A Field Guide to The Limb Progenitor

A Field Guide to The Limb Progenitor

A dissertation presented

by

Alan Rodrigues

to

The Department of Molecular and Cellular Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biology

Harvard University

Cambridge, Massachusetts

March 2016

Dissertation Advisor: Dr. Clifford J. Tabin Alan Rodrigues

A Field Guide to The Limb Progenitor

Abstract

The primary goal of this thesis was to characterize the embryonic limb progenitor, a cell type that populates the early limb bud during the onset of limb morphogenesis. Two features of the limb progenitor were explored 1) the identity of a potential set of minimal transcription factors that are sufficient to instantiate limb progenitor identity and 2) the limb progenitor gene expression response to signaling molecules that are critical for regulating the ultimate morphology of the developing limb.

The limb bud consists of mesenchymal progenitor cells, or limb progenitors, that originate from the lateral plate mesoderm (LPM). In addition to the limb progenitors of the limb bud, the LPM also gives rise to mesenchyme that will form the mesodermal components of the trunk, flank and neck. However, only the cells of the limb are capable of forming the patterned skeletal structures observed in arms and legs. The other LPM-derived mesenchymal progenitor cells are unable to form such structures, even when placed in the context of the organizing signals found in the limb bud. To determine which genes are responsible for specifying a limb progenitor identity that is distinct from the identity of neighboring mesenchymal tissues, a screen of transcription

factors capable of imparting limb progenitor-like properties on non-limb embryonic and post embryonic fibroblasts was conducted. A minimal subset of genes was identified with the capability of producing limb progenitor-like cells that are similar in gene expression and functional capacity as endogenous limb progenitors.

During embryonic development, fields of progenitor cells form complex spatial structures through a dynamic interaction with external signaling molecules (morphogens). It has been posited that morphogens instruct cells to adopt morphogenetic fates through the activation of gene expression in a graded manner, a so called morphogen gradient model. The morphogen Sonic hedgehog (Shh) has been extensively studied in the developmental context of limb patterning and has been thought to be a dose dependent regulator of skeletal number. In this study, a limb progenitor culture system was used to directly and quantitatively study the limb progenitor gene expression response to Shh and Fgf signals. The limb progenitor response to Shh was found to be a simple ON/OFF switch, a response that is far simpler than those predicted by morphogen gradient models. However, additional complexity in the limb progenitor response was uncovered when the Shh response was assessed in conjunction with Fgf dose. These results highlight the importance of studying morphogen mediated response in context with other signals.

Table of Contents

List of Figures and Tables

Chapter One: Introduction: The Limb Progenitor in Context

Chapter Two: An Operational Characterization of Limb Progenitor Identity

Chapter Three: A limb Progenitor Response Model for the Molecular Initiation of Morphogenesis

Acknowledgements

I would like to thank Williams College and its community of administrators and faculty for educating me in the liberal arts, an experience that has allowed me to continue to build the kaleidoscopic set of skills required to think through deep questions in the life sciences.

I would like to thank those who contributed to my graduate training program through the MCB department at Harvard, including Mike Lawrence, Patty Perez, Andrew Murray, Konrad Hochedlinger, Arkhat Abzhanov, and Iain Drummond.

I would like to thank my thesis advisor, Cliff Tabin for allowing me to explore my fascination with cells of multicellular origin through many conceptual lenses and techniques. While in Cliff's lab, I had the opportunity to interact with the members of his diverse lab, and also visit labs across the country and world to learn techniques. I appreciate the openness with which I was able to explore questions, and credit the environment of Cliff's lab for creating the opportunity to do so. As an extension of Cliff's lab, I would like to thank other key members of the Genetics Department at Harvard Medical School including, Terri Broderick, Kim Burman, James deMelo, Hongyun Li, Vonda Shannon, Sophia Zhao and the NRB Biopolymers Facility. Outside of Boston, I would like to thank John Fallon, for ozd mutant eggs and his

stewardship of the field of limb development. I would like to thank Denis Duboule and Guillaume Andrey for sharing their technical expertise and knowledge of Hox genes. I would like to thank Charlotte Colle, whose partnership on the reprogramming project was essential for its progress.

I was fortunate to learn the fundamentals of molecular and developmental biology through the generosity of others. To that end, I would like to thank Jimmy Hu, Jessica Whited, Kimberly Cooper, Patrick Allard, Mark Emerson, and Jenna Galloway. I was fortunate to learn about development alongside an energetic group of graduate student comrades that included Jimmy, Amy Shyer, Aysu Uygur, Matt Schwartz, Jessica Chen, and Tyler Huycke. I also would like to thank my graduate school classmates, especially Jamie Webster and Kyle McElroy.

I would like to thank my core group of kindred spirits – Nikhil Sharma, Bumjin Namkoong, John McGee, Patrick Schorderet, Sandy Klemm and Jonathan Tang. Thank you for being open to the big picture and the big problems.

I would like to thank my family. My wife, Amy, inspires me with her vision into the essentials that often lay obscured in front of us – thank you for keeping us all safe and for making sure I stay tethered to the ground. My brother has been a constant lifeline through difficult moments, and I treasure the weekends I spent with him in New York City away from the lab in Boston. Finally, I would like to thank my parents

who brought our family to the United States for better opportunities and trusted that I would make the best of their sacrifices despite my only having a coarse intuition for the road ahead.

Chapter 1

Introduction: The Limb Progenitor in Context

The emergence of form in organisms, a process known as morphogenesis, has intrigued investigators since the origins of biological inquiry (Aristotle, 1979). Embryonic morphogenesis in vertebrates initiates with a fertilized egg and proceeds through a series of morphogenetic events until birth. This complex process can be representatively caricatured into three steps: 1) the emergence of the overall body plan 2) the emergence of organ structure and 3) organ maturation and integration into a functional being.

Of these three stages, organ structure emergence, or organogenesis, begins at a stage of morphogenesis that is most conserved across vertebrate species, and is often called the phylotypic stage (Fig. 1.1)

At the onset of this period, every proto-organ is composed of one of two embryonic cell types – motile mesenchymal cells and connected epithelial cells. The proto-limb, for example, begins as a hemisphere of mesenchymal cells encased in an epithelial peel of cells. This simple structure gives rise to an entire limb, with the mesenchyme contributing to the skeleton and the epithelium contributing to the skin. The mesenchymal cells can be thought of as the progenitor cells of the limb, or limb progenitors, because they are functionally equivalent to each other but restricted with respect to neighboring embryonic mesenchymal cells (Wolpert, 2007).

Embryonic limb progenitors are situated within a bud-like structure

The proto-limb or limb bud, as it is also called, emerges from the lateral plate mesoderm at around 4 days, 9 days, or 4 weeks after fertilization in the

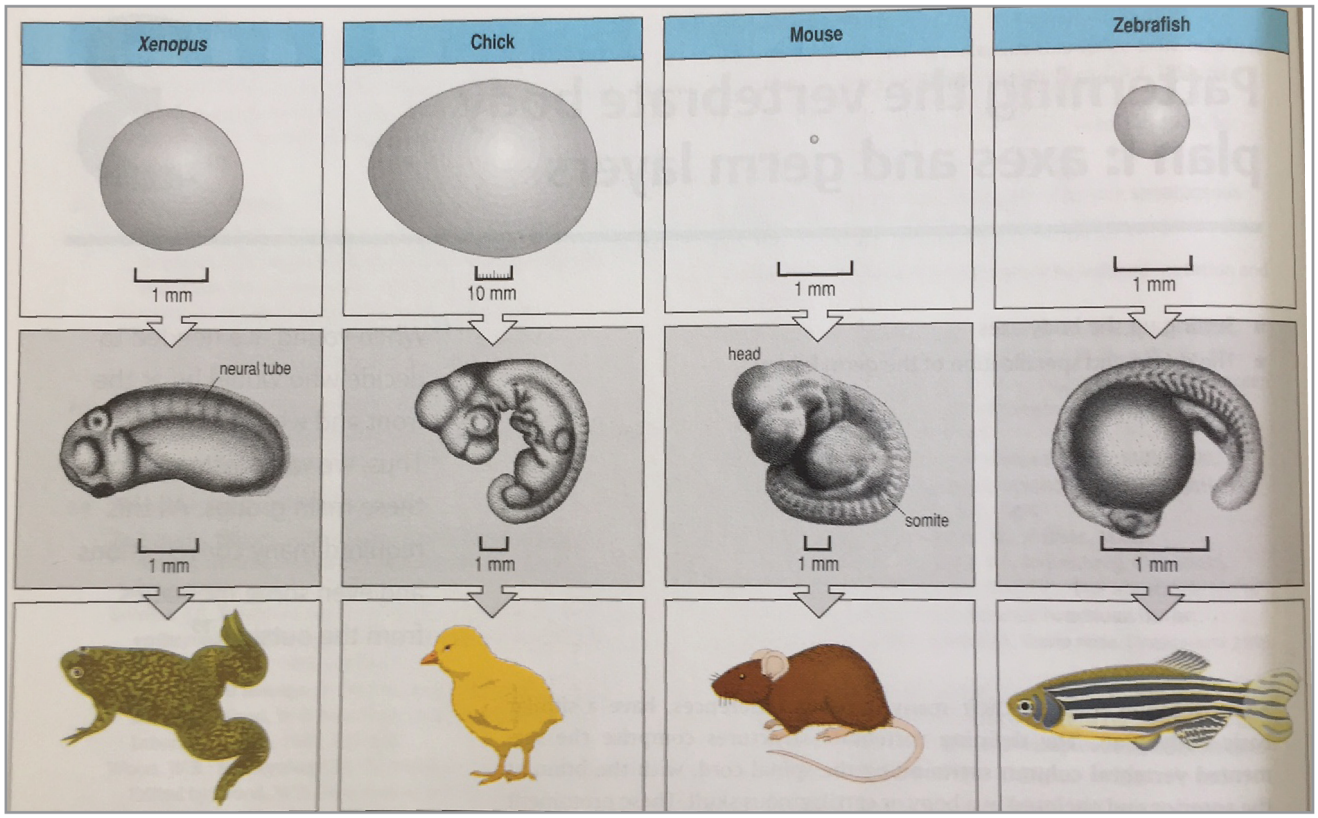


Fig 1.1 The phylotypic stage of vertebrate embryogenesis (middle row)

While the eggs (top row) and final body structure (bottom row) of the frog, chick, mouse, and zebrafish are drastically different in size and form, all these species pass through an embryonic stage (middle row) in which they possess striking morphological similarity. Adapted from Wolpert, 2007.

chicken, mouse, or human embryo, respectively (Hill, 2016). As the bud emerges, the mesenchymal cells within them, or limb progenitors, are under the influence of a number of signaling molecules that are critical for their emergence and maintenance. Wnts and Fgf are secreted from the epithelial peel that encases the limb bud and retinoic acid is secreted from the body axis from which the bud emerges (ten Berge et al., 2008, Cunningham and Duester, 2015). When limb progenitors are harvested from the bud, they must be cultured in the presence of these external signals in order to remain in an undifferentiated state (Cooper et al., 2011). Without these signals, these progenitors, if plated at sufficient density, will spontaneously undergo chondrogenesis (ten Berge et al., 2008).

The orientation of the limb bud with respect to the body naturally provides a spatial axis system. The region of the bud closest to the body is the proximal end of the bud, whereas the region furthest away is the distal end. This axis, called the proximo-distal, is the axis through which the limb elongates in order to create volume for the skeletal elements of the limb. Orthogonal to the proximo-distal axis is the anterior-posterior axis, which extends from the anterior top of the bud to the posterior bottom of the bud (Fig. 1.2).

From an experimental perspective, the limb bud is an ideal model system for the study of organogenesis because its simple initial structure and accessible location make it amenable to surgical manipulation, especially in the chick embryo, which matures inside an egg rather than a maternal host.

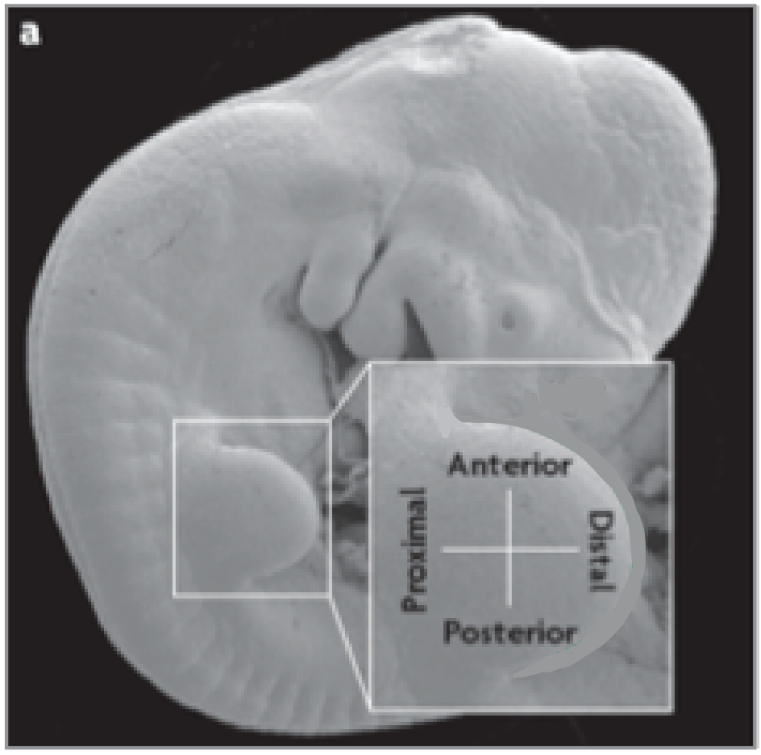


Fig 1.2: Embryonic mouse 10.5 days post fertilization

The proximo-distal and antero-posterior axis of the fore-limb bud.Adapted from Zeller et al., 2009.

Moreover, extreme manipulations at the genetic or tissue level rarely result in lethal consequences for the embryo, allow for one to gauge their full impact on limb morphogenesis.

Functional uniformity of cells in the early limb bud

One could argue that the emerging heterogeneity in form seen in the limb could be pre-specified in progenitor cells of the embryo. However, functional studies with limb progenitors suggest otherwise. This crucial feature of functional homogeneity has been shown through 'recombinant limb' experiments, in which dissociated limb bud cells are re-aggregated, placed within a jacket of limb bud ectoderm, and then grafted onto a host embryo (Fernandez-Teran et al., 1999). Limb bud progenitors are able to produce recombinant limbs with all three segments along the proximodistal axis. These structures have a one bone, one bone, 2 bone pattern. While this structure is complete along the proximo-distal axis, it is less complete along the antero-posterior axis. This is consistent with the fact that Fgf, which is needed for proximo-distal outgrowth, is present from the epithelium, but Shh, which is needed for antero-posterior growth is not present. If a recombinant limb is supplied with a source of Shh at the posterior base of the limb, it is capable of forming a relatively complete limb across both the antero-posterior axis of the limb (Ros et al., 1994). Taken together, these experiments suggest that limb progenitors, when encased in the confines of an epithelial peel, are capable of self-organizing, provided key sources of influence of Fgf and Shh (Fig 1.3).

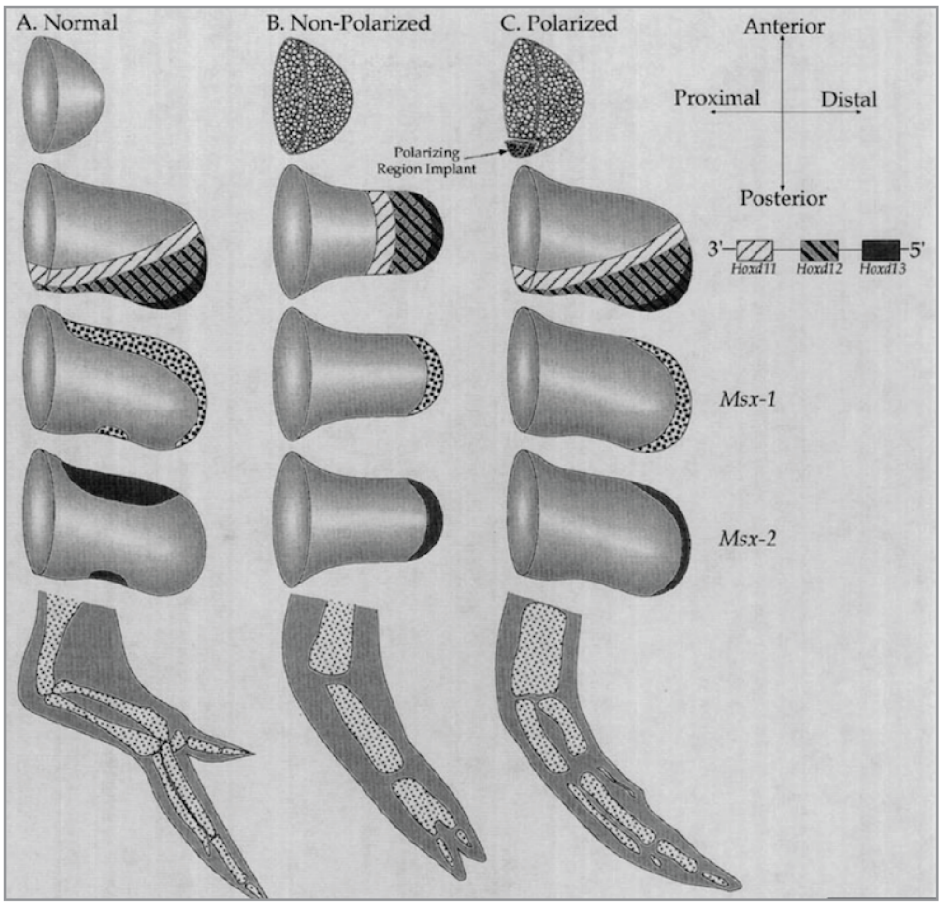


Fig. 1.3: Comparison of in vivo limb development (A) to recombinant limb development (B, C)

Recombinant limbs that do not include Shh polarizing cells lose skeletal bones along the antero-posterior axis (B). These lost structures are rescued by the reintroduction of a Shh polarizing region to the posterior portion of the recombinant limb (C). Adapted from Ros et al., 1994.

In the developmental context, limb bud progenitors exist during a limited time window (day 4 of chick embryogenesis) and in a specific location (the 250 micron region closest to the distal limb bud tip). One broad class of models argues that limb bud progenitors are transitory due to an autonomous clock-based mechanism that leads to their progressive pattern commitment (Summerbell et al., 1973). A competing class of models argues that the fate of limb bud progenitors is dependant on the signaling environment that they encounter. In this view, limb bud progenitors are transient due to changes in signaling environment caused by limb bud outgrowth (Tabin & Wolpert, 2007). In opposition to the first class of models, the second class of models predicts that it may be possible to maintain limb bud progenitors in vitro if the appropriate signaling environment is recapitulated. While limb bud progenitors differentiate into cartilage when cultured in vitro without exogenously added signals, they can be maintained for up to 36 hours when cultured in the presence of Wnt3a, Fgf8, and RA. After being cultured with these signals, limb bud progenitors can still give rise to a patterned limb with three segments across the PD axis when grafted as a recombinant limb (Cooper et al., 2011). In this in vitro setting, Wnt3a and Fgf8, which are endogenously secreted from the limb bud ectoderm, prevent differentiation of limb bud progenitors (ten Berge et al., 2008). RA, which is released from the flank, is believed to prevent the onset of pattern commitment (Cooper & Hu et al., 2011).

