9–22.
- Renfree, M.B. et al., 2011. Genome sequence of an Australian kangaroo, *Macropus eugenii*, provides insight into the evolution of mammalian reproduction and development. *Genome Biology*, 12(8), p.R81.
- Renfree, M.B. & Tyndale-Biscoe, C.H., 1973. Intrauterine Development after Diapause in the Marsupial *Macropus eugenii*. *Developmental biology*, 32, pp.28–40.
- Reynolds, R., Niemiller, M. & Revell, L., 2014. Toward a Tree-of-Life for the boas and pythons: Multilocus species-level phylogeny with unprecedented taxon sampling. *Molecular Phylogenetics and Evolution*, 71, pp.201–213.
- Riddle, R.D. et al., 1993. Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell*, 75(7), pp.1401–1416.
- Rieppel, O. et al., 2003. The anatomy and relationships of *Haasiophis terrasanctus*, a fossil snake with well-developed hind limbs from the Mid-Cretaceous of the middle east. *Journal of Paleontology*, 77(3), pp.536–558.
- Ringrose, L. & Paro, R., 2007. Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development*, 134(2), pp.223–232.

- Rinn, J.L. et al., 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, 129(7), pp.1311–1323.
- Roberts, D.J. et al., 1995. Sonic hedgehog is an endodermal signal inducing Bmp-4 and Hox genes during induction and regionalization of the chick hindgut. *Development*, 121(10), pp.3163–74.
- Rombel, I.T. et al., 2002. ORF-FINDER: a vector for high-throughput gene identification. *Gene*, 282(1-2), pp.33–41.
- Rudnicki, M. a & Jaenisch, R., 1995. The MyoD family of transcription factors and skeletal myogenesis. *BioEssays*, 17(3), pp.203–209.
- Rundle, H.D. et al., 2000. Natural Selection and Parallel Speciation in Sympatric Sticklebacks. *Science*, 287, pp.306–308.
- Sadleir, R.M.F.S. & Tyndale-Biscoe, C.H., 1977. Photoperiod and the Termination of Embryonis Diapause in the Marsupial eugenii. *Biology of reproduction*, 16, pp.605–608.
- Sahni, M. et al., 1999. FGF signaling inhibits chondrocyte proliferation and regulates bone development through the STAT-1 pathway FGF signaling inhibits chondrocyte proliferation and regulates bone development through the STAT-1 pathway. *Genes & Development*, 13, pp.1361–1366.
- Saleh, M. et al., 2000. Cell Signaling Switches HOX-PBX Complexes from Repressors to Activators of Transcription Mediated by Histone Deacetylases and Histone Acetyltransferases. *Molecular and Cellular Biology*, 20(22), pp.8623–8633.
- Sanchez-Elsner, T., 2006. Noncoding RNAs of Trithorax Response Elements Recruit Drosophila Ash1 to Ultrabithorax. *Science*, 311(5764), pp.1118–1123.
- Sanger, T.J. & Gibson- Brown, J.J., 2004. The developmental bases of limb reduction and body elongation in squamates. *Evolution*, 58(9), pp.2103–2106.
- Santa Barbara, P. & Roberts, D., 2002. Tail gut endoderm and gut/genitourinary/tail development: a new tissue-specific role for Hoxa13. *Development*, 129(3), pp.551–61.
- Saxena, A. & Carninci, P., 2011. Long non-coding RNA modifies chromatin: epigenetic silencing by long non-coding RNAs. *BioEssays*, 33(11), pp.830–839.
- Scanlon, J.D. et al., 1999. The palaeocology of the primitive snake Pachyrhachis. *Historical Biology*, 13, pp.127–152.
- Schmitt, S., Prestel, M. & Paro, R., 2005. Intergenic transcription through a Polycomb group response element counteracts silencing. *Genes and Development*, 19(6), pp.697–708.

