Causes and Consequences of Lung Loss in Salamanders

Causes and Consequences of Lung Loss in Salamanders

A dissertation presented

by

Zachary Robert Lewis

to

The Department of Organismic and Evolutionary Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biology

Harvard University

Cambridge, Massachusetts

December 2015

Dissertation Advisor: Professor James Hanken Zachary Robert Lewis

Causes and Consequences of Lung Loss in Salamanders

Abstract

Lungs were once thought to be a universal feature of tetrapods and essential for vertebrate life on land. This view changed in the late 19th century with the discovery of several salamander species that lack lungs. Since these species are descendants of lunged ancestors, the absence of lungs must represent an instance of evolutionary loss. Further study has revealed several independent losses of lungs across amphibians, including at the base of the salamander family Plethodontidae. Plethodontids comprise over two thirds of all living salamander species, yet many features of lungless salamander biology are unknown. My dissertation investigates the evolution and development of lung loss, including the genetic basis for lunglessness and the consequences of lunglessness for respiration and the circulatory system.

I determined that plethodontid salamanders are not entirely lungless; lungs actually begin to form in the embryo. This includes lung specification, the formation of a lung primordium, and conserved expression of genetic markers of lung differentiation. However, the lung primordium subsequently regresses by apoptosis, yielding adults with no trace of a vestigial lung. Transcriptome sequencing of the lung primordia of lunged and lungless salamanders suggests a role for increased Tgfβ signaling in lung regression. I established

that Tgfβ represses lung development in salamanders with lungs, providing support for its role in lung loss.

Plethodontid salamanders perform gas exchange through their skin and lining of the mouth (extrapulmonary respiration). I discovered a novel gene in salamanders that is potentially neofunctionalized for extrapulmonary respiration in plethodontids. The lung-specific gene encoding surfactant-associated protein C (SPC) is duplicated in salamanders. Both paralogs of this gene are expressed in the lung of lunged salamanders, representing the ancestral expression pattern of SPC in tetrapods. In contrast, lungless salamanders express a paralog of SPC in extrapulmonary sites of gas exchange. These sites include the skin during the aquatic larval stage and the lining of the mouth in terrestrial adults. I propose that extrapulmonary expression of this paralog in salamanders reduces the thickness of the mucus layer that covers the respiratory surfaces and aids gas exchange.

The lungs function as part of an integrated cardiopulmonary system. In animals with lungs, the atrial septum helps to separate oxygenated and deoxygenated blood in the heart. I characterized cardiac anatomy within a broad sample of lunged and lungless salamanders. I found that independent lineages of lungless salamanders convergently evolved a reduced and non-functional atrial septum, resulting in blood flow between the two atrial chambers. In mammals, formation of the atrial septum is dependent on morphogens secreted from the developing lungs. I provide evidence that atrial septum reduction in plethodontid salamanders is a direct consequence of loss of these signaling interactions due to lung regression. Developmental interaction between the heart and lungs may mediate the coordinated evolution of the cardiopulmonary system, ensuring that the atrial septum develops in the presence of lungs but does not fully form in lungless species, where it would be disadvantageous.

Contents

List of Figures

List of Tables

Acknowledgments

I have received an extraordinary amount of help and support during the course of my PhD research. These interactions have enriched my dissertation and helped to shape my research, writing and relationships.

I first would like to acknowledge the support and help of my advisor, Jim Hanken. Jim has offered advocacy, advice, help and feedback from the very beginning of my PhD research up until the very end. His encouragement has both pushed this research forward and helped to make it worthwhile.

My dissertation committee has been critical in this process. I would like to thank Arkhat Abzhanov, Cassandra Extavour, Jay Rajagopal and Cliff Tabin for their extremely helpful criticism, input and enthusiasm.

The Hanken Laboratory has been a source of knowledge, insight, friendship and entertainment. I want to especially thank Liz Sefton, Hillary Maddin and Nadine Piekarski, who have been crucial sources of practical and experimental advice. Breda Zimkus, Mara Laslo, Yunke Wu and Brent Hawkins have also been immensely helpful in terms of providing comments on written work and oral presentations and lending experimental help. I have had the pleasure of working with Jorge Dorantes, a post-baccalaureate volunteer in the lab. Jorge contributed

to cloning and characterizing the expression of SPC and SPC-like in salamanders, as well as an large amount of collecting and fieldwork.

Many people have accompanied me on collecting trips and this support has been key to obtaining sufficient numbers of specimens. These individuals include: Melissa Aja, Bob Cooke, Jorge Dorantes, Carolyn Eng, Jim Hanken, Brent Hawkins, Alyssa Koomas, Laura Lagomarsino, Mara Laslo, Hillary Maddin, Joe Martinez, Matt McCarroll, Nadine Piekarski, Apratim Sahay, Arpiar Saunders, Liz Sefton, Steve Tilley, Sandra Triepel, David Woolf and Yunke Wu.

Several people have provided technical advice and support for this work. Jennifer Couget helped with sample preparation for transcriptome sequencing. Bob Freeman, Chris Laumer, Shane Campbell-Staton and Lauren O'Connell provided extraordinary assistance for bioinformatic analyses. Carolyn Marks assisted with transmission electron microscopy processing and imaging. Hillary Maddin taught me μ-CT imaging and visualization. Chris Laumer taught me how to perform histological sectioning and staining.

Thank you to a fantastic herpetology department including José Rosado, Joe Martinez, Tsuyoshi Takahashi and Jon Woodward. Thank you to Hanken Lab managers past and present including Anne Everly, Caroline DeVane and Matt Gage. I want to acknowledge the staff of the Ernst Mayr Library including Mary Sears, Ronnie Broadfoot and Dorothy Barr for research assistance. Several agencies provided access to living collections for this research including Massachusetts Audubon Society, Harvard Forest, Cape Cod National Seashore, the Massachusetts Department of Conservation and Recreation, and the Arnold Arboretum. I received funding for this research from the National Science Foundation Graduate Research Fellowship Program, the William F. Milton Fund (Harvard Medical School), Sigma Xi, and

the Society for the Study of Amphibians and Reptiles. Support for fieldwork was provided by the Barbour Award, the Robert G. Goelet Award and the Kenneth Miyata Award, all administered by the Museum of Comparative Zoology, Harvard University.

My family has been immeasurably supportive throughout my PhD research. Thanks to my parents and siblings for listening to me and encouraging me.

Carolyn Eng played an outsized role in the entirety of this project and continues to play one in my life. It was through our interactions in the lab and in the field that we got to know each other. Carolyn encouraged me to redouble my collection efforts after two summers of discouraging returns. Her encouragement and support did not end there. She accompanied me on many collection trips across Massachusetts. Our trips to the field together to look for salamanders were some of my favorite times in graduate school. Her careful reading and intelligent comments have greatly improved this dissertation. Her love and support made this research possible, gratifying and fun.

Chapter 1

Introduction

1.1 The origins of lunglessness

Aristotle wrote, "life is bound up with the passage of the breath outwards and inwards" (Parva Naturalia, 480a20-23). As much as the cycling of air and a beating heart define life, death is the cessation of both. This seems especially true for the tetrapods, where aerial respiration through the lungs has allowed the conquest of land. However, many tetrapods live successful lives without breathing in the traditional sense. These physiological oddities are the lungless amphibians, in which respiration occurs over the skin and the lining of the mouth. The implications of lunglessness are broad: How can lungs, which develop within an integrated embryonic system, become fully lost? How do lungless animals compensate for loss of this respiratory surface? How might studying this phenomenon aid our understanding of human disease and birth defects? In the course of my dissertation research I set out to answer these questions.

Lung loss was discovered independently by two researchers: Harris H. Wilder and Lorenzo

Camerano (Wilder, 1894, 1896). Despite the predictable detriments of lung loss, it has evolved several times in different groups of tetrapods within the three modern orders of Amphibia (Fig. 1.1). These lungless species include one species of caecilian (Nussbaum and Wilkinson, 1995), one species of frog (Bickford et al., 2008), one genus of hynobiid salamanders, and the entirety of the largest family of salamanders, Plethodontidae (Fig. 1.1; Noble, 1925). The convergent evolution of lung loss across many taxa and the successful adaptive radiation of plethodontids suggests that lung loss has some selective advantage. Determining precisely what this advantage is has been difficult; there are competing adaptive hypotheses to explain why lung loss may have occurred.

The first-proposed and most broadly accepted hypothesis for lung loss is that proto-plethodontid larvae lived in fast-moving streams where lung reduction (and eventual loss) would have been selected for based on buoyancy reduction, thereby preventing these salamanders from being washed downstream (Beachy and Bruce, 1992; Bruce et al., 1994; Dunn, 1926; Wilder and Dunn, 1920). This so-called rheotropic hypothesis depends on several assumptions, including the assumed ancestral distribution of proto-plethodontids in cold torrential brooks of the Appalachian highlands, the selective advantage of lung loss in stream dwelling, and the fact that convergently derived lungless (and most lung-reduced) salamanders can be found in cold brooks. This hypothesis has been criticized because of varying interpretations of the topology of the Appalachians during the Late Mesozoic, when plethodontids were thought to originate. Some researchers believe that the Appalachians were flat at that time, and that plethodontids instead lost lungs as a secondary effect of reduction in the width of the head and reduced efficiency of ventilating the lungs through a buccal pumping system (Ruben and Boucot, 1989; Ruben et al., 1993). Ruben and Boucot (1989)

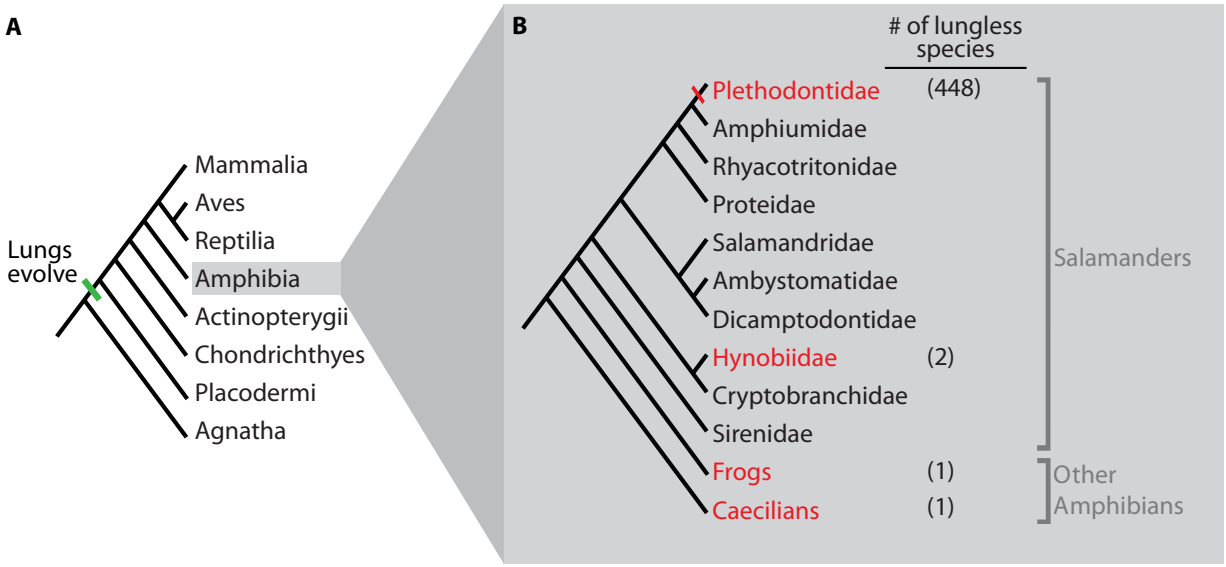


Figure 1.1: Instances of lung loss in amphibians. (A) Lungs may have evolved in placoderms and were subsequently lost in several lineages (including Chondrichthyes, members of Actinopterygii and certain amphibians). Alternatively, lungs evolved in bony fishes and have been lost in members of Actinopterygii and certain amphibians. Several lunged sarcopterygian groups, including Actinistia (coelocanths) and Dipnomorpha (lungfish), are omitted from the phylogeny for simplicity. (B) Lung loss in amphibians has evolved at least four times independently. Red text indicates families or orders containing at least one lungless species. The plethodontid salamanders all lack lungs as adults, and it is most parsimonious to assume that this loss occurred once at the base of the lineage. Of over 680 described salamander species, over 448 are within Plethodontidae. Tree topology from Zhang and Wake (2009).

cite divergent functional (e.g., feeding) demands and increased terrestriality as responsible for reduced head width, resulting in inefficient buccal pumping. Reagan and Verrell (1991) propose that a shift towards terrestrial mating (due to decreased energy costs) led to selection for terrestrial feeding systems, causing pulmonary ventilation to become inefficient. Terrestrial feeding involves tongue projection using the hyobranchial musculature and skeleton. Specialization of the hyobranchial apparatus for tongue projection is thought to be incompatible with a strong buccal respiratory pump (Lombard and Wake, 1986).

The hypothetical biogeographic origin of plethodontids was predicted based on extant taxonomic diversity in the southern Appalachians. However, recent phylogenetic and paleontological evidence supports a younger origin of plethodontids in western North America (Shen et al., 2016). If a western origin is correct, then this should prompt a re-examination of the hypotheses regarding the adaptive significance of lung loss.

When taking into account the lungless caecilian *Atretochoana eiselti*, it becomes likely that there may be no single adaptive hypothesis to explain lung loss. The elongate aquatic *A. eiselti* has a large buccal cavity (Nussbaum and Wilkinson 1995). The presence of a (presumably) strong buccal force pump in aquatic *A. eiselti* does not support the hypotheses that terrestriality preceded lung loss (Wilkinson and Nussbaum 1997). *Atetochoana eiselti* is not found in cold upland streams, arguing against the rheotropic hypothesis (Hoogmoed et al., 2011). However, reduced buoyancy may still be advantageous in the native environment of *A. eiselti* (Wilkinson et al., 2014). Bickford et al. (2008) discovered the first known lungless frog, *Barbourula kalimantanensis*, in fast-flowing mountain streams of Borneo—interpreted as support for the rheotropic hypothesis of lung loss. The rheotropic hypothesis is also supported by the ecology of certain amphibians with reduced or rudimentary lungs, such as salamanders within *Rhyacotriton* and the tailed frog *Ascaphus truei*. Both of these species live in fast-flowing streams.

With some exceptions (including *Atretochoana eiselti*), lunglessness has been documented in small species with a high surface-area-to-volume ratio (Bickford et al. 2008; Wake and Donnelly 2010). The efficient cutaneous and buccal respiration found in small amphibians may enable the evolution of lung loss. However, due to the few consistently shared ecological and morphological traits uniting lungless amphibians, and uncertainty as to the types of

environments in which plethodontids evolved, it is possible that diverse selective pressures brought about independent instances of lung loss. The difficult nature of reconstructing the precise ancestral habitat, behavior and morphology of proto-plethodontids makes directly determining why lung loss has occurred in this group particularly challenging. While the adaptive explanation for lunglessness has essentially reached an impasse, it is still an open question as to how lung loss occurred. In addition, the morphological and physiological consequences of lunglessness have not received much attention.

In Chapter 2, I investigate the developmental basis for lung loss. I find that lungless salamanders are not entirely lungless and instead begin to develop lungs during embryogenesis. I demonstrate that the pulmonary rudiment is homologous to the lung based on morphological and molecular criteria. Lunglessness in adults occurs by the apoptotic regression of the lung rudiment during late embryogenesis.

In Chapter 3, I use transcriptome sequencing and experimental approaches to identify the genes responsible for lung loss. I find a number of known regulators of lung development differentially expressed in the lung rudiment of the lungless species *Plethodon cinereus* compared to the lunged salamander *Ambystoma mexicanum*. Among these, I posit that increased Tgfβ signaling is responsible for lung loss through the downregulation of the critical lung gene Nkx2.1.