Limb progenitor cells form a complex skeletal structure under the influence of critical signaling molecules

From a high-level morphogenetic view, the limb bud elongates along the proximo-distal axis and then assumes the emerging shape of the final limb structure. Critically, as the limb emerges it assumes a pattern of a single bone (upper-arm or stylopod), two bones (forearm or zeugopod), and many bones (hand or autopod). This pattern is conserved across vertebrates, suggesting deep conservation in the fundamental constructions rules that guide limb progenitors aggregation into three-dimensional bone structures.

Although the distal tip of the limb bud continues to maintain a population of undifferentiated mesenchymal cells from stages 18-25 (Tabin and Wolpert, 2007), these cells become increasingly committed with respect to patterning potential as the limb bud elongates. Limb bud progenitors, which are the undifferentiated cells that exist in the limb bud at the beginning of this time window (stage 18-20), are not restricted in terms of their patterning potential. This crucial feature has been shown through 'recombinant limb' experiments, in which dissociated limb bud cells are re-aggregated, placed within a jacket of limb bud ectoderm, and then grafted onto a host embryo (Fernandez-Teran et al., 1999). Limb bud progenitors are able to produce recombinant limbs with all three segments along the proximodistal axis whereas undifferentiated distal tip cells from later limb buds (> stage 21) produce recombinant limbs with 2 or fewer segments (Dudley et al., 2002). Thus, while genes involved in patterning, most notably Hox genes, are

expressed in complex spatial patterns concurrently with limb bud emergence (Nelson et al., 1996), limb bud progenitors are able to reestablish their patterning gene expression based on their new positions within the recombinant limb (Ros et al., 1994).

Growth along the proximo-distal axis of the limb is dependent on the presence of Fgf signals that are secreted from the distal tip of the epithelial peel. If Fgfs are inactivated, limb bud outgrowth arrests and only a truncated limb is formed (Martin, 1998). Conversely, growth along the antero-posterior axis is dependent on the secreted signal Sonic hedgehog. If Shh is inactivated in the limb, the forarm loses a bone and the hand loses almost all of its digits (Ros et al., 2003). In addition to loss of function phenotypes, both signals have potent gain-of-function effects on limb morphogenesis. If Fgf is ectopically provided to a region of the lateral plate mesoderm that is not normally fated to produce a limb, an entire ectopic limb will be fated to grow (Martin, 1998). If Shh is ectopically provided to the anterior region of the limb bud, a second pair of hands will emerge that mirrors the hand that forms from the posterior portion of the limb bud (Riddle et al., 1993).

A number of key patterning genes are not only targets of Shh and Fgf, but require both in concert for their expression. 5' HoxD genes (Hoxd11, Hoxd12, Hoxd13) and secondary signals such as Bmp2, can be ectopically induced within the limb by application of Shh and Fgf together, but not by either one alone (Laufer et al., 1994). The necessity of endogenous Fgf activity for HoxD gene expression cannot be easily tested, as in the absence of both

Fgf4 and Fgf8 activity limb development itself fails (Martin, 1998). However, a requirement for Shh in regulating limb HoxD gene expression has been verified genetically (Ros et al., 2003). HoxD genes are normally expressed in two distinct phases in the developing limb bud. Early (during so-called Phase 1), Hoxd11, Hoxd12, and Hoxd13 are expressed in a set of posteriorly-biased nested domains, with Hoxd13 being the most restricted. However, as the hand plate forms and the distal limb segment is patterned, there is a transition in Hox gene regulation (Phase 2) wherein Hoxd11, Hoxd12 and Hoxd13 expression spread across the AP axis of the distal limb bud, with Hoxd13 becoming the broadest in expression (Nelson et al., 1996). The first phase of Hox gene expression is largely, but not completely, dependent on Shh activity, as Phase 1 expression occurs, but at a significantly reduced level, in the absence of Shh activity. Additionally, in absence of Shh, Phase 2 of Hoxd expression is entirely lost (Ros et al., 2003).

Motivation for Chapter 2: Cellular reprogramming to a limb progenitor state

Limb progenitors give rise to all of the skeleton, tendons, ligaments, muscle connective tissue and dermis in the limb. These progenitor cells are specified at specific points along the lateral plate mesoderm and first become visible in the embryo as a mound of proliferating mesenchymal cells enveloped by an ectodermal jacket (mouse e9.5 - e10.5). As they emerge, limb progenitors are committed to a limb fate and a developmental potential distinct from

neighboring neck and interlimb mesenchyme. Tbx5 and Tbx4 were the first transcription factors found to be necessary for forelimb and hindlimb bud emergence, respectively. However, it is unlikely that they are the sole master regulators of limb progenitor identity since Tbx5 is also essential for heart development . Instead, 'limbness' is potentially encoded by a set of transcription factors. This set of transcription factors could be revealed through cellular reprogramming experiments.

The transcriptional reprogramming of differentiated cells into limb could be fruitful on a number of levels. First, it applies reprogramming to the problem of limb development, a context where the approach has not been previously exploited. Second, this study adds a new layer to the definition of a progenitor cell. In previous studies, progenitor cells have been defined in terms of the cell types that can be derived from them. Limb progenitor cells can indeed give rise to an array of tissue types (skeletal tissues, connective tissues, etc.). However, early limb progenitors have a second property as well: they can form all the structures of the limb. As the limb bud grows out, progenitors are progressively restricted in the proximodistal pattern of the structures they can produce (although they maintain the ability to differentiate into the same range of cell types).

Motivation for Chapter 3: Gauging the limb progenitor gene expression response to Shh and Fgf

Sonic hedgehog (Shh) has been extensively studied in the developmental context of limb patterning. As previously mentioned, it is expressed in the distal posterior mesenchyme of the limb bud and was established as the key mediator of AP patterning through experiments that showed that it could ectopically induce the AP axis. Normal chick limbs possess three digits – 2-3-4 (anterior to posterior). However, if Shh expressing cells are grafted into the distal anterior mesenchyme, the limb develops 6 digits with a 4-3-2-2-3-4 pattern (Riddle et al., 1993).

A number of experiments have illustrated that both Shh concentration and time of exposure impact digit patterning. With respect to concentration, chick limb bead implants of high concentrations of Shh induce the ectopic formation of digit 4 (most posterior digit in chick) whereas bead implants of low concentrations of Shh induce the ectopic formation of digit 2 (most anterior digit in chick). With respect to time of exposure, at a fixed concentration, a longer period of Shh bead exposure ectopically induces a more posterior digit as compared to the digit induced by a shorter period of Shh bead exposure (Yang et al., 1997). Thus, high concentration of and long exposure to Shh appear to be linked to more posterior digits. These results were further supported by fate mapping experiments of descendents of Shh producing cells in mouse limb bud. These experiments revealed that Shh producing descendents contribute to all of the two most posterior digits (digits 4 and 5 of a 1-2-3-4-5 digit pattern) and a portion of the third most posterior digit (digit

3). Thus, digits 3, 4, and 5 are exposed to high concentrations of Shh but for differing amounts of time (Harfe et al., 2004).

Despite our understanding of the impact of concentration and time of Shh in digit patterning, it remains unclear as to how limb bud progenitors integrate differences in concentration and temporal exposure to Shh at the cellular and molecular level. Two experiments suggest that some form of integration of concentration and length of exposure occurs at the molecular level. First, high concentrations of Shh for short periods of time can elicit similar digit duplications as low concentrations of Shh for long periods of time. Second, limb bud progenitors appear to possess a 'memory' for Shh exposure. If the distal anterior portion of a chick limb bud is exposed to low concentrations of Shh for a very short period of time (10 hours), no ectopic digits arise. However, if these 'primed' limbs are then exposed to another dose of low Shh for an additional period of time (16 hours), they can form digits with a more posterior identity than 'non-primed' limbs (limbs that receive a buffer bead for 10 hours followed by a low concentration Shh bead for 16 hours) (Harfe et al., 2004).

While these two experiments provide evidence for the integration of concentration and time of Shh exposure, they still rely on morphological rather than molecular readouts. This has been in part due to the technical limitations of bead implantations studies, which do not allow for precise delivery of concentration and temporal doses of Shh in a spatially and developmentally

controlled context. Recently, however, a limb bud progenitor culture system has been developed which allows for such precise control (Cooper et al., 2011). The following proposal will take advantage of this novel limb bud progenitor culture system to study the effects of different concentrations and durations of Shh exposure on the limb progenitor gene expression response. It will also study the influence that Fgf has on this response, given the knowledge that Shh and Fgf are key co-regulators of 5' HoxD and Bmp2 gene expression. Assessing the influence of Shh and Fgf on the limb progenitor gene expression response will provide insights into how differences in extrinsic signal dose can elicit differential gene expression programs that set the stage for digit patterning.

References

Aristotle. Generation of Animals, translated by A. L. Peck, Cambridge, MA: Harvard University Press (1979).

Cooper K.L., Hu J.K., ten Berge D., Fernandez-Teran M., Ros M.A., Tabin C.J. Initiation of proximal-distal patterning in the vertebrate limb by signals and growth. Science. 332(6033):1083-6 (2011).

Cunningham T.J. and Duester G.Mechanisms of retinoic acid signalling and its roles in organ and limb development. Nature Reviews Molecular Cell Biology. 16, 110–123 (2015).

Dudley A.T., Ros M.A., Tabin C.J. A re-examination of proximodistal patterning during vertebrate limb development. Nature. 418(6897):539-44 (2002).

Fernandez-Teran M., Piedra M.E., Ros M.A., Fallon J.F. The recombinant limb as a model for the study of limb patterning, and its application to muscle development. Cell Tissue Res. 296(1):121-9 (1999).

Harfe B.D., Scherz P.J., Nissim S., Tian H., McMahon A.P., Tabin C.J.

Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. Cell. 118(4):517-28 (2004).

Hill, M.A. Embryology Lecture - Limb Development. https://embryology.med.unsw.edu.au/embryology/index.php/Lecture_-_Limb_Development (2016).

Laufer E., Nelson C.E., Johnson R.L., Morgan B.A., Tabin C. Sonic hedgehog and Fgf-4 act through a signaling cascade and feedback loop to integrate growth and patterning of the developing limb bud. Cell. 79(6):993-1003 (1994).

Martin G.R. The roles of FGFs in the early development of vertebrate limbs. Genes Dev. 12(11):1571-86 (1998).

Nelson C.E., Morgan B.A., Burke A.C., Laufer E., DiMambro E., Murtaugh L.C., Gonzales E., Tessarollo L., Parada L.F., Tabin C. Analysis of Hox gene expression in the chick limb bud. Development. 122(5):1449-66 (1996).

Riddle R.D., Johnson R.L., Laufer E., Tabin C. Sonic hedgehog mediates the polarizing activity of the ZPA. Cell. 75(7):1401-16 (1993).

Ros M.A., Dahn R.D., Fernandez-Teran M., Rashka K., Caruccio N.C., Hasso S.M., Bitgood J.J., Lancman J.J., Fallon J.F. The chick oligozeugodactyly (ozd) mutant lacks sonic hedgehog function in the limb. Development. 130(3):527-37 (2003).

Ros M.A., Lyons G.E., Mackem S., Fallon J.F. Recombinant limbs as a model to study homeobox gene regulation during limb development. Dev Biol. 166(1):59-72 (1994).

Summerbell D., Lewis J.H., Wolpert L. Positional information in chick limb morphogenesis. Nature. 244(5417):492-6 (1973).

Tabin C., Wolpert L. Rethinking the proximodistal axis of the vertebrate limb in the molecular era. Genes Dev. 21(12):1433-42 (2007).

ten Berge D., Brugmann S.A., Helms J.A., Nusse R. Wnt and FGF signals interact to coordinate growth with cell fate specification during limb development. Development. 135(19):3247-57 (2008).

Yang Y., Drossopoulou G., Chuang P.T., Duprez D., Marti E., Bumcrot D., Vargesson N., Clarke J., Niswander L., McMahon A., Tickle C. Relationship between dose, distance and time in Sonic Hedgehog-mediated regulation of anteroposterior polarity in the chick limb. Development. 124(21):4393-404 (1997).

Zeller R., López-Ríos J., Zuniga A. Vertebrate limb bud development: moving towards integrative analysis of organogenesis Nature Reviews Genetics. Vol.10(12), 845 (2009).

Chapter 2

An Operational Characterization of

Limb Progenitor Identity

An Operational Characterization of Limb Progenitor Identity

Alan Rodrigues (1,2), Charlotte Colle (1),
Tia DiTommaso (3), Johanna Kowalko (1, 4), James DeMelo (1), Danos Christodoulou (1), Josh Gorham (1), Patrick Tschopp (1),
Christine E. Seidman (1), J.G. Seidman (1),
Jessica L. Whited (3) and Clifford J. Tabin (1)

(1) Department of Genetics
Harvard Medical School
77 Avenue Louis Pasteur
Boston, MA 02115

(2) Department of Molecular and Cellular Biology
Harvard University
52 Oxford Street
Cambridge, MA 02138

(3) Brigham Regenerative Medicine Center and the Department of Orthopedic Surgery
Brigham & Women's Hospital
Harvard Medical School
75 Francis Street
Boston, MA 02115

(4) Current affiliation:
Department of Genetics, Development and Cell Biology
Iowa State University
Ames, Iowa 50011

Abstract

The limb bud consists of mesenchyme derived from the lateral plate mesoderm (LPM). The LPM also gives rise to mesenchyme that will form the mesodermal components of the trunk, flank and neck. However, these other LPM-derived tissues cannot produce the variety of cell types found in the limb bud nor can they be directed to form a patterned appendage-like structure, even when placed in the context of the organizing signals normally found in the limb bud. We have conducted a screen to identify a set of transcription factors capable of imparting limb progenitor-like properties on non-limb embryonic and postembryonic fibroblasts. Reprogrammed limb progenitors express similar genes and can differentiate into similar cell types as legitimate limb bud cells.

The cells that will ultimately give rise to the majority of tissues within the mature limb originate from the somatopleural layer of the lateral plate mesoderm, a continuous epithelium lining the embryonic coelom. Mesenchymal limb bud progenitors are generated through localized epithelial-mesenchymal transition (EMT). Under the influence of transcription factors such at Tbx5 and signals such as Fgf10,a large pool of mesenchymal limb bud progenitors are produced specifically at the level of the future limbs (Gros and Tabin, 2014). However, EMT also occurs elsewhere along the coelomic epithelium, for example, in the future neck and flank regions, to generate progenitors of the dermis of the flank. These tissues are fundamentally different from the limb bud progenitor cells. For example, when limb bud mesenchyme is dissociated, pelleted and then placed into an empty jacket of limb mesenchyme, it will give rise to a patterned limb-like structure when grafted back onto a host embryo (Ros et al., 1994). Such “recombinant limbs” contain all the tissue types normally derived from the limb bud progenitors including bones, cartilage, tendon, ligament, dermis, etc. However, when an equivalent experiment is carried out using flank mesenchyme, a limb bud does not grow out (Fig. 2.1). Moreover, the grafted tissue that remains does not differentiate into the cell types that typify the limb. Thus limb progenitors clearly have distinct properties, differing both in

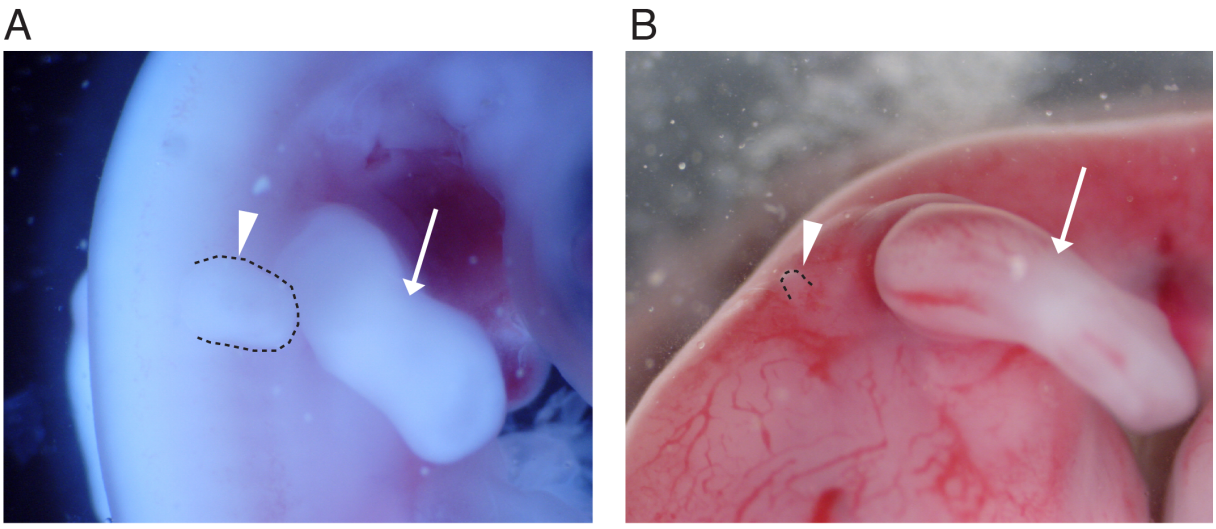


Figure 2.1: Recombinant limb grafts with limb and flank mesenchyme

(A) An implanted recombinant limb (arrowhead) made with chick limb mesenchyme grows in mass at a rate similar to the endogenous limb. Image taken 2 days after graft of recombinant limb. Endogenous limb marked with arrow.

(B) An implanted recombinant limb (arrowhead) made with chick flank mesenchyme fails to grow after implantation. Image taken 2 days after graft of recombinant limb. Endogenous limb marked with arrow.

terms of their response to patterning signals and in their developmental potential, from non-limb lateral plate-derived mesenchyme.

To understand what it really means to be a limb bud progenitor, we decided to see if we could identify a set of transcription factors, expressed ubiquitously in the early limb field, that might be responsible for establishing and maintaining the unique transcriptional characteristics and differentiation potential of limb bud progenitor cells. To that end, we took a reprogramming approach (Takahashi and Yamanaka, 2006), reasoning that the full set of transcription factors giving early limb mesenchyme its unique properties might be sufficient to convert ("reprogram") non-limb mesenchyme into limb mesenchyme.

As a first step, we wanted to generate a list of transcription factors expressed specifically in the early limb fields. We already had at our disposal a transcriptome-level analysis of the developing mouse hindlimb bud (Tschopp et al., 2014). We decided to compare this to the transciptome of the chick limb buds, reasoning that the key genes in establishing limb-specific mesenchyme would be evolutionarily conserved. Moreover the chick offered greater ease in dissection to obtain other, related tissues for transcriptional comparisons. To that end, we harvested tissue from chick embryos at the times when the mesenchyme first arises by EMT. We performed RNA-seq analysis on the forelimb bud mesenchyme (at HH Stage 17-19), hindlimb bud mesenchyme (HH Stage 17-19), presumptive neck mesenchyme (HH

Figure 2.2: Transcriptomic comparison of the early limb bud to neighboring mesenchymal tissue

(A) Schematic of chick embryos: stage 15 (left), stage 19/20 (right). Regions of embryo that were use for transcriptomic analysis are labeled.