- Schneider, I. et al., 2011. Appendage expression driven by the Hoxd Global Control Region is an ancient gnathostome feature. *Proceedings of the National Academy of Sciences of the United States of America*, 108(31), pp.1–5.
- Schneider, I. & Shubin, N.H., 2013. The origin of the tetrapod limb: from expeditions to enhancers. *Trends in genetics*, 29(7), pp.419–426.
- Schneider, N.Y. et al., 2009. The olfactory system of the tammar wallaby is developed at birth and directs the neonate to its mother's pouch odours. *Reproduction*, 138(5), pp.849–857.
- Schorderet, P. & Duboule, D., 2011. Structural and Functional Differences in the Long Non-Coding RNA Hotair in Mouse and Human. *PLoS Genetics*, 7(5), p.e1002071.
- Scott, V., Morgan, E.A. & Stadler, H.S., 2005. Genitourinary functions of Hoxa13 and Hoxd13. *Journal of biochemistry*, 137(6), pp.671–6.
- Sears, K.E., 2009. Differences in the timing of prechondrogenic limb development in mammals: the marsupial-placental dichotomy resolved. *Evolution; international journal of organic evolution*, 63(8), pp.2193–2200.
- Sela-Donenfeld, D. & Kalcheim, C., 2000. Inhibition of noggin expression in the dorsal neural tube by somitogenesis: a mechanism for coordinating the timing of neural crest emigration. *Development*, 127, pp.4845–4854.
- Selleri, L. et al., 2001. Requirement for Pbx1 in skeletal patterning and programming chondrocyte proliferation and differentiation. *Development*, 128(18), pp.3543–57.
- Sensiate, L.A. et al., 2014. Dact gene expression profiles suggest a role for this gene family in integrating Wnt and TGF- β signaling pathways during chicken limb development. *Developmental Dynamics*, 243(3), pp.428–439.
- Shahi, M.H., Rey, J.A. & Castresana, J.S., 2012. The sonic hedgehog-GLI1 signaling pathway in brain tumor development. *Expert Opinion on Therapeutic Targets*, 16(12), pp.1227–1238.
- Shapiro, M.D. et al., 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature*, 428(6984), pp.717–723.
- Shapiro, M.D., Bell, M.A. & Kingsley, D.M., 2006. Parallel genetic origins of pelvic reduction in vertebrates. *Proceedings of the National Academy of Sciences of the United States of America*, 103(37), pp.13753–13758.
- Shapiro, M.D., Hanken, J. & Rosenthal, N., 2003. Developmental basis of evolutionary digit loss in the Australian lizard *Hemiergis*. *Journal of experimental zoology. Part B*, 297(1), pp.48–56.
- Shaw, G. & Renfree, M.B., 2006. Parturition and perfect prematurity- birth in marsupials. *Australian Journal of Zoology*, 54, pp.139–149.

- Sheth, R. et al., 2012. Hox genes regulate digit patterning by controlling the wavelength of a Turing-type mechanism. *Science (New York, N.Y.)*, 338(6113), pp.1476–80.
- Singarete, M. et al., 2015. Molecular evolution of HoxA13 and the multiple origins of limbless morphologies in amphibians and reptiles. *Genetics and Molecular Biology*, 38(3), p.ISSN 1678–4685.
- Sivanantharajah, L. & Percival-Smith, A., 2015. Differential pleiotropy and HOX functional organization. *Developmental Biology*, 398(1), pp.1–10.
- Small, K.M. & Potter, S.S., 1993. Homeotic transformations and limb defects in Hox A11 mutant mice. *Genes & Development*, 7(12), pp.2318–2328.
- Soshnikova, N., 2014. Hox genes regulation in vertebrates. *Developmental dynamics*, 243(1), pp.49–58.
- Soshnikova, N. & Duboule, D., 2009. Epigenetic temporal control of mouse Hox genes in vivo. *Science*, 324(5932), pp.1320–3.