1.2 Amphibian lungs

Scientists frequently marvel at the human lung with its 10 million highly stereotyped branches, which terminate in alveoli and surface area half the size of a tennis court (Metzger et al.,

2008). Amphibian lungs, typically described as simple and saccular, have few admirers, but they do vary remarkably in their degree of vascularization and internal septation (Maina, 2002). Generally speaking, pulmonary vascularization is inversely related to the degree of cutaneous respiration. All amphibians respire through their skin to some extent, but salamanders typically display higher rates of cutaneous gas exchange than frogs. Therefore, frogs tend to have more vascularized and subdivided lungs while most salamander lungs have fewer capillary meshes per unit area. Salamander lungs are unicameral and typically very thin (Miller, 1893). The pulmonary vasculature runs through thickened septa, but the walls of the large intervening airspaces are only several cell diameters thick (Fig. 1.2).

All lungs, and all respiratory organs (e.g., skin, gills), share a similar structure: a thin tissue layer, which separates the external respiratory media from the blood and allows gas exchange through a partial pressure gradient between the blood and respiratory media (Weibel, 1984). A thin tissue layer, while beneficial for gas exchange, presents several problems. It is susceptible to pathogens, dehydration and injury. Mammalian lungs have multiple cell types in order to clear particles and aid tissue compliance in the alveoli and airways. Type I and type II alveolar epithelial cells (AT1 and AT2) comprise most of the cells within the mammalian alveoli. AT1 cells are extremely squamous, with thin cytoplasmic projections and few organelles; their thin morphology enables efficient gas exchange to blood vessels. AT2 cells are cuboidal and play a secretory role, such as the secretion of pulmonary surfactant (Weibel, 1984). The lungs of anamniotes, like amphibians and fish, have multiple cell types, including ciliated cells, but the functions of the AT1 and AT2 cells are likely combined into a single morphologically intermediate cell type (Hightower et al., 1975; Meban, 1979; Pattle et al., 1977; Podkowa and Goniakowska-Witalińska, 1998; Power et al., 1999).

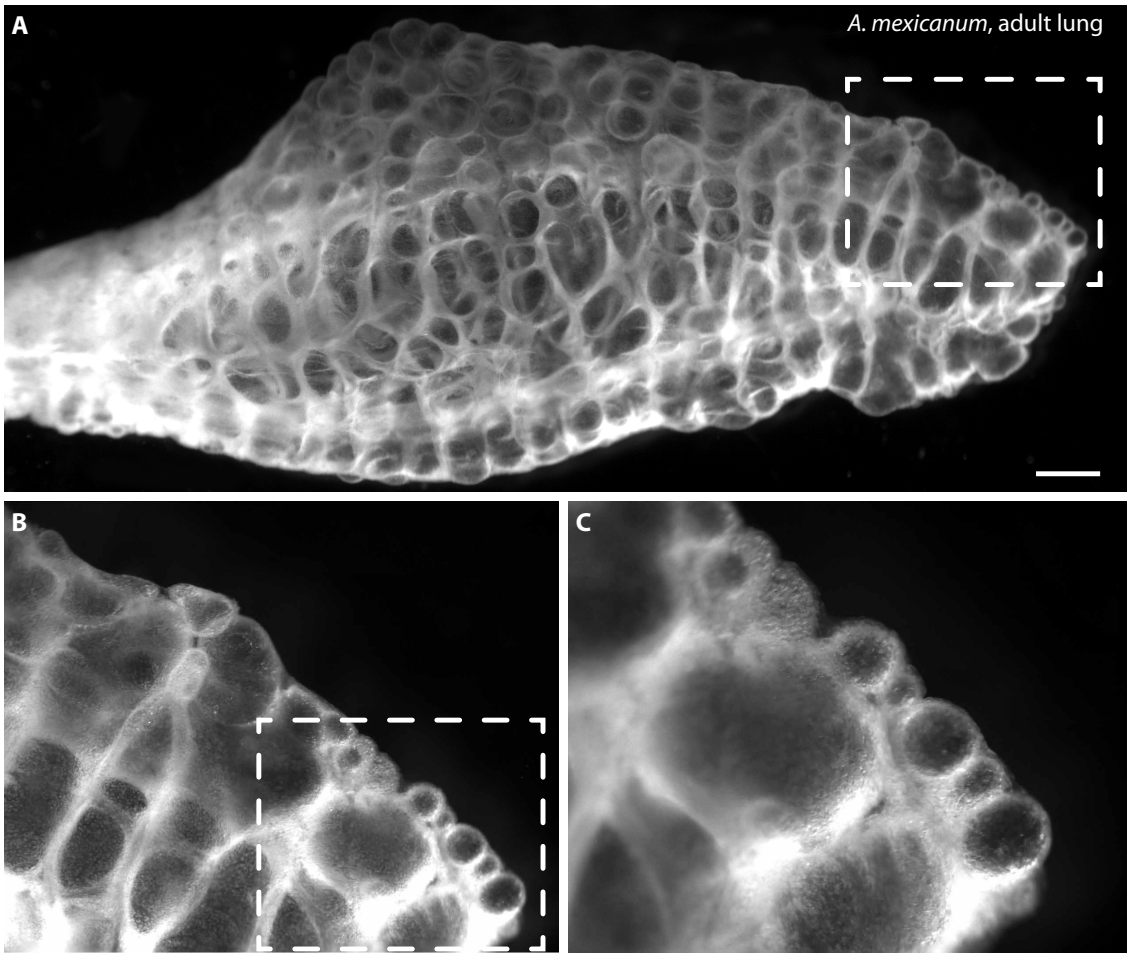


Figure 1.2: The unicameral lungs of *Ambystoma mexicanum*. (A) The adult lung is saccular. Primary alveolar septa are externally visible as a white mesh. Intervening epithelium is thin and nearly transparent. Anterior is to the left. (B,C) Enlargements of boxed regions within (A) and (B), respectively. Scale bar: 1 mm.

Pulmonary surfactant is a mixture of phospholipids, cholesterol and proteins and is found in lungs or lung homologs across all lunged vertebrates (Daniels et al., 1998). Proteins make up about 10% of pulmonary surfactant volume. These proteins include the four pulmonary surfactant-associated proteins (A, B, C, and D). SPA and SPD have mainly innate immune functions, while the lung-specific hydrophobic surfactant proteins SPB and SPC help to regulate the adsorption, spreading, production and re-uptake of surfactant phospholipids

(Whitsett and Weaver, 2002). Therefore, SPC and SPB play a role in reducing surface tension of the mucous layer of the lung and increasing pulmonary compliance. Amphibian lungs are both intrinsically compliant and do not possess microscopic airspaces prone to collapse. Therefore, pulmonary surfactant presence in amphibians and other animals with saccular lungs has been explained as functioning to prevent the thin lungs from self-adhering (Daniels et al., 1995, 1998). I put forward an alternative hypothesis in Chapter 4, focusing on SPC in particular, but applicable to surfactant in general, that SPC has a direct role in aiding gas exchange.

In Chapter 4, I describe the discovery of a paralog of SPC, which is found so far only in salamanders. This paralog, SPC-like, maintains a conservative expression pattern solely in the lungs in a lunged species of salamander. However, SPC-like expression is surprisingly divergent within a species of lungless salamander. Instead of lung-specific expression, SPC-like is expressed in the skin during embryogenesis and larval life, and in the mouth during and after metamorphosis. These expression patterns correlate with the respiratory sites preferentially employed during each life history stage, and I suggest that expression of SPC-like aids gas exchange across cutaneous and oral epithelia. I identify the potential secretion of surfactant vesicles from larval lungless salamander skin by transmission electron microscopy.

1.3 Lung development

The gene expression patterns and molecular interactions that underlie lung development are remarkably conserved across vertebrates, ranging from fish (Cass et al., 2013; Zheng et al., 2011), to amphibians (Rankin et al., 2015), to mammals (Morrisey and Hogan, 2010). There

are no published studies on the gene expression patterns within developing salamander lungs, but mature salamander lungs express many key pulmonary proteins, suggesting conservation of lung development programs within this order (Miller et al., 2001). Pulmonary development commences in amphibians just prior to the branching of the external gills (Greil, 1905; Mekeel, 1930; Tsuda, 1924).

The morphogenetic process of early lung and trachea development is remarkably similar between mammals and amphibians (Hogan, 1999; Rankin et al., 2015; Spooner and Wessells, 1970). There is some debate about the first morphological indications of lung development and whether the pulmonary anlage is paired or medial (Mekeel, 1930). Modern authors tend to agree that the first indication of lung development is the formation of a medial indentation in the ventral pharynx, the laryngotracheal groove (Rankin et al., 2015), which becomes a distinct diverticulum, the laryngotracheal tube. The laryngotracheal tube has slight lateral swellings at its distal (caudal) extent, which correspond to the lung buds. Subsequently, the esophagotracheal ridges, corresponding to the two lateral surfaces of the pharynx, begin to join and septate the esophagus (dorsal) from the trachea and lung buds (ventral) (Que et al., 2006). Morphogenesis of the lungs diverges between amphibians and mammals following lung bud outgrowth. Mammalian lung development involves extensive branching (Metzger et al., 2008), while amphibian lungs develop by proximal-to-distal elongation followed by the formation of internal septa (Rose and James, 2013). Branching morphogenesis of the lung may be common to many amniotes (Flint, 1906; Maina, 2011; Schachner et al., 2013).

The influence of the pulmonary mesoderm on the development of the lungs has been known since the 1930s (Rudnick, 1933). Nearly all aspects of early lung development occur through coordinated crosstalk between the endoderm-derived epithelium and the splanchnic

mesoderm-derived pulmonary mesenchyme (McCulley et al., 2015). These signaling interactions are extensively reviewed in Chapter 3. As discussed below, recent research highlights reciprocal interactions between endoderm and mesoderm not only in patterning the morphology of the lung epithelium, but in directing the formation of portions of the heart and the pulmonary vasculature (Hoffmann et al., 2014; Peng et al., 2013).

1.4 The developmental evolution of the cardiopulmonary system

Most textbooks explain evolution of the lungs as an adaptation for aerial respiration in hypoxic environments. However, based on the phylogenetic distribution of lungs and lung function in extant fishes, this interpretation is probably incorrect. Vertebrate ancestors and early vertebrates possessed avascular hearts, which relied solely on blood within their lumen for oxygen. The ancestral function of the lungs in early-branching vertebrates was likely to provide oxygen to the heart, not to the body (Farmer, 1999). According to Aristotle, "the lungs be no part for themselves, but for the heart; and therefore it [is] superfluous for those creatures to have lungs that have no hearts; but nature is never wanting in things necessary, nor abounds in superfluities" (Problems, 859a). Although Aristotle considered the lungs necessary for cooling the heart, not for providing oxygen, his perceptive hypothesis that lungs function to support the heart is confirmed by recent paleontological and physiological evidence (Burggren, 1988; Farmer, 1999). Indeed, the heart and lungs are integrated physiologically, evolutionarily and developmentally.

Mechanisms for separating oxygenated and deoxygenated blood evolved in sarcopterygian fishes, such as lungfish and tetrapods. Septation of the cardiac chambers represents one such mechanism to maintain a dual circulatory pattern (Fig. 1.3). Lungfish and amphibians accomplish separation of oxygenated and deoxygenated blood with the atrial septum in the inflow tract and the spiral septum in the conus arteriosus (outflow tract) (Davies and Francis, 1941; Icardo et al., 2005). In amphibians, blood from the lungs enters the left atrium through the pulmonary veins. This oxygenated blood then passes into the ventricle, where it is pumped through the outflow tract and to the head, body and skin. Blood in the systemic circulation is a mix of oxygenated and deoxygenated blood, because some of it travels through respiratory cutaneous and buccopharyngeal capillaries. Relatively deoxygenated blood returns to the sinus venosus and then empties into the right atrium before being pumped into the ventricle and then out of the heart to the lungs (Fig. 1.3). The problem that early researchers observed in amphibians is that while the atrium is divided by the atrial septum, the ventricle has no obvious septation. While amphibians are susceptible to mixing of bloodstreams within the ventricle, the blood streams are kept mostly separate (Haberich, 1965; Johansen, 1962; Noble, 1925). The precise mechanism by which blood remains separate through the undivided ventricle is undetermined, but it likely pertains to the extensive trabecular network found in the ventricle (Johansen and Hanson, 1968)

Separation of systemic and pulmonary circuits by cardiac septation enables reliance on lungs as the principal sites of gas exchange, dominating the gills and the skin as organs of respiration. Therefore, cardiac septation enabled the conquest of land. Recent research suggests that lung development is induced by signals from the heart; subsequent signals emanating from the lungs pattern the development of the atrial septum (Goddeeris et al.,

2008; Hoffmann et al., 2009, 2014; Serls et al., 2005). Specifically, early Fibroblast growth factor (Fgf) signaling from the cardiac mesoderm is important for lung specification (Serls et al., 2005). After lung formation, Sonic hedgehog (Shh) signaling from the lungs via the transcription factor Gli1, combined with mesenchymal Tbx5 expression, helps to specify a population of migratory multipotent mesenchymal cells. These progenitor cells go on to form the atrial septum, portions of the pulmonary mesenchyme, and the pulmonary vascular plexus (Hoffmann et al., 2014; Peng et al., 2013).

Molecular interactions between the heart and lungs provide a developmental genetic paradigm with which to understand the evolution of atrial septation. The developmental integration of the heart and lungs results in a co-evolutionary relationship, whereby the developing lung dictates the formation of structures (such as the atrial septum and pulmonary veins) necessary for reliance on pulmonary respiration.

In Chapter 5, I explore the developmental and evolutionary integration of the heart and lungs in independent lineages of lungless salamanders. Plethodontid salamanders have long been known to have either absent or partial atrial septa (Bruner, 1900; Hopkins, 1896; Putnam and Kelly, 1978). However, these anatomical descriptions conflict with respect to septum presence and configuration of the inflow tract. In contrast to past studies, I examine a broad comparative data set using three-dimensional computed tomographic methods. I determine that plethodontid salamanders and the convergently lungless species *Onychodactylus japonicus* have convergently evolved a reduced atrial septum compared to lunged ancestors.

I find that Gli1 and Tbx5 expression are significantly downregulated in plethodontid salamanders relative to lunged salamanders. Given their known role in patterning the atrial septum, Tbx5 and Gli1 downregulation is a potential mechanism by which the atrial septum

is reduced. Human birth defects resulting in single source of mixed blood to the right atrium are lethal unless treated due to pressure differentials between the right and left atria. I argue that confluence between the left and right atria in lungless salamanders is not just a consequence of lung loss, but also an indispensible feature of hearts with a single source of mixed blood because of the adverse effects of pressure differentials between the right and left atria.

