Evolution of Morphological Novelty Through the Repurposing of Cell Programs

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy
of Freie Universität Berlin

by

Magdalena Schindler
Berlin, September 2024

The present work was conducted from June 2019 to September 2024 at the Max Planck Institute for Molecular Genetics in the research group of Prof. Dr. Stefan Mundlos.

1st reviewer: Prof. Dr. Stefan Mundlos

2nd reviewer: Prof. Dr. Sigmar Stricker

Date of defense: 31.01.2025

Declaration of Authorship

I hereby declare that I alone am responsible for the content of my doctoral dissertation and that I have only used the sources or references cited in the dissertation.

Magdalena Schindler

Table of contents

1. Abstract

The remarkable morphological diversity observed across the animal kingdom is a result of the evolutionary modifications of developmental processes. A striking example of such diverse morphologies is the tetrapod limb where profound phenotypic adaptations enabled species to colonize diverse terrestrial, aquatic, and aerial habitats. This organ thus provides a compelling model for examining the cellular mechanisms driving such diversification.

Among mammals, powered flight evolved uniquely within the order Chiroptera, commonly known as bats. This remarkable ability is attributed to their wing structure, which is characterized by massively elongated digits connected by a specialized tissue membrane, the chiropatagium, a distinctive feature found only in this species. However, the molecular origins and the extent of developmental reprogramming necessary to achieve this dramatic morphological adaptation remain poorly understood.

To address this, we employed comparative single-cell and functional genomics approaches to investigate limb developmental cell states in both mice and bats. Our analysis revealed a significant conservation of cell states and processes across species. In contrast to previous hypotheses, this also included the process of interdigital apoptosis, which typically eliminates interdigital tissue in species with separated digits. Through microdissection of embryonic chiropatagium and subsequent single-cell transcriptomics profiling of this tissue, we discovered that the chiropatagium originates from fibroblasts that are distinct from apoptosis-related interdigital cells. We furthermore revealed that these fibroblasts in the distal, autopodial limb region repurpose a developmental program otherwise found in the proximal limb. Genes expressed in this chiropatagium cell program showed enrichment in functions related to cell proliferation, migration, and extracellular matrix organization, highlighting the significance of these processes in wing development. Functional genomics and gene network analyses identified the developmental transcription factors *MEIS2* and *TBX3* as key regulators of this cell program. Importantly, ectopic expression of these genes in distal mesenchymal cells of transgenic mice led to the activation of genes linked to chiropatagium development. As a result, mutant limbs exhibited significant morphological changes, including increased cell number, enhanced extracellular matrix content, and retention of interdigital tissue. Lastly, the bat *MEIS2* and *TBX3* regulatory landscapes revealed significant differences in enhancer activity and 3D chromatin structure, suggesting species-specific genomic features that mediate their activity during wing formation.

Altogether, our findings unravel fundamental molecular mechanisms underlying bat wing development and demonstrate how drastic morphological changes, such as the emergence of a wing, can arise from the repurposing of existing developmental programs.

2. Zusammenfassung

Die bemerkenswerte morphologische Vielfalt im Tierreich ist das Ergebnis evolutionärer Veränderungen in Entwicklungsprozessen. Die Extremitäten der Tetrapoden stellen ein eindrucksvolles Beispiel dieser Vielfalt dar, da sie tiefgreifende phänotypische Anpassungen durchlaufen haben, um die Besiedelung verschiedenster Lebensräume – an Land, im Wasser und in der Luft – im Laufe der Evolution zu ermöglichen. Daher dienen sie als einzigartiges Modell zur Untersuchung der zellulären Mechanismen, die dieser Diversifizierung zugrunde liegen.

Unter den Säugetieren hat sich der aktive Flug ausschließlich in der Ordnung der Chiroptera, besser bekannt als Fledermäuse, entwickelt. Diese bemerkenswerte Fähigkeit verdanken sie der Struktur ihrer Flügel, die durch stark verlängerte Finger geprägt sind, welche von einer spezialisierten Flughaut, dem Chiropatagium, verbunden werden – ein Merkmal, das nur bei dieser Spezies vorkommt. Die molekularen Ursprünge und das Ausmaß der Entwicklungsveränderungen, die notwendig sind, um diese außergewöhnliche morphologische Anpassung zu ermöglichen, sind jedoch weitgehend unbekannt.

Um dies zu untersuchen, nutzten wir komparative Ansätze der Einzelzell- und funktionellen Genomik, um Gliedmaßen-Entwicklungszellzustände sowohl bei Mäusen als auch bei Fledermäusen zu analysieren und miteinander zu vergleichen. Unsere Analyse offenbarte eine signifikante Konservierung von Zellzuständen und -prozessen zwischen den beiden Spezies, einschließlich der interdigitalen Apoptose - einem Vorgang, der bei Spezies mit getrennten Fingern das interdigitale Gewebe eliminiert. Entgegen früheren Hypothesen findet dieser Prozess auch bei Fledermäusen statt, was auf einen anderen molekularen Ursprung der Flughaut zwischen den Fingern hindeutet. Durch Mikrodissektion des embryonalen Chiropatagiums und anschließende Einzelzell-Transkriptom-Analyse dieses Gewebes entdeckten wir, dass das Chiropatagium aus Fibroblasten besteht, die sich deutlich von den apoptotischen interdigitalen Zellen unterscheidet. Zudem konnten wir zeigen, dass diese Fibroblasten im distalen, autopodialen Bereich der Extremitäten ein Entwicklungsprogramm umfunktionieren, das normalerweise im proximalen Bereich der Extremitäten aktiv ist. Dieses Chiropatagium-Zellprogramm ist gekennzeichnet durch Prozesse, die an Zellproliferation, Migration und der Organisation der extrazellulären Matrix beteiligt sind, was deren zentrale Bedeutung für die Flügelentwicklung unterstreicht.

Funktionelle Genomik- und Gen-Netzwerk-Analysen identifizierten die entwicklungsbedingten Transkriptionsfaktoren *MEIS2* und *TBX3* als Schlüsselfaktoren dieses Zellprogramms. Interessanterweise löste die ektopische Expression dieser Gene in distalen mesenchymalen Zellen transgener Mäuse die Aktivierung von Genen aus, die mit der Entwicklung des Chiropatagiums verbunden sind. Infolgedessen zeigten die Extremitäten der Mutanten

signifikante morphologische Veränderungen, einschließlich einer erhöhten Zellzahl, eines erhöhten Gehalts an extrazellulärer Matrix und der Beibehaltung von interdigitalem Gewebe.

Zuletzt zeigten die regulatorischen Domänen von *MEIS2* und *TBX3* in Fledermäusen signifikante Unterschiede in der Enhancer-Aktivität und der 3D-Chromatinstruktur, was auf artspezifische genomische Merkmale hinweist, die die Aktivität dieser Faktoren während der Flügelentwicklung steuern.

Zusammenfassend enthüllen unsere Erkenntnisse die grundlegenden molekularen Mechanismen der Flügelentwicklung in Fledermäusen und verdeutlichen, wie drastische morphologische Veränderungen, wie die Entstehung eines Flügels, durch die Umnutzung bestehender Entwicklungsprogramme entstehen können.

3. Introduction

Throughout embryogenesis, a complex organism with hundreds of different cell types develops from a single cell. This process involves intricate transitions of cell states, controlled by the precise expression of distinct genes in time and space (Bentovim, Harden, & DePace, 2017). Perturbation of these tightly controlled gene activities can therefore result in various diseases, ranging from congenital malformations to cancer (Anania & Lupiáñez, 2020; Chakraborty & Ay, 2019). However, in an evolutionary context, these modifications are considered to be a major driver of phenotypic innovation leading to the remarkable morphological diversity observed throughout the animal kingdom (H. K. Long, Prescott, & Wysocka, 2016). Although important elements controlling gene expression have been characterized, the alterations they undergo during evolution and how this impacts cell states during development remain elusive. Altogether, this raises the puzzling question, how changes to gene activities and cell programs during embryogenesis drive cell type evolution and ultimately the development of novel traits. Are entirely novel gene activities established and thereby new cell types created? Or are existing cell states and their gene programs modified and repurposed? Here, I address this by investigating how changes to gene programs alter developmental cell fates to drive morphological novelty.

In metazoans, spatiotemporal gene expression is controlled by complex regulatory landscapes, ensuring a sophisticated interplay of regulatory components (Bolt & Duboule, 2020). This is mainly regulated by cis-regulatory elements of the non-coding genome, including gene promoters and their controlling enhancers (Wittkopp & Kalay, 2011). Communication between both is facilitated through precise chromatin folding, guiding their physical contact within restricted regulatory domains (Remeseiro, Hörnblad, & Spitz, 2016; Robson, Ringel, & Mundlos, 2019). The output of such regulatory interactions establishes cell type-specific behaviors within the developing embryo - such as cell proliferation, homoeostasis, apoptosis or differentiation - which collectively drive morphogenesis. Moreover, during evolution, regulatory components and their interactions are substrates to modifications, bearing the potential to induce changes to developmental gene activities and resulting cell states (H. K. Long et al., 2016).

The developing limb is a well-characterized model for studying gene regulation in development and evolution. During limb morphogenesis, precisely regulated expression of genes establishes spatial axes which define limb patterning and ultimately appendage morphology, a process considered highly conserved across limbed species (McQueen & Towers, 2020; Petit, Sears, & Ahituv, 2017). Nonetheless, this organ was extensively modified in evolution which resulted in a myriad of different appendage morphologies, equipping animals for diverse lifestyles (Varga & Varga, 2022; Zhu & Tabin, 2023). The bat wing forelimb represents a

remarkable example of limb adaptation, as its massively elongated digits connected by wing membrane tissue make bats the only mammal capable of powered flight (K. L. Cooper & Tabin, 2008). However, the exact mechanisms driving the evolution of limb developmental cell types leading to the development of novel morphological structures remain elusive.

The following introductory sections will summarize the major players of gene regulation in development, how their interplay translates into cell behaviors and how their modification during evolution can influence developmental cell states.

3.1. Developmental cell states in morphogenesis

Embryogenesis requires the production of hundreds of different specialized cell types from a single cell. To ensure this, cells undergo various transitions to collectively produce a wide cellular variety at precise proportions and patterns within the developing embryo (Belmonte-Mateos & Pujades, 2021; Bentovim et al., 2017). This fundamental principle of unidirectional development, first described in 1957, remains central to our current understanding of embryogenesis to this day (Waddington, 1957). Importantly, understanding developmental cell states and how they drive embryo patterning and morphology requires their systematic characterization. Furthermore, to investigate how these processes are modified in evolution, they must be compared within a suitable model system.

The next sections will introduce the developing limb as model system for cell differentiation, spatial patterning and morphogenesis in development and evolution. They will specifically highlight the main phases of limb development, the central factors orchestrating these processes as well as how their alterations in evolution influence limb morphology.

3.1.1. The developing limb as Evo-Devo model system

During embryogenesis, a tightly regulated interplay of cell proliferation, differentiation or controlled elimination collectively shapes embryo structures and therefore ultimately determines the morphological phenotypes observed in the adult organism. Furthermore, evolutionary modifications to patterning processes during embryogenesis bear the potential to drive morphological diversification (Feregrino & Tschopp, 2022). Therefore, evolutionary and developmental processes are deeply connected, making developmental systems a common model to study the modifications of molecular programs and their impact on cellular and morphological phenotypes (Cardoso-Moreira et al., 2019; Pantalacci & Sémon, 2015; Silbereis, Pochareddy, Zhu, Li, & Sestan, 2016).

The vertebrate limb exemplifies a commonly used model system to study gene regulation, patterning and tissue differentiation in developmental biology (Cohn & Tickle, 1996; Petit et al., 2017). Because this organ is non-essential for survival, embryonic limbs carrying mutations affecting its development can in many cases be studied throughout embryogenesis (Cohn &

Tickle, 1996; Feregrino, Sacher, Parnas, & Tschopp, 2019; Feregrino & Tschopp, 2022). Furthermore, as the limb develops exteriorly, it is easily accessible for molecular and morphological characterization. Therefore, signaling pathways and underlying gene regulatory networks governing initiation of its outgrowth and patterning have been well characterized through genetic and experimental studies, mainly in mouse and chicken model systems. The fundamental limb developmental pathways are furthermore believed to be largely conserved between species (Maier et al., 2017; McQueen & Towers, 2020; Zeller, López-Ríos, & Zuniga, 2009). However, and very interestingly, developmental patterning of the vertebrate limb also experienced various modifications along tetrapod evolution, resulting in a myriad of morphological phenotypes (Petit et al., 2017; Varga & Varga, 2022; Zhu & Tabin, 2023). This makes this organ a well-suited system for studying the mechanisms driving phenotypic innovation in evolutionary developmental ("evo-devo") biology.

Generally, limb development can be categorized in four phases: limb bud initiation, pattern specification, tissue differentiation and limb shaping and finally its growth to adult size (Figure 1) (Cohn & Tickle, 1996). Here, I will outline these fundamental processes of these phases, especially focusing on patterning and shaping of the mouse limb, which is the best characterized to date. Furthermore, I will give an overview on how alterations to those processes can influence limb cell types and morphology and how this potentially contributes to the emergence of novel phenotypic traits during vertebrate evolution.

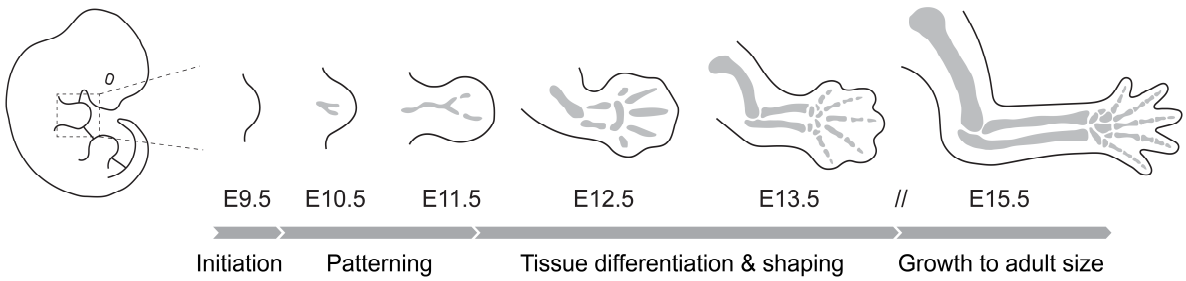

Figure 1: Overview of murine limb development.

Shown are the main phases of limb development - initiation, patterning, tissue differentiation, and growth to adult size - along with corresponding embryonic stages and their characteristic limb morphologies. Mesenchymal condensations and skeletal elements are depicted in grey.

3.1.2. Limb outgrowth initiation and patterning along three spatial axes

In mice, limb buds emerge through proliferation of mesenchymal progenitor cells derived from the lateral plate mesoderm (LPM) at embryonic day 9.5 days post coitum (E9.5, Figure 1). Subsequently, limbs develop along three spatial axes, each controlled by a distinct signaling center. Outgrowth from the LPM is initiated by the limb-specific T-box transcription factors (TFs) TBX5 and TBX4 (Gibson-Brown et al., 1996; Rallis et al., 2003; Rodriguez-Esteban et al., 1999), which activate *fibroblast growth factor (Fgf) 10* expression in the fore- and hindlimb

field, respectively (Agarwal et al., 2003; Ng et al., 2002). Secreted FGF10 diffuses from the mesenchyme to the surrounding ectodermal tissue, where it activates expression of *Fgf8* in an epithelial-mesenchymal feedback loop, which is a crucial driver for mesenchymal cell proliferation and ultimately limb bud outgrowth (Figure 2) (Petit et al., 2017). Furthermore, expression of *Fgf8* promotes the emergence of the apical ectodermal ridge (AER), a cellular epithelial structure located at the most distal part of the limb (Ohuchi et al., 1997; Xu et al., 1998). The AER is the main signaling center of the proximo-distal (PD) axis, running from shoulder to digits. Its FGF signals keep the underlying mesenchymal cells in a proliferative state, thereby promoting limb outgrowth (Sekine et al., 1999). Eventually, the AER will promote the appearance of a second signaling center in the posterior part of the limb, the zone of polarizing activity (ZPA). The ZPA secretes the morphogen Sonic hedgehog (SHH), which provides spatial information for establishing the anterior-posterior (AP) axis, which runs from digit I to V (Riddle, Johnson, Laufer, & Tabin, 1993; Tickle & Towers, 2017). Importantly, SHH secreted from the ZPA and FGFs from the AER engage in a positive regulatory loop to maintain their expression during patterning of the developing limb bud (Figure 2). Finally, dorsal-ventral (DV) patterning from the back of the hand to the palm is established by expression of Wnt-family member *Wnt7a* and *Engrailed 1* (*En1*) in the dorsal and ventral limb ectoderm, respectively (C. Logan, Hornbruch, Campbell, & Lumsden, 1997; Parr & McMahon, 1995). Altogether, the precise spatiotemporal expression of these patterning factors is crucial for proper tissue differentiation by instructing surrounding cell fates to ultimately define limb element identity.

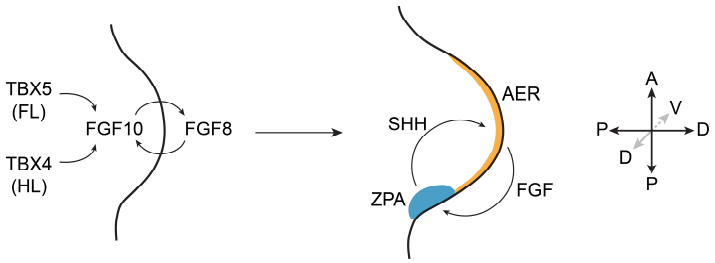

Figure 2: Early limb development and patterning is governed by distinct signaling centers.

Schematic representation of limb initiation and early outgrowth (left) and patterning along PD and AP axes (right). The limb specific transcription factors TBX5 and TBX4 initiate expression of *Fgf10* in the limb mesenchyme, which induces expression of *Fgf8* and emergence of the AER (yellow). Presence of this structure promotes establishment of the ZPA (blue), which secretes SHH to the surrounding mesenchyme. The positive regulatory loop between FGFs from the AER and SHH from the ZPA is essential for proper limb outgrowth.

3.1.3. Limb segment specification along the proximal-distal axis

Outgrowth along the PD axis eventually originates three distinct limb elements. The stylopod is located at the most proximal side of the limb and will give rise the humerus (upper arm) and

femur (thigh) in the fore- and hindlimb, respectively. The zeugopod, found more distally in the central part of the limb, consists of radius and ulna (forearm) in the forelimb and tibia and fibula (lower leg) in the hindlimb. The autopod is the most distal limb segment and comprises carpals, metacarpals, and phalanges (wrist, ankle and the digits) (Figure 3). Establishment of the proximo-distal axis during limb outgrowth and therefore definition of these limb segments is organized by complex gene networks and self-regulatory feedback systems. Based on genetic and experimental studies in mouse and chicken limbs, the key PD patterning factors have been characterized.

An essential signaling network regulating limb outgrowth along the proximo-distal axis is established by FGFs from the AER and SHH in the ZPA. Mesenchymal progenitor cells directly underlying the AER are kept in an undifferentiated state by FGFs, which stimulates their proliferation and promotes distal identity (Lewandoski, Sun, & Martin, 2000; ten Berge, Brugmann, Helms, & Nusse, 2008). Furthermore, SHH signaling from the posterior limb mesenchyme in the ZPA, together with bone morphogenic protein (BMP) 4, upregulate expression of *Gremlin1* (*Grem1*), a BMP antagonist (Zúñiga, Haramis, McMahon, & Zeller, 1999). GREM1-mediated BMP antagonism in the posterior-distal mesenchyme enhances FGF expression in the AER, which maintains expression of *Shh* and mesenchymal cell proliferation (Bénazet et al., 2009; Khokha, Hsu, Brunet, Dionne, & Harland, 2003). This signaling system is eventually terminated by spatial separation of *Shh* and *Grem1* expressing cells, leading to BMP overcoming inhibition by GREM1 (Scherz, Harfe, McMahon, & Tabin, 2004). As BMP signaling in the limb eventually increases, it becomes a major driver of subsequent developmental processes such as chondrogenesis as well as apoptosis of the interdigital mesenchymal cell population (Barna & Niswander, 2007; Bénazet et al., 2012; Gañan, Macias, Duterque-Coquillaud, Ros, & Hurle, 1996) (see chapter 3.1.4).

The patterning of proximal to distal order of limb segments is furthermore promoted by temporal, successive expression of *HoxA* and *HoxD* genes of groups 9-13 (Andrey et al., 2013). While *Hox9* and *Hox10* define the most proximal portion in the stylopod, *Hox11* expression determines the zeugopod, and the most distal autopodial segment is assigned by expression of *Hox12* and *Hox13* (Figure 3) (Zakany & Duboule, 2007). Loss-of-function of these *HoxA* and *HoxD* components have been shown to strongly impact limb morphology of their respective segments, indicating that proper limb morphogenesis strongly depends on *Hox* gene function (Davis, Witte, Hsieh-Li, Potter, & Capecchi, 1995; Dollé, Dierich, et al., 1993; Fromental-Ramain et al., 1996; Wellik & Capecchi, 2003).

Lastly, proximal-distal identity is determined by two opposing gradients of FGFs secreted from the AER and retinoic acid (RA) secreted from the embryo flank, which jointly pattern developing limb elements. The opposing FGF-RA signals are translated into a gradient of the transcription factors MEIS1 and MEIS2, with high levels of MEIS1/2 defining proximal, and gradually

decreasing levels determining more distal identity (Figure 3). MEIS1 and MEIS2 are homeobox transcription factors belonging to the TALE (three-amino-acid loop extension) superclass of transcription factors. Both proteins are highly similar and have been reported to have redundant, yet crucial functions during limb development (Delgado et al., 2021). Specifically, the important role of MEIS1/2 in limb PD patterning has been shown by several studies reporting disturbed skeletal segment formation upon their absence. Depletion of MEIS1/2 during limb development has been reported to result in the complete loss or massive reduction of proximal limb elements, while digit development was unaffected (Delgado et al., 2021; Delgado et al., 2020). Complete knock-out of both, *Meis1* and *Meis2*, results in elimination of limb bud growth altogether (Delgado et al., 2021). Furthermore, ectopic expression of *Meis* factors in growing limbs disrupts proper development of distal elements in the zeugopod and autopod and leads to distal-to-proximal transformations (Capdevila, Tsukui, Rodríquez Esteban, Zappavigna, & Izpisúa Belmonte, 1999; Mercader et al., 1999; Mercader et al., 2009). Altogether, this strongly emphasizes the essential role of MEIS transcription factors in defining limb outgrowth and proximal and distal segment identity.

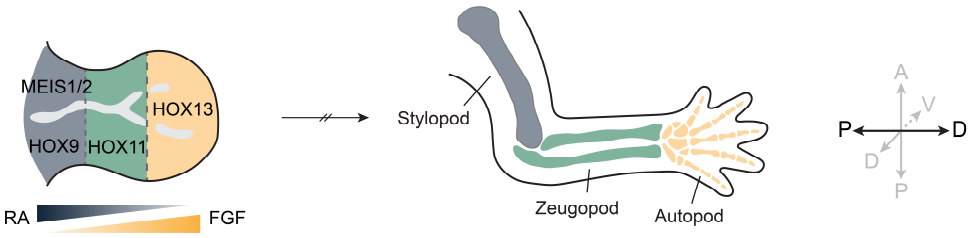

Figure 3: Proximo-distal patterning results in three distinct limb segments.

Schematic representation of proximo-distal patterning factors (left) and resulting limb element identities (right). During limb outgrowth along the PD axis, the interplay of several patterning components instructs proximal or distal identities of limb segments. These include the establishment of a RA-FGF gradient along the PD axis, as well as expression of *Meis1/2* and *HoxA/D* transcription factors, with *Meis1/2* and *HoxA/D9* defining proximal stylopodial (dark blue), *HoxA/D11* zeugopodial (green) and *HoxA/D13* distal autopodial (yellow) limb segments.

3.1.4. Shaping the limb through cell differentiation and controlled cell death

As limb development progresses, a complex heterogeneous mixture of numerous different limb cell states is generated. While skeletal muscle progenitors migrate into the limb bud from the somitic mesoderm (Chevallier, Kieny, & Mauger, 1977; Christ, Jacob, & Jacob, 1977), LPM-derived progenitor cells of the limb bud differentiate into tendons, connective tissue and smooth muscle populations (M. Logan et al., 2002; Pearse, Scherz, Campbell, & Tabin, 2007; Wachtler, Christ, & Jacob, 1981). Furthermore, the condensation of mesenchymal cells initiates the development of the skeletal elements of stylopod, zeugopod and autopod (Pearse et al., 2007). Importantly, in species with free digits, the formation of autopodial chondrogenic condensates

occurs simultaneously with the removal of cells in the interdigital mesenchyme mediated by controlled cell death, leading to digit separation (Figure 4). Therefore, proper spatiotemporal control of this process is a crucial mechanism for shaping the developing limb into its adult form.

Interdigital tissue removal is driven by various cell death processes, such as caspase-dependent apoptosis, lysosomal-mediated cell death, cellular senescence, and necrosis (Chautan, Chazal, Cecconi, Gruss, & Golstein, 1999; Montero & Hurlé, 2010; Stewart et al., 2000; Zuzarte-Luis, Berciano, Lafarga, & Hurlé, 2006), indicating the complexity of the machinery at play. Among these, apoptotic processes are the most well-documented form of cell death in the interdigital tissue. For one, dying cells in this tissue have been shown to exhibit the cellular and morphological pattern of apoptosis, characterized by features such as caspase activity, chromatin condensation and cell fragmentation, among others (Garcia-Martinez et al., 1993; Hernández-Martínez & Covarrubias, 2011; Mori et al., 1995). Moreover, several secreted factors have been identified as key signals involved in regulating apoptotic cell fates within the interdigital tissue.

Studies in mouse and chick limb proposed BMPs as the main drivers directly triggering the apoptotic signal in the interdigital mesenchyme (Kaltcheva, Anderson, Harfe, & Lewandoski, 2016; Zuzarte-Luis & Hurle, 2005; Zuzarte-Luís & Hurlé, 2002). Indeed, AER-specific inactivation of BMPs results in prolonged FGF signals from the AER, leading to failed interdigital tissue regression (K. S. Choi, Lee, Maatouk, & Harfe, 2012; Maatouk, Choi, Bouldin, & Harfe, 2009). However, application of FGFs to or their genetic overexpression, for instance, also results in syndactyly through loss of interdigital cell death (Hernández-Martínez, Castro-Obregón, & Covarrubias, 2009; Lu, Minowada, & Martin, 2006; Macias, Gañan, Ros, & Hurle, 1996; Ngo-Muller & Muneoka, 2000). Therefore, interdigital cell death seems to require a concerted action of both, the removal of FGFs promoting cell survival as well as induction of cell death by BMPs signaling (Figure 4) (Kaltcheva et al., 2016).

Another crucial apoptotic signaling component is retinoic acid (RA, Figure 4). Its role in inducing interdigital cell death was shown in several studies reporting accelerated interdigital tissue removal upon application of RA to mouse limbs (Hernández-Martínez et al., 2009; Lussier, Canoun, Ma, Sank, & Shuler, 1993; Rodriguez-Leon et al., 1999). Accordingly, application of a RA antagonist results in inhibition of apoptosis and retention of interdigital tissue (Hernández-Martínez et al., 2009; Rodriguez-Leon et al., 1999). Furthermore, mouse mutants lacking ALDH1A2, the enzyme catalyzing RA synthesis, or its receptors exhibit syndactyly as a consequence of failing tissue regression (Ghyselinck et al., 1997; Kastner et al., 1997; Zhao, Brade, Cunningham, & Duester, 2010).

Interestingly, besides cell death, other mechanisms have been reported as essential for digit separation. A recent study reported that active migration of epidermal cells into the interdigital

mesenchyme is required for digit separation (Kashgari et al., 2020). The authors observed that, as the interdigital mesenchyme regresses, the epidermis invaginates from the dorsal and ventral side, forming a multi-layered structure of epithelial tissue, which eventually converges and separates (Kashgari et al., 2020). This process therefore relies on cell proliferation and migration guided by chemotactic cues in the extracellular matrix.

Taken together, the simultaneous processes of digit formation and interdigital tissue removal represent a tightly regulated, complex interplay of cell differentiation, migration and controlled elimination which together define adult limb morphology.

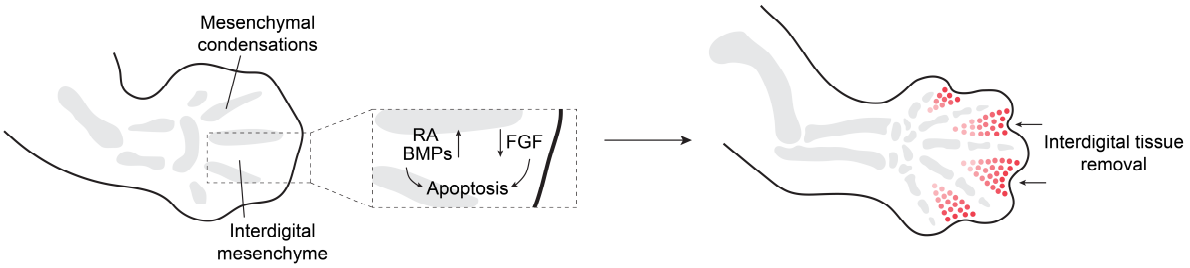

Figure 4: Limb morphology is shaped through interdigital tissue removal via apoptosis.

Schematic representation of limb shaping through digit formation and interdigital tissue removal. When mesenchymal condensations form, BMP and RA levels in the interdigital mesenchyme increase while FGF levels from the AER decline (left). This collectively induces apoptosis in the interdigital regions, resulting in tissue removal in a distal-to-proximal manner (right).

3.1.5. Appendage diversification in evolution

The previous sections have summarized the key components orchestrating limb bud initiation and outgrowth and how the cell states and spatial patterns defined during this process ultimately determine adult limb morphology. While the previously described signaling pathways and morphogenic components are largely conserved in tetrapods, modifications to these processes can influence the phenotypic outcome of limb morphology. The following sections will therefore outline examples of such alterations in appendage evolution and comment on their potential molecular drivers.

Through evolutionary variations in limb element number or proportion, tetrapods have evolved a fascinating diversity in appendage morphologies, equipping them for life in diverse habitats and lifestyles. Interestingly, limb evolution has been studied in an increasing number of non-model organisms, suggesting mechanisms for how limb patterning is modified during early and advanced stages of development (Petit et al., 2017). On a molecular and cellular level, this is achieved through modification of signaling centers, resulting in altered limb patterning and therefore variation in element length, number or shape. Basically, these modifications can be classified into the ones truncating limb developmental trajectories, as well as the ones

enhancing or expanding them, resulting in reduced or increased limb features, respectively (Zhu & Tabin, 2023).

One example of the former is the reduction in the number of limb elements. A reduced number of digits, for instance, evolved repeatedly in the evolution of tetrapods, in jerboas, horses, pigs, lizards, cattle and camels, among others (K. L. Cooper et al., 2014; Lopez-Rios et al., 2014; Shapiro, Hanken, & Rosenthal, 2003). Different mechanisms affecting distinct stages of limb development have been described to cause these phenotypes. Digit number reduction in pigs, cattle and lizards, for instance, has been reported to occur through disturbance of SHH signaling, thereby influencing early limb patterning (K. L. Cooper et al., 2014; Lopez-Rios et al., 2014; Sears et al., 2011; Shapiro et al., 2003; Tissières et al., 2020). On the other hand, digit loss in jerboas, horses and camels was shown to be a result of increased and spatially expanded apoptosis surrounding the remaining digits, thereby influencing limb shape in post-patterning stages (K. L. Cooper et al., 2014).

A more extreme example of element reduction if represented by the loss of whole limb structures, exemplified by hindlimb loss in cetaceans or the loss of both fore- and hindlimbs in snakes. In dolphins, for instance, hindlimb buds initially develop and form an *Fgf8* expressing AER. However, this structure is not maintained. Moreover, *Shh*, and therefore the ZPA, is absent in hindlimb structures. Therefore, outgrowth of the hindlimb bud is disturbed due to lacking key signaling centers and it degenerates early during limb development (Thewissen et al., 2006). Snakes have lost both their fore- and hindlimbs. Interestingly, similar to dolphins, limb loss in these animals has also been linked to a failure of AER and ZPA maintenance and a resulting failure in limb outgrowth (Cohn & Tickle, 1999; Leal & Cohn, 2016). A more recent study furthermore reported snake-specific genomic sequence changes within a *Shh* limb-specific enhancer, the ZPA regulatory sequence (ZRS). This enhancer and its activity are highly conserved, however its function was progressively lost throughout snake evolution (Kvon et al., 2016).

In contrast to element loss, appendages have also evolved through their expansion. The jerboa, for instance, represents an example of elongated limb skeletal elements, as they develop extremely elongated metatarsals in their feet. This has been reported to be a result of species-specific expression of genes, which have previously not been associated with skeletal elongation during development. One major transcription regulator identified in this study is SHOX2. This gene is usually involved in elongation of proximal limb elements in other vertebrates but gained jerboa-specific expression in their autopodial metatarsals (Saxena et al., 2022). Lastly, the expansion of limb structures can be a result of decreased elimination of cells during development thereby retaining otherwise removed limb tissues. This has, for instance, been reported to be the case in the webbed hindlimbs of ducks. Here, expression of BMP-antagonist *Grem1* inhibits BMP-induced interdigital apoptosis, resulting in digit fusion by

soft tissue (Merino et al., 1999; Weatherbee, Behringer, Rasweiler, & Niswander, 2006; Yokouchi et al., 1996).

3.1.6. Bat wing development

One remarkable example of appendage adaptation that has fascinated researchers for decades is found in the order Chiroptera, commonly known as bats. They are the only mammals having evolved powered flight, which is facilitated by a highly specialized forelimb morphology (Figure 5). This includes massively elongated skeletal elements, especially digits II-V in the autopod, and the presence of several wing membranes (propatagium, chiropatagium, plagiopatagium and uropatagium) (K. L. Cooper & Tabin, 2008; Gunnell & Simmons, 2005; Simmons, Seymour, Habersetzer, & Gunnell, 2008). In contrast, their hindlimb elements are comparatively short with completely separated digits (K. L. Cooper & Tabin, 2008). The morphological adaptations arising in bat forelimbs can be observed early on during limb development. While limb morphology is indiscernible between mice and bats during early stages of limb initiation, at Carnegie stage (CS) 15, equivalent to E11.5 in mice, the bat forelimb bud loses its symmetry and expands on the posterior side (Cretekos et al., 2005). Subsequently, skeletal condensations form, and digits II-V begin to elongate. Furthermore, while the interdigital tissue between digits I and II is removed, it is retained between digits II-V, leading to the formation of the chiropatagium (Cretekos et al., 2005) (Figure 5).

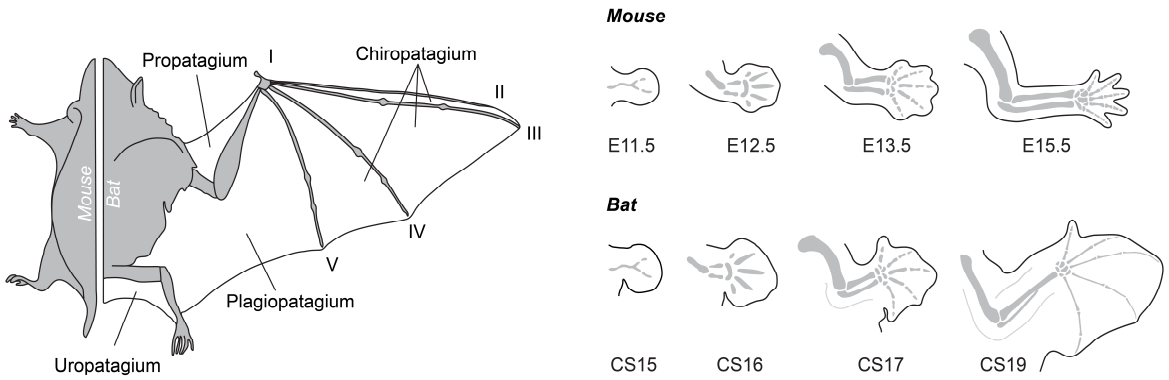

Figure 5: Comparison of bat and mouse forelimb morphology and development.

Schematic representation of limb morphologies in adult mice and bats (left) and during limb development (right). Bat wings are characterized by elongated skeletal elements and the presence of several wing membranes (propatagium, plagiopatagium, chiropatagium and uropatagium). These phenotypic features are most prominent in the forelimb autopod with elongation of digits II-V connected by the chiropatagium, while digit I is short and free. Bat hindlimbs exhibit a similar morphology as mouse hindlimbs, with short and separated digits. Forelimb morphological differences are visible early during development (E11.5/CS15) and become more pronounced with digit elongation and interdigital tissue retention (E12.5/CS16-E13.5/CS17). This figure was adapted from (Cretekos, Rasweiler, & Behringer, 2001).

The specific wing forelimb adaptation of chiropterans is believed to have allowed them to colonize unique ecological niches, thereby enabling their great evolutionary success and diversification into over 1,400 species, comprising roughly 20% of all extant mammalian

species (Burgin, Colella, Kahn, & Upham, 2018; Gunnell & Simmons, 2005; Sadier et al., 2021). Consequently, bats offer a fascinating model for studying the connections between developmental processes and functional morphology, as well as for uncovering the mechanisms that drive evolution.

The earliest known bat fossil from approximately 52.5 million years ago shows remarkable morphological similarity to extant bats. Importantly, they already exhibit the defining features of a wing forelimb, suggesting that the ability to fly emerged only once at the root of this taxonomical group (Gunnell & Simmons, 2005; Simmons et al., 2008). Furthermore, analysis of fossil and extant bat limbs suggest that the relative length of their forelimb elements was not significantly altered throughout this time (Sears, Behringer, Rasweiler, & Niswander, 2006). Therefore, the absence of fossils documenting their transition from a terrestrial to an aerial life has significantly hindered our understanding of bat evolution. To solve this problem, evo-devo studies comparing mouse and bat limbs have started to investigate the underlying developmental mechanisms driving this transition (Sadier et al., 2021).

Initially, those studies were limited to candidate gene approaches, focusing on known limb developmental genes and patterning factors. Investigation of early developmental stages during limb bud outgrowth, for instance, showed enlarged AER and ZPA signaling centers in bat forelimbs. Furthermore, these studies proposed a novel expression domain of *FGF8*, which reactivates *SHH* at a later developmental stage when it is inactive in mouse limbs (L. N. Cooper, Cretekos, & Sears, 2012; Cretekos, Deng, Green, Rasweiler, & Behringer, 2007; Hockman et al., 2008). Both, the prolonged and enlarged expression domains of these genes were suggested to contribute to cell survival and proliferation and ultimately limb elongation (Hockman et al., 2008). Moreover, it was suggested that the continued presence of *FGF8* leads to inhibition of BMP signaling through *GREM1* activation, resulting in decreased levels or inhibition of apoptosis and persistence of interdigital tissue (L. Cooper, Sears, Adams, & Pedersen, 2013; Hockman et al., 2008; Weatherbee et al., 2006). Other studies investigating skeletal element elongation reported elevated *BMP2* and *PRX1* expression levels in the perichondrium of bat forelimb digits, suggesting that this promotes cartilage proliferation and therefore digit elongation (Cretekos et al., 2008; Sears et al., 2006). Furthermore, on a gene regulatory level, replacement of a mouse *PRX1* enhancer with the orthologous bat sequence resulted in increased *PRX1* gene expression and elongated forelimb bones, thereby suggesting regulatory modification as source of evolutionary variation in bats (Cretekos et al., 2008).

More recent evo-devo studies utilized genome-wide approaches to characterize bat limb development through more systematic analyses. Transcriptomic characterizations of embryonic bat limbs, for instance, have highlighted gene expression differences as well as unique expression profiles in bat wings. Generally, this revealed specific expression of

developmental transcription factors, especially members of the t-box and homeobox family, among others (Dai et al., 2014; Mason et al., 2015; Z. Wang et al., 2014; Z. Wang et al., 2010). Also, the expression of genes involved in BMP and FGF signaling pathways has been shown to be differentially enriched in bat wings (Eckalbar et al., 2016; Z. Wang et al., 2014; Z. Wang et al., 2010). Furthermore, genome and epigenetic comparisons suggest potential regulatory elements specific to the bat lineage. One study, for instance, investigated gene regulatory element variation and identified bat accelerated regions (BARs), elements which evolved at increased rates in bats compared to other species (Booker et al., 2016). Furthermore, functional ChIP-seq data for active and repressive histone marks was produced from developing bat limbs at several developmental stages and intersected with this data (Eckalbar et al., 2016). A subset of the identified bat specific regions was found to be located in the vicinity of key limb developmental genes (Booker et al., 2016; Eckalbar et al., 2016). Although these studies hypothesized a role of the described regions in bat forelimb evolution, their functional characterization remains to be investigated.

3.1.7. Challenges and limitations of Evo-Devo studies

The number of evo-devo studies on non-model organisms, including bats, is constantly increasing. Although these approaches have highlighted species-specific gene expression as well as regulatory differences and proposed mechanisms involved in phenotypic evolution, many limitations and challenges remain.

Firstly, these studies typically investigate and describe isolated examples, focusing on single genes or loci. Gene regulation, the precise spatiotemporal transcriptional output and the translation into cell behaviors are ultimately interconnected features of developmental systems. Therefore, to understand their modification during evolution and how this impacts embryogenesis, these features must be studied collectively. However, due to the complexity of large datasets generated through genome-wide approaches, integration of these complex modalities can be challenging. Moreover, the lacking availability of rare biological material of non-model organisms can be a limiting factor for generating such datasets. Secondly, the morphological adaptations that have evolved in the bat limb, for instance, result from a combination of several phenotypic innovations and involve a myriad of complex cellular interactions. Consequently, investigating them in whole tissue bulk assays creates a diluted readout averaging the transcriptomic and epigenetic information of many individual cell types. Therefore, the technical resolution of many approaches limits the molecular mapping of such sophisticated developmental systems. Lastly, functional validations directly linking the observed findings with the phenotypic consequences are rarely undertaken in this kind of studies, therefore preventing any causality. Because of this, the complexity of morphological innovation and the technical limitations of analytical approaches, linking identified gene

regulatory differences to novel biological functions at cell type and tissue level remains a major challenge in evo-devo research.

3.2. Profiling gene activities and cell behaviors in development and evolution

Proper spatiotemporal gene expression is the foundation of embryogenesis. During this process, cells are precisely instructed to produce a distinct differentiation output to shape embryonic tissues and structures. Importantly, the modification of spatiotemporal expression patterns can alter such cell states and result in novel cellular or morphological phenotypes during evolution. Therefore, the ultimate goal – as well as the challenge – of developmental and evolutionary studies is, to link altered cell behaviors to the development of novel morphological traits.

The following sections will introduce a molecular toolbox to collectively study the complex interplay of gene regulatory features and how they orchestrate cell behaviors during development. Specifically, they will outline the latest advancements on studying developmental cell states and their dynamics during embryogenesis, highlighting the relevance of such approaches for mapping their alterations. Furthermore, they will introduce the major components of gene regulatory landscapes, how their interplay is regulated in the nucleus to generate precise gene activity outputs and how this can be profiled to study developmental regulatory networks.

3.2.1. Comparing transcriptomes and cell behaviors

Complex gene regulatory interactions collectively drive distinct gene activities, which define a cell's specific function. On a larger scale, spatiotemporal patterns of such activities within the developing embryo drive cell transitions and ultimately tissue differentiation. Furthermore, modifications to these patterns during evolution has been shown to result in altered cellular and morphological outputs. Indeed, it has been reported that differences in development between species correlate with changes in gene expression (Cardoso-Moreira et al., 2019; Romero, Ruvinsky, & Gilad, 2012). Consequently, comparing gene expression between tissues and species is an important aspect to consider when investigating the molecular evolution of altered cell states driving novel phenotypic traits.

During the last decades, the development of comparative transcriptomic approaches, especially RNA sequencing (RNA-seq), massively advanced evo-devo studies, as it allows for the detection of gene expression in a genome-wide manner (Bainbridge et al., 2006). The generated data can be used to compare gene expression levels between tissues, developmental stages or species, including non-model organisms, to identify tissue- or

species-specific gene regulatory programs (Pantalacci & Sémon, 2015; Roux, Rosikiewicz, & Robinson-Rechavi, 2015).

Interestingly, transcriptome-based comparative evo-devo studies have outlined fundamental characteristics of species evolution. As a general trend, organ developmental transcriptomes seem to be most similar between species during early development, coinciding with their similarity in morphology (Abzhanov, 2013; Cardoso-Moreira et al., 2019; Kalinka & Tomancak, 2012). This mainly involves genes with relevant functions during embryogenesis, such as developmental transcription factors or morphogens (Davidson & Erwin, 2010; Morov, Ukizintambara, Sabirov, & Yasui, 2016). As development progresses and morphological differences between species increase, transcription level become increasingly more distinct as well (Abzhanov, 2013; Cardoso-Moreira et al., 2019; Kalinka & Tomancak, 2012). This strongly suggests genes active during early development presumably function in multiple organs and stages, making them more resistant to alteration (Cardoso-Moreira et al., 2019; Garfield & Wray, 2009). As organs differentiate, active genes become more restricted in their temporal and spatial expression. This potentially reduces functional constraints, thereby enabling evolutionary changes in their expression patterns (Cardoso-Moreira et al., 2019; Garfield & Wray, 2009).

Developmental transcription programs ultimately result in specialized cell states. Therefore, they not only represent a fundamental level of driving the development of a complex, multicellular organism, but also an important link between molecular and phenotypic evolution (Tanay & Sebé-Pedrós, 2021). However, cell states are highly heterogeneous at individual levels and are constantly remodeled during embryogenesis (Jiang, Xu, & Guo, 2021; Zeng, 2022). Therefore, although bulk RNA-sequencing has been an indispensable tool in evo-devo research, its major limitation when investigating tissue cell types, their dynamics and their modification during evolution, is the resolution of the transcriptomic readout.

A relatively recent technical innovation, single-cell RNA sequencing (scRNA-seq), allows the profiling of the transcriptome of single cells from a given sample individually (Figure 6) (Macosko et al., 2015; Zheng et al., 2017). This enables the characterization of molecular dynamics within complex tissues, organs or organisms at cellular resolution, preserving their heterogeneity (Y. H. Choi & Kim, 2019). Therefore, scRNA-seq has become the state-of-the-art approach for investigating the compositions and trajectories of developmental cell populations during pattern formation and organogenesis (Medina-Jiménez, Budd, & Janssen, 2024; Tominaga, Nishitsuji, & Satoh, 2023). Importantly, in an evo-devo context, this technology facilitates the comparison of cell states as evolutionary units as well as their specification programs both within and between tissues and species. Recent studies have applied scRNA-seq in various tissues and model as well as non-model species (Han et al., 2022; Jiang, Xiao, et al., 2021; Jones et al., 2022; Mayshar et al., 2023; Musser et al., 2021;

Vergara et al., 2021; F. Wang et al., 2022). This generated high-resolution cell atlases composed of millions of cells that enable comparative studies of cell programs. Interestingly, these approaches reliably detected both, homologous and conserved populations as well as highly diverse species-specific expression patterns and differentiation trajectories, thereby revealing various states of evolution on cellular level (Bakken et al., 2021; Mayshar et al., 2023; Shafer, Sawh, & Schier, 2022; Woych et al., 2022).

Importantly, recent single-cell transcriptomic studies have provided high-resolution reference datasets of developing limbs of several species, including mouse, chicken and human (Bastide et al., 2022; Cao et al., 2019; Desanlis, Paul, & Kmita, 2020; Feregrino et al., 2019; He et al., 2020; Kelly, Huynh, & Guilak, 2020; Markman et al., 2023; B. Zhang et al., 2023). As previously mentioned, the developing limb has long served as a model system for developmental biology (Cohn & Tickle, 1996) and the key players and fundamental mechanisms orchestrating limb development have been well characterized at tissue level. The recent development of single-cell technologies, however, enables studying the processes of pattern formation and tissue differentiation during embryogenesis at cellular level (Figure 6). These studies showed that major limb developmental cell populations - such as lateral plate mesoderm-derived, ectodermal or muscle cells - can confidently be identified and annotated based on previously characterized marker gene co-expression within them. However, deeper analysis of this data furthermore enables investigation of cell differentiation dynamics and trajectories, the identification of previously unknown cell populations and their markers, as well as extraction of gene programs of specific cell populations.