- Spitz, F., 2010. Control of Vertebrate Hox Clusters by Remote and Global Cis -Acting Regulatory Sequences. In J. S. Deutsch, ed. *Hox Genes: Studies from the 20th to the 21st Century*. Heidelberg: Landes Bioscience and Springer+Business Media, pp. 63–78.
- Spitz, F. et al., 2005. Inversion-induced disruption of the Hoxd cluster leads to the partition of regulatory landscapes. *Nature genetics*, 37(8), pp.889–93.
- Spitz, F. et al., 2001. Large scale transgenic and cluster deletion analysis of the HoxD complex separate an ancestral regulatory module from evolutionary innovations. *Genes & development*, 15(17), pp.2209–2214.
- Spitz, F. & Duboule, D., 2008. Global control regions and regulatory landscapes in vertebrate development and evolution. In *Advances in genetics*. pp. 176–205.
- Spitz, F., Gonzalez, F. & Duboule, D., 2003. A Global Control Region Defines a Chromosomal Regulatory Landscape Containing the HoxD Cluster. *Cell*, 113, pp.405–417.
- Stanfel, M.N. et al., 2005. Regulation of organ development by the NKX-homeodomain factors: an NKX code. *Cellular and molecular biology*, 51, pp.OL785–799.
- Stricker, S. et al., 2002. Role of Runx genes in chondrocyte differentiation. *Developmental biology*, 245(1), pp.95–108.
- Sucena, E. et al., 2003. Regulatory evolution of shavenbaby/ovo underlies multiple cases of morphological parallelism. *Nature*, 424(6951), pp.935–938.
- Suryamohan, K. & Halfon, M.S., 2015. Identifying transcriptional cis - regulatory modules in animal genomes. *WIREs Developmental Biology*, 4(4), pp.59–84.

- Szalay, F., 1994. *Evolutionary history of the marsupials and an analysis of osteological characters*, New York: Cambridge University Press.
- Takagi, T. et al., 1998. DeltaEF1, a zinc finger and homeodomain transcription factor, is required for skeleton patterning in multiple lineages. *Development*, 125(1), pp.21–31.
- Takahashi, Y. et al., 2004. Expression profiles of 39 HOX genes in normal human adult organs and anaplastic thyroid cancer cell lines by quantitative real-time RT-PCR system. *Experimental Cell Research*, 293(1), pp.144–153.
- Tarchini, B. & Duboule, D., 2006. Control of Hoxd genes' collinearity during early limb development. *Developmental cell*, 10(1), pp.93–103.
- Taylor, H.S., Vanden Heuvel, G.B. & Igarashi, P., 1997. A conserved Hox axis in the mouse and human female reproductive system: late establishment and persistent adult expression of the Hoxa cluster genes. *Biology of reproduction*, 57(6), pp.1338–1345.
- Tchernov, E. et al., 2000. A fossil snake with limbs. *Science*, 287, pp.2010–2012.
- The UniProt Consortium, 2008. The universal protein resource (UniProt). *Nucleic acids research*, p.36; D190–D195.
- Thompson, J.D., Higgins, D.G. & Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*, 22(22), pp.4673–4680.
- Thornton, B. & Basu, C., 2011. Real-time PCR (qPCR) primer design using free online software. *Biochemistry and Molecular Biology Education*, 39(2), pp.145–154.
- Tickle, C., 2006. Making digit patterns in the vertebrate limb. *Nature reviews. Molecular cell biology*, 7(1), pp.45–53.
- Trapnell, C. et al., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols*, 7(3), pp.562–78.
- Trapnell, C. et al., 2010. Transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms. *Nature Biotechnology*, 28(5), pp.511–515.
- Trapnell, C., Pachter, L. & Salzberg, S.L., 2009. TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics*, 25(9), pp.1105–1111.
- Tschopp, P. et al., 2014. A relative shift in cloacal location repositions external genitalia in amniote evolution. *Nature*, 516, pp.391–394.