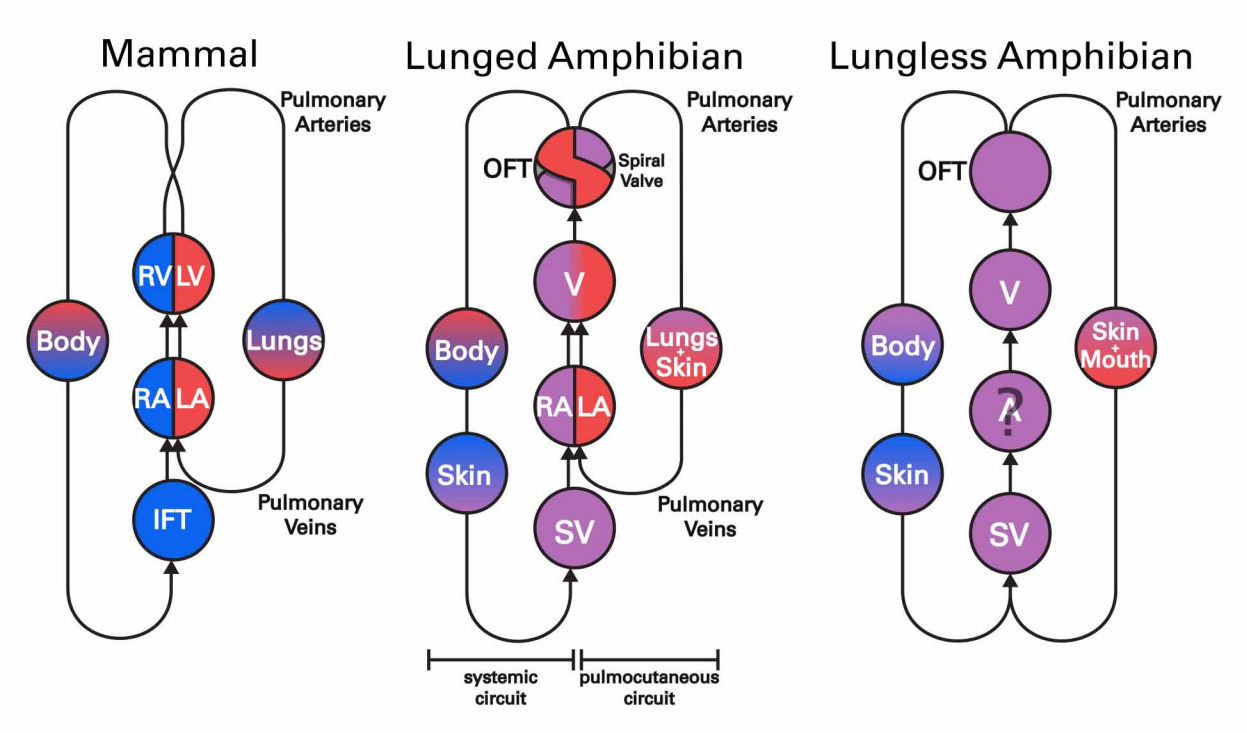


Figure 1.3: Configuration of the circulatory system across mammals and amphibians. This schematic compares the dual circulatory pathways in mammals and lunged amphibians to the configuration in lungless amphibians. The atrium of lungless amphibians is marked by a question mark, because the morphology is unresolved. In some lungless salamanders, the pulmonary arch is maintained and provides blood to a portion of the mouth and pharynx and a small portion of skin (McMullen, 1938).

1.5 Summary

There are surprisingly few known instances of the development of a transient internal organ rudiment. This may be due to the fact that careful ontogenetic characterization is necessary to recognize the development and regression of an internal organ, while it is straightforward to recognize externally visible rudiments such as the human tail or persistent vestigial structures such as the human appendix. Alternatively, internal organ loss may be rare due to the tight integration of internal organ formation and the possibility for outright loss to impact the function of organ systems that are developmentally or functionally interlinked.

The study of lung loss not only offers a view towards the mechanisms of organ loss, but also may inform our understanding of lung development across vertebrates. Little is known about lung development in salamanders. Additionally, a number of characterized human birth defects affect the formation of the lungs, trachea and esophagus, and lungless salamanders may offer insight into the etiology of such malformations. Moreover, the effects of lung loss on the circulatory system and on respiratory physiology are not well known. Here I use a multifaceted approach combining morphology, transcriptomics and developmental genetics in an attempt to describe the causes and consequences of lung loss.

1.6 References

Beachy, C.K., and Bruce, R.C. (1992). Lunglessness in plethodontid salamanders is consistent with the hypothesis of a mountain stream origin: a response to Ruben and Boucot. Am. Nat. 139, 839–847.

Bickford, D., Iskandar, D., and Barlian, A. (2008). A lungless frog discovered on Borneo. Curr. Biol. 18, 374–375.

Bruce, R.C., Beachy, C.K., Lenzo, P.G., Pronych, S.P., and Wassersug, R.J. (1994). Effects of lung reduction on rheotactic performance in amphibian larvae. J. Exp. Zool. 268, 377–380.

Bruner, H.L. (1900). On the heart of lungless salamanders. J. Morphol. 16, 323–336.

Burggren, W.W. (1988). Cardiac design in lower vertebrates: what can phylogeny reveal about ontogeny? Experientia 44, 919–930.

Cass, A.N., Servetnick, M.D., and McCune, A.R. (2013). Expression of a lung developmental cassette in the adult and developing zebrafish swimbladder. Evol. Dev. 132, 119–132.

Daniels, C.B., Orgeig, S., and Smits, A.W. (1995). The evolution of the vertebrate pulmonary surfactant system. Physiol. Zool. 68, 539–566.

Daniels, C.B., Lopatko, O. V, and Orgeig, S. (1998). Evolution of surface activity related functions of vertebrate pulmonary surfactant. Clin. Exp. Pharmacol. Physiol. 25, 716–721.

Davies, F., and Francis, E. (1941). The heart of the salamander (*Salamandra salamandra*, L.), with special reference to the conducting (connecting) system and its bearing on the phylogeny of the conducting systems of mammalian and avian hearts. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 231, 99–130.

Dunn, E. (1926). The Salamanders of the Family Plethodontidae (Northampton: Smith College).

Farmer, C. (1999). Evolution of the vertebrate cardio-pulmonary system. Annu. Rev. Physiol. 61, 573–592.

Flint, J.M. (1906). The development of the lungs. Am. J. Anat. 6, 1–137.

Goddeeris, M.M., Rho, S., Petiet, A., Davenport, C.L., Johnson, G.A., Meyers, E.N., and Klingensmith, J. (2008). Intracardiac septation requires hedgehog-dependent cellular contributions from outside the heart. Development 135, 1887–1895.

Greil, A. (1905). Bemerkungen zur frage nach dem ursprünge der lungen. Anat. Anz. 29, 447–506.

Haberich, F. (1965). The functional separation of venous and arterial blood in the univentricular frog heart. Ann. N. Y. Acad. Sci. 127, 459–476.

Hightower, J.A., Burke, J.D., and Haar, J.L. (1975). A light and electron microscopic study of the respiratory epithelium of the adult aquatic newt, *Notophthalmus viridescens*. Can. J. Zool. 53, 465–472.

Hoffmann, A.D., Peterson, M.A., Friedland-Little, J.M., Anderson, S.A., and Moskowitz, I.P. (2009). Sonic hedgehog is required in pulmonary endoderm for atrial septation.

Development 136, 1761–1770.

Hoffmann, A.D., Yang, X.H., Burnicka-Turek, O., Bosman, J.D., Ren, X., Steimle, J.D., Vokes, S.A., McMahon, A.P., Kalinichenko, V. V., and Moskowitz, I.P. (2014). Foxf genes integrate Tbx5 and Hedgehog pathways in the second heart field for cardiac septation. PLoS Genet. 10, e1004604.

Hogan, B.L.M. (1999). Morphogenesis. Cell 96, 225–233.

Hoogmoed, M.S., Maciel, A.O., and Coragem, J.T. (2011). Discovery of the largest lungless tetrapod, *Atretochoana eiselti* (Taylor, 1968) (Amphibia: Gymnophiona: Typhlonectidae), in its natural habitat in Brazilian Amazonia. Bol. Do Mus. Para. Emilio Goeldi, Ciencias Nat. 6, 241–262.

Hopkins, G. (1896). The heart of some lungless salamanders. Am. Nat. 30, 829–833.

Icardo, J.M., Ojeda, J.L., Colvee, E., Tota, B., Wong, W.P., and Ip, Y.K. (2005). Heart inflow tract of the African lungfish *Protopterus dolloi*. J. Morphol. 263, 30–38.

Johansen, K. (1962). Double circulation in the amphibian *Amphiuma tridactylum*. Nature 194, 991–992.

Johansen, K., and Hanson, D. (1968). Functional anatomy of the hearts of lungfishes and amphibians. Am. Zool. 8, 191–210.

Lombard, R.E., and Wake, D.B. (1986). Tongue evolution in the lungless salamanders, family Plethodontidae IV. Phylogeny of plethodontid salamanders and the evolution of feeding dynamics. Syst. Zool. 35, 532–551.

Maina, J. (2011). Bioengineering Aspects in the Design of Gas Exchangers (Heidelberg: Springer).

Maina, J.N. (2002). Amphibian lung. In Functional Morphology of the Vertebrate Respiratory Systems (Enfield, NH: Science Publishers), pp. 70–84.

McCulley, D., Wienhold, M., and Sun, X. (2015). The pulmonary mesenchyme directs lung development. Curr. Opin. Genet. Dev. 32, 98–105.

McMullen, E.C. (1938). The morphology of the aortic arches in four genera of plethodontid salamanders. J. Morphol. 62, 559–597.

Meban, C. (1979). An electron microscope study of the respiratory epithelium in the lungs of the fire salamander (*Salamandra salamandra*). J. Anat. 128, 215–224.

Mekeel, A.G. (1930). Pulmonary development in the lungless salamanders. Unpubl. PhD Diss. Cornell Univ. 1–115.

Metzger, R.J., Klein, O.D., Martin, G.R., and Krasnow, M.A. (2008). The branching

programme of mouse lung development. Nature 453, 745–750.

Miller, L.D., Wert, S.E., and Whitsett, J.A. (2001). Surfactant proteins and cell markers in the respiratory epithelium of the amphibian, *Ambystoma mexicanum*. Comp. Biochem. Physiol. A. Mol. Integr. Physiol. 129, 141–149.

Miller, W.S. (1893). The structure of the lung. J. Morphol. 8, 165–189.

Morrisey, E.E., and Hogan, B.L.M. (2010). Preparing for the first breath: genetic and cellular mechanisms in lung development. Dev. Cell 18, 8–23.

Noble, G.K. (1925). The integumentary, pulmonary, and cardiac modifications correlated with increased cutaneous respiration in the amphibia: A solution of the "hairy frog" problem. J. Morphol. 40, 341–416.

Nussbaum, R.A., and Wilkinson, M. (1995). A new genus of lungless tetrapod: A radically divergent caecilian (Amphibia: Gymnophiona). Proc. R. Soc. London. Ser. B Biol. Sci. 261, 331–335.

Pattle, R.E., Schock, C., Creasey, J.M., and Hughes, G. (1977). Surpellic films, lung surfactant, and their cellular origin in newt, caecilian, and frog. J. Zool. London 182, 125–136.

Peng, T., Tian, Y., Boogerd, C.J., Lu, M.M., Kadzik, R.S., Stewart, K.M., Evans, S.M., and Morrisey, E.E. (2013). Coordination of heart and lung co-development by a multipotent cardiopulmonary progenitor. Nature 500, 589–592.

Podkowa, D., and Goniakowska-Witalińska, L. (1998). The structure of the airbladder of the catfish *Pangasius hypophthalmus* Roberts and Vidthayanon 1991 (previously *P. sutchi* Fowler 1937). Folia Biol. 46, 189–196.

Power, J.H., Doyle, I.R., Davidson, K., and Nicholas, T.E. (1999). Ultrastructural and protein analysis of surfactant in the Australian lungfish *Neoceratodus forsteri*: evidence for conservation of composition for 300 million years. J. Exp. Biol. 202, 2543–2550.

Putnam, J., and Kelly, D. (1978). A new interpretation of interatrial septation in the lungless salamander, *Plethodon glutinosus*. Copeia 1978, 251–254.

Que, J., Choi, M., Ziel, J.W., Klingensmith, J., and Hogan, B.L.M. (2006). Morphogenesis of the trachea and esophagus: current players and new roles for noggin and Bmps. Differentiation 74, 422–437.

Rankin, S.A., Thi Tran, H., Wlizla, M., Mancini, P., Shifley, E.T., Bloor, S.D., Han, L., Vleminckx, K., Wert, S.E., and Zorn, A.M. (2015). A molecular atlas of *Xenopus* respiratory system development. Dev. Dyn. 244, 69–85.

Reagan, N.L., and Verrell, P.A. (1991). The evolution of plethodontid salamanders: did

terrestrial mating facilitate lunglessness? Am. Nat. 138, 1307–1313.

Rose, C.S., and James, B. (2013). Plasticity of lung development in the amphibian, *Xenopus laevis*. Biol. Open 2, 1324–1335.

Ruben, J., and Boucot, A. (1989). The origin of the lungless salamanders (Amphibia: Plethodontidae). Am. Nat. 134, 161–169.

Ruben, J., Reagan, N., Verrell, P., and Boucot, A. (1993). Plethodontid salamander origins: A response to Beachy and Bruce. Am. Nat. 142, 1038–1051.

Rudnick, D. (1933). Developmental capacities of the chick lung in chorioallantoic grafts. J. Exp. Zool. 66, 125–153.

Schachner, E.R., Hutchinson, J.R., and Farmer, C. (2013). Pulmonary anatomy in the Nile crocodile and the evolution of unidirectional airflow in Archosauria. PeerJ 1, e60.

Serls, A.E., Doherty, S., Parvatiyar, P., Wells, J.M., and Deutsch, G.H. (2005). Different thresholds of fibroblast growth factors pattern the ventral foregut into liver and lung. Development 132, 35–47.

Shen, X.-X., Liang, D., Chen, M.-Y., Mao, R.-L., Wake, D.B., and Zhang, P. (2016). Phylogeny, time and biogeography of plethodontids. Syst. Biol. 65, 66–81.

Spooner, B.S., and Wessells, N.K. (1970). Mammalian lung development: interactions in primordium formation and bronchial morphogenesis. J. Exp. Zool. 175, 455–466.

Tsuda, S. (1924). Über die entwicklungsgeschichte und morphologie des kehlkopfes von *Hynobius nebulosus*. Folia Anat. Jpn. 2, 5–82.

Weibel, E.R. (1984). The Pathway for Oxygen (Cambridge: Harvard University Press).

Whitsett, J.A., and Weaver, T.E. (2002). Hydrophobic surfactant proteins in lung function and disease. N. Engl. J. Med. 347, 2141–2148.

Wilder, H.H. (1894). Lungenlose salamandriden. Anat. Anz. 9, 216–220.

Wilder, H.H. (1896). Lungless salamanders. Anat. Anz. 12, 182–192.

Wilder, I.W., and Dunn, E. (1920). The correlation of lunglessness in salamanders with a mountain brook habitat. Copeia 1920, 63–68.

Wilkinson, M., Kok, P.J.R., Ahmed, F., and Gower, D.J. (2014). *Caecilita* Wake & Donnelly, 2010 (Amphibia: Gymnophiona) is not lungless: Implications for taxonomy and for understanding the evolution of lunglessness. Zootaxa 3779, 383–388.

Zhang, P., and Wake, D.B. (2009). Higher-level salamander relationships and divergence dates inferred from complete mitochondrial genomes. Mol. Phylogenet. Evol. 53, 492–508.

Zheng, W., Wang, Z., Collins, J.E., Andrews, R.M., Stemple, D., and Gong, Z. (2011). Comparative transcriptome analyses indicate molecular homology of zebrafish swimbladder and mammalian lung. PLoS One 6, e24019.

Chapter 2

Lung development in lungless salamanders

Collaborators: Ryan R. Kerney, James Hanken

2.1 Abstract

Although lungs have been crucial to the remarkable adaptive radiation of tetrapods, they evolved well before tetrapods emerged onto land. Over the course of evolution, certain amphibians have dispensed with pulmonary respiration altogether, relying instead on cutaneous gas exchange through the skin and the lining of the mouth. While lung loss was long thought to be restricted to salamanders, recent work has demonstrated lung loss in all three modern orders of amphibians—Anura, Caudata, and Gymnophiona. Yet, it remains unclear why or how lung loss occurs. I have discovered that lungless salamanders (family Plethodontidae) complete early stages of lung morphogenesis and begin to develop lung primordia during

embryogenesis. Lungless salamanders express the lung specification gene Wnt2b. Lung rudiments are only transiently present and their regression results in lung loss by the time of hatching. Prior to regression, the lung rudiment expresses pulmonary markers Nkx2.1 and Sox9. The morphological lung rudiment likely represents retention of the ancestral lunged state. In addition to recapitulation of lung morphogenesis in plethodontid salamanders there is conservation of lung developmental genetic pathways. Expression of lung development and differentiation markers offers additional evidence that lung developmental genetic pathways are conserved in lungless salamanders, despite absence of functional adult lungs for over 60 million years.