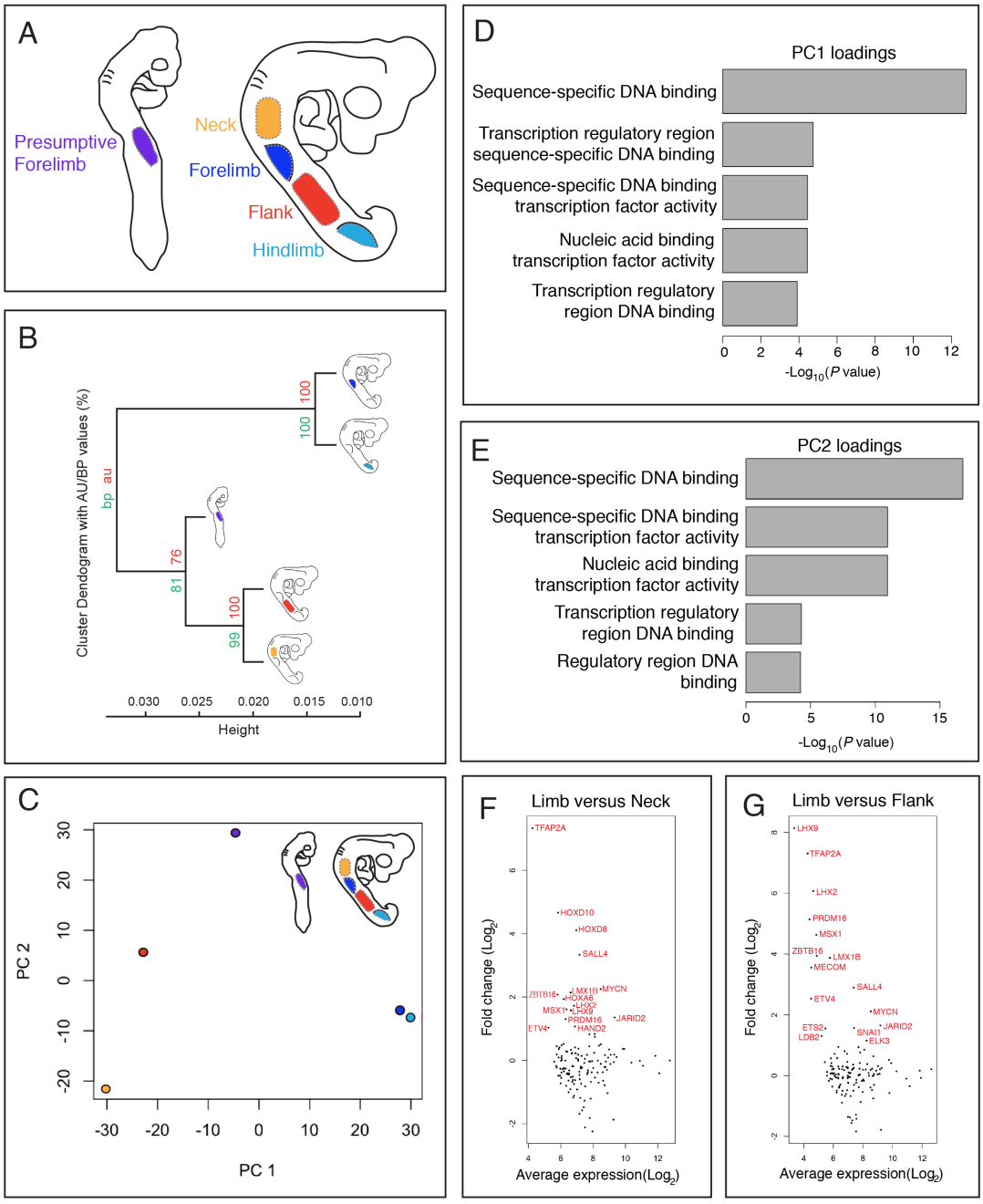
(B) Hierarchical clustering of transcriptomic data from chick forelimb bud, hindlimb bud, presumptive forelimb, neck mesenchyme precursors, and flank mesenchyme precursors. Forelimb and hindlimb bud tissue form a cluster independent of the other three tissues.

(C) Principal component analysis of five transcriptomic data sets. Forelimb and hindlimb bud expression values cluster closely together. Separation of other three data sets occurs across principal (PC1) and principal component 2 (PC2).

(D,E) Top five most statistically significant enriched gene ontology classifications for top 100 genes associated with PC1 and PC2

(F,G) Differential expression analysis (MA-plot) of core gene set. Limb expression (average of forelimb and hindlimb) over neck expression (F). Limb expression over flank expression (G). Labeled points indicate genes with greater than two-fold overexpression in limb tissue.

Figure 2.2 (Continued)



Stage 19-20), and presumptive flank mesenchyme (HH Stage 19-20). Additionally, we profiled the epithelial lateral plate mesoderm prior to forelimb bud emergence (HH 15). Encouragingly, since we assume the core set of transcription factors will be common to forelimbs and hindlimbs, hierarchical clustering of the transcriptomes of all five profiled tissues resulted in two broad clades, one corresponding to forelimb and hindlimb bud tissues and another to the remaining three tissues sampled (Fig. 2.2a).

While we carried out this analysis to identify transcription factors differentiating these tissues, in principle other classes of genes could differ between them as well. To determine the classes of genes that indeed differentiate these five tissues from each other, we carried out a principal component analysis (PCA). The first and second principal components (PC) accounted for 48% and 28% of the variance in the five data sets. When plotted in the principle component space, the forelimb and hindlimb bud tissues cluster together tightly. PC 1 separates the remaining three tissues from the limb tissues whereas PC 2 separates presumptive forelimb and neck mesenchyme from the limb tissues (Fig. 2.2b). To determine the key drivers of this separation in PC space, the top 100 genes contributing to each principal component were used in a gene set enrichment analysis. For both PC1 and PC2, the top five most significant classes of gene function were all related to transcription or transcriptional regulation (Fig. 2.2c,d). This result suggests that the key

drivers of difference between limb and non-limb lateral plate mesenchyme are, not unexpectedly, transcription factors.

We then intersected our existing mouse hindlimb bud transcriptional data set with our chick data to generate an evolutionarily conserved set of candidate genes we could use in a reprogramming assay. Of the 1806 mouse transcription factors, chromatin remodeling factors, and transcription co-factors in the mouse genome, 303 are expressed at appreciable levels in the mouse hindlimb bud. Of these 303 transcriptionally related genes, 142 are co-expressed in both the chick forelimb and hindlimb bud (Fig. 2.2e). Of this core set of 142 transcription factors, co-factors and chromatin remodelers that were expressed in the limb, we were particularly interested in those that were differentially expressed relative to the neck and/or flank mesenchyme. Only 15 of the 142 factors were more than two-fold over-expressed in the limb as compared to the neck and 16 of the 142 factors were more than two-fold overexpressed when compared to the flank. In these cohorts, 11 genes were shared across both lists (Fig. 2.2 f,g). We selected all genes in these lists (Fig. 2.3a), except for Lhx9 and Hoxa6, which were deemed potentially redundant to Lhx2 and other Hox genes, respectively. We also included several genes that were not differentially expressed, but which were expressed in both the chick and mouse limb bud mesenchyme and had been previously implicated functionally as being important for limb

Figure 2.3: A cellular reprogramming screen to identify critical limb progenitor transcription factors

(A) Candidate transcription factors used in the screen and locations of their *in vivo* expression.

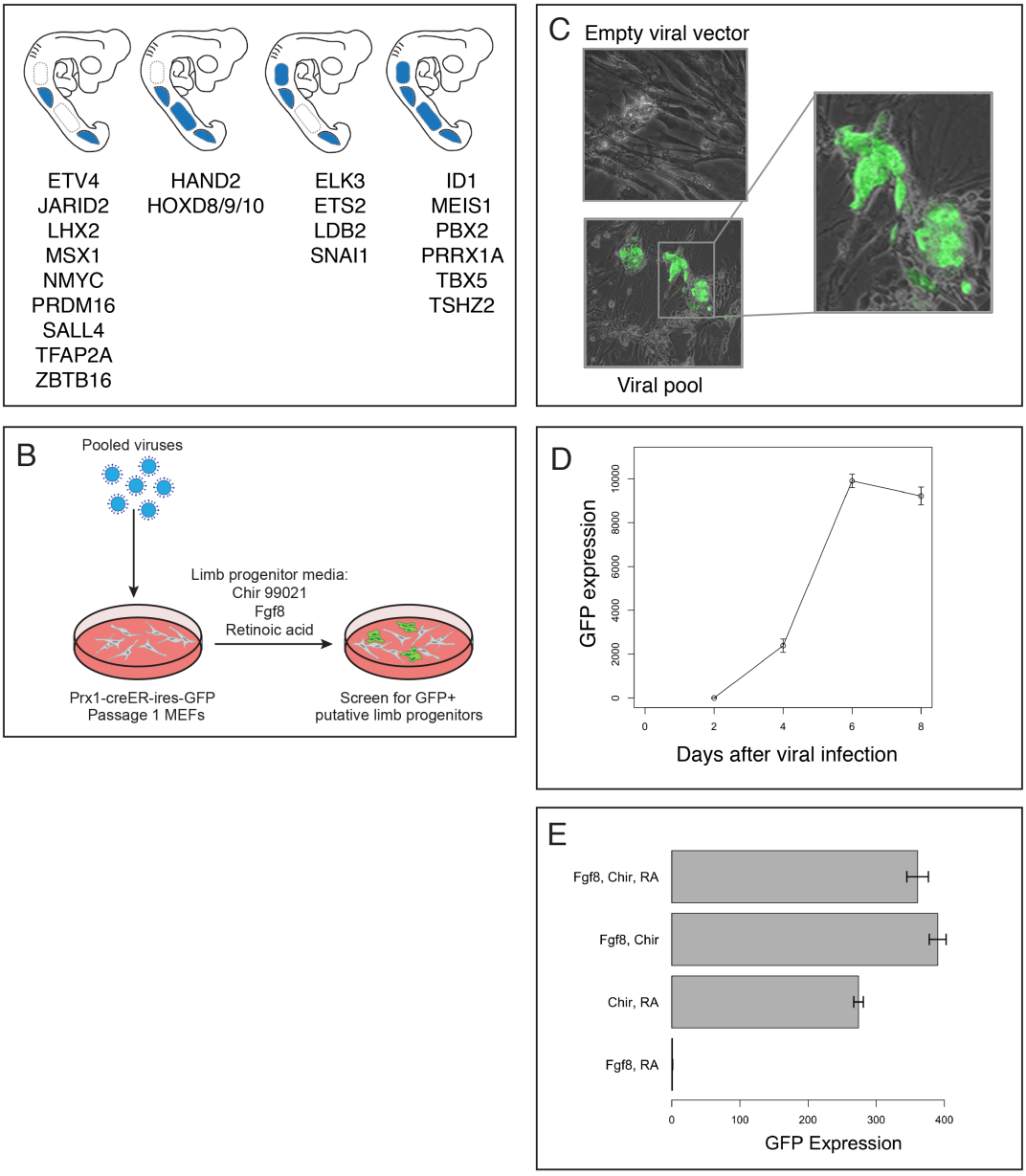
(B) Outline of reprogramming assay. Prx-creER-ires-eGFP MEFs are infected with a viral pool of all candidate master transcription factors. Two days post infection, MEFs are supplemented with limb progenitor media containing Chir 99021, Fgf8, and retinoic acid, after which they are screened for GFP.

(C) 6 days after viral infection, GFP positive cells emerge from initially GFP negative Prx-creER-ires-eGFP MEFs. MEFs express GFP and also exhibit a flat to rounded change in cell morphology.

(D) Time course of reporter activation, measured by quantitative PCR. Robust activation of reporter is observed six days after viral infection.

(E) Repetition of reprogramming assay with different media formulations. Efficiency of reporter activation is measured by quantitative PCR six days after viral infection. Removal of Fgf8 results in a moderate loss of reporter induction. Removal of Chir 99021 results in a total loss of reporter induction.

FIgure 2.3 (Continued)



bud outgrowth. Ultimately, our list of candidate reprogramming transcriptions factors included 21 genes (Fig. 2.3a).

The 21 candidate genes were inserted into retroviral vectors that were grown and pooled. We chose a viral vector (pMXs) capable of infecting murine cells, as the mouse system was the most amenable to setting up an assay for reprogramming. To that end, we took advantage of the limb-specific GFP activity seen in Prx1-creER-ires-GFP transgenic mice. Activity of this transgene, thus, served as a proxy for limb-specific transcriptional regulation in an initial assay. The transgenic line was used to generate mouse embryonic fibroblasts and also post-natal tail tip fibroblasts. While the Prx1 promoter strongly drives expression in limb buds, it is also expressed in some other regions of the embryo, such as the head mesoderm. Therefore, we first cell-sorted for GFP-negative cells before plating. Primary cells were expanded through one passage and then infected with pooled retroviral vectors transducing our transcription factor library and the cells were plated in media already known to support limb progenitors (a cocktail that includes Fgf8, retinoic acid and the Wnt agonist Chir99021) (Cooper et al., 2011) (Fig. 2.3b).

Six days after infection, the emergence of GFP positive cells became apparent. Importantly, a large fraction of these GFP positive cells were also associated with a change in morphology from flat MEFs to a more rounded, and less contact-inhibited, morphology reminiscent of freshly harvested limb progenitors plated at high density (Fig. 2.3c). GFP

transcript levels were measured by qPCR 2, 4, 6 and 8 days after reprogramming. Maximum levels of GFP expression were reached 6 days after infection (Fig. 2.3d). The reprogramming process was repeated, but with the removal extrinsic signals from the limb progenitor-supporting media. While the removal of Fgf8 resulted in a modest decrease in the levels of transgene expression, the removal of the Wnt agonist, Chir99021, lead to complete loss of expression – suggesting the critical nature of Wnt pathway stimulation for the conversion of fibroblasts into putative limb progenitors.

To identify which transcription factors in our pool were responsible for the induction of GFP-positive cells in our cultures, we repeated the experiment in a series of separate cultures, in each case removing one of the transcription factors from the mix and assaying GFP expression quantitatively by qPCR. Removal of most factors had no effect on the efficiency of reprogramming (Fig. 2.4a). Surprisingly, the efficiency actually went up if either Hoxd genes or Tbx genes were removed from the pool, indicating that although these genes are critical for early limb patterning, their expression actually interferes with the reprogramming of non-limb mesenchyme to a putative limb-like state (Fig. 2.4a). Another gene that was not critical for inducing green GFP-positive cells in our assay was Nmyc. However, in the absence of this gene the GFP-positive cells did not take on the rounded morphology reminiscent of legitimate limb progenitors. Nmyc was therefore deemed important for this process.

Figure 2.4: Identification of a minimal set of transcription factors essential for the induction of putative limb progenitor-like cells

(A) Quantitative drop out experiments assessed by qPCR show that, in addition to Lhx2 and Sall4, Prdm16, Plzf, and Msx1 also contribute to prx reporter expression.

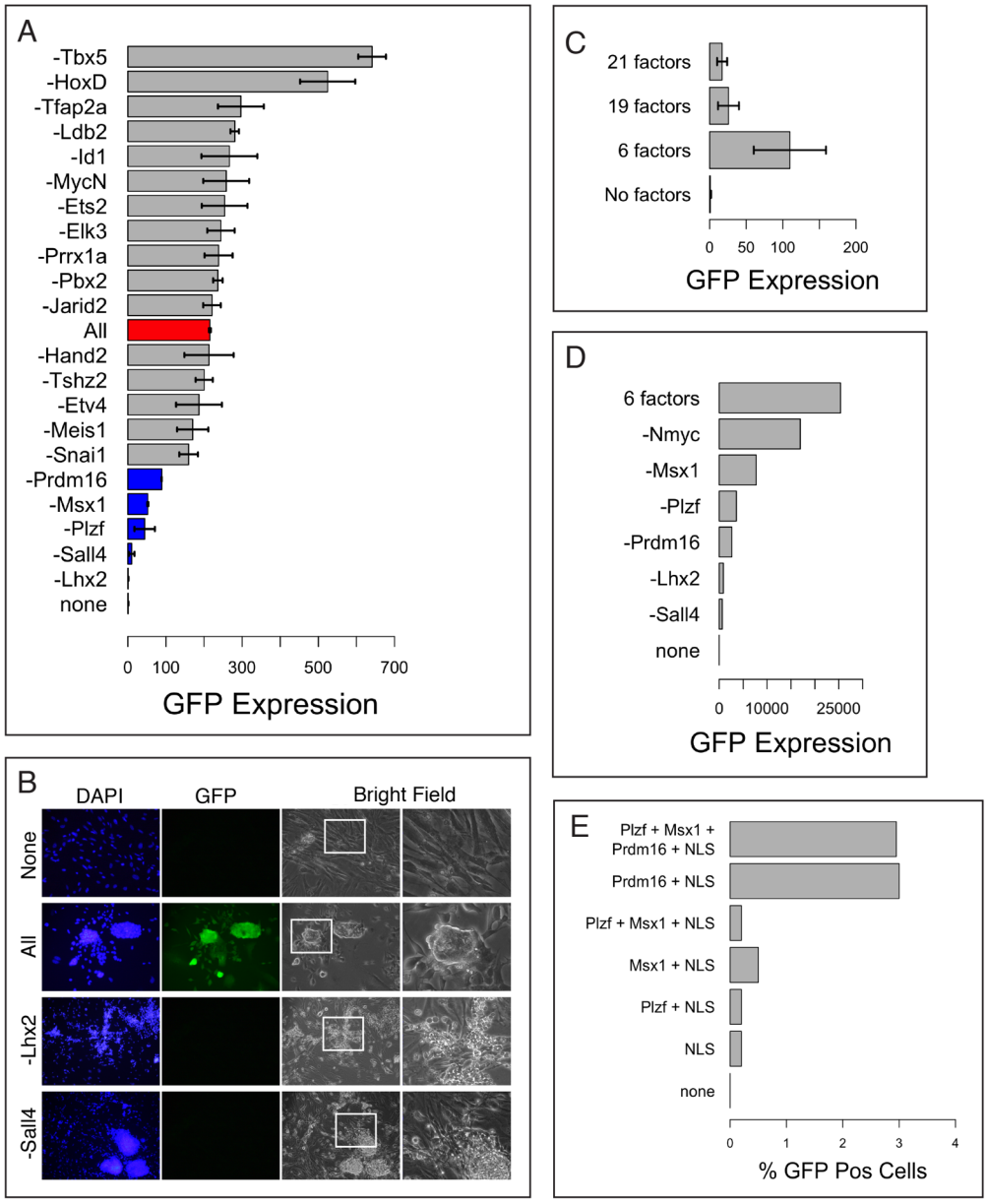
(B) Qualitative drop out experiments reveal that Lhx2 and Sall4 are essential for induction of GFP reporter positive cells.

(C) Reprogramming was conducted with 3 viral pools: 21 factors – All candidate transcription factors; 19 factors – All candidates excluding Tbx5 and HoxD genes; 6 factors – Lhx2, Sall4, Prdm16, Plzf, Msx1, and Nmyc. The 6 factor viral cocktail induces the highest level of reporter expression by qPCR.

(D) Drop out experiments with 6 factor cocktail. Removal of any of the 6 factors results in a loss of GFP transgene expression.

(E) Add one in experiments: Nmyc, Lhx2, and Sall4 (NLS) as a 3 factor cocktail. Plzf, Msx1, or Prdm16 are added all together or individually to the 3 factor cocktail. The 3 factor cocktail (Nmyc, Lhx2, Sall4) with Prdm16 is sufficient to produce the same percentage of GFP reporter positive MEFs as all six factors. GFP % assessed through flow cytometry.

Figure 2.4 (Continued)



Five other genes had a significant effect on the reprogramming. Lhx2 and Sall4 were absolutely required (Fig. 2.4a,b) while Plzf, Msx1 and Prdm16 were needed for full efficiency (Fig. 2.4a).