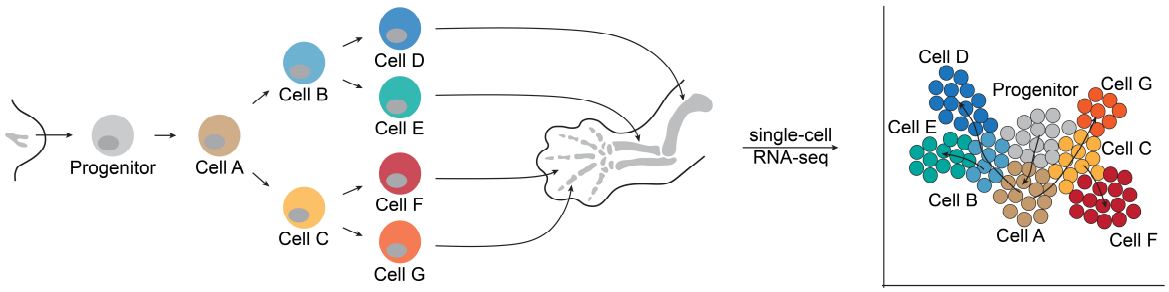

Figure 6: Single-cell RNA-seq enables studying heterogeneous tissues and developmental cell states.

Schematic representation limb tissue differentiation from a less differentiated structure into a heterogeneous organ composed of a myriad of different cell states. During this process, progenitor cells (grey) undergo various transitions to produce a wide cellular variety (left). By applying scRNA-seq individual cells of complex tissues can be analyzed and their differentiation trajectories and dynamics can be mapped to study developmental cell states (right).

Single-cell RNA-sequencing technologies provide a powerful resource for studying the molecular processes of embryogenesis. Moreover, these technologies can be used to investigate how modifications to developmental programs impact cell states, their patterns and their dynamics during embryogenesis. This reveals important insights into the evolution of cell

populations, thereby emphasizing the great potential of single-cell technologies in evolutionary comparative biology. Altogether, single-cell technologies have not only greatly expanded our knowledge of embryonic development but also paved the way towards better understanding the molecular origin of phenotypic diversity across the animal kingdom.

3.2.2. The relevance of CREs in spatiotemporal gene regulation

As highlighted in the previous paragraph, our ability to profile the transcriptomes of individual cells during development has greatly advanced our knowledge of the cellular processes driving embryogenesis. Importantly, these processes rely on precise control of spatiotemporal gene expression through the coordination of different regulatory elements and their interactions, which is organized at several chromatin scales in the nucleus. Therefore, their modification is considered a potent source for altering expression patterns within the embryo during evolution and thereby creating novel cell outputs.

In metazoans, spatiotemporal expression of genes during embryogenesis has been shown to be tightly regulated by different cis-regulatory elements (CREs), which are sequences of the non-coding genome controlling gene transcription. Different types of CREs have been identified during the last decades, including promoters, enhancers, silencers, and insulators (Jindal & Farley, 2021; Y. Li, Chen, Kaye, & Wasserman, 2015). Among these, promoters and enhancers are the most thoroughly characterized and will therefore be described in more detail.

Promoters are found directly upstream of a gene body and - because of their predictable location – these elements are relatively well defined. They generally contain two components, the proximal and core promoter element. The latter is usually located within 50 base pairs (bp) of the transcription start site (TSS) (S. J. Cooper, Trinklein, Anton, Nguyen, & Myers, 2006). Here, general transcription factors assemble to form the pre-initiation complex, which includes the recruitment of RNA polymerase II, so that transcription along the gene body can begin (Figure 7) (Haberle & Stark, 2018). The proximal promoter is typically found around 250 bp upstream of the transcription start site (TSS) and contains binding sites for sequence-specific TFs (Haberle & Stark, 2018). Together, these elements facilitate initiation of gene transcription and are therefore indispensable regulatory elements for gene expression.

However, even though promoters are sufficient to drive initiation of transcription, they generally exhibit a low basal activity (Haberle & Stark, 2018). Therefore, additional elements are needed in metazoan gene regulation to govern gene activities at different levels, times and locations (Petit et al., 2017). The main drivers of this specificity are CREs called enhancers, which bind gene promoters to instruct their specific spatiotemporal activation (Figure 7) (Shlyueva, Stampfel, & Stark, 2014). In contrast to promoters, enhancers are less well defined. Their sequences highly vary in size (typically 0.1 - 1 kilobase pair) and can be found widely

distributed throughout regulatory landscapes in inter- as well as intra-genic regions (Field & Adelman, 2020; Kvon, 2015). In addition, they can activate promoters independently of their orientation or distance (Ong & Corces, 2012). Enhancer function requires the binding of multiple specific transcription factors (Buecker & Wysocka, 2012). Therefore, each enhancer sequence contains clusters of transcription factor binding sites, the distinct combination and organization of which determines its cell type-specific and spatiotemporal activity (Figure 7) (Spitz & Furlong, 2012). Here, factors such as binding affinity as well as number, spacing and orientation of motifs can functionally influence the activity and robustness of enhancer elements (H. K. Long et al., 2016). In this way, their distinct enhancer grammar represents a major aspect in regulating complex gene activities during the development of numerous different cell types in embryogenesis from the same limited set of genes (Heinz, Romanoski, Benner, & Glass, 2015; Jindal & Farley, 2021; Zabidi & Stark, 2016).

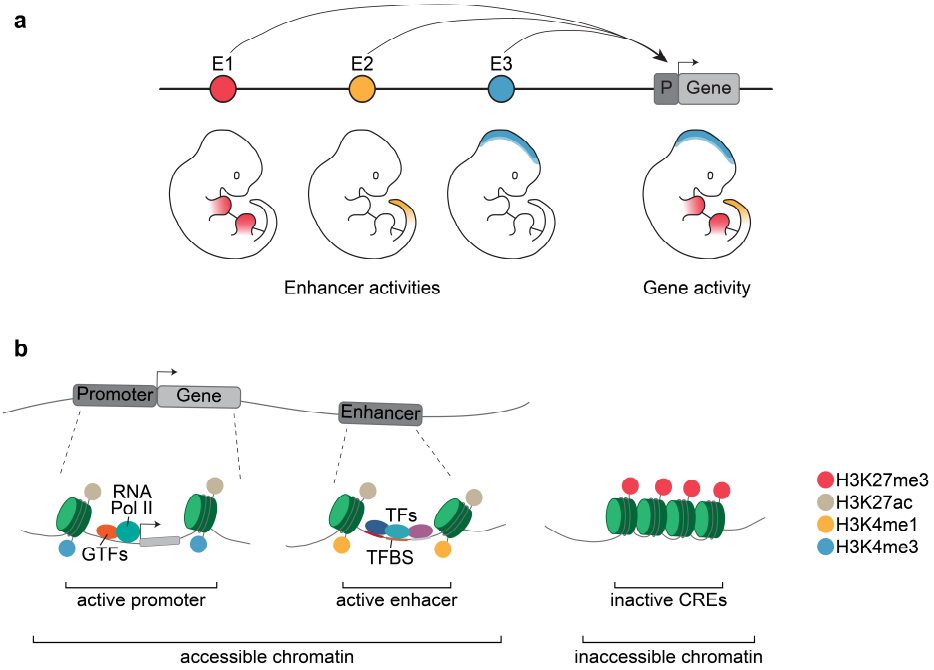

Figure 7: CRE activities collectively drive tissue-specific gene activity and can be identified by their distinct chromatin modifications.

a Shown are three enhancers (E1-3) with distinct tissue-specific activities in embryonic limb buds (E1), the tail tip (E2) or the brain (E3), instructing a gene promoter to drive gene expression. Their concerted action results in the diverse spatiotemporal activity of one gene in various tissues. This figure was adapted from (Robson et al., 2019). **b** Schematic representation of promoter and enhancer elements and their distinct features. Promoters are bound by general TFs (GTFs), recruiting RNA polymerase II to initiate transcription. Enhancers are composed of TF binding sites (TFBS) enabling binding of specific TFs. CREs can be identified by distinct chromatin states, including chromatin accessibility and histone modifications. Active CREs are found in highly accessible chromatin regions and are typically enriched in H3K27ac. Promoters and enhancers can furthermore be distinguished by the methylation state of lysine 4 at histone 3 (H3K4me3 or H3K4me1, respectively). H3K27me3 is considered a repressive histone modification, indicating an inactive chromatin state.

3.2.3. Identifying CREs

As outlined above, the non-coding genome contains important functional elements regulating gene expression during development. Therefore, identifying these features is crucial to study their activity and function. Generally, CREs can be identified by various characteristics, among those their distinctive epigenetic state of chromatin modifications and accessibility (Figure 7).

In the nucleus, deoxyribonucleic acid (DNA) is wrapped around histone proteins, which forms eu- or heterochromatic regions of high or low accessibility, representing an active or repressive state, respectively (Mansisidor & Risca, 2022). Furthermore, histone proteins carry different post-translational modifications, such as acetylation or methylation, which indicate a certain functional state (Bannister & Kouzarides, 2011; Y. Li et al., 2015). Typically, active CREs are found in highly accessible, open chromatin regions enriched in acetylated lysine 27 of histone H3 (H3K27ac). Tri-methylation of the same residue (H3K27me3) on the other hand indicates their repression by the polycomb machinery (Calo & Wysocka, 2013; Tsompana & Buck, 2014). Promoter and enhancer regions can furthermore be distinguished by different levels of methylation of lysine 4 on histone H3. While promoters show high levels of tri-methylated lysine 4 (H3K4me3), enhancer elements are enriched in mono-methylated lysine 4 (H3K4me1) (Calo & Wysocka, 2013). Lastly, active regulatory elements are bound by a distinct mixture of general or specific transcription factors, which furthermore functionally characterizes promoter and enhancer elements, respectively (Palstra & Grosveld, 2012).

Different genomics approaches have been developed in the last years to profile and map DNA-protein interactions as well as epigenetic states of CREs. Chromatin immunoprecipitation DNA-sequencing (ChIP-seq), for example, identifies genomic regions surrounding histones carrying a certain epigenetic modification or DNA stretches bound by a protein of interest. Furthermore, accessibility of chromatin can be detected by Assay for Transposase-Accessible Chromatin (ATAC) sequencing. Importantly, although these genomic techniques can identify putative CREs, functional assays are crucial for validating their actual activities. This is especially important for characterizing the diverse and complex activities of enhancer elements. Typically, enhancer activity is assessed by its ability to drive transcription of a reporter gene. In this method, a candidate enhancer is cloned adjacent to a reporter gene under the control of a minimal promoter (Kvon, 2015). The resulting enhancer-reporter constructs can then be introduced into the genome and its spatiotemporal activity can be visualized through the reporter gene's expression *in vivo* (Kvon, 2015; Visel, Bristow, & Pennacchio, 2007). Therefore, by collectively mapping epigenetic features of protein-DNA interactions and chromatin accessibility, CREs can be identified and characterized to study their specific activities in various developmental cell types and tissues.

Interestingly, profiling of CREs revealed that the non-coding genome is full of regulatory elements and therefore its vast majority is involved in gene regulation. Mammalian genomes,

for instance, have been shown to comprise more that 100,000 of such elements, by far outnumbering the approximately 20,000 coding genes (Lupiáñez, Spielmann, & Mundlos, 2016; Pennacchio, Bickmore, Dean, Nobrega, & Bejerano, 2013). This, together with their ability to activate promoters over large genomic distances, allows genes to be controlled by multiple enhancers with different activities (H. K. Long et al., 2016). One the one hand, the resulting combinatorial complexity is thought to account for the great pleiotropy of genes observed in higher organisms, allowing for precise cell type- and tissue-specific gene expression (de Laat & Duboule, 2013). On the other hand, gene promoters being controlled by multiple distal enhancers raises the question of how this accurate communication is realized. The following sections will therefore introduce the levels of chromatin organization of regulatory landscapes and how this influences enhancer-promoter interaction and ultimately gene regulation.

3.2.4. Regulatory landscapes are shaped by the 3D genome

The physical interplay between enhancers and promoters is a crucial aspect to successfully drive precise spatiotemporal gene expression during complex processes like embryogenesis (Palstra et al., 2003). Yet, regulatory elements can be separated by large distances in the genome. In metazoans, distal enhancers can lie tens to thousands of kilobases, or even megabases, from their target promoters (Mora, Sandve, Gabrielsen, & Eskeland, 2016; Schoenfelder & Fraser, 2019; Williamson, Hill, & Bickmore, 2011). To overcome these large genomic distances, chromatin is highly organized on several levels in the nucleus to ensure the required physical proximity of enhancer elements and their promoter targets (Robson et al., 2019). Therefore, studying chromatin architecture is fundamental to understand its function in gene regulation during complex developmental cell transitions.

The development of chromosome conformation capture (3C)-based techniques has significantly increased the ability to map chromatin architecture (Dekker, Rippe, Dekker, & Kleckner, 2002). Here, chromatin interactions of one (3C), multiple (4C and 5C) or all (Hi-C) genomic loci are quantified based on proximity ligation in the nucleus (Dekker et al., 2002; Dostie et al., 2006; Lieberman-Aiden et al., 2009; Simonis et al., 2006). The data obtained with these techniques, especially through the genome-wide mapping of chromatin interactions through Hi-C, has revealed that mammalian genomes are non-randomly and hierarchically organized in the 3D nuclear space (Dekker et al., 2002; Lieberman-Aiden et al., 2009).

For one, distal enhancers have been shown to contact their target promoters across large genomic distances through chromatin looping (Figure 8) (Schoenfelder & Fraser, 2019). In vertebrates, these contacts are restricted within self-interacting regulatory landscapes of submegabase size, so-called topologically associated domains (TADs, Figure 8) (Dixon et al., 2012; Nora et al., 2012). At larger chromosome scale, TADs with similar epigenetic features

are spatially segregated in multi-megabase interactions forming A and B compartments of active or inactive chromatin, respectively (Lieberman-Aiden et al., 2009; Schoenfelder & Fraser, 2019). Finally, in nuclear space, active A compartments are localized in the nuclear interior, where they are permissive to transcription. In contrast, transcriptionally inactive B compartments are positioned at the nuclear periphery (Hoskins, Smith, & Reddy, 2021; Rao et al., 2014).

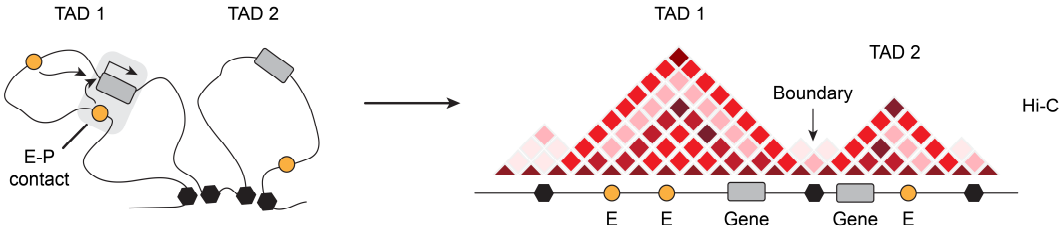

Figure 8: Enhancer-Promoter contacts are restricted by topologically associated domains.

Enhancers instruct gene promoters by physical contact, potentially throughout large genomic distances (left). These enhancer-promoter contacts preferentially occur within, but not between, topologically associated domains (TADs). These self-interacting domains are separated by boundary elements, creating structural blocks providing functional gene regulatory units. Chromatin interactions within TADs can be visualized by 3C-derived methods, such as Hi-C (right). In these heatmaps, the intensity of each pixel represents the interaction frequency between the respective ends of the diagonal. This figure was adapted from (Robson et al., 2019).

Enhancer-promoter communication is ultimately influenced by each of the above-mentioned structural features which together shape regulatory landscapes. In the context of developmental gene expression, this overview will mainly focus on TADs as the genomic units containing genes and their regulatory elements.

3.2.5. TADs act as functional units in developmental gene regulation

TADs are highly self-interacting chromatin structures of 100 kb to several Mb in size conserved across vertebrates (Dixon, Gorkin, & Ren, 2016; Dixon et al., 2012; D. Li et al., 2022). These structural units are separated by insulating boundary elements enriched in CCCTC-binding factor (CTCF) and structural maintenance of chromosome (SMC)-complex cohesin (Fudenberg et al., 2016; Rao et al., 2014). The physical separation of neighboring domains restricts enhancer-promoter contacts by promoting intra-domain interactions while impeding those across boundary regions, thereby contributing to their specificity (Figure 8) (Fudenberg et al., 2016; Furlong & Levine, 2018; Symmons et al., 2014).

The view of TADs as genomic scaffolds guiding and isolating enhancer-promoter contacts is debated, and it cannot be excluded that they merely are a representation of the contacts forming within them (Lupiáñez et al., 2016; Nora et al., 2017; Schwarzer et al., 2017). However, TAD structures were observed to be highly conserved between species, tissues and cell types, thereby connoting their biological relevance (Dixon et al., 2012; Fraser et al., 2015; Harmston et al., 2017). Moreover, their role in gene regulation during embryogenesis has been

addressed by several studies. For instance, genes singly occupying a TAD were shown to be especially associated with developmental processes (H. S. Long et al., 2022). Consistent with this, regulatory elements of developmental genes were reported to strongly overlap with identified TAD structures (Symmons et al., 2014). Importantly, the relevance of TADs in regulating enhancer-promoter communication has furthermore and especially been highlighted by studies investigating the effect of their disruptions. This revealed that structural variations such as deletions, inversions or duplications affecting boundary elements can result in ectopic interactions between regulatory elements of neighboring domains leading to misexpression of genes within (Figure 9) (Lupiáñez et al., 2015). This phenomenon known as enhancer adoption or hijacking can drive a wide set of diseases, ranging from congenital malformations to cancer (Akdemir et al., 2020; Anania & Lupiáñez, 2020; Lettice et al., 2011). Taken together, this suggest that TADs represent genomic units providing an insulated environment for gene regulation during complex processes such as development.

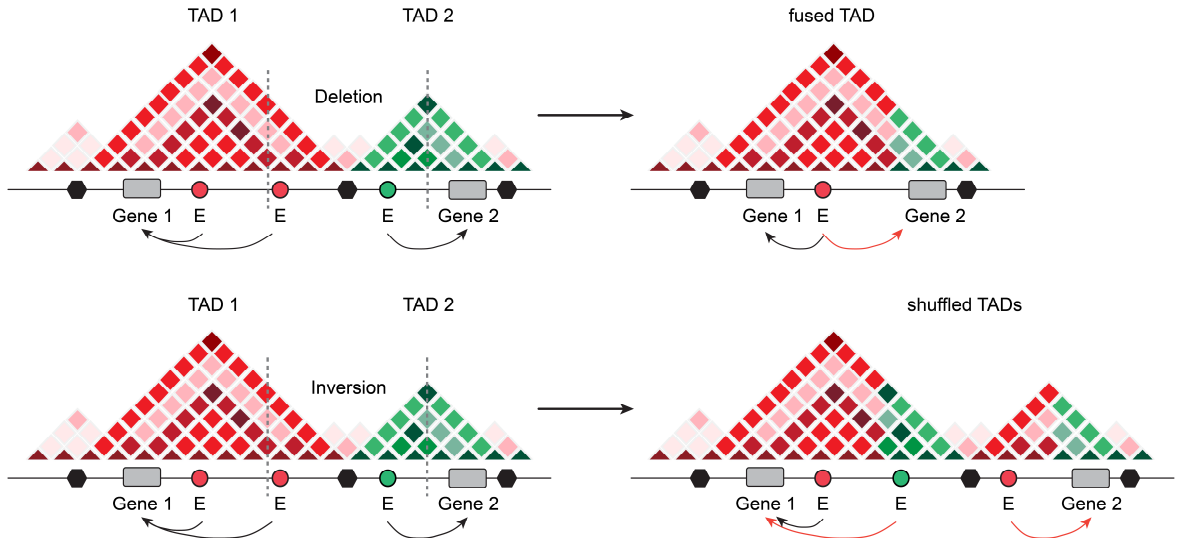

Figure 9: Structural variations can result in misexpression of genes through TAD rearrangements and enhancer adoption.

Schematic representation of Hi-C maps and the corresponding genomic loci with genes, their enhancers and boundary elements. Structural variations such as a genomic deletion (upper panel) or inversion (lower panel) affecting TAD boundary elements result in rearranged domains (fused TAD or shuffled TAD) and ectopic enhancer-promoter contacts (red arrows). This phenomenon, known as enhancer adoption, can result in misexpression of genes within affected TADs.

The previous sections have highlighted the importance of regulatory elements and their 3D interactions for precise gene regulation during embryogenesis and how their disruptions can result in gene misexpression and disease. However, in an evolutionary context, modifications to the different layers of gene regulation are thought to be a major driver of phenotypic innovation. Therefore, the next chapters will outline how gene regulatory variation can be achieved in the context of evolution.

3.2.6. CREs and 3D landscapes as subjects to evolution

The previous chapters have outlined, how CREs and their 3D interplay with their target genes represent a central aspect in orchestrating spatiotemporally precise gene expression programs during embryogenesis. Interestingly, in nearly all metazoans, multiple core developmental processes are shared, resulting in commonalities in body plans, tissues and structures (Maeso, Irimia, Tena, Casares, & Gómez-Skarmeta, 2013). Those include, for instance, the establishment of body axes or the formation of germ layers during gastrulation (Anlas & Trivedi, 2021; Castanon-Ortega & Heisenberg, 2005).

Advances in studying the basis of metazoan evolution through whole genome sequencing revealed, that a fundamental set of genes is shared by all animals, indicating that basic developmental processes are instructed by conserved gene networks (Maeso et al., 2013). The large number of deeply conserved spatial expression patterns during embryogenesis across many phyla furthermore suggests, that not only developmental genes themselves, but also the regulatory elements controlling their activities can be extremely stable throughout evolution (Peter & Davidson, 2011). Indeed, a high level of sequence as well as functional conservation can be found in developmental promoters and enhancers, even across large evolutionary distances (Phan et al., 2024; Visel et al., 2007). Furthermore, functionally homologous enhancers often share a conserved genomic location (Cande, Goltsev, & Levine, 2009; Wittkopp & Kalay, 2011). It is therefore believed that those conserved developmental programs represent a crucial aspect of establishing common structures and cell types.

At the same time, changes in spatiotemporal gene expression by modulation of CREs, especially enhancers, and their activities during development are believed to be a potent source of evolutionary change (Carroll, 2008). Indeed, enhancers were reported to exhibit a high rate of turnover and variability between species (Arnold et al., 2014; Villar et al., 2015). Furthermore, because of their modular organization and spatiotemporal specificity, changes in enhancers allow for the modulation of gene activities in specific cell types, without affecting others (H. K. Long et al., 2016; Wray, 2007). Therefore, these elements are considered to be major drivers of cis-regulatory variation (Wittkopp & Kalay, 2011).

Several mechanisms have been proposed on how enhancer activity can emerge during evolution. For one, new enhancers can be generated *de novo* from a non-regulatory DNA sequence (Figure 10). This has been reported for mammalian species-specific enhancer elements, suggesting that rapid enhancer evolution is a fundamental aspect of mammalian genomes (Villar et al., 2015). Also, transposable elements are considered to be a prominent source for regulatory novelty. After being introduced into the host genome, their regulatory information can acquire mutations yielding new enhancer activities (Bourque et al., 2008; Feschotte, 2008). On the other hand, enhancer activities can be modified from pre-existing ancestral regulatory elements through genomic mutations (Figure 10). Point mutations in the

enhancer sequence, for instance, can alter transcription factor binding as well as result in targeting of different promoters (Rebeiz & Tsiantis, 2017). Moreover, local deletions or duplications affecting regulatory elements can have dosage effects by reducing or increasing enhancer activity, respectively (H. K. Long et al., 2016).

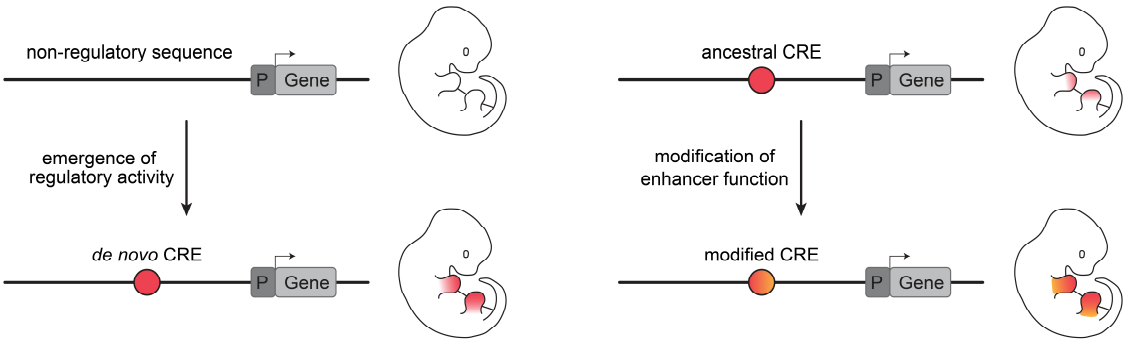

Figure 10: Emergence of novel or modified enhancer activity.

Schematic representation of novel generation (left) or modification (right) of enhancer activity throughout evolution. *De novo* elements can emerge from non-regulatory DNA or existing, ancestral elements can be modified to generate novel or altered activity.

On a larger scale, structural rearrangements affecting TAD boundaries can also influence enhancer activity by changing the regulatory landscapes and target promoters they are exposed to. It generally is believed that TAD restructuring during evolution is rare and structural rearrangements rather occur at boundaries keeping highly syntenic domains intact (Farré, Robinson, & Ruiz-Herrera, 2015; Harmston et al., 2017; Vietri Rudan et al., 2015). Nevertheless, cases have been reported where species-specific TAD rearrangements caused enhancer-promoter rewiring resulting in distinct expression patterns of the affected developmental genes (Marlétaz et al., 2023; Real et al., 2020).

Taken together, the *de novo* generation of enhancers, the modification of their activity or their 3D interactions can be considered potent sources of modulating gene regulation in evolution and therefore driving phenotypic innovation (Erwin, 2020). Therefore, mapping regulatory elements, profiling gene expression and characterizing developmental cell states provide powerful approaches for collectively investigating gene regulatory programs during embryogenesis. Furthermore, they facilitate the framework for comparing these programs between species to enhance our understanding of how alterations to developmental gene regulation can drive the emergence of morphological novelty.

4. Aim of this study

During embryogenesis, the tightly regulated interplay of cell proliferation, differentiation or controlled elimination shapes embryo structures and therefore determines the morphologies in the adult organism. Modifications of these developmental processes are thus considered to be a major driver of phenotypic innovation. However, it is unclear how changes to gene activities during embryogenesis drive cell type evolution and ultimately the development of novel traits. Are new gene programs generated to create entirely novel cell types? Are developmental cell fates altered to remove or retain structures? Or are existing cell states and their gene programs modified to repurpose their function?

Here, I use the developing limb and its modifications throughout evolution as a model to address this central question. Specifically, I investigate the molecular origin of a unique mammalian trait, the interdigital wing membrane in bat forelimbs. By systematically comparing limb cell types between mice and bats I explore, how developmental cell states are modified in evolution to drive such dramatic morphological changes.

5. Materials

5.1. Chemicals and Reagents

All chemicals used in this study were obtained from Sigma-Aldrich, Roth or Merck in analytical grade quality unless stated differently.

5.2. Instruments and Software

A list of instruments and software used in this thesis can be found in section 10.5 (Table 8 and Table 9).

5.3. Plasmids and bacterial strains

For cloning experiments, 5-alpha Competent *E. coli* cells (NEB, Cat. #C2987) and Top10 *E.coli* cells were used.

A list of recombinant DNA and plasmids used in this study can be found in section 10.4 (Table 6).

5.4. Primers

Primers were synthesized by IDT Eurofins. Primer sequences used in this study can be found in section 10.4 (Table 2, Table 3, Table 4 and Table 5).

6. Methods

All wet-lab experiments in this work were performed by myself, except for light sheet and confocal laser-scanning microscopy of embryonic limbs which was performed by Rose Yinghan Behncke and René Hägerling, and Juliane Glaser, respectively.

Computational analyses were performed as follows: Single-cell RNA-seq analyses by Christian Feregrino, bulk RNA-seq analyses by Anna A. Monaco, ChIP-seq analyses by Bai-Wei Lo, pySCENIC TF network analysis by Alexander Barclay and enhancer prediction and conservation analysis by Tobias Zehnder. In all cases I was closely involved in defining the goals of each analysis and the resulting exchange to interpret results and design next steps.

6.1. Molecular biological methods

Molecular biological methods such as polymerase chain reaction (PCR), agarose gel electrophoresis, DNA cloning or transformation of competent *E.coli* bacterial cells were conducted according to standard procedures from Green and Sambrook Green (2012).

6.2. Isolation of DNA and RNA

6.2.1. Isolation of plasmid DNA

Plasmid DNA was isolated using NucleoSpin Plasmid EasyPure (Macherey-Nagel, #740727) or Nucleobond Xtra Midi EF kit (Macherey-Nagel, #740420) according to the manufacturer's instructions. Plasmid DNA quantity and quality were assessed by NanoDrop.

6.2.2. Isolation of genomic DNA

For isolation of genomic DNA from 96 well plates, cells were washed ones with DPBS (Gibco, #14190144) and lysed by adding 50 µL lysis buffer (10 mM Tris-HCl pH 7.3, 10 mM EDTA pH 8.0, 10 mM NaCl, 0.3 % SDS, 0.02 µg/µL Proteinase K) to each well and incubated at 56°C overnight. DNA was extracted using the KingFisher Flex System with Magnetic beads (MagAttract Suspension G, QUIAGEN #1026901). 8 µL of magnetic beads were pipetted into each well of a KingFisher 96 well plate and the 50 µL cell lysate as well as 35 µL of 85% ethanol were added. Additionally, two plates containing 100 µL of 85 % Ethanol for washing steps and a third plate containing 35 µL of H_2O for the final elution were prepared. The extraction was performed according to the manufacturer's instructions.

DNA for copy number analysis by qPCR and DNA from mouse and bat tissues was extracted using the DNeasy Blood & Tissue Kit (QUIAGEN, #69504) according to the manufacturer's instructions.

6.2.3. RNA Isolation

Mouse and bat tissues for RNA isolation were dissected in cold PBS and immediately snap-frozen and stored at −80°C until further processing. Total RNA was extracted using the RNeasy Micro or Mini Kit (QUIAGEN, #74004 and #74104) according to the manufacturer's instructions. In short, tissue samples derived from E11.5, E12.5 and E13.5 as well as CS15 and CS17 embryos were directly homogenized in RTL lysis buffer supplemented with β-Mercaptoethanol using a syringe. Limb tissues from older embryonic stages (E15.5 and CS19) were pulverized in liquid nitrogen using a Bel-Art SP Scienceware Liquid Nitrogen-Cooled Mini Mortar (Fisher Scientific) prior to homogenization. Genomic DNA was removed by digestion using the RNase-free DNase Set (QUIAGEN, #79254). Total RNA was eluted in an appropriate volume of EB buffer and RNA purity and concentration were measured by NanoDrop.

6.3. Animal Models

6.3.1. Mice

Wildtype mouse embryonic tissues were derived from crossings of CD1 x CD1 or C57BL/6J x 129. Transgenic embryos were generated by tetraploid aggregation (Artus & Hadjantonakis,

2011) by the in-house Transgenics Core Unit of the Max-Planck-Institute for Molecular Genetics. Female mice of CD1 strain were used as foster mothers.

Mice were kept in a controlled environment (12 h light and 12 h dark cycle, temperature of 20–22.2 °C, humidity of 30–50%) in the in-house animal facility of the Max-Planck-Institute for Molecular Genetics. Food, water and bedding were changed routinely.

All animal experiments and procedures were conducted as approved by LAGeSo Berlin under the following license numbers: ZHV120, G0176/19-MaS1_Z, G0243/18-AldS1_G and G0098/23-AldS1_G.

6.3.2. Bats (*Carollia perspicillata*)

Bat samples (*Carollia perspicillata)* were obtained from a captive population maintained at the Papiliorama Zoo in Kerzers, Switzerland, in collaboration with Nicolas Fasel. To control population growth, a controlled number of individuals was occasionally euthanized by cervical dislocation performed by trained personnel, following general guidelines for animal handling and in vivo research (Kilkenny, Browne, Cuthill, Emerson, & Altman, 2010). Euthanized females were inspected for pregnancy, in which case embryos were dissected and preserved for molecular studies. Females in late stages of gestation were not euthanized for ethical reasons.

Additional bat samples from *Carollia perspicillata* were obtained from a breeding colony housed at the Institute for Cell Biology and Neuroscience, at the Goethe University in Frankfurt am Main under a permit authorized by the RP Darmstadt in collaboration with Julio Hechavarria. Samples collected from this colony originated from bats that were euthanized for collecting brain tissue without any further experimental manipulation (following § 4 Abs. 3 of the German TierSchG). Female bats, after euthanizing, were additionally inspected for possible pregnancies (undetectable from the outside) and embryos were dissected as above if present.

Bat embryonic stages were determined based on crown-rump length and morphological features based on a previously described embryonic staging system (Cretekos et al., 2005).

6.4. *Carollia perspicillata* genome and annotation

The chromosome-scale *Carollia perspicillata* genome assembly and annotation were generated and kindly provided by Michael Hiller and Ariadna Morales at the Centre for Translational Biodiversity Genomics, Senckenberg Society for Nature Research & Goethe University, in Frankfurt am Main, Germany (michael.hiller@senckenberg.de; ariadna.morales@senckenberg.de). The genome annotation was further refined by Igor Ulitsky at Weizmann Institute of Science, Rehovot, Israel (igor.ulitsky@weizmann.ac.il) and Stefan Haas at the Max Planck Institute for Molecular Genetics (haas@molgen.mpg.de) to

correct mislabeling of orthologous genes and artificial grouping of individual coding sequences and to add long non-coding transcripts.

6.5. Single-cell RNA-sequencing

6.5.1. Single cell isolation and methanol preservation

For single-cell gene expression analysis, mouse embryonic limbs at E11.5, E12.5 and E13.5 and bat embryonic limbs at CS15 and CS17 as well as chiropatagium specimens were dissected and limb cells were dissociated with trypsin at 37°C for 10 minutes. Trypsinization was stopped by adding 1 volume of 5% BSA in 1x DPBS, the cell suspension was filtered through a 40 µm Flowmi Cell Strainer (Merck, #BAH136800040) and centrifuged at 300 x g at 4°C for 5 minutes. The cell pellet was resuspended in 100 µL 0.04 % BSA in 1x DPBS and 900 µL of 100 % methanol were added dropwise to de-hydrate the cells. After incubation on ice for 15 minutes, samples were stored at -80 °C until library preparation.

All single-cell RNA-seq experiments were performed in biological duplicates.

6.5.2. Re-hydration of cells

De-hydrated cells were equilibrated to 4 °C and centrifuged at 1000 x g at 4 °C for 10 minutes. The cell pellet was washed twice in 1 mL of re-hydration buffer (1 % BSA, 0.4 U/µL Ambion RNse Inhibitor (Invitrogen, #AM1682) and 0.2 U/µL SUPERaseIn RNase Inhibitor (Invitrogen, #AM2696) in 1x DPBS). After the second wash, cells were resuspended in an appropriate volume of re-hydration buffer, counted using a Countess cell counter and cell suspension was diluted to a final concentration of 1000 cells/µL. The target cell recovery aimed for in each experiment was 10,000 cells.

6.5.3. 10X Genomics single-cell RNA-sequencing library preparation

Single cell gene expression libraries were prepared using the 10X Genomics Chromium Next GEM Single Cell 3’ GEM, Library & Gel Bead Kit v3.1 (10X Genomics, #PN-1000121) according to the manufacturer’s instructions. In short, Gel Beads-in-emulsion (GEMs) were generated by loading a Master Mix containing reverse transcription (RT) reaction mix and 16,500 cells of the prepared cell suspension, Single Cell 3’ v3.1 Gel Beads and Partitioning Oil into designated wells of a Chromium Next GEM Chip G (10X Genomics, #PN-1000121). The chip was run on a Chromium Controller X/iX immediately after loading. The generated GEMs were carefully aspirated, transferred into a tube strip and incubated in a Bio-Rad C1000 Touch thermal cycler. Inside the individual GEMs, cells were lysed and primers containing an Illumina read 1 sequencing primer, a 16 nt 10X barcode, a 12 nt unique molecular identifier (UMI) and a 30 nt poly(dT) sequence were released and incubated with the cell lysate and RT reaction mix, producing full-length, poly-A-selected, barcoded cDNA. After incubation, GEMs were

dissolved by adding Recovery Agent and cDNAs were purified using Silane magnetic beads. The cDNAs were amplified by PCR, purified using AMPure XP Reagent (Beckmann, #A63881) and concentration was measured using Qubit dsDNA HS Assay Kit (Invitrogen, #Q33231). To optimize cDNA amplicon size, cDNAs were fragmented enzymatically and purified by double-sided size selection using AMPure XP beads (Beckmann, #A63881). Subsequently, ends were repaired, A-tailed and adaptors containing an Illumina read 2 primer sequence were ligated to the fragments. Finally, single or dual sample indices were added to the cDNAs via PCR using the Single Index Kit T Set A (10X Genomics, #PN-1000213) or Dual Index Kit TT Seat A (10X Genomics, #PN-1000215). Library concentration was measured by Qubit dsDNA HS Assay Kit (Invitrogen, #Q33231) and quality was assessed using Bioanalyzer High Sensitivity DNA Analysis Kit (Agilent, #5067-4626). Finally, scRNA-seq libraries were sequenced as asynchronous 28 bp / 90 bp paired-end reads using the Illumina NovaSeq 6000 system with 300 million reads per sample. Read 1 contained the 16 bp 10X barcode and the 12 bp UMI, while the cDNA fragment was encoded in read 2. Library sequencing was performed by the in-house Sequencing Core Facility of the Max Planck Institute for Molecular Genetics.

6.6. Single-cell transcriptomic analysis

Analysis of single-cell RNA-seq data was performed by Christian Feregrino at the Max-Planck-Institute for Molecular Genetics (feregrino@molgen.mpg.de) in close exchange with myself to define the goals of the analysis, interpret results and design next steps.

6.6.1. Filtering, normalization and data integration

Single-cell RNA-sequencing libraries were processed with 10X Genomics CellRanger v6.0.2 (Zheng et al., 2017), using custom genome annotations for *C. perspicillata* and *M. musculus*. Quality filtering of the individual count matrices was based on relative UMI counts, excluding those above 4 times the mean or below 0.2 times the sample median. Additional filters removed cells with ribosomal UMI percentages greater than the median plus 3 MAD or lower than the median minus 3 MAD, and cells with a UMI count/genes detected ratio below 0.15 and UMI counts less than two-thirds the sample median. The filtered datasets were integrated by species and limb (e.g., all mouse forelimb samples were combined). Seurat v4.3.0 (Hao et al., 2024) was used for integration, with datasets first Log-normalized at a scale factor of 10,000 and each cell's cell cycle state scored. SCTransform was applied to regress out ribosomal UMI percentage, UMI count, and S and G2M cell cycle scores. The top 25% of integration features were used to find integration anchors, with SCT normalization, the first 20 dimensions, and a k.filter of 100. These anchors were then applied with the IntegrateData tool.

6.6.2. Dimensionality reduction and clustering

Using the integrated data from the most variable features (stand. Variance > median + MAD of the dat), we performed a PCA and selected the first 20 principal components (18 for the Chiropatagium samples) for further analysis. To maintain and visualize the global structure of the data, we generated t-distributed stochastic neighbor embedding (tSNE) plots using FFT-accelerated Interpolation-based t-SNE (Linderman, Rachh, Hoskins, Steinerberger, & Kluger, 2019) following the approach outlined by Kobak et al. (Kobak & Berens, 2019). Clusters were identified using the FindNeighbors and FindClusters functions, with a resolution of 0.7 and a random seed of 42. Marker genes for each cluster were determined from the un-integrated expression data, applying a minimum expression percentage of 0.25 and a log-fold change (lfc) threshold of 0.5. Cells identified as LPM-derived were isolated, and the same method was used to create the individual datasets presented. A cluster from the MmHl dataset with a high proportion of hemoglobin UMIs was excluded from the analysis.

6.6.3. Generation of an inter-species atlas

The integration of all bat and mouse limb LPM-derived cells was conducted using a similar approach. We first isolated the LPM cells from each dataset, separated the individual libraries, and applied the same workflow. For this integration, the top 12.5% of integration features were utilized. Clusters were identified using a resolution of 0.6. The identity labels from this integration were then applied to the individually clustered datasets. Each cluster was initially classified into one of the three primary groups of LPM cells based on a simple majority, and then assigned the identity of the most frequently represented label within that group.

6.6.4. Apoptosis-related gene expression comparison

Cells from the integrated cluster 3 RA-Id were compared to the rest of the cells within each species and limb dataset using the FindMarkers function. This comparison employed a minimum expression percentage and a log-fold change (lfc) threshold of 0.0001, considering all genes related to apoptosis, the marker genes specific to the cluster, and 20 randomly selected marker genes from other clusters.

6.6.5. Label transfer

To transfer annotation labels from the bat forelimb LPM-derived mesenchyme dataset to the Chiropatagium LPM-derived cells, we identified transfer anchors using the first 20 principal components (PCs) and the SCT normalized data. The TransferData function was then used to apply these labels to the Chiropatagium LPM-derived cells.

6.6.6. Pseudotime analysis

We calculated RNA velocities for our single-cell libraries using velocyto (La Manno et al., 2018), applying the stricter mode for bat samples. From each individual LPM-derived dataset, we selected the *Hoxd13*-positive cells (those with >0 UMIs). After re-clustering and annotating the dataset, we exported it to an AnnData format and integrated the RNA velocity for further analysis. Using scvelo v0.3.2 (Bergen, Lange, Peidli, Wolf, & Theis, 2020) we filtered the data and calculated the first and second order moments using the first 20 PCs. We then applied the dynamical model to compute RNA velocity, allowing for differential kinetics. Guided by the observed RNA velocity trajectories and cluster identities, we further refined the dataset to exclude as much of the chondrogenic lineage as possible. We then generated diffusion maps using the first 15 PCs, selecting diffusion eigenvectors 1 and 2. With the slingshot package, we inferred differentiation trajectories and pseudotime values for each cell based on the Seurat-calculated clusters and the apparent start and end clusters identified by RNA velocity. We then recalculated RNA velocity without the dynamical model. Finally, using CellRank v2.0.4 (Lange et al., 2022), we computed velocity, connectivity, and pseudotime kernels, combining them in proportions of 0.2, 0.4, and 0.4, respectively.

6.6.7. Proximo-distal computational dissection

Each cell of species and limb datasets got assigned a proximal or distal score using Seurat. The AddModuleScore function was employed to score each cell based on sets of genes: distal ("*MSX1*", "*HOXD13*"), proximal ("*SHOX2*"), chondrogenic ("*SOX9*", "*COL2A1*"), and fibroblastic ("*DCN*", "*ZFHX3*"). The same method was applied to other gene sets as well. For each cluster, we calculated the mean difference between the proximal and distal scores. Clusters with the top one-third most extreme score differences were classified as very distal or very proximal. Genes were categorized as typical proximal or distal if, in both species, they were expressed in at least 20% of the cells in 20% of the corresponding clusters and exhibited a difference of >0.15 lfc compared to the opposite cells. Genes highly expressed in the chondrogenic lineage were excluded. Co-expression of genes was determined when $UMIs_x > 0$ and $UMIs_y > 0$.

6.6.8. Proximal/Distal fibroblast cell program expression

To understand the expression profiles defining proximal and distal fibroblasts without focusing on their differences, distal fibroblasts (10 FbI1) were compared to the rest of the cells, excluding proximal fibroblasts (specifically, cells from 8 FbA and 9 FbL labeled as 9 FbL in the main integration), and vice versa. The fibroblast program 2 was analyzed similarly, comparing cells from cluster 7 FbIr to cells from 8 FbA, labeled as 8 FbA in the main integration. This analysis was conducted in two stages: first using highly variable genes, and then including all genes identified as differentially expressed (DE) in both comparisons. GO Terms enrichment

analyses were performed using clusterProfiler (Wu et al., 2021), with all genes expressed in at least 9 cells within the sample serving as the background universe. The function simplify was used with a cut-off of 0.6 on adjusted p-values.

6.6.9. Differential gene expression analysis in mutant limbs

The single-cell datasets from mutant mice were filtered for LPM-derived cells expressing *Hoxd13* and integrated with corresponding cells from the wild-type mouse forelimb. Each newly identified cluster was annotated on the wild-type mouse forelimb dataset using the same logic as before. Differential gene expression between wild-type and mutant cells was tested on a cluster-wise basis using MAST (Finak et al., 2015). This analysis focused on highly variable genes, all genes from our distal/proximal fibroblast program, and excluded genes located on the X chromosome, mitochondrial genes, and ribosomal genes. Only genes expressed in at least 15% of cells from either genotype were tested using MAST, applying a zlm model with the formula “genotype + orig.ident + percent.rp.” An lrTest was then conducted on the genotype coefficient, followed by p-value adjustment using p.adjust. A hypergeometric test was used to evaluate the overrepresentation of DE genes (p.val < 0.01, lfc > 0.25) within our program. We calculated the mean expression of genes in our distal/proximal fibroblast program within cluster 10 FbI1 of the mutant forelimb data and then determined the Pearson correlation of each cell to this mean using the same genes. This analysis focused on clusters where we observed overexpression of *MEIS2* and *TBX3* (p.val < 0.01, lfc > 0.15) in the mutants, as well as the corresponding clusters in the mouse forelimb and bat forelimb datasets.

6.7. Microscopy analysis of apoptosis

6.7.1. Lysotracker Assay

Bat embryonic limbs were dissected in cold DPBS and transferred immediately into 5 µM LysoTracker Red DND-99 (Invitrogen, #12090146) in DPBS. Samples were incubated at 37 °C for 45 minutes, washed four times in DPBS and fixed overnight at 4°C in 4% PFA/DPBS. The following day, samples were washed with PBS for 10 minutes, dehydrated through increasing methanol concentrations (25%, 50%, 75%, 100% (v/v) methanol in PBS) and stored at -20°C until imaging.

6.7.2. Cleaved-Caspase 3 immunofluorescence assay

For immunofluorescence staining of cleaved caspase 3, samples were fixed in 4% PFA/DPBS for 2 hours at 4 °C. Samples were then washed three times and stored in DPBS at 4 °C until immunofluorescent staining was performed. For staining, samples were washed twice in DPBS for 5 minutes and subsequently incubated 3 times for 1 hour in 0.5% Triton-X-PBS (PBST) for permeabilization. After that, samples were incubated in blocking solution (5% FCS/PBST)

overnight at 4 °C. Anti Cleaved Caspase-3 (D175) Antibody (rabbit polyclonal, Cell Signaling Technology, #9661, Lot 47) was diluted in blocking solution (1:400) and incubated for 72 hours at 4 °C. Samples were washed three times with blocking solution and three times with PBST and subsequently incubated in blocking solution overnight at 4 °C. Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ 568 (Invitrogen, #A10042, Lot 2306809) and DAPI were diluted in blocking solution (1:1000) and incubated for 48 hours at 4°C. Samples were washed three times with blocking solution, three times with PBST, three times with DPBS, and post-fixed in 4% PFA/DPBS for 20 minutes.

6.7.3. Microscopy imaging

Sample imaging and image processing were performed together with Juliane Glaser (jglaser@molgen.mpg.de) at Max Planck Institute for Molecular Genetics.

Prior to imaging, samples from both experiments were washed three times with 0.02M phosphate buffer (pH 7.4) and cleared in RIMS (Refractive Index Matching Solution, 13% Histodenz (Sigma-Aldrich D2158) in 0.02M PB) at 4 °C for at least one day. Whole-mount limbs were then imaged with a Zeiss LSM880 confocal laser-scanning microscope in fast-Airyscan mode. Images were processed with the ZEN software (Airyscan processing and Max intensity projection of the Z-stacks) and with Fiji (Schindelin et al., 2012).

6.8. Gene regulatory network analysis using pySCENIC

Gene network analysis was performed by Alexander Barclay at Structural Genomics Group, Centre Nacional d'Anàlisis Genòmic (CNAG-CRG) in Barcelona, Spain (alexander.barclay@cnag.eu) in close exchange with myself to define the goals of the analysis, interpret results and design next steps.

Gene regulatory networks were built using pySCENIC, the Python version of SCENIC (Van de Sande et al., 2020). Raw counts and cell-type identities were extracted from a Seurat object without SCT transformation and filtered per pySCENIC guidelines, removing cells with fewer than 200 or more than 6000 genes, and genes present in fewer than 3 cells.