2.2 Introduction

Loss of organs is a common theme in evolutionary history. From tooth loss in birds to loss of limbs in snakes and cetaceans, these losses often have profound impacts on animal ecology, locomotion, and physiology (Cohn and Tickle, 1999; Herrera et al., 2013; Thewissen et al., 2006). Microevolutionary losses, such as pelvic girdle loss in threespine sticklebacks and trait loss in the cavefish *Astyanax mexicanus*, have furthered our understanding of the evolution of development and the genetic underpinnings of trait loss (Chan et al., 2010; Gross et al., 2009). However, these studies have not addressed macroevolutionary morphological transitions during the origin of new species or families of animals. Studying these transitions in detail is difficult because extant transitional species rarely exist. Examining rudimentary structures can help to circumvent this problem. Organ rudiments are structures transiently retained in embryos, which hint at an animal's evolutionary history. Examples include the

transient appearance of limb buds in some limbless lizards, or hind limb buds in some marine mammals (Bejder and Hall, 2002; Rahmani, 1974; Sedmera et al., 1997). Rudiments offer an opportunity to study the mechanisms by which organs are lost, and the evolutionary developmental basis of evolutionary transitions.

One instance of organ loss at the base of a major adaptive radiation is the loss of lungs in plethodontid salamanders (Fig. 1.1). Plethodontid salamanders are the most speciose family of salamanders and account for over two thirds of salamander diversity (Wake and Hanken, 1996). All plethodontid species are lungless; they have no lungs and respire entirely through non-pulmonary tissues, mainly the skin and buccopharyngeal mucosa. Lung loss has convergently occurred at least three additional times in distantly related amphibians, and lung reduction is common (Fig. 1.1; Bickford et al., 2008; Noble, 1925; Nussbaum and Wilkinson, 1995). The adaptive significance of lung loss is highly debated (Beachy and Bruce, 1992; Reagan and Verrell, 1991; Ruben et al., 1993; Wilder and Dunn, 1920). Little is known about the developmental underpinnings of lung loss.

The search for a vestigial lung began immediately after the discovery of lunglessness. Wilder (1896) and Ludike (1955) both observed a pharyngeal invagination in adults that they identified as a retained glottis. In an unpublished doctoral dissertation, Amy Grace Mekeel compared lung morphogenesis in lunged salamanders to several species of plethodontids. Although she found no trace of a glottis rudiment in adults, Mekeel claimed that lungless salamanders display embryonic indications of lung development. However, the lung rudiment was present transiently and disappeared during embryogenesis (Mekeel, 1926, 1930).

If a transient rudiment is present, then this implies that lung loss occurs through degeneration of a lung rudiment and that plethodontid salamanders demonstrate some conserved

aspects of lung development, despite absence of adult lungs. Aspects of the genetic mechanisms underlying lung development must then be present in plethodontids, and not lost to drift during the 66 Ma or more since this family diverged from their last lunged ancestor (Shen et al., 2016).

It is unknown whether portions of the gene regulatory network involved in lung specification, morphogenesis and differentiation are expressed in lungless salamanders. One possibility is that lungless salamanders have somehow mutated, downregulated or lost a master regulator gene involved in lung development, resulting in the failure of expression of lung-specific genes. An alternative possibility is that some genes in the lung gene regulatory network maintain expression in plethodontids due to their pleiotropic roles in the development of other organs. Several critical, early markers of lung development have been identified in mammals and frogs, including Wnt2b and Nkx2.1 (Goss et al., 2009; Minoo et al., 1999; Rankin et al., 2012, 2015). These genes may serve as indicators of lung gene regulatory network activity.

I present evidence that lung developmental genetic pathways are maintained in plethodontid salamanders. In particular, early development of the lung is conserved. I show conservation of lung specification and the expression of pulmonary markers. However, the lung anlage is present only transiently and regresses by apoptosis.

2.3 Materials and methods

2.3.1 Embryo collection and husbandry

All plethodontid embryos were collected under Massachusetts Division of Fisheries and Wildlife permits and local permits, where applicable: DFW permit numbers 181.10SCRA (2010), 080.11SCRA (2011), 080.11SCRA (2012), 027.13SCRA (2013), 083.14SCRA (2014), and 022.15SCRA (2015); Cape Cod National Seashore permit number CACO-00214 (*Hemidactylium scutatum*). Locality data is provided in Table A.1. *Plethodon cinereus* embryos were typically found underneath moss atop fallen hemlock logs or inside rotting hemlock logs. *Plethodon cinereus* embryos were also obtained by injecting gravid females with 0.1 ml 10 mg/ml human chorionic gonadotropin (#CG10-1VL, Sigma, St. Louis, MO) two times one week apart, then allowing them to lay eggs in the dark at 17°C. *Desmognathus fuscus* embryos were found in seepage areas underneath mats of moss atop fallen logs or large rocks.

Embryos were kept at 15–17°C on filter paper moistened with 0.1X MMR (Marc's Modified Ringer solution: 0.01 M NaCl, 0.2 mM KCl, 0.1 mM $MgSO_4$, 0.2 mM $CaCl_2$, 0.5 mM HEPES pH 7.4) and 100 µg/ml gentamicin (Sigma) or fully immersed in the same solution. Better success was obtained raising embryos in solution. Embryos were monitored daily for fungal infections and to check their developmental stage. Fungal infections were treated by dechorionating embryos and rearing in solution. Alternate approaches to treating fungal infections, such as washes in hydrogen peroxide solution or use of methylene blue, did not improve mortality rate. Staging was performed as described for *Plethodon cinereus* (Kerney 2011) and for *Hemidactylium scutatum* (Hurney et al., 2015). *Desmognathus fuscus* stage numbers correspond to comparable stages in *P. cinereus* (Kerney, 2011). Developmental

timing for *P. cinereus* was obtained by raising three clutches of embryos from laying to hatching at a constant 15°C and photographing these embryos daily or every other day.

Ambystoma mexicanum embryos were obtained from the *Ambystoma* Genetic Stock Center, University of Kentucky. Embryos were reared at 17°C immersed in 20% Holtfreter solution and staged according to Bordzilovskaya et al. (1989).

2.3.2 Histology

Embryos were fixed in neutral-buffered formalin or Bouin fixative, washed with 1x phosphate-buffered saline (PBS), dehydrated then cleared with xylene, and embedded in Paraplast (McCormick Scientific, Wetzlar, Germany). Serial sections (7 or 8 µm) were stained using Mallory trichrome method (Presnell et al., 1997). The following modifications were made: A 10-min stain in Mayer hematoxylin followed rehydration of sections. Sections were washed 10 min in dH_2O, immersed 30 sec in 1% acid fuchsin, rinsed several times with dH_2O, immersed in 1% phosphomolybdic acid for 5 min, then in a modified Mallory II (1% orange G, 1% aniline blue, and 2% oxalic acid) for 3 min, followed by dehydration and mounting. Photographs were taken using a Leica DMRE microscope (Wetzlar, Germany) equipped with a QImaging Retiga 2000r camera and a QImaging RGB slider (Model: RGB-HM-S-IR; Surrey, Canada) and Volocity 6.0 software (PerkinElmer, Waltham, MA). Adjustment of exposure, levels and color balance were performed in Adobe Photoshop CS5.

2.3.3 Micro-CT imaging and histological reconstruction

µ-CT imaging was performed on specimens stained for several days with Lugol's iodine solution (Metscher, 2011). 3-D reconstructions were completed in Amira 6.0 (FEI Visualizations Sciences Group, Hillsboro, OR). For histological reconstruction, serial sections were imaged as above. Monochrome images were imported into Amira and aligned using the Align Slices module. Endocasts of pharyngeal and foregut spaces were segmented using the segmentation editor. Additional details on µ-CT imaging and analysis are presented in Chapter 5 Materials and methods.

2.3.4 PCR and probe design

Amplified fragments were cloned into the pcrII-TOPO vector (Invitrogen, Carlsbad, CA). Clones were sequenced and *in situ* hybridization probes were constructed using Ambion Maxiscript SP6/T7 enzymes (Grand Island, NY) and incorporating digoxigenin-labeled UTP (Roche, Penzberg, Germany). Primers are provided in Table A.2. *In situ* probe clones obtained from other researchers are listed in Table A.3.

2.3.5 *In situ* hybridization

Briefly, embryos were fixed overnight in 4% paraformaldehyde of MEMFA (0.1 M MOPS (pH 7.4), 2 mM EGTA, 1 mM $MgSO_4$, 3.7% formaldehyde), dehydrated and stored in 70 or 100% MeOH at -20°C. Wholemount mRNA *in situ* hybridization (ISH) was performed by rehydrating samples, then samples were treated with between 5 and 10 µg/ml proteinase K for 30 to 60 min, washed with PBTw (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4,

1.8 mM KH_2PO_4, 0.2% Tween-20), post-fixed in 4% paraformaldehyde, washed with PBTw, and pre-hybridized in hybridization buffer for 2 hr at 65°C (hybridization buffer: 50% formamide, 5x SSC, 0.1 mg/ml heparin, 1x Denhardt's, 0.01% CHAPS, 0.2 mg/ml tRNA, 0.1% Tween-20; all solutions RNase-free). DIG-labeled riboprobes were diluted approximately 1:40 in hybridization buffer then denatured at 85°C for 10 min before adding to specimens. Hybridization was carried out overnight at 65°C. Posthybridization washes were performed with a solution of 50% formamide, 5x SSC and 0.2% Tween-20 at 65°C for 8 changes of 30 min each. Specimens were washed with maleic acid buffer plus 0.2% Tween-20 (MABT) prior to blocking and antibody incubation. Antibody block solution included 20% heat-inactivated sheep serum, 2% blocking reagent (Roche, Penzberg, Germany) in MABT. Samples were incubated with 1:2500 anti-DIG-AP Fab fragments (Roche) diluted in blocking solution overnight at 4°C. Extensive washes with MABT were performed prior to color development using BM-Purple (Roche) or NBT/BCIP (Sigma). Color development occurred over several hours to several days and was stopped by administration of 1 mM EDTA in PBS, followed by postfixing in 4% PFA. Embryos were then embedded in OCT for cryosectioning at 14–16 µm section thickness. Sections were cut at −23°C, dried, rehydrated in PBS and then mounted using Fluoromount-G (eBioscience, San Diego, CA).

2.3.6 Measuring apoptosis

Apoptotic rates were examined at st. 20/21 in *Plethodon cinereus* and st. 41 in *A. mexicanum*. *Plethodon* lung rudiments show no histological signs of regression by these stages. Standard terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) was per-

formed using an *in situ* apoptosis detection kit (Trevigen #4810-30-K). To calculate the percentage of lung rudiment cells undergoing apoptosis, the number of apoptotic and normal cells were counted in sections processed using TUNEL. Cell counting was performed in ImageJ by distinguishing cell nuclei stained with Methyl Green and DAB positive apoptotic cells. Statistical differences were assessed using a two-tailed Mann-Whitney U-Test. Figures are reported as the mean ± SEM. Apoptosis was only quantified in the endoderm, but *Plethodon* showed a large amount of mesenchymal apoptosis that was not quantified. Endodermal and mesenchymal tissues are easily distinguished based on their physical separation and the epithelial morphology of endoderm.

2.3.7 Transcriptome sequencing

See Appendix B.

2.4 Results

2.4.1 Lungless salamanders develop a lung rudiment

I histologically analyzed developmental series of lungless salamanders to determine whether a lung rudiment forms, primarily focusing on the eastern red-backed salamander, *Plethodon cinereus*. Mexican axolotl, *Ambystoma mexicanum*, was used as a lunged outgroup. The gross morphology of these two species is quite distinct, but it is possible to compare similar stages of development by noting the formation of several relatively invariant features. These include the formation and branching of the external gills and the development of the heart,

eyes and limbs (Bordzilovskaya et al., 1989; Kerney, 2011). I performed all comparisons at equivalent stages based on these criteria. These stages have different corresponding numbers. Stages are abbreviated in the text using the first letters of the Latin binomial followed by the stage number.

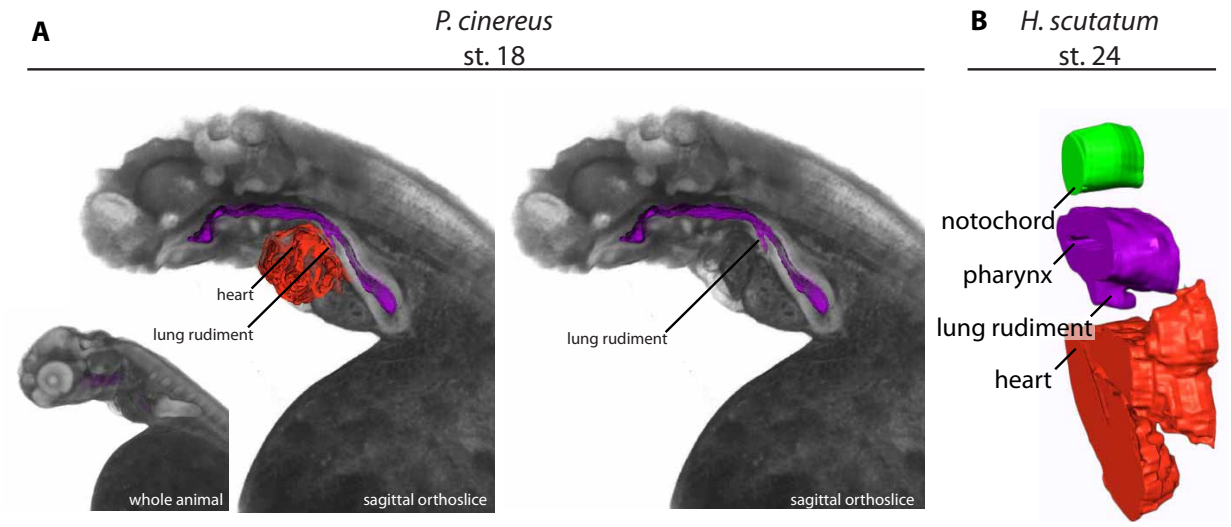


Figure 2.1: Lungless salamanders develop a lung rudiment. (A, B) While lungless as adults, plethodontid salamanders such as *Plethodon cinereus* (A) and *Hemidactylium scutatum* (B) develop a lung rudiment as embryos. Anterior is to the left. (A) *Plethodon cinereus* embryo stained with Lugol's iodine and subjected to X-ray µ-CT; reconstruction and segmentation were performed in Amira. (B) The *H. scutatum* specimen was reconstructed from serial histological sections in Amira. Endodermal tissue is purple, cardiac mesoderm is red, and the notochord is green.

The first morphological sign of lung development *Ambystoma mexicanum* is the formation of a median laryngotracheal groove in the ventral foregut at stage 38 ("Am38"; Fig. 2.3A,D,E). In *A. mexicanum*, the laryngotracheal groove becomes a distinct diverticulum—the laryngotracheal tube (LT, or lung rudiment), which emerges from the ventral foregut at the posterior margin of the developing heart at the axial level of the division between the sinus venosus and atrium during stage Am39 (Fig. 2.3B,F,G).

Similar morphogenetic events occur in the lungless salamander *Plethodon cinereus* (Fig. 2.2). The laryngotracheal groove forms during Pc17 and early Pc18 (Fig. 2.4). The LT is present by late Pc18 as a posteroventral diverticulum from the foregut, which has a lumen continuous with the pharynx (Figs. 2.1B, 2.5A,B). The *P. cinereus* LT is similar to that of *A. mexicanum*, which features a simple cuboidal epithelium surrounded by ventral mesenchyme (Fig. 2.2A,B). The LT appears less distinct in *A. mexicanum* because of the mesolecithal distribution of yolk, which slightly obscures morphology at early stages, as opposed to the telolecithal embryos of *P. cinereus*. The laryngotracheal tubes of *A. mexicanum* and *P. cinereus* have slight distal lateral buds at early stages.