To determine whether a smaller set of factors was sufficient to generate GFP-positive cells, we repeated the reprogramming assay using all 21 factors, 19 factors (removing Hoxd genes and Tbx5), and 6 factors (Sall4, Lhx2, Prdm16, Plzf, Msx1, and Nmyc). Of the three cohorts of viral cocktails, the 6 factor cocktail induced the highest level of GFP-expression, indicating that a smaller group of genes were not only sufficient, but also more efficient at reprogramming than the original 21 factor cocktail (Fig. 2.4c). We then assessed the importance of these remaining six factors in two complementary ways. First, we repeated the one factor removal experiment starting with just the six factors. Removal of any one of the six factors resulted in a drop in reprogramming efficiency; however, removal of either Sall4, Lhx2, Prdm16 or Plzf resulted in the greatest loss (Fig. 2.4d). Second, we tested to see if Sall4 and Lhx2, the two most critical factors for GFP induction, along with Nmyc could induce GFP expression. The three factors ("NLS") were unable to produce GFP positive cells, as measured by flow cytometry. Combining NLS with Plzf and/or Msx1 was also not able to produce GFP positive cells. However, combining NLS with Prdm16 was able to produce GFP positive cells at the same rate as all six factors together (Fig. 2.4e). Thus, we concluded that Sall4, Lhx2 and Prdm16 are absolutely critical

for the induction of GFP positive cells and that Nmyc was critical for the proper changes in cellular morphology. We also noticed that removal of Plzf, while not resulting in the loss of GFP-positive cells, did reduce the size and number of colonies formed (Fig. 2.5). Taken together, Sall4, Lhx2, Prdm16 and Nmyc form a critical core of factors necessary to generate putative limb progenitors and Plzf may enhance the production of fully functional progenitors. Strikingly, we obtained similar reprogramming efficiencies and changes in morphology whether using mouse embryo fibroblasts or post-natal tail-tip fibroblasts. In both cases, infection of the core four factors resulted in approximately 15% GFP-positive cells in culture at day 6 or 7 (Fig. 2.6).

As noted above, however, the limb buds are not the only tissues where the Prx1 promoter is active. Thus, although induced by a set of transcription factors normally co-expressed in the limb mesenchyme, the GFP-positive cells need not have been reprogrammed to a limb-like state. This was especially a concern since the four transcription factors in our final pool, Nmyc, Lhx2, Sall4, and Prdm16, are also expressed in other regions of the embryo; although, importantly, the combinatorial expression of all four is indeed unique to the limb. We, therefore, tested whether GFP-positive cells, infected with our pool of viruses, also exhibited up-regulation of other known limb-specific genes. A panel of limb markers was measured in RNA lysates from infected MEFs versus

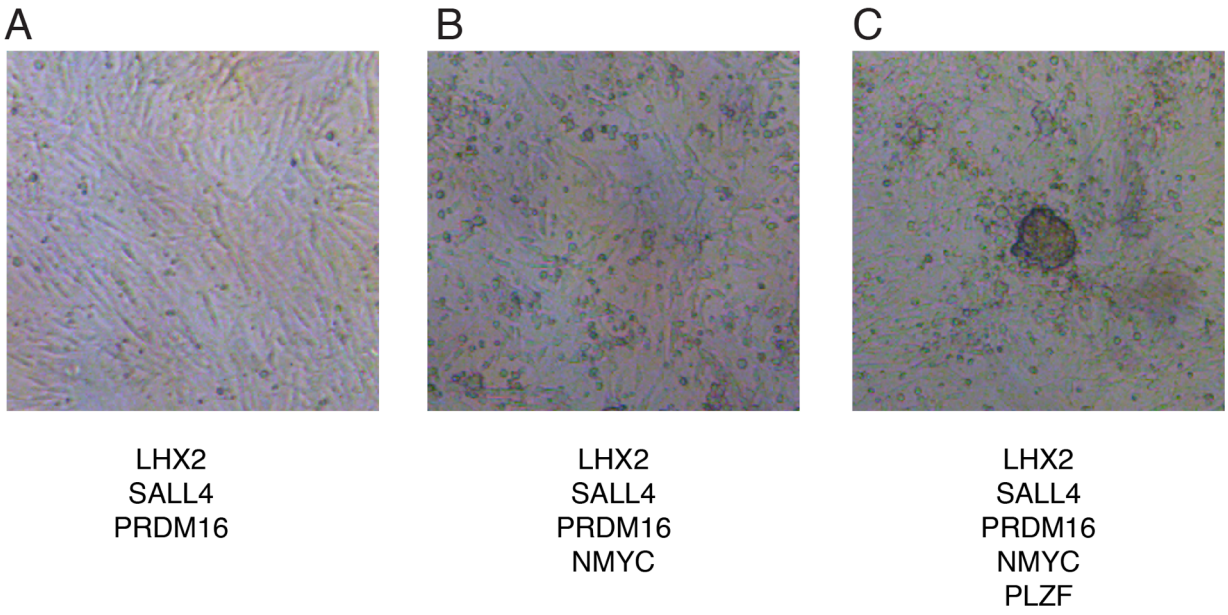


Figure 2.5: Morphology changes induced after viral infection

(A) 5 days after viral infection with *lhx2*, *sall4*, and *prdm16*. No morphology changes from normal fibroblast morphology.

(B) 5 days after viral infection with *lhx2*, *sall4*, *prdm16*, and *Nmyc*. Morphology changes apparent. Small aggregates of contact inhibited cells present.

(C) 5 days after viral infection with *lhx2*, *sall4*, *prdm16*, *Nmyc*, and *plzf*. Addition of *plzf* to four-factor cocktail creates much larger aggregates of contact-inhibited cells.

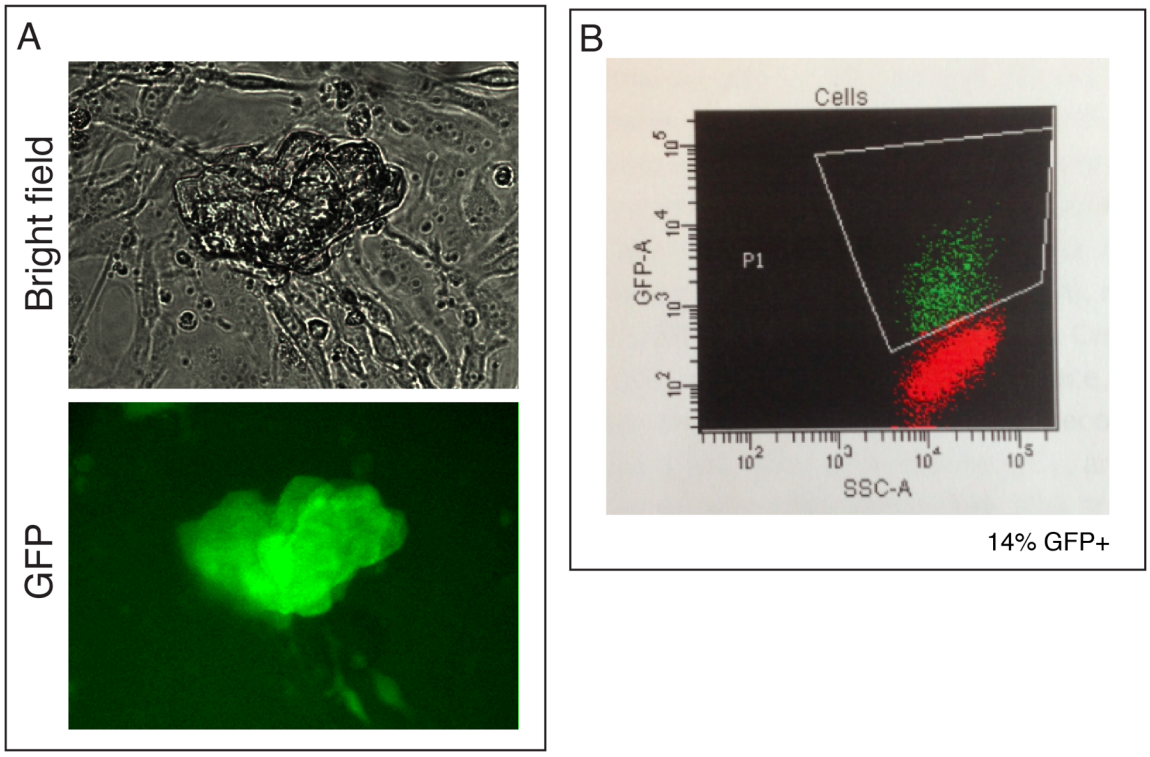


Figure 2.6: GFP-positive cells induced in tail tip fibroblasts carrying reporter allele

(A) Bright field and GFP images of a GFP positive colony of tail tip fibroblasts infected with Lhx2, Sall4, Prdm16, and Nmyc (5 days after viral infection).

(B) Flow cytometry analysis of tail tip fibroblasts infected with Lhx2, Sall4, Prdm16, and Nmyc (7 days after viral infection). 14% of cells are GFP positive.

non-infected MEFs. Strikingly each of the markers was up-regulated in infected conditions as compared to MEFs (Fig. 2.7a).

To further validate this conclusion at a global transcriptome-level, we examined the expression values of GFP-positive sorted cells following RNAseq. GFP-positive cells were sorted seven days after infection with the four-core reprogramming factors (Sall4, Lhx2, Prdm16, and Nmyc). Indeed, almost every known limb bud marker, including all of our initial pool of candidate transcription factors known to mark the limb bud in a tissue-specific manner, were expressed in the GFP-positive cells. Further, we compared transcriptome expression values of the GFP-positive sorted reprogrammed cells with similarly sorted limb bud progenitors harvested from embryos. Interestingly, when comparing the expression values of all transcription factors in the mouse genome mouse forelimb bud cells have a perhaps surprisingly low correlation coefficient to hindlimb bud cells (0.82). However, GFP-positive reprogrammed cells have a similar correlation coefficient to mouse forelimb bud cells (0.77) as (Fig. 2.7b,c). Taken together, these results suggest that Lhx2, Sall4, Nmyc, and Prdm16 are sufficient to globally alter the transcriptome of mouse embryonic fibroblasts to one that bears close resemblance to endogenous limb progenitors.

If the GFP-positive cells are truly reprogrammed, they need to not only have a transcriptional profile reminiscent of limb bud mesenchyme but additionally they have to behave like limb bud mesenchyme. In

Figure 2.7: Characterization of limb progenitor-like cells and the expression of reprogramming genes in the zebrafish and axolotl

(A) qPCR analysis of limb marker genes in MEFs infected with Lhx2, Sall4, Prdm16, and Nmyc, 6 days after infection. Expression values are reported as a log2 ratio of expression in infected MEFs versus non-infected MEFs.

(B) Comparison of transcription factor and co-factor expression values (log2 values, 1299 genes) across mouse forelimb (E9.5) and mouse hindlimb (E10) buds.

(C) Comparison of transcription factor and co-factor expression values (log2 values, 1299 genes) across mouse forelimb (E9.5) and GFP positive MEFs infected with Lhx2, Sall4, Prdm16, and Nmyc.

(D-F) qPCR analysis of differentiation markers Sox9 (D), Gdf5 (E), and Scleraxis (F). Reprogrammed MEFs were infected with Lhx2, Sall4, Prdm16, Nmyc, and Plzf. Undifferentiated reprogrammed MEFs (7 days after viral infection) did not have appreciably different levels of markers than MEFs. Differentiated reprogrammed MEFs expressed appreciably higher levels of all three marker genes.

(G) Alcian stain of differentiated reprogrammed MEFs, undifferentiated reprogrammed MEFs, and MEFs. Only differentiated reprogrammed MEFs stained positively for alcian blue.

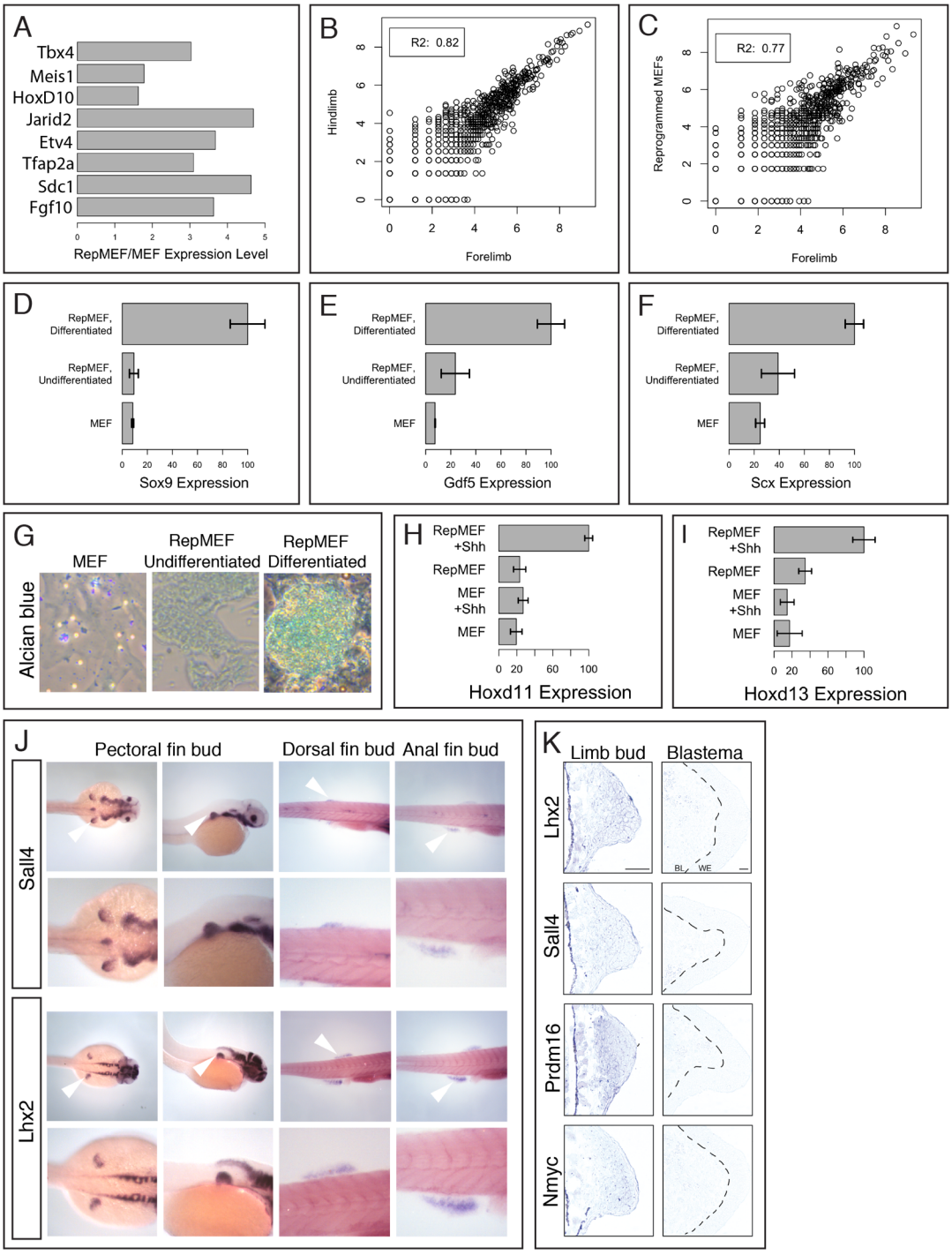
Figure 2.7 (Continued)

(H,I) qPCR analysis of patterning markers, Hoxd11 (H) and Hoxd13 (I) in MEFs and reprogrammed MEFs treated or untreated with Shh.. MEFs treated with Shh did not induce either marker. Reprogrammed MEFs (infected with Lhx2, Sall4, Prdm16, Nmyc, and Plzf) treated with Shh six days after viral infection did induce both markers.

(J) Whole mount *in situ* hybridization of Sall4 and Lhx2 in zebrafish embryos (40 hpf) (pectoral fin bud) and larvae (14 dpf) (dorsal and anal fin buds). Both genes are detected in the pectoral, dorsal, and anal fin buds.

(K) Section *in situ* hybridization of Lhx2, Sall4, Prdm16, and Nmyc in axolotl hatchling limb buds (stage 44-45) and axolotl blastemas (23 days post-amputation). All four genes are detected in the limb bud but not in the blastema. BL: Blastema. WE: Wound epidermis.

Figure 2.7 (Continued)



general, in the reprogramming field, this would involve the capacity to differentiate into the progenitor derivative cell types, however, in the context of limb mesenchyme this is not sufficient; an important additional criteria is that they also be capable of responding properly to patterning signals. Thus, we tested both the differentiation and patterning potential of our putative reprogrammed limb progenitors.

To test differentiation potential, mouse embryonic fibroblasts were reprogrammed with the cocktail of four genes as well as Plzf to ensure maximal levels of functionality. Seven days after reprogramming, the signals necessary to maintain limb progenitors in an undifferentiated state (Chir99021, Fgf8, retinoic acid) were removed. The cultures were maintained for 5 more days to allow the putative reprogrammed cells in the culture dish to differentiate. Five days after culturing, the cells were analyzed with qPCR for the expression of markers for cell types normally derived from limb bud mesenchyme, including Sox9, a marker for chondrogenic mesenchyme; Gdf5, a marker for cells of the joint interzone; and Scleraxis a marker for tendon and ligament precursors. The level of expression of these markers in these differentiated reprogrammed cells was compared to that seen in reprogrammed cells prior to differentiation (7 days after viral infection) and also to MEFs treated under identical conditions but not infected by virus. The expression of all three markers was upregulated in only the condition in which reprogrammed cells were differentiated (Fig. 2.7d-f). The

differentiated reprogrammed cells also stained positively with Alcian blue, a marker for chondrocytes, whereas the MEFs and undifferentiated reprogrammed cells did not (Fig. 2.7g).

To additionally determine whether the reprogrammed cells would respond to patterning signals in a manner similar to endogenous limb progenitors, reprogrammed cells were treated with Sonic hedgehog (Shh), a polarizing signal in the limbs that induces expression of Hoxd11 and Hoxd13, among other targets. 6 days after viral infection, the reprogrammed cells, and MEFs with no viral infection, were treated with Shh. After two days in culture, expression of Hoxd11 and Hoxd13 was assessed by qPCR. The upregulation of both genes was observed in reprogrammed cells exposed to Shh, relative to reprogrammed cells or MEFs that did not receive Shh. Importantly, this activation of patterning genes in response to Shh was not seen in non-reprogrammed MEFs treated with this factor (Fig. 2.7h,i). In sum, the reprogrammed cells appear to share differentiation as well as patterning potential with endogenous limb progenitors.

We have identified a set of four-core genes capable of reprogramming non-limb mesenchyme to a limb progenitor-like state. Of these, Lhx2, Nmyc and Sall4 have been previously characterized in the limb and have critical roles in limb bud outgrowth. Forced Lhx2 down-regulation in the chick limb causes a down-regulation of genes associated with limb progenitor maintenance that leads to the cessation

of limb outgrowth (Rodriguez-Esteban et al., 1998). The conditional deletion of Nmyc in the developing mouse limb leads to decreased proliferation of limb progenitors and, at the tissue level, uniformly small skeletal elements (Ota et al., 2007). Finally, knockdown of Sall4 (along with its homolog Sall1) in zebrafish embryos results in the complete lack of pectoral fins (analogous to forelimbs in mouse and human) (Harvey and Logan, 2006).

While it might be expected that the same, or related, genes would play a role in establishing paired fin mesenchyme in fish, given the high level of conservation between the fin bud and limb bud in their expression of early patterning genes, fish also possess additional unpaired fins such as the median dorsal fin and anal fin. While different in structure and function, the median fins consist of the same cell types as the paired appendages, and moreover are patterned by the same suite of signaling molecules (e.g. Shh, Fgf, RA) as the paired fins (Freitas et al., 2006, Dahn et al., 2007). Thus at a cellular level the early cells of the two types of fins share both developmental potential and responsiveness properties. However, at later stages they express distinct suites of patterning genes and form distinct structures.