Vertebrate motifs were downloaded from JASPAR (Rauluseviciute et al., 2023) and converted to clusterbuster motifs using Biopython for input into pySCENIC. Transcription factor and motif names were extracted, and adjacencies between transcription factors were calculated from the filtered counts using pySCENIC's GRN function.

Regulons, highlighting enriched motifs, were calculated with filtered counts with the CTX function. Cells in which TF or target gene expression was 0 were masked when calculating the correlation between a TF-target pair. Then, both positive and negative regulons were calculated without pruning due to incompatibility with the novel bat genome annotation.

The area under the curve (AUC) for regulon enrichment was computed with the AUCELL function and combined with cell-type labels to generate regulon specificity scores (RSS) for both positive and negative regulons across cell clusters. Gene regulatory networks were constructed based on these RSS values, keeping only transcription factor-target gene connections with a weight greater than 10, and excluding target genes with mean expression below 0.05 across all cells.

6.9. Gene expression analysis with RNA-sequencing

For gene expression analysis, mouse and bat tissues were dissected, snap frozen and total RNA was extracted as described in 6.2.3. Libraries were prepared from polyA-enriched RNAs using KAPA HyperPrep Kit (Roche, #07962347001). RNA-seq libraries were sequenced as 100 bp paired-end reads using the Illumina NovaSeq 6000 system with 100 million reads per sample. Library preparation and sequencing was performed by the in-house Sequencing Core Facility of the Max Planck Institute for Molecular Genetics.

RNA-seq experiments were performed at least in biological duplicates.

6.9.1. RNA-seq mapping

Read mapping to mm39 and *Carollia perspicillata* reference genomes was performed using the STAR_2.6.1d software (Dobin et al., 2013) with the following options: "*--chimSegmentMin 10 --alignIntronMin 20 --outFilterMismatchNoverReadLmax 0.05 --outSAMmode NoQS --outFilterMismatchNmax 10*".

To analyze *MEIS2* and *TBX3* expression from knock-in constructs in transgenic animals, the transgene sequence was temporarily merged with the genome annotation using the option "*-add*".

For each sample, read counts per gene were obtained via the R function "featureCounts" (Liao, Smyth, & Shi, 2014), with the parameter "*--countReadPairs -s 2*". For visualization in UCSC browser, CPM normalized bigwig files were created using the "bamCoverage" tool from deepTools (Ramírez et al., 2016). For representation of gene expression in bar graphs, read counts were normalized using DESeq2 v1.38.3 (R v4.2.2) (Love, Huber, & Anders, 2014) (see 6.9.2).

6.9.2. Differential gene expression analysis

DEG analyses were performed by Anna Alessandra Monaco at the Max-Planck-Institute for Molecular Genetics (amonaco@molgen.mpg.de) in close exchange with myself to define the goals of the analysis, interpret results and design next steps.

Differentially expressed genes (DEGs) were identified from featureCounts (Liao et al., 2014) count matrices using the tool DESeq2 v1.38.3 (R v4.2.2) (Love et al., 2014). For each

comparison, the quality of and correlation between replicates was analyzed using PCA and Euclidean distance between samples. Outlier replicates were removed from DEG analysis if needed. Lowly expressed genes were then filtered using the edgeR tool v3.40.2 filterbyexpr() (Chen, Lun, & Smyth, 2016). Furthermore, non-annotated transcripts were removed to ensure one-to-one comparability between both species. Read counts were then normalized and differential expression calculated as log2 fold change (lfc) from the mean of normalized counts. Genes with a lfc larger than ±0.5 and an adjusted Wald test p value below 0.05 were finally assigned as differentially expressed.

6.10. Profiling histone modifications and transcription factor binding by ChIP-sequencing

6.10.1. Chromatin immunoprecipitation experimental procedures

ChIP-seq experiments were performed using iDeal ChIP-seq Kit for Histones (Diagenode, #C01010059) and iDeal ChIP-seq Kit for Transcription Factors (Diagenode, #C01010055) according to the manufacturer's instructions with minor modifications during cell or tissue fixation and chromatin preparation. In short, cells and tissues of embryos until E13.5 were dissociated with trypsin at 37 °C for 10 minutes, centrifuged at 1,100 rpm at 4 °C for 5 minutes and resuspended in 500 µL of cold 10 % FCS (PANSera ES, #P30-2600) in DPBS. Tissues derived from embryos older than E13.5 were directly transferred into 500 µL of cold 10 % FCS in DPBS. 500 µL of fixation solution (2% Formaldehyde and 10% FCS in DPBS) were added to the samples (final concentration of 1 % formaldehyde) and incubated shaking at 4°C for exactly 10 minutes. The fixation was stopped by adding 100 µL of 1.425 mM glycine and samples were immediately centrifuged at 500 x g (cells) or 850 x g (tissue) at 4 °C for 5 minutes. Samples were washed once in cold DPBS and cell pellets or tissue pieces were snap frozen in liquid nitrogen and stored until chromatin preparation.

For chromatin shearing, snap frozen cell pellets were resuspended in 1 mL Lysis Buffer 1 and incubated shaking at 4 °C for 20 minutes. Snap frozen tissue pieces were chopped in Lysis Buffer 1 and then incubated shaking at 4 °C for 20 minutes. Samples were pelleted by centrifugation at 500 x g (cells) or 850 x g (tissue). 1 mL Lysis Buffer 2 was added and samples were incubated shaking at 4 °C for 10 minutes. After pelleting samples as before, 300 µl of Shearing Buffer supplemented with 1x Proteinase Inhibitor Cocktail were added and cells and tissue pieces were sheared by sonication using the Bioruptor with the following settings: 4 °C, 30 seconds "on", 30 seconds "off", High Power setting, total of 45 cycles. Every 10 cycles, samples were mixed by vortexing, and liquid was briefly spun down. After shearing, samples were centrifuged at 16,000 x g at 4°C for 10 minutes. 10 µL were used for analysis of shearing

efficiency and chromatin amount and the remaining 290 µL were snap frozen in liquid nitrogen and stored at -80 °C until chromatin immunoprecipitation.

For analyzing shearing efficiency and chromatin amount, 10 µL of sheared chromatin were mixed with 1 µL of 5 M NaCl and incubated shaking at 65 °C overnight for reverse crosslinking. 1 µL of RNase A (10 mg/mL, Sigma #R4875) was added and samples were incubated at 37 °C for 1 hour. Following this, samples were mixed with 1 µL of Proteinase K (20 mg/mL) and incubated at 55 °C for 1 hour. DNA was precipitated by adding 1 µL of glycogen (Thermo Fisher Scientific, #R0561), 3 volumes of 100 % ethanol and 1/10 volume of 3 M NaAc. After incubation at -20 °C for 1 hour, samples were centrifuged at 4 °C at full speed for 30 minutes, washed once with 70 % ethanol and pellets were dissolved in 20 µL H_2O. The sample concentration was measured by NanoDrop, and the remaining sample was analyzed on a 1 % agarose gel. Chromatin immunoprecipitations were performed with 5 µg of chromatin for histones and 20 µg of chromatin for transcription factors. The antibodies used for chromatin immunoprecipitations in this study are listed in Table 1.

All ChIP-seq experiments were performed in biological duplicates.

Table 1: Antibodies used in this study for ChIP-seq experiments.

Antibody	*Amount per IP*	*Manufacturer and catalogue number*
Anti-H3K27ac	2.4 µg	Diagenode, #C15410174
Anti-H3K4me1	2.9 µg	Diagenode, #C15410037
Anti-H3K4me3	2.5 µg	Merck, #07-473
Anti-CTCF	1.2 µg	Diagenode, #C15410210
Anti-Meis1/2a + Anti-Meis2	2 µg each (4 µg total)	Kindly provided by Irene Delgado and Miguel Torres, as described in (Delgado et al., 2021)

E10.5 Meis1/2 ChIP data from mouse forelimbs was obtained from previously published data (Delgado et al., 2021).

6.10.2. ChIP-seq mapping

Read mapping of the sequenced samples to indexed mouse and bat reference genomes (mm39 / carper2) was performed using the STAR_2.6.1d software (Dobin et al., 2013). Reads were filtered, sorted and duplicates were removed using SAMtools (H. Li et al., 2009). Lastly, CPM normalized bigwig files were created using the “bamCoverage” tool from deepTools (Ramírez et al., 2016).

6.10.3. Peak calling and differential acetylation

Peak calling, differential acetylation and transcription factor motif analysis were performed by Bai-Wei Lo at the Max-Planck-Institute for Molecular Genetics (lo@molgen.mpg.de) in close exchange with myself to define the goals of the analysis, interpret results and design next steps. Visualization of differentially H3K27-acetylated regions was done by Anna Alessandra Monaco at the Max-Planck-Institute for Molecular Genetics (amonaco@molgen.mpg.de).

Differentially acetylated regions between distal forelimb and distal hindlimb and acetylated regions common to both tissues were predicted from H3K27ac ChIP-seq alignments using macs2 (Y. Zhang et al., 2008) bdgdiff command with the following parameters: *-l 800 -g 500* and a likelihood ratio cut-off of 1000. For each TAD, determined by Hi-C, the coverage of acetylated regions in the distal forelimb relative to the regions shared between distal forelimb and distal hindlimb was calculated. The acetylated regions specific to the distal forelimb, and the acetylation regions shared between distal forelimb and distal hindlimb were first intersected with accessible regions in bat forelimbs (CS17). The distal forelimb-specific accessible acetylated regions were given as input and the commonly accessible acetylated regions as background for motif enrichment analysis done by homer2 (Heinz et al., 2010) with a q-value cut-off of 0.01. The list of significantly enriched motifs was manually curated to only retain the most significant TF motifs for each gene family. For visualization of peak distribution, Tornado plots were generated using deeptools v3.5.4. (Ramírez et al., 2016).The scores per regions were calculated using computeMatrix in scale-regions mode. Here, the scores were based on the ChIP bigwig pileup files as well as the regions based on BED files defining the determined ChIP-seq peaks.

MEIS binding signal of one of the distal bat forelimb replicates was aggregated by TADs. The signal (area under the curve) was calculated using deepTools pyBigwig (Ramírez et al., 2016). By analyzing the second derivative of the density distribution of the signal by TAD, the dividing point to the subpopulation of enriched TADs was determined. The same procedure was carried out for the coverage length of the acetylation peaks identified above. Finally, genes were categorized as "Cytoskeleton”, "ECM”, or "Transcription Factor” using the Uniprot database (Consortium, 2022), if keywords included the terms "Cytoskeleton”, "Extracellular matrix”, or ("DNA-binding” OR “Transcription regulation”), respectively.

6.11. Profiling chromatin accessibility with ATAC-sequencing

6.11.1. ATAC-seq experimental procedure

To profile chromatin accessibility in embryonic tissue cells were dissociated with trypsin at 37°C for 10 minutes. Trypsinization was stopped by adding 10 % FCS (PANSera ES, #P30-2600) in DMEM (Gibco, #10829-018) and samples were centrifuged at 300 x g at 4 °C for 5 minutes.

Cell pellets were resuspended in 500 µL freezing media 1 (20 % FCS in DMEM) and transferred into a cryovial. 500 µL of freezing media 2 (20 % FCS, 20 % DMSO (Sigma-Aldrich, #D2650) in DMEM) were added and cells were cryo-frozen in a Mr. Frosty overnight at -80 °C and stored in liquid nitrogen until library preparation.

Prior to library preparation, cryo-preserved cells were thawed at 37 °C, transferred into pre-warmed cell culture media (10 % FSC in DMEM) and centrifuged at 1,100 rpm at room temperature for 5 minutes to remove residual DMSO. After resuspension in culture media, cells were counted using Countess cell counter and 60,000 - 75,000 cells were transferred into a 1.5 mL reaction tube.

ATAC-seq library preparation was performed according to (Buenrostro, Wu, Chang, & Greenleaf, 2015). In short, cells were washed with cold DPBS, resuspended in 50 µL cold lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.1 % (v/v) Igepal CA-630) and immediately centrifuged at 500 x g at 4 °C for 10 minutes to isolate nuclei. For transposition, nuclei pellets were resuspended in transposition reaction mix (25 µL TD buffer, 2.5 µL TDE1 Tagment DNA Enzyme (Illumina, # 20034198) and 22.5 µL nuclease -free H_2O) and incubated at 37 °C for 30 minutes. DNA fragments were purified using MiniElute PCR Purification Kit (QUIAGEN, #28004) according to the manufacturer's instructions and eluted in 10 µL elution buffer EB. The transposed fragments were amplified and indexed with barcoded adaptors via PCR (reaction mix: 10 µL transposed DNA, 10 µL nuclease-free H_2O, 2.5 µL Custom Nextera PCR Primer 1 and 2 (each 25 µM), 25 µL NEBNext High Fidelity 2x PCR Master Mix (NEB, #M0541S) / thermocycler: 1 cycle (5 minutes at 72 °C, 30 seconds at 98 °C), 5 cycles (10 seconds at 98 °C, 30 seconds at 63 °C, 1 minute at 72°C)). To determine the appropriate number of PCR cycles, a qPCR side reaction is run with 5 µL of the amplified DNA (reaction mix: 4.41 µL nuclease free H_2O, 0.025 µL Custom Nextera Primer 1 and 2 (25 µM), 0.09 µL SYBR Green I (Invitrogen, #S-7563), 5 µL NEBNext High Fidelity 2x PCR Master Mix (NEB, #M0541S) / thermocycler: 1 cycle (30 seconds at 98 °C), 20 cycles (10 seconds at 98 °C, 30 seconds at 63 °C, 1 minute at 72°C)). The remaining 45 µL PCR reaction were run with the determined additional cycle number. Finally, libraries were purified by double-sided size selection (0.5 X / 1.3 X) using AMPure XP beads (Beckmann, #A63881) and eluted in 20 µL of 10 mM Tris-HCl. Concentration was measured using Qubit dsDNA HS Assay Kit (Invitrogen, #Q33231) and quality was assessed with Bioanalyzer High Sensitivity DNA Analysis Kit (Agilent, #5067-4626).

Libraries were sequenced as 100 bp paired-end reads using the Illumina NovaSeq 6000 system with 100 million reads per sample. Library sequencing was performed by the in-house Sequencing Core Facility of the Max Planck Institute for Molecular Genetics.

All ATAC-seq experiments were performed as biological duplicates.

Mouse ATAC-seq data of embryonic limbs at E11.5 and E13.5 was obtained from previously published data (Paliou et al., 2019) and data was mapped as described above.

6.11.2. ATAC-seq mapping

For processing of sequenced ATAC-seq files, adapters were trimmed using the cutadapt tool (Martin, 2011). Bowtie2 was used for read mapping to the indexed reference genomes (mm39/carPer2) (Langmead & Salzberg, 2012). Reads were filtered, sorted and duplicates were removed using SAMtools (H. Li et al., 2009). Lastly, for visualization, CPM normalized bigwig files were created using the “bamCoverage” tool from deepTools (Ramírez et al., 2016).

6.12. Analyzing 3D genome architecture by Hi-C

6.12.1. Hi-C experimental procedure

Hi-C libraries from bat embryonic fibroblast cells or bat and mouse embryonic forelimbs were prepared as previously described (Rao et al., 2014).

$1x10^6$ bat fibroblast cells and 500,000 limb cells were fixed in 2% formaldehyde in 10%FCS/PBS for 10 minutes at room temperature. Fixation was stopped by adding 1/10 volume of 1.425 M glycine and cells immediately pelleted at 400 x g for 8 minutes at 4°C. Pellets were resuspended in lysis buffer (50 mM Tris/HCl pH 7.5, 1250 mM NaCl, 5 mM EDTA, 0.5 % NP-40, 1.15% TX-100, 1x protease inhibitors (cOmplete EDTA-free Protease Inhibitor Cocktail, Merck, #04693259001)) and incubated for 10 minutes on ice. Nuclei were pelleted at 2400 rpm for 2 minutes at 4°C, snap frozen in liquid nitrogen and stored until processing.

For restriction digest, nuclei pellets were thawed and resuspended in 500 µl 1X *DpnII* buffer with 50 µl 0.5% SDS/ H_2O and incubated for 10 min at 62°C. 145 µl of H_2O were added and SDS was quenched by adding 10% Triton-X 100 (Sigma, #93443). Chromatin was subsequently digested with two installments of *DpnII* enzyme (100 U, Biolabs, #R0543L) for 2 hours at 37°C. After digestion, the enzyme was inactivated for 20 minutes at 65°C. Samples were centrifuged at 5000 x g for 5 minutes and resuspended in 200 µL of 1.2X NEB buffer 2 (Biolabs, #B70025). DNA fragment ends were marked with biotin-14-dATP (Life Technologies, #19524-016) using the following master mix: 37.5 µL 0.04 mM biotin-14-dATP, 1.5 µL 10 mM dCTP, 1.5 µL 10 mM dGTP, 1.5 µL 10 mM dTTP, 8 µL 5U/µL DNA Polymerase I, Large (Klenow), Fragment (NEB, #M0210). The reaction mix was incubated for 45 minutes to 1.5 hours at 37°C. For proximity ligation of marked DNA fragments, the following master mix was assembled and added to samples: 663 µL H2O, 120 µL 10X NEB T4 DNA ligase buffer (NEB, #B0202), 100 µL 10% Triton X-100, 12 µL BSA (10 mg/mL), 5 µL T4 DNA Ligase (NEB, #M0202, 400 U/mL). Samples were incubated over night at 18°C. The following day, samples were cnentrifuged at 5000 x g for 5 minutes, resuspended in 600 µL 10 mM Tris/HCl (pH 7.5) and incubated with 25 µL Proteinase K and 60 µL 10% SDS at 55°C for 30 minutes.

Subsequently, crosslinking was reversed by adding 65 µL of 5M NaCl and incubation at 65°C for 2 hours. DNA was precipitated by adding 750 µL of 100% Ethanol and 75 µL 3M NaAc (pH 5.2). After incubating at -80°C for 15 minutes, samples were centrifuged at maximum speed at 4°C for 15 minutes. DNA pellets were washed with 70% Ethanol and resuspended in 130 µL of 10 mM Tris/HCl (pH 8.0) and incubated for 15 minutes at 37°C to fully dissolve DNA. DNA concentration was measured by Qubit High Sensitivity Assay and fragments were sheared to a fragment size of 300-500 bp using a S-Series 220 Covaris sonicator with the following settings: 1 cycle, 55 seconds, Duty cycle 10%, Intensity 4, Cycle/burst 200. Samples were transferred into DNA low-bind tubes and biotin-containing fragments were pulled down with Dynabeads MyOne Streptavidin T1 beads (Invitrogen, #65602). For this, Dynabeads were washed with 400 µL 1X Tween Washing Buffer (15mL 10mM Tris pH 7.5, 6mL 5M NaCl, 30µL 0.5M EDTA, 15µL of Tween 20, 8.96mL H_2O), resuspended in 300 µL 2X Binding Buffer (5mL 5M NaCl, 100µL 1X TrisHCl, 20µL 0.5M EDTA, 5.9mL of H_2O) and added to sheared samples. After incubation for 15 minutes at room temperature, beads were washed twice with Tween Washing buffer for 2 minutes at 55°C. Finally, beads were resuspended in 100 µL 1X NEB T4 DNA Ligase Buffer (NEB, #B0202). The buffer was discarded, and beads were resuspended in the following master mix and incubated for 30 minutes at room temperature: 8.8µL 10X NEB T4 DNA ligase buffer with 10mM ATP, 2µL of 25mM dNTP mix, 5µL 10U/µL NEB T4 PNK (NEB, #M0201), 4µL 3U/µL NEB T4 DNA Polymerase (NEB, #M0203), 1µL 5U/µL NEB DNA Polymerase I, Large (Klenow) Fragment (NEB, #M0210), 78.2µL H_2O. Beads were washed in 100µL 1X NEB buffer 2 and 600 µL of 1X Tween Washing Buffer for 2 minutes at 55°C and resuspended in 100µL of the following master mix and incubated for 30 minutes at 37°C: 90µL 1X NEB Buffer 2, 5µL 10 mM dATP, 5µL NEB Klenow exo minus (NEB, M0212, 5 U/µL). Beads were washed in 600µL 1X Tween Washing buffer, 100µL 1X NEB Quick Ligation buffer (NEB, #B6058) and finally resuspended in 50µL 1X NEB Quick Ligation buffer. For adaptor ligation, 2µL of NEB DNA Quick Ligase (NEB, #M2200) and 3µL of the universal adapter from NEBNext Ultra DNA Library Prep Kit for Illumina (Biolabs, #E7601A) was added. After incubation for 15 minutes at room temperature, 3µL user enzyme were added and mixture was incubated for another 15 minutes. Subsequently, beads were washed twice in 600µL Tween Washing Buffer for 2 minutes at 55°C, once in 100µL 10mM Tris/HCl (pH7.5) and finally resuspended in 50µL 10mM Tris/HCl (pH7.5). Final amplification and adding of indices via PCR were performed using following PCR set up: 15µL adaptor-ligated DNA fragments on beads, 25µL NEBNext Ultra II Q5 Master Mix (Biolabs, #M05445), 5µL Index Primer/i7 Primer, 5 µL Universal PCR Primer/i5 Primer. Cycle number was determined based on previously measured DNA concentration. Afterwards, samples were purified and size selected using AMPure XP beads (Beckmann, #A63881). Finally, DNA concentration was measured using Qubit DNA High

Sensitivity Kit and library quality was assessed with Bioanalyzer using High Sensitivity DNA Reagents (Agilent Technologies, #5067-4626).

Hi-C libraries were generated as three technical replicates and sequenced on a NovaSeq2 with 100bp paired-end reads.

6.12.2. Hi-C mapping and TAD calling

For processing of Hi-C samples, the reference genomes (mm39/carPer2) were first indexed with the short read aligner BWA 0.7.17 (H. Li & Durbin, 2010). Raw reads from each sequenced Hi-C library were then processed using the Juicer pipeline v1.5.6 as previously described (Durand et al., 2016). The three technical replicates were processed independently and merged after filtering and deduplication. For visualization, Hi-C maps were generated using Juicer tools 1.11.09 (Durand et al., 2016) using the parameter "*pre -q 30*". Hi-C maps were displayed as KR-normalized heatmaps with 5 kb bin size.

For analysis of regulatory regions, topologically associated domains (TADs) were determined using Hi-C data of mouse limbs or *Carollia perspicillata* embryonic fibroblasts using the software TopDom (Shin et al., 2016) with the following parameters: *KR normalized, resolution: 50 kb, window size: 10*.

6.13. Plasmid design and cloning

6.13.1. Design and cloning of single guide RNAs for CRISPR/Cas9-mediated genome editing

sgRNA constructs were generated as previously described (Ran et al., 2013). Briefly, complementary sgRNA oligos were annealed, phosphorylated and cloned into *BbsI* digested and dephosphorylated pSpCas9(BB)-2A-Puro (PX459) V2.0 vector (Addgene; #62988).

sgRNA constructs targeting the *H11* and *Rosa26* locus were cloned and generously provided by Francisca Martinez Real and Konrad Chudzik. scRNA sequences can be found in section 10.4 (Table 4).

6.13.2. Design and cloning of expression constructs

CRISPR-mediated, site-specific knocking of expression constructs at H11 and Rosa26 loci were performed by homology-mediated insertion based on a plasmid donor. For this, a sgRNA was used to introduce a single double strand break at the site of interest. The plasmid-derived donor DNA sequence is flanked by homology arms surrounding the insertion site to mediate homology directed repair (HDR).

For cloning of donor expression constructs, necessary backbone and fragment components were assembled from several plasmids. A pUC-Amp plasmid containing symmetric homology arms (0.7 kb) designed on the H11 knock-in site as well as ampicillin antibiotic resistance and

plasmid backbone was ordered from Twist Bioscience. The plasmid containing Rosa26 asymmetric homology arms (2 kb and 0.8 kb) serving as PCR template was generated and generously provided by Konrad Chudzik. The Dathe-*Bmp2* enhancer (Dathe et al., 2009) was amplified from G4 wildtype DNA. *Carollia perspicillata MEIS2* cDNA was ordered from Twist Bioscience as a fragment and *TBX3* cDNA was ordered from GeneWiz (Azenta Life Sciences) in a pUC-Amp vector. Hsp68 promoter, Kozak sequence (GAGTGG), SV40 polyA signal were included in the design of both ordered overexpression constructs.

Backbones and fragments were amplified by PCR using PrimeSTAR GXL Polymerase (Takara, Cat #R050A), introducing overlapping homologous sequences for Gibson assembly. Fragments were assembled using Gibson Assembly Master Mix (NEB, Cat #M5510) and cloned into 5-alpha Competent *E. coli* cells (NEB, Cat. #C2987). Products were validated via restriction digestion and Sanger sequencing (EuroFins).

Recombinant DNA and primers used for cloning can be found in section 10.4 (Table 2 and Table 6).

For alignment of MEIS2 and TBX3 protein sequences, bat and mouse coding sequences were translated into amino acid sequences using ExPASy (Gasteiger et al., 2003) and protein sequences were aligned using the Multiple Sequence Alignment tool MultAlin (Corpet, 1988).

6.14. Cell culture

6.14.1. Culturing of bat embryonic fibroblasts

Head- and organ-free tissue from a CS16/CS17 bat embryo (*Carollia perspicillata*) was minced in DMEM supplemented with 15% FCS and cryo-frozen in DMEM containing 10% DMSO and 15% FCS until further processing.

Bat embryonic fibroblasts were established based on previously published instructions (Qiu, Lai, Stumpo, & Blackshear, 2016).

In short, cryo-frozen tissue pieces were thawed and subsequently digested with trypsin for 20 minutes at 37°C. Dissociated cells were centrifuged at 1,000 rcf for 5 minutes, resuspended in fibroblast culture media (DMEM high glucose (Gibco, #10829-018), 15% fetal calf serum (FCS,PANSera ES, #P30-2600), 1% penicillin/streptomycin (Lonza, #DE17-603), 1% L-glutamine (Lonza, #BE17-605E) and transferred to a 6 well plate for cell attachment and expansion. Remaining tissue pieces were transferred into a separate well and fixated with a glass coverslip for fibroblast cell outgrowth. When cells reached a density of approximately 80%, they were split into a new culturing flask or cryo-frozen in freezing media (DMEM high glucose supplemented with 15% FCS and 10% DMSO (Sigma-Alrich, #D2650) in 1-3x10^6 cells/vial aliquots. Fibroblast cells were cultured at 37°C and 5% CO_2.

6.14.2. Culturing of mouse embryonic stem cells (mESCs)

Culturing of mouse embryonic stem cells (mESCs) was performed according to procedures previously described (Andrey & Spielmann, 2017; Behringer, 2014; George et al., 2007; Kraft et al., 2015).

For the experiments in this study mouse embryonic stem cells of the strain G4 (129/Sv x C57BL/6 F1 hybrid genetic background) have been used. G4 cells were cultured on a monolayer of mitotically inactivated feeder cells (mouse embryonic fibroblasts, MEFs) derived from E13.5 and E14.5 CD1 and DR4 (puromycin/ neomycin/ hygromycin resistant) embryos. MEFs were prepared in house by Asita C. Stiege. Feeder cells were seeded at least 6 hours prior to ESC culture on gelatinized cell culture dishes (0.01 % gelatin (Sigma, #G-1393), 37 °C for 15 minutes). Mouse ESCs were cultured in mESC medium (Knockout Dulbecco's Modified Eagle's Medium (DMEM) 4,500 mg/ml glucose, with sodium pyruvate (Gibco, #10829-018), 15 % fetal calf serum (FCS, PANSera ES, #P30-2600), 10 mM Glutamine (Lonza, #BE17-605E), 1x penicillin/streptomycin (Lonza, #DE17-603), 1x non-essential amino acids (Gibco, #11140-35), 1x nucleosides (Chemicon, #ES-008D), 0.1 mM β-Mercaptoethanol (Gibco, #3150-010) and 1000 U/mL Murine Leukemia Inhibitory Factor (Chemicon, #ESG1107)) at 37 °C and 7.5% CO_2. The medium was changed every 24 hours and cells were passaged onto a fresh dish every 2-3 days or cryo-frozen in a density of 0.5-1x10^6 cells/vial.

For passaging, cells were washed with DPBS and dissociated with Trypsin-EDTA (Gibco, #25300-054) at 37 °C for 10 minutes. The cell suspension was resuspended in mESC culture medium to inactivate the trypsin and centrifuged at 1,100 rpm at room temperature for 5 minutes. Cells were then seeded in the desired density on a fresh feeder coated dish. For freezing, cells were dissociated as described. After centrifugation, 0.5-1x10^6 cells were resuspended in 500 µL mESC medium containing 20 % FCS and transferred to a cryo-vial containing 500 µL of mESC freezing medium (Knockout Dulbecco's Modified Eagle's Medium (DMEM) 4,500 mg/ml glucose, with sodium pyruvate with 20 % FCS and 20 % DMSO (SIGMA #D-2650)). Vials were frozen in a Mr. Frosty over night at -80 °C and stored in liquid nitrogen.

6.14.3. CRISPR/Cas9-based Genome Engineering of mESCs

Genome editing of G4 mouse embryonic stem cells was performed as described in (Kraft et al., 2015) and for transfections FuGENE HD Transfection Agent (Promega, #E2311) was used according to the manufacturer’s instructions. In short, 3x10^5 G4 cells were seeded in a 6 cm dish onto a CD1 feeder monolayer and cultured for 24 hours. 2 hours before transfection, medium was exchanged with mESC medium without penicillin/streptomycin. 8 µg of the sgRNA containing pX459 vector and 4 µg of the targeting construct were combined in 125 µL of OptiMEM (Gibco, #51985-026) and 100 µL OptiMEM were mixed with 25 µL of FuGENE HD Agent in a separate tube. Both mixtures were combined, incubated at room temperature for 15

minutes. The mixture was then added dropwise to G4 cells. After 16-18 hours the transfection medium was changed to mESC medium and cells were allowed to recover for 24 hours. After recovery, transfected G4 cells were passaged onto 3 6 cm dishes with puromycin resistant DR4 feeders and subjected to antibiotic selection for 48 hours using mESC medium supplemented with puromycin (2 µg/mL, Sigma-Aldrich, #P8833). After selection, cells were allowed to recover in mESC medium until mESC colonies were of adequate size for picking. For this, single colonies were picked, dissociated into single cells and transferred into a 96 well plate with CD1 feeder cells. Clones were grown for 2-3 days and plates were then split into 3 replicates, 2 of which were frozen in U-bottom 96 well plates in plate freezing medium (Bicarbonate free DMEM (DMEM powder (Gibco #52100) in H_20) containing 10 mM HEPES (Sigma #H-0887), 20 % FCS, 10 % DMSO (Sigma-Aldrich, #D2650)). The third plate was cultured for 2-3 days and genomic DNA was extracted from expanded clones for genotyping by PCR. Positive clones were thawed from the frozen 96 well plates, expanded to 6 cm dishes and then cryo-frozen as described above. An aliquot of ~$1x10^5$ cells was feeder depleted, cultured for 2-3 days and then used for extraction of genomic DNA to verify the genotype by PCR and assess the copy number by qPCR.

6.14.4. Genotyping of Genome-Engineered mESCs

For genotyping via PCR or copy number analysis via qPCR genomic DNA of cultured cells was extracted as described in 6.2.2. All primers used in this study can be found in the appendix in table Table 2.

Genome-engineered mESCs were screened for the desired mutation and verified via genotyping PCR using the following reaction mix and thermal cycler program: 2 µL PCR buffer (10x, Fermentas, 750 mM TrisHCl pH 8.8, 200 mM $(NH_4)_2SO_4$, 0.1% Tween20, 15 mM $MgCl_2$), 0.5 µL dNTP Mix (1.25 mM), 0.5 µL Primer Mix (100 µM each), 0.35 µL *Taq* Polymerase (prepared by Asita Carola Stiege in house), 50 ng mESC DNA, H_2O ad 20 µL / 1 cycle (94°C 5 min), 35 cycles (94°C 5 min, 55°C 30 sec, 72°C 1 min/1 kb), 1 cycle (72°C 10 min, 4°C ∞). PCR products were analyzed by agarose gel electrophoresis according to standard procedures.

Copy number analysis from genomic DNA via qPCR was performed using the Biozym Blue S'Green qPCR Kit (Biozym, #331416). The following reaction mix was assembled in a 384-well-plate in triplicates for each sample: 1 µL Primer Mix (2.5 µM each), 6 µL Blue S'Green qPCR 2x Mix (supplemented with 4 µL ROX Additive (50 µM)), 3 µL H_2O, 2 µL DNA (2 ng/µL). qPCRs were run with the following thermal cycler program: 1 cycle (95°C 10 min), 40 cycles (95°C 15 sec, 50°C 1 min, 72°C 1 min/1 kb), followed by melting curve). For analysis, Ct values were normalized to a control region outside the targeted site and compared either to a wildtype control or between tested clones to assess the copy number.

6.15. Tetraploid morula complementation to generate mutant embryos

Transgenic embryos were generated through tetraploid morula aggregations. For aggregation mutant ESCs were seeded on a CD1 feeder layer and cultured until colonies reached appropriate size. These colonies were then used for morula complementation as previously described (Artus & Hadjantonakis, 2011).

Aggregation experiments were conducted by the in-house transgenics core unit at the Max Planck Institute for Molecular Genetics.

6.16. Whole mount *in situ* hybridization (WISH)

6.16.1. Generation of DIG-labelled *in situ* probes

mRNA expression from *MEIS2* expression constructs in whole embryos was visualized by WISH using digoxigenin (DIG)-labelled antisense RNA probes. A probe specifically targeting construct expression was generated by PCR amplification using the generated plasmid as template. The PCR reaction was performed with *Taq* polymerase using the following reaction mix and cycler program: reaction mix 5 µL Pfu buffer (Fermentas buffer with 20 mM $MgSO_4$), 1 µL 1.25 mM dNTPs, 1 µL primer mix (10 µM each), 1 µL polymerase mix (*Taq*:*Pfu* in a ratio of 9:1), 2 µL DNA template, 40 µL H_2O / thermocycler: 1 cycle (5 minutes at 96 °C), 20 cycles (30 seconds at 96 °C, 30 seconds at 58 °C, 2 minutes at 72°C), 1 cycle (30 minutes at 72 °C), 4°C ∞.

The PCR products contained A-overhangs which allowed sub-coning into a pTA vector provided by Asita C. Stiege at the Max Planck Institute for Molecular Genetics. For subcloning into competent Top10 *E.coli* cells, a ligation mix of 1 µL T4 DNA ligase, 1 µl 10 mM ATP, 1 µL PEG4000, 2 µL T4 ligase buffer, 50 ng digested pTA vector, 300 ng insert and H_2O ad 20 µL were combined and incubated at room temperature for 2 hours. 10 µL of ligation mix were used for transformation of *E.coli* cells. The following day, colonies were checked for correct plasmid sequences by restriction digest and probe sequences and orientation were confirmed by Sanger sequencing. The resulting plasmid DNA served as template for PCR amplification using T7 or SP6 primers and subsequent *in vitro* transcription. For PCR amplification the following reaction mix was combined: 2 ng/µL DNA template, 10 µL 10x fermentas buffer, 8 µL 1,25 mM dNTPs, 4 µL SP6 and T7 primer (10 µM), 2 µL *Taq*, H_20 ad 100 µL. PCR fragments were purified using MinElute Reaction Cleanup Kit (QIAGEN, #28204). 200 ng purified product were used as template for *in vitro* transcription and DIG labelling. For this, the template was combined with DIG RNA labelling mix (Roche, #11277073910), transcription buffer and RNA polymerases SP6 or T7 (Roche, #10999644001) depending on template orientation to generate an anti-sense probe. This reaction mixture was incubated for 2 hours at 37°C.

Remaining DNA was digested with *DNaseI* treatment (Roche, #04716728001) for 15 minutes at 37°C and the reaction was stopped by adding 2 mM EDTA/ H_2O -DEPC (pH 8.0). Finally, Finally, DIG-labeled RNA probes were precipitated with 1/10 volume 0.4 M LiCl and 3 volumes 100% ethanol. After incubation for 20 minutes at -80 °C, probes were pelleted by centrifugation for 20 minutes at maximum speed at 4°C, pellets washed with 70% ethanol and resuspended in 10mM Tris.

Primer sequences can be found in section 10.4 (Table 5).

6.16.2. WISH experimental procedure

For WISH experiments, whole embryos were dissected in cold PBS, fixed in 4% PFA/PBS (pH 7.4) for overnight at 4°C and de-hydrated the following day by an increasing methanol series (25%, 50% 75% and 100% methanol in PBS (v/v)). Each incubation was performed for 30 minutes on ice. De-hydrated samples were stored in 100% methanol at -20°C until further processing.

Prior to WISH procedure, embryos were re-hydrated through a stepwise incubation in 75%, 50%, and 25% methanol in PBST ((v/v), PBST: PBS with 0.1% Tween20) according to de-hydration. Embryos were then bleached for 1 hour on ice with 6% H_2O_2 in PBST. Embryos were washed with PBST at room temperature and subsequently treated with proteinase K (10 µg/ml) for 5 minutes. Digestion was quenched with glycine (2 mg/ml). After three washes in PBST, embryos were fixed in 4% PFA for 20 minutes at room temperature. Next, embryos were incubated in L1 buffer (50% formamide, 5X SSC, 1% SDS, 0.1% Tween20 in DEPC-H2O, pH 4.5) for 15 minutes at 68°C. To block unspecific binding sites, 0.1% transfer RNA (RNA type III from baker's yeast, Sigma, #R6750) and 0.05% heparin (Sigma, #H3149) were added to L1 buffer, and the embryos were incubated for an additional 2 hours at 68°C. DIG-labeled RNA probes (1.5 µg probe/embryo) were then added to the reaction mixture, and embryos were incubated overnight. The following day, unbound probes were removed through multiple washing steps in the following order: three washes of 30 minutes each in L1 at 68°C, three washes of 30 minutes each in L2 (50% formamide, 2X SSC pH 4.5, 0.1% Tween20) at 68°C, one wash of 20 minutes in L3 (2X SSC pH 4.5, 0.1% Tween20) at 68°C, and one wash of 20 minutes at room temperature. RNA was digested by treatment with RNase solution (0.1 M NaCl, 0.01 M Tris pH 7.5, 0.2% Tween20, 100 µg/ml RNase A in H2O) for 1 hour, followed by several washes with PBST. The embryos were then incubated in blocking solution (1X TBS, 1% Tween20, 2% lamb serum (Gibco, #16070), 0.2% BSA (Sigma, #A2153)) for 2 hours at room temperature and then further incubated overnight at 4°C in the same blocking solution containing anti-digoxigenin-alkaline phosphatase antibody (1:5000, Roche, #11093274910). The next day, unbound antibodies were removed by two washes for 1 hour and two washes for 30 minutes in TBST2 buffer (1X TBS, 0.1% Tween20, 0.05% levamisole/tetramisole

(Sigma, #L9756)). Embryos were then prepared for staining by a 15-minute incubation in alkaline phosphatase buffer (20 mM NaCl, 50 mM MgCl2, 0.1% Tween20, 0.1 M Tris-HCl, 0.05% levamisole/tetramisole). Staining was performed using BM Purple AP Substrate (Roche, #1442074) until the desired color intensity was achieved (~2 hours). Upon completion of staining, embryos were washed with PBST, fixed in 4% PFA/PBS supplemented with 0.2% glutaraldehyde and 5 mM EDTA for long-term storage at 4°C. Imaging of the embryos was performed using the ZEISS SteREO DiscoveryV12 with a cold light source CL9000 microscope and Leica DFC420 digital camera.

6.17. 3D imaging of mouse limbs

Staining and imaging of mouse limbs was performed by Rose Yinghan Behncke and René Hägerling at Berlin-Brandenburg Center for Regenerative Therapies at Charité Universitätsmedizin Berlin (rose.behncke@charite.de; rene.haegerling@charite.de) in close exchange with myself to define the goals of the analysis, interpret results and design next steps.

For 3D imaging of mouse limbs, E15.5 limbs were dissected in cold PBS and limbs subsequently fixed in 4% PFA in PBS overnight. The following day, limb tissues were dehydrated in increasing methanol concentrations (25%, 50%, 75%, 100% (v/v) methanol in PBS) and stored at -20°C until processing.

Before processing, limbs were re-hydrated in decreasing methanol concentrations (75%, 50%, 25%, 0% (v/v) in PBS) and kept in PBS on ice. Limb specimens were then incubated in a solution of 25 µM DraQ5, dissolved in PermBlock solution (1% BSA, 0.5% Tween 20 in PBS), for 12 hours. After three washes with PBS-T, specimens were dehydrated in increasing methanol concentrations (50%, 70%, 95%, and 99% (v/v) methanol in H_2O) and stained in a solution of 1.5 µM eosin Y, dissolved in a 1:1 methanol:BABB (benzyl alcohol:benzoate, ratio 1:2) for 4 hours. Specimens were then treated twice with an optical clearing procedure with BABB solution for 4 hours each.

After fluorescence whole-mount staining, optically cleared embryo limbs were imaged using Lightsheet 7 microscope (Zeiss, Oberkochen, Germany). Individual stacks were captured with a step size of 2.5 µm and at 5x magnification. For the operation of the microscope and acquisition of images the ZEN 3.1 (black edition) software was used. Digital 3D reconstruction of light sheet image stacks was done using IMARIS Microscopy Image Analysis Software (Oxford Instruments, Abingdon, UK).

For image analysis and quantifications, IMARIS Microscopy Image Analysis Software was used. For all measurements, the autopod region excluding the digits was considered. For the measurement of the total volume of mouse limbs, the Volume function in IMARIS was used. The total number of cells was quantified within DraQ5-positive nuclei using the Spots function

in IMARIS. For the quantification of connective tissue in the limbs, Eosin Y-positive structures of the limbs were analyzed using the Volume function in IMARIS. 3D imaging was done for 4 individuals of each genotype. Testing for statistical significance was done by one-sided ANOVA followed by Dunnett's multiple comparisons test using GraphPad Prism software.

6.18. Enhancer prediction and conservation

Enhancer prediction and conservation was analyzed by Tobias Zehnder at Max Planck Institute for Molecular Genetics (zehnder@molgen.mpg.de) in close exchange with myself to define the goals of the analysis, interpret results and design next steps.

Enhancers in mouse and bat limb tissues were predicted by intersecting open chromatin regions mapped by ATAC-seq and enhancer/promoter prediction based on histone modifications profiled by ChIP-seq. First, Genrich tool was used to call ATAC-seq peaks (Gaspar, 2018). Peaks overlapping an enhancer predicted by CRUP with a probability score >0.3 (Ramisch et al. 2019) and not overlapping an annotated TSS (UCSC) or a promoter predicted by eHMM (T. Zehnder, Benner, & Vingron, 2019) were annotated as enhancer.

Enhancer conservation analysis was performed as previously described (Tobias Zehnder, 2021). The corresponding code can be accessed at GitHub (tobiaszehnder/IPP, published 15.09.2021, GitHub, https://github.com/tobiaszehnder/IPP.git, last accessed 06.09.2024).

In short, predicted enhancers were projected between mouse and bat using a stepped pairwise sequence alignment approach across multiple bridging species. The set of bridging species used here can be found in Table 7 in section 10.5. Projected elements from predicted enhancers were then classified into directly (DC), indirectly (IC) or not conserved (NC) CREs according to the following criteria: DC elements overlap a direct sequence alignment between the reference and target species; IC elements do not overlap a direct sequence alignment but are either directly overlapping or in vicinity to a multi-species anchor. The remaining enhancer elements are classified as non-conserved (NC).

6.19. Identification of bat accelerated regions (BARs)

Analysis of bat accelerated regions was performed in collaboration with Ariadna Morals and Michael Hiller at the Centre for Translational Biodiversity Genomics, Senckenberg Society for Nature Research & Goethe University, in Frankfurt am Main, Germany (ariadna.morales@senckenberg.de; michael.hiller@senckenberg.de).

To identify regions with higher evolutionary rates in bats compared to other mammals (i.e. Bat Accelerated Regions or BARs), we used phyloP and screened a set of intergenic conserved neutral elements (CNEs) identified by (Hecker & Hiller, 2020) within a 508 multi-genome alignment (508 mammalian species including 83 bats) . However, we applied additional filters to merge elements spaced within less than 20 bp, each with a minimum length of 50 bp,

resulting in combined elements of at least 100 bp. The above filters resulted in a set of 355361 CNEs. Then, we ran PhyloP using a likelihood ratio test to compute p-values and test conservation/acceleration (--mode CONACC) in the ancestral Chiroperan branch and its descendant lineages (--subtree). The resulting p-values were filtered to retain only negative values to indicate acceleration, and then transformed into -log10 to apply a False Discovery Rate Correction (FDR). However, setting a statistical threshold to keep or discard BARs is challenging. Therefore, we ranked the BARs based on the corrected values after FDR, resulting in 66006 BARs. As a positive control for the BARs identified in our 508 mammalian alignment, we overlapped the 166 BARs previously found by (Booker et al., 2016) and identified 149 BARs that overlapped with our set. Notably, BAR116 highlighted in this study, which has relevant implications for limb development, is outside our set, as the values of acceleration are not statistically significant. We thoroughly checked this region (different settings and even different multi-genome alignments) and always got the same result: no significant acceleration. Finally, to place our BARs into the genome coordinates of our model organism, we liftOver the coordinates from human (hg38) to mouse (mm10) to *Carollia perspicillata* (carPer2). The final set of BARs mapped to the genome of *C. perspicillata* includes 55974 genomic elements.

7. Results

The following part will present the results of this study in three sections. First, limb developmental cell populations of mice and bats are mapped at single cell level to annotate and compare cell populations and their features during limb differentiation, with particular emphasis on the interdigital mesenchyme. Section two reveals the molecular signature of the chiropatagium tissue in bat wings, specifically focusing on its cell composition and transcriptomic profile. Furthermore, it outlines potential regulatory features involved in shaping the bat autopod. Finally, section three illustrates how limb developmental cell types are altered and influence morphology upon ectopic expression of chiropatagium markers in transgenic mouse limbs.

7.1. Comparing mouse and bat limb development at single-cell resolution

To study how developmental cell states are altered to drive emergence of novel morphological phenotypes, we used bat wing development as a model system. The bat forelimb exhibits various skeletal and tissue adaptations throughout its whole forelimb, which become most prominent in the distal autopodial segment. A central feature of wing morphology is the chiropatagium tissue, connecting the elongated digits and building the wing membrane. The chiropatagium beautifully exemplifies a morphological structure unique to the order Chiroptera. From an evolutionary perspective, the presence of this unique tissue proposes different cellular mechanisms for its emergence: (1) the innovation of a novel cell type, (2) the persistence of a cell state that is removed in other species or (3) the repurposing of an existing limb cell state at a different anatomical position. Importantly, profiling of its development on a molecular level allows the comparison to a model species of limb development such as mice, where the interdigital tissue is removed. This comparative approach allows the identification of conserved and divergent features of limb development, potentially indicative of the molecular origin of this novel trait. Crucially, the embryonic stages during which digit separation in mice and tissue retention in bats occur have been defined and can therefore confidently be mapped and aligned for comparison. Furthermore, the chiropatagium tissue is absent in the hindlimbs of bats, which therefore provide an ideal internal control as well as an informative morphological contrast. Altogether, these features make the chiropatagium a well-suited tissue to study morphological innovation.

In this first section an integrated single-cell RNA-seq atlas of mouse and bat embryonic limbs was generated to characterize developmental cell populations. We specifically focused on an interdigital cell population to investigate similarities and differences during the process of tissue regression. It was found that limb developmental cell populations, including the interdigital

mesenchyme, are of high similarity in both species, both at the level of gene expression and cell proportion. Furthermore, in contrast to previous hypotheses proposing interdigital apoptosis inhibition as source of retained wing tissue (Weatherbee et al., 2006), we observed active cell death in the interdigital regions of bat forelimbs. This demonstrates that cell death inhibition in interdigital mesenchymal cells is not the major driver of chiropatagium development, suggesting a different cellular origin of this tissue.

7.1.1. Mouse and bat limbs share homologous cell populations during development

To characterize the molecular mechanisms driving phenotypic changes during embryogenesis, we aimed to compare bat and mouse limb differentiation at single-cell resolution during developmental stages where limb morphology is defined. This enables the characterization of similarities and differences in limb cell populations and their transcriptional profiles between both species. To do so, I generated single-cell transcriptome datasets from developing mouse fore- and hindlimbs at E11.5, E12.5 and E13.5. Furthermore, I generated scRNA-seq data from bat fore- and hindlimbs at CS15 (equivalent to E11.5) and CS17 (equivalent to E13.5) (Figure 11a and b). Limbs of both species were dissected at the base of the limb (E11.5/CS15) or at the elbow (E13.5/CS17). Importantly, the chosen developmental timepoints cover stages of both morphological similarity (E11.5/CS15) as well as divergence (E13.5/CS17), thereby capturing the manifestation of species-specific phenotypes, including the removal or retention of the interdigital tissue (van den Eijnde et al., 1997; Zakeri, Quaglino, & Ahuja, 1994).