The LT elongates following its formation in *A. mexicanum*. Lung buds at the distal end of the LT are clearly visible at Am40 (Figs. 2.3C, 2.2A). Embryos at stage Am41 show an enlarged trachea, triangular in cross-section, and enlarged lung buds (Fig. 2.2C). The trachea and glottis are well developed by hatching (Am 44; Fig. 2.3J). At hatching the lungs are simple saccular structures with walls approximately two cell-layers thick (Fig. 2.3K).

After the LT forms in *Plethodon cinereus* at Pc18 it elongates through Pc21 (Figs. 2.2, 2.5). During late Pc21 the LT begins to show histological signs of regression, including reduced length and collapse of the tube (Fig. 2.2D). The rates of apoptosis in the lung rudiment were quantified at Pc20, prior to morphological indications of lung rudiment regression and compared to a comparable stage of *A. mexicanum*. *Plethodon cinereus* have significantly higher levels of apoptosis in the lung rudiment than *A. mexicanum*, as assessed by TUNEL ($p < 0.05$; $6.3\% \pm 0.64\%$, $N = 4$ vs. $1.2\% \pm 0.14\%$, $N = 4$; Fig. 2.6C). Apoptotic figures are found mainly in the proximal LT and the surrounding anteroventral mesenchyme (Fig. 2.6B). No apoptosis was observed in the mesenchyme surrounding the distal tip of the LT

and very few apoptotic cells were identified throughout the rest of the embryo. During late Pc21 and Pc22 the lung rudiment begins to morphologically regress. There are no signs of an epithelial evagination and a thin but long mesenchymal strip of cells occupies the former location of the LT (Fig. 2.5C,D). By stage 23, there is no sign of either the condensed mesenchyme or the epithelial diverticulum (Fig. 2.5E).

The lung rudiment was also found in other plethodontid species, including *Hemidactylium scutatum* and *Desmognathus fuscus* (Figs. 2.1B, 2.7). In *D. fuscus*, LT development and regression closely resemble that of *P. cinereus* (Fig. 2.7). In *H. scutatum*, the rudiment appears to be present for a shorter duration of time. In all plethodontid species, the LT emerges at the stage when the gills begin to branch and at the same axial position as the LT in *A. mexicanum*—dorsal to the division between the atrium and sinus venosus of the heart.

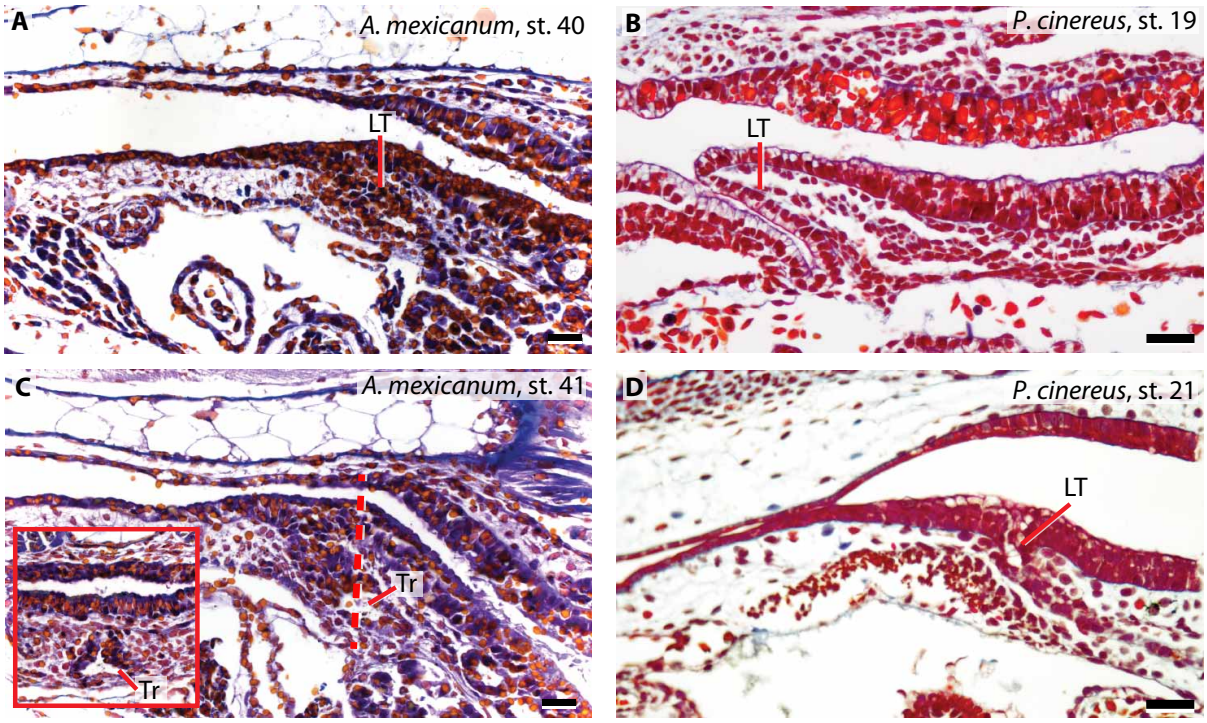
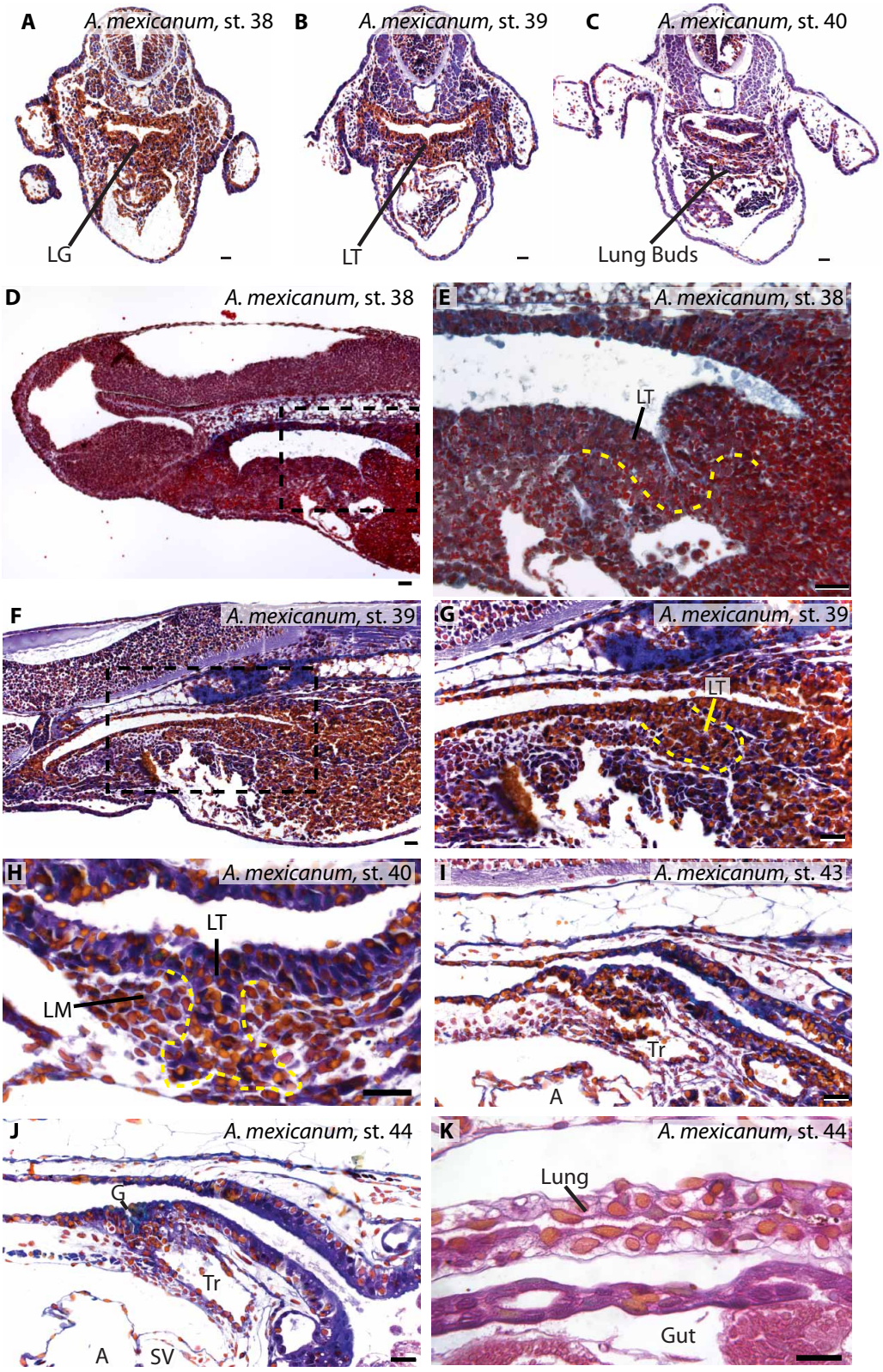


Figure 2.2: The embryonic lung rudiment regresses in plethodontids. Pulmonary development begins to diverge between lunged (A, C) and lungless (B, D) species following the formation of the laryngotracheal tube (LT). Mid-sagittal sections; anterior is to the left. (A) *Ambystoma mexicanum* shows elongation of the LT and the presence of histologically distinguishable lung buds by stage 40 (lung buds not shown). (B) *Plethodon cinereus* develops an LT at stage 18; stage 19 (pictured) is approximately three days after LT emergence. The LT elongates through stage 20 and early stage 21. (C) A developing trachea (Tr) is present in *A. mexicanum* at stage 41, surrounded by lung mesenchyme. Inset is a transverse section at the axial level indicated by the dashed line. (D) The LT displays histological indications of regression during late stage 21 in *P. cinereus*. Scale bars, 50 µm.

Figure 2.3: Lung development in *Ambystoma mexicanum*, a salamander with lungs. (A-C): Transverse sections showing key features of lung development. (D-K): The development of the laryngotracheal tube (LT) and trachea as revealed by mid-sagittal and transverse sections. (A) Similar to *P. cinereus* (Fig. 2.4), lung development begins with the appearance laryngotracheal groove (LG) by stage 38. (B,D,E) The fold extends ventrally to form the LT at stage 38. (E) Represents the boxed region in (D). (C) A more posterior section reveals the presence of the lung buds by stage 40, which immediately begin to extend caudally. (F) The LT continues to extend at stage 39. (G) Close-up of the boxed region in (F). (H) Transverse view of the bifurcated LT surrounded by lung mesenchyme (LM) at the anterior portion of the LT at stage 40. (I) The trachea has fully formed by stage 43. (J) By stage 44 the glottis (G) becomes distinct. The lungs (K) and trachea are fully formed. Additional abbreviations: A, atrium; SV, sinus venosus; TA, truncus arteriosus; V, ventricle. Scale bars, 50 µm

Figure 2.3, continued



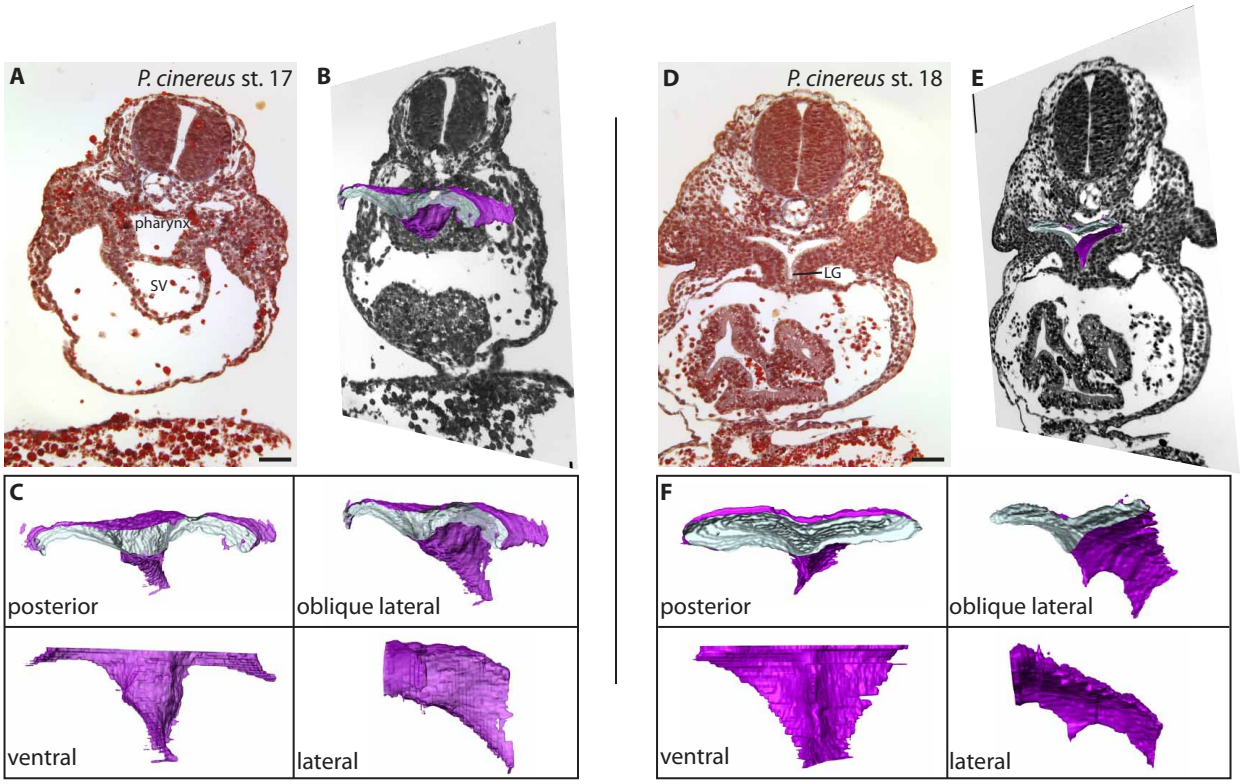


Figure 2.4: Formation of the laryngotracheal groove in the lungless salamander *Plethodon cinereus*. (A-C): Pharyngeal and foregut morphology at stage 17. (A) Transverse histological section through the putative lung region shows a wide, rectangular pharynx dorsal to the sinus venosus (SV). (B) Oblique lateral view of a histological reconstruction of an endocast of the pharynx (purple) overlaid on a representative orthoslice. (C) Views of the pharyngeal endocast from anterior-posterior, oblique lateral, ventral and lateral perspectives. (D-F): Pharyngeal and foregut morphology at st. 18. (D) Transverse histological section through the putative lung region shows the laryngotracheal groove (LG). (E) Oblique lateral view of a histological reconstruction of an endocast of the pharynx (purple) overlaid on a representative orthoslice. (F) Views of the pharyngeal endocast from anterior-posterior, oblique lateral, ventral and lateral perspectives. Scale bars, 100 µm. Histological reconstructions are not to scale.