- Tschopp, P. et al., 2011. Reshuffling genomic landscapes to study the regulatory evolution of Hox gene clusters. *Proceedings of the National Academy of Sciences of the United States of America*, 108(26), pp.10632–10637.

- Tschopp, P. & Duboule, D., 2011. A genetic approach to the transcriptional regulation of Hox gene clusters. *Annual review of genetics*, 45, pp.145–166.
- Uetz, P., 2007. *The EMBL reptile database*, Heidelberg: European Molecular Biology Laboratory.
- Ulitsky, I. et al., 2011. Conserved Function of lincRNAs in Vertebrate Embryonic Development despite Rapid Sequence Evolution. *Cell*, 147(7), pp.1537–1550.
- Untergasser, A. et al., 2007. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research*, 35, pp.71–74.
- Vargas, A.O. et al., 2008. The evolution of HoxD-11 expression in the bird wing: insights from Alligator mississippiensis. *PloS one*, 3(10), p.e3325.
- Van de Ven, C. et al., 2011. Concerted involvement of Cdx/Hox genes and Wnt signaling in morphogenesis of the caudal neural tube and cloacal derivatives from the posterior growth zone. *Development*, 138(16), pp.3451–3462.
- Vidal, N. et al., 2010. Blindsnake evolutionary tree reveals long history on Gondwana. *Biology letters*, 6(4), pp.558–561.
- Vidal, N. & Hedges, S.B., 2004. Molecular evidence for a terrestrial origin of snakes. *Proceedings of the Royal Society B: Biological Sciences*, 271, pp.S226–S229.
- Vidal, N. & Hedges, S.B., 2009. The molecular evolutionary tree of lizards, snakes, and amphisbaenians. *Evolution*, 332, pp.129–139.
- Vidal, N. & Hedges, S.B., 2005. The phylogeny of squamate reptiles (lizards, snakes, and amphisbaenians) inferred from nine nuclear protein-coding genes. *Comptes rendus biologies*, 328(10-11), pp.1000–1008.
- Villarejo-Balcells, B. et al., 2011. Expression pattern of the FoxO1 gene during mouse embryonic development. *Gene Expression Patterns*, 11(5-6), pp.299–308.
- Villavicencio, E.H. et al., 2002. Cooperative E-box regulation of human GLI1 by TWIST and USF. *Genesis*, 32(4), pp.247–258.
- Wake, D.B., Wake, M.H. & Specht, C.D., 2011. Homoplasy: from detecting pattern to determining process and mechanism of evolution. *Science*, 331(6020), pp.1032–5.
- Wang, J. et al., 2004. Mouse transcriptome: Neutral evolution of “non-coding” complementary DNAs. *Nature*, 431.
- Wang, K.C. et al., 2011. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*, 472(7341), pp.120–124.
- Wang, L. et al., 2013. CPAT: Coding-potential assessment tool using an alignment-free logistic regression model. *Nucleic Acids Research*, 41(6), pp.1–7.

- Wang, X. & Chamberlin, H.M., 2002. Multiple regulatory changes contribute to the evolution of the *Caenorhabditis* lin-48 ovo gene. *Genes & development*, 16(18), pp.2345–9.
- Warot, X. et al., 1997. Gene dosage-dependent effects of the *Hoxa-13* and *Hoxd-13* mutations on morphogenesis of the terminal parts of the digestive and urogenital tracts. *Development*, 124(23), pp.4781–91.
- Warren, R. et al., 1994. Evolution of homeotic gene regulation and function in flies and butterflies. *Nature*, 372, pp.458–461.
- Weisbecker, V. & Nilsson, M., 2008. Integration, heterochrony, and adaptation in pedal digits of syndactylous marsupials. *BMC evolutionary biology*, 8(160), pp.1–14.
- Wellik, D.M. & Capecchi, M.R., 2003. *Hox10* and *Hox11* Genes Are Required to Globally Pattern the Mammalian Skeleton. *Science*, 301, pp.363–367.