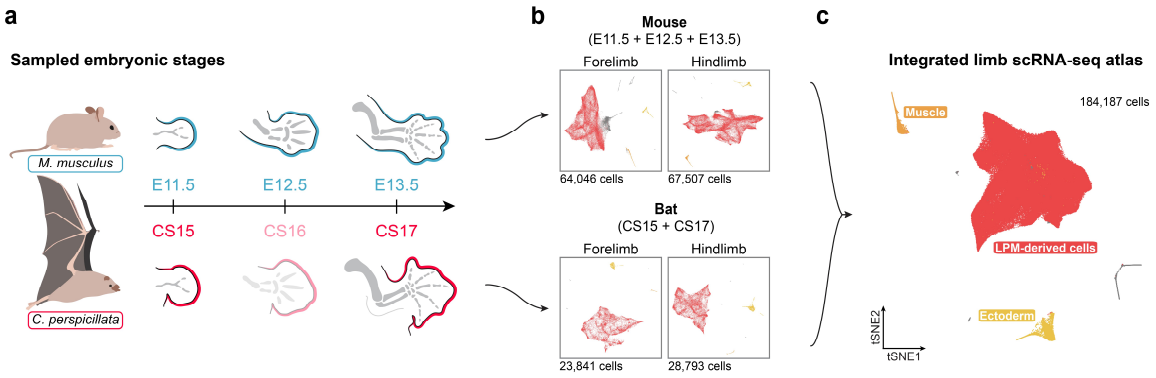

Figure 11: Single cell RNA sequencing from mouse and bat embryonic limbs identifies major known cell populations of the developing limb.

a Overview of the key embryonic stages of mouse (blue) and bat (red) limb development collected for scRNA-seq experiments this study (Mouse: E11.5, E12.5 and E13.5; Bat: CS15 and CS17). Shown is a schematic representation of forelimb morphology of equivalent mouse and bat stages. **b** Individual integrated tSNE plots of mouse and bat fore- and hindlimb cells. The colors indicate the major limb developmental cell populations (red: LPM-derived limb cells, orange: muscle cells, yellow: ectodermal cells). **c** tSNE plot of the integrated inter-species limb single cell atlas. The colors indicate the major limb developmental cell populations as in **b**.

Using the SCTtransform tool from Seurat, which has been successfully used to integrate scRNA-seq data from different species while preserving meaningful biological information

(Song, Miao, Brazma, & Papatheodorou, 2023), cells from both species, tissues and developmental stages were combined. This generated an integrated single-cell limb atlas containing 184,187 representative cells, enabling collective clustering, annotation and comparison of cell populations of high similarity (Figure 11c). Clustering of the integrated data resulted in distinct populations which can be annotated based on well-characterized marker gene expression. Importantly, the identified clusters represent the major cell populations known to be present in developing limbs, including lateral plate mesoderm-derived cells as well as cells of non-LPM origin, namely ectodermal and muscle cells (Feregrino et al., 2019; Feregrino & Tschopp, 2022; Markman et al., 2023).

These populations were further subdivided into distinct cell clusters which were annotated based on previously reported marker genes identified by performing differential gene expression analysis. This identified seven cell clusters for the muscle lineage, including myogenic progenitors, myoblasts, myocytes and myotubes, representing the central cell states of myogenesis (Figure 30, see section 10.1) (Bentzinger, Wang, & Rudnicki, 2012). Furthermore, nine clusters were determined for the limb ectoderm, including epidermal and dermal populations as well as basal keratinocytes and the *Fgf8*-positive AER population (Figure 30, see section 10.1).

As the LPM-derived cell population is the biggest contributor to limb developmental tissue at these stages, we chose to focus our following analyses on these cells. The 168,918 LPM-derived cells were subdivided into 18 subpopulations (Figure 12a). The identified subclusters recapitulated the major differentiation lineages of the developing limb. These include mesenchymal progenitor cells (red) marked by expression of *Msx1* and *Top2a*, chondrogenic cells (blue) marked by *Sox9* and *Col2a1* as well as fibroblast cells (green) defined by *Col3a1* and *Col1a1*, among other genes (Feregrino et al., 2019; Kelly et al., 2020; Markman et al., 2023). Interestingly, no major differences in gene expression of top marker genes used to define these clusters were identified between mouse and bat (Figure 12b, see also Figure 30). Furthermore, visualizing LPM-derived cells by their species origin revealed that cells from both species contributed equally to the clusters across the integrated datasets (Figure 12c). This indicates high transcriptional similarity of mouse and bat cells as well as their even distribution across LPM-derived clusters. Finally, the visualization of LPM-derived cells by developmental stage revealed that cells from earlier, less differentiated stages (E11.5/CS15) cluster together. In contrast, cells corresponding to later developmental stages (E13.5/CS17) predominantly are part of clusters which tend to separate from others, situated along the periphery (Figure 12d). Altogether, this suggests that our integrated limb atlas recapitulates limb development across different stages of limb morphogenesis. Furthermore, our data indicates that the identified cell populations and their defining marker genes are largely conserved during limb development between species, despite their diverse morphologies.

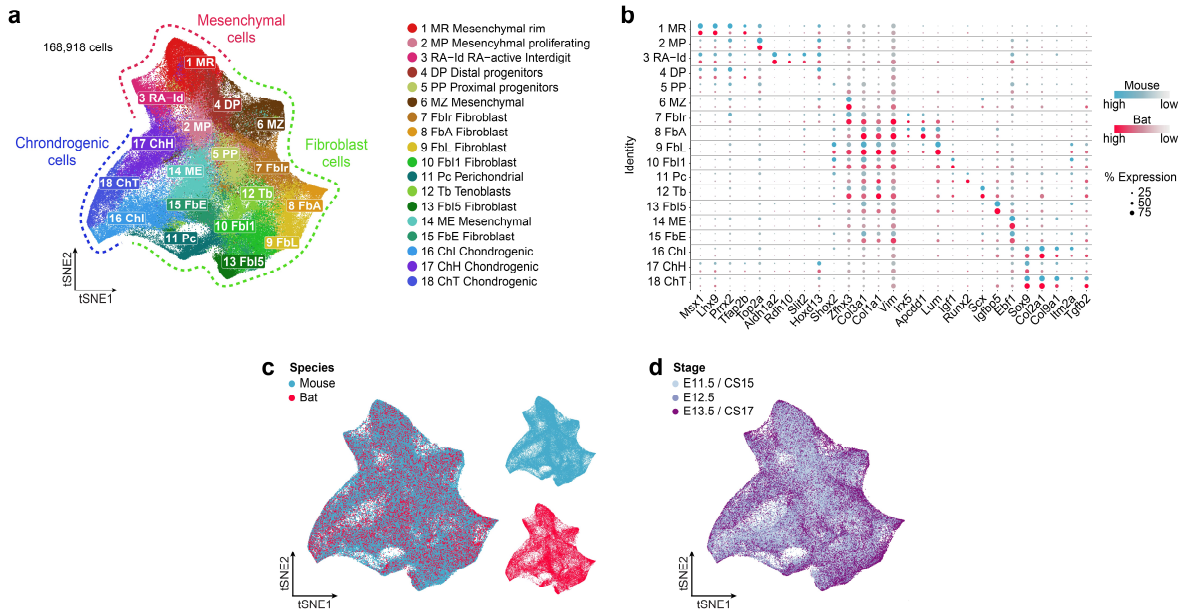

Figure 12: Sub-clustering of LPM-derived cells reveals homologous cell populations recapitulating the differentiation lineages of limb development.

a tSNE plot of sub-clustered integrated LPM-derived limb cells. Shown are the 18 identified clusters and their annotations. Dashed lines highlight the main differentiation lineages within LPM-derived cells (red: mesenchymal cells, green: fibroblast cells, blue: chondrogenic cells). **b** Dot plot showing marker gene expression used for integrated cluster annotation. Dot colors represent the species (blue: mouse, red: bat) and color intensity represents expression levels. The dot size represents the percentage of cells expressing the respective marker gene. **c** and **d** tSNE plot of LPM-derived limb cells showing cell contribution by species (c, mouse: blue, bat: red) and timepoint (d, light blue: E11.5/CS15, lavender: E12.5, purple: E13.5/CS17) to the integrated atlas.

7.1.2. Relative cell proportions reflect tissue differentiation during limb morphogenesis

Tissue differentiation during embryogenesis involves dynamic cellular processes leading to changes in cell proportions over developmental time (Markman et al., 2023). Furthermore, variations in cell proportions or compositions of distinct populations could impact developing morphological structures and, in an evolutionary context, could therefore be a result of species-specific differentiation programs (Huang et al., 2023). To investigate whether differences in morphology are reflected at cell proportion level, relative proportions for each identified cell cluster across developmental stages in fore- and hindlimbs of both species were calculated (Figure 13). At stages of high morphological similarity (E11.5/CS15), highly similar cell proportions in the main developmental lineages (~30% mesenchymal cells, ~50% fibroblast cells, ~20% chondrogenic cells) as well as distinct cell populations in both species and tissues were observed. During differentiation into a stage of more distinct morphology (E13.5/CS17) most cell populations exhibited significant changes in their proportions. Specifically, the relative amount of mesenchymal cells decreased, while a higher proportion of differentiated fibroblast cells is observed. The chondrogenic population underwent the least changes in all examined

samples. Overall, the changes in relative cell proportions showed highly similar trends in developing limbs of both species. This might thus reflect cell population changes during limb tissue differentiation rather than the emergence of species-specific, divergent morphologies. Therefore, these findings highlight a high similarity of mouse and bat limb cell populations, both at transcriptional level as well as relative proportion. Overall, these observations suggest that limbs of both species are constructed from conserved, highly similar cell populations, regardless of the resulting morphological phenotype.

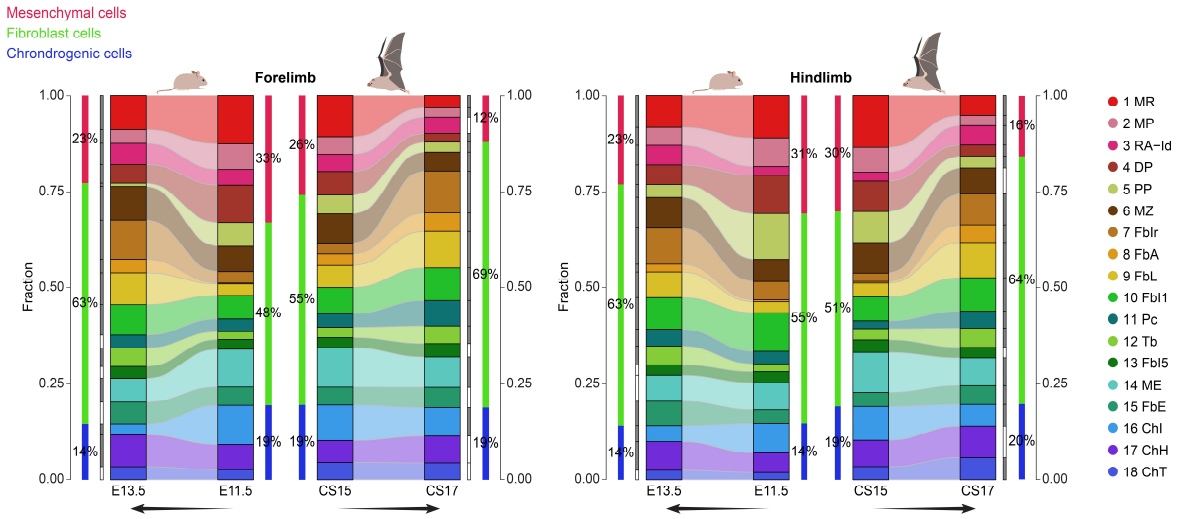

Figure 13: Relative cell proportions are highly similar between tissues and species and reflect differentiation dynamics during limb development.

Alluvial plots showing relative cell proportions of mouse and bat forelimb (left) and hindlimb (right) LPM-derived mesenchymal cell clusters between developmental stages E11.5/CS15 and E13.5/CS17. Bars on each side show the relative amount of cells in the main differentiation lineages (red: mesenchymal cells, green: fibroblast cells, blue: chondrogenic cells). A statistically significant change in relative cell proportion per cluster is indicated by a dark grey bar.

7.1.3. Comparable apoptosis patterns in mouse and bat forelimbs indicate an independent cellular origin for wing membrane formation

The generated limb single-cell atlas represents a powerful resource to compare gene expression signatures between species at cellular level. It therefore allows the characterization of differences potentially indicative of species-specific cell state changes leading to the development of distinct phenotypes, such as the wing membrane. Our analyses did not reveal evidence of a novel cell state responsible for the emergence of this trait. Therefore, we sought to explore the possibility that the wing membrane may arise from the persistence of a cell state that is removed in species with free digits.

Indeed, several studies have hypothesized that the presence of the chiropatagium tissue is a result of failed interdigital tissue regression through the inhibition of apoptosis (Hockman et al., 2008; Z. Wang et al., 2014; Weatherbee et al., 2006). Importantly, this process has been characterized in model systems like mouse and chick limbs and involves the interplay of

several components. The major regulators of interdigital cell death include members of the BMP family (especially BMP2 and BMP7) as well as retinoic acid and the enzymes involved in its metabolism, such as ALDH1A2 and RDH10 (Hernández-Martínez et al., 2009). Therefore, we next aimed to compare apoptotic processes between limbs of both species and map interdigital cell death at cellular resolution. To identify the cells involved in and most likely affected by interdigital cell death, the expression of BMP and RA signaling components were mapped in our integrated LPM-derived cells. This revealed the retinoic-acid active interdigital mesenchymal cells (cluster 3 RA-Id) as central population of apoptotic gene co-expression (Figure 14). Importantly, these findings agree with previously characterized spatial expression profiles of these genes (G. Luo et al., 1995; Mason et al., 2015; Mueller, Hellmann, & Nickel, 2012; B. Zhang et al., 2023). Based on this, we identified this cell population as the signaling source for interdigital cell death in developing mouse and bat limbs, establishing it as the focus of further comparative analyses.

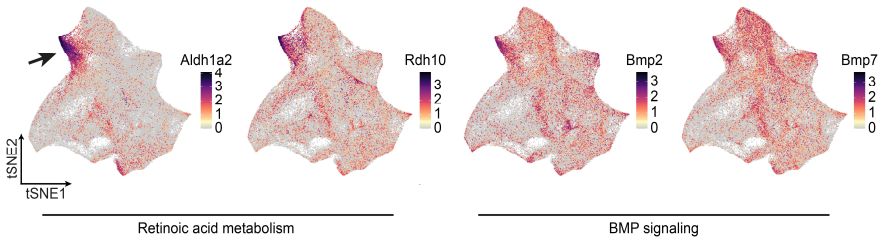

Figure 14: Expression of central interdigital cell death components highlights cells of the retinoic acid active interdigit population as central apoptotic signaling cluster.

tSNE plots of integrated LPM-derived limb cells showing the expression of essential components of retinoic acid metabolism (*Aldh1a2* and *Rdh10*) and BMP signaling (*Bmp2* and *Bmp7*) involved in the regulation of interdigital cell death. The arrow highlights the retinoic acid active interdigital cluster (3 RA-Id).

To characterize species-specific features related to apoptotic activity within this cluster, an extended set of genes associated with cell death during limb development was defined, including apoptosis through *Bcl2*, senescence, RA, BMP or FGF signaling (Dupé et al., 1999; Lorda-Diez et al., 2015; Pajni-Underwood, Wilson, Elder, Mishina, & Lewandoski, 2007). The expression of these genes was then compared within the homologous RA-Id cluster between both species by calculating differential gene expression between this cluster and the remaining LPM-derived cells (Figure 15a and d). To our great surprise, we observed no significant relative gene expression differences between mouse and bat limb tissues, suggesting no inhibition or reduction of apoptosis on a transcriptional level.

To further investigate the presence or absence of interdigital cell death in bat forelimb tissues, embryonic limbs were stained for two markers indicating apoptosis execution and lysosomal activity. For one, we used LysoTracker, a dye marking acidic subcellular compartments such as the lysosome, thereby indicating cell death (Fogel, Thein, & Mariani, 2012). Furthermore, we performed immunofluorescence staining for cleaved caspase 3, an effector protein

activated during the caspase cascade triggering the execution of apoptosis (Porter & Jänicke, 1999). These assays were performed with bat limbs at CS17, equivalent to E13.5 in mouse, when apoptosis of the interdigital tissue occurs (van den Eijnde et al., 1997; Zakeri et al., 1994). Intriguingly, we observed positive cell death and cleaved caspase 3 staining in the interdigital region of the bat forelimb at levels comparable to those found between digits I and II and in the hindlimb regions (Figure 15b, c, e and f), which get separated. The strongest signal in both assays was observed at the most distal interdigital regions, decreasing proximally. Importantly, structures positive for LysoTracker and cleaved caspase 3 staining were also detected throughout the remaining interdigital tissue.

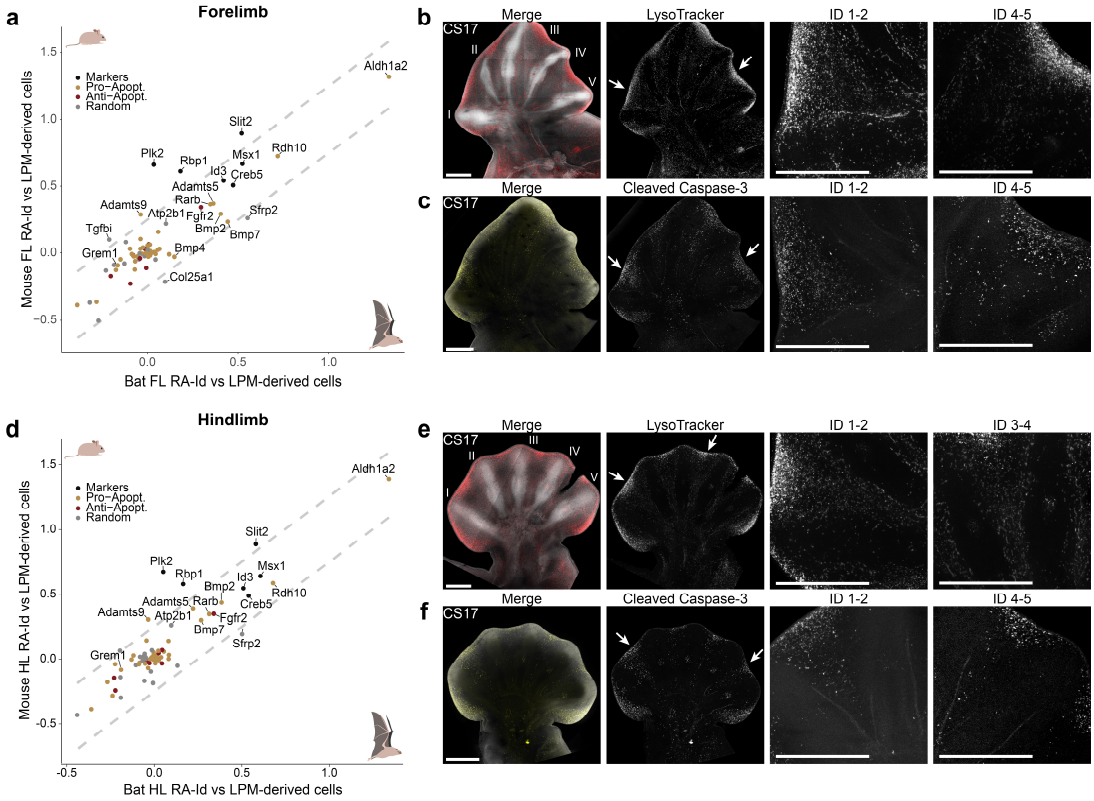

Figure 15: Gene expression and fluorescent microscopy analysis reveal apoptosis-induced interdigital cell death in bat forelimbs.

a and **d** Dot plots showing correlation of relative expression levels of genes involved in pro- (yellow) or anti-apoptotic (red) processes during limb development in the integrated retinoic acid active interdigital cluster (3 RA-Id) against the remaining LPM-derived mesenchymal cells of mouse and bat forelimb (a) and hindlimb (d) cells. Relative expression levels are depicted as log-fold change (lfc). Marker genes of this cluster are highlighted in dark blue, random genes included as control are highlighted in grey. Dashed lines indicate a deviation of ±0.25 of a linear regression of 1. Microscopy images in **b**, **c**, **e** and **f** show bat forelimb (b and c) and hindlimb (e and f) at CS17 stained with LysoTracker dye (b and e) and against cleaved caspase 3 (c and f). Shown are whole limb images as well as magnified interdigital regions between digit I and II as well as III and IV and IV and V. Merge: DAPI and LysoTracker/cleaved caspase 3. The scale bar represents 500 µm.

The obtained results revealed that apoptotic gene expression in the homologous retinoic acid active interdigital mesenchyme does not exhibit species-specific differences. Therefore, there is no indication of cell death inhibition or survival of cells of this cluster. This furthermore

coincides with the finding that cell death, mediated by a caspase-triggered apoptotic process, is present in the interdigital regions in both, bat fore- and hindlimbs. Altogether, these findings strongly suggest that the chiropatagium tissue is not formed through inhibition of apoptosis and resulting retention of retinoic-acid active mesenchymal cells. Therefore, we concluded that the chiropatagium wing tissue is of a different cellular origin and composition.

7.2. Investigating the cellular features of the chiropatagium

In the previous section, we have highlighted a high level of similarity between mouse and bat limb developmental cell populations and the processes they undergo despite their divergent morphologies. Therefore, we next aimed to investigate, how species-specific phenotypes, such as the chiropatagium, emerge from such highly similar cell populations. To do so, we first characterized the cellular composition and developmental origin of the chiropatagium cells. We furthermore investigated whether the transcriptional profile of these cells is unique to bat forelimbs or shares features with bat hindlimb or mouse limb tissues. With this, we discovered a cellular program of proximal fibroblasts that has been redeployed to the distal bat limb. Moreover, we mapped underlying regulatory features contributing to the gene regulation in the distal bat limb, thereby identifying central regulatory components of the wing fibroblast cell program.

7.2.1. The chiropatagium is composed of *MEIS2*-positive fibroblast cells

In order to characterize species-specific features of limb development in our single-cell data, mouse and bat limb tissues were individually analyzed by separately clustering fore- and hindlimb cells of both species (Figure 16a, b, d and e). The independently identified cell populations recapitulated the previously determined clusters to a high degree and were thus annotated using our integrated data as a reference. Individual clusters were first categorized as mesenchymal, fibroblast or chondrogenic cluster and then labelled according to the identity of the majority of cells in the respective cluster in the integration. Consistent with our previous findings, we observed strong correlation of gene expression between species within those clusters, suggesting overall comparable cell identities of the individually identified populations (Figure 16c and f). The strongest correlation was observed between the clusters of each differentiation lineage, indicating distinct gene expression signatures between but also high transcriptional similarity within each lineage.

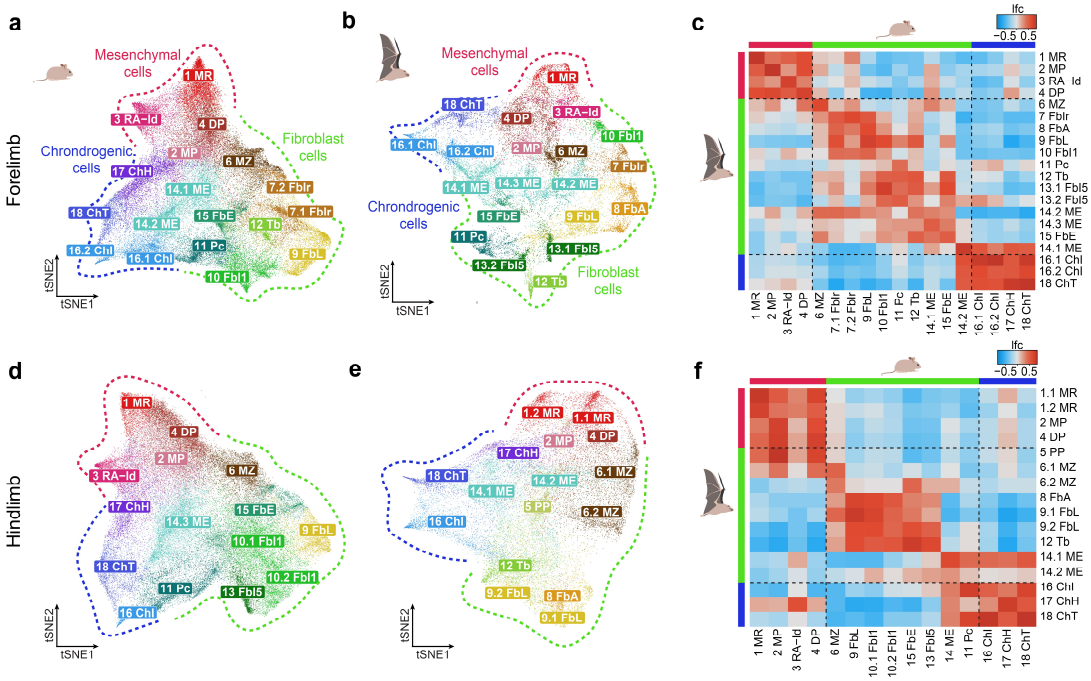

Figure 16: Individual clustering of species and tissue data confirms similarity of limb cell populations.

a, **b**, **d** and **e** show tSNE plots of individual clustering of mouse and bat LPM-derived fore- and hindlimb mesenchymal cells. The main developmental lineages are highlighted (red: mesenchymal cells, green: fibroblast cells, blue: chondrogenic cells). The cluster labels and colors are derived from the integrated cluster annotation from Figure 12 a. **c** and **f** Clustering correspondence between mouse and bat forelimbs (c) and hindlimbs (f). Shown is the Pearson correlation of the lfc value between the focus cluster and the remaining limb cells of the top 10 marker genes of all clusters. The main developmental lineages are highlighted as in a.

To specifically investigate the cellular composition of the bat chiropatagium, we microdissected this tissue from further developed bat limbs at CS18/CS19 and performed single-cell RNA-sequencing. The obtained cells were individually clustered and its LPM-derived cell populations were annotated by label transferring using the LPM-derived bat forelimb data as reference (Figure 17a and b). This process pairs each chiropatagium cell with its most transcriptionally similar counterpart in the reference dataset (Stuart et al., 2019; Stuart & Satija, 2019). Interestingly, this label transfer analysis revealed that the analyzed chiropatagium tissue mainly consist of fibroblast cells. Specifically, these cells show the highest transcriptional similarity to clusters 7 FbIr, cluster 8 FbA and cluster 10 FbI1 in the LPM-derived bat forelimb data of earlier stages (Figure 17c). Interestingly, and in agreement with our previous results, the retinoic acid-active interdigital cluster (3 RA-Id) was only minimally represented in the chiropatagium (~1%). This furthermore confirmed that these cells are not persisting to form the wing membrane tissue.

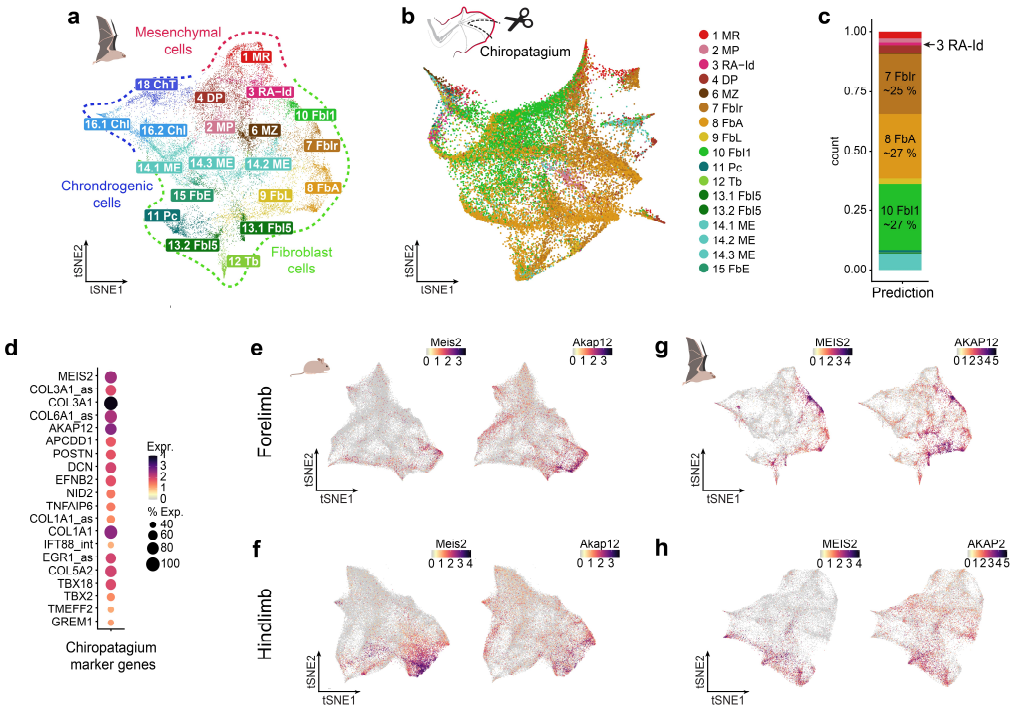

Figure 17: Microdissection of the bat chiropatagium identifies its cellular and transcriptional composition and reveals *MEIS2* and fibroblast components as cellular markers.

a tSNE plot of individual clustering of bat LPM-derived forelimb cells (same as in Figure 16b). **b** tSNE plot of the cells derived from the mesenchymal portion of microdissected chiropatagium tissue from CS18/19 bat forelimbs. Cells are colored and labelled by their transcriptional correspondence to the bat LPM-derived forelimb dataset. **c** Quantification of cluster correspondence between chiropatagium cells and LPM-derived forelimb cells from b. The three major contributing clusters (7 Fblr, 8 FbA and 10 FbI1) as well as the RA-Id cluster are highlighted. **d** Dot plot showing marker genes of chiropatagium cells based on differential gene expression between chiropatagium cells and forelimb LPM-derived mesenchymal cells. Dot color represents normalized expression levels and dot size indicates the percentage of cells expression the respective marker gene. **e**-**h** tSNE plots showing the expression pattern of the chiropatagium marker genes *MEIS2* and *AKAP12* in mouse forelimbs (e) and hindlimbs (f) as well as bat forelimbs (g) and hindlimbs (h).

Next, the distinct transcriptional profile of the chiropatagium tissue was examined by analyzing differential gene expression between these cells and the LPM-derived cells of the bat forelimb (Figure 17d). We observed high expression of several collagens (*COL3A1*, *COL1A1*, *COL5A1*) as well as factors associated with extracellular matrix organization such as *AKAP12* and *DCN*, confirming the fibroblast identity of these cells (H. Li, 2022; Muhl et al., 2020). This analysis moreover revealed chiropatagium enriched expression of the transcription factor *MEIS2* as well as the BMP inhibitor *GREM1*, two factors that have previously been described to be expressed in the bat wing (Dai et al., 2014; Mason et al., 2015; Weatherbee et al., 2006). Notably, the marker genes identified as defining chiropatagium cell identity were not wing- or species-specific but also showed expression in the developing bat hindlimb as well as in both mouse limbs, however at a different location, as exemplarily shown for *MEIS2* and *AKAP12* (Figure 17e-h). Taken together, these results reveal the transcriptional identity of the cells constituting the chiropatagium tissue. Importantly, these cells display a fibroblast gene expression signature composed of genes expressed in the limb of both species. Therefore,

our observations do not identify a unique species- and wing-specific expression pattern in this tissue.

7.2.2. Chiropatagium fibroblasts originate from an independent developmental trajectory in the distal bat limb

To further characterize species-specific features of the chiropatagium fibroblast cells, we aimed to compare their developmental origin and differentiation. To do so, cell trajectories in mouse and bat distal limbs were analyzed, specifically focusing on distal non-skeletal cells expressing *Hoxd13*, a well-characterized marker of the autopodial lineage in the developing limb (Desanlis et al., 2020; Dollé, Izpisúa-Belmonte, Brown, Tickle, & Duboule, 1993; Morgan & Tabin, 1993). By conducting RNA velocity and pseudotime analyses of the selected distal populations, their most likely cell-state transitions were determined (Figure 18a-d). Consistent with previously reported cell differentiation patterns in the developing limb, this analysis highlighted distinct differentiation trajectories sharing a common putative origin (2 MP and 4 DP) (Feregrino et al., 2019; Markman et al., 2023). Each trajectory was furthermore defined by marker genes exhibiting increased expression over pseudotime (Figure 18a-d). Based on this we observed trajectories present in both species, such as autopodial progenitors marking the mesenchymal rim expressing *Tfap2b* as well as different types of non-skeletal connective tissue differentiation lineages marked by *Lum* or *Zfhx3* (Feregrino et al., 2019; Markman et al., 2023). Moreover, the retinoic acid active interdigital cell population marked by increasing expression of *Aldh1a2* was also detected as a distinct cell trajectory in both species. In line with our previous results, the cells forming this trajectory co-expressed genes involved in the interdigital apoptotic process, such as *Bmp2* and *Bmp7*, in both mouse and bat distal limbs (Figure 18a-d). Intriguingly, this furthermore identified a distinct cell trajectory of fibroblast cells in the distal bat forelimb marked by increasing expression of the homeobox transcription factor *MEIS2* (Figure 18b). This trajectory was not detected in mouse limbs and was absent in bat hindlimbs, suggesting an independent cellular differentiation path unique to the bat forelimb. Additionally, cells forming this trajectory expressed *GREM1*, but did not exhibit expression of *ALDH1A2*, *BMP2* or *BMP7*. Therefore, based on the transcriptional profile shared with the previously characterized chiropatagium fibroblasts, these cells can be assumed to mark the cellular origin of this tissue. Altogether, these results demonstrate that the mesenchymal portion of the bat chiropatagium develops independently and distinctly from the retinoic acid active interdigital cell population and is primarily composed of fibroblast cells highly expressing *MEIS2*.

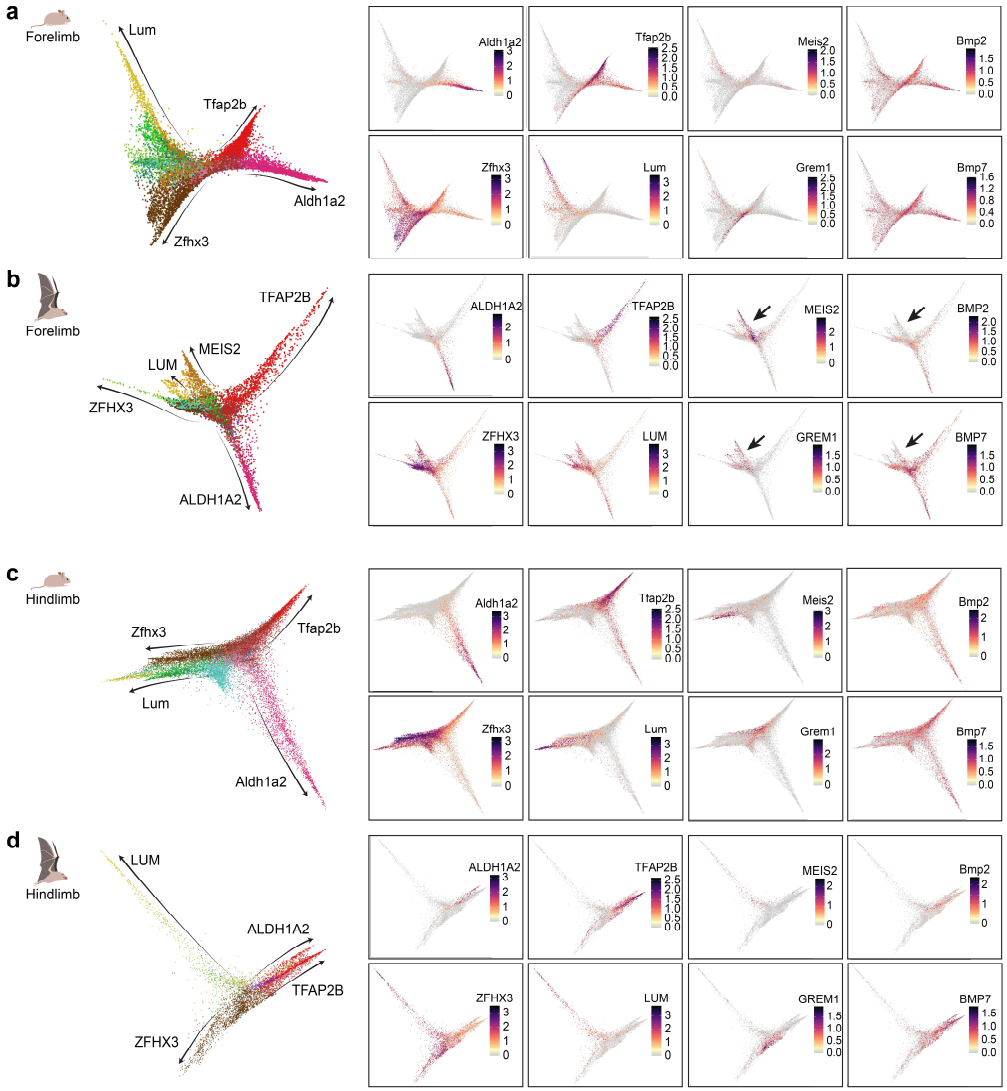

Figure 18: Differentiation trajectories of distal mouse and bat limbs suggest independent identities and developmental paths of RA-Id and *MEIS2*-positive fibroblast cells.

a-**d** Differentiation trajectories from *Hoxd13*-positive, non-chondrogenic mouse and bat forelimb (a and b) and hindlimb (b and c) cells derived from RNA velocity and pseudotime analysis. Trajectories are indicated by arrows and were annotated based on increasing expression of marker genes (*Aldh1a2*, *Lum*, *Tfap2b* and *Zfhx3*). The colors in the left panels correspond to cluster colors from Figure 12a and Figure 16a, b, d and e. Note the additional trajectory marked by *MEIS2* expression in bat forelimbs. Furthermore, shown is the expression of marker genes (*Aldh1a2*, *Lum*, *Tfap2b*, *Zfhx3*) as well as *Meis2*, *Bmp2*, *Grem1* and *Bmp7* in differentiation trajectories.

7.2.3. A proximal limb fibroblast cell program is repurposed in the distal bat wing

Consistent with previous studies, our findings identified *MEIS2* as a prevalent marker gene of the chiropatagium tissue (Dai et al., 2014; Mason et al., 2015). The functions of this transcription factor during limb development have mainly been studied during early developmental stages, where it has a crucial role in patterning by defining proximal identity of limb segments (Delgado et al., 2021; López-Delgado, Delgado, Cadenas, Sánchez-Cabo, & Torres, 2021). However, as our analyses highlighted a species-specific distal expression

pattern of this gene during later stages of limb development, we aimed to characterize the role of *MEIS2* in autopod differentiation and morphogenesis.

To distinguish autopodial from proximal cells within our single-cell transcriptomic data, the proximo-distal axis was first defined by assigning a distal or proximal identity to each cell of the four datasets (Figure 19a). This assignment was based on calculating the gene expression ratio between the distal markers *Hoxd13* and *Msx1* and the proximal marker *Shox2* (Desanlis et al., 2020; Dollé, Izpisúa-Belmonte, et al., 1993; Markman et al., 2023) (Figure 19b). Most limb clusters could be unambiguously identified as either proximal or distal, with only a few exceptions due to similar marker expression levels. Next, the co-expression of *Hoxd13* and *Meis2* was assessed to specifically identify *Meis2*-positive cells in the distal autopod (Figure 19c). The highest fraction (16.4 %) of cells co-expressing *Hoxd13* and *Meis2* was observed in the bat forelimb, most of which belonged to fibroblast cluster 10 FbI1 (Figure 19c, indicated with an arrow). We also detected co-expression of these genes in cluster 7 FbIr, although at lower cellular fraction. Our previous results showed that the cells within these clusters exhibit significant transcriptional similarity to chiropatagium tissue and contribute to a differentiation trajectory unique to the bat wing forelimb. Therefore, this observation not only confirmed these findings but furthermore suggested these cells as cellular component of chiropatagium development in the distal bat forelimb.

To characterize the transcriptional profile of these distal *MEIS2*-positive cells within bat cluster 10 more thoroughly, differential gene expression analysis was performed against the remaining LPM-derived bat forelimb cells. This identified 20 marker genes of this population (Figure 19d). These include markers of fibroblast differentiation (*AKAP12*, *COL3A1*, *DCN*) (Muhl et al., 2020) and members of the T-box (Tbx) transcription factor family (*TBX18* and *TBX3*), among others. To assess whether this gene set represents a species- or tissue-specific cellular feature, its co-expression was analyzed in fore- and hindlimb cells of both species. Intriguingly, we found this set of genes highly co-expressed in proximal fibroblast cells of all four samples (mostly clusters 8 FbA and 9 FbL), while its distinct distal expression was unique to the bat forelimb (Figure 19e). Notably, highly similar results were found for the marker genes of cluster 7 (Figure 31, see section 10.1).

In summary, our findings identified a fibroblast cell cluster marked by high expression of *MEIS2*, among other genes, which forms a distinct developmental trajectory unique to the distal bat forelimb. These cells highly transcriptionally resemble the distal chiropatagium tissue, as well as fibroblast cells present in the proximal portion of fore- and hindlimbs of both species. We therefore hypothesized that a proximal fibroblast cell state and its gene expression program were repurposed in the distal bat limb, where it contributes to the formation of the wing membrane (schematically depicted in Figure 19f).

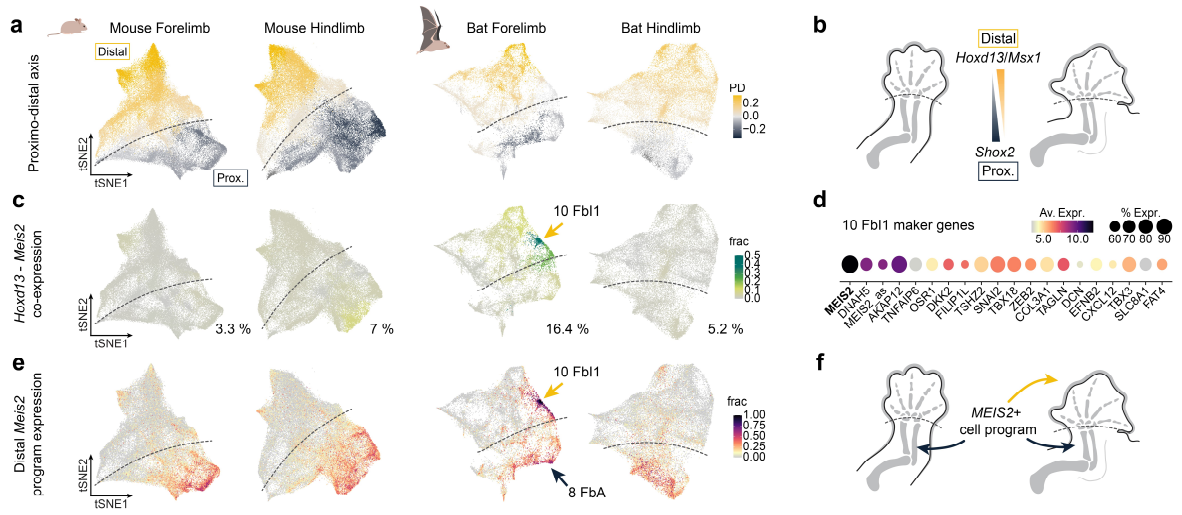

Figure 19: *MEIS2*-positive fibroblasts of the distal bat limb share gene expression signatures with proximal fibroblast cells common to both species.

a Assignment of a proximal (dark blue) or distal (yellow) identity to each cell of mouse and bat fore- and hindlimb LPM-derived mesenchymal datasets based on the ratio between *Hoxd13* and *Msx1* (distal) and *Shox2* (proximal) expression per cell. Shown are the mean expression scores per cluster. **b** Schematic representation of distal and proximal markers used in a. **c** Co-expression of *Hoxd13* and *Meis2* in LPM-derived mesenchymal fore- and hindlimb datasets. Shown is the fraction of cells co-expressing both genes. Cells in the bat forelimb dataset highly co-expressing *HOXD13* and *MEIS2* (green) to a large degree overlap with cluster 10 FbI1 and are highlighted with an arrow. **d** Dot plot showing marker genes of cluster 10 FbI1 calculated based on differential gene expression between this cluster and the remaining cells of the bat forelimb LPM-derived mesenchyme. Dot color intensity represents expression levels and dot size indicates the percentage of cells expression the respective marker gene. **e** Co-expression of FbI1 marker genes in mouse and bat fore- and hindlimbs. Shown is the fraction of cells co-expressing the gene set from d. Distal and proximal cells in the bat forelimb highly co-expressing this program are highlighted with an arrow. **f** Schematic representation of differences in spatial expression of the identified *MEIS2*-positive cell program common to mouse and bat proximal fibroblast cells, while its distal expression is wing-specific.

To test this hypothesis and further characterize the level of similarity between the shared proximal and wing forelimb-specific fibroblast cells, differentially expressed genes were identified in the proximal (mainly 8 FbA) or distal (mainly 10 FbI1) bat fibroblast population against the remaining non-fibroblast LPM-derived cells and compared the enriched genes between both populations (Figure 20a). This resulted in 169 and 208 genes significantly enriched in the distal or proximal fibroblast cell population, respectively (Figure 20c). Interestingly, 144 (~62%) of these genes exhibited high relative expression in both clusters, indicating a high transcriptional similarity between the investigated populations and thus suggesting a repurposing of proximal fibroblasts during bat wing development. 25 genes (~11%) were exclusively expressed in the distal fibroblast cells, while 64 genes (27%) were exclusive to the proximal fibroblast cluster. The expression pattern of a representative gene for each group (distal, proximal, shared) is shown in Figure 20b.

To further assess the level of evolutionary conservation of this fibroblast cell program, the transcriptional similarity with proximal fibroblasts of mouse developing limb was examined. Interestingly, a subset of the differentially expressed genes present in both bat clusters was

shared with mouse proximal fibroblasts (34 genes, ~24%) (Figure 20a and c). These genes include fibroblast markers such several collagens, *Fbln1*, *Lum* and *Dcn*, amongst others (Muhl et al., 2020) and might therefore indicate an evolutionary conserved function of this gene set in limb fibroblast cell populations. To further evaluate possible cellular functions this shared gene set is involved in, we performed gene ontology (GO) enrichment analysis (Figure 20d). The most highly enriched GO terms involved processes related to extracellular matrix and structure organization, mesenchymal cell proliferation and ameboidal type cell migration, all of which can be considered essential components of fibroblast cell identity and function (Plikus et al., 2021). It can therefore be speculated that these fibroblast processes serve as crucial functions for interdigital mesenchyme remodeling during wing development.

Altogether, our findings showed that the gene expression profile of the identified distal *MEIS2*-positive cell state in the bat forelimb to a large degree overlaps with that of a proximal fibroblast population in mouse and bat limbs, while also exhibiting unique gene expression characteristics. This highlights a common and conserved fibroblast cell identity but importantly also reflects its dual spatial identity within the developing wing.