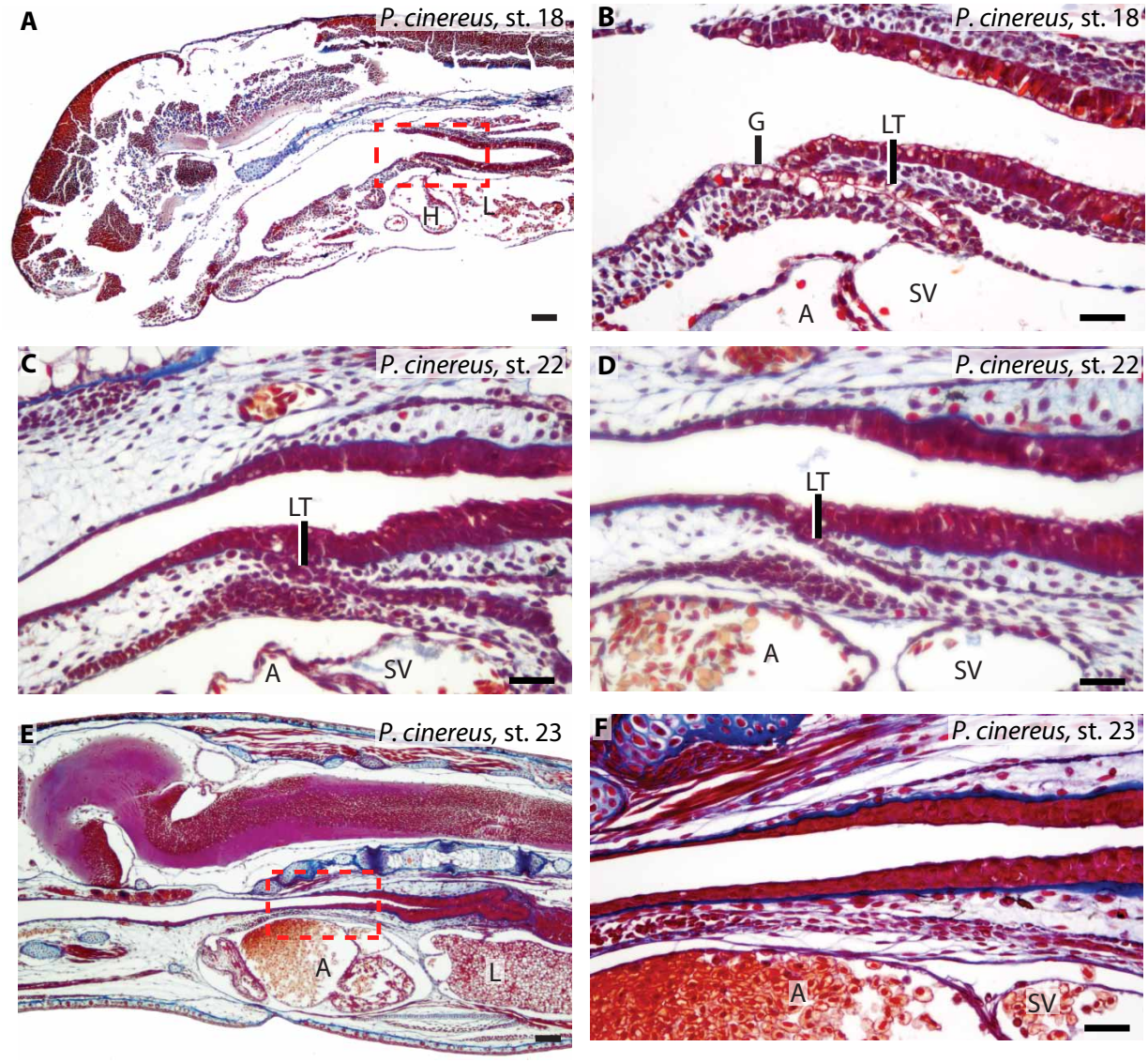


Figure 2.5: Formation and regression of the laryngotracheal tube in the lungless salamander *Plethodon cinereus*. (A-F): Sagittal sections through *P. cinereus* embryos over three developmental stages; anterior is to the left. (A) A distinct glottis (G) and laryngotracheal tube (LT) form at stage 18. (B) Enlargement of boxed region in (A). (C, D) The LT has collapsed by stage 22 and forms a band of thin mesenchyme with only a small endodermal outpocketing. (E) The LT is completely absent by stage 23, prior to hatching. (F) Enlargement of boxed region in (E). Additional abbreviations: A, atrium; L, liver; SV, sinus venosus. Scale bars, 50 µm (B,C,D,F); 150 µm (A,E).

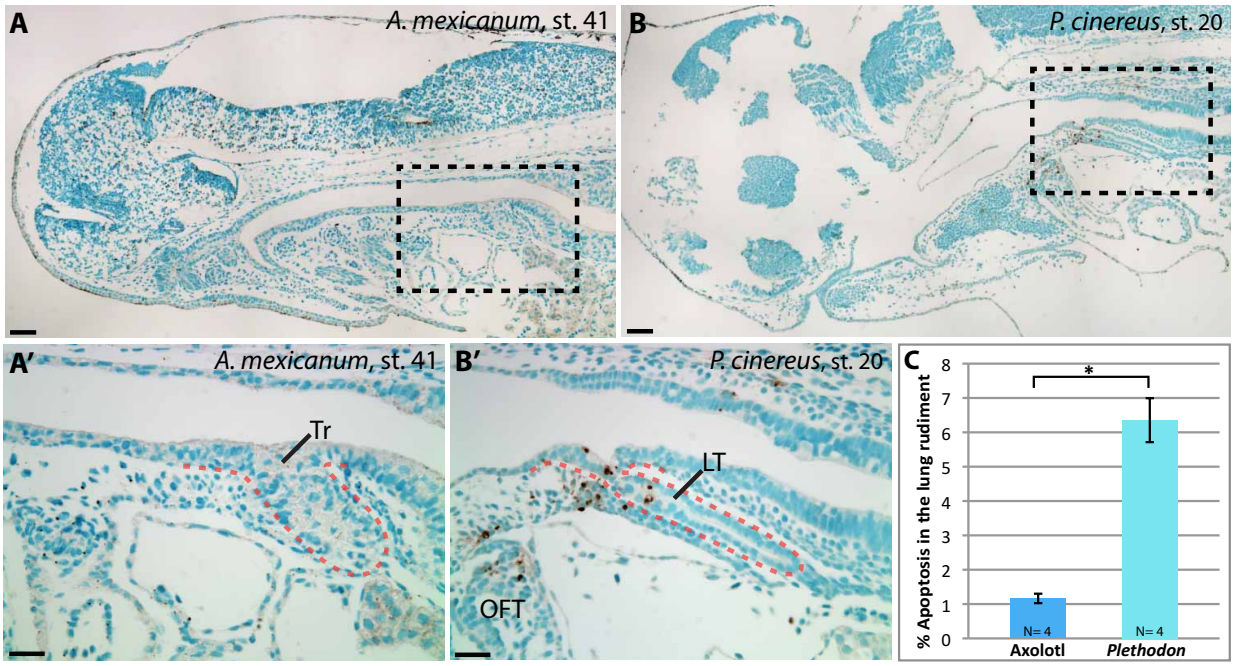


Figure 2.6: The lung rudiment regresses by apoptosis. Mid-sagittal sections; anterior is to the left. (A, A') *Ambystoma mexicanum* display few apoptotic figures, as assessed by TUNEL using DAB chromogen. The trachea (Tr) is outlined by the red dashed line. Black specks in the integument and central nervous system are melanocytes. (B, B') At stages just preceding histological regression (early stage 20), lungless *Plethodon cinereus* embryos have numerous apoptotic cells (brown staining) in the laryngotracheal tube (LT) and surrounding mesenchyme. Few apoptotic cells are present in other tissues. (C) *P. cinereus* have significantly higher rates of apoptosis in the endoderm of the lung rudiment compared to *A. mexicanum* (Mann-Whitney U-test, $p < 0.05$). Additional abbreviations: OFT, outflow tract. Scale bars, 100 µm (A, B); 50 µm (A', B').

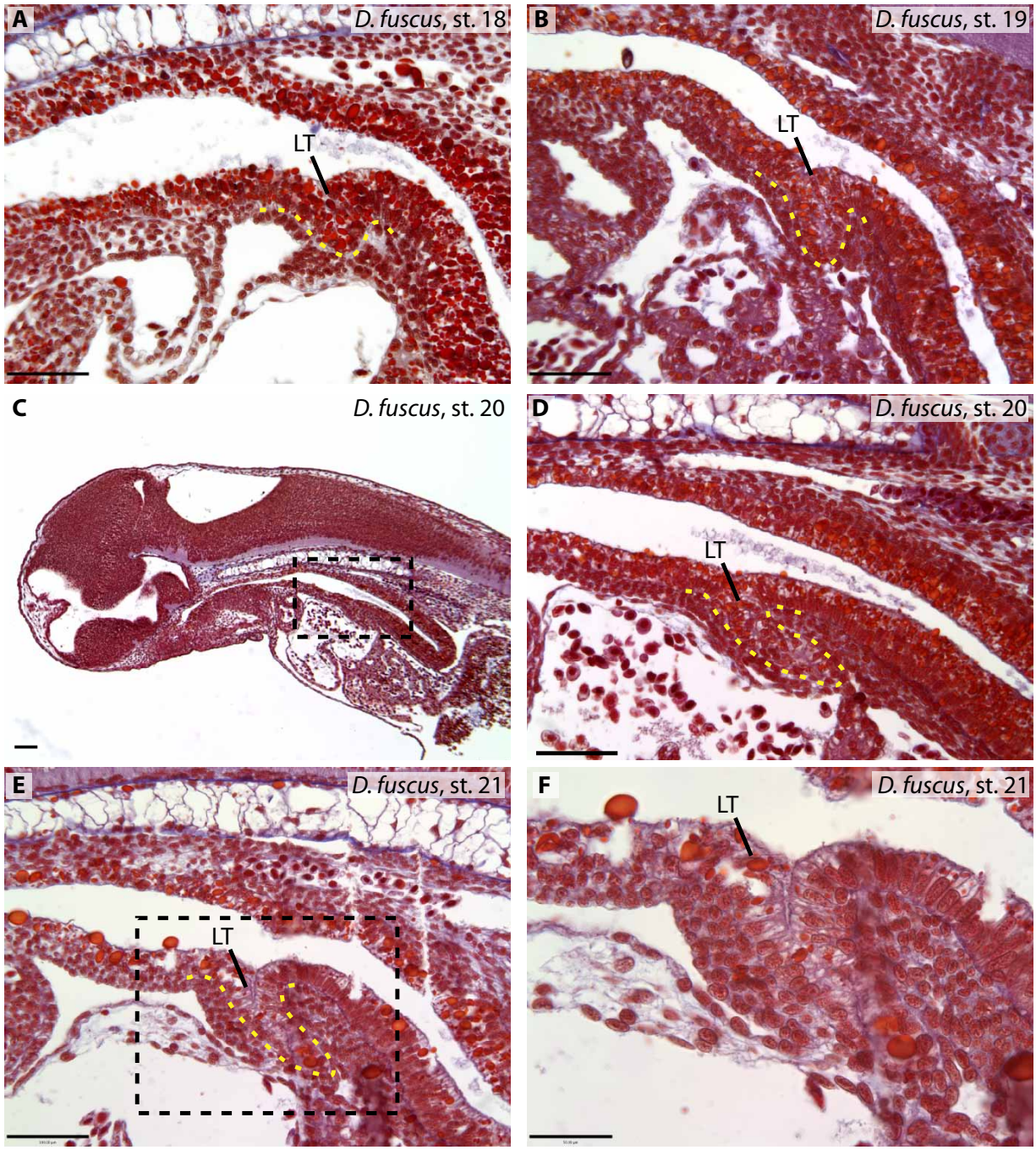


Figure 2.7: The laryngotracheal tube develops in *Desmognathus fuscus*. Mid-sagittal sections; anterior is to the left. Stages are matched to *P. cinereus* stages (Kerney, 2011). (A) The formation of the laryngotracheal tube (LT) at stage 18. The endodermal extent of the LT is outlined by a yellow dashed line. (B-D) The LT diverticulum elongates from stage 19 through 20. (E, F) At stage 21, the LT appears to collapse distally and cease elongation. Scale bars: 100 µm, except (F): 50 µm.

2.4.2 Genetic markers of lung development are expressed in plethodontid salamanders

The presence of a transient morphological rudiment is evidence that pathways underlying early lung development remain intact in plethodontids, despite absence of lungs in adults. If the corresponding genetic networks are intact, then I expect that lung-specific genes will be expressed in plethodontids. Prior to lung outgrowth, the lungs are specified by the expression of mesenchymal Wnt2 and Wnt2b (Goss et al., 2009; Harris-Johnson et al., 2009). In mammals and frogs the first marker of the developing lungs is the homeobox gene Nkx2.1 (Hines and Sun, 2014; Rankin et al., 2015). Subsequently, lungs begin expressing markers such as Sox9, which markers the distal airways (Perl et al., 2005; Rankin et al., 2015). I examined the expression of Wnt2b, Nkx2.1 and Sox9 over a range of stages in lungless salamanders by *in situ* hybridization.

In *Plethodon cinereus* embryos, Wnt2b is expressed prior to the formation of the lungs in two mesenchymal domains surrounding the laryngotracheal groove (Fig. 2.8A). This expression pattern precisely matches that observed in *Xenopus laevis* (Rankin et al., 2015) and mammals (Goss et al., 2009). Expression of Wnt2b persists during the formation of the LT (Fig. 2.8B). Subsequently, the lung rudiment begins to express the lung marker Nkx2.1 specifically along the luminal surface of the LT at stages Pc19 through Pc21 (Fig. 2.9). Expression of Nkx2.1 persists at least until Pc21 (Fig. 2.9). Sox9, a marker for distal lung identity, is expressed in *P. cinereus* broadly at stage 19, including in the lung rudiment (data not shown). Sox9 expression is stronger and more specific to the LT at Pc21 (Fig. 2.10B,D). Expression is also observed in the head and the heart. In *A. mexicanum*, Sox9 is weakly

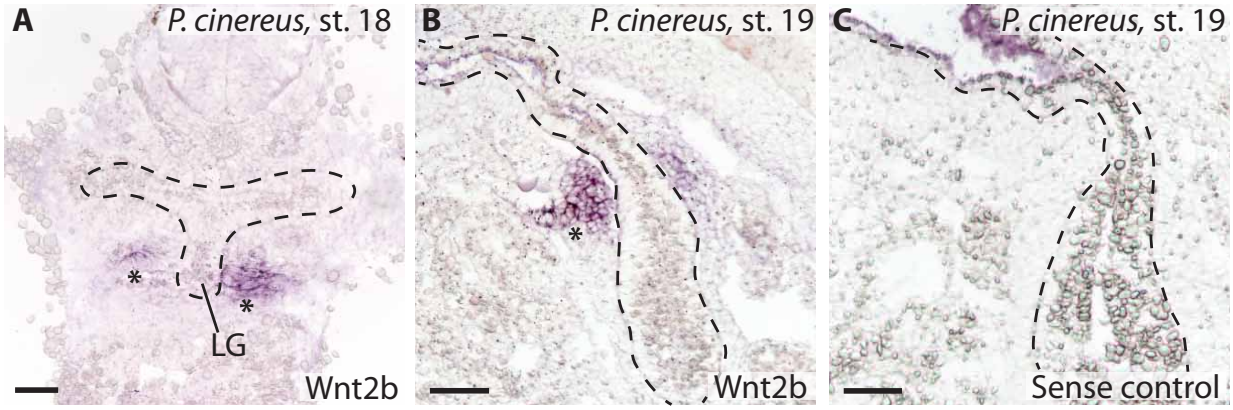


Figure 2.8: The lungless salamander *Plethodon cinereus* expresses Wnt2b, a marker of lung specification. (A-C) Wnt2b expression by *in situ* hybridization. A dashed line outlines the pharynx and foregut. (A) A transverse section through a stage-18 *P. cinereus* embryo shows the laryngotracheal groove (LG) flanked by lateral foci of Wnt2b-expressing mesenchyme (dorsal to the asterisks). (B) A sagittal section lateral to the laryngotracheal tube shows Wnt2b expression at stage 19. Anterior is to the left. (C) Negative control. Sagittal section at the same position as (B). Anterior is to the left. While some probe trapping has occurred in the pharynx, note the lack of staining in mesenchyme. Scale bars, 100 µm..

expressed at Am42 in the lungs (not shown) and expression becomes strong in the luminal surface of the lung by Am44 (Fig. 2.10A,C).