To address the possibility that formation of mesenchyme of the median fins might involve the same factors as the paired fins, we examined the expression of two of the genes identified in our study. Using *in situ* hybridization, we detected Lhx2 and Sall4 not only in

pectoral fins (as was already known), but also in the dorsal and anal fins buds of zebrafish (Fig 2.7j). These results suggest that the initial formation of appendage buds uses an ancient transcriptional program that predated the divergence in patterning of the paired and median fins.

More generally, Sall4 has been implicated in maintaining pluripotency in ES cells as well as in reprogramming differentiated cells into iPS cells, including playing a role in repressing differentiation (Yang et al., 2008; Buganim et al., 2014; Yuri et al., 2009). This raised the possibility that Sall4 might also play a role in the dedifferentiation stages of amphibian limb regeneration. Indeed previous studies have shown that Sall4 is expressed in the mid-stage blastema during Xenopus limb regeneration, although expression at the earliest stages of regeneration has not previously been examined (Neff et al., 2005). As the developmental limb bud and the regenerative blastema give rise to similar structures containing the same cell types, we wondered if all four genes we identified as limbness factors would be expressed equivalently in the two systems. Accordingly, we examined their expression in the axolotls. All four genes were detected via *in situ* hybridization in the embryonic axolotl limb bud (stage 44-45); however, they were not detected in the early blastema (Fig. 2.7k). This indicates a fundamental difference between limb bud mesenchyme and the cells of the regenerating limb blastema. These results are not entirely surprising. The cells of the early developing limb bud are multipotent, giving rise to

clonally related bone, cartilage, tendon, ligament, muscle connective tissue, and dermis (Pearse et al., 2007), while the regenerating blastema consists of a mixed population of lineage-restricted tissue-specific progenitors derived from distinct differentiated cell types in the limb (Rinkevich et al, 2011, Lehoczky et al., 2011, Kragl et al., 2009, Stewart and Stankunas, 2012, Tu and Johnson 2011).

Among other motivations, limb regeneration has long been studied in the hopes that lessons from studying animals where this extraordinary phenomenon takes place can be applied to higher vertebrates. However, as highlighted by the differences in transcription factor expression, the process of regeneration is quite distinct from limb development and we are evolutionarily so far removed from these creatures that it is not at all clear that the process of amphibian limb regeneration can ever be recapitulated in a mammal. From a future regenerative medicine perspective, an alternative might be to try to make use of the fact that, at one time, each mammalian organism knew how to make a limb—when it was an embryo. Thus, rather than mimicking the regenerative capacity of a different species, the idea would be to recapitulate embryonic limb development in an adult setting. The potential feasibility of such an approach was shown by a recent study from Jonathan Slack's laboratory (Lin et al., 2013). They worked with frogs that, although capable of full limb regeneration as larva, only form single spikes following amputation as adults. They found that grafting

embryonic limb bud cells to an adult amputated stump, in the context of relevant signaling molecules, can result in regeneration of a segmented limb-like structure with multiple, albeit malformed, digits. While far from being a functional limb, these regenerates greatly exceeded the pattern and anatomical complexity of the spikes that normally form on an amputated adult frog. However, the possibility of some day using embryonic limb bud cells to enhance regeneration only begs the question of whether one could get sufficient numbers of embryonic limb bud cells to grow a new limb on an adult mammal. Our work suggests that, in principle, this might be achievable by reprogramming of adult fibroblasts to cells with the properties of early embryonic limb bud progenitors.

Materials and Methods

Tissue and RNA collection

Fertilized white leghorn chicken eggs (Charles River Laboratories) were incubated at 38 °C. Chicken embryos were staged according to the Hamburger and Hamilton stage series. Forelimb and hindlimb buds were dissected from stage 18 embryos. Flank and neck mesenchyme were dissected from stage 19 embryos (to allow for enough accumulation of mesenchymal cells). Flank tissue was located in the lateral plate mesoderm derived mesenchyme between forelimb and hindlimb buds. Neck tissue was located in the mesenchyme directly above the forelimb

bud. After dissection, cells were incubated in 0.25% trypsin on ice for 5-10 minutes to loosen attached ectodermal tissue. After trypsin incubation, tissue was transferred to 10% chick serum (Life Technologies) in PBS. Loosed ectodermal tissues were removed and remaining mesenchyme was placed in Trizol (Life Technologies) for RNA extraction.

RNA-Seq library construction and analysis

RNA-Seq on chick RNA was carried out as previously described (Christodoulou et al. 2014, Christodoulou et al. 2009). Libraries were constructed without RNA or cDNA fragmentation and did not include normalization. Uniform amplification was achieved with amplification cycling before the reaction reached saturation, as determined by quantitative PCR. Following Hi-Seq (Illumina) sequencing, reads were aligned using Tophat (version 1.4.0) (Trapnell et al. 2009)

Computational analysis

Analysis on transcriptome gene expression was conducted in R. The pvclust package (Suzuki et al. 2006) was used to perform hierarchical clustering. The stats package (prcomp function) was used to perform principle component analysis. The AnimalTFDB online resource (Zhang et al. 2012) was used to select transcription factors from the chick and mouse genomes.

Embryonic and tail tip fibroblast isolation

Homozygous Prx1-creER-ires-GFP mice were acquired from the lab of Shun Murakami (Kawanami et al. 2009) and bred with CD1 mice (Charles River) to generate heterozygous Prx1-creER-ires-GFP embryos or 2 day-old pups. Embryonic fibroblasts were derived from E11.5-12.5 embryos (the head, limbs, lateral plate mesoderm derived tissues, and internal organs were discarded). The dissected embryos were minced with a razor blade and incubated in 0.25% trypsin (Sigma) for 10-15 minutes. The resulting suspension was plated in 10-cm tissue culture dishes in DMEM/10% FBS/1%Pen-Strep media. The cells were grown at 37°C in 3% O_2 (Heracell) until confluent, and then split once before being frozen. Tail tip fibroblasts were derived from tail clippings of 2 day-old pups. The tail clippings were minced with a razor blade and then incubated in 0.25% trypsin for 45 minutes. The resulting suspension was plated in 10-cm tissue culture dishes in DMEM/10% FBS/1%Pen-Strep media. The cells were grown at 37°C in 3% O_2 (Heracell) until confluent, and then split once before being frozen.

Plasmid construction

The coding regions of candidate genes were PCR-amplified from mouse embryo derived cDNA or purchased clones (ThermoScientific). The PCR-amplified sequences were cloned into pDONR221 using the Gateway BP

reaction mix (Invitrogen). The resulting entry clones were then recombined with pMXs-gw (Gift from Shinya Yamanaka – Addgene plasmid #18656) using the Gateway LR reaction mix (Invitrogen).

Production of retroviruses

Plat-E cells (Morita et al., 2000) were grown to 40% confluency in 10-cm dishes. pMXs-based retroviral vectors were introduced to Plat-E cells using Fugene 6 transfection reagent (Roche). 27 µl of Fugene 6 transfection reagent was diluted in 300 µl Opti-MEM (Life Technologies) and incubated for 5 minutes at room temperature. 9 µg plasmid DNA was added to the mixture and then incubated for 15 minutes at room temperature. The Fugene/DNA mixture was added to the Plat-E cells. The cells were incubated overnight at 37°C. The next day, the media was changed with 8 mL of fresh media and the cells were incubated for another 12-24 hours. 36-48 hours after the initial transfection, the media was collected and filtered.

Reprogramming experiments

Passage 1 Prx1-creER-ires-GFP MEFs or TTFs were plated in DMEM/10% FBS/ 1% Pen-Strep and infected with retroviruses supplemented with polybrene (8 µg ml-1). After 12-24 hours, the media containing the retroviruses was replaced with fresh DMEM/10% FBS/1% Pen-Strep media. 48 hours after viral infection, the media was

supplemented with 7.5 μM Chir99021 (Tocris), 150 ng/mL Fgf8 (R&D Systems), and 25 nM retinoic acid (Tocris). Subsequently, the media was changes every two days. GFP expression and morphology changes were assessed using an inverted microscope (Leica). In cases where Shh was added to reprogrammed cells, 2 ng/ul of recombinant protein was used (R&D Systems). For analysis at day 7 after viral infection, cells were fixed 4% paraformaldehyde for 20 minutes at room temperature. After washing three times with PBS, cells were incubated with chick anti-GFP (Abcam) (1:2000 dilution in PBST and goat serum) at 4°C overnight. After incubation, cells were washed 3 times with PBS and then incubated in Alexa-488 conjugated secondary antibody (Invitrogen) (1:500 dilution in PBST in PBS) at room temperature for 1 hour (in dark). Cells were washed 3 times with PBS and then incubated in PBS/DAPI (1:1000 dilution) for 5 minutes. Cells were then wash 3 times with PBS and imaged using an inverted microscope.

Flow cytometry

For GFP and/or mCherry expression analyses, cells were trypsinized (0.25% trypsin) from culture dishes and analyzed on a FACS Aria II SORP (BD Biosciences) with FlowJo software.

SCRB-seq and analysis

Mouse forelimb and hindlimb bud mesenchyme was collected from E9-9.5 CD1 mouse embryos (time pregnant mice were purchased from Charles River). Reprogrammed MEFs (Prx1-GFP positive cells) were collected 7 days after viral infection. Individual cells were FAC sorted into wells of 384-well plates. (HSCI Flow Cytometry core, Cambridge, MA). Library preparation, sequencing, read mapping, and gene expression quantification were carried out by the Broad Technology Labs (Cambridge, MA) as previously described (Soumillon et al. 2014). Unique molecular identifier counts were pooled across cells of a given condition (forelimb, hindlimb, reprogrammed cells) to generate population level averages of gene expression.

Quantitative PCR

RNA was extracted from cells using the RNeasy Mini Kit (Qiagen) and then reverse transcribed with Oligo(dT) primers (Invitrogen) using Superscript III (Invitrogen). cDNA was analyzed by quantitative PCR (qPCR) on the Stratagene MX3000P cycler. Each sample was run in triplicate and quantified by comparison to a standard curve (10-fold serial dilutions of a sample or synthesized gene fragment).

Alcian Staining

Staining was carried out as described previously (Seo et al. 2007). Briefly, media was withdrawn from high-density cell cultures and fixed

with fixative solution (30% EtOH, 0.4% PFA, 4% acetic acid) for 15 minutes at room temperature. Cells were stained with Alcian blue staining solution (0.5% Alcian blue in 75% EtOH, 0.1M HCl) overnight at 37°C.

Zebrafish whole mount in situs

Whole-mount RNA in situ hybridization was performed as described previously (Thisse et al., 1993). The *sall4* probe was described previously (Harvey and Logan, 2006). The *lhx2* probe was described previously (Seth et al., 2006). Pectoral fin buds were analyzed in 30-36 hpf embryos. Anal and dorsal fin buds were analyzed in 5.4 mm and 5.7 mm larval fish (Parichy et al. 2009).

Axolotl section in situs

All axolotl experimentation was performed in accordance with institutional IACUC guidelines (protocol 04160). Sequences encoding axolotl orthologs to *sall4*, *prdm16*, *lhx2*, and *nmyc* were derived from a tissue-coded RNA-sequencing dataset comprising >1.5 billion 100-bp paired sequencing reads. These multi-tissue RNA-seq reads were used in conjunction with Trinity software (Haas et al., 2013) to create a de novo annotated reference transcriptome. Reconstructed transcripts from RNA-seq were also annotated using similarities in identities to orthologous genes in other species using OrthoMCL software (Li et al., 2003).

Fragments of the 3' UTR's were amplified from axolotl limb bud cDNA using the following primers: *sall4*: 5'-TGAAGGTAACCCGCTTCTTG-3' and 5'GATGTGCTAAAGCCGAAAGG-3'; *prdm16*: 5'-CCATTCGTTGTACCCAGCTT-3' and 5'-CTGACATCTGGGGGTGAAAG-3'; *lhx2*: 5'-CCACTGTTTGCGTCACTGTT-3' and 5'-TCAATTGATTGGAGGGGTTC-3'; *nmyc*: 5'-TGCTATAAGATGCAGCACCAA-3' and 5'-TGCTTCTGTTCTGACGGATG-3' and subcloned into pGEM-T-easy. In situ probes were generated as previously described (Whited et al., 2011). Limb buds were harvested at stage 44-45 (Nye et al., 2003), and blastemas were harvested at 23 days post amputation as previously described (Whited et al., 2011). Section *in situ* hybridization was preformed as previously described (Murtaugh et al., 2001) with some modifications (Whited et al., 2011).

References

Y. Buganim et al. The developmental potential of iPSCs is greatly influenced by reprogramming factor selection. Cell Stem Cell. 15(3):295-309 (2014).

D.C. Christodoulou, J.M. Gorham, D.S. Herman, J.G. Seidman. Construction of normalized RNA-seq libraries for next-generation sequencing using the crab duplex-specific nuclease. Curr. Protoc. Mol. Biol. Chapter 4:Unit4.12 (2011).

D.C. Christodoulou et al. 5'RNA-Seq identifies Fhl1 as a genetic modifier in cardiomyopathy. J. Clin. Invest. 124(3):1364-70 (2014).

K.L. Cooper *et al.* Initiation of proximal-distal patterning in the vertebrate limb by signals and growth. Science. 332(6033):1083-6 (2011).

R.D. Dahn, M.C. Davis, W.N. Pappano, N.H. Shubin. Sonic hedgehog function in chondrichthyan fins and the evolution of appendage patterning. Nature. 445(7125):311-4 (2007).

R. Freitas, G. Zhang, M.J. Cohn. Evidence that mechanisms of fin development evolved in the midline of early vertebrates. Nature. 442(7106):1033-7 (2006).

J. Gros, C.J. Tabin. Vertebrate limb bud formation is initiated by localized epithelial-to-mesenchymal transition. Science. 343(6176):1253-6 (2014).

B.J. Haas et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8(8):1494-512 (2013).

V. Hamburger, H.L. Hamilton. A series of normal stages in the development of the chick embryo. Dev. Dyn. 195(4):231-72 (1992).

S.A. Harvey, M.P. Logan. sall4 acts downstream of tbx5 and is required for pectoral fin outgrowth. Development. 133(6):1165-73 (2006).

A. Kawanami, T. Matsushita, Y.Y. Chan, S.A. Murakami. Mice expressing GFP and CreER in osteochondro progenitor cells in the periosteum. Biochem. Biophys. Res. Commun. 386(3):477-82 (2009).

M. Kragl et al. Cells keep a memory of their tissue origin during axolotl limb regeneration. Nature. 460(7251):60-5 (2009).

J.A. Lehoczky, B. Robert, C.J. Tabin. Mouse digit tip regeneration is mediated by fate-restricted progenitor cells. Proc. Natl. Acad. Sci. U. S. A. 108(51):20609-14 (2011).

L. Li, C.J. Stoeckert, D.S. Roos. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. 13(9):2178-89 (2003).

G. Lin, Y. Chen, J.M. Slack. Imparting regenerative capacity to limbs by progenitor cell transplantation. Dev. Cell. 24(1):41-51 (2013).

S. Morita, T. Kojima, T. Kitamura. Plat-E: an efficient and stable system for transient packaging of retroviruses. Gene Ther. 7(12):1063-6 (2000).

L.C. Murtaugh, L. Zeng, J.H. Chyung, A.B. Lassar. The chick transcriptional repressor Nkx3.2 acts downstream of Shh to promote BMP-dependent axial chondrogenesis. Dev. Cell. 1(3):411-22 (2001).

A.W. Neff et al. Expression of Xenopus XlSALL4 during limb development and regeneration. Dev. Dyn. 233(2):356-67 (2005).

H.L. Nye, J.A. Cameron, E.A. Chernoff, D.L. Stocum. Extending the table of stages of normal development of the axolotl: limb development. Dev. Dyn. 226(3):555-60 (2003).

S. Ota, Z.Q. Zhou, D.R. Keene, P. Knoepfler, P.J. Hurlin. Activities of N-Myc in the developing limb link control of skeletal size with digit separation. Development. 134(8):1583-92 (2007).

D.M. Parichy, M.R. Elizondo, M.G. Mills, T.N. Gordon, R.E. Engeszer. Normal table of postembryonic zebrafish development: staging by externally visible anatomy of the living fish. Dev. Dyn. 238(12):2975-3015 (2009).

R.V. Pearse, P.J. Scherz, J.K. Campbell, C.J. Tabin. A cellular lineage analysis of the chick limb bud. Dev. Biol. 310(2):388-400 (2007).

Y. Rinkevich, P. Lindau, H. Ueno, M.T. Longaker, I.L. Weissman. Germ-layer and lineage-restricted stem/progenitors regenerate the mouse digit tip. Nature. 476(7361):409-13 (2011).

C. Rodriguez-Esteban et al. Lhx2, a vertebrate homologue of apterous, regulates vertebrate limb outgrowth. Development. 125(20):3925-34 (1998).

M.A. Ros, G.E. Lyons, S. Mackem, J.F. Fallon. Recombinant limbs as a model to study homeobox gene regulation during limb development. Dev. Biol. 166(1):59-72 (1994).

H.S. Seo, R. Serra. Deletion of Tgfbr2 in Prx1-cre expressing mesenchyme results in defects in development of the long bones and joints. Dev. Biol. 310(2):304-16 (2007).

A. Seth et al. belladonna/(Ihx2) is required for neural patterning and midline axon guidance in the zebrafish forebrain. Development. 133(4):725-35 (2006).

S. Stewart, K. Stankunas. Limited dedifferentiation provides replacement tissue during zebrafish fin regeneration. Dev. Biol. 365(2):339-49 (2012).

M. Soumillon , D. Cacchiarelli , S. Semrau , A.V. Oudenaarden , T.S. Mikkelsen. Characterization of directed differentiation by high-throughput single-cell RNA-Seq. bioRxiv. http://dx.doi.org/10.1101/003236 (2014).

R. Suzuki, H. Shimodaira. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics. 22, 1540–1542 (2006)

K. Takahashi, S. Yamanaka. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 126(4):663-76 (2006).

C. Thisse, B. Thisse, T.F. Schilling, J.H. Postlethwait. Structure of the zebrafish snail1 gene and its expression in wild-type, spadetail and no tail mutant embryos. Development. 119(4):1203-15 (1993).

C. Trapnell, L. Pachter, S.L. Salzberg. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 25(9):1105-11 (2009).

P. Tschopp *et al.* A relative shift in cloacal location repositions external genitalia in amniote evolution. Nature. 516(7531):391-4 (2014).

S. Tu, S.L. Johnson. Fate restriction in the growing and regenerating zebrafish fin. Dev. Cell. 20(5):725-32 (2011).

J.L. Whited, J.A. Lehoczky, C.A. Austin, C.J. Tabin. Dynamic expression of two thrombospondins during axolotl limb regeneration. Dev. Dyn. 240(5):1249-58 (2011).

J. Yang et al. Genome-wide analysis reveals Sall4 to be a major regulator of pluripotency in murine-embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. 105(50):19756-61 (2008).

S. Yuri et al. Sall4 is essential for stabilization, but not for pluripotency, of embryonic stem cells by repressing aberrant trophectoderm gene expression. Stem Cells. 27(4):796-805 (2009).

H.M. Zhang et al. AnimalTFDB: a comprehensive animal transcription factor database. Nucleic Acids Res. 40(Database issue):D144-9 (2012).