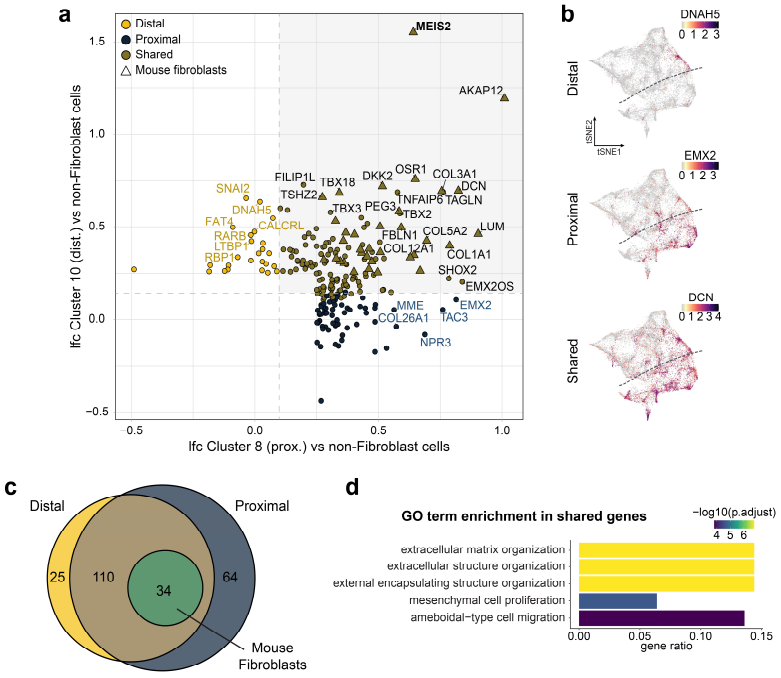

Figure 20: Distal and proximal bat fibroblasts share a cell program involved in central features of fibroblasts identity and function.

a Correlation between differentially expressed genes from distal (10 FbI1) and proximal (8 FbA) *MEIS2*-positive clusters in bat forelimb data identified in Figure 19. Shown is the log-fold change (lfc) of differentially expressed genes between the respective cluster and remaining LPM-derived cells of non-fibroblast identity. Yellow: genes expressed in the distal cluster; dark blue: genes expressed in proximal cluster; brown: genes expressed in both clusters. Genes shared with mouse fibroblasts are depicted as triangles. **b** tSNE plots of representative genes expressed in the distal (*DNAH5*), proximal (*EMX2*) or both clusters (*DCN*) of the bat forelimb. **c** Venn diagram showing the gene number and their overlap (brown) between genes enriched in the distal (yellow) and proximal (dark blue) fibroblast clusters. The genes shared with mouse fibroblast are depicted in green. **d** Top 5 enriched GO terms of the 144 shared genes between the distal and proximal cluster.

7.2.4. Regulatory features of the distal bat limb identify MEIS2 and TBX3 as central transcription factors of the wing network

The previous sections have characterized a wing specific distal fibroblast cell population as well as the transcriptomic profile defining its spatial and functional identity. To further elucidate the central components driving this cell program, we next aimed to identify the regulatory relationships between the identified gene markers. To do so, a gene regulatory network analysis was performed for each bat forelimb cluster using the SCENIC tool (Van de Sande et al., 2020). This analysis identifies co-expression modules to infer potential regulatory relationships between genes based on their expression patterns across single cells. To determine which transcription factors are likely regulating which genes in each module, enriched motifs in the regions of the genes within each module are identified. This establishes a regulatory connection and hierarchy between expressed genes within each cell cluster.

Interestingly, this analysis placed *MEIS2* as the regulon with the highest positive regulatory score within bat cluster 10 FbI1 (RSS >0.23, see Figure 32 in section 10.1), suggesting a significant regulatory function of this TF within these cells. Consistent with this observation, network representation of cluster 10 FbI1 further highlights MEIS2 as a central node within this cluster, exhibiting both direct and indirect connections to several components involved in the previously identified gene program (e.g. AKAP12, RARB, SNAI2) (Figure 21). Notably, these connections include multiple other transcription factors, further underscoring its prominent position within the network hierarchy. MEIS2 furthermore displays negative connectivity with chondrogenic markers such as SOX9 and FOXC1, reinforcing the non-chondrogenic identity of these cells (Akiyama, Chaboissier, Martin, Schedl, & de Crombrugghe, 2002; Almubarak et al., 2024). Altogether, these results suggest MEIS2 as a central transcription factor of the regulatory architecture in this cell populations and thus a potential driver of the chiropatagium cell program.

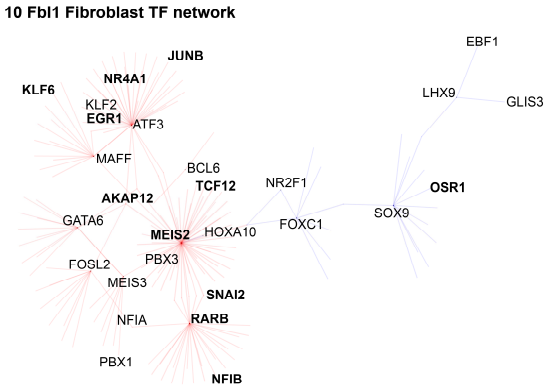

Figure 21: Gene regulatory networks highlight central cluster-specific transcription factors. SCENIC network.

SCENIC transcription factor network for genes enriched in the distal *MEIS2*-positive cluster 10 FbI1. Positive and negative regulatory connections are represented by red and blue lines, respectively. Genes from the previously described *MEIS2*-positive cell program are highlighted in bold.

To further elucidate transcriptomic and regulatory relationships in the distal limb, I generated additional bulk transcriptomic and functional genomics datasets. Importantly, I specifically dissected distal mouse and bat fore- and hindlimbs at a later developmental stage (E15.5 and CS18/19), cutting at the metacarpals in the autopod.

First, differential gene expression between the distal forelimb and hindlimb was performed to highlight tissue specific gene expression features of distal limbs in both species using the informative morphological contrast between these tissues (Figure 22a and b). Interestingly, this revealed only small transcriptional differences between distal mouse limbs, the most prominent ones being *Tbx5* as well as *Tbx4* and *Pitx1*, specific markers of fore- and hindlimb development, respectively (Figure 22a) (M. Logan & Tabin, 1999; Rodriguez-Esteban et al., 1999). In contrary, bat distal fore- and hindlimbs exhibited a myriad of additional DEGs (Figure 22b). Importantly, this revealed several components of the distal fibroblast cell program upregulated in the distal forelimb, including *MEIS2*, *TBX3*, *DNAH5* and *TBX18*, confirming the distal identity of these genes in the bat wing, even at later developmental stages. Furthermore, we observed several members of the HOX family of transcription factors differentially expressed in the distal part of the forelimb, such as *HOXD9*, *HOXD10* and *HOXA2*. Interestingly, these TFs have been described to play a crucial role during early limb development, especially in the patterning of limb structures (Pineault & Wellik, 2014; Zakany & Duboule, 2007). Importantly, these findings highlight higher transcriptional diversity with emerging morphological differences, suggesting a broad range of gene expression profiles - including several developmental TFs - contributing to the development of wing phenotypes. Moreover, these observations confirm the distal identity of previously determined factors, such as *MEIS2* and TBX family members, in the wing.

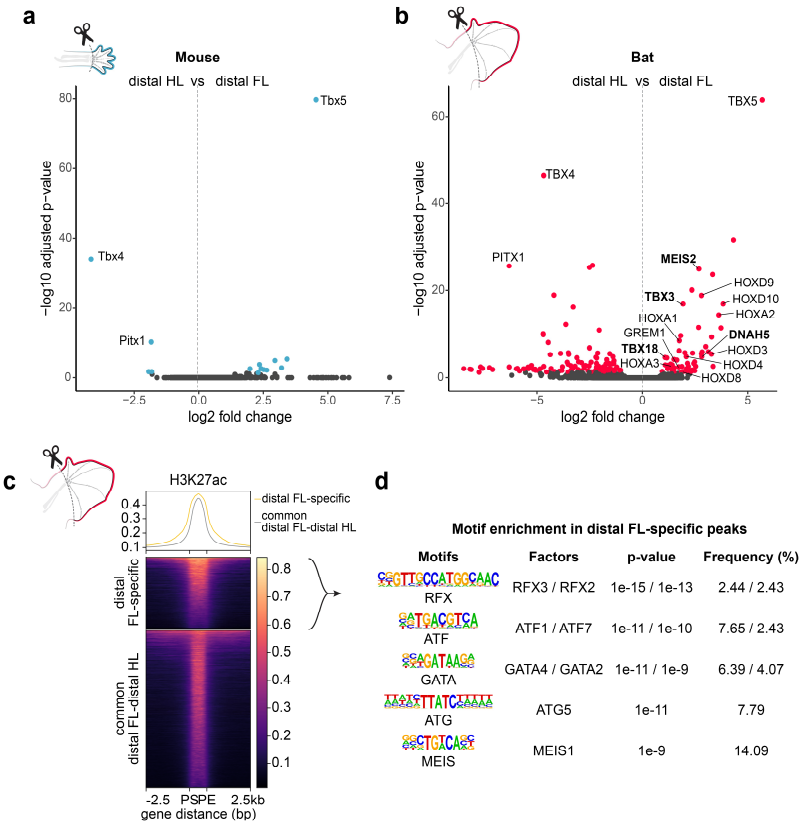

Figure 22: Transcriptomic and epigenetic features of the distal bat limb suggest developmental patterning genes and transcription factors contributing to wing formation.

a and **b** Volcano plot showing differentially expressed genes between distal forelimbs and hindlimbs in mice (blue, a) and bats (red, b) at E15.5 and CS18/19. Genes also present in the *MEIS2*-positive cell program from Figure 20 are highlighted in bold. **c** Tornado plot showing H3K27ac peaks intersected with open chromatin regions specific to the distal bat forelimb and peaks common to distal forelimb and hindlimb. Shown are the regions from peak start (PS) to peak end (PE). **d** Motif enrichment analysis in the distal forelimb specific peaks. Shown are the top 5 enriched binding sites identified with HOMER2 with p-value and frequency per transcription factor family.

Next, I generated ChIP-seq data from distally dissected fore- and hindlimbs to profile the histone mark H3K27ac as indicator of active cis-regulatory elements. To identify key regulators potentially driving wing-related gene expression, genome-wide distal forelimb-specific differential acetylation peaks were analyzed to find enriched transcription factor binding motifs within these regions (Figure 22c). The top enriched binding motifs suggested binding of transcription factors with crucial roles during embryogenesis, such as member of the RFX and GATA family (Figure 22d). Members of these TF families have, for instance, been shown to be involved in patterning processes or nerve migration in the developing limb (Feenstra et al., 2012; Hayashi et al., 2016; Kozhemyakina, Ionescu, & Lassar, 2014). Intriguingly, this analysis furthermore revealed MEIS binding motif enriched in the differentially acetylated regions (Figure 22d). This is in agreement with our previous observations and further suggest a central function of this TF in driving wing specific gene network.

To specifically investigate the function of MEIS2 in the distal bat forelimb, a functional genomics approach was applied to define its potential downstream regulatory relationships. Transcription

factors are known to bind not only to promoters but also to enhancers distributed across gene regulatory domains, or TADs (Koji & Takashi, 2024; Robson et al., 2019). To identify these regulatory domains, I generated Hi-C maps from bat limbs to visualize 3D chromatin interactions and identify topologically associated domains (TADs). Next, to determine putative binding targets of MEIS2, I conducted ChIP-seq for MEIS TFs on distally dissected bat limbs using a dual antibody assay (Delgado et al., 2021) detecting the binding of MEIS1 and MEIS2. Notably, both TFs recognize the same binding motif (Schulte & Geerts, 2019) and have been reported to have redundant functions during limb development (Delgado et al., 2021). Furthermore, MEIS1 is only minimally expressed in the distal bat forelimb (Figure 33, see section 10.1). It was therefore inferred that the majority of the detected signal originates from binding of MEIS2.

Regulatory domains were then searched for enrichment with MEIS2 binding sites and active promoter and enhancer marks (H3K27ac) by analyzing the respective ChIP-seq signal coverage per TAD (Figure 23a). This analysis revealed 656 TADs enriched in distal forelimb-specific H3K27ac coverage and 257 TADs distinctly enriched in MEIS2 binding signal. Of these, 173 TADs, encompassing 1885 genes, were enriched in both features. Intersecting the genes within these domains with the ones from the previously identified wing fibroblast cell program highlighted 71 genes. Notably, the 20 genes with the highest MEIS2 signal in their regulatory domains showed several ECM components (*COL3A1*, *COL5A1*, *DCN*, *LUM*) and developmental transcription factors, including several from the T-box TF family (*TBX3*, *TBX2*, *TBX18*) (Figure 23b). Of those, the *TBX3* domain exhibited the highest distal forelimb-specific acetylation coverage. The striking pattern of MEIS2 binding and H3K27ac coverage in the distal forelimb indicating its activity and regulatory relationship with MEIS2 is exemplarily shown for this gene locus (Figure 23c). Interestingly, *TBX3* has previously been suggested to play a role in bat wing development (Dai et al., 2014; Z. Wang et al., 2014), which is further supported by our observations.

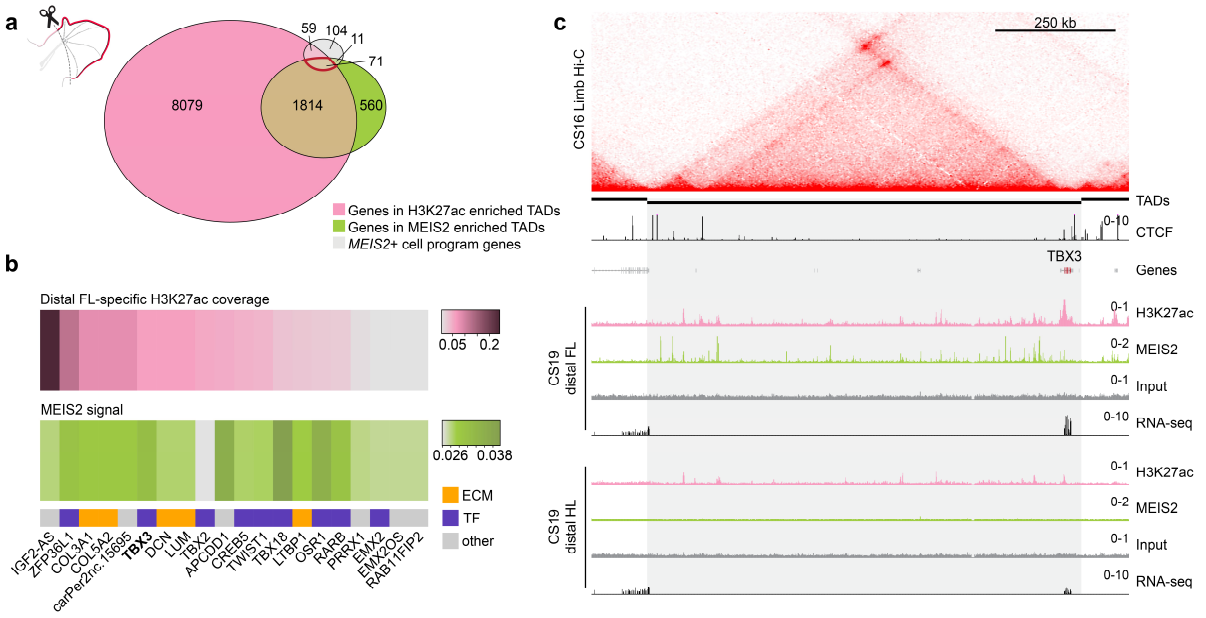

Figure 23: Distal forelimb specific H3K27ac and MEIS2 binding signatures reveal potential regulatory components of the MEIS2 network.

a Venn diagram showing the overlap between genes in H3K27ac enriched (pink) and MEIS2-binding enriched (green) TADs and genes expressed in the previously characterized *MEIS2*-positive fibroblast cell program (grey, see Figure 20) **b** Heatmaps showing the portion of each TAD covered by H3K27ac peaks (pink) and the mean signal per TAD of MEIS2 binding (green). Shown are the top 20 genes by MEIS2-binding signal ordered by unique acetylation peak coverage. Genes are classified by their function in extracellular matrix organization (orange) or as transcription factor (purple). **c** *TBX3* regulatory domain identified based on Hi-C. Shown is Hi-C based TAD calling as well as CTCF ChIP-seq from whole bat forelimbs at CS19. Furthermore, H3K27ac and MEIS2 ChIP-seq with input track and RNA-seq from distal bat forelimbs at CS18/19 are shown.

As previously mentioned, the functions of MEIS TFs have mainly been studied during early limb development. Therefore, we lastly aimed to elucidate whether MEIS2 regulates a similar set of genes at early and late stages of limb development and thus activates a common downstream program. To test this, bat MEIS2 binding patterns in gene promoters in the distal wing at CS18/19 were compared with MEIS binding in mouse embryonic limbs at E10.5 (Delgado et al., 2021), when these TFs have a crucial function in patterning of the limb. While 127 MEIS bound promoters in E10.5 mouse forelimbs were identified, 266 promoters were bound by MEIS2 in distal bat forelimbs at later stages of development. Interestingly, we observed only little overlap (21 promoters) in bound promoter regions between both tissues (Figure 24). This strongly suggests that MEIS2 activates a different set of gene components and therefore likely drives different functions in both developmental stages.

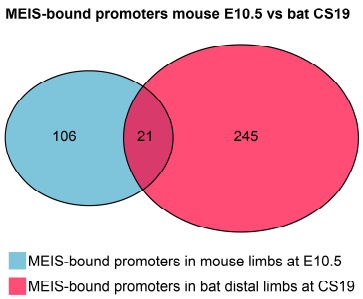

Figure 24: MEIS TFs bind a different set of gene promoters in early mouse (E10.5) and late bat (CS19) limb development.

Venn diagram showing the overlap of MEIS-bound promoters in mouse forelimbs at E10.5 (blue) and bat distal forelimbs at CS19 (orange).

Combined, these results not only confirm an essential role of MEIS2 in gene regulation shaping the bat distal forelimb and defining its wing fibroblast program, but also highlight other central factors of this network presumably contributing to the formation of the wing membrane. Specifically, our analyses reveal members of the TBX transcription factor family as potential regulators of the distal bat forelimb. Based on our observations and previous studies, we consider TBX3 as a central component playing a significant role in the wing-related regulatory network.

7.3. Distalization of *MEIS2* and *TBX3* expression in transgenic mice induces developmental processes related to wing formation

In the previous sections we have identified a wing specific distal fibroblast cell population and its cellular program, which is shared with common proximal limb fibroblast cells. We have furthermore revealed central gene regulatory components of the distal bat limb, suggesting an important role in driving the expression of this cell program. In the next section we therefore aimed to specifically assess the biological function of this distal fibroblast cell program. Using transgenic mouse models, we investigated how ectopic expression of two central regulatory factors of the wing fibroblast program, *MEIS2* and *TBX3*, influence mouse limb developmental cell types and limb morphology. On a cellular level we showed that ectopic distal expression of the respective components induces different parts of the distal fibroblast program and therefore recapitulates cellular bat fibroblast phenotypes. Morphological characterization of transgenic mouse limbs furthermore revealed the presence of essential features of bat wing development. These analyses provide proof that the identified wing fibroblast program and its central regulatory components play an important role in chiropatagium formation.

7.3.1. Distal expression of *MEIS2* and *TBX3* partially induces the transcriptomic profile of wing fibroblast cells

The previous results have identified a wing specific distal fibroblast cell population, its gene expression signature as well as its central regulators. In this section, we aimed to specifically assess the biological function of this program during limb development. For this, we used transgenic mouse models to express two central regulatory components of this program, the TFs *MEIS2* and *TBX3*, ectopically in the distal mouse limb, individually as well as in combination. As the protein sequences of both transcription factors are highly conserved between species (Figure 34, see section 10.1), I generated transgenic constructs where the expression of *MEIS2* or *TBX3* bat coding sequences are driven by a previously characterized *Bmp2* enhancer (Dathe et al., 2009) (Figure 25a). Importantly, this enhancer exhibits specific activity in the distal, non-skeletal region of mouse limbs. I introduced the expression constructs into the *H11* and *Rosa26* loci by site-specific insertions using the CRISPR/Cas9 system. These genome loci are generally considered safe harbor loci, meaning they allow for intended transgene expression but their insertion does not perturb surrounding endogenous gene activities and functions (Hippenmeyer et al., 2010; Ma et al., 2017; Tasic et al., 2011; Zambrowicz et al., 1997). I verified the specific spatial activity of our constructs in the distal limb of mutant mice at E12.5 by exemplarily visualizing the expression of transgenic *MEIS2* through whole mount in situ hybridization (WISH) (Figure 25b). Furthermore, I validated transgene expression by bulk RNA-seq analysis of dissected wildtype and mutant distal limbs at E12.5 (Figure 25c).

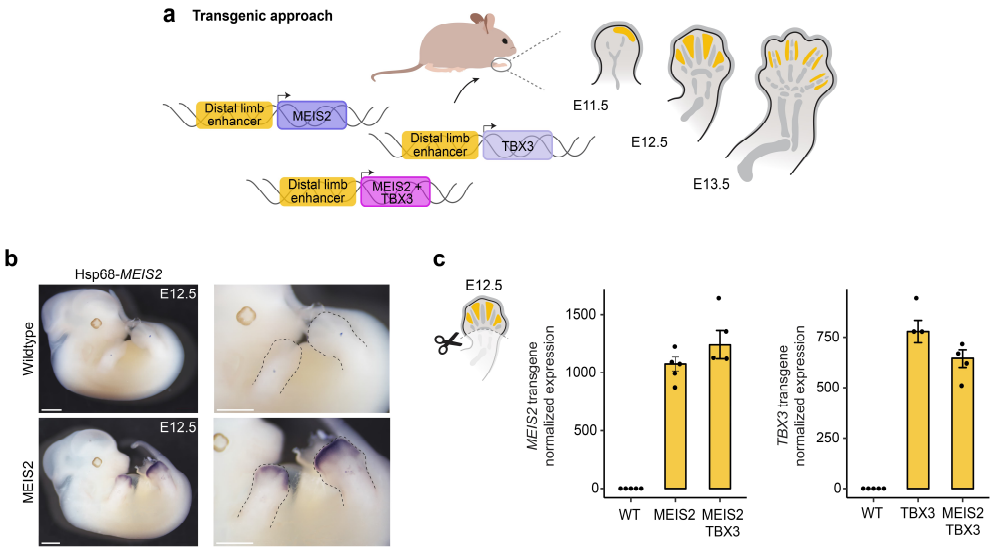

Figure 25: A transgenic system to induce ectopic distal expression of *MEIS2* and *TBX3* in mouse limbs.

a Schematic representation of distal *MEIS2* and *TBX3* expression strategy. *MEIS2* and *TBX3* expression was controlled by a previously characterized *Bmp2* enhancer (Dathe et al., 2009) with specific activity in the distal and interdigital mesenchyme (indicated in yellow and schematically shown in E11.5-E13.5

mouse limbs). **b** WISH showing *MEIS2* transgene expression in E12.5 wildtype control and *MEIS2* mutant embryos. Fore- and hindlimbs are highlighted by dashed lines. Scale bars represent 1 cm. n=2. **c** RNA-seq showing *MEIS2* and *TBX3* transgene expression in dissected distal forelimbs at E12.5 of single and double mutants. Shown are normalized expression levels with standard error. n=3 (*TBX3*), 4 (*MEIS2-TBX3*) or 5 (wildtype and *MEIS2*).

To investigate the impact of ectopic distal *MEIS2* and *TBX3* expression on mouse limb developmental cell types, I generated scRNA-seq data of mutant forelimbs at E12.5, where enhancer activity is most prominent throughout the distal and interdigital mesenchyme (Dathe et al., 2009). The subsequent analysis was focused on distal autopodial cell clusters by using the subset of *Hoxd13* positive cells. These cells were integrated with corresponding E12.5 wildtype mouse forelimb cells from our reference data and clustered and annotated accordingly. Next, the populations exhibiting increased expression of *MEIS2* and *TBX3* in each mutant compared to the wildtype cells were identified to determine the relevant clusters for our analysis. This revealed distinct populations in the distal limb of all three mutants, the mesenchymal rim and the retinoic acid active interdigital cells and distal progenitors (1 MR, 3 RA-Id and 4 DP) (Figure 26a). These results validated distal mesenchymal enhancer activity on single-cell level and more importantly provided the framework for further characterizing gene expression changes within these cells upon transgene expression.

To do so, differential gene expression analysis was conducted between wildtype and mutant cells in the affected populations. Notably, subsets of genes from the previously identified bat fibroblast cell program were found among the DEGs in each mutant. Statistical analysis further revealed that these bat fibroblast genes are significantly overrepresented in every examined mutant cell cluster. (Figure 26a). Moreover, we observed a distinct set of bat fibroblast cell program genes affected in each mutant, suggesting the activation of different parts of the gene program by each respective TF or their combination. For instance, increased expression of the collagens *Col1a2* and *Col2a1* were only observed upon *MEIS2* expression, whereas upregulation of *Lgals1*, *Tmsb4x* or *Tpm4* was restricted to *TBX3* mutant clusters. Interestingly, a subset of genes including *Tbx18*, *Tshz1* and *Tshz2*, *Col3a1*, *Twist1*, *Alcam* and *Vim* were specifically upregulated in one of the single mutants but appeared jointly upregulated in the double mutant. This could indicate a partially additive effect of *MEIS2* and *TBX3*. Another set of genes, such as *Tbx2*, *Ccn1*, *Marcks*, *Adgrl3*, *Rdh10*, *Shox2*, *Rspo3* and *Snai2*, showed increased gene expression in the double mutant clusters only, suggesting that a concerted action of *MEIS2* and *TBX3* might be needed to induce expression of these genes. Notably, among the genes differentially upregulated upon distal *MEIS2* expression, we observed several factors that have been identified as potential MEIS2 targets within the distal bat limb network in our previous analyses, such as *Tbx2*, *Tbx18* or *Col3a1* (Figure 26a, asterisk). This confirmed our epigenetic characterization of the distal bat limb and suggests potential direct as well as indirect targets of these transcription factors in distal limb cell types.

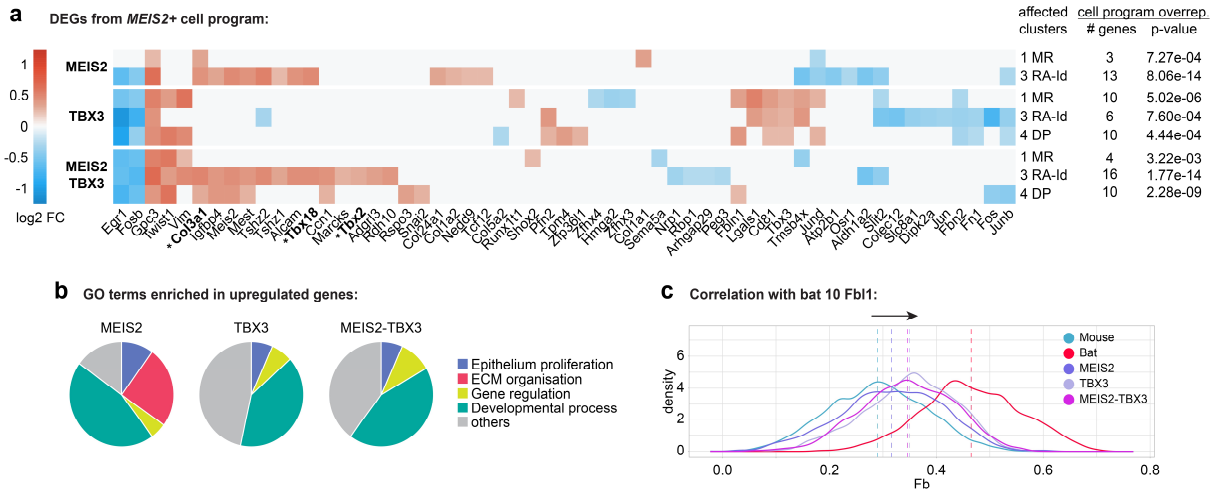

Figure 26: Ectopic distal expression of *MEIS2* and *TBX3* partially induces expression of wing fibroblast cell program in affected limb clusters.

a Expression heatmaps showing differentially expressed genes from the identified *MEIS2*-positive cell program (Figure 20) in clusters of mouse mutant limbs at E12.5 affected by transgene expression (1 MR and 3 RA-Id in the *MEIS2* mutant, 1 MR, 3 RA-Id and 4 DP in *TBX3* and double mutants). The number of genes from the *MEIS2*-positive fibroblast cell program (see Figure 20) differentially upregulated and well as the p-value of the cell program overrepresentation in the upregulated genes are shown per cluster. Genes previously identified as potential MEIS2 targets are highlighted with an asterisk. **b** GO terms enriched in upregulated genes in transgenic mutants. Shown is the proportion of the top 10 GO terms enriched in the affected clusters of each mutant grouped by their biological function (Epithelium proliferation, ECM organization, Gene regulation or Developmental process). The full list of GO terms can be found in section 10.1. **c** Correlation of mouse bat cells with bat cluster 10 Fbl1 based on the expression of genes from the *MEIS2*-positive fibroblast cell program. Shown is the density of the correlation of all cells in the mouse clusters affected by transgenic expression, as well as corresponding clusters of wildtype mouse and bat forelimbs.

Our results revealed that *MEIS2* and *TBX3* activate distinct parts of the bat fibroblast cell program in a mouse genomic context. Furthermore, their combined expression resulted in partially shared but mostly individual features. To further gain insight into the biological processes specifically regulated and induced by *MEIS2* and *TBX3*, GO term enrichment analysis was performed on the differentially upregulated genes of each affected cluster (Figure 26b, see also Figure 35, Figure 36 and Figure 37 in section 10.1). Intriguingly, this revealed GO terms related to epithelium proliferation and morphogenesis as well as regulation of gene expression in all three mutants, among others. Furthermore, several GO terms related to ECM organization were enriched in the *MEIS2* mutant specifically. Overall, more than half of the top enriched GO terms could be related to biological processes recapitulating the ones enriched in the characterized wing fibroblast cells (Figure 20d), supporting a central role of *MEIS2* and *TBX3* in driving this cell program.

Finally, we sought to specifically assess whether the induced cellular phenotypes in the transgenic mouse limb populations recapitulate a bat wing fibroblast cell state. For this, the mean expression of the identified fibroblast cell program genes across the affected mouse cell clusters and their bat counterparts as well as the wing fibroblast cells was calculated (cluster

10 FbI1). Their transcriptional correlation was then assessed based on the relative expression level of this gene set (Figure 26c). Intriguingly, the cell states from all three mutants exhibited an increased gene expression correlation with the examined bat cells, with the *TBX3* and double mutant showing the highest correlations. Therefore, we concluded that the cellular phenotypes induced by ectopic distal expression of *MEIS2* and *TBX3* partially recapitulate the cellular bat fibroblast cell state as well as the biological functions enriched in the transcriptional profile of these cells.

7.3.2. Transgenic mouse limbs exhibit morphological features recapitulating central components of wing formation

We next aimed to evaluate not only the cellular but also the morphological consequences of ectopic distal expression of *MEIS2* and *TBX3*. To do so, tissue phenotypes of transgenic mouse limbs were analyzed by light sheet microscopy at a later developmental stage where limb morphology is fully developed (E15.5). The 3D surface volume and nuclei staining with the fluorescent dye DRAQ5 were used to quantify overall limb autopod volume and cell number, respectively (Figure 27a and b). Furthermore, limbs were stained with EosinY, a dye marking extracellular matrix and cell cytoplasm (Fischer, Jacobson, Rose, & Zeller, 2008), as a proxy to quantify connective tissue and collagen content (Figure 27a and b). At this developmental stage we observed a missing third digit in mutant limbs expressing *MEIS2*. Indeed, it has previously been reported that overexpression of *Meis* transcription factors in the distal limb interferes with the development of autopodial skeletal elements (Mercader et al., 2009). Therefore, to avoid a resultant bias in autopod tissue quantification, subsequent quantifications were focused on autopodial limb regions excluding the digits.

Interestingly, these analyses revealed a significant increase in volume and cell number at similar levels in all three mutants (Figure 27d). The connective tissue content was significantly increased in examined limbs of all genotypes as well, however it was most pronounced upon distal expression of *TBX3* (Figure 27d). Moreover, *TBX3* mutant limbs exhibited fusion of digits two and three (Figure 27a and c). Although of varying manifestation, it was observed in all examined limbs of a total of four individuals. Transversal sections of these limbs confirmed the retention of interdigital tissue, resembling cutaneous syndactyly. It can be hypothesized that the induction of increased levels of cell proliferation and connective tissue production observed in this mutant might contribute to digit fusion through interdigital tissue growth. However, the underlying molecular mechanism is yet to be investigated.

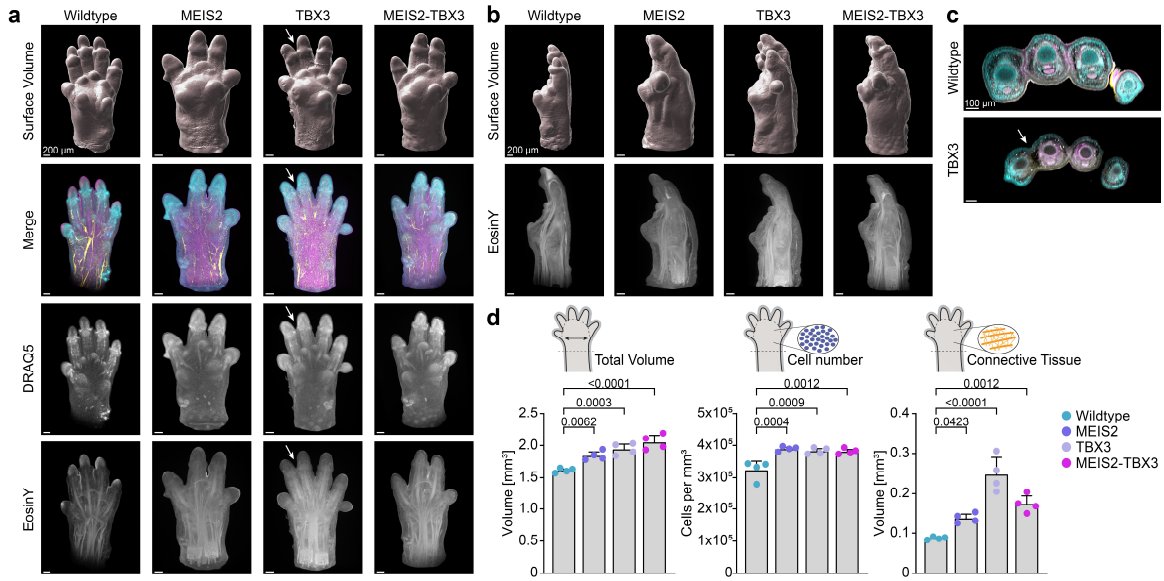

Figure 27: Transgenic limbs exhibit morphological features related to chiropatagium tissue development.

a 3D imaging of mouse wildtype and transgenic mutant limbs at E15.5. Shown is a surface representation as well as a cross-section of DRAQ5 and EosinY staining and a merged image. The syndactyly between digit 2 and 3 in the *TBX3* mutant is highlighted with an arrow. The scale bars represent 200 µm. n=4. **b** Side-view of wildtype and mutant limbs. Shown is the surface representation and a cross-section of the EosinY staining. **c** Cross-section of wildtype and *TBX3* mutant limb with an arrow indicating tissue between digits II and III. **d** Quantification of the total surface volume, the cell number and the connective tissue volume in wildtype and mutant limbs. Bars represent mean and standard deviation; one-way ANOVA with Dunnett's multiple comparisons test was performed; shown is adjusted p-value; n=4.

Taken together, our analyses revealed the manifestation of higher connective tissue content and increased cell numbers upon distal expression of *MEIS2* and *TBX3*, confirming our previous cellular phenotypic characterization of mutant limb cell populations. We therefore conclude that the observed cellular and morphological phenotypes in the autopodial portion of mutant limbs recapitulate essential aspects of bat wing development. These results provide evidence that the wing fibroblast cell program characterized in this work is in part mediated by the transcription factors *MEIS2* and *TBX3* and plays a central role in the formation of the bat chiropatagium tissue.

7.3.3. *Meis2* and *Tbx3* regulatory domains in mouse and bat limbs exhibit species-specific 3D chromatin structure and regulatory differences

The previous paragraphs have highlighted the developmental transcription factors MEIS2 and TBX3 as prominent regulators involved in driving a wing specific limb developmental program. However, how species- and forelimb-specific expression of these factors arises during limb development and how this may be implemented on a genomic and regulatory level, is yet to be investigated. Therefore, we sought to compare the chromatin domains and regulatory components of these genes between limbs and species, to assess their evolutionary

conservation and species-specific modification. To do so, I generated bulk transcriptomic and functional genomic datasets of mouse and bat limbs at developmental stages during which limb morphology is established (E11.5/CS15 and E13.5/CS17). Specifically, I generated ChIP-seq data profiling H3K27ac, H3K4me1 and H3K4me3 histone modifications as well as CTCF binding and chromatin accessibility data using ATAC-seq.

According to our previous approach, regulatory domains using Hi-C data from mouse and bat limbs were identified. Furthermore, active putative cis-regulatory elements in E13.5/CS17 fore- and hindlimb tissues were determined by intersecting ATAC-seq to profile chromatin accessibility and specific chromatin modifications (H3K27ac, H3K4me1 and H3K4me3). Specifically, enhancer probabilities were calculated using the tool CRUP (Ramisch et al., 2019), which assesses enhancer activity by integrating H3K27ac, H3K4me1, and H3K4me3 histone modification signals identified through ChIP-seq. Regions with a high enhancer probability (> 0.3) and an open chromatin signal were then classified as promoters if they overlapped with a transcription start site or as enhancers if they did not. Furthermore, the functional conservation of these regulatory regions between mouse and bat limbs was investigated by classifying directly, indirectly and non-conserved elements using Interspecies Point Projection (Phan et al., 2024). Lastly, bat accelerated regions (BARs) were calculated. BARs are regulatory regions within bat genomes that, while highly conserved throughout vertebrate evolution, exhibit an elevated substitution rate specifically within the bat lineage. Consequently, these regions are likely to reflect regulatory modifications unique to these species and may play a role in the development of bat-specific traits.

This comprehensive functional dataset was then used to compare regulatory domains of *Meis2* and *Tbx3* between mouse and bat limbs. Notably, overall conservation of the 3D structure of the approximately 2 Mb-sized *Meis2* regulatory domain was observed in both species (Figure 28). Within this conserved structure, distinct and prominent interactions between the *Meis2* gene promoter and putative regulatory elements throughout the domain were identified. Consistent with this 3D architecture, enriched binding of the architectural transcription factor CTCF at TAD boundary regions was detected. Additionally, individual CTCF peaks overlapping the *Meis2* promoter and putative regulatory elements were found, likely facilitating regulatory loops within the domain. The pattern of CTCF binding was remarkably similar between species, with only minor differences identified, primarily in the region upstream of *Meis2*.

However, despite this architectural conservation, regulatory features within this domain exhibited distinct species- and tissue-specific differences (Figure 28b and d), corresponding to *Meis2* gene expression levels in mouse and bat limb tissues. In mice, *Meis2* gene expression and its regulatory landscape were remarkably similar between the forelimb and hindlimb at both developmental stages examined (Figure 28a and b). This similarity was particularly evident in gene expression and CRUP signals at E13.5, where only minimal differences were

observed between the two tissues. In contrast, the bat *MEIS2* locus revealed several forelimb-specific enhancer regions (Figure 28d). This forelimb regulatory specificity corresponded with elevated *MEIS2* expression in bat forelimbs at CS17 (Figure 28c). Moreover, the bat *MEIS2* locus contained several enhancer elements without sequence or functional orthologues in mice, and can thus be considered both, species- and forelimb-specific. Notably, a high density of BAR elements was detected downstream of *MEIS2*, overlapping with a cluster of forelimb-specific accessible regions at CS15 and CS17. These elements may therefore represent bat-specific regulatory sequences linked to wing development.

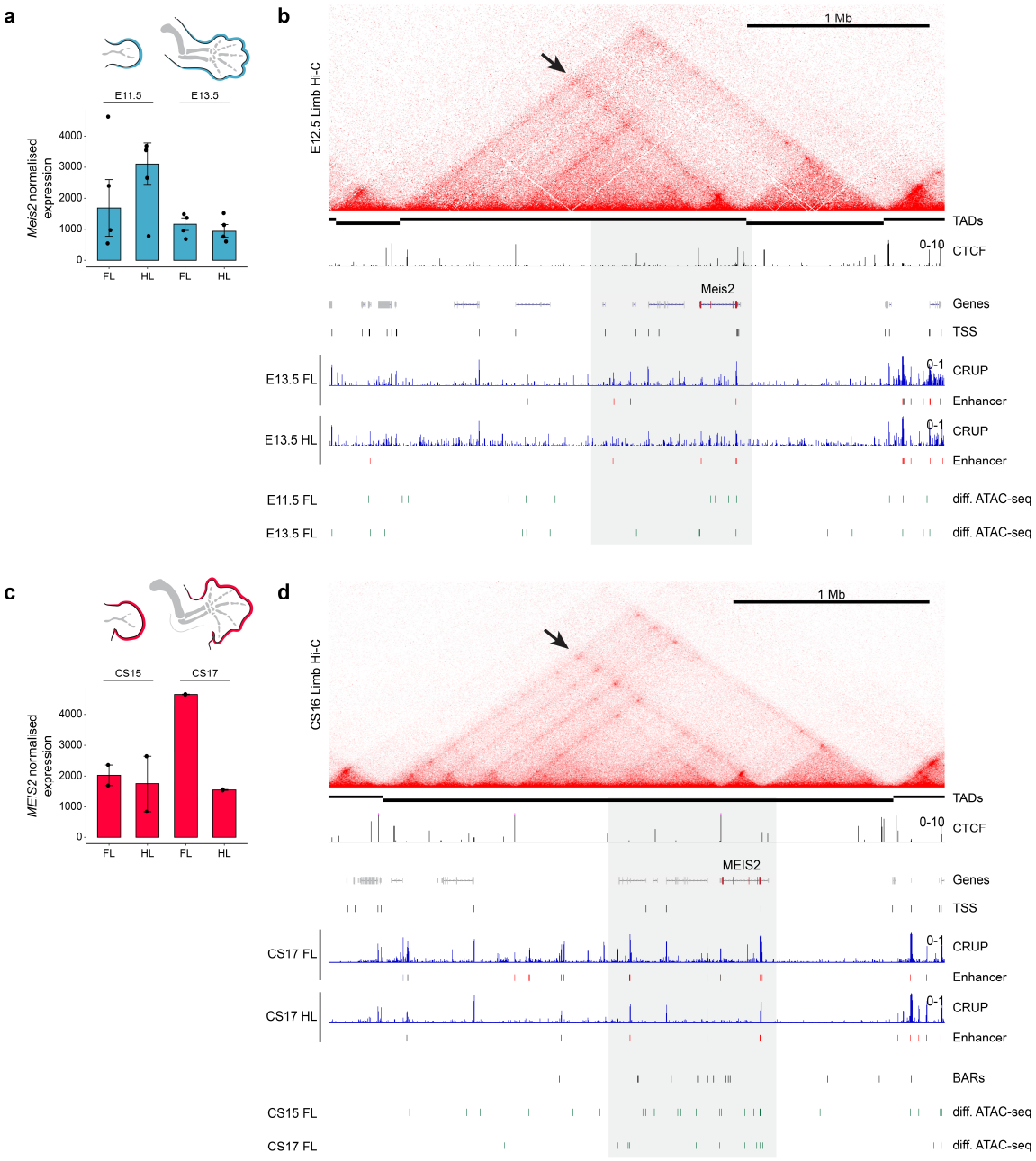

Figure 28: The *Meis2* locus in mouse and bat forelimbs exhibits chromatin structure conservation but increased regulatory forelimb activity in bat wings.

a and **c** Normalized expression levels of *Meis2* in mouse (a) and bat (b) forelimbs and hindlimbs. Mouse: n=4; bat: n=2. Error bars show the standard error. **b** and **d** Shown is the *Meis2* locus in mice (b) and bats (d) identified by Hi-C as well as CTCF ChIP-seq and fore- and hindlimb epigenetic features. CRUP enhancer prediction scores are shown in blue with predicted enhancer elements below (red: directly and indirectly conserved enhancers with functional orthologue in the other species; black: directly and indirectly conserved enhancers with non-functional orthologue in the other species, grey: non-conserved enhancers). Differential, forelimb-specific ATAC-seq peaks are shown in green. Furthermore, bat accelerated regions (BARs) are shown in the lower panel. The *Meis2* gene region and the region downstream in both species is highlighted by a grey box. Arrows indicate putative regulatory interactions with the *Meis2* promoter.

Similar to the *Meis2* mouse locus, only minor differences in gene expression and regulatory features were observed between the forelimb and hindlimb in the mouse *Tbx3* locus (Figure 29a and b). This contrasts with the neighboring domain of the forelimb-specifying transcription factor *Tbx5* (Rallis et al., 2003), which is highly enriched with forelimb-specific regulatory elements and accessible chromatin regions. In the bat, however, the *TBX3* regulatory domain displayed not only a myriad of forelimb-specific regulatory elements but also a striking enrichment of enhancer elements across the entire domain, reminiscent of the strong forelimb specificity of *TBX5* (Figure 29d). Importantly, this correlates with the strong forelimb-specific expression of *TBX3* in the bat wing at both developmental stages examined (Figure 29c). The region surrounding the gene, as well as the downstream TAD boundary, showed a pronounced enrichment of enhancer elements compared to the bat hindlimb and the corresponding mouse region. Additionally, several of these putative enhancer elements lack sequence or functional orthologues in mouse limbs. Notably, the central domain also contains numerous BAR elements, many of which overlap with forelimb-specific accessible regions at CS15 and CS17.

Interestingly, analysis of the 3D chromatin architecture at this locus revealed notable structural differences between mouse and bat. In the mouse, the *Tbx3* and *Tbx5* domains are largely self-interacting, resulting in their isolation from one another (Figure 29b). Within the *Tbx3* regulatory domain, two prominent interactions in the Hi-C contact map were observed, likely corresponding to the boundary elements of this TAD and an intra-TAD interaction between the *Tbx3* gene promoter and a downstream region. In contrast, the bat locus showed increased contact between the *Tbx3* and *Tbx5* domains, indicated by the formation of additional chromatin interactions between both regions (Figure 29d). These interactions primarily involved the outermost TAD boundary elements and a region surrounding *TBX3*. Additionally, enhanced contact between the promoter regions of *Tbx3* and *Tbx5* was detected, indicating a stronger interaction between these loci in the bat. Consistent with these structural differences, variations in CTCF binding patterns at the *TBX3* domain boundaries were observed between the two species, with the most pronounced differences occurring at both boundary regions. Importantly, this suggests potential alterations in these boundary elements, which may contribute to the modification of overall chromatin architecture.

Altogether, these results reveal a myriad of regulatory modifications in *MEIS2* and *TBX3* regulatory domains specific to bat forelimbs. Interestingly, this suggests evolutionary adaptations of conserved enhancer elements as well as the emergence of novel regulatory regions underlying the emergence of wing-specific expression patterns. These regulatory changes seem to have evolved within conserved 3D chromatin domains in the case of *MEIS2*, while our observations suggest that they coincide with additional architectural modifications in the *TBX3* locus. The generated datasets furthermore provide useful tools to further explore how gene regulation was modified throughout bat evolution in a genome wide manner. Overall, they confirm our previous findings that *MEIS2* and *TBX3* are prominent subjects in bat wing evolution and therefore likely contribute to the development of forelimb-specific traits. Yet, deeper investigation of the observed differences is needed to fully assess their functional consequences during limb development.