A number of markers of developing and mature mammalian lungs have been identified through bioinformatic and genetic methods (Chang et al., 2013; Treutlein et al., 2014). The expression patterns of many of these markers are conserved in amphibians (Rankin et al., 2015). As described in Appendix B, I sequenced the transcriptome of the LT at Pc19 and Pc21 in order to determine whether the gene expression profile of the LT overlaps with that of a developing lung. LT transcriptomes for *A. mexicanum* at Am40 and a single juvenile lung sample were compared to *Plethodon cinereus* LT transcriptomes to assess lung marker expression (Table 2.1). *Plethodon cinereus* expresses a number of markers for alveolar type II cells in the LT, including Lamp3, Abca3, Aquaporin 4 and all four pulmonary surfactant proteins A, B, C, and D. Markers of proximal and distal lung identity such as Sox2, FoxJ1

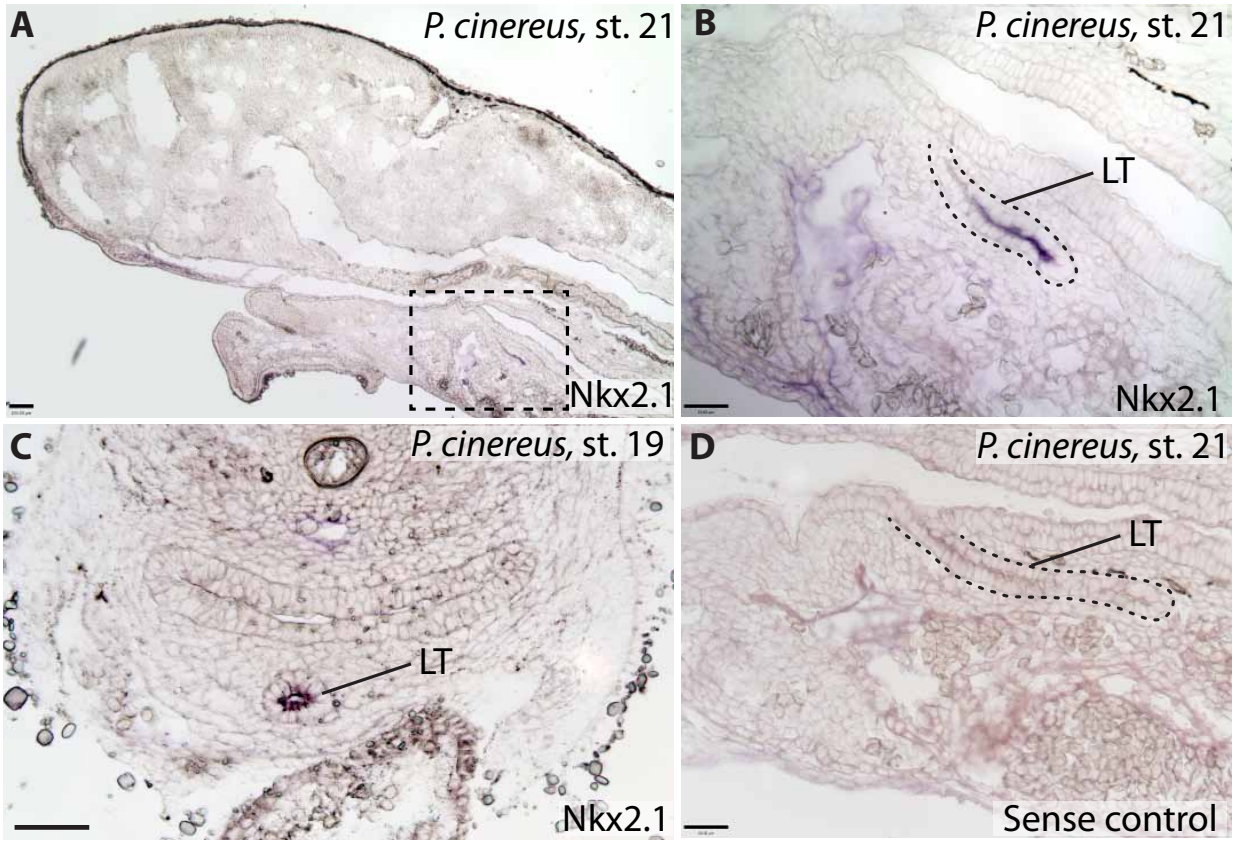


Figure 2.9: The lungless salamander *Plethodon cinereus* expresses Nkx2.1, an early marker of pulmonary identity. (A-C) Nkx2.1 expression by *in situ* hybridization. Sagittal sections; anterior is to the left. (A) Low magnification mid-sagittal section through a stage-21 *P. cinereus* embryo displays the specificity of Nkx2.1 expression to the laryngotracheal tube (LT). (B) Magnification of boxed region in (A). A dashed line highlights the laryngotracheal tube. Nkx2.1 expression is found specifically in the luminal cells lining the LT. (C) Transverse section at stage 19 illustrates the specific expression of Nkx2.1 in the LT. (D) Sense negative control. Scale bars, 100 µm (A, C); 50 µm (B, D).

and the aforementioned Sox9 are also found in the *P. cinereus* LT at both early and late developmental stages. Lung development markers such as Fgf signaling pathway members are also expressed in *P. cinereus* LT (Table 2.1). These expression data offer further evidence that lungs begin to form in lungless salamanders and of the homology of the LT from *P. cinereus* to the LT from *A. mexicanum* and other lunged tetrapods.

Table 2.1: Lung marker expression in the lung rudiments of *Plethodon cinereus* (lungless) and *Ambystoma mexicanum* (lunged). Gene expression was characterized by transcriptome sequencing of the lung rudiment, as described in Appendix B. AT2, alveolar type II cell.

	Lungless early (st. 19)	Lungless late (st. 21)	Lunged early (st. 40)	Adult lung	Cell type
SP-A	+	+	+	+	AT2
SP-B	+	+	+	+	AT2
SP-C	+		+	+	AT2
SP-D	+	+		+	AT2
LAMP-3	+	+	+	+	AT2
ABCA-3	+	+	+	+	AT2
NAPSA	+	+	+	+	AT2
Aquaporin 4		+	+	+	AT2
Cathespin-H	+	+	+	+	AT2
Sox2	+	+	+	+	Proximal
Sox9	+	+	+	+	Distal
FoxJ1	+	+	+	+	Ciliated
Nkx2.1	+		+	+	
GATA-6	+	+	+	+	
FOXA2	+	+	+	+	
Wnt2b	+	+	+	+	
Acidic chitinase		+	+		
Wnt7b	+	+	+	+	
Etv4	+	+	+		
Etv5	+	+	+		

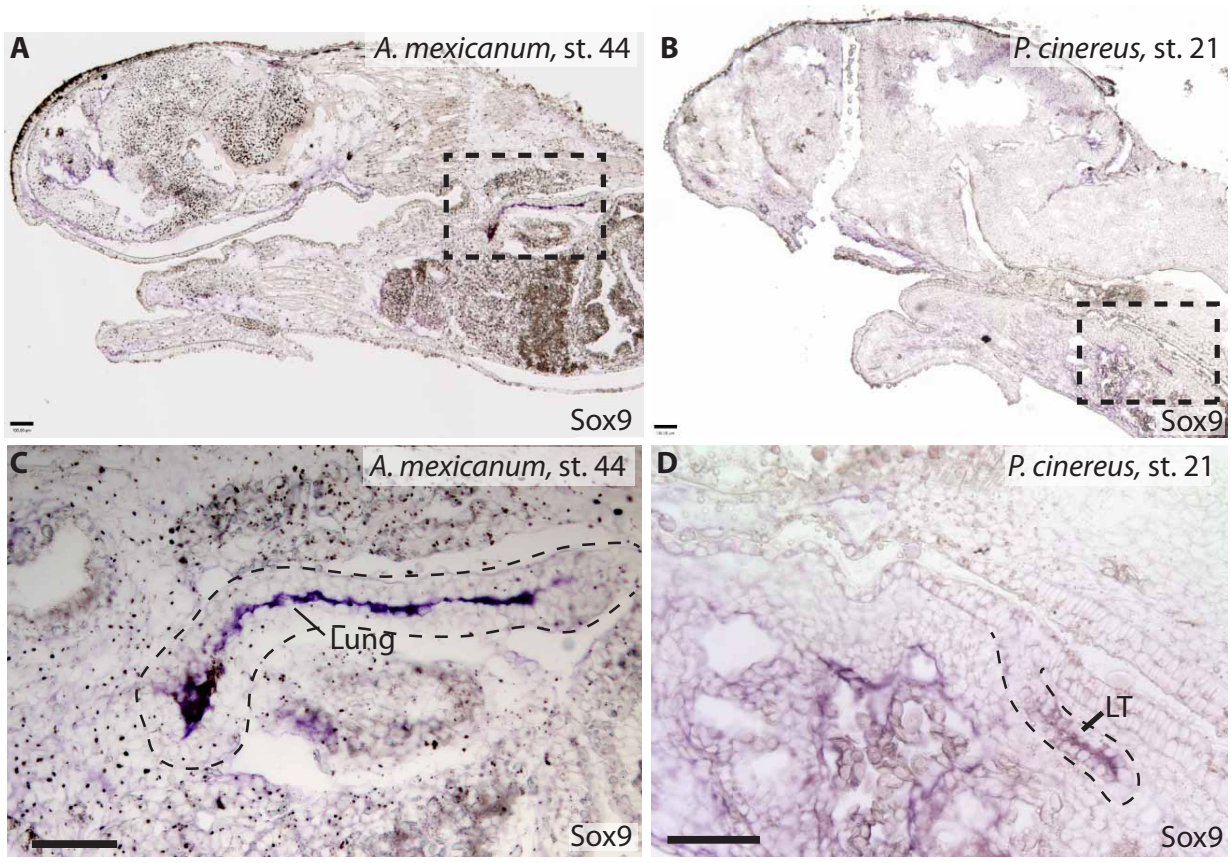


Figure 2.10: Sox9, a marker of pulmonary differentiation, is expressed in the laryngotracheal tube of *Plethodon cinereus*. (A, C) Expression of Sox9 in the lung of *Ambystoma mexicanum*. (A) Sagittal section, lateral to the midline, at stage 44. (C) Magnification of the boxed region in (A) highlights Sox9 expression in the cells lining the lumen of the lung, which is outlined by a dashed line. (B, D) Expression of Sox9 in the laryngotracheal tube (LT) at stage 21 in *P. cinereus*. Mid-sagittal section; anterior is to the left. (D) Magnification of the boxed region in (B) demonstrates Sox9 expression through most of the LT, which is outlined by a dashed line. Scale bars, 100 µm.

2.5 Discussion

Vestigial structures and rudiments provide windows into the evolutionary history of animals (Bejder and Hall, 2002; Hall, 1984; Kurten, 1963). Darwin (1859) placed considerable focus on organ rudiments as support for evolution by natural selection: "[rudimentary] organs may be compared with the letters in a word, still retained in the spelling, but become useless

in the pronunciation, but which serve as a clue in seeking for its derivation" (pp. 455-456). A common example is the transient limb buds or pelvic girdles found in tetrapods that have undergone limb loss. Vestiges of internal organs have received far less attention, likely because detailed anatomical characterization is necessary to observe them in the first place. Internal organ rudiments have the potential to reveal how organ loss can occur in highly pleiotropic genetic networks and within tightly integrated organ systems.

The search for a lung rudiment in plethodontid salamanders began soon after lunglessness was discovered. Wilder (1896) and Ludike (1955) claimed to observe a glottis rudiment in adult plethodontids. From my developmental evidence and examination of adult animals, it is unlikely likely that this structure is a pulmonary vestige. Instead it likely represents a folding of the pharynx. Mekeel's (1930) dissertation was the first discovery of a transient lung rudiment in plethodontids, but her results were never published in an academic journal. Although my results are largely in concordance with hers, I failed to find the early bilateral outpocketings from the pharynx that she observed. Mekeel's morphological descriptions are excellent. However, her work suffers from a lack of convincing figures. Mekeel does not fully describe the regression of the lung rudiment. In addition, Mekeel claims that the caudal portion of the rudimentary lung forms a vestigial endodermal tissue in the dorsal wall of the sinus venosus following regression. I find no support for this claim, and it is difficult to imagine how an endoderm-derived epithelium would become incorporated into the heart. In spite of some potential pitfalls, it is surprising that Mekeel's discovery has not been more broadly publicized.

Mekeel (1930) relied solely on developmental morphology as an indicator of lung development in plethodontids. In contrast, I demonstrate the presence of a plethodontid lung

rudiment morphologically and genetically. I find that the development of the lungs in a lunged salamander and the lung rudiment in plethodontid salamanders occurs at equivalent stages. In both instances, pulmonary development begins at the onset of gill branching (Mekeel, 1930). In addition, the LT of *A. mexicanum* and plethodontids arise at precisely the same axial location, viz., dorsal to the division between atria and sinus venosus of the heart (Figs. 2.1, 2.2). I find this transient morphological structure present in *Plethodon cinereus*, *Desmognathus fuscus* and *Hemidactylium scutatum*, which together represent a broad phylogenetic sampling of plethodontid salamanders (Pyron and Wiens, 2011; Shen et al., 2016). Lungs are specified in *P. cinereus*, as evidenced by mesenchymal expression of Wnt2b flanking the laryngotracheal groove, which is the precursor to the LT (Fig. 2.8). This same expression pattern can be observed in frogs (Rankin et al., 2015) and mice (Goss et al., 2009). Once the LT forms, it expresses the pulmonary markers Nkx2.1 and Sox9 (Figs. 2.9, 2.10). Nkx2.1 is the earliest specific marker of pulmonary identity (Lazzaro et al., 1991).

Sox9 is expressed embryonically in lung mesenchyme surrounding the trachea and in distal lung epithelium (Perl et al., 2005). Sox9 plays a role in tracheal cartilage development and distal lung morphogenesis (Chang et al., 2013; Turcatel et al., 2013). Sox9 expression in *P. cinereus* suggests a distal identity of the LT. Distal identity of the LT may hint at a mechanism of lung loss because mutations affecting lung development may act modularly to inhibit either proximal or distal development. Examples include knockouts of Shh and Bmp4, which specifically abrogate tracheal development, while partial lungs still form (Li et al., 2008; Litingtung et al., 1998). Both Shh and Bmp4 knockout phenotypes are likely mediated through downregulation of Nkx2.1 expression relative to the expression of Sox2 (Domyan et al., 2011). Shh and Bmp4 mutants stand in contrast to Fgf10 knockout models,

in which the trachea develops but lung development is mostly inhibited (Min et al., 1998). Due to the expression of Sox9 throughout the LT (Fig. 2.10), it is likely that lungless salamanders are missing the proximal portion of their lungs, corresponding to defects leading to decreased Nkx2.1 expression.

Transcriptome profiling of the LT reveals that many lung markers are expressed within this tissue in addition to the candidates assessed by *in situ* hybridization, offering further support for the molecular homology of the *A. mexicanum* lung and the plethodontid lung rudiment (Table 2.1). Lung rudiment specification, morphogenesis and expression of pulmonary markers illustrate that lung developmental pathways are conserved in plethodontids and not lost to genetic drift. These data support the conclusion that early lung development is conserved in plethodontids despite millions of years of evolutionary history since lungs were lost in adults.

The plethodontid and lunged salamander laryngotracheal tubes are homologous based on shared timing of emergence, morphology, axial position and gene expression profiles. The ultimobranchial body of the thyroid is also a transient endodermal diverticulum. However, the LT cannot be the ultimobranchial body, which fuses to the thyroid, because the LT is not derived from the pharyngeal pouch and the ultimobranchial body develops far anterolateral to the lung rudiment (Dent, 1942; Kusakabe et al., 2006; Saxén and Toivonen, 1955; Wang et al., 2011; Wilder, 1929).