Chapter 3

A Limb Progenitor Response Model for the Molecular Initiation of Morphogenesis

A Limb Progenitor Response Model for the Molecular Initiation of Morphogenesis

Abstract

During embryonic development, fields of progenitor cells form complex spatial structures through a dynamic interaction with external signaling molecules (morphogens). For instance, the limb bud, which initially begins as a hemisphere, elongates and ultimately gives rise to the skeletal pattern observed in vertebrate limbs. To initiate this morphological response, progenitor cells must translate information about extracellular levels of morphogen into appropriate gene expression responses. The morphogen Sonic hedgehog (Shh) has been extensively studied in the developmental context of limb patterning. It is expressed in the distal posterior mesenchyme of the limb bud and was established as the key mediator of anterior-posterior (AP) patterning through experiments that showed that it could ectopically induce the AP axis. In this study, a limb progenitor culture system was used to study the limb progenitor gene expression response to Shh and Fgf signals. The limb progenitor response to Shh was found to be a simple ON/OFF switch, a response that is far simpler than those predicted by morphogen gradient models. However, additional complexity in the limb progenitor response was uncovered when the Shh response was assessed in conjunction with Fgf dose. These results highlight the importance of studying morphogen mediated response in context with other signals.

Introduction

Molecular models of morphogenesis

A number of models have focused on understanding how key molecular components work in concert with each other to instantiate the process of morphogenesis. It has been well appreciated that protein signals are secreted by source cells to neighboring cells than then influence morphogenetic behavior (Rogers and Schier, 2011). In what manner do these extrinsic signals influence the molecular stoichiometry of genes within cells through changes in gene expression? Perhaps the most well known model for this process is the French flag model. In this model a secreted molecule (or morphogen), is secreted from a point source, as the molecule diffuses from this source, it creates a gradient of concentration (Wolpert, 1969). This gradient of concentration across a field of cells results in different swatches of cells experiencing different levels of the morphogen. Within the cell, different genes are activated by different thresholds of morphogen concentrations. In a slight variant, different concentrations could result in different levels of expression, also creating a heterogeneity in gene expression across the field of limb progenitors. This heterogeneity in gene expression across the field of cells creates the complexity that is then translated to the complex forms that emerge out of the uniform tissue. Thus, in the case of this model, morphogens transmit information to cells through gradations of dose (Wolpert, 2007). A number of previous studies have attempted to use in vivo grafting of Shh into

the early limb bud to test the validity of this model (Yang et al., 1997, Wada et al., 1999, Drossopoulou et al., 2000). Although embryological experiments have provided clues as to how limb progenitors interpret morphogens, their reliance on morphological and non-quantitative molecular readouts have obscured a decisive understanding of morphogen interpretation. Furthermore, the primary method of ectopically applying morphogens has been through bead implantation, which does not allow for precise delivery of concentration and temporal doses of Shh in a spatially and developmentally controlled context.

More recent models have also suggested that morphogens can transmit information to cells through the time of exposure (Cohen et al., 2013). For instance, studies from the developing neural tube have led to a model in which cells respond to Shh by a temporal adaptation mechanism (Dessaud et al., 2007). This form of response is characterized by two phases. In the first phase (0-12 hrs), the output of the signal transduction pathway, as measured by GLI activity, is high and similar across high and low concentrations of Shh. In the second phase (12-24 hrs), the output of the pathway decays at a rate that is inversely proportional to the concentration of Shh delivered. Thus, neural tube progenitors are able to integrate concentration and duration of Shh exposure such that the concentration of Shh produces a time-limited period of signal transduction that is proportional to ligand concentration. It is unclear whether a temporal adaptation mechanism occurs in the limb, given the lack of a quantitative understanding between Shh dose and gene expression output.

Key molecular components at play during limb morphogenesis

A number of genetic analyses, originating with the Weicshaus and Nüsslein-Volhard morphogenetic mutant screens in fruit flies, have consistently identified conserved sets of genes that when inactivated result in pattern formation disruption. These genes are enriched for function in cell-cell communication (secreted ligands) and gene regulation (transcription factors) (Nüsslein-Volhard and Wieschaus, 1980).

In the limb a number of genes have been shown to be critical for proper skeleton formation in the limb. From the perspective of signaling, the Sonic hedgehog (Shh) pathway components have proved to be critical for generating complexity across the anterior posterior axis of the limb (Riddle et al., 1993, Harfe et al., 2004). With regard to its expression, Shh is secreted from the posterior region of the proto-limb and influences the expression of a number of genes, including directly pathway targets such as Ptch1 and the Gli proteins (Marigo and Tabin, 1996, Marigo et al., 1996a). In the limb, Ptch1 has been used as a spatial read out to determine the range of Shh influence (Marigo et al., 1996b).

When Shh is inactivated in the limb, a significant number of bones in the limb are lost (Pagan et al., 1996). Conversely, when Shh is ectopically added to the anterior proto-limb, a significant number of bones in the limb are added (Riddle et al., 1993).

The pathway through which Shh influences gene transcription in limb progenitors has been characterized through molecular analyses. Briefly, Shh

binds to a receptor on the cell surface, Patched1 (Ptch1), which then through a series of molecular steps results in the cleavage of a transcriptional repressor, Gli3. The absence of repression results in gene activation of a number of genes, including the Shh receptor, Ptch1 (Hooper and Scott, 2005). The primary piece of evidence supporting the role of Gli3 as the effector of Shh signaling in the limb is that the loss of bones in a Shh mutant can be rescued with the addition of a Gli3 mutation (Litingtung et al., 2002).

Shh has a critical role in activating the expression of Hoxd11, Hoxd12, and Hoxd13 (Laufer et al., 1994). These genes, which a together referred to as the 5' HoxD genes, are expressed in a pattern in which Hoxd11 and Hoxd12 occupy a larger swath of the posterior limb as compared to Hoxd13 (Nelson et al., 1996). A later phase of 5' HoxD expression occurs once the pro-limb establishes the hand region. In this phase, Hoxd13 dominates spatial expression (Nelson et al., 1996).

This change in expression is now understood to be due to a change in enhancer landscapes (Montavon and Duboule, 2013). The early phase of HoxD expression, which is most relevant to understanding the initiation of gene expression asymmetry in the limb progenitor, is dependent on a set of enhancers proximal to HoxD11. In contrast, the enhancers necessary for late phase HoxD expression are proximal to HoxD13 (Andrey et al., 2013)

In the limb, it is known that Shh is necessary to maintain 5'HoxD expression. In mutants lacking Shh expression, 5' HoxD expression is absent in the developing limb bud (Ros et al., 2003). Additionally, ectopic application

of Shh in the anterior, Shh-free portion of the proto-limb leads to ectopic expression of 5'HoxD expression (Laufer et al., 1994). Importantly, it should be noted that for Shh to induce 5'HoxD expression, the ectopic expression must be placed in close enough proximity to the epithelial layer encasing the limb mesenchyme in order to receive Fgf signals.

Open questions and approaches concerning limb morphogenesis:

The goal of this work is to develop a compact model of molecular regulation between external signals and gene expression of key genes critical for limb morphogenesis. To do so, empirical observations of the relationship between Shh (and Fgf) and 5' HoxD genes will be made primarily using a limb progenitor culture system that allows for quantitative dosing of ligands and quantitative measurement of gene expression (Cooper et al., 2011). Thus, these naïve progenitors serve as a model for the initiation of tissue polarization in the proto-limb.

These observations will be used to question whether predictions made by prevalent molecular models of morphogen regulation of gene expression hold true in the limb. For instance, do limb progenitors exhibit differential gene activation on the basis of either Sonic hedgehog concentration or duration?

In addition, insights from these observations will better define how morphogens work in concert with other signaling molecules to regulate gene expression. More specifically, given the in vivo knowledge that Shh requires the

presence of Fgf to activate HoxD gene expression, how can one model their control of HoxD expression together, rather than in isolation?

Results

Characterizing the limb progenitor direct-gene-target expression response to Shh ligand

Before establishing the relationship between Shh exposure and limb progenitor 5' Hoxd gene response, an independent baseline response was characterized by assaying the response of Gli1 and Ptch1, both of which are established direct targets of the Shh pathway in the limb. To determine the concentration dose dependency to Shh, limb progenitors were treated with increasing doses of Shh ligand (Active N-terminal fragment) and then assayed for transcript levels of Gli1 or Ptch1 with quantitative PCR (qPCR). In response to increasing Shh dose, both genes display a sharp inflection point of expression increase, indicating a switch-like response to Shh (Fig. 3.1a, b). To determine temporal dependency to Shh, primary limb progenitors were treated with a fixed amount of Shh (a high dose of 2 ng/ul, as determined by the titration curve in Fig. 3.1a) for increasing periods of time ranging from 12 to 36 hours and then assayed for levels of Ptch1 transcripts. Over the time course, levels of Ptch1 expression increased linearly with respect to time of Shh exposure (Fig. 3.1c). This increase in Ptch1 expression over time was observed across a range of Shh concentrations (Fig. 3.1d).

These results suggest that the limb progenitor direct-gene-target response to Shh is not graded, as would be predicted by morphogen gradient

models, but is instead switch-like. Additionally, the Ptch1 response does not exhibit temporal adaptation with respect to signal sensing (as observed in neural tube progenitors), but rather continues to amplify with respect to time.

Characterizing the limb progenitor 5' HoxD gene expression response to Shh or Fgf dose

5' HoxD (Hoxd11, Hoxd12, and Hoxd13) expression levels in limb progenitors were measured in response to increasing Shh dose. In a manner similar to Ptch1, all three genes were activated in a switch-like manner. Additionally, there was no difference appreciable difference in the threshold of activation among the three genes (Fig. 3.1e). This switch-like profile was also observed for BMP2 (Fig. 3.1f). Thus, all genes tested showed that the information transmitted from Shh ligand to gene expression is limited to two options, an ON state and an OFF state. Furthermore, the activation dose required for all these genes are similar.

While Shh appears to exhibit simple dose and temporal dependence characteristics when considered in isolation, additional complexity could be created by coupling Shh regulation of genes with another signaling ligand. Indeed, based on grafting experiments in the embryo, Shh requires the presence of Fgf to activate gene expression. To understand the co-regulation of Shh and Fgf of gene expression in limb progenitors, additional dose titration experiments were conducted by varying Fgf in conjunction with Shh. When Fgf concentration is varied, neither Ptch1 nor Gli1 respond to Fgf at any supplied

Figure 3.1: Characterization of gene expression responses of limb progenitors to either Shh or Fgf ligands

(A-B) Shh ligand dose and Gli1/Ptch1 expression titration curve: Naïve limb progenitors were dosed with varying concentrations of Shh for 24 hours (3 biological replicates for each concentration dose). Expression levels of Gli1 or Ptch1 transcripts were measured via qPCR and normalized to Actin transcript expression levels (levels reported in relative expression units). Fgf was present at a fixed concentration of 120 ng/ul.

(C) Shh ligand time of exposure and Ptch1 expression titration curve: Naïve limb progenitors were treated with Shh ligand for 12, 24, and 36 hours (3 biological replicates for each temporal dose). Expression levels of Ptch1 transcripts were measured via qPCR and normalized to Actin transcript expression levels (levels reported in relative expression units).

(D) Ptch1 temporal titration curves with varying levels of Shh ligand. Limb progenitors were treated for two periods of time (10 and 24 hours) at one of four doses of Shh ligand (0, 0.03, 0.25, 2 ng/ul) (3 biological replicated for each unique temporal and concentration dose). Expression levels of Ptch1 transcripts were measured via qPCR and normalized to Actin transcript expression levels (levels reported in relative expression units).

(E-F) Shh ligand dose and 5' HoxD/Bmp2 gene expression titration curves. Limb progenitors were treated for 24 hours with varying levels of Shh ligand (3 biological replicates for each dose). Expression levels of Hoxd11, Hoxd12, Hoxd13, and Bmp2 were measured via qPCR and normalized to Actin

Figure 3.1 (Continued)

transcript expression levels (levels reported in relative expression units). Fgf8 ligand was supplied at a fixed dose of 120 ng/ul.

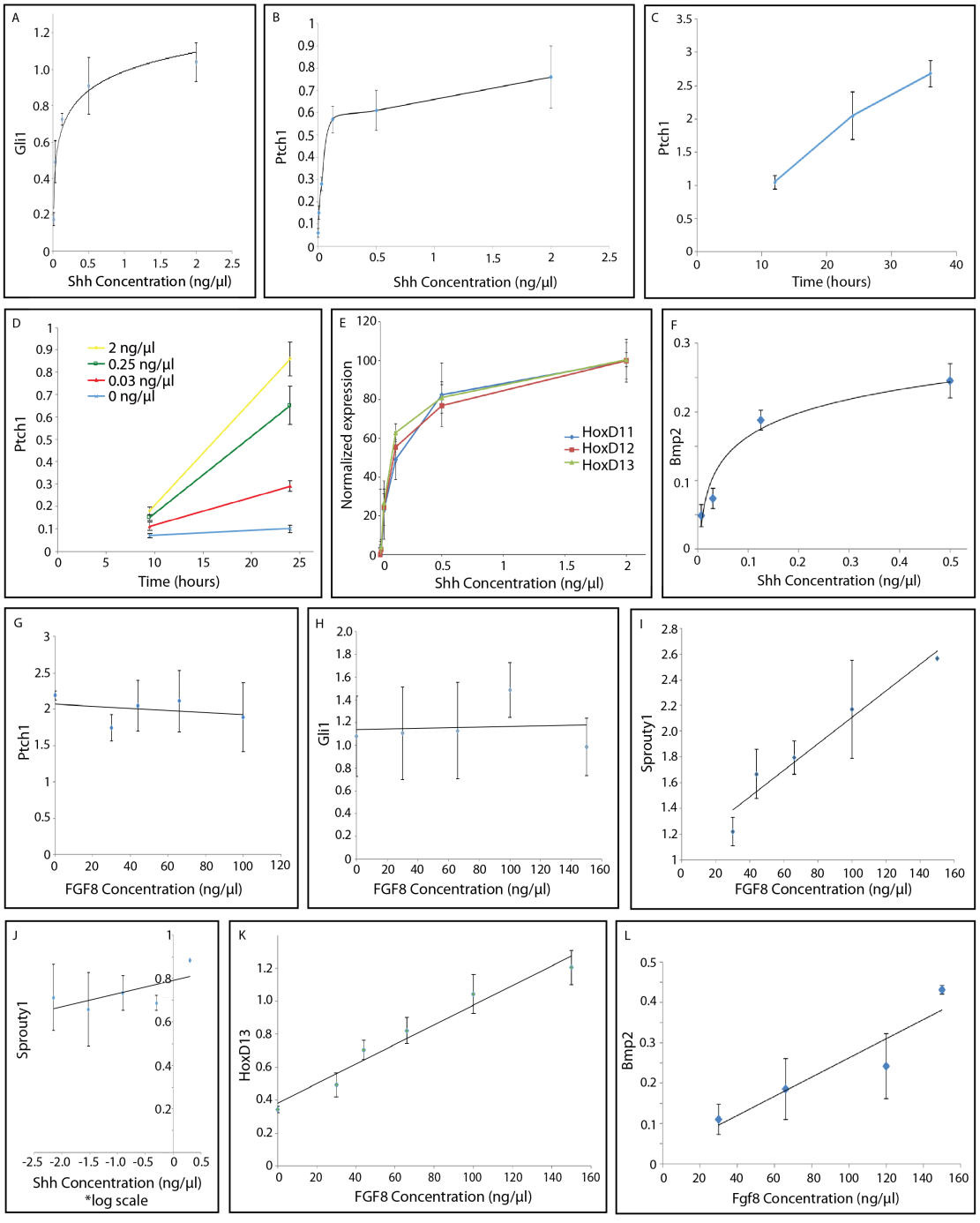
(G-H) Fgf ligand dose and Ptch1/Gli1 expression titration curve. Limb progenitors were treated with increasing doses of Fgf8 (3 biological replicates for each dose) for 12 hours. Expression levels of Ptch1/Gli1 were measured via qPCR and normalized to Actin transcript levels.

(I) Fgf ligand dose and Sprouty1 expression titration curve. Limb progenitors were treated with increasing doses of Fgf8 (3 biological replicates for each dose) for 12 hours. Expression levels of Sprouty1 were measured via qPCR and normalized to Actin transcript levels.

(J) Shh ligand dose and Sprouty1 expression titration curve. Limb progenitors were treated with increasing doses of Shh (3 biological replicates for each dose) for 24 hours. Expression levels of Sprouty1 were measured via qPCR and normalized to Actin transcript levels.

(K-L) Fgf ligand dose and Hoxd13/Bmp2 expression titration curve. Limb progenitors were treated with increasing doses of Fgf8 (3 biological replicates for each dose) for 12 hours. Expression levels of Hoxd13 or Bmp2 were measured via qPCR and normalized to Actin transcript levels. Shh was supplied at a fixed dose of 1 ng/ul.

Figure 3.1 (Continued)



dose (Fig. 3.1g, h). Conversely, as Fgf concentration increases, Sprouty, an established direct target of the Fgf pathway, exhibits a linear response in expression (Fig. 3.1i). However, when Shh dose is increased, Sprouty does not respond at any dose level (Fig. 3.1j). Thus, in limb progenitors, the direct targets of the Fgf pathway are insulated from exposure to Shh and the direct targets of the Shh pathway are insulated from the Fgf pathway.

With this baseline understanding of direct target response, the activation dynamics of 5' HoxD genes by Fgf was explored. In the presence of fixed amounts of Shh, when Fgf dose was increased, the expression of Hoxd13 increased in a linear fashion (Fig. 3.1k). In addition, the response of Bmp2 responded to Fgf dose in a linear fashion (Fig. 3.1l). Thus, Fgf influences the expression levels of Hoxd13 in a manner similar to its direct pathway target, Sprouty. When the Shh dose titration was repeated in two variations, with a high dose of Fgf and a low dose of Fgf, the lower dose of Fgf cut maximal expression of the Hoxd13 in a linear manner. Thus, no amount of Shh could rescue the loss of expression brought about by lower amounts of Fgf (Fig. 3.2a). Conversely, when the Fgf dose titration curve was repeated with either a high dose of Shh or a low dose of Shh, the low dose flattened the Hoxd13 response curve to marginal expression levels, indicating a drastic switch in expression levels (Fig. 3.2b).