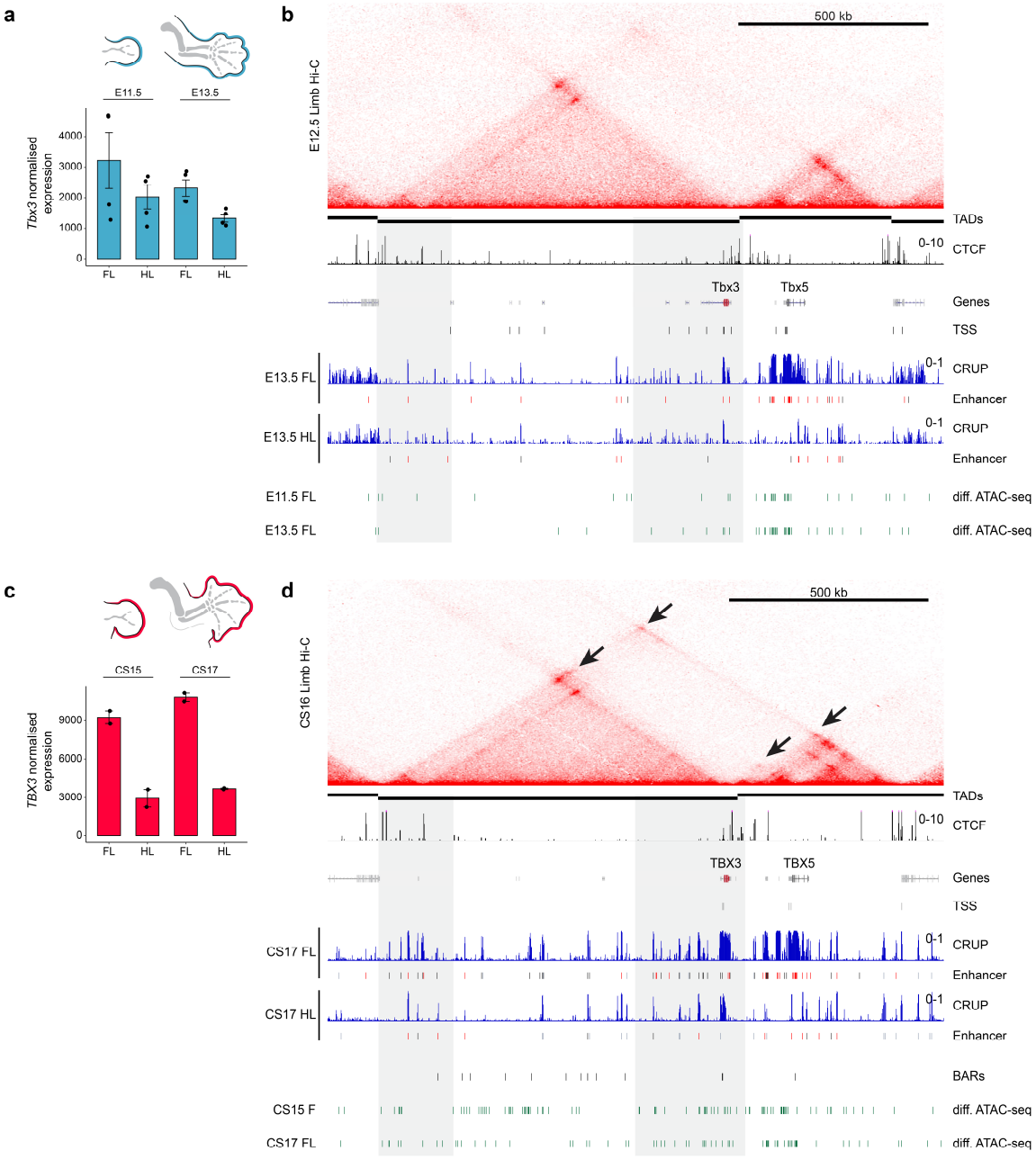

Figure 29: The *Tbx3* locus displays 3D chromatin interaction changes and strong wing-specific regulatory activity.

a and **c** Normalized expression levels of *Tbx3* in mouse (a) and bat (b) forelimbs and hindlimbs. Mouse: n=4; bat: n=2. **b** and **d** Shown is the *Tbx3* locus in mice (b) and bats (d) identified by Hi-C as well as CTCF ChIP-seq and fore- and hindlimb epigenetic features. CRUP enhancer prediction scores are shown in blue with predicted enhancer elements below (red: directly and indirectly conserved enhancers with functional orthologue in the other species; black: directly and indirectly conserved enhancers with non-functional orthologue in the other species, grey: non-conserved enhancers). Forelimb-specific ATAC-seq peaks are shown in green. Furthermore, bat accelerated regions (BARs) are shown in the lower panel. The *Tbx3* gene region and the region upstream, as well as the upstream boundary region in both species is highlighted by a grey box. Arrows indicate bat-specific 3D chromatin interactions.

8. Discussion

The processes governing embryogenesis rely on precise spatiotemporal gene expression that drive the transition of undifferentiated cells into distinct states. The resulting cell type-specific behaviors – like proliferation, homeostasis, differentiation, or elimination – collectively drive morphogenesis and shape the developing embryo. In evolutionary terms, modifications to these processes have the potential to be source of phenotypic innovation. However, how such alterations influence developmental programs and cell types and to what extent these are modified to drive novel traits remains a longstanding question in evolutionary developmental biology.

Here, I addressed this question by using the morphological diversity of the limb as a model to study the molecular and cellular basis of evolutionary change during development. Specifically, we made use of an extreme example of morphological adaptation, the emergence of a forelimb interdigital wing membrane in bats. By comparing mouse and bat limb developmental cell states we identified a wing fibroblast cell program, which emerges from a common autopodial progenitor and shares the transcriptional signature of a conserved proximal limb fibroblast cell type. As such, we demonstrated that morphological novelty is driven by repurposing of existing gene activities and cell types rather than establishing novel ones from scratch.

Here, I will discuss how the morphological adaptation of an ancestral organ constructed by conserved cell types reveals important aspects of the mechanisms driving evolution, thereby expanding our understanding of evo-devo biology. In this context, I will comment on the significance of repurposing regulatory networks in the modification of developmental cell states. Finally, I will discuss the challenges of comparative genomics studies in molecularly disentangling developmental cell populations across species and future research directions proposed by this work.

8.1. Conserved cell states despite diverse morphologies

Bats, members of the order Chiroptera, have undergone remarkable evolutionary adaptations, one of which is the transformation of their forelimbs into wings. This forelimb morphology is characterized by elongated skeletal elements and the presence of several wing membranes (Sadier et al., 2021). Importantly, while other species, such as flying squirrels, colugos and marsupial possums have also developed various forms of patagia (Casanovas-Vilar et al., 2018; Meredith, Westerman, & Springer, 2009; Schmitz, Ohme, Suryobroto, & Zischler, 2002), these species only use them for non-powered flight, or gliding. The unique appendage adaptations of bats however make them the only mammals capable of true, powered flight which contributed to their great evolutionary success (Burgin et al., 2018; Gunnell & Simmons, 2005; Sadier et al., 2021). The defining morphological features of the bat wing appear most

pronounced in the autopodial region, where digits II-V are dramatically elongated and connected by an elastic wing membrane, the chiropatagium. It is this distinctive feature that inspired the name “Chiroptera”, derived from the Greek words *cheir* (hand) and *pteron* (wing), reflecting this specialized limb structure.

Importantly, in species with separated digits, the interdigital tissue is removed during the process of limb development. Therefore, from an evolutionary perspective, different cellular mechanisms can be proposed for the unique emergence of the chiropatagium tissue: (1) the innovation of a novel cell type, (2) the persistence of a cell population that regresses in other species or (3) the modification of an existing cell state to repurpose its function.

The tetrapod limb represents an ancestral organ that – despite its conservation – has undergone a myriad of phenotypic changes throughout evolution. Many studies have been investigating the gene expression and regulatory changes responsible for its diversification. However, these studies were mostly limited to candidate gene approaches and lacked cellular resolution in a heterogeneous developing tissue. Therefore, assessing the degree of conservation of cellular phenotypes and how this is influenced by gene activity changes has been challenging. Only recently the development of single-cell studies has enabled the characterization of cell state homology during development by integrating limb cells from different species. This revealed a high degree of cell state similarity, for instance when comparing mouse and human limbs covering different stages of patterning and morphogenesis (B. Zhang et al., 2023). Here, the authors reported highly similar expression patterns of genes controlling fore- and hindlimb as well as proximal–distal identity in both species. A similar trend was observed in an even more diverse multi-species atlas composed of cells from human, mouse, chicken, frog and axolotl. Despite large evolutionary distances of these species and significant differences in limb morphogenesis, a high degree of transcriptional similarity was detected between their limb cell populations (Zhong et al., 2023). Interestingly, when comparing limb developmental cell states between mice and bats in our inter-species limb atlas, we observed an equally high degree of similarity. In line with the mentioned studies, this indicates that limbs of both species are constructed from the same set of conserved limb cell populations, despite their diverging morphologies. This, however, raises the question, how these highly similar cell states are modified to result in such phenotypic differences.

Prominent evolutionary mechanisms producing phenotypic innovation include, for instance, gene duplications or losses (Albalat & Cañestro, 2016; Moore & Purugganan, 2003). However, a previous study used the genomes of six bat species for a genome-wide screen for gene changes that occurred in these species and did not identify any candidates connected to limb development (Jebb et al., 2020). Alternatively, genes can be newly incorporated into existing regulatory networks thereby gaining activity in a different cell type or tissue. This principle of co-option has been described in several examples and is considered an important source of

phenotypic novelty (Holland, 2013; McLennan, 2008; Sanetra, Begemann, Becker, & Meyer, 2005; Schindler et al., 2023). However, our observations indicated that the genes expressed the bat forelimb are not species- or wing-specific but are expressed in mouse limb tissues as well. Importantly, this suggests that changes within these transcriptionally highly similar populations involve modifications of existing developmental programs instead of driving novel cell types or incorporating new genes.

In the present work, we especially exemplify the degree of cell state similarity with our investigations of the cellular origin of the bat chiropatagium. Several studies have been proposing the retention of the interdigital mesenchyme through the inhibition of cell death during limb development as origin of this tissue (L. Cooper et al., 2013; Hockman et al., 2008; Weatherbee et al., 2006). However, contrary to these previous findings, we not only found a highly similar interdigital cell population but revealed the same apoptotic fate of those cells in both species. Importantly, this rules out the persistence of this developmental cell population in forming the chiropatagium, and furthermore emphasizes the level of cell state similarity and conservation of key developmental processes.

Interestingly, previous studies already suggested that the expression of factors involved in apoptosis signaling, such as components of the RA pathway and BMP signaling factors, did not exhibit significant differences (Mason et al., 2015; Z. Wang et al., 2014; Weatherbee et al., 2006). However, observations such as the expression of an anti-apoptotic component, namely *Grem1*, have led to the hypothesis that interdigital apoptosis might be inhibited and cell survival promoted in bat forelimbs (Weatherbee et al., 2006). With our single-cell approach we revealed two distinct cell populations present in the interdigital region of bat forelimbs. One population corresponded to the apoptotic interdigital mesenchyme, while the other consisted of wing fibroblasts expressing *Grem1*, indicating that the investigated apoptosis signaling components and *Grem1* are not co-expressed within the same cells. Therefore, this finding provides a coherent explanation for previously reported discrepancies and further emphasizes the essential utility of single-cell technologies in delineating complex expression patterns at cellular resolution.

Characterization of the interdigital chiropatagium cells identified in this study revealed that these cells exhibit the expression of genes involved in mesenchymal proliferation, extracellular matrix organization and cell migration. Interestingly, these processes not only indicated the fibroblast identity of these cells (Plikus et al., 2021) but also have been shown to play a role in the remodeling of interdigital tissue during limb development (Kashgari et al., 2020). Here, the authors described a mechanism distinct from interdigital cell death involved in digit separation, relying on active migration of epidermal cells into the interdigital mesenchyme. As the interdigital mesenchyme is regressed, the epidermis invaginates into the interdigital space, forming a multi-layered structure of epithelial tissue, which eventually converges and

separates. Interestingly, disruption of this process resulted in fusion of the migrating epithelium and manifested as syndactyly of digits. Based on these results and our observations, it can be hypothesized similar processes of active interdigital tissue remodeling through fibroblast cell proliferation and migration are involved in wing development. Yet, the precise underlying mechanism remains to be investigated.

As discussed here, our comparative single-cell approach highlights the advantages of single-cell technologies for understanding developmental processes and their modification or conservation during evolution. However, when investigating cell state conservation or diversification through integrative single-cell approaches, it is crucial to consider the technical limitations when interpreting given results. While data integration allows the direct comparison of transcriptionally similar cell types, a significant technological challenge lies in preserving biological heterogeneity (Song et al., 2023). Indeed, a primary concern when integrating single-cell data across multiple species is the risk of overcorrection, which can mask species-specific biological information. We encountered this issue during the integration of our bat and mouse limb data, where distal wing fibroblasts clustered with other limb fibroblasts from both species. Importantly, our findings based on individually analyzed data demonstrated that this clustering was not a technical artifact but instead reflected the high degree of transcriptional similarity between these two populations. Therefore, our comparative approach combining data integration and individual analyses not only highlights a high degree of cell state conservation in the limb, but also emphasizes the importance of employing diverse analytical methodologies when comparing data derived from different species.

8.2. Repurposing of conserved developmental cell states drives phenotypic novelty

The astonishing degree of transcriptional similarity between mouse and bat limb cells raises the question, how and to what extent these cell states are modified to create novel morphological output. Importantly, our work has revealed that the emergence of a novel limb structure is driven by the repurposing of a conserved proximal limb cell state and its gene expression program in the distal bat autopod. This points towards a mechanism where a similar developmental path is initiated at different anatomical locations, thereby making minimal modifications to existing programs.

Interestingly, elucidating the developmental origin of this cell population suggested that the chiropatagium fibroblast cells emerge from a common distal progenitor and subsequently follow a novel differentiation trajectory in the autopod. Recent studies in the developing mouse limb have proposed a model in which a naïve limb progenitor differentiates into a proximal and autopodial limb progenitor state, giving rise to proximal and distal limb segments, respectively (Markman et al., 2023). According to this model, the resulting distal cell state might therefore

be transcriptionally similar but epigenetically distinct to its proximal cell homologue. Intriguingly, a similar case has recently been reported for transcriptionally homologous skeletogenic cells originating from different germ layers (M. Wang et al., 2024). This study proposes a convergent specification of cell types from different origins induced by lineage-specific transcription factors. On the one hand, this implies the activation of a highly similar downstream transcriptional program, while requiring a different upstream set of regulatory elements. On the other hand, the potentially diverse epigenetic profiles of cells from different origins likely result in the activation of additional gene sets and therefore distinct transcriptional characteristics. Importantly, this is in line with our observations of gene subsets expressed uniquely in the distal or proximal fibroblast populations. Our work, consistent with other studies, therefore, proposes the convergence of cell fates through distinct regulatory profiles executing a functional core program as a prominent mechanism in modulating developmental cell types in evolution.

Further investigations are however necessary to fully characterize the different upstream regulators initiating this conserved program in the proximal and distal limb and determine how extensively existing networks were adapted. Therefore, it would be exciting to not only compare the transcriptomes, but also the epigenetic profiles of both cell populations at single cell level using single-cell ATAC-seq, a method detecting open chromatin regions in single cells. This would not only allow to decipher their distinct epigenetic profiles but also provide insight into potential distinct upstream regulators of both, distal and proximal, fibroblast cell states.

In this context, it is crucial to consider that the proximal and distal fibroblast populations are likely modulated by the signaling molecules present in their microenvironment, which vary according to their anatomical location and cellular origin. For instance, the interdigital space is ultimately sculpted by the simultaneously developing distal mesenchymal condensations. Moreover, distal limb mesenchymal cells are exposed to signals from the AER (Mercader et al., 2000). Therefore, signaling crosstalk between these populations may influence cellular fates in the interdigital region that proximal cells are not receiving (Murgai, Altmeyer, Wiegand, Tylzanowski, & Stricker, 2018; Villacorte et al., 2010). However, the intricate interactions between interdigital chiropatagium cells and neighboring populations, namely chondrocytes or ectodermal cells, as well as the specific signaling pathways they engage with, remain to be elucidated. Ultimately, it is important to bear in mind that the bat wing represents a complex phenotype shaped throughout different stages of limb development, starting with the posterior expansion of the limb bud (Cretekos et al., 2005), followed by the elongation of all three skeletal elements and the formation of multiple patagia tissues (Sadier et al., 2021). This suggests a coordinated interplay of several molecular mechanism that partly progressively, partly simultaneously orchestrate the patterning and differentiation of these structures.

Taken together, in line with the high level of similarity observed in limb developmental cell types of mice and bats, our observations reveal the re-purposing of an existing and conserved limb cell program in chiropatagium development. Importantly, this highlights subtle evolutionary adjustments rather than a brute-force alteration of cell behaviors as evolutionary strategy to drive novel traits.

8.3. Putative cell program regulators exhibit conserved gene function with modified activity

Our studies highlighted two transcription factors as prominent regulators of the here identified repurposed wing fibroblast cell program, *MEIS2* and *TBX3*. Both genes have been described to be expressed in the distal bat forelimb in different bat species (*Miniopterus natalensis*, *Miniopterus schreibersii*), confirming this as conserved feature of wing development (Dai et al., 2014; Z. Wang et al., 2014).

Importantly, the distal interdigital expression of *Meis2* is not unique to the bat forelimb but has been reported to occur in mouse limbs and the bat hindlimb as well (Dai et al., 2014; Mason et al., 2015). Indeed, our quantification of autopodial *Meis2* expression revealed *Meis2*-positive cells in the distal limb regions of the mouse and the bat hindlimb, although at lower fraction. Concordantly, the above mentioned *in situ* hybridization studies have shown expression of *Meis2* in the interdigital region of mouse limb tissues, however at lower levels than in bat forelimbs. Importantly, bat limbs additionally exhibited prolonged expression of *MEIS2* compared to stage matched mouse limbs. Interestingly, these observations point towards a spatially conserved expression pattern of this gene in the interdigital space in both species. However, its temporal activity and expression levels likely were modified, resulting in prolonged and elevated expression during the development of bat wings.

Interestingly, our transgenics approach demonstrated that sustained distal expression of *MEIS2*, *TBX3* and their combination partially recapitulates different parts of the identified wing cell program in mouse limbs. Although their sole expression is not sufficient to recreate a complex wing phenotype, this indicates a degree of functional conservation of these transcription factors in the distal limb and further emphasizes the lineage specific modification of an existing limb cell state to adapt limb morphology. Importantly, a similar case has been described for the evolutionary origins of the gliding membrane of the sugar glider (Moreno et al., 2024). Here, the authors identified an upstream regulator of lateral patagium development, *Emx2*, exhibiting conserved spatial expression and function across mammals. Similar to our observations, the authors provide evidence that expression timing and levels were likely modified from a pre-existing developmental program. Taken together, our observations point towards an at least partially conserved function of *Meis2* and *Tbx3* in autopod morphogenesis while their expression levels and duration were modified in bat wing development. Importantly,

this might point towards a mechanism modifying already existing and active developmental pathways by adapting the activity of their key regulators.

Regulatory variation is considered a common source of such evolutionary changes which can, in principle, arise from different molecular mechanisms. For example, changes to regulatory elements can modify the expression of genes within gene networks, affecting expression levels, duration or spatial domains (Marlétaz et al., 2023; Moreno et al., 2024). Alternatively, gene expression can be modified through genomic rearrangements re-wiring regulatory interactions and thereby altering the functional content of gene regulatory domains (Marlétaz et al., 2023; Real et al., 2020). Therefore, it would be exciting to further investigate how these modifications are implemented on a genomic and gene regulatory level to alter gene activities.

When inspecting *MEIS2* and *TBX3* regulatory domains in mouse and bat limbs as well as regulatory elements within them, we noticed prominent regulatory differences between bat fore- and hindlimb but also between species. Specifically, both gene landscapes in bat forelimbs exhibited a large number of forelimb- and species-specific putative regulatory elements. Furthermore, we noticed the presence of accelerated regions specific to the bat lineage in both domains. Bat accelerated regions represent conserved elements that acquired many sequence changes in the bat lineage (Capra, Erwin, McKinsey, Rubenstein, & Pollard, 2013). These fast-evolving regions might therefore represent modified or novel enhancer elements contributing to bat-specific gene activities. A similar case has been reported in the *Emx2* regulatory domain in the sugar glider. Here, the authors reported significant enrichment of glider accelerated regions in this domain and investigate their regulatory activity as potential drivers of changed gene expression (Moreno et al., 2024). It would thus be of great interest to further characterize the activities of a selected set of functionally conserved as well as limb- or species-specific enhancer elements. Moreover, the analysis of enriched binding motifs within regulatory elements exhibiting altered spatiotemporal activity would provide valuable insights into the networks driving wing development. Such an investigation could reveal the transcription factors interacting with these elements and suggest potential regulators orchestrating gene expression during species-specific developmental processes.

On a larger genomic scale, we observed alterations in 3D chromatin structure in the *TBX3* regulatory domain in bat forelimbs. Specifically, we noticed additional chromatin interactions with regions of the neighboring TAD harboring *TBX5*, a regulator specifically involved in forelimb development. Although highly speculative, this could be due to a weakened boundary element separating the two domains and indicate a mechanism where *TBX3* adapted regulatory forelimb specificity from the neighboring *TBX5* domain. The concept of subtle changes in 3D chromatin folding during evolution has been described in a study comparing human and macaque brain tissues. Here, the authors not only discovered the existence of divergent loops, but also revealed the disappearance of weak boundary elements in the human

lineage (X. Luo et al., 2021). Another study compared different human populations and Neanderthals to emphasize changes in chromatin folding, such as variations in looping patterns and the strengthening of sub-TADs (McArthur et al., 2022). Therefore, the observed subtle modifications to chromatin architecture propose additional mechanisms at play in the modification of gene activities during bat wing development. Yet, if and how this influences regulation of *TBX3* gene activity is to be determined.

Lastly, as our study proposes both, *MEIS2* and *TBX3*, as putative regulators of chiropatagium development, it would be exciting to further study their individual, but especially also their concerted actions in autopod morphogenesis. Intriguingly, a recent study has identified a compound binding motif of MEIS and TBX factors, thereby proposing their concerted action during limb development (Delgado et al., 2021). To better understand the downstream effects and hierarchical relationships of these factors, further functional genomic analyses of bat and transgenic mouse limbs may be needed. For instance, ChIP-seq analysis of MEIS and TBX binding in developing limbs could reveal regions bound by both factors, thereby revealing their joint downstream targets. Moreover, by integrating scRNA-seq and scATAC-seq data, cells co-expressing both transcription factors can be isolated and compared with cells lacking their expression, thereby uncovering distinct chromatin profiles associated with the presence of these TFs.

Taken together, further insight into the nature of modifications to *MEIS2* and *TBX3* activities is needed to answer how extensively existing networks were adapted to modify limb cell programs. Nonetheless, our observations highlight bat-specific regulatory features as well as their potentially altered chromatin interactions, proposing exciting further directions to disentangle regulatory innovation associated with this trait.

8.4. Developmental TFs *MEIS2* and *TBX3* likely play distinct roles in early limb development and later autopod morphogenesis

In the context of examining conserved functions during limb development, it is important to emphasize that the transcription factors MEIS2 and TBX3 play pivotal roles during early limb formation. However, these functions likely are distinct and therefore to distinguish from those observed during the later stages of autopod morphogenesis described here. MEIS1/2 homeobox transcription factors are specifically known for their patterning functions during early limb development, where they determine the proximal identity of zeugopod and stylopod. Interestingly, this proximal element specification by *Meis* genes represents an ancient limb developmental function conserved across vertebrates, as shown for several species including mice, birds and amphibians (Coy & Borycki, 2010; Delgado et al., 2021; Mercader et al., 1999; Mercader, Tanaka, & Torres, 2005). Previous studies in bat limbs also highlight *Meis2* expression in the same proximal regions during early stages limb development (Dai et al.,

2014). Importantly, comparison of MEIS binding in early mouse limb buds at E10.5 with late distal wings at CS19 revealed only minimal overlap in MEIS bound promoters, therefore suggesting a different set of downstream targets. Furthermore, previous studies have reported that *Meis2* expression is activated through distinct mechanisms in early proximal and later distal limb development. During early limb development, *Meis2* expression is activated by retinoic acid, a signaling molecule synthesized in the embryonic flank (Mercader et al., 2000). Although this molecule also plays a crucial role in the distal limb, *Meis2* expression in this region has been shown to be activated independently of RA (Mason et al 2015). We therefore conclude that distinct upstream activators and downstream targets are involved at each developmental timepoint and anatomical position. Consequently, this suggests that the limb programs at play are most likely unrelated and regulated through independent mechanisms.

Tbx3 exhibits similar, yet less characterized features. This transcription factor is expressed in proximal regions during early stages of limb development and plays a crucial role in anterior-posterior pattering of the limb (Soussi et al., 2024). Interestingly, evidence suggests that *Tbx3* expression is regulated by RA as well, indicating a regulatory mechanism similar to that of *Meis2* in the early proximal limb bud (Ballim, Mendelsohn, Papaioannou, & Prince, 2012). In contrast to *Meis2*, however, much less is known about its spatiotemporal expression pattern during later stages of autopod morphogenesis. Nevertheless, it can be hypothesized that *Tbx3* as well has distinct roles during early and late limb development.

Finally, it is also important to acknowledge that the developmental stages examined in this study (E11.5, E13.5 and E15.5 and equivalent bat stages) not fully encompass the earlier stages (~E10.5) when proximal expression of *Meis2* and *Tbx3* is initiated and contributes to early limb patterning. Consequently, a comparative analysis of the cell states associated with these factors during early limb patterning and autopod morphogenesis is lacking, which limits our ability to conclusively distinguish their functions at each respective developmental timepoint.

8.5. Chiropatagium development is a polygenic trait

As discussed above, our work highlighted the critical roles of *MEIS2* and *TBX3* in chiropatagium development. Specifically, the observed gene expression profiles in transgenic limbs, along with alterations in morphology, cell number, and matrix production, are characteristic of the chiropatagium cell gene program, supporting their central role. However, *MEIS2* and *TBX3* expression alone, or even in combination, is not sufficient to fully recapitulate a wing fibroblast cell state or morphology. Consistent with this, our ChIP-seq analysis shows that only a subset of the gene program components is enriched in MEIS2 binding, suggesting that this transcription factor does not target the whole cell program.

Importantly, this study not only identifies individual components, but a fibroblast cell program that as a whole was repurposed in bat wing evolution. This strongly suggests that this tissue represents a polygenic trait, and its formation is induced and patterned by the concerted action of multiple genes and their species-specific interactions. Indeed, genome-wide association studies have reported that complex traits tend to be highly polygenic (Bergey et al., 2018; O'Connor et al., 2019). Although it is often unclear how critical each component is for its manifestation, it is reasonable to assume that their individual as well as joint action have important biological effects. Predicting these effects on the other hand, especially when reference data is limited, can be challenging.

Genetic techniques, such as gene knockout or transgenic expression as utilized in this study are classical approaches to study the functional role of candidate genes and their relationships. Here, examining the effects of gene loss or gain provides insight into its role within a cell, tissue or a whole organism. However, the knockout or misexpression of key developmental factors may interfere with and disrupt normal molecular or morphological processes, potentially resulting in misleading results. We encountered this difficulty for one of our transgenic mutants, where distal *MEIS2* expression resulted in the partial disruption of autopod development in form of a missing digit. Indeed, several studies have reported that ectopic distal expression of MEIS factors inhibits the development of distal limb segments (Capdevila et al., 1999; Mercader et al., 2009). Therefore, we cannot rule out the possibility that these disruptions impact both molecular and phenotypic outcomes, which may limit our ability to interpret the functional roles of *MEIS2*, both on its own and in combination with *TBX3*.

Alternatively, it can be helpful to use nature itself as a guide, especially when studying evolutionary processes. Our work has provided evidence that essential gene functions in chiropatagium development likely are conserved, and ancestral programs are reused and modified to drive this trait. Therefore, evolutionary comparison of similar traits can outline functional similarities of central components and networks driving them. Interestingly, the formation of patagia represents a trait shared between a number of extant species and can therefore be used for comparative studies. Besides bats, which evolved powered flight, several other species from diverse taxa use unpowered flight, or gliding. These include flying squirrels, colugos and marsupial possums, amongst others (Casanovas-Vilar et al., 2018; Meredith et al., 2009; Schmitz et al., 2002). What all these species anatomically have in common is a lateral membrane (lateral patagium, or plagiopatagium in bats) between their fore- and hindlimbs. Notably, this core structure of flight arose independently across large evolutionary distances. Comparing molecular profiles driving its development might therefore reveal shared ancestral aspects of patagium emergence.

A recent study has addressed the morphological and molecular features during patagium outgrowth in the sugar glider (Feigin et al., 2023). Interestingly, the authors delineate specific

characteristics of the lateral patagium, including increased cell density and elevated levels of cellular proliferation, which closely align with our observations in the interdigital patagium. This strongly suggests a high degree of morphological similarity between both patagia tissues. Furthermore, on a molecular level, the authors profiled gene expression signatures enriched in the sugar glider and bat lateral patagium. Intriguingly, a subset of those genes with known functions in limb development is specifically expressed in the chiropatagium fibroblast cells as well. Among those are *Tbx3* and *Tbx2*, *Grem1*, *Osr1*, *Shox2* and *Prrx1*. Exploration of their function might thus give information on their role in skin membrane development.

Tbx2, *Tbx3* and *Grem1*, for instance, are developmental factors involved in establishing the anterior-posterior patterning in the developing limb (Panman et al., 2006; Sheeba & Logan, 2017; Tümpel et al., 2002). *Shox2* has a crucial role in chondrocyte proliferation of proximal limb segments (Ling et al., 2007). However, it has been reported as a marker for skin fibroblasts as well (Hsia et al., 2016). Similarly, the mesenchymal limb marker *Prrx1* has been shown to be a central transcription factor in stromal fibroblasts (Lee et al., 2022). Finally, *Osr1* has been reported to mark interstitial muscle connective tissue in the limb (Vallecillo-García et al., 2017). These previously characterized functions of this gene set highlight their roles in transcriptional regulation and patterning during development, but importantly also propose a function in determining fibroblast identity. Moreover, we observed differential upregulation of shared patagium genes, *Shox2* and *Tbx2*, in the transgenic mutant expressing both, *MEIS2* and *TBX3*. This not only indicates a role of these genes in driving patagium development, but on a larger scale, points towards their molecular convergence in the emergence of flight membranes through the proliferation of fibroblast cells.

Besides shared molecular features, this comparative approach importantly also highlights cell program components specific to the chiropatagium, such as *MEIS2*. One could hypothesize that the genes common to the different patagia function as downstream effectors, orchestrating the development of a similar tissue by executing a “patagium core program”. Meanwhile, tissue-specific factors may act as upstream regulators, determining the precise anatomical position where this developmental process is initiated. The conserved expression pattern of *MEIS2* in the interdigital space or *Emx2* in the interlimb skin region (Moreno et al., 2024), for instance, might highlight their role in defining the cellular origin for program initiation. Therefore, our findings not only point towards a common program activated by specific regulators, but also further highlight the need for concerted actions and therefore a polygenic basis for this trait.

Taken together, further investigation is needed to determine the precise functions of the identified components of the cell program identified in this study and their interplay during developmental processes. Nonetheless, the exploration and comparison of convergent developmental programs promise great potential to gain further insight into the mechanisms

driving the modification of developmental cell types and morphological adaptation. In this regard our work provides evidence further supporting the roles of previously described genes in the development of wing membranes. Moreover, it adds species- and cell-specific factors to the regulators driving wing membrane development in limb evolution. Finally, it outlines the use of a whole cell program instead of individual components as molecular basis of phenotypic evolution.

8.6. Limitations of this study

In this work, we examine how alterations to developmental cell states drive the emergence of novel phenotypes. Using state-of-the-art techniques including single-cell, omics and transgenic approaches we provide evidence that an existing, conserved fibroblast cell program was repurposed in the distal bat limb during chiropatagium development. However, it is important to consider potential technical and experimental limitations of this study.

Comparative single-cell studies, while invaluable, present significant technical challenges. One key issue is the reliance on previously published marker genes for cluster annotation across diverse species, which restricts our ability to fully assess the extent of cell state similarity and conservation. Additionally, single-cell data clustering and analysis ultimately is sensitive to and influenced by batch effects, depth of profiling or the technology used (Tran et al., 2020). Lastly, clustering single-cell data from dissociated cell suspensions, as used in this study, inherently lacks true spatial information, which can only be inferred from the gene expression patterns of known marker genes. As previously discussed, we were facing such technical limitations in our integrated limb dataset, where the high transcriptional similarity of cell states led to loss of spatial information and therefore the masking of biological information. The interplay of experimental and computational innovation is driving a constantly advancing repertoire of methodologies, allowing the profiling and analysis of multiple modalities, alone or in combination, at unprecedented scale and resolution (Domcke & Shendure, 2023; Heumos et al., 2023). Nonetheless, our observations highlight the necessity to interpret given results with caution and in full recognition of the inherent experimental and technical limitations.

What also needs to be considered are limitations originating from the organism used in this study itself. For one, working with a non-model species, such as bats, means restricted availability of biological material, particularly embryonic specimens at the precise developmental stage required. This scarcity arises from the limited number of breeding colonies available, and orchestrating planned matings in captivity has proven difficult. Furthermore, bat gestation periods are with approximately 120 days much longer than in the mouse and they produce only a single offspring per reproductive cycle (Rasweiler & Badwaik, 1997; Rasweiler & de Bonilla, 1992). Secondly, the resources available for this organism present notable limitations. Genomic data and annotations for non-model species are often

less complete or of lower quality compared to the well-established reference genomes of model organisms such as the mouse or human. While our research benefits from a high-quality genome and improved annotations, the possibility of missing or inaccurately defined information cannot be ruled out. Furthermore, the relatively limited availability of reference data further restricts the scope of comparative analyses. Lastly, the inability to perform genetic manipulation studies in this organism leaves a largely descriptive and speculative approach, constraining the depth of mechanistic insights that can be drawn from our research.

Lastly, the limitations of genetic validation systems pose significant challenges, even within well-established model organisms like the mouse. These systems must meet multiple criteria, such as accurate expression patterns and timing, sufficient expression levels, and non-detrimental effects on embryonic development. Consequently, fully recapitulating bat expression patterns at cellular level depends strongly on the activity of the chosen regulatory elements and is therefore biologically limited. Furthermore, the system of genetic manipulation is experimentally constrained. As our data suggests, the trait in question is polygenic, making it difficult to determine not only which but also how many factors are crucial need to be expressed. While co-expressing two genes within safe harbor loci is experimentally feasible, the complexity increases significantly as additional factors are introduced as safe harbor environments are limited. These constraints underscore the complexity of this approach and the difficulties in achieving precise genetic manipulation.

Importantly, recognizing these limitations, our study combines multiple methodologies rather than relying on a singular approach, thereby gathering evidence from a range of perspectives. This strategy creates a comprehensive framework that combines both cellular and molecular insights, thereby revealing complex developmental features and, as a result, serving as a powerful tool in the field of evolutionary developmental research.

8.7. Conclusion and outlook

Exploring cell program repertoires and their molecular characteristics can significantly enhance our understanding of cell type diversity, development, and evolution across Metazoa (Tanay & Sebé-Pedrós, 2021). Using a comparative single cell and genomic approach, we uncover the cellular origin of the chiropatagium, a wing membrane uniquely gained in chiropterans. Our findings indicate a striking similarity in gene expression across species, despite diverse morphologies. Rather than generating entirely new cell states, we describe the repurposing of a gene program of an existing, conserved cell state in a different anatomical context as the basis of driving this novel trait. Thus, even a phenotypic adaptation as dramatic as the development of wings can be achieved through relatively minor modifications to established pathways, illustrating how evolution introduces novelty by refining existing elements.

These findings not only advance our understanding of the evolutionary mechanisms that drive the emergence of novel traits but also pave the way for numerous promising directions of further research. For example, it would be exciting to investigate how the here identified cellular modifications are implemented on a genomic and gene regulatory level. A promising approach is, for instance, the use of synthetic biology, which could enable the introduction and replacement of entire bat genomic landscapes in the mouse genome. This offers an approach to investigate whole landscape regulatory behaviors together with phenotypic outcomes. A significant advantage of this strategy is that, once such a functional system is established, it becomes possible to genetically manipulate and dissect this system and systematically study the resulting effects.

Moreover, the ongoing advancements in the single-cell era are producing an increasingly abundant collection of datasets, providing an invaluable resource for investigating cell type modifications in the context of evolution and comparative biology. Moreover, the expanding array of technologies and the potential for their integration unlocks numerous opportunities to collectively examine gene expression, regulatory and epigenetic factors, and their precise spatial dynamics. Collectively, these advances open exciting new avenues for evolutionary developmental biology research, each contributing incrementally to our understanding of the molecular and morphological diversity shaped by evolution.

9. Bibliography

Abzhanov, A. (2013). von Baer's law for the ages: lost and found principles of developmental evolution. *Trends Genet, 29*(12), 712-722. doi:10.1016/j.tig.2013.09.004

Agarwal, P., Wylie, J. N., Galceran, J., Arkhitko, O., Li, C., Deng, C., . . . Bruneau, B. G. (2003). Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. *Development, 130*(3), 623-633. doi:10.1242/dev.00191

Akdemir, K. C., Le, V. T., Chandran, S., Li, Y., Verhaak, R. G., Beroukhim, R., . . . Futreal, P. A. (2020). Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer. *Nat Genet, 52*(3), 294-305. doi:10.1038/s41588-019-0564-y

Akiyama, H., Chaboissier, M. C., Martin, J. F., Schedl, A., & de Crombrugghe, B. (2002). The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. *Genes Dev, 16*(21), 2813-2828. doi:10.1101/gad.1017802

Albalat, R., & Cañestro, C. (2016). Evolution by gene loss. *Nature Reviews Genetics, 17*(7), 379-391. doi:10.1038/nrg.2016.39

Almubarak, A., Zhang, Q., Zhang, C. H., Abdelwahab, N., Kume, T., Lassar, A. B., & Berry, F. B. (2024). FOXC1 and FOXC2 regulate growth plate chondrocyte maturation towards hypertrophy in the embryonic mouse limb skeleton. *Development, 151*(16). doi:10.1242/dev.202798

Anania, C., & Lupiáñez, D. G. (2020). Order and disorder: abnormal 3D chromatin organization in human disease. *Brief Funct Genomics, 19*(2), 128-138. doi:10.1093/bfgp/elz028

Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., . . . Duboule, D. (2013). A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science, 340*(6137), 1234167. doi:10.1126/science.1234167

Andrey, G., & Spielmann, M. (2017). CRISPR/Cas9 Genome Editing in Embryonic Stem Cells. *Methods Mol Biol, 1468*, 221-234. doi:10.1007/978-1-4939-4035-6_15

Anlas, K., & Trivedi, V. (2021). Studying evolution of the primary body axis in vivo and in vitro. *Elife, 10*. doi:10.7554/eLife.69066

Arnold, C. D., Gerlach, D., Spies, D., Matts, J. A., Sytnikova, Y. A., Pagani, M., . . . Stark, A. (2014). Quantitative genome-wide enhancer activity maps for five Drosophila species show functional enhancer conservation and turnover during cis-regulatory evolution. *Nat Genet, 46*(7), 685-692. doi:10.1038/ng.3009

Artus, J., & Hadjantonakis, A. K. (2011). Generation of chimeras by aggregation of embryonic stem cells with diploid or tetraploid mouse embryos. *Methods Mol Biol, 693*, 37-56. doi:10.1007/978-1-60761-974-1_3

Bainbridge, M. N., Warren, R. L., Hirst, M., Romanuik, T., Zeng, T., Go, A., . . . Jones, S. J. (2006). Analysis of the prostate cancer cell line LNCaP transcriptome using a sequencing-by-synthesis approach. *BMC Genomics, 7*, 246. doi:10.1186/1471-2164-7-246

Bakken, T. E., Jorstad, N. L., Hu, Q., Lake, B. B., Tian, W., Kalmbach, B. E., . . . Lein, E. S. (2021). Comparative cellular analysis of motor cortex in human, marmoset and mouse. *Nature, 598*(7879), 111-119. doi:10.1038/s41586-021-03465-8

Ballim, R. D., Mendelsohn, C., Papaioannou, V. E., & Prince, S. (2012). The ulnar-mammary syndrome gene, Tbx3, is a direct target of the retinoic acid signaling pathway, which regulates its expression during mouse limb development. *Mol Biol Cell, 23*(12), 2362-2372. doi:10.1091/mbc.E11-09-0790

Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res, 21*(3), 381-395. doi:10.1038/cr.2011.22

Barna, M., & Niswander, L. (2007). Visualization of cartilage formation: insight into cellular properties of skeletal progenitors and chondrodysplasia syndromes. *Dev Cell, 12*(6), 931-941. doi:10.1016/j.devcel.2007.04.016

Bastide, S., Chomsky, E., Saudemont, B., Loe-Mie, Y., Schmutz, S., Novault, S., . . . Spitz, F. (2022). TATTOO-seq delineates spatial and cell type-specific regulatory programs in the developing limb. *Sci Adv, 8*(50), eadd0695. doi:10.1126/sciadv.add0695

Behringer, R., Gertsenstein, M., Nagy, K. V., Nagy, A. . (2014). *Manipulating the Mouse Embryo: A Laboratory Manual* (4th ed.): Cold Spring Harbor Laboratory Press.

Belmonte-Mateos, C., & Pujades, C. (2021). From Cell States to Cell Fates: How Cell Proliferation and Neuronal Differentiation Are Coordinated During Embryonic Development. *Front Neurosci, 15*, 781160. doi:10.3389/fnins.2021.781160

Bénazet, J. D., Bischofberger, M., Tiecke, E., Gonçalves, A., Martin, J. F., Zuniga, A., . . . Zeller, R. (2009). A self-regulatory system of interlinked signaling feedback loops controls mouse limb patterning. *Science, 323*(5917), 1050-1053. doi:10.1126/science.1168755

Bénazet, J. D., Pignatti, E., Nugent, A., Unal, E., Laurent, F., & Zeller, R. (2012). Smad4 is required to induce digit ray primordia and to initiate the aggregation and differentiation of chondrogenic progenitors in mouse limb buds. *Development, 139*(22), 4250-4260. doi:10.1242/dev.084822

Bentovim, L., Harden, T. T., & DePace, A. H. (2017). Transcriptional precision and accuracy in development: from measurements to models and mechanisms. *Development, 144*(21), 3855-3866. doi:10.1242/dev.146563

Bentzinger, C. F., Wang, Y. X., & Rudnicki, M. A. (2012). Building muscle: molecular regulation of myogenesis. *Cold Spring Harb Perspect Biol, 4*(2). doi:10.1101/cshperspect.a008342

Bergen, V., Lange, M., Peidli, S., Wolf, F. A., & Theis, F. J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nature Biotechnology, 38*(12), 1408-1414. doi:10.1038/s41587-020-0591-3

Bergey, C. M., Lopez, M., Harrison, G. F., Patin, E., Cohen, J. A., Quintana-Murci, L., . . . Perry, G. H. (2018). Polygenic adaptation and convergent evolution on growth and cardiac genetic pathways in African and Asian rainforest hunter-gatherers. *Proc Natl Acad Sci U S A, 115*(48), E11256-e11263. doi:10.1073/pnas.1812135115

Bolt, C. C., & Duboule, D. (2020). The regulatory landscapes of developmental genes. *Development, 147*(3). doi:10.1242/dev.171736

Booker, B. M., Friedrich, T., Mason, M. K., VanderMeer, J. E., Zhao, J., Eckalbar, W. L., . . . Ahituv, N. (2016). Bat Accelerated Regions Identify a Bat Forelimb Specific Enhancer in the HoxD Locus. *PLoS Genet, 12*(3), e1005738. doi:10.1371/journal.pgen.1005738

Bourque, G., Leong, B., Vega, V. B., Chen, X., Lee, Y. L., Srinivasan, K. G., . . . Liu, E. T. (2008). Evolution of the mammalian transcription factor binding repertoire via transposable elements. *Genome Res, 18*(11), 1752-1762. doi:10.1101/gr.080663.108

Buecker, C., & Wysocka, J. (2012). Enhancers as information integration hubs in development: lessons from genomics. *Trends Genet, 28*(6), 276-284. doi:10.1016/j.tig.2012.02.008

Buenrostro, J. D., Wu, B., Chang, H. Y., & Greenleaf, W. J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol, 109*, 21.29.21-21.29.29. doi:10.1002/0471142727.mb2129s109

Burgin, C. J., Colella, J. P., Kahn, P. L., & Upham, N. S. (2018). How many species of mammals are there? *Journal of Mammalogy, 99*(1), 1-14. doi:10.1093/jmammal/gyx147 %J Journal of Mammalogy

Calo, E., & Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? *Mol Cell, 49*(5), 825-837. doi:10.1016/j.molcel.2013.01.038

Cande, J., Goltsev, Y., & Levine, M. S. (2009). Conservation of enhancer location in divergent insects. *Proc Natl Acad Sci U S A, 106*(34), 14414-14419. doi:10.1073/pnas.0905754106

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., . . . Shendure, J. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature, 566*(7745), 496-502. doi:10.1038/s41586-019-0969-x

Capdevila, J., Tsukui, T., Rodríquez Esteban, C., Zappavigna, V., & Izpisúa Belmonte, J. C. (1999). Control of vertebrate limb outgrowth by the proximal factor Meis2 and distal antagonism of BMPs by Gremlin. *Mol Cell, 4*(5), 839-849. doi:10.1016/s1097-2765(00)80393-7

Capra, J. A., Erwin, G. D., McKinsey, G., Rubenstein, J. L., & Pollard, K. S. (2013). Many human accelerated regions are developmental enhancers. *Philos Trans R Soc Lond B Biol Sci, 368*(1632), 20130025. doi:10.1098/rstb.2013.0025

Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., . . . Kaessmann, H. (2019). Gene expression across mammalian organ development. *Nature, 571*(7766), 505-509. doi:10.1038/s41586-019-1338-5

Carroll, S. B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell, 134*(1), 25-36. doi:10.1016/j.cell.2008.06.030

Casanovas-Vilar, I., Garcia-Porta, J., Fortuny, J., Sanisidro, Ó., Prieto, J., Querejeta, M., . . . Alba, D. M. (2018). Oldest skeleton of a fossil flying squirrel casts new light on the phylogeny of the group. *Elife, 7*. doi:10.7554/eLife.39270

Castanon-Ortega, I., & Heisenberg, C.-P. (2005). A stern view of gastrulation. *Nat Cell Biol, 7*(1), 19-19. doi:10.1038/ncb0105-19

Chakraborty, A., & Ay, F. (2019). The role of 3D genome organization in disease: From compartments to single nucleotides. *Semin Cell Dev Biol, 90*, 104-113. doi:10.1016/j.semcdb.2018.07.005

Chautan, M., Chazal, G., Cecconi, F., Gruss, P., & Golstein, P. (1999). Interdigital cell death can occur through a necrotic and caspase-independent pathway. *Curr Biol, 9*(17), 967-970. doi:10.1016/s0960-9822(99)80425-4

Chen, Y., Lun, A. T., & Smyth, G. K. (2016). From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. *F1000Res, 5*, 1438. doi:10.12688/f1000research.8987.2

Chevallier, A., Kieny, M., & Mauger, A. (1977). Limb-somite relationship: origin of the limb musculature. *J Embryol Exp Morphol, 41*, 245-258.

Choi, K. S., Lee, C., Maatouk, D. M., & Harfe, B. D. (2012). Bmp2, Bmp4 and Bmp7 are co-required in the mouse AER for normal digit patterning but not limb outgrowth. *PLoS One, 7*(5), e37826. doi:10.1371/journal.pone.0037826

Choi, Y. H., & Kim, J. K. (2019). Dissecting Cellular Heterogeneity Using Single-Cell RNA Sequencing. *Mol Cells, 42*(3), 189-199. doi:10.14348/molcells.2019.2446

Christ, B., Jacob, H. J., & Jacob, M. (1977). Experimental analysis of the origin of the wing musculature in avian embryos. *Anat Embryol (Berl), 150*(2), 171-186. doi:10.1007/bf00316649

Cohn, M. J., & Tickle, C. (1996). Limbs: a model for pattern formation within the vertebrate body plan. *Trends Genet, 12*(7), 253-257. doi:10.1016/0168-9525(96)10030-5

Cohn, M. J., & Tickle, C. (1999). Developmental basis of limblessness and axial patterning in snakes. *Nature, 399*(6735), 474-479. doi:10.1038/20944

Consortium, T. U. (2022). UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res, 51*(D1), D523-D531. doi:10.1093/nar/gkac1052 %J Nucleic Acids Research

Cooper, K. L., Sears, K. E., Uygur, A., Maier, J., Baczkowski, K. S., Brosnahan, M., . . . Tabin, C. J. (2014). Patterning and post-patterning modes of evolutionary digit loss in mammals. *Nature, 511*(7507), 41-45. doi:10.1038/nature13496

Cooper, K. L., & Tabin, C. J. (2008). Understanding of bat wing evolution takes flight. *Genes Dev, 22*(2), 121-124. doi:10.1101/gad.1639108

Cooper, L., Sears, K., Adams, R., & Pedersen, S. (2013). Bat evolution, ecology, and conservation. In.

Cooper, L. N., Cretekos, C. J., & Sears, K. E. (2012). The evolution and development of mammalian flight. *Wiley Interdiscip Rev Dev Biol, 1*(5), 773-779. doi:10.1002/wdev.50

Cooper, S. J., Trinklein, N. D., Anton, E. D., Nguyen, L., & Myers, R. M. (2006). Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. *Genome Res, 16*(1), 1-10. doi:10.1101/gr.4222606

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res, 16*(22), 10881-10890. doi:10.1093/nar/16.22.10881

Coy, S. E., & Borycki, A. G. (2010). Expression analysis of TALE family transcription factors during avian development. *Dev Dyn, 239*(4), 1234-1245. doi:10.1002/dvdy.22264

Cretekos, C. J., Deng, J. M., Green, E. D., Rasweiler, J. J., & Behringer, R. R. (2007). Isolation, genomic structure and developmental expression of Fgf8 in the short-tailed fruit bat, Carollia perspicillata. *Int J Dev Biol, 51*(4), 333-338. doi:10.1387/ijdb.062257cc

Cretekos, C. J., Rasweiler, J. J., & Behringer, R. R. (2001). Comparative studies on limb morphogenesis in mice and bats: a functional genetic approach towards a molecular

understanding of diversity in organ formation. *Reprod Fertil Dev, 13*(7-8), 691-695. doi:10.1071/rd01115

Cretekos, C. J., Wang, Y., Green, E. D., Martin, J. F., Rasweiler, J. J. t., & Behringer, R. R. (2008). Regulatory divergence modifies limb length between mammals. *Genes Dev, 22*(2), 141-151. doi:10.1101/gad.1620408

Cretekos, C. J., Weatherbee, S. D., Chen, C. H., Badwaik, N. K., Niswander, L., Behringer, R. R., & Rasweiler, J. J. t. (2005). Embryonic staging system for the short-tailed fruit bat, Carollia perspicillata, a model organism for the mammalian order Chiroptera, based upon timed pregnancies in captive-bred animals. *Dev Dyn, 233*(3), 721-738. doi:10.1002/dvdy.20400

Dai, M., Wang, Y., Fang, L., Irwin, D. M., Zhu, T., Zhang, J., . . . Wang, Z. (2014). Differential expression of Meis2, Mab21l2 and Tbx3 during limb development associated with diversification of limb morphology in mammals. *PLoS One, 9*(8), e106100. doi:10.1371/journal.pone.0106100

Dathe, K., Kjaer, K. W., Brehm, A., Meinecke, P., Nürnberg, P., Neto, J. C., . . . Mundlos, S. (2009). Duplications involving a conserved regulatory element downstream of BMP2 are associated with brachydactyly type A2. *Am J Hum Genet, 84*(4), 483-492. doi:10.1016/j.ajhg.2009.03.001

Davidson, E. H., & Erwin, D. H. (2010). Evolutionary innovation and stability in animal gene networks. *J Exp Zool B Mol Dev Evol, 314*(3), 182-186. doi:10.1002/jez.b.21329

Davis, A. P., Witte, D. P., Hsieh-Li, H. M., Potter, S. S., & Capecchi, M. R. (1995). Absence of radius and ulna in mice lacking hoxa-11 and hoxd-11. *Nature, 375*(6534), 791-795. doi:10.1038/375791a0

de Laat, W., & Duboule, D. (2013). Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature, 502*(7472), 499-506. doi:10.1038/nature12753

Dekker, J., Rippe, K., Dekker, M., & Kleckner, N. (2002). Capturing chromosome conformation. *Science, 295*(5558), 1306-1311. doi:10.1126/science.1067799

Delgado, I., Giovinazzo, G., Temiño, S., Gauthier, Y., Balsalobre, A., Drouin, J., & Torres, M. (2021). Control of mouse limb initiation and antero-posterior patterning by Meis transcription factors. *Nat Commun, 12*(1), 3086. doi:10.1038/s41467-021-23373-9

Delgado, I., López-Delgado, A. C., Roselló-Díez, A., Giovinazzo, G., Cadenas, V., Fernández-de-Manuel, L., . . . Torres, M. (2020). Proximo-distal positional information encoded by an Fgf-regulated gradient of homeodomain transcription factors in the vertebrate limb. *Sci Adv, 6*(23), eaaz0742. doi:10.1126/sciadv.aaz0742

Desanlis, I., Paul, R., & Kmita, M. (2020). Transcriptional Trajectories in Mouse Limb Buds Reveal the Transition from Anterior-Posterior to Proximal-Distal Patterning at Early Limb Bud Stage. *J Dev Biol, 8*(4). doi:10.3390/jdb8040031

Dixon, J. R., Gorkin, D. U., & Ren, B. (2016). Chromatin Domains: The Unit of Chromosome Organization. *Mol Cell, 62*(5), 668-680. doi:10.1016/j.molcel.2016.05.018

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., . . . Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature, 485*(7398), 376-380. doi:10.1038/nature11082

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., . . . Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics, 29*(1), 15-21. doi:10.1093/bioinformatics/bts635

Dollé, P., Dierich, A., LeMeur, M., Schimmang, T., Schuhbaur, B., Chambon, P., & Duboule, D. (1993). Disruption of the Hoxd-13 gene induces localized heterochrony leading to mice with neotenic limbs. *Cell, 75*(3), 431-441. doi:10.1016/0092-8674(93)90378-4

Dollé, P., Izpisúa-Belmonte, J. C., Brown, J., Tickle, C., & Duboule, D. (1993). Hox genes and the morphogenesis of the vertebrate limb. *Prog Clin Biol Res, 383a*, 11-20.

Domcke, S., & Shendure, J. (2023). A reference cell tree will serve science better than a reference cell atlas. *Cell, 186*(6), 1103-1114. doi:10.1016/j.cell.2023.02.016

Dostie, J., Richmond, T. A., Arnaout, R. A., Selzer, R. R., Lee, W. L., Honan, T. A., . . . Dekker, J. (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res, 16*(10), 1299-1309. doi:10.1101/gr.5571506

Dupé, V., Ghyselinck, N. B., Thomazy, V., Nagy, L., Davies, P. J., Chambon, P., & Mark, M. (1999). Essential roles of retinoic acid signaling in interdigital apoptosis and control of BMP-7 expression in mouse autopods. *Dev Biol, 208*(1), 30-43. doi:10.1006/dbio.1998.9176

Durand, N. C., Shamim, M. S., Machol, I., Rao, S. S., Huntley, M. H., Lander, E. S., & Aiden, E. L. (2016). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst, 3*(1), 95-98. doi:10.1016/j.cels.2016.07.002

Eckalbar, W. L., Schlebusch, S. A., Mason, M. K., Gill, Z., Parker, A. V., Booker, B. M., . . . Ahituv, N. (2016). Transcriptomic and epigenomic characterization of the developing bat wing. *Nat Genet, 48*(5), 528-536. doi:10.1038/ng.3537

Erwin, D. H. (2020). Evolutionary dynamics of gene regulation. *Curr Top Dev Biol, 139*, 407-431. doi:10.1016/bs.ctdb.2020.02.006

Farré, M., Robinson, T. J., & Ruiz-Herrera, A. (2015). An Integrative Breakage Model of genome architecture, reshuffling and evolution: The Integrative Breakage Model of genome evolution, a novel multidisciplinary hypothesis for the study of genome plasticity. *Bioessays, 37*(5), 479-488. doi:10.1002/bies.201400174

Feenstra, J. M., Kanaya, K., Pira, C. U., Hoffman, S. E., Eppey, R. J., & Oberg, K. C. (2012). Detection of genes regulated by Lmx1b during limb dorsalization. *Dev Growth Differ, 54*(4), 451-462. doi:10.1111/j.1440-169X.2012.01331.x

Feigin, C. Y., Moreno, J. A., Ramos, R., Mereby, S. A., Alivisatos, A., Wang, W., . . . Mallarino, R. (2023). Convergent deployment of ancestral functions during the evolution of mammalian flight membranes. *9*(12), eade7511. doi:doi:10.1126/sciadv.ade7511

Feregrino, C., Sacher, F., Parnas, O., & Tschopp, P. (2019). A single-cell transcriptomic atlas of the developing chicken limb. *BMC Genomics, 20*(1), 401. doi:10.1186/s12864-019-5802-2

Feregrino, C., & Tschopp, P. (2022). Assessing evolutionary and developmental transcriptome dynamics in homologous cell types. *Dev Dyn, 251*(9), 1472-1489. doi:10.1002/dvdy.384

Feschotte, C. (2008). Transposable elements and the evolution of regulatory networks. *Nat Rev Genet, 9*(5), 397-405. doi:10.1038/nrg2337

Field, A., & Adelman, K. (2020). Evaluating Enhancer Function and Transcription. *Annu Rev Biochem, 89*, 213-234. doi:10.1146/annurev-biochem-011420-095916

Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A. K., . . . Gottardo, R. (2015). MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biology, 16*(1), 278. doi:10.1186/s13059-015-0844-5

Fischer, A. H., Jacobson, K. A., Rose, J., & Zeller, R. (2008). Hematoxylin and eosin staining of tissue and cell sections. *CSH Protoc, 2008*, pdb.prot4986. doi:10.1101/pdb.prot4986

Fogel, J. L., Thein, T. Z., & Mariani, F. V. (2012). Use of LysoTracker to detect programmed cell death in embryos and differentiating embryonic stem cells. *J Vis Exp*(68). doi:10.3791/4254

Fraser, J., Ferrai, C., Chiariello, A. M., Schueler, M., Rito, T., Laudanno, G., . . . Nicodemi, M. (2015). Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. *Mol Syst Biol, 11*(12), 852. doi:10.15252/msb.20156492

Fromental-Ramain, C., Warot, X., Messadecq, N., LeMeur, M., Dollé, P., & Chambon, P. (1996). Hoxa-13 and Hoxd-13 play a crucial role in the patterning of the limb autopod. *Development, 122*(10), 2997-3011. doi:10.1242/dev.122.10.2997

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., & Mirny, L. A. (2016). Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep, 15*(9), 2038-2049. doi:10.1016/j.celrep.2016.04.085

Furlong, E. E. M., & Levine, M. (2018). Developmental enhancers and chromosome topology. *Science, 361*(6409), 1341-1345. doi:10.1126/science.aau0320

Gañan, Y., Macias, D., Duterque-Coquillaud, M., Ros, M. A., & Hurle, J. M. (1996). Role of TGF beta s and BMPs as signals controlling the position of the digits and the areas of

interdigital cell death in the developing chick limb autopod. *Development, 122*(8), 2349-2357. doi:10.1242/dev.122.8.2349

Garcia-Martinez, V., Macias, D., Gañan, Y., Garcia-Lobo, J. M., Francia, M. V., Fernandez-Teran, M. A., & Hurle, J. M. (1993). Internucleosomal DNA fragmentation and programmed cell death (apoptosis) in the interdigital tissue of the embryonic chick leg bud. *J Cell Sci, 106 (Pt 1)*, 201-208. doi:10.1242/jcs.106.1.201

Garfield, D. A., & Wray, G. A. (2009). Comparative embryology without a microscope: using genomic approaches to understand the evolution of development. *J Biol, 8*(7), 65. doi:10.1186/jbiol161

Gaspar, J. M. J. G. r. G. (2018). Genrich: detecting sites of genomic enrichment.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res, 31*(13), 3784-3788. doi:10.1093/nar/gkg563

George, S. H., Gertsenstein, M., Vintersten, K., Korets-Smith, E., Murphy, J., Stevens, M. E., . . . Nagy, A. (2007). Developmental and adult phenotyping directly from mutant embryonic stem cells. *Proc Natl Acad Sci U S A, 104*(11), 4455-4460. doi:10.1073/pnas.0609277104

Ghyselinck, N. B., Dupé, V., Dierich, A., Messaddeq, N., Garnier, J. M., Rochette-Egly, C., . . . Mark, M. (1997). Role of the retinoic acid receptor beta (RARbeta) during mouse development. *Int J Dev Biol, 41*(3), 425-447.

Gibson-Brown, J. J., Agulnik, S. I., Chapman, D. L., Alexiou, M., Garvey, N., Silver, L. M., & Papaioannou, V. E. (1996). Evidence of a role for T-box genes in the evolution of limb morphogenesis and the specification of forelimb/hindlimb identity. *Mech Dev, 56*(1-2), 93-101. doi:10.1016/0925-4773(96)00514-x

Green, M. R. (2012). Molecular cloning : a laboratory manual. In J. Sambrook (Ed.), (4th ed. ed.). Cold Spring Harbor, N.Y. :: Cold Spring Harbor Laboratory Press.

Gunnell, G. F., & Simmons, N. B. (2005). Fossil Evidence and the Origin of Bats. *Journal of Mammalian Evolution, 12*(1), 209-246. doi:10.1007/s10914-005-6945-2

Haberle, V., & Stark, A. (2018). Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Biol, 19*(10), 621-637. doi:10.1038/s41580-018-0028-8

Han, L., Wei, X., Liu, C., Volpe, G., Zhuang, Z., Zou, X., . . . Liu, L. (2022). Cell transcriptomic atlas of the non-human primate Macaca fascicularis. *Nature, 604*(7907), 723-731. doi:10.1038/s41586-022-04587-3

Hao, Y., Stuart, T., Kowalski, M. H., Choudhary, S., Hoffman, P., Hartman, A., . . . Satija, R. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nature Biotechnology, 42*(2), 293-304. doi:10.1038/s41587-023-01767-y

Harmston, N., Ing-Simmons, E., Tan, G., Perry, M., Merkenschlager, M., & Lenhard, B. (2017). Topologically associating domains are ancient features that coincide with Metazoan clusters of extreme noncoding conservation. *Nat Commun, 8*(1), 441. doi:10.1038/s41467-017-00524-5

Hayashi, S., Akiyama, R., Wong, J., Tahara, N., Kawakami, H., & Kawakami, Y. (2016). Gata6-Dependent GLI3 Repressor Function is Essential in Anterior Limb Progenitor Cells for Proper Limb Development. *PLoS Genet, 12*(6), e1006138. doi:10.1371/journal.pgen.1006138

He, P., Williams, B. A., Trout, D., Marinov, G. K., Amrhein, H., Berghella, L., . . . Wold, B. J. (2020). The changing mouse embryo transcriptome at whole tissue and single-cell resolution. *Nature, 583*(7818), 760-767. doi:10.1038/s41586-020-2536-x

Hecker, N., & Hiller, M. (2020). A genome alignment of 120 mammals highlights ultraconserved element variability and placenta-associated enhancers. *Gigascience, 9*(1). doi:10.1093/gigascience/giz159

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., . . . Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell, 38*(4), 576-589. doi:10.1016/j.molcel.2010.05.004

Heinz, S., Romanoski, C. E., Benner, C., & Glass, C. K. (2015). The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol, 16*(3), 144-154. doi:10.1038/nrm3949

Hernández-Martínez, R., Castro-Obregón, S., & Covarrubias, L. (2009). Progressive interdigital cell death: regulation by the antagonistic interaction between fibroblast growth factor 8 and retinoic acid. *Development, 136*(21), 3669-3678. doi:10.1242/dev.041954

Hernández-Martínez, R., & Covarrubias, L. (2011). Interdigital cell death function and regulation: new insights on an old programmed cell death model. *Dev Growth Differ, 53*(2), 245-258. doi:10.1111/j.1440-169X.2010.01246.x

Heumos, L., Schaar, A. C., Lance, C., Litinetskaya, A., Drost, F., Zappia, L., . . . Single-cell Best Practices, C. (2023). Best practices for single-cell analysis across modalities. *Nature Reviews Genetics, 24*(8), 550-572. doi:10.1038/s41576-023-00586-w

Hippenmeyer, S., Youn, Y. H., Moon, H. M., Miyamichi, K., Zong, H., Wynshaw-Boris, A., & Luo, L. (2010). Genetic mosaic dissection of Lis1 and Ndel1 in neuronal migration. *Neuron, 68*(4), 695-709. doi:10.1016/j.neuron.2010.09.027

Hockman, D., Cretekos, C. J., Mason, M. K., Behringer, R. R., Jacobs, D. S., & Illing, N. (2008). A second wave of Sonic hedgehog expression during the development of the bat limb. *Proc Natl Acad Sci U S A, 105*(44), 16982-16987. doi:10.1073/pnas.0805308105

Holland, L. Z. (2013). Evolution of new characters after whole genome duplications: insights from amphioxus. *Semin Cell Dev Biol, 24*(2), 101-109. doi:10.1016/j.semcdb.2012.12.007

Hoskins, V. E., Smith, K., & Reddy, K. L. (2021). The shifting shape of genomes: dynamics of heterochromatin interactions at the nuclear lamina. *Curr Opin Genet Dev, 67*, 163-173. doi:10.1016/j.gde.2021.02.003

Hsia, L. T., Ashley, N., Ouaret, D., Wang, L. M., Wilding, J., & Bodmer, W. F. (2016). Myofibroblasts are distinguished from activated skin fibroblasts by the expression of AOC3 and other associated markers. *Proc Natl Acad Sci U S A, 113*(15), E2162-2171. doi:10.1073/pnas.1603534113

Huang, X., Henck, J., Qiu, C., Sreenivasan, V. K. A., Balachandran, S., Amarie, O. V., . . . Spielmann, M. (2023). Single-cell, whole-embryo phenotyping of mammalian developmental disorders. *Nature, 623*(7988), 772-781. doi:10.1038/s41586-023-06548-w

Jebb, D., Huang, Z., Pippel, M., Hughes, G. M., Lavrichenko, K., Devanna, P., . . . Teeling, E. C. (2020). Six reference-quality genomes reveal evolution of bat adaptations. *Nature, 583*(7817), 578-584. doi:10.1038/s41586-020-2486-3

Jiang, M., Xiao, Y., E, W., Ma, L., Wang, J., Chen, H., . . . Guo, G. (2021). Characterization of the Zebrafish Cell Landscape at Single-Cell Resolution. *Front Cell Dev Biol, 9*, 743421. doi:10.3389/fcell.2021.743421

Jiang, M., Xu, X., & Guo, G. (2021). Understanding embryonic development at single-cell resolution. *Cell Regen, 10*(1), 10. doi:10.1186/s13619-020-00074-0

Jindal, G. A., & Farley, E. K. (2021). Enhancer grammar in development, evolution, and disease: dependencies and interplay. *Dev Cell, 56*(5), 575-587. doi:10.1016/j.devcel.2021.02.016

Jones, R. C., Karkanias, J., Krasnow, M. A., Pisco, A. O., Quake, S. R., Salzman, J., . . . Wyss-Coray, T. (2022). The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science, 376*(6594), eabl4896. doi:10.1126/science.abl4896

Kalinka, A. T., & Tomancak, P. (2012). The evolution of early animal embryos: conservation or divergence? *Trends Ecol Evol, 27*(7), 385-393. doi:10.1016/j.tree.2012.03.007

Kaltcheva, M. M., Anderson, M. J., Harfe, B. D., & Lewandoski, M. (2016). BMPs are direct triggers of interdigital programmed cell death. *Dev Biol, 411*(2), 266-276. doi:10.1016/j.ydbio.2015.12.016

Kashgari, G., Meinecke, L., Gordon, W., Ruiz, B., Yang, J., Ma, A. L., . . . Andersen, B. (2020). Epithelial Migration and Non-adhesive Periderm Are Required for Digit Separation during Mammalian Development. *Dev Cell, 52*(6), 764-778.e764. doi:10.1016/j.devcel.2020.01.032

Kastner, P., Mark, M., Ghyselinck, N., Krezel, W., Dupé, V., Grondona, J. M., & Chambon, P. (1997). Genetic evidence that the retinoid signal is transduced by heterodimeric RXR/RAR functional units during mouse development. *Development, 124*(2), 313-326. doi:10.1242/dev.124.2.313

Kelly, N. H., Huynh, N. P. T., & Guilak, F. (2020). Single cell RNA-sequencing reveals cellular heterogeneity and trajectories of lineage specification during murine embryonic limb development. *Matrix Biol, 89*, 1-10. doi:10.1016/j.matbio.2019.12.004

Khokha, M. K., Hsu, D., Brunet, L. J., Dionne, M. S., & Harland, R. M. (2003). Gremlin is the BMP antagonist required for maintenance of Shh and Fgf signals during limb patterning. *Nat Genet, 34*(3), 303-307. doi:10.1038/ng1178

Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M., & Altman, D. G. (2010). Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *J Pharmacol Pharmacother, 1*(2), 94-99. doi:10.4103/0976-500X.72351

Kobak, D., & Berens, P. (2019). The art of using t-SNE for single-cell transcriptomics. *Nat Commun, 10*(1), 5416. doi:10.1038/s41467-019-13056-x

Koji, K., & Takashi, F. (2024). Regulatory landscape of enhancer-mediated transcriptional activation. *Trends in Cell Biology*. doi:https://doi.org/10.1016/j.tcb.2024.01.008

Kozhemyakina, E., Ionescu, A., & Lassar, A. B. (2014). GATA6 is a crucial regulator of Shh in the limb bud. *PLoS Genet, 10*(1), e1004072. doi:10.1371/journal.pgen.1004072

Kraft, K., Geuer, S., Will, A. J., Chan, W. L., Paliou, C., Borschiwer, M., . . . Andrey, G. (2015). Deletions, Inversions, Duplications: Engineering of Structural Variants using CRISPR/Cas in Mice. *Cell Rep, 10*(5), 833-839. doi:10.1016/j.celrep.2015.01.016

Kvon, E. Z. (2015). Using transgenic reporter assays to functionally characterize enhancers in animals. *Genomics, 106*(3), 185-192. doi:10.1016/j.ygeno.2015.06.007

Kvon, E. Z., Kamneva, O. K., Melo, U. S., Barozzi, I., Osterwalder, M., Mannion, B. J., . . . Visel, A. (2016). Progressive Loss of Function in a Limb Enhancer during Snake Evolution. *Cell, 167*(3), 633-642.e611. doi:10.1016/j.cell.2016.09.028

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., . . . Kharchenko, P. V. (2018). RNA velocity of single cells. *Nature, 560*(7719), 494-498. doi:10.1038/s41586-018-0414-6

Lange, M., Bergen, V., Klein, M., Setty, M., Reuter, B., Bakhti, M., . . . Theis, F. J. (2022). CellRank for directed single-cell fate mapping. *Nature Methods, 19*(2), 159-170. doi:10.1038/s41592-021-01346-6

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods, 9*(4), 357-359. doi:10.1038/nmeth.1923

Leal, F., & Cohn, M. J. (2016). Loss and Re-emergence of Legs in Snakes by Modular Evolution of Sonic hedgehog and HOXD Enhancers. *Curr Biol, 26*(21), 2966-2973. doi:10.1016/j.cub.2016.09.020

Lee, K.-W., Yeo, S.-Y., Gong, J.-R., Koo, O.-J., Sohn, I., Lee, W. Y., . . . Kim, S.-H. (2022). PRRX1 is a master transcription factor of stromal fibroblasts for myofibroblastic lineage progression. *Nat Commun, 13*(1), 2793. doi:10.1038/s41467-022-30484-4

Lettice, L. A., Daniels, S., Sweeney, E., Venkataraman, S., Devenney, P. S., Gautier, P., . . . FitzPatrick, D. R. (2011). Enhancer-adoption as a mechanism of human developmental disease. *Hum Mutat, 32*(12), 1492-1499. doi:10.1002/humu.21615

Lewandoski, M., Sun, X., & Martin, G. R. (2000). Fgf8 signalling from the AER is essential for normal limb development. *Nat Genet, 26*(4), 460-463. doi:10.1038/82609

Li, D., He, M., Tang, Q., Tian, S., Zhang, J., Li, Y., . . . Li, M. (2022). Comparative 3D genome architecture in vertebrates. *BMC Biol, 20*(1), 99. doi:10.1186/s12915-022-01301-7

Li, H. (2022). Physiologic and pathophysiologic roles of AKAP12. *Sci Prog, 105*(3), 368504221109212. doi:10.1177/00368504221109212

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics, 26*(5), 589-595. doi:10.1093/bioinformatics/btp698

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics, 25*(16), 2078-2079. doi:10.1093/bioinformatics/btp352

Li, Y., Chen, C. Y., Kaye, A. M., & Wasserman, W. W. (2015). The identification of cis-regulatory elements: A review from a machine learning perspective. *Biosystems, 138*, 6-17. doi:10.1016/j.biosystems.2015.10.002

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics, 30*(7), 923-930. doi:10.1093/bioinformatics/btt656

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., . . . Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science, 326*(5950), 289-293. doi:10.1126/science.1181369

Linderman, G. C., Rachh, M., Hoskins, J. G., Steinerberger, S., & Kluger, Y. (2019). Fast interpolation-based t-SNE for improved visualization of single-cell RNA-seq data. *Nature Methods, 16*(3), 243-245. doi:10.1038/s41592-018-0308-4

Ling, Y., Hongbing, L., Mingquan, Y., Jing, Y., Fanxin, L., Ken, M., & YiPing, C. (2007). Shox2 is required for chondrocyte proliferation and maturation in proximal limb skeleton. *Dev Biol, 306*(2), 549-559. doi:https://doi.org/10.1016/j.ydbio.2007.03.518

Logan, C., Hornbruch, A., Campbell, I., & Lumsden, A. (1997). The role of Engrailed in establishing the dorsoventral axis of the chick limb. *Development, 124*(12), 2317-2324. doi:10.1242/dev.124.12.2317

Logan, M., Martin, J. F., Nagy, A., Lobe, C., Olson, E. N., & Tabin, C. J. (2002). Expression of Cre Recombinase in the developing mouse limb bud driven by a Prxl enhancer. *Genesis, 33*(2), 77-80. doi:10.1002/gene.10092

Logan, M., & Tabin, C. J. (1999). Role of Pitx1 upstream of Tbx4 in specification of hindlimb identity. *Science, 283*(5408), 1736-1739. doi:10.1126/science.283.5408.1736

Long, H. K., Prescott, S. L., & Wysocka, J. (2016). Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell, 167*(5), 1170-1187. doi:10.1016/j.cell.2016.09.018

Long, H. S., Greenaway, S., Powell, G., Mallon, A. M., Lindgren, C. M., & Simon, M. M. (2022). Making sense of the linear genome, gene function and TADs. *Epigenetics Chromatin, 15*(1), 4. doi:10.1186/s13072-022-00436-9

López-Delgado, A. C., Delgado, I., Cadenas, V., Sánchez-Cabo, F., & Torres, M. (2021). Axial skeleton anterior-posterior patterning is regulated through feedback regulation between Meis transcription factors and retinoic acid. *Development, 148*(1). doi:10.1242/dev.193813

Lopez-Rios, J., Duchesne, A., Speziale, D., Andrey, G., Peterson, K. A., Germann, P., . . . Zeller, R. (2014). Attenuated sensing of SHH by Ptch1 underlies evolution of bovine limbs. *Nature, 511*(7507), 46-51. doi:10.1038/nature13289

Lorda-Diez, C. I., Garcia-Riart, B., Montero, J. A., Rodriguez-León, J., Garcia-Porrero, J. A., & Hurle, J. M. (2015). Apoptosis during embryonic tissue remodeling is accompanied by cell senescence. *Aging (Albany NY), 7*(11), 974-985. doi:10.18632/aging.100844

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology, 15*(12), 550. doi:10.1186/s13059-014-0550-8

Lu, P., Minowada, G., & Martin, G. R. (2006). Increasing Fgf4 expression in the mouse limb bud causes polysyndactyly and rescues the skeletal defects that result from loss of Fgf8 function. *Development, 133*(1), 33-42. doi:10.1242/dev.02172

Luo, G., Hofmann, C., Bronckers, A. L., Sohocki, M., Bradley, A., & Karsenty, G. (1995). BMP-7 is an inducer of nephrogenesis, and is also required for eye development and skeletal patterning. *Genes Dev, 9*(22), 2808-2820. doi:10.1101/gad.9.22.2808

Luo, X., Liu, Y., Dang, D., Hu, T., Hou, Y., Meng, X., . . . Su, B. (2021). 3D Genome of macaque fetal brain reveals evolutionary innovations during primate corticogenesis. *Cell, 184*(3), 723-740.e721. doi:10.1016/j.cell.2021.01.001

Lupiáñez, D. G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., . . . Mundlos, S. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell, 161*(5), 1012-1025. doi:10.1016/j.cell.2015.04.004

Lupiáñez, D. G., Spielmann, M., & Mundlos, S. (2016). Breaking TADs: How Alterations of Chromatin Domains Result in Disease. *Trends Genet, 32*(4), 225-237. doi:10.1016/j.tig.2016.01.003

Lussier, M., Canoun, C., Ma, C., Sank, A., & Shuler, C. (1993). Interdigital soft tissue separation induced by retinoic acid in mouse limbs cultured in vitro. *Int J Dev Biol, 37*(4), 555-564.

Ma, Y., Yu, L., Pan, S., Gao, S., Chen, W., Zhang, X., . . . Zhang, L. (2017). CRISPR/Cas9-mediated targeting of the Rosa26 locus produces Cre reporter rat strains for monitoring Cre-loxP-mediated lineage tracing. *Febs j, 284*(19), 3262-3277. doi:10.1111/febs.14188

Maatouk, D. M., Choi, K. S., Bouldin, C. M., & Harfe, B. D. (2009). In the limb AER Bmp2 and Bmp4 are required for dorsal-ventral patterning and interdigital cell death but not limb outgrowth. *Dev Biol, 327*(2), 516-523. doi:10.1016/j.ydbio.2009.01.004

Macias, D., Gañan, Y., Ros, M. A., & Hurle, J. M. (1996). In vivo inhibition of programmed cell death by local administration of FGF-2 and FGF-4 in the interdigital areas of the embryonic chick leg bud. *Anat Embryol (Berl), 193*(6), 533-541. doi:10.1007/bf00187925

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., . . . McCarroll, S. A. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell, 161*(5), 1202-1214. doi:10.1016/j.cell.2015.05.002

Maeso, I., Irimia, M., Tena, J. J., Casares, F., & Gómez-Skarmeta, J. L. (2013). Deep conservation of cis-regulatory elements in metazoans. *Philos Trans R Soc Lond B Biol Sci, 368*(1632), 20130020. doi:10.1098/rstb.2013.0020

Maier, J. A., Rivas-Astroza, M., Deng, J., Dowling, A., Oboikovitz, P., Cao, X., . . . Sears, K. E. (2017). Transcriptomic insights into the genetic basis of mammalian limb diversity. *BMC Evol Biol, 17*(1), 86. doi:10.1186/s12862-017-0902-6

Mansisidor, A. R., & Risca, V. I. (2022). Chromatin accessibility: methods, mechanisms, and biological insights. *Nucleus, 13*(1), 236-276. doi:10.1080/19491034.2022.2143106

Markman, S., Zada, M., David, E., Giladi, A., Amit, I., & Zelzer, E. (2023). A single-cell census of mouse limb development identifies complex spatiotemporal dynamics of skeleton formation. *Dev Cell, 58*(7), 565-581.e564. doi:10.1016/j.devcel.2023.02.013

Marlétaz, F., de la Calle-Mustienes, E., Acemel, R. D., Paliou, C., Naranjo, S., Martínez-García, P. M., . . . Gómez-Skarmeta, J. L. (2023). The little skate genome and the evolutionary emergence of wing-like fins. *Nature, 616*(7957), 495-503. doi:10.1038/s41586-023-05868-1

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011, 17*(1), 3 %J EMBnet.journal. doi:10.14806/ej.17.1.200

Mason, M. K., Hockman, D., Curry, L., Cunningham, T. J., Duester, G., Logan, M., . . . Illing, N. (2015). Retinoic acid-independent expression of Meis2 during autopod patterning in the developing bat and mouse limb. *Evodevo, 6*, 6. doi:10.1186/s13227-015-0001-y

Mayshar, Y., Raz, O., Cheng, S., Ben-Yair, R., Hadas, R., Reines, N., . . . Stelzer, Y. (2023). Time-aligned hourglass gastrulation models in rabbit and mouse. *Cell, 186*(12), 2610-2627.e2618. doi:10.1016/j.cell.2023.04.037

McArthur, E., Rinker, D. C., Gilbertson, E. N., Fudenberg, G., Pittman, M., Keough, K., . . . Capra, J. A. (2022). Reconstructing the 3D genome organization of Neanderthals reveals that chromatin folding shaped phenotypic and sequence divergence. 2022.2002.2007.479462. doi:10.1101/2022.02.07.479462 %J bioRxiv

McLennan, D. A. (2008). The Concept of Co-option: Why Evolution Often Looks Miraculous. *Evolution: Education and Outreach, 1*(3), 247-258. doi:10.1007/s12052-008-0053-8

McQueen, C., & Towers, M. (2020). Establishing the pattern of the vertebrate limb. *Development, 147*(17). doi:10.1242/dev.177956

Medina-Jiménez, B. I., Budd, G. E., & Janssen, R. (2024). Single-cell RNA sequencing of mid-to-late stage spider embryos: new insights into spider development. *BMC Genomics, 25*(1), 150. doi:10.1186/s12864-023-09898-x

Mercader, N., Leonardo, E., Azpiazu, N., Serrano, A., Morata, G., Martínez, C., & Torres, M. (1999). Conserved regulation of proximodistal limb axis development by Meis1/Hth. *Nature, 402*(6760), 425-429. doi:10.1038/46580

Mercader, N., Leonardo, E., Piedra, M. E., Martínez, A. C., Ros, M. A., & Torres, M. (2000). Opposing RA and FGF signals control proximodistal vertebrate limb development through regulation of Meis genes. *Development, 127*(18), 3961-3970. doi:10.1242/dev.127.18.3961

Mercader, N., Selleri, L., Criado, L. M., Pallares, P., Parras, C., Cleary, M. L., & Torres, M. (2009). Ectopic Meis1 expression in the mouse limb bud alters P-D patterning in a Pbx1-independent manner. *Int J Dev Biol, 53*(8-10), 1483-1494. doi:10.1387/ijdb.072430nm

Mercader, N., Tanaka, E. M., & Torres, M. (2005). Proximodistal identity during vertebrate limb regeneration is regulated by Meis homeodomain proteins. *Development, 132*(18), 4131-4142. doi:10.1242/dev.01976

Meredith, R. W., Westerman, M., & Springer, M. S. (2009). A phylogeny of Diprotodontia (Marsupialia) based on sequences for five nuclear genes. *Mol Phylogenet Evol, 51*(3), 554-571. doi:10.1016/j.ympev.2009.02.009

Merino, R., Rodriguez-Leon, J., Macias, D., Gañan, Y., Economides, A. N., & Hurle, J. M. (1999). The BMP antagonist Gremlin regulates outgrowth, chondrogenesis and programmed cell death in the developing limb. *Development, 126*(23), 5515-5522. doi:10.1242/dev.126.23.5515

Montero, J. A., & Hurlé, J. M. (2010). Sculpturing digit shape by cell death. *Apoptosis, 15*(3), 365-375. doi:10.1007/s10495-009-0444-5

Moore, R. C., & Purugganan, M. D. (2003). The early stages of duplicate gene evolution. *Proc Natl Acad Sci U S A, 100*(26), 15682-15687. doi:10.1073/pnas.2535513100

Mora, A., Sandve, G. K., Gabrielsen, O. S., & Eskeland, R. (2016). In the loop: promoter-enhancer interactions and bioinformatics. *Brief Bioinform, 17*(6), 980-995. doi:10.1093/bib/bbv097

Moreno, J. A., Dudchenko, O., Feigin, C. Y., Mereby, S. A., Chen, Z., Ramos, R., . . . Mallarino, R. (2024). Emx2 underlies the development and evolution of marsupial gliding membranes. *Nature, 629*(8010), 127-135. doi:10.1038/s41586-024-07305-3

Morgan, B. A., & Tabin, C. J. (1993). The role of homeobox genes in limb development. *Curr Opin Genet Dev, 3*(4), 668-674. doi:10.1016/0959-437x(93)90105-x

Mori, C., Nakamura, N., Kimura, S., Irie, H., Takigawa, T., & Shiota, K. (1995). Programmed cell death in the interdigital tissue of the fetal mouse limb is apoptosis with DNA fragmentation. *Anat Rec, 242*(1), 103-110. doi:10.1002/ar.1092420114

Morov, A. R., Ukizintambara, T., Sabirov, R. M., & Yasui, K. (2016). Acquisition of the dorsal structures in chordate amphioxus. *Open Biol, 6*(6). doi:10.1098/rsob.160062

Mueller, T., Hellmann, T., & Nickel, J. (2012). Missense Mutations in GDF-5 Signaling: Molecular Mechanisms Behind Skeletal Malformation. doi:10.5772/35195

Muhl, L., Genové, G., Leptidis, S., Liu, J., He, L., Mocci, G., . . . Betsholtz, C. (2020). Single-cell analysis uncovers fibroblast heterogeneity and criteria for fibroblast and mural cell identification and discrimination. *Nat Commun, 11*(1), 3953. doi:10.1038/s41467-020-17740-1

Murgai, A., Altmeyer, S., Wiegand, S., Tylzanowski, P., & Stricker, S. (2018). Cooperation of BMP and IHH signaling in interdigital cell fate determination. *PLoS One, 13*(5), e0197535. doi:10.1371/journal.pone.0197535

Musser, J. M., Schippers, K. J., Nickel, M., Mizzon, G., Kohn, A. B., Pape, C., . . . Arendt, D. (2021). Profiling cellular diversity in sponges informs animal cell type and nervous system evolution. *Science, 374*(6568), 717-723. doi:10.1126/science.abj2949

Ng, J. K., Kawakami, Y., Büscher, D., Raya, A., Itoh, T., Koth, C. M., . . . Izpisúa Belmonte, J. C. (2002). The limb identity gene Tbx5 promotes limb initiation by interacting with Wnt2b and Fgf10. *Development, 129*(22), 5161-5170. doi:10.1242/dev.129.22.5161

Ngo-Muller, V., & Muneoka, K. (2000). Influence of FGF4 on digit morphogenesis during limb development in the mouse. *Dev Biol, 219*(2), 224-236. doi:10.1006/dbio.2000.9612

Nora, E. P., Goloborodko, A., Valton, A. L., Gibcus, J. H., Uebersohn, A., Abdennur, N., . . . Bruneau, B. G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell, 169*(5), 930-944.e922. doi:10.1016/j.cell.2017.05.004

Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., . . . Heard, E. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature, 485*(7398), 381-385. doi:10.1038/nature11049

O'Connor, L. J., Schoech, A. P., Hormozdiari, F., Gazal, S., Patterson, N., & Price, A. L. (2019). Extreme Polygenicity of Complex Traits Is Explained by Negative Selection. *Am J Hum Genet, 105*(3), 456-476. doi:10.1016/j.ajhg.2019.07.003

Ohuchi, H., Nakagawa, T., Yamamoto, A., Araga, A., Ohata, T., Ishimaru, Y., . . . Noji, S. (1997). The mesenchymal factor, FGF10, initiates and maintains the outgrowth of the chick limb bud through interaction with FGF8, an apical ectodermal factor. *Development, 124*(11), 2235-2244. doi:10.1242/dev.124.11.2235

Ong, C. T., & Corces, V. G. (2012). Enhancers: emerging roles in cell fate specification. *EMBO Rep, 13*(5), 423-430. doi:10.1038/embor.2012.52

Pajni-Underwood, S., Wilson, C. P., Elder, C., Mishina, Y., & Lewandoski, M. (2007). BMP signals control limb bud interdigital programmed cell death by regulating FGF signaling. *Development, 134*(12), 2359-2368. doi:10.1242/dev.001677

Paliou, C., Guckelberger, P., Schöpflin, R., Heinrich, V., Esposito, A., Chiariello, A. M., . . . Andrey, G. (2019). Preformed chromatin topology assists transcriptional robustness of Shh during limb development. *Proc Natl Acad Sci U S A, 116*(25), 12390-12399. doi:10.1073/pnas.1900672116

Palstra, R. J., & Grosveld, F. (2012). Transcription factor binding at enhancers: shaping a genomic regulatory landscape in flux. *Front Genet, 3*, 195. doi:10.3389/fgene.2012.00195

Palstra, R. J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., & de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. *Nat Genet, 35*(2), 190-194. doi:10.1038/ng1244

Panman, L., Galli, A., Lagarde, N., Michos, O., Soete, G., Zuniga, A., & Zeller, R. (2006). Differential regulation of gene expression in the digit forming area of the mouse limb bud by SHH and gremlin 1/FGF-mediated epithelial-mesenchymal signalling. *Development, 133*(17), 3419-3428. doi:10.1242/dev.02529

Pantalacci, S., & Sémon, M. (2015). Transcriptomics of developing embryos and organs: A raising tool for evo-devo. *J Exp Zool B Mol Dev Evol, 324*(4), 363-371. doi:10.1002/jez.b.22595

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*(6520), 350-353. doi:10.1038/374350a0

Pearse, R. V., 2nd, Scherz, P. J., Campbell, J. K., & Tabin, C. J. (2007). A cellular lineage analysis of the chick limb bud. *Dev Biol, 310*(2), 388-400. doi:10.1016/j.ydbio.2007.08.002

Pennacchio, L. A., Bickmore, W., Dean, A., Nobrega, M. A., & Bejerano, G. (2013). Enhancers: five essential questions. *Nat Rev Genet, 14*(4), 288-295. doi:10.1038/nrg3458

Peter, I. S., & Davidson, E. H. (2011). Evolution of gene regulatory networks controlling body plan development. *Cell, 144*(6), 970-985. doi:10.1016/j.cell.2011.02.017

Petit, F., Sears, K. E., & Ahituv, N. (2017). Limb development: a paradigm of gene regulation. *Nat Rev Genet, 18*(4), 245-258. doi:10.1038/nrg.2016.167

Phan, M. H. Q., Zehnder, T., Puntieri, F., Lo, B.-W., Lenhard, B., Mueller, F., . . . Ibrahim, D. M. (2024). Conservation of Regulatory Elements with Highly Diverged Sequences Across Large Evolutionary Distances. *bioRxiv*, 2024.2005.2013.590087. doi:10.1101/2024.05.13.590087

Pineault, K. M., & Wellik, D. M. (2014). Hox genes and limb musculoskeletal development. *Curr Osteoporos Rep, 12*(4), 420-427. doi:10.1007/s11914-014-0241-0

Plikus, M. V., Wang, X., Sinha, S., Forte, E., Thompson, S. M., Herzog, E. L., . . . Horsley, V. (2021). Fibroblasts: Origins, definitions, and functions in health and disease. *Cell, 184*(15), 3852-3872. doi:10.1016/j.cell.2021.06.024

Porter, A. G., & Jänicke, R. U. (1999). Emerging roles of caspase-3 in apoptosis. *Cell Death Differ, 6*(2), 99-104. doi:10.1038/sj.cdd.4400476

Qiu, L. Q., Lai, W. S., Stumpo, D. J., & Blackshear, P. J. (2016). Mouse Embryonic Fibroblast Cell Culture and Stimulation. *Bio Protoc, 6*(13). doi:10.21769/BioProtoc.1859

Rallis, C., Bruneau, B. G., Del Buono, J., Seidman, C. E., Seidman, J. G., Nissim, S., . . . Logan, M. P. (2003). Tbx5 is required for forelimb bud formation and continued outgrowth. *Development, 130*(12), 2741-2751. doi:10.1242/dev.00473

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., . . . Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res, 44*(W1), W160-165. doi:10.1093/nar/gkw257

Ramisch, A., Heinrich, V., Glaser, L. V., Fuchs, A., Yang, X., Benner, P., . . . Vingron, M. (2019). CRUP: a comprehensive framework to predict condition-specific regulatory units. *Genome Biology, 20*(1), 227. doi:10.1186/s13059-019-1860-7

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols, 8*(11), 2281-2308. doi:10.1038/nprot.2013.143

Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., . . . Aiden, E. L. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell, 159*(7), 1665-1680. doi:10.1016/j.cell.2014.11.021

Rasweiler, J. J. t., & Badwaik, N. K. (1997). Delayed development in the short-tailed fruit bat, Carollia perspicillata. *J Reprod Fertil, 109*(1), 7-20. doi:10.1530/jrf.0.1090007

Rasweiler, J. J. t., & de Bonilla, H. (1992). Menstruation in short-tailed fruit bats (Carollia spp.). *J Reprod Fertil, 95*(1), 231-248. doi:10.1530/jrf.0.0950231

Rauluseviciute, I., Riudavets-Puig, R., Blanc-Mathieu, R., Castro-Mondragon, Jaime A., Ferenc, K., Kumar, V., . . . Mathelier, A. (2023). JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. *Nucleic Acids Res, 52*(D1), D174-D182. doi:10.1093/nar/gkad1059 %J Nucleic Acids Research

Real, F. M., Haas, S. A., Franchini, P., Xiong, P., Simakov, O., Kuhl, H., . . . Lupiáñez, D. G. (2020). The mole genome reveals regulatory rearrangements associated with adaptive intersexuality. *Science, 370*(6513), 208-214. doi:10.1126/science.aaz2582

Rebeiz, M., & Tsiantis, M. (2017). Enhancer evolution and the origins of morphological novelty. *Curr Opin Genet Dev, 45*, 115-123. doi:10.1016/j.gde.2017.04.006

Remeseiro, S., Hörnblad, A., & Spitz, F. (2016). Gene regulation during development in the light of topologically associating domains. *Wiley Interdiscip Rev Dev Biol, 5*(2), 169-185. doi:10.1002/wdev.218

Riddle, R. D., Johnson, R. L., Laufer, E., & Tabin, C. (1993). Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell, 75*(7), 1401-1416. doi:10.1016/0092-8674(93)90626-2

Robson, M. I., Ringel, A. R., & Mundlos, S. (2019). Regulatory Landscaping: How Enhancer-Promoter Communication Is Sculpted in 3D. *Mol Cell, 74*(6), 1110-1122. doi:10.1016/j.molcel.2019.05.032

Rodriguez-Esteban, C., Tsukui, T., Yonei, S., Magallon, J., Tamura, K., & Izpisua Belmonte, J. C. (1999). The T-box genes Tbx4 and Tbx5 regulate limb outgrowth and identity. *Nature, 398*(6730), 814-818. doi:10.1038/19769

Rodriguez-Leon, J., Merino, R., Macias, D., Gañan, Y., Santesteban, E., & Hurle, J. M. (1999). Retinoic acid regulates programmed cell death through BMP signalling. *Nat Cell Biol, 1*(2), 125-126. doi:10.1038/10098

Romero, I. G., Ruvinsky, I., & Gilad, Y. (2012). Comparative studies of gene expression and the evolution of gene regulation. *Nat Rev Genet, 13*(7), 505-516. doi:10.1038/nrg3229

Roux, J., Rosikiewicz, M., & Robinson-Rechavi, M. (2015). What to compare and how: Comparative transcriptomics for Evo-Devo. *J Exp Zool B Mol Dev Evol, 324*(4), 372-382. doi:10.1002/jez.b.22618

Sadier, A., Urban, D. J., Anthwal, N., Howenstine, A. O., Sinha, I., & Sears, K. E. (2021). Making a bat: The developmental basis of bat evolution. *Genet Mol Biol, 43*(1 Suppl 2), e20190146. doi:10.1590/1678-4685-gmb-2019-0146

Sanetra, M., Begemann, G., Becker, M. B., & Meyer, A. (2005). Conservation and co-option in developmental programmes: the importance of homology relationships. *Front Zool, 2*, 15. doi:10.1186/1742-9994-2-15

Saxena, A., Sharma, V., Muthuirulan, P., Neufeld, S. J., Tran, M. P., Gutierrez, H. L., . . . Cooper, K. L. (2022). Interspecies transcriptomics identify genes that underlie disproportionate foot growth in jerboas. *Curr Biol, 32*(2), 289-303.e286. doi:10.1016/j.cub.2021.10.063

Scherz, P. J., Harfe, B. D., McMahon, A. P., & Tabin, C. J. (2004). The limb bud Shh-Fgf feedback loop is terminated by expansion of former ZPA cells. *Science, 305*(5682), 396-399. doi:10.1126/science.1096966

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., . . . Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods, 9*(7), 676-682. doi:10.1038/nmeth.2019

Schindler, M., Osterwalder, M., Harabula, I., Wittler, L., Tzika, A. C., Dechmann, D. K. N., . . . Real, F. M. (2023). Induction of kidney-related gene programs through co-option of SALL1 in mole ovotestes. *Development, 150*(17). doi:10.1242/dev.201562

Schmitz, J., Ohme, M., Suryobroto, B., & Zischler, H. (2002). The colugo (Cynocephalus variegatus, Dermoptera): the primates' gliding sister? *Mol Biol Evol, 19*(12), 2308-2312. doi:10.1093/oxfordjournals.molbev.a004054

Schoenfelder, S., & Fraser, P. (2019). Long-range enhancer-promoter contacts in gene expression control. *Nat Rev Genet, 20*(8), 437-455. doi:10.1038/s41576-019-0128-0

Schulte, D., & Geerts, D. (2019). MEIS transcription factors in development and disease. *Development, 146*(16). doi:10.1242/dev.174706

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., . . . Spitz, F. (2017). Two independent modes of chromatin organization revealed by cohesin removal. *Nature, 551*(7678), 51-56. doi:10.1038/nature24281

Sears, K. E., Behringer, R. R., Rasweiler, J. J. t., & Niswander, L. A. (2006). Development of bat flight: morphologic and molecular evolution of bat wing digits. *Proc Natl Acad Sci U S A, 103*(17), 6581-6586. doi:10.1073/pnas.0509716103

Sears, K. E., Bormet, A. K., Rockwell, A., Powers, L. E., Noelle Cooper, L., & Wheeler, M. B. (2011). Developmental basis of mammalian digit reduction: a case study in pigs. *Evol Dev, 13*(6), 533-541. doi:10.1111/j.1525-142X.2011.00509.x

Sekine, K., Ohuchi, H., Fujiwara, M., Yamasaki, M., Yoshizawa, T., Sato, T., . . . Kato, S. (1999). Fgf10 is essential for limb and lung formation. *Nat Genet, 21*(1), 138-141. doi:10.1038/5096

Shafer, M. E. R., Sawh, A. N., & Schier, A. F. (2022). Gene family evolution underlies cell-type diversification in the hypothalamus of teleosts. *Nat Ecol Evol, 6*(1), 63-76. doi:10.1038/s41559-021-01580-3

Shapiro, M. D., Hanken, J., & Rosenthal, N. (2003). Developmental basis of evolutionary digit loss in the Australian lizard Hemiergis. *J Exp Zool B Mol Dev Evol, 297*(1), 48-56. doi:10.1002/jez.b.19

Sheeba, C. J., & Logan, M. P. O. (2017). Chapter Twelve - The Roles of T-Box Genes in Vertebrate Limb Development. *Curr Top Dev Biol, 122*, 355-381. doi:https://doi.org/10.1016/bs.ctdb.2016.08.009

Shin, H., Shi, Y., Dai, C., Tjong, H., Gong, K., Alber, F., & Zhou, X. J. (2016). TopDom: an efficient and deterministic method for identifying topological domains in genomes. *Nucleic Acids Res, 44*(7), e70. doi:10.1093/nar/gkv1505

Shlyueva, D., Stampfel, G., & Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. *Nat Rev Genet, 15*(4), 272-286. doi:10.1038/nrg3682

Silbereis, J. C., Pochareddy, S., Zhu, Y., Li, M., & Sestan, N. (2016). The Cellular and Molecular Landscapes of the Developing Human Central Nervous System. *Neuron, 89*(2), 248-268. doi:10.1016/j.neuron.2015.12.008

Simmons, N. B., Seymour, K. L., Habersetzer, J., & Gunnell, G. F. (2008). Primitive Early Eocene bat from Wyoming and the evolution of flight and echolocation. *Nature, 451*(7180), 818-821. doi:10.1038/nature06549

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., . . . de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet, 38*(11), 1348-1354. doi:10.1038/ng1896

Song, Y., Miao, Z., Brazma, A., & Papatheodorou, I. (2023). Benchmarking strategies for cross-species integration of single-cell RNA sequencing data. *Nat Commun, 14*(1), 6495. doi:10.1038/s41467-023-41855-w

Soussi, G., Girdziusaite, A., Jhanwar, S., Palacio, V., Notaro, M., Sheth, R., . . . Zuniga, A. (2024). TBX3 is essential for establishment of the posterior boundary of anterior genes and upregulation of posterior genes together with HAND2 during the onset of limb bud development. *Development, 151*(11). doi:10.1242/dev.202722

Spitz, F., & Furlong, E. E. (2012). Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet, 13*(9), 613-626. doi:10.1038/nrg3207

Stewart, S., Yi, S., Kassabian, G., Mayo, M., Sank, A., & Shuler, C. (2000). Changes in expression of the lysosomal membrane glycoprotein, LAMP-1 in interdigital regions during embryonic mouse limb development, in vivo and in vitro. *Anat Embryol (Berl), 201*(6), 483-490. doi:10.1007/s004290050335

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., 3rd, . . . Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell, 177*(7), 1888-1902.e1821. doi:10.1016/j.cell.2019.05.031

Stuart, T., & Satija, R. (2019). Integrative single-cell analysis. *Nat Rev Genet, 20*(5), 257-272. doi:10.1038/s41576-019-0093-7

Symmons, O., Uslu, V. V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., . . . Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. *Genome Res, 24*(3), 390-400. doi:10.1101/gr.163519.113

Tanay, A., & Sebé-Pedrós, A. (2021). Evolutionary cell type mapping with single-cell genomics. *Trends Genet, 37*(10), 919-932. doi:10.1016/j.tig.2021.04.008

Tasic, B., Hippenmeyer, S., Wang, C., Gamboa, M., Zong, H., Chen-Tsai, Y., & Luo, L. (2011). Site-specific integrase-mediated transgenesis in mice via pronuclear injection. *Proc Natl Acad Sci U S A, 108*(19), 7902-7907. doi:10.1073/pnas.1019507108

ten Berge, D., Brugmann, S. A., Helms, J. A., & Nusse, R. (2008). Wnt and FGF signals interact to coordinate growth with cell fate specification during limb development. *Development, 135*(19), 3247-3257. doi:10.1242/dev.023176

Thewissen, J. G., Cohn, M. J., Stevens, L. S., Bajpai, S., Heyning, J., & Horton, W. E., Jr. (2006). Developmental basis for hind-limb loss in dolphins and origin of the cetacean bodyplan. *Proc Natl Acad Sci U S A, 103*(22), 8414-8418. doi:10.1073/pnas.0602920103

Tickle, C., & Towers, M. (2017). Sonic Hedgehog Signaling in Limb Development. *Front Cell Dev Biol, 5*, 14. doi:10.3389/fcell.2017.00014

Tissières, V., Geier, F., Kessler, B., Wolf, E., Zeller, R., & Lopez-Rios, J. (2020). Gene Regulatory and Expression Differences between Mouse and Pig Limb Buds Provide Insights into the Evolutionary Emergence of Artiodactyl Traits. *Cell Rep, 31*(1), 107490. doi:10.1016/j.celrep.2020.03.054

Tominaga, H., Nishitsuji, K., & Satoh, N. (2023). A single-cell RNA-seq analysis of early larval cell-types of the starfish, Patiria pectinifera: Insights into evolution of the chordate body plan. *Dev Biol, 496*, 52-62. doi:10.1016/j.ydbio.2023.01.009

Tran, H. T. N., Ang, K. S., Chevrier, M., Zhang, X., Lee, N. Y. S., Goh, M., & Chen, J. (2020). A benchmark of batch-effect correction methods for single-cell RNA sequencing data. *Genome Biology, 21*(1), 12. doi:10.1186/s13059-019-1850-9

Tsompana, M., & Buck, M. J. (2014). Chromatin accessibility: a window into the genome. *Epigenetics Chromatin, 7*(1), 33. doi:10.1186/1756-8935-7-33

Tümpel, S., Sanz-Ezquerro, J. J., Isaac, A., Eblaghie, M. C., Dobson, J., & Tickle, C. (2002). Regulation of Tbx3 expression by anteroposterior signalling in vertebrate limb development. *Dev Biol, 250*(2), 251-262.

Vallecillo-García, P., Orgeur, M., Vom Hofe-Schneider, S., Stumm, J., Kappert, V., Ibrahim, D. M., . . . Stricker, S. (2017). Odd skipped-related 1 identifies a population of embryonic fibro-adipogenic progenitors regulating myogenesis during limb development. *Nat Commun, 8*(1), 1218. doi:10.1038/s41467-017-01120-3

Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., . . . Aerts, S. (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nature Protocols, 15*(7), 2247-2276. doi:10.1038/s41596-020-0336-2

van den Eijnde, S. M., Luijsterburg, A. J., Boshart, L., De Zeeuw, C. I., van Dierendonck, J. H., Reutelingsperger, C. P., & Vermeij-Keers, C. (1997). In situ detection of apoptosis during embryogenesis with annexin V: from whole mount to ultrastructure. *Cytometry, 29*(4), 313-320. doi:10.1002/(sici)1097-0320(19971201)29:4<313::aid-cyto8>3.0.co;2-a

Varga, Z., & Varga, M. (2022). Gene expression changes during the evolution of the tetrapod limb. *Biol Futur, 73*(4), 411-426. doi:10.1007/s42977-022-00136-1

Vergara, H. M., Pape, C., Meechan, K. I., Zinchenko, V., Genoud, C., Wanner, A. A., . . . Arendt, D. (2021). Whole-body integration of gene expression and single-cell morphology. *Cell, 184*(18), 4819-4837.e4822. doi:10.1016/j.cell.2021.07.017

Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D. T., Tanay, A., & Hadjur, S. (2015). Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep, 10*(8), 1297-1309. doi:10.1016/j.celrep.2015.02.004

Villacorte, M., Suzuki, K., Hayashi, K., de Sousa Lopes, S. C., Haraguchi, R., Taketo, M. M., . . . Yamada, G. (2010). Antagonistic crosstalk of Wnt/beta-catenin/Bmp signaling within the Apical Ectodermal Ridge (AER) regulates interdigit formation. *Biochem Biophys Res Commun, 391*(4), 1653-1657. doi:10.1016/j.bbrc.2009.12.109

Villar, D., Berthelot, C., Aldridge, S., Rayner, T. F., Lukk, M., Pignatelli, M., . . . Odom, D. T. (2015). Enhancer evolution across 20 mammalian species. *Cell, 160*(3), 554-566. doi:10.1016/j.cell.2015.01.006

Visel, A., Bristow, J., & Pennacchio, L. A. (2007). Enhancer identification through comparative genomics. *Semin Cell Dev Biol, 18*(1), 140-152. doi:10.1016/j.semcdb.2006.12.014

Wachtler, F., Christ, B., & Jacob, H. J. (1981). On the determination of mesodermal tissues in the avian embryonic wing bud. *Anat Embryol (Berl), 161*(3), 283-289. doi:10.1007/bf00301826

Waddington, C. J. (1957). The Strategy of the Genes.

Wang, F., Ding, P., Liang, X., Ding, X., Brandt, C. B., Sjöstedt, E., . . . Luo, Y. (2022). Endothelial cell heterogeneity and microglia regulons revealed by a pig cell landscape at single-cell level. *Nat Commun, 13*(1), 3620. doi:10.1038/s41467-022-31388-z

Wang, M., Di Pietro-Torres, A., Feregrino, C., Luxey, M., Moreau, C., Fischer, S., . . . Tschopp, P. (2024). Distinct Gene Regulatory Dynamics Drive Skeletogenic Cell Fate Convergence During Vertebrate Embryogenesis. 2024.2003.2026.586769. doi:10.1101/2024.03.26.586769 %J bioRxiv

Wang, Z., Dai, M., Wang, Y., Cooper, K. L., Zhu, T., Dong, D., . . . Zhang, S. (2014). Unique expression patterns of multiple key genes associated with the evolution of mammalian flight. *Proc Biol Sci, 281*(1783), 20133133. doi:10.1098/rspb.2013.3133

Wang, Z., Dong, D., Ru, B., Young, R. L., Han, N., Guo, T., & Zhang, S. (2010). Digital gene expression tag profiling of bat digits provides robust candidates contributing to wing formation. *BMC Genomics, 11*, 619. doi:10.1186/1471-2164-11-619

Weatherbee, S. D., Behringer, R. R., Rasweiler, J. J. t., & Niswander, L. A. (2006). Interdigital webbing retention in bat wings illustrates genetic changes underlying amniote limb diversification. *Proc Natl Acad Sci U S A, 103*(41), 15103-15107. doi:10.1073/pnas.0604934103

Wellik, D. M., & Capecchi, M. R. (2003). Hox10 and Hox11 genes are required to globally pattern the mammalian skeleton. *Science, 301*(5631), 363-367. doi:10.1126/science.1085672

Williamson, I., Hill, R. E., & Bickmore, W. A. (2011). Enhancers: from developmental genetics to the genetics of common human disease. *Dev Cell, 21*(1), 17-19. doi:10.1016/j.devcel.2011.06.008

Wittkopp, P. J., & Kalay, G. (2011). Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet, 13*(1), 59-69. doi:10.1038/nrg3095

Woych, J., Ortega Gurrola, A., Deryckere, A., Jaeger, E. C. B., Gumnit, E., Merello, G., . . . Tosches, M. A. (2022). Cell-type profiling in salamanders identifies innovations in vertebrate forebrain evolution. *Science, 377*(6610), eabp9186. doi:10.1126/science.abp9186

Wray, G. A. (2007). The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet, 8*(3), 206-216. doi:10.1038/nrg2063

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., . . . Yu, G. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb), 2*(3), 100141. doi:10.1016/j.xinn.2021.100141

Xu, X., Weinstein, M., Li, C., Naski, M., Cohen, R. I., Ornitz, D. M., . . . Deng, C. (1998). Fibroblast growth factor receptor 2 (FGFR2)-mediated reciprocal regulation loop between FGF8 and FGF10 is essential for limb induction. *Development, 125*(4), 753-765. doi:10.1242/dev.125.4.753

Yokouchi, Y., Sakiyama, J., Kameda, T., Iba, H., Suzuki, A., Ueno, N., & Kuroiwa, A. (1996). BMP-2/-4 mediate programmed cell death in chicken limb buds. *Development, 122*(12), 3725-3734. doi:10.1242/dev.122.12.3725

Zabidi, M. A., & Stark, A. (2016). Regulatory Enhancer-Core-Promoter Communication via Transcription Factors and Cofactors. *Trends Genet, 32*(12), 801-814. doi:10.1016/j.tig.2016.10.003

Zakany, J., & Duboule, D. (2007). The role of Hox genes during vertebrate limb development. *Curr Opin Genet Dev, 17*(4), 359-366. doi:10.1016/j.gde.2007.05.011

Zakeri, Z., Quaglino, D., & Ahuja, H. S. (1994). Apoptotic cell death in the mouse limb and its suppression in the hammertoe mutant. *Dev Biol, 165*(1), 294-297. doi:10.1006/dbio.1994.1255

Zambrowicz, B. P., Imamoto, A., Fiering, S., Herzenberg, L. A., Kerr, W. G., & Soriano, P. (1997). Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc Natl Acad Sci U S A, 94*(8), 3789-3794. doi:10.1073/pnas.94.8.3789

Zehnder, T. (2021). Computational Approaches for the Prediction of Gene Regulatory Elements and the Analysis of their Evolutionary Conservation. doi:http://dx.doi.org/10.17169/refubium-28880

Zehnder, T., Benner, P., & Vingron, M. (2019). Predicting enhancers in mammalian genomes using supervised hidden Markov models. *BMC Bioinformatics, 20*(1), 157. doi:10.1186/s12859-019-2708-6

Zeller, R., López-Ríos, J., & Zuniga, A. (2009). Vertebrate limb bud development: moving towards integrative analysis of organogenesis. *Nat Rev Genet, 10*(12), 845-858. doi:10.1038/nrg2681

Zeng, H. (2022). What is a cell type and how to define it? *Cell, 185*(15), 2739-2755. doi:10.1016/j.cell.2022.06.031

Zhang, B., He, P., Lawrence, J. E. G., Wang, S., Tuck, E., Williams, B. A., . . . Teichmann, S. A. (2023). A human embryonic limb cell atlas resolved in space and time. *Nature*. doi:10.1038/s41586-023-06806-x

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., . . . Liu, X. S. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biology, 9*(9), R137. doi:10.1186/gb-2008-9-9-r137

Zhao, X., Brade, T., Cunningham, T. J., & Duester, G. (2010). Retinoic acid controls expression of tissue remodeling genes Hmgn1 and Fgf18 at the digit-interdigit junction. *Dev Dyn, 239*(2), 665-671. doi:10.1002/dvdy.22188

Zheng, G. X., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., . . . Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nat Commun, 8*, 14049. doi:10.1038/ncomms14049

Zhong, J., Aires, R., Tsissios, G., Skoufa, E., Brandt, K., Sandoval-Guzmán, T., & Aztekin, C. (2023). Multi-species atlas resolves an axolotl limb development and regeneration paradox. *Nat Commun, 14*(1), 6346. doi:10.1038/s41467-023-41944-w

Zhu, M., & Tabin, C. J. (2023). The role of timing in the development and evolution of the limb. *Front Cell Dev Biol, 11*, 1135519. doi:10.3389/fcell.2023.1135519

Zúñiga, A., Haramis, A. P., McMahon, A. P., & Zeller, R. (1999). Signal relay by BMP antagonism controls the SHH/FGF4 feedback loop in vertebrate limb buds. *Nature, 401*(6753), 598-602. doi:10.1038/44157

Zuzarte-Luis, V., Berciano, M. T., Lafarga, M., & Hurlé, J. M. (2006). Caspase redundancy and release of mitochondrial apoptotic factors characterize interdigital apoptosis. *Apoptosis, 11*(5), 701-715. doi:10.1007/s10495-006-5481-8

Zuzarte-Luis, V., & Hurle, J. M. (2005). Programmed cell death in the embryonic vertebrate limb. *Semin Cell Dev Biol, 16*(2), 261-269. doi:10.1016/j.semcdb.2004.12.004

Zuzarte-Luís, V., & Hurlé, J. M. (2002). Programmed cell death in the developing limb. *Int J Dev Biol, 46*(7), 871-876.

10. Appendix

10.1. Extended figures

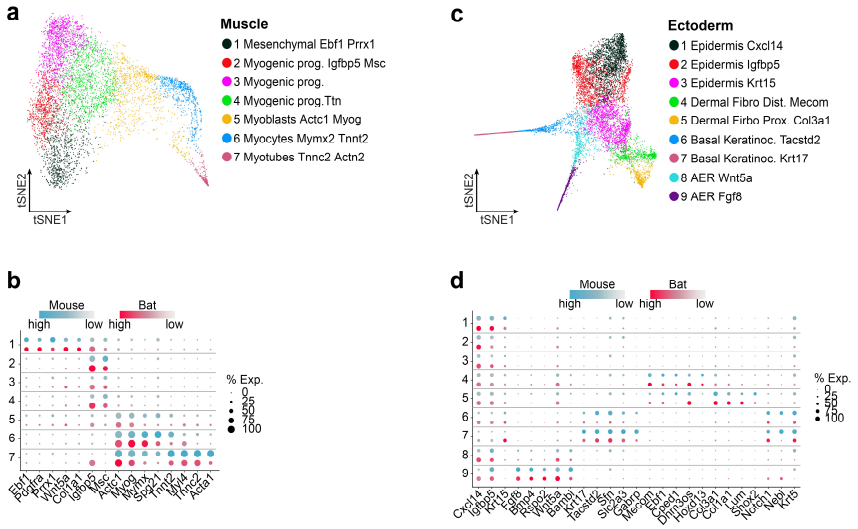

Figure 30: Muscle and ectoderm sub-clusters.

a Sub-clustering of muscle cells. Marker gene expression used for integrated cluster annotation (blue: mouse, red: bat). **b** Dot size represents percentage of cells expressing the respective marker gene; dot intensity indicates expression level. **c** Sub-clustering of ectodermal cells. **d** Dot size represents percentage of cells expressing the respective marker gene; dot intensity indicates expression level.

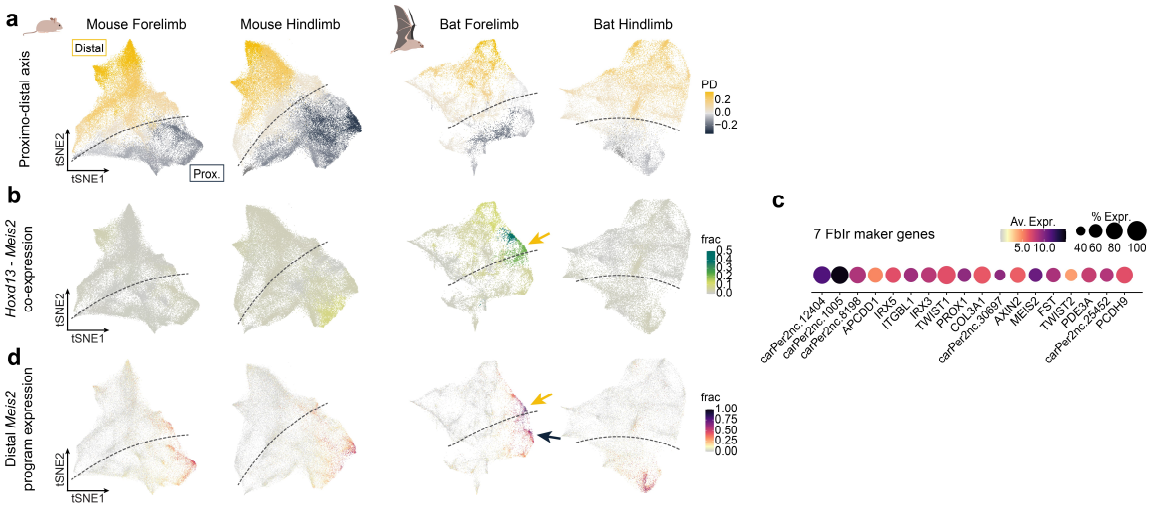

Figure 31: Bat fibroblast cluster 7 Fblr gene program.

a Proximal and distal cell identity as in Figure 19. **b** Co-expression of Hoxd13 and Meis2 as in Figure 19. The arrow indicates cluster 7 Fblr. **c** Marker genes of bat cluster 7 Fblr based on differential gene expression of this cluster versus the remaining LPM-derived forelimb cells. **d** Expression of gene set of bat cluster 7 Fblr in mouse and bat fore- and hindlimbs. Shwon is the fraction of co-expression score of the whole marker set shown in c.

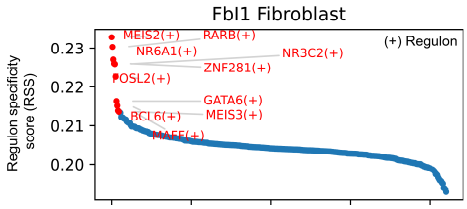

Figure 32: Positive regulons of bat fibroblast cluster 10 Fbl1.

Positive regulons from pySCENIC gene regulatory network analysis. Shown is the regulon specificity score (RSS) for bat cluster 10 Fbl1.

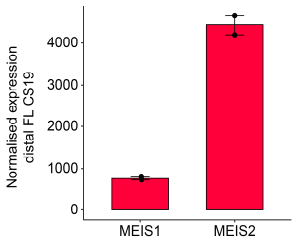

Figure 33: Normalized expression levels of *MEIS1* and *MEIS2* in the distal bat forelimb at CS19.

Shown are the normalized expression levels of *MEIS1* and *MEIS2* in dissected distal bat forelimbs at CS19. The error bars show the standard error. n=2.

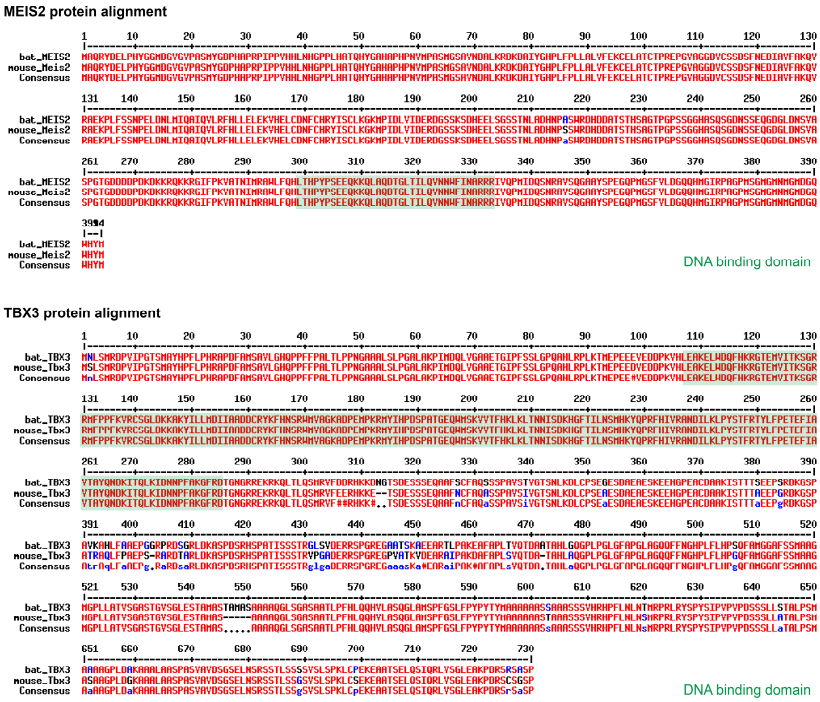

Figure 34: MEIS2 and TBX3 protein alignment.

Sequence alignment of mouse and bat MEIS2 and TBX3 protein sequences. The DNA binding domains are highlighted in blue.

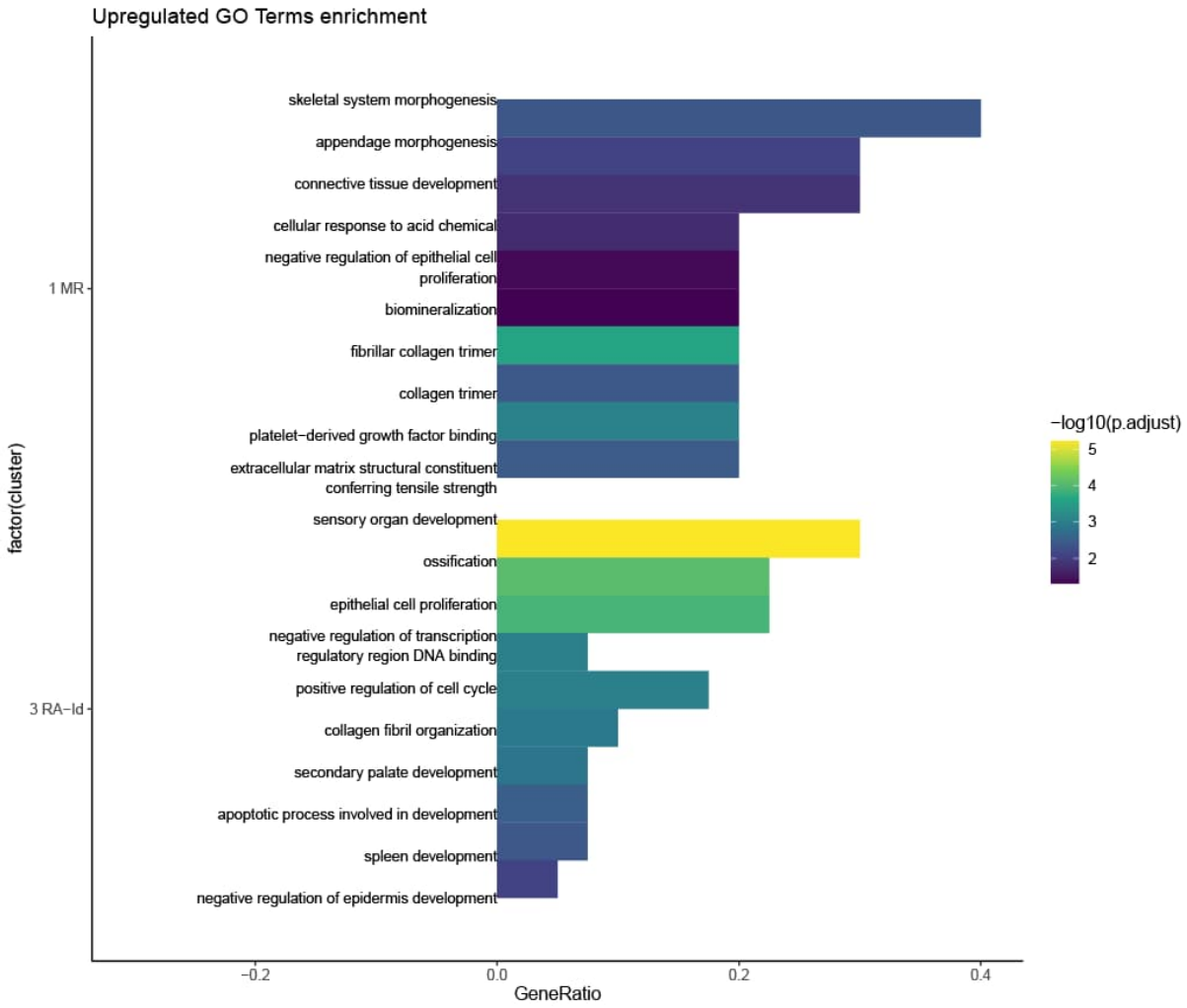

Figure 35: GO terms enriched in genes upregulated in MEIS2 mutant.

Shown are the top 10 enriched GO terms in the affected clusters 1 MR and 3 RA-Id of the MEIS2 mutant with GeneRatio and the adjusted p-value.

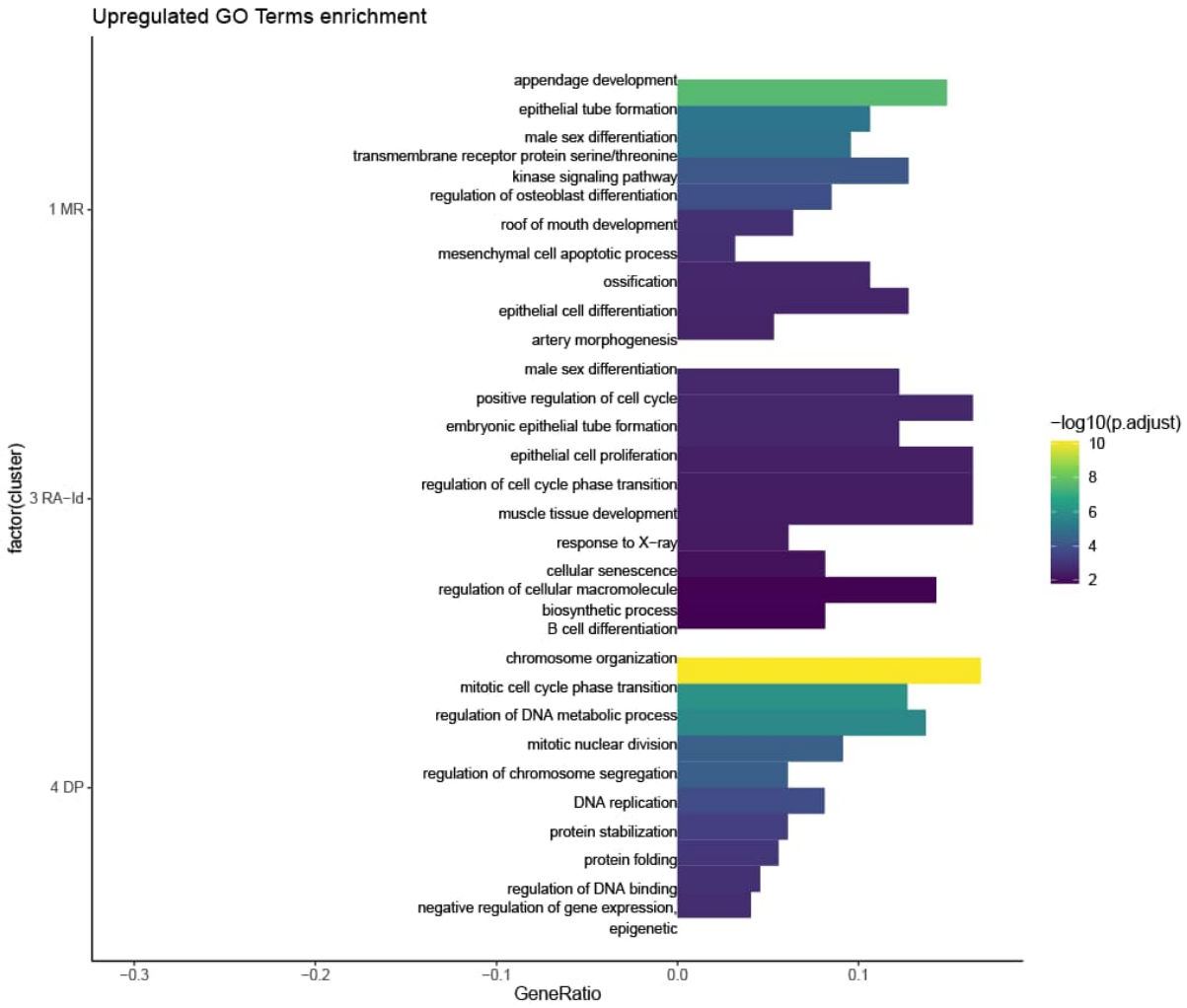

Figure 36: GO terms enriched in genes upregulated in TBX3 mutant.

Shown are the top 10 enriched GO terms in the affected clusters 1 MR, 3 RA-Id and 4 DP of the MEIS2 mutant with GeneRatio and the adjusted p-value.

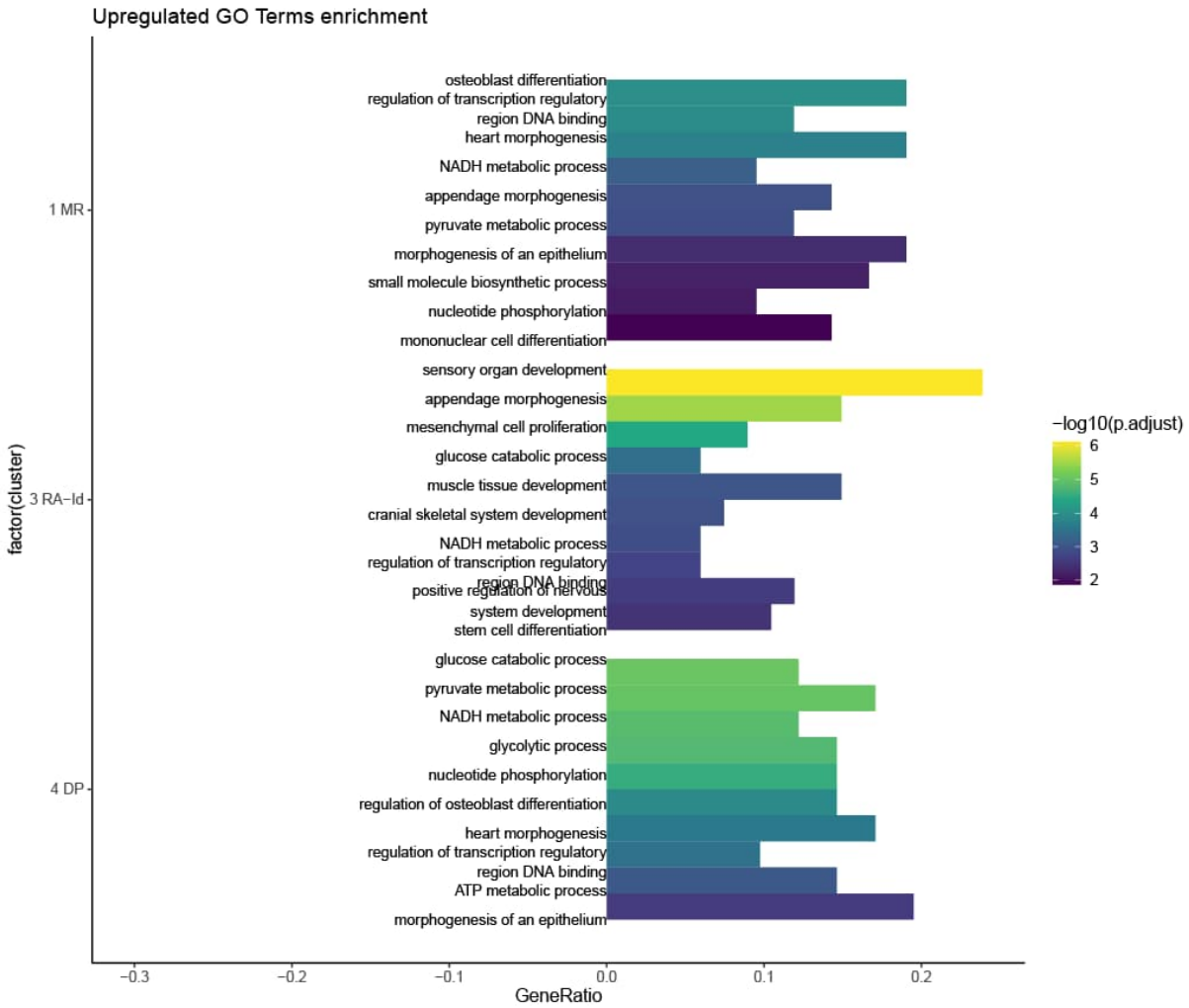

Figure 37: GO terms enriched in genes upregulated in MEIS2-TBX3 mutant.

Shown are the top 10 enriched GO terms in the affected clusters 1 MR and 3 RA-Id of the MEIS2-TBX3 mutant with GeneRatio and the adjusted p-value.

10.2. List of figures

10.3. List of tables

10.4. Primers, oligonucleotides and recombinant DNA

The primers used for cloning expression constructs are listed in Table 2.

Table 2: Cloning primer sequences used in this study.

Primer	*Sequence (5' → 3')*
H11_Bmp2_h_er_FW	GCTGAAGCTGATGGAACAgcGCCATGGCATTAATCAGACA
Bmp2_h_er_Hsp68_RV	ggctgctcagtttggatgttTTCAGCACACCGTGCTTATC
Bmp2_h_er_Hsp68_FW	GATAAGCACGGTGTGCTGAAaacatccaaactgagcagcc
Hsp68_meis2_RV	ATCGTACCTTTGCGCCATcagtctGGCGCCGCGCTCTGCT
Hsp68_Meis2_FW	AGCAGAGCGCGGCGCCagactgATGGCGCAAAGGTACGAT
H11_Bmp2_h_er_RV	TGTCTGATTAATGCCATGGCgcTGTTCCATCAGCTTCAGC
cpTbx3_Dathe_FW	CCTTCCAGAAGCAGAGCGCGGCGCCGAGTGGATGAACCTCTCCAT
cpTbx3_Dathe_RV	TCTTGCTGATCATGATTAGTGTTTGgatccagacatgataagata
cpTbx3_Dathe_BB_FW	tatcttatcatgtctggatcCAAACACTAATCATGATCAGCAAGA
cpTbx3_Dathe_BB_RV	ATGGAGAGGTTCATCCACTCGGCGCCGCGCTCTGCTTCTGGAAGG
cpTbx3-DatheR26_fw	TGGGAGAATCCCTTCcccctcttccGCCATGGCATTAATCAGACA
cpTbx3-DatheR26_rv	AAAGACTGGAgttgcagatcacgaggatccagacatgataagata
cpTbx3-DatheR26_BB_fw	caatgtatcttatcatgtctggatcctcgtgatctgcaacTCCAG
cpTbx3-DatheR26_BB_rv	ACATATGTCTGATTAATGCCATGGCggaagaggggGAAGGGATTC

The primer sequences used for genotyping of genome edited mESCs are listed in Table 3.

Table 3: Genotyping primer sequences used in this study.

Primer	*Sequence (5' → 3')*
Gen_Chr13_qPCR_F	GGGAGCTGACACCACTATTTAC
Gen_Chr13_qPCR_R	ACATTAAACCCTGGGGGAAG
qPCR_cpMeis2_Dathe_FW	TGGTTCCATCCAGAGACAAGC
qPCR_cpMeis2_Dathe_RV	ATCGTACCTTTGCGCCATCA
qPCR_cpTbx3_Dathe_FW	TCCAGAGACAAGCGAAGACA
qPCR_cpTbx3_Dathe_RV	GAGGTAGGAACGGATGGTAGG
GenoES-H11-3HA-FW	TTCACTGCATTCTAGTTGTGGTTTG
GenoES-H11-3HA-RV	GTTGGCAGTTTTGGCCAGTT
GenoES-H11-5HA-FW	GTTAGGCTTGTGTCAACTGTTTG
GenoES_DatheEn_RV	CCTTTCTGGTCATTGAAAGGCAC
geno_cpTbx3_Dathe5HA_RV	TCTTGTGGCTTACAAATGGAGT
R265HA_DatheTbx_fwd	GGAAAGACAACAAACACCTGAA

R265HA_DatheTbx_rv	GCGTCCCCATAAAAACaac

The single guide RNA sequences used for knock-ins at H11 and Rosa26 loci are listed in Table 4.

Table 4: sgRNA oligonucleotide sequences used in this study.

Primer	*Sequence (5' → 3')*
sg_H11_FW	GCTGATGGAACAGGTAACAA
sg_H11_RV	TTGTTACCTGTTCCATCAGC
sg_R26_FW	GACTGGAGTTGCAGATCACGA
sg_R25_RV	TCGTGATCTGCAACTCCAGTC

The primer sequences used to generate antisense RNA WISH probes are listed in Table 5.

Table 5: Primer sequences used for the generation of WISH probes.

Primer	*Sequence (5' → 3')*
WISH_Hsp68-cpMeis2_FW	TCTGGTTCCATCCAGAGACAAG
WISH_Hsp68-cpMeis2_RV	ATGGTCAGCGAGGTTTGTGG
SP6	ATTTAGGTGACACTATAG
T7	TAATACGACTCACTATAGGG

Table 6: Recombinant DNA and constructs used in this study.

Recombinant DNA	*Source*
pSpCas9(BB)-2A-Puro (PX459) V2.0 vector	Addgene, #62988
H11-back-pTwist+Amp+High+Copy	Twist Bioscience
cp-Meis2-cDNA	Twist Bioscience
cp-Tbx3-pUC-GW-Amp	GeneWiz, Azenta Life Science
H11-Hsp68-LacZ	NA

10.5. Bridging species

Table 7: Bridging species used for IPP.

Ciona intestinalis
Carollia perspicillata
Zebra fish
Chicken
Human
Opossum

Mouse
Rat
Pig
Mole
Platypus
Red eared slider turtle

10.6. Instruments and Software

The instruments used in this study are listed in Table 8.

Table 8: List of instruments used in this study.

Instrument	*Type*	*Manufacturer*
Bioruptor	UCD 300	Diagenode
Camera	DFC420	Leica
Chromium Controller	iX	10X Genomics
CO_2 Incubator	HEPA Class 100	Thermo Scientific
Cooling centrifuge	5417R	Eppendorf
Cooling centrifuge	Avanti J-E	Beckman-Coulter
Cooling centrifuge rotor	JLA16.250	Beckman-Coulter
Light source	CL9000	Zeiss
qPCR thermocycler	ABIPrism HAT 79000 RT	Applied Biosystems
Sonicator	S220	Covaris
Stereomicroscope	MZ12 Discovery V12	Zeiss
Tabletop centrifuge	5414D	Eppendorf
Thermocycler	GeneAmp PCR System 2700, 2720, and 9700	Applied Biosystems

Table 9: Software, algorithms and scripts used in this study.

Softwares and algorithms	*Source*
Adobe Illustrator and Photoshop	https://www.adobe.com/
Bowtie2	(Langmead & Salzberg, 2012)
BWA 0.7.17	(H. Li & Durbin, 2010)
CellRanger v6.0.2	(Zheng et al., 2017)
CellRank v2.0.4	(Lange et al., 2022)
clusertProfiler	(Wu et al., 2021)
CRUP	(Ramisch et al., 2019)
cutadapt	(Martin, 2011)

deepTools	(Ramírez et al., 2016)
DESeq2 v1.38.3 (R v4.2.2)	(Love et al., 2014)
edgeR v3.40.2	(Chen et al., 2016)
eHMM	(T. Zehnder et al., 2019)
ExPASy	(Gasteiger et al., 2003)
featureCounts	(Liao et al., 2014)
FFT-accelerated Interpolation-based t-SNE	(Linderman et al., 2019)
Fiji	(Schindelin et al., 2012)
Genrich	(Gaspar, 2018)
Graphpad Prism	https://www.graphpad.com/scientific-software/prism/
Homer2	(Heinz et al., 2010)
IMARIS Analysis Software	Oxford Instruments
IPP	https://github.com/tobiaszehnder/IPP
JASPAR	(Rauluseviciute et al., 2023)
Juicer v1.5.6	(Durand et al., 2016)
Macs2	(Y. Zhang et al., 2008)
MAST	(Finak et al., 2015)
MultiAlin	(Corpet, 1988)
pySCENIC	(Van de Sande et al., 2020)
SAMtools	(H. Li et al., 2009)
Scvelo v0.3.2	(Bergen et al., 2020)
Seurat v4.3.0	(Hao et al., 2024)
Snapgene	https://www.snapgene.com/
STAR_2.6.1d	(Dobin et al., 2013)
TopDom	(Shin et al., 2016)
UCSC genome browser	https://genome.ucsc.edu
Uniprot	(Consortium, 2022)
Velocyto	(La Manno et al., 2018)
ZEN software	Zeiss

10.7. List of Abbreviations

°C *Degree celcius*

µL *microliter*

µm *micrometer*

µM *Micromolar*

3C *Chromosome conformation capture*

AER *Apical ectodermal ridge*

ATAC *Assay for transposase-accessible chromatin*

ATP *Adenosine triphosphate*

BAR *Bat accelerated region*

BMP *Bone morphogenic protein*

bp *base pair*

BSA *Bovine serum albumin*

cDNA *Complementary DNA*

ChIP-seq *Chromatin immunoprecipitation DNA-sequencing*

CO_2 *Carbon dioxide*

CRE *cis-regulatory element*

CS *Carnegie stage*

CTCF *CCCTC-binding factor*

dATP *Deoxyadenosine triphosphate*

DC *Directly conserved*

dCTP *Deoxycytosine triphosphate*

DEG *Differentially expressed gene*

dGTP *Deoxyguanine triphosphate*

DIG *Digoxigenin*

DMEM *Dulbeccos Modified Eagle Medium*

DMSO *Dimethyl sulfoxide*

DNA *deoxyribonucleic acid*

DPBS *Dulbecco's phosphate buffered saline*

dsDNA *double stranded DNA*

dTTP *Deoxythyrosine triphosphate*

DV *Dorsal-ventral*

E *Embryonic day post coitum*

E. coli *Escherichia coli*

EDTA *Ethylenediaminetetraacetic acid*

En1 *Engrailed 1*

ESC *Embryonic stem cell*

evo-devo *Evolutionary developmental biology*

FCS *Fetal calf serum*

Fgf Fibroblast growth factor

GEM *Gel Beads-in-emulsion*

GO *Gene ontology*

Grem1 Gremlin 1

GTF *General transcription factor*

HCl *Hydrogen chloride*

HDR *Homology directed repair*

Hsp68 *Heat shock protein 68*

IC *Indirectly conserved*

kb *kilobase pair*

LAGeSo *Landesamt für Gesundheit und Soziales*

lfc *log-fold change*

LPM *Lateral plate mesoderm*

Mb *Megabase pair*

mESC *Mouse embryonic stem cell*

$MgCl_2$ *Magnesium chloride*

min *minute*

mRNA *Messenger RNA*

NaCl *Sodium chloride*

NC *Not conserved*

PBS *Phosphate buffered saline*

PC *Principal component*

PCR *polymerase chain reaction*

PD *Proximo-distal*

PFA *Paraformaldehyde*

RA *Retinoic acid*

RNA *ribonucleic acid*

RNA-seq *RNA sequencing*

RT *Reverse transcription*

scRNA-seq *single-cell RNA-sequencing*

SDS *Sodium dodecyl sulfate*

sec *second*

sgRNA *Single guide RNA*

SHH *Sonic hedgehog*

TAD *Topologically associated domain*

TF *Transcription factor*

TFBS *Transcription factor binding site*

tSNE *t-distributed stochastic neighbor embedding*

TSS *transcription start site*

UMI *Unique molecular identifier*

WISH *Whole mount in situ hybridisation*

ZPA *Zone of polarizing activity*

11. Acknowledgements

First and foremost, I would like to express my gratitude to Prof. Dr. Stefan Mundlos for granting me the opportunity to work on this exciting project. I also want to thank Prof. Dr. Sigmar Stricker for kindly agreeing to serve as the second reviewer for this thesis.

A special thank you goes to Fany for her guidance throughout my doctorate. Her expertise and mentorship have been invaluable, not only in shaping this work but also in contributing significantly to my personal and professional development.

I would also like to express my gratitude to all collaborators for their help and contributions to this work. I would especially like to thank Nicolas Fasel and the Papiliorama Foundation and Julio Hechavarria and his team, without whom this project and all future research related to it would not have been possible.

Of course, I also want to thank all former and current members of the Mundlos team:
Alex, Alicia, Alessa, Anand, Andrea, Andreas, Anna, Asita, Bai-Wei, Bjørt, Blanka, Carola, César, Christian, Christina, Daniel, Darío, Fany, Fiona, Gabi, Gianluca, Giulia, Guillaume, Jana, Jessy, Josh, Juliane, Konrad, Lila, Lion, Masha, Mike, Mikie, Milan, Natalia, Natalia, Norbert, Philipp, Robert, Sala, Silvia, Tobi, Uirá, Ute, Vanessa and Vera. Thank you all for your help and input, the fruitful discussions, the many cakes, celebrations and afterwork drinks. I especially want to thank Asita and Ute for the experimental and organizational support of the entire team. The lab would cease to exist without you.
A very special thank you goes to Christian for all the countless plots and scribbles, for patiently answering all of my questions and for his support, not only as co-author but also as a trusted ally and advocate.

I am deeply grateful to my friends, both old and new, for their unwavering support, for every unforgettable experience, and for always being there when I needed them.
My deepest gratitude, however, is reserved for my family - especially my parents, Angelika and Peter, and my husband Mike. Your unconditional support throughout all the highs and lows means everything to me. I would not be who I am, nor where I am today, without you.