Approximately three weeks after forming (when embryos are incubated at 15°C), the lung rudiment begins to regress in *P. cinereus*. Regression occurs by apoptosis, particularly in the proximal epithelium (Fig. 2.6). Since a rudiment is specified and develops in plethodontids, lung regression may represent a cessation of lung growth, a loss of lung maintenance or loss of

both processes. Alternatively, cues for active degeneration may cause developmental arrest and regression. A number of mouse mutant models recapitulate lung loss or arrest in lung development (Morrisey and Hogan, 2010). Broadly speaking, disruption of Fgf, Shh, Wnt, Tgfβ, RA and Bmp signaling can all cause partial or complete lung loss (Table 3.1). The fact that signaling pathways play an outsized role in lung development reflects the necessity of reciprocal interactions between the endoderm and underlying mesoderm to pattern the lungs (Hines and Sun, 2014; McCulley et al., 2015). The same signaling mechanisms are conserved across vertebrates, and perturbation of these pathways results in lung loss in frogs (Rankin et al., 2012). Due to the many pleiotropic roles that intercellular signaling pathways play during development, it will be interesting to uncover the genetic mechanism by which lungs are lost but other organs (such as the limbs and heart) remain unaffected (Chapter 3).

Pleiotropy may be one explanation for the conservation of lung development in plethodontids despite the absence of adult lungs. It is possible that the structure itself plays important roles in the development of other organs, such as the heart (Devi and More, 1966; Goddeeris et al., 2008; Hoffmann et al., 2009; Peng et al., 2013; Xie et al., 2012), or the gene regulatory networks governing initial lung formation are conserved because of their pleiotropic roles in the development of organs such as the limbs (e.g., Shh, Bellusci et al., 1997; Riddle et al., 1993). Previous research in multiple taxa suggests that tooth developmental modules may remain conserved and functional following evolutionary tooth loss, but that key inductive factors are missing (Harris et al., 2006; Jackman and Stock, 2006; Kurten, 1963; Mitsiadis et al., 2003, 2006; Wiens, 2011). Lung development modules may be similarly maintained in plethodontids, with cues for continued lung development perturbed, such as maintenance of the reciprocal interaction between the endoderm and surrounding splanchic mesoderm.

Despite many conserved aspects of early lung development, lungs have not been regained since the distant origin of plethodontids.

By applying both morphological and molecular insight to lung development in plethodontids, I demonstrate that lungless salamanders undergo early stages of lung development and form a lung rudiment. This rudiment regresses by apoptosis. Surprisingly, lungless salamanders express markers of lung specification and differentiation. Presence of the transient rudiment and expression of Wnt2b, Nkx2.1 and Sox9 suggest conservation of inductive interactions that govern lung formation. By studying the developmental basis of lung loss, we have an opportunity to reveal mechanisms of macroevolutionary organ loss at the base of a major evolutionary radiation. This work opens future avenues to explore the genetic basis of lung loss and determine whether lung loss in other lineages occurs by similar or different means.

2.6 References

Beachy, C.K., and Bruce, R.C. (1992). Lunglessness in plethodontid salamanders is consistent with the hypothesis of a mountain stream origin: a response to Ruben and Boucot. Am. Nat. 139, 839–847.

Bejder, L., and Hall, B.K. (2002). Limbs in whales and limblessness in other vertebrates: mechanisms of evolutionary and developmental transformation and loss. Evol. Dev. 4, 445–458.

Bellusci, S., Furuta, Y., Rush, M.G., Henderson, R., Winnier, G., and Hogan, B.L. (1997). Involvement of Sonic hedgehog (Shh) in mouse embryonic lung growth and morphogenesis. Development 124, 53–63.

Bickford, D., Iskandar, D., and Barlian, A. (2008). A lungless frog discovered on Borneo. Curr. Biol. 18, 374–375.

Bordzilovskaya, N., Dettlaff, T., Duhon, S., and Malacinski, G. (1989).

Developmental-stage series of axolotl embryos. In Developmental Biology of the Axolotl, J. Armstrong, and G. Malacinski, eds. (Oxford: Oxford University Press), pp. 201–219.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J., et al. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. Science. 327, 302–305.

Chang, D., Alanis, D., Miller, R., Ji, H., Akiyama, H., McCrea, P.D., and Chen, J. (2013). Lung epithelial branching program antagonizes alveolar differentiation. Proc. Natl. Acad. Sci. U. S. A. 110, 18042–18051.

Cohn, M.J., and Tickle, C. (1999). Developmental basis of limblessness and axial patterning in snakes. Nature 399, 474–478.

Darwin, C. (1859). On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life (London: John Murray).

Dent, J.N. (1942). The embryonic development of *Plethodon cinereus* as correlated with the differentiation and functioning of the thyroid gland. J. Morphol. 71, 577–601.

Devi, B., and More, J.R.S. (1966). Total tracheopulmonary agenesis associated with asplenia, agenesis of umbilical artery and other anomalies. Acta Paediatr. Scand. 55, 107–116.

Domyan, E.T., Ferretti, E., Throckmorton, K., Mishina, Y., Nicolis, S.K., and Sun, X. (2011). Signaling through BMP receptors promotes respiratory identity in the foregut via repression of Sox2. Development 138, 971–981.

Goddeeris, M.M., Rho, S., Petiet, A., Davenport, C.L., Johnson, G.A., Meyers, E.N., and Klingensmith, J. (2008). Intracardiac septation requires hedgehog-dependent cellular contributions from outside the heart. Development 135, 1887–1895.

Goss, A.M., Tian, Y., Tsukiyama, T., Cohen, E.D., Zhou, D., Lu, M.M., Yamaguchi, T.P., and Morrisey, E.E. (2009). Wnt2/2b and beta-catenin signaling are necessary and sufficient to specify lung progenitors in the foregut. Dev. Cell 17, 290–298.

Gross, J.B., Borowsky, R., and Tabin, C.J. (2009). A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish *Astyanax mexicanus*. PLoS Genet. 5, e1000326.

Hall, B.K. (1984). Developmental mechanisms underlying the formation of atavisms. Biol. Rev. 59, 89–122.

Harris, M.P., Hasso, S.M., Ferguson, M.W.J., and Fallon, J.F. (2006). The development of archosaurian first-generation teeth in a chicken mutant. Curr. Biol. 16, 371–377.

Harris-Johnson, K.S., Domyan, E.T., Vezina, C.M., and Sun, X. (2009). Beta-catenin promotes respiratory progenitor identity in mouse foregut. Proc. Natl. Acad. Sci. U. S. A. 106, 16287–16292.

Herrera, A.M., Shuster, S.G., Perriton, C.L., and Cohn, M.J. (2013). Developmental basis of phallus reduction during bird evolution. Curr. Biol. 23, 1065–1074.

Hines, E.A., and Sun, X. (2014). Tissue crosstalk in lung development. J. Cell. Biochem. 115, 1469–1477.

Hoffmann, A.D., Peterson, M.A., Friedland-Little, J.M., Anderson, S.A., and Moskowitz, I.P. (2009). Sonic hedgehog is required in pulmonary endoderm for atrial septation. Development 136, 1761–1770.

Hurney, C.A., Babcock, S.K., Shook, D.R., Pelletier, T.M., Turner, S.D., Maturo, J., Cogbill, S., Snow, M.C., and Kinch, K. (2015). Normal table of embryonic development in the four-toed salamander, *Hemidactylium scutatum*. Mech. Dev. 136, 99–110.

Jackman, W.R., and Stock, D.W. (2006). Transgenic analysis of Dlx regulation in fish tooth development reveals evolutionary retention of enhancer function despite organ loss. Proc. Natl. Acad. Sci. U. S. A. 103, 19390–19395.

Kerney, R. (2011). Embryonic staging table for a direct-developing salamander, *Plethodon cinereus* (Plethodontidae). Anat. Rec. 294, 1796–1808.

Kurten, B. (1963). Return of a lost structure in the evolution of the felid dentition. Comment. Biol. Soc. Sci. Fenn. 26, 1–12.

Kusakabe, T., Hoshi, N., and Kimura, S. (2006). Origin of the ultimobranchial body cyst: T/ebp/Nkx2.1 expression Is required for development and fusion of the ultimobranchial body to the thyroid. Dev. Dyn. 235, 1300–1309.

Lazzaro, D., Price, M., de Felice, M., and Di Lauro, R. (1991). The transcription factor TTF-1 is expressed at the onset of thyroid and lung morphogenesis and in restricted regions of the foetal brain. Development 113, 1093–1104.

Li, M., Li, C., Liu, Y., Xing, Y., Hu, L., Borok, Z., Kwong, K.Y.-C., and Minoo, P. (2008). Mesodermal deletion of transforming growth factor-beta receptor II disrupts lung epithelial morphogenesis: cross-talk between TGF-beta and Sonic hedgehog pathways. J. Biol. Chem. 283, 36257–36264.

Litingtung, Y., Lei, L., Westphal, H., and Chiang, C. (1998). Sonic hedgehog is essential to foregut development. Nat. Genet. 20, 58–61.

Ludike, R. (1955). Über den respirationsapparat vershiedener urodelen und seine beziehungen zum herzen. Z. Morph. U. Okol. Tiere, Bd. 43, 578–615.

McCulley, D., Wienhold, M., and Sun, X. (2015). The pulmonary mesenchyme directs lung development. Curr. Opin. Genet. Dev. 32, 98–105.

Mekeel, A.G. (1926). A pulmonary vestige in the lungless salamanders. Anat. Rec. 34, 141.

Mekeel, A.G. (1930). Pulmonary development in the lungless salamanders. Unpubl. PhD Diss. Cornell Univ. 1–115.

Metscher, B.D. (2011). X-ray microtomographic imaging of intact vertebrate embryos. Cold Spring Harb. Protoc. 2011, 1462–1471.

Min, H., Danilenko, D.M., Scully, S. a., Bolon, B., Ring, B.D., Tarpley, J.E., DeRose, M., and Simonet, W.S. (1998). Fgf-10 is required for both limb and lung development and exhibits striking functional similarity to *Drosophila* branchless. Genes Dev. 12, 3156–3161.

Minoo, P., Su, G., Drum, H., Bringas, P., and Kimura, S. (1999). Defects in tracheoesophageal and lung morphogenesis in Nkx2.1(-/-) mouse embryos. Dev. Biol. 209, 60–71.

Mitsiadis, T.A., Chéraud, Y., Sharpe, P., and Fontaine-Pérus, J. (2003). Development of teeth in chick embryos after mouse neural crest transplantations. Proc. Natl. Acad. Sci. U. S. A. 100, 6541–6545.

Mitsiadis, T.A., Caton, J., and Cobourne, M. (2006). Waking-up the sleeping beauty: recovery of the ancestral bird odontogenic program. J. Exp. Zool. B. Mol. Dev. Evol. 306, 227–233.

Morrisey, E.E., and Hogan, B.L.M. (2010). Preparing for the first breath: genetic and cellular mechanisms in lung development. Dev. Cell 18, 8–23.

Noble, G.K. (1925). The integumentary, pulmonary, and cardiac modifications correlated with increased cutaneous respiration in the amphibia: A solution of the "hairy frog" problem. J. Morphol. 40, 341–416.

Nussbaum, R.A., and Wilkinson, M. (1995). A new genus of lungless tetrapod: A radically divergent caecilian (Amphibia: Gymnophiona). Proc. R. Soc. London. Ser. B Biol. Sci. 261, 331–335.

Peng, T., Tian, Y., Boogerd, C.J., Lu, M.M., Kadzik, R.S., Stewart, K.M., Evans, S.M., and Morrisey, E.E. (2013). Coordination of heart and lung co-development by a multipotent cardiopulmonary progenitor. Nature 500, 589–592.

Perl, A.-K.T., Kist, R., Shan, Z., Scherer, G., and Whitsett, J.A. (2005). Normal lung development and function after Sox9 inactivation in the respiratory epithelium. Genesis 41, 23–32.

Presnell, J.K., Schreibman, M.P., and Humason, G.L. (1997). Humason's Animal Tissue

Techniques (Baltimore: Johns Hopkins University Press).

Pyron, R.A., and Wiens, J.J. (2011). A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. Mol. Phylogenet. Evol. 61, 543–583.

Rahmani, T.M. (1974). Morphogenesis of the rudimentary hind-limb of the Glass Snake (*Ophisaurus apodus* Pallas). J. Embryol. Exp. Morphol. 32, 431–443.

Rankin, S.A., Gallas, A.L., Neto, A., Gómez-Skarmeta, J.L., and Zorn, A.M. (2012). Suppression of Bmp4 signaling by the zinc-finger repressors Osr1 and Osr2 is required for Wnt/β-catenin-mediated lung specification in *Xenopus*. Development 139, 3010–3020.

Rankin, S.A., Thi Tran, H., Wlizla, M., Mancini, P., Shifley, E.T., Bloor, S.D., Han, L., Vleminckx, K., Wert, S.E., and Zorn, A.M. (2015). A molecular atlas of *Xenopus* respiratory system development. Dev. Dyn. 244, 69–85.

Reagan, N.L., and Verrell, P.A. (1991). The evolution of plethodontid salamanders: did terrestrial mating facilitate lunglessness? Am. Nat. 138, 1307–1313.

Riddle, R.D., Johnson, R.L., Laufer, E., and Tabin, C. (1993). Sonic hedgehog mediates the polarizing activity of the ZPA. Cell 75, 1401–1416.

Ruben, J.A., Reagan, N.L., Verrell, P.A., and Boucot, A.J. (1993). Plethodontid salamander origins: a response to Beachy and Bruce. Am. Nat. 142, 1038–1051.

Saxén, L., and Toivonen, S. (1955). The development of the ultimobranchial body in *Xenopus laevis* Daudin and its relation to the thyroid gland and epithelial bodies. J. Embryol. Exp. Morphol. 3, 376–384.

Sedmera, D., Misek, I., and Klima, M. (1997). On the development of cetacean extremities: I. Hind limb rudimentation in the spotted dolphin (*Stenella attenuata*). Eur. J. Morphol. 35, 25–30.

Shen, X.-X., Liang, D., Chen, M.-Y., Mao, R.-L., Wake, D.B., and Zhang, P. (2016). Phylogeny, time and biogeography of plethodontids. Syst. Biol. 65, 66–81.

Thewissen, J., Cohn, M.J., Stevens, L.S., Bajpai, S., Heyning, J., and Horton Jr., W.E. (2006). Developmental basis for hind-limb loss in dolphins and origin of the cetacean bodyplan. Proc. Natl. Acad. Sci. U. S. A. 103, 8414–8418.

Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M. A, and Quake, S.R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature 509, 371–375.

Turcatel, G., Rubin, N., Menke, D.B., Martin, G., Shi, W., and Warburton, D. (2013). Lung mesenchymal expression of Sox9 plays a critical role in tracheal development. BMC

Biol. 11, 117.

Wake, D.B., and Hanken, J. (1996). Direct development in the lungless salamanders: what are the consequences for developmental biology, evolution and phylogenesis? Int. J. Dev. Biol. 40, 859–869.

Wang, J.H., Deimling, S.J., D'Alessandro, N.E., Zhao, L., Possmayer, F., and Drysdale, T.A. (2011). Retinoic acid is a key regulatory switch determining the difference between lung and thyroid fates in *Xenopus laevis*. BMC Dev. Biol. 11, 75.

Wiens, J.J. (2011). Re-evolution of lost mandibular teeth in frogs after more than 200 million years, and re-evaluating Dollo's law. Evolution. 65, 1283–1296.

Wilder, H.H. (1896). Lungless salamanders. Anat. Anz. 12, 182–192.

Wilder, M.C. (1929). The significance of the ultimobranchial body (postbranchial body, suprapericardial body): A comparative study of its occurrence in urodeles. J. Morphol. 47, 283–333.

Wilder, I.W., and Dunn, E. (1920). The correlation of lunglessness in salamanders with a mountain brook habitat. Copeia 1920, 63–68.

Xie, L., Hoffmann, A.D., Burnicka-Turek, O., Friedland-Little, J.M., Zhang, K., and Moskowitz, I.P. (2012). Tbx5-hedgehog molecular networks are essential in the second heart field for atrial septation. Dev. Cell 23, 280–291.

Chapter 3