Taken together it appears that Shh activates limb progenitor HoxD gene expression in a switch-like manner that is distinctly different from the linear manner exhibited by Fgf. This fact, coupled with the fact that first responders

Figure 3.2: Characterization of Shh and Fgf co-regulation of Hoxd13 gene expression in limb progenitors

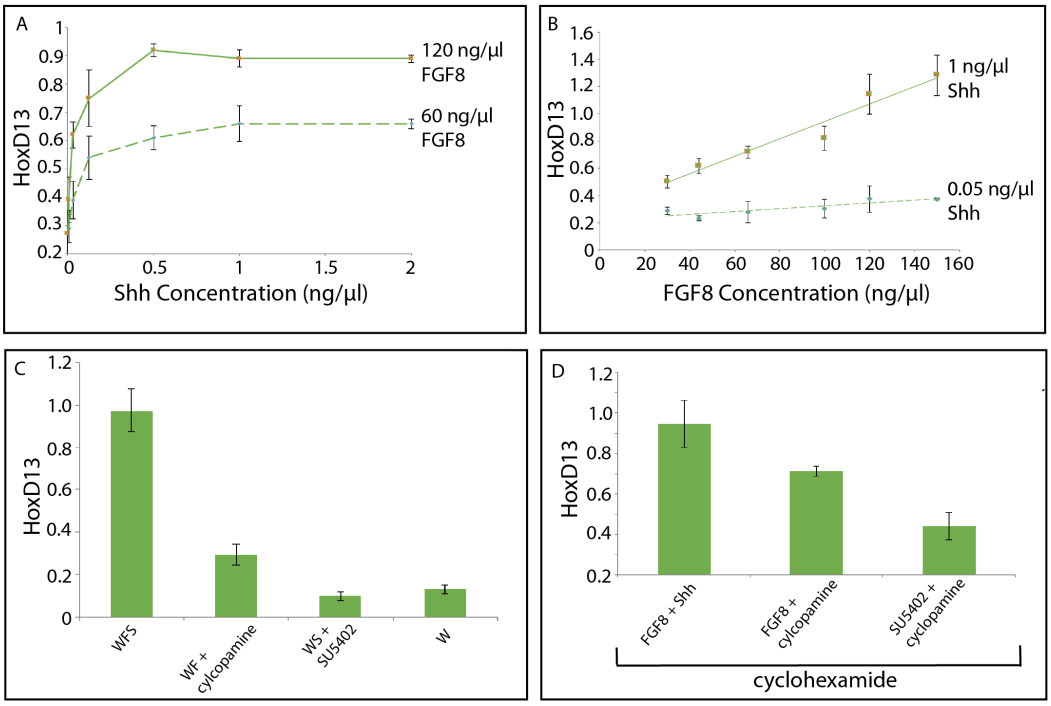
(A) Shh ligand and Hoxd13 titration curve, at two different doses of Fgf8. Limb progenitors were treated with increasing amounts of Shh at one of two concentrations of Fgf8 (3 biological replicates for each condition). Expression levels of Hoxd13 were measured via qPCR and normalized to Actin transcript levels.

(B) Fgf8 ligand and Hoxd13 titration curve, at two different doses of Shh. Limb progenitors were treated with increasing amounts of Fgf8 at one of two concentrations of Shh (3 biological replicates for each condition). Expression levels of Hoxd13 were measured via qPCR and normalized to Actin transcript levels.

(C) Assessment of the response of Hoxd13 gene expression in limb progenitors to Shh and/or Fgf ligand. Limb progenitors were treated with 1) Fgf8 and Shh 2) Fgf8 and Shh antagonist cyclopamine 3) Shh and Fgf antagonist SU5402 4) no signal for 12 hours (3 biological replicates for each dose). Expression levels of Hoxd13 were measured via qPCR and normalized to Actin transcript levels.

(D) Assessment of the response of Hoxd13 gene expression in limb progenitors to Shh and/or Fgf ligand in the presence of cyclohexamide. Limb progenitors were treated with cyclohexamide in conjunction with one of the following three conditions for 12 hours 1) Fgf8 and Shh 2) Fgf8 and cyclopamine 3) SU5402 and cyclopamine (3 biological replicates for each dose). Expression levels of Hoxd13 were measured via qPCR and normalized to Actin transcript levels.

Figure 3.2 (Continued)



to each pathway are regulated in a similar manner but are nonetheless insulated from the other pathway, suggests that mechanisms of action within the limb progenitor are independent.

To determine the necessity and sufficiency of Shh and Fgf for Hoxd13 gene expression activation, limb progenitor cells were dosed with Shh and Fgf, Shh and a Fgf pathway inhibitor, Fgf and a Shh pathway inhibitor, or both pathway inhibitors (Fig. 3.2c). The addition of inhibitors ensured the blocking of trace amounts of Shh and Fgf in the culture medium. Fgf was sufficient to activate a modest level of Hoxd13, in the absence of Shh, as compared to the baseline condition of Fgf and Shh pathway inhibitors. Conversely, Shh alone was unable to activate any Hoxd13 expression as compared to the baseline condition. Lastly, Shh and Fgf together were able to synergistically activate Hoxd13 expression levels. These results are consistent with a model in which Fgf potentiates positive regulators of 5' HoxD transcription whereas Shh relieves the repression of a negative, switch-like, regulator. This model is consistent with published literature showing that the main effector of the Shh pathway in the limb is Gli3, which acts as a repressor in the absence of Shh and is destroyed in the presence of Shh.

These necessity and sufficiency experiments were repeated in the presence of cylcohexamide to determine the extent to which Hoxd13 activation was dependent on positive transcriptional feedback (Fig. 3.2d). Fgf alone was able to modestly activate transcription, with a slight increase in expression when Shh was present. Thus, the system is primed to respond to Shh and Fgf

in the absence of translation, however, at a dampened level, suggesting the need for transcriptional feedback to reach maximal levels of HoxD expression.

Characterizing the limb progenitor 5' HoxD gene expression maintenance dependency on Shh or Fgf dose

While the previous experiments provided an understanding of Shh and Fgf dose dependency on 5' HoxD gene *activation*, they do not provide information about the need of those signals for the maintenance of expression. To test the need of Shh and Fgf for HoxD expression maintenance after activation, limb progenitor cells were treated with Shh and Fgf for a period of 12 hours after which they were treated with media lacking both signals and containing Shh and Fgf inhibitors (cyclopamine and SU5402, respectively). Expression levels of Ptch1, Sprouty1, Hoxd11, Hoxd12, and Hoxd13 were measured at time points after the switch to media lacking signal (Fig. 3.3a). Baseline null expression was determined by sampling expression in progenitor cells treated with only signal inhibitors for 12 hours. Once Shh and Fgf are withdrawn, both Ptch1 response and Sprouty response decay to baseline levels in two hours, indicating the loss of active signaling influence of both Shh and Fgf (Fig. 3.3b, c). In the absence of both signals, Hoxd13 expression decays and reaches baseline levels in four hours (Fig. 3.3d). Conversely, Hoxd11 and Hoxd12 expression remains the same for the first four hours and then increases rapidly (Fig. 3.3e, f). These results suggest that the temporal regulation of Hoxd13 is distinct from Hoxd11 and Hoxd12. Additionally, it

suggests that Hoxd11 and Hoxd12 expression levels are suppressed by Hoxd13 expression. To determine whether Hoxd13 maintenance is dependent on a single signal or a combination of the two, limb progenitor cells were treated with Shh and Fgf for 4, 11, or 18 hours after which they were treated with either Fgf, Shh, both signals or neither signal (Fig. 3.3h-j). The general trend observed across these variants was that while Hoxd13 expression fell despite the presence of Shh, it was maintained by the presence of Fgf alone. Thus, taken together it appears that while the activation dynamics across Hoxd11, Hoxd12 and Hoxd13 exhibit similar features, the maintenance of their expression diverges.

In vivo characterization of Shh positive and negative regions of the early limb bud

To better contextualize the relationship dynamics observed between external signals (Shh, Fgf) and gene targets (5' HoxD genes), Shh positive regions of the limb bud were profiled in comparison to Shh negative regions. For these experiments, tissues were collected from stage 21-22 embryos, a stage comparable to the cells harvested for in vitro titration experiments described above.

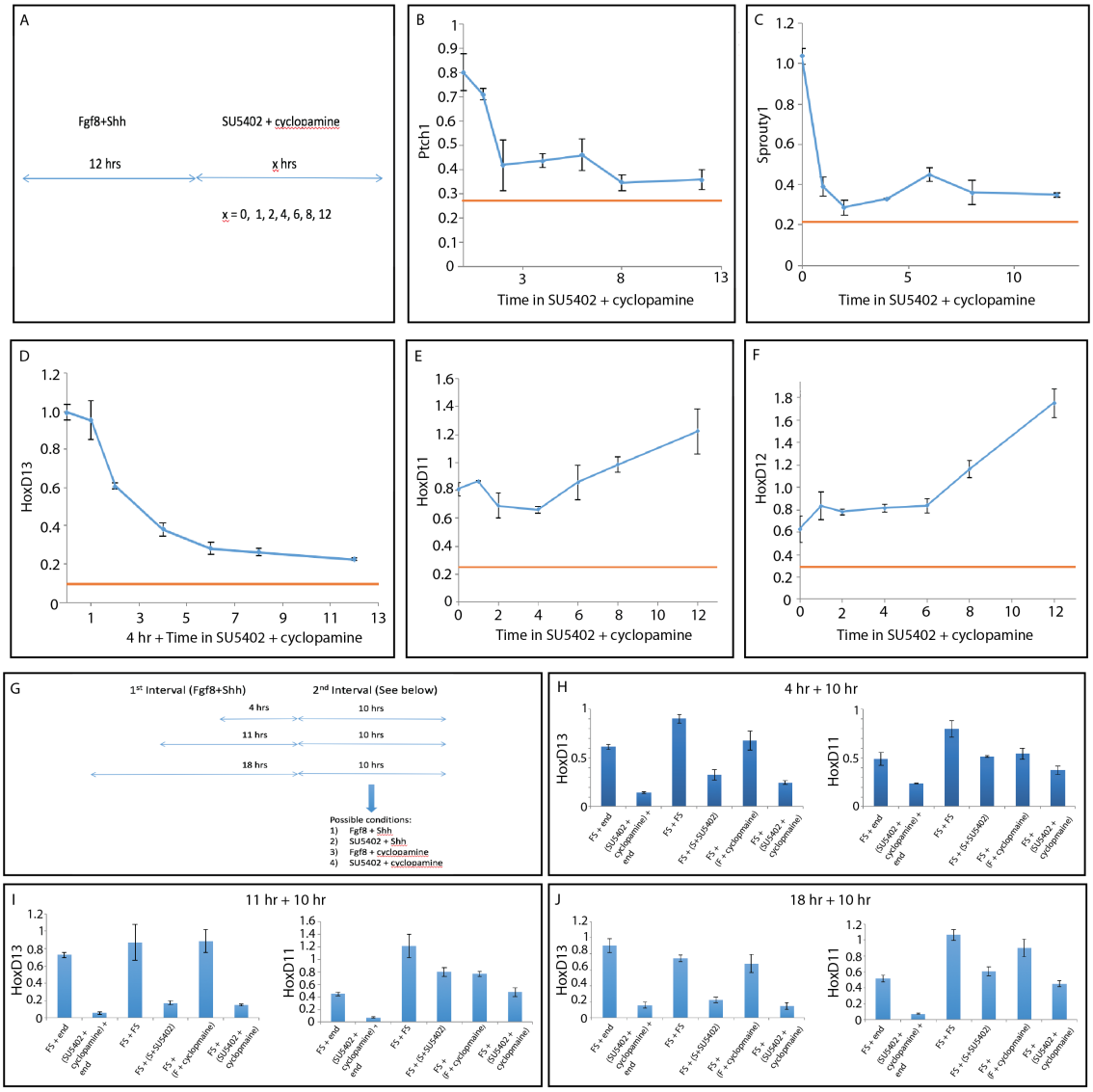
To get a better sense of the gene regulatory differences between Shh positive and Shh negative limb progenitors, the HoxD cluster was biochemically characterized using chromosomal capture and histone modification analysis. Chromosomal capture was used to determine which portions of regulatory DNA

Figure 3.3: Temporal dependency of limb progenitor gene expression on Fgf and Shh ligands.

(A-F) Limb progenitors were treated with Fgf8 and Shh for 12 hours, after which both signals were withdrawn from media and replaced with SU5402 (Fgf) and cyclopamine (Shh) antagonists. Samples were collected 0, 1, 2, 4, 6, 8, 12 hours after withdrawal of Fgf8 and Shh (3 biological replicates for each time point. A schematic of the experiment is presented in (A). Levels of expression were measured via qPCR for Ptch1 (B), Sprouty1 (C), Hoxd13 (D), Hoxd11 (E), and Hoxd12 (F).

(G-J) Limb progenitors were treated with in two distinct phases. In the first phase, progenitors were treated with Fgf8 and Shh for either 4, 11, or 18 hours. In the second phase, progenitors were treated for 10 hours under one of four possible conditions: 1) Fgf8 and Shh 2) Fgf antagonist SU5402 and Shh 3) Fgf8 and Shh antagonist cyclopamine 4) Fgf antagonist SU5402 and Shh antagonist cyclopamine. A schematic of the experiment is presented in (G). Each distinct combination was carried out in biological triplicate. Levels of expression were measured for Hoxd13 and Hoxd11 via qPCR.

Figure 3.3 (Continued)



were in physical association with the Hoxd11 promoter. Using 4C, which couples the biochemical assay of chromosomal capture with next-generation sequencing, Shh postive cells (Hoxd11-13 positive) were profiled and compared to Shh negative cells (Hoxd11-13 negative) (Fig. 3.4a). Both samples had strikingly similar binding profiles, indicating that the lack of expression in Shh negative cells is not due to the lack of physical association between regulatory DNA and the Hoxd11 promoters. The 4C profiles were used to identify the genomic regions containing the key distal enhancer regions involved in 5' HoxD expression. With these regions in mind, H3K27Ac and H3K27Me3 histone modification marks were profiled via ChIP-seq to determine their chromatin state (active - H3K27Ac, repressed - H3K27Me3). In both Shh positive and Shh negative limb regions, the distant regulatory elements were marked with active H3K27Ac marks, albeit with a lower level of magnitude in the Shh negative regions of the limb bud (Fig 4.4b). Conversely, the regulatory regions adjacent to the 5' HoxD promoters contained high levels of H3K27Ac marks in the Shh positive limb bud tissue and low levels in the Shh negative region (Fig 4.4b). When H3K27Me3 marks were profiled in the same two tissues, high levels of H3K27Me3 were observed only in the proximal DNA region of Shh negative tissue (Fig 4.4c). Taken together, these results suggest that HoxD gene expression is dependent on the release of repression of proximal DNA regions. Once this release occurs, active transcription can occur under, in part, the influence of enhancers that are already active.

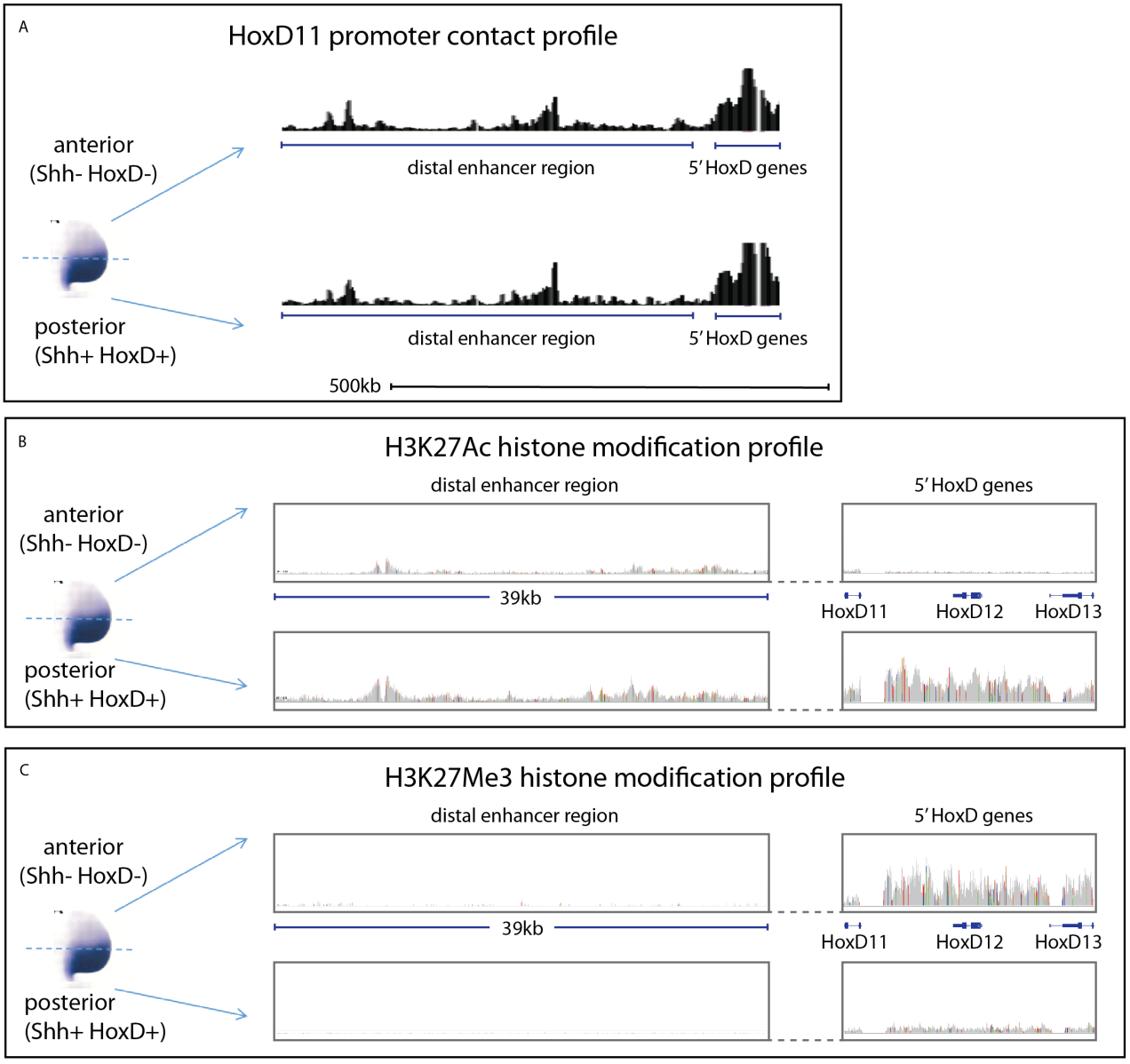
Figure 3.4: In vivo characterization of Shh positive and Shh negative regions of the early limb bud

(A) Profile of Hoxd11 promoter contacts with neighboring DNA regions in the HoxD locus. Samples from the distal anterior (Shh negative, Hoxd11/12/13 negative) and the distal posterior (Shh positive, Hoxd11/12/13 positive) limb bud were profiled. DNA contacts were observed between the Hoxd11 promoter and the 500 kb gene desert area flanking the HoxD cluster.

(B) K27Ac histone modification profiles of the HoxD locus in distal anterior (Shh negative, Hoxd11/12/13 negative) and distal posterior (Shh positive, Hoxd11/12/13 positive) limb bud tissue. A 39 kb sub-portion of the 500 kb gene desert presented in (A) is presented as well as the genomic region containing the 5' HoxD genes.

(C) K27Me3 histone modification profiles of the HoxD locus in distal anterior (Shh negative, Hoxd11/12/13 negative) and distal posterior (Shh positive, Hoxd11/12/13 positive) limb bud tissue. A 39 kb sub-portion of the 500 kb gene desert presented in (A) is presented as well as the genomic region containing the 5' HoxD genes.

Figure 3.4 (Continued)



Discussion

Comparison of results to predictions of morphogen gradient models

Classic morphogen gradient models have been proposed in order to understand how external signals are involved in creating spatial heterogeneity in gene expression, and subsequently form. These models rest on the notion that spatial heterogeneity can occur through differences in dose exposure to an external signal. Differences in dose could manifest themselves in creating different activation thresholds of different sets of genes or by tuning the level of gene expression. In the case of Hoxd11, Hoxd12, and Hoxd13 (as well as direct pathway targets), no differences were observed with respect to the Shh dose required to activate a gene expression response. In addition to the lack of difference in activation, the activation kinetics were switch-like, making the possibility of achieving variance in spatial expression in vivo less likely. Indeed, in vivo, it appears that the direct Shh target, Ptch1, exhibits uniform high expression across the posterior half of the limb bud and minimal uniform expression across the anterior half. One could imagine that while Shh has a relatively simple relationship to the genes explored in this study, there could be additional genes that have more complex responses that are responsible for generating differences in pattern. However, genome wide expression profiling of the anterior (Shh negative) and posterior (Shh positive) halves of the limb bud indicate that the only key developmental genes that are significantly (2-fold difference) differentially expressed are Ptch1/2, Shh, Lmo1, Bmp2, Hoxd11,

Hoxd12, Hoxd13 (expression data provided by personal communication with J. Young and M. Schwartz). Thus, the gene expression response of limb progenitors to Shh appears to be far simpler than morphogen gradient models would predict.