- Welscher, P. et al., 2002. Progression of vertebrate limb development through SHH-mediated counteraction of *GLI3*. *Science*, 298(5594), pp.827–830.
- Wiens, J.J., Brandley, M.C. & Reeder, T.W., 2006. Why does a trait evolve multiple times within a clade? Repeated evolution of snakelike body form in squamate reptiles. *Evolution*, 60(1), pp.123–41.
- Wiens, J.J. & Slingluff, J.L., 2001. How lizards turn into snakes : A phylogenetic analyses of body-form evolution in anguillid lizards. *Evolution*, 55(11), pp.2303–2318.
- Williams, J.G., 2000. STAT signalling in cell proliferation and in development. *Current Opinion in Genetics and Development*, 10(5), pp.503–507.
- Wolf, C. et al., 1991. The *M-twist* gene of *Mus* is expressed in subsets of mesodermal cells and is closely related to the *Xenopus X-twi* and the *Drosophila twist* genes. *Developmental biology*, 143(2), pp.363–73.
- Woltering, J.M. et al., 2009. Axial patterning in snakes and caecilians: evidence for an alternative interpretation of the Hox code. *Developmental biology*, 332(1), pp.82–89.
- Woltering, J.M. et al., 2014. Conservation and divergence of regulatory strategies at Hox Loci and the origin of tetrapod digits. *PLoS biology*, 12(1), p.e1001773.
- Woltering, J.M., 2012. From lizard to snake; behind the evolution of an extreme body plan. *Current genomics*, 13(4), pp.289–99.
- Xue, S. et al., 2015. RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation. *Nature*, 517(7532), pp.33–8.
- Yasutake, J., Inohaya, K. & Kudo, A., 2004. Twist functions in vertebral column formation in medaka, *Oryzias latipes*. *Mechanisms of development*, 121(7-8), pp.883–94.

- Yekta, S., Tabin, C.J. & Bartel, D.P., 2008. MicroRNAs in the Hox network: an apparent link to posterior prevalence. *Nature reviews. Genetics*, 9(10), pp.789–96.
- Yokouchi, Y., Sakiyama, J. & Kuroiwa, A., 1995. Coordinated expression of Abd-B subfamily genes of the HoxA cluster in the developing digestive tract of chick embryo. *Developmental biology*, 169, pp.76–89.
- Yoshida, C. a. et al., 2004. Runx2 and Runx3 are essential for chondrocyte maturation, and Runx2 regulates limb growth through induction of Indian hedgehog. *Genes and Development*, 18(8), pp.952–963.
- Young, T. et al., 2009. Cdx and Hox genes differentially regulate posterior axial growth in mammalian embryos. *Developmental cell*, 17(4), pp.516–26.
- Yu, H. et al., 2012. Evolution of coding and non-coding genes in HOX clusters of a marsupial. *BMC Genomics*, 13(251), pp.1–15.
- Zakany, J. & Duboule, D., 2007. The role of Hox genes during vertebrate limb development. *Current opinion in genetics & development*, 17(4), pp.359–366.
- Zákány, J. & Duboule, D., 1999. Hox genes in digit development and evolution. *Cell and tissue research*, 296(1), pp.19–25.
- Zákány, J., Kmita, M. & Duboule, D., 2004. A dual role for Hox genes in limb anterior-posterior asymmetry. *Science (New York, N. Y.)*, 304(5677), pp.1669–1672.
- Zinzen, R.P. et al., 2006. Evolution of the ventral midline in insect embryos. *Developmental cell*, 11(6), pp.895–902.
- Zuniga, A. et al., 2004. Mouse limb deformity mutations disrupt a global control region within the large regulatory landscape required for Gremlin expression. *Genes & development*, 18(13), pp.1553–64.
- Zuniga, A. et al., 2002. Mouse Twist is required for fibroblast growth factor-mediated epithelial-mesenchymal signalling and cell survival during limb morphogenesis. *Mechanisms of development*, 114(1-2), pp.51–9.