While the morphogen gradient model appears to over-complicate gene expression responses between one signal and gene expression, it is under-developed with respect to thinking about other critical aspects of morphogen mediated gene expression. It does not provide much in the way of understanding how multiple morphogens work together to control gene expression, and how this coupling could be employed to generate spatial variation in gene expression. The co-regulation of gene expression by multiple signaling sources could help explain how rich spatial variation in gene expression could occur, despite the fact that each signal in isolation could have rather simple regulative effects.

A dual perspective model of morphogen mediated co-regulation of genes

The results obtained above can be assimilated to create a simple model of 5' HoxD regulation in limb progenitors. In contrast to classic morphogen models, the model will incorporate two signaling inputs (Shh and Fgf) and will consider two perspectives (gene expression levels, the state of the HoxD locus). At the level of gene expression, Fgf and Shh have distinct influences, Fgf serves to initiate a positive (and titratable) force of gene expression potential. In contrast, Shh acts to release internal barriers that resist the positive force of

gene expression brought about by Fgf. Critically, once this internal barrier is released by Shh, Shh is no longer needed to allow for positive transcriptional flow to occur.

At the level of gene regulation within the progenitor, the key release of transcriptional barriers occurs through activity located at the proximal (promoter) regulatory DNA sequence. Furthermore, the fact that relief of the Gli3 repressor is the primary mode of Shh action at the level of gene regulation is consistent with the model that the key role of Shh is to relieve Gli3 repression at the promoter of genes to allow for active enhancers and positive transcriptional forces to initiate and maintain 5' HoxD expression. Once this release occurs, additional complexity in expression patterns may emerge through competitive dynamics between Hoxd13 and Hoxd11/12 expression.

Materials and Methods:

Limb progenitor culture system

Fertilized white leghorn chicken eggs (Charles River Laboratories) were incubated at 38 °C. Chicken embryos were staged according to the Hamburger and Hamilton stage series. Tissue from the distal anterior portion of forelimb buds were dissected from stage 20-21 embryos (4 days post fertilization). After dissection, buds were incubated in 0.25% trypsin on ice for 5-10 minutes to loosen attached ectodermal tissue. After trypsin incubation, tissue was transferred to 10% chick serum (Life Technologies) in PBS. Loosened ectodermal tissues were removed and remaining mesenchyme was

mechanically dissociated with manual pipetting. Cells were plated in 96 well plates at high density (10^7 cells per milliliter). Baseline media in culture contained 250 ng/ml Wnt3a protein (R&D Systems) in DMEM/F12 media (Life Technologies). Additional ligands added to media include Shh (1 ng/ul unless otherwise noted, N-terminal fragment, R&D Systems), Fgf8 (120 ng/ul unless otherwise noted, Human Fgf8 isoform B, R&D Systems), SU5402 (Tocris), cyclopamine (Tocris), and cycloheximide solution (R&D Systems).

Quantitative PCR

RNA was extracted from progenitor cells using the RNeasy Mini Kit (Qiagen) and then reverse transcribed with Oligo(dT) primers (Invitrogen) using Superscript III (Invitrogen). cDNA was analyzed by quantitative PCR (qPCR) on the Stratagene MX3000P cycler. Each sample was run in triplicate and quantified by comparison to a standard curve (10-fold serial dilutions of a sample or synthesized gene fragment). Primers for Hoxd11, Hoxd12, Hoxd13, Ptch1, Gli1, and Sprouty1 were designed to span intronic regions and also tested for amplification efficiency as well as the capacity to produce a single amplification product.

Chromosomal capture (4C-seq)

4C was carried out in collaboration with the lab of Denis Duboule (with Guillaume Andrey) as described in Andrey et al., 2013 and Noordermeer et al., 2011. Hoxd11 primer set was designed for the chick HoxD11 promoter. PCRs

were multiplexed and sequenced at the EPFL, Switzerland. De-multiplexing, mapping and 4C-analysis were performed through HTSstation (http://htsstation.epfl.ch) according to procedure described in Andrey et al., 2013.

Histone modification analysis (ChIP-seq)

The anterior and posterior portions of 5 dozen chick forelimbs were dissected and fixed in 1% formaldehyde for 15 minutes. Cells were then washed three times in cold PBS and stored at negative 80 degrees Celsius before being processed. Antibodies against H3K27me3 (Millipore, 17-622) and H3K27Ac (Active Motif, 39685) were used as previously described in Schorderet et al., 2011. ChIP-seq library construction and sequencing was carried out by the NRB Biopolymers facility.

References

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

Cohen M., Briscoe J., Blassberg R. Morphogen interpretation: the transcriptional logic of neural tube patterning. Curr Opin Genet Dev. 23(4):423-8 (2013).

Cooper K.L., Hu J.K., ten Berge D., Fernandez-Teran M., Ros M.A., Tabin C.J. Initiation of proximal-distal patterning in the vertebrate limb by signals and growth. Science. 332(6033):1083-6 (2011).

Dessaud E., Yang L.L., Hill K., Cox B., Ulloa F., Ribeiro A., Mynett A., Novitch B.G., Briscoe J. Interpretation of the sonic hedgehog morphogen gradient by a temporal adaptation mechanism. Nature. 450(7170):717-20 (2007).

Drossopoulou G., Lewis K.E., Sanz-Ezquerro J.J., Nikbakht N., McMahon A.P., Hofmann C., Tickle C. A model for anteroposterior patterning of the vertebrate limb based on sequential long- and short-range Shh signalling and Bmp signalling. Development. 127(7):1337-48 (2000).

Harfe B.D., Scherz P.J., Nissim S., Tian H., McMahon A.P., Tabin C.J. Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. Cell. 118(4):517-28 (2004).

Hooper J.E., Scott M.P. Communicating with Hedgehogs. Nat Rev Mol Cell Biol. 6(4):306-17 (2005).

Laufer E., Nelson C.E., Johnson R.L., Morgan B.A., Tabin C. Sonic hedgehog and Fgf-4 act through a signaling cascade and feedback loop to integrate growth and patterning of the developing limb bud. Cell. 79(6):993-1003 (1994).

Litingtung Y., Dahn R.D., Li Y., Fallon J.F., Chiang C. Shh and Gli3 are dispensable for limb skeleton formation but regulate digit number and identity. Nature 418: 979–983 (2002).

Marigo V., Johnson R.L., Vortkamp A., Tabin C.J. Sonic hedgehog differentially regulates expression of GLI and GLI3 during limb development. Dev Biol. 180(1):273-83 (1996a).

Marigo V., Scott M.P., Johnson R.L., Goodrich L.V., Tabin C.J. Conservation in hedgehog signaling: induction of a chicken patched homolog by Sonic hedgehog in the developing limb. Development. 122(4):1225-33 (1996b).

Marigo V., Tabin C.J. Regulation of patched by sonic hedgehog in the developing neural tube. Proc Natl Acad Sci U S A. 93(18):9346-51 (1996).

Montavon T., Duboule D. Chromatin organization and global regulation of Hox gene clusters.
Philos Trans R Soc Lond B Biol Sci. 368(1620):20120367 (2013).

Nelson C.E., Morgan B.A., Burke A.C., Laufer E., DiMambro E., Murtaugh L.C., Gonzales E., Tessarollo L., Parada L.F., Tabin C. Analysis of Hox gene expression in the chick limb bud. Development. 122(5):1449-66 (1996).

Noordermeer D., Leleu M., Splinter E., Rougemont J., De Laat W., Duboule D. The dynamic architecture of Hox gene clusters. Science. 334(6053):222-5 (2011).

Nüsslein-Volhard C., Wieschaus E. Mutations affecting segment number and polarity in Drosophila. Nature. 287 (5785): 795–801 (1980).

Pagan S.M., Ros M.A., Tabin C., Fallon J.F. Surgical removal of limb bud Sonic hedgehog results in posterior skeletal defects. Dev. Biol. 180, 35–40 (1996)

Riddle R.D., Johnson R.L., Laufer E., Tabin C. Sonic hedgehog mediates the polarizing activity of the ZPA. Cell. 31;75(7):1401-16 (1993).

Rogers K.W., Schier A.F. Morphogen gradients: from generation to interpretation. Annu Rev Cell Dev Biol. 27:377-407 (2011).

Ros M.A., Dahn R.D., Fernandez-Teran M., Rashka K., Caruccio N.C., Hasso S.M., Bitgood J.J., Lancman J.J., Fallon J.F. The chick oligozeugodactyly (ozd) mutant lacks sonic hedgehog function in the limb. Development. 130(3):527-37 (2003).

Schorderet P., Duboule D. Structural and functional differences in the long non-coding RNA hotair in mouse and human. PLoS Genet. 7(5):e1002071 (2011).

Wada N., Kawakami Y., Nohno T. Sonic hedgehog signaling during digit pattern duplication after application of recombinant protein and expressing cells. Dev Growth Differ. 41(5):567-74 (1999).

Wolpert L. Positional information and the spatial pattern of cellular differentiation. Journal of Theoretical Biology. Volume 25, Issue 1, (1969).

Wolpert L. et al. Principles of development (3rd ed.). Oxford University Press. ISBN 0-19-927536-X (2007).

Yang Y., Drossopoulou G., Chuang P.T., Duprez D., Marti E., Bumcrot D., Vargesson N., Clarke J., Niswander L., McMahon A., Tickle C. Relationship between dose, distance and time in Sonic Hedgehog-mediated regulation of anteroposterior polarity in the chick limb. Development. 124(21):4393-404 (1997).

Chapter 4:

Potential Further Explorations of the Limb Progenitor

A continuation of the work presented in Chapter 2

Additional experiments would be beneficial to strengthen body of the current set of experiments presented. More thorough genome wide expression analysis of reprogrammed Prx positive cells would be helpful in determining the extent to which Prx positive cells are fully reprogrammed. Additionally, while in vitro functional evidence of the differentiation potential of reprogrammed cells have been presented, in vivo functional evidence still remains the gold standard in the reprogramming field. Both these avenues of improvement are currently being pursued.

If more thorough validation of the current reprogrammed Prx positive cells indicates that they are not fully reprogrammed, additional modification to the current reprogramming protocol may be worth making. First, the current reprogramming protocol involves reprogramming cells for 7 to 8 days before reprogramming. It may be advantageous to reprogram for longer periods of time, as is needed for the reprogramming of other cell types such as pluripotent stem cells and neurons (Takahashi and Yamanaka, 2006, Vierbuchen et al., 2010). Second, the Prx reporter was used in this study because it is an established driver of gene expression in the limb development community. However, it is also known that Prx expression in the embryo extends out of the limb bud into mesenchymal cells in the interlimb regions, making it an imprecise readout of limb progenitor status. Other markers such as the factors found to be critical in the generation of Prx positive cells (Lhx2 or Sall4) and also specific to limb buds as compared to the interlimb mesenchyme may be

more effective in generating a pure population of reprogrammed limb progenitors. Third, it may be the case that the current set of factors found to be minimally sufficient to activate the Prx reporter are not a viable minimal core set of factors for reprogramming limb progenitors. If this is the case, a repeat drop-out screen could be conducted using an alternate marker for limb progenitor identification such as Lhx2 or Sall4.

A continuation of the work presented in Chapter 3

Much of the current work in building models of morphogenesis has focused on the relationship between morphogens (external signals) and the genes that they control in a potentially dose-dependent manner. In the limb, it appears that this relationship may be simpler than expected. Shh, a key influencer of morphology across the antero-posterior axis of the limb controls 5' HoxD, Bmp2, and its direct targets in a simple ON/OFF manner. Moreover, it has been appreciated in the literature that, despite extensive searching, there are no genes that correspond to particular skeletal elementals in the limb (Welten et al., 2011). Given these facts, it appears that a purely molecular view of the emergence of form in the limb is limited.

The work in Chapter 3 focused on the individual (or population level average) limb progenitor and its response to extrinsic signals. Future work may benefit from taking a collective view of the limb progenitor – to focus less on what occurs within the nucleus of each progenitor and rather focus more on the multicellular dynamics that arise when limb progenitors interact with each

other. Given that the gene expression “codes” specified in the early limb by Shh and Fgf are relatively simple, its stands to reason that additional critical complexities are “encoded” at the emergent level of multicellular dynamics.

An integration of the work presented in Chapter 2 and Chapter 3

Early in limb development, progenitor cells occupy the entire limb bud and continue to occupy the distal tip of the bud as limb outgrowth occurs. Two key features of these cells have been appreciated for decades but have remained poorly understood at the level of transcriptional regulation: (1) these cells continue to remain in an undifferentiated progenitor state as long as they lie under the influence of signals provided by the overlying ectoderm (ten Berge et al., 2008) and (2) while these cells remain undifferentiated, they successively pass through distinct states of patterning potential (Tabin and Wolpert, 2007). These features cannot be addressed without having, as a starting point, an understanding of how the limb progenitor state itself is established.

The results in Chapter 2 suggest that that Lhx2 and Sall4 play a central role in this, providing a starting point to examine progenitor cell maintenance and the dynamic changes in patterning that take place within the progenitors over time. There are multiple lines of evidence that suggest that Lhx2 and Sall4 play critical roles in regulating the multipotent, undifferentiated status of limb progenitors. First, they are co-expressed throughout the early limb bud and expression is maintained in the undifferentiated distal tip of the limb as outgrowth occurs (Bell et al., 2004, Darnell et al., 2007).

Second, the reprogramming experiments of Chapter 2 show that these two genes appear to be essential for reverting adult fibroblasts to a limb progenitor-like state. Finally, both genes have been implicated in maintaining 'stemness' properties in other cells types. Sall4 is enriched across many embryonic stem and progenitor cells (hematopoetic progenitors) as well as adult stem/stem-like cells (Tatetsu et al., 2016). Similarly, Lhx2 has been shown to maintain the undifferentiated properties of hair follicle progenitors (Rhee et al., 2006).

In addition to their critical role in maintaining the limb progenitor state, it may be likely that Lhx2 and Sall4 also concurrently play crucial roles in regulating the patterning potential of limb progenitors. The inactivation of Lhx2 (in combination with functionally redundant genes Lhx9 and Lmx1b) in the limb results in severe patterning defects across the proximodistal (PD), anteroposterior (AP), and dorsoventral (DV) axes (Tzchori et al., 2009). There is less established evidence for Sall4 and its involvement in patterning of the limb. However, its homologs, Sall1 and Sall3 influence the expression of Hox genes, which are key transcriptional read-outs of patterning state in the developing limb (Kawakami et al., 2009).

While Lhx2 and Sall4 have not been mechanistically studied in the context of the limb, studies in other systems have shed light on their specific functions during transcriptional regulation. In ES cells, Sall4 was shown to bind the mouse genome more pervasively than either Oct4 or Nanog (Yang et al., 2008). Additionally, Sall4 has been implicated in both activating genes

necessary for the stem cell state and repressing genes that specify other competing lineages (trophectoderm) (Zhang et al., 2006). Thus, Sall4 has a pervasive and dual role in regulation transcription in ES cells. Based on these facts, it is possible to hypothesize that Sall4's dual role as an activator and repressor can be brought to bear during the maintenance of the progenitor state as well as the emergence of pattern across the anteroposterior and dorsoventral axes of the early limb bud.

Recently, the transcriptional role of Lhx2 has been interrogated in hair follicle stem cells through profiling and ChIP-seq (Folgueras et al., 2013). In those cells, Lhx2 was found to be a positive regulator of a large number of hair follicle signature stem cell genes as well as genes that code for components of cytoskeletal and adhesion networks. Perhaps most importantly, the authors found that Lhx2 binding in hair follicle stem cells drastically differed from its binding in embryonic skin tissue. Thus, while the critical biochemical features of Lhx2 and Sall4 may be similarly deployed across different cellular contexts, their functional roles are likely to be divergent and therefore require study in specific developmental contexts. Ultimately, while Chapter 2 and 3 addressed limb progenitor identity and limb progenitor patterning as separate phenomena, it may be very likely that the genes critical for maintaining limb progenitor identity may also be the very ones that poise genes such as the 5' HoxD genes and Bmp2 to respond to external signals such as Shh.

References:

Bell G.W., Yatskievych T.A., Antin P.B. GEISHA, a high throughput whole mount in situ hybridization screen in chick embryos. Devel. Dynamics 229: 677-687 (2004).

Darnell D.K., Kaur S., Stanislaw S., Davey S., Konieczka J.H., Yatskievych T.A., Antin P.B. GEISHA: An In situ hybridization gene expression resource for the chicken embryo. Cytogenet. Genome Res. 117:30-35 (2007).

Folgueras A.R., Guo X., Pasolli H.A., Stokes N., Polak L., Zheng D., Fuchs E. Architectural niche organization by LHX2 is linked to hair follicle stem cell function. Cell Stem Cell. 13(3):314-27 (2013).

Kawakami Y., Uchiyama Y., Rodriguez Esteban C., Inenaga T., Koyano-Nakagawa N., Kawakami H., Marti M., Kmita M., Monaghan-Nichols P., Nishinakamura R., Izpisua Belmonte J.C. Sall genes regulate region-specific morphogenesis in the mouse limb by modulating Hox activities. Development. 136(4):585-94 (2009).

Rhee H., Polak L., Fuchs E. Lhx2 maintains stem cell character in hair follicles. Science. 312(5782):1946-9 (2006).

Tabin C., Wolpert L. Rethinking the proximodistal axis of the vertebrate limb in the molecular era. Genes Dev. 21(12):1433-42 (2007).

Takahashi K., Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 126(4):663-76 (2006).

Tatetsu H., Kong N.R., Chong G., Amabile G., Tenen D.G., Chai L. SALL4, the missing link between stem cells, development and cancer. Gene. S0378-1119(16)30072-5 (2016).

ten Berge D., Brugmann S.A., Helms J.A., Nusse R. Wnt and FGF signals interact to coordinate growth with cell fate specification during limb development. Development. 135(19):3247-57 (2008).

Tzchori I., Day T.F., Carolan P.J., Zhao Y., Wassif C.A., Li L., Lewandoski M., Gorivodsky M., Love P.E., Porter F.D., Westphal H., Yang Y. LIM homeobox transcription factors integrate signaling events that control three-dimensional limb patterning and growth. Development. 136(8):1375-85 (2009).

Vierbuchen T., Ostermeier A., Pang Z.P., Kokubu Y., Südhof T.C., Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. Nature. 463(7284):1035-41 (2010).

Welten M., Pavlovska G., Chen Y., Teruoka Y., Fisher M., Bangs F., Towers M., Tickle C. 3D expression patterns of cell cycle genes in the developing chick wing and comparison with expression patterns of genes implicated in digit specification. Dev Dyn. 240(5):1278-88 (2011).

Yang J., Chai L., Fowles T.C., Alipio Z., Xu D., Fink L.M., Ward D.C., Ma Y. Genome-wide analysis reveals Sall4 to be a major regulator of pluripotency in murine-embryonic stem cells. Proc Natl Acad Sci U S A. 105(50):19756-61 (2008).

Zhang J., Tam W.L., Tong G.Q., Wu Q., Chan H.Y., Soh B.S., Lou Y., Yang J., Ma Y., Chai L., Ng H.H., Lufkin T., Robson P., Lim B. Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nat Cell Biol. 8(10):1114-23 (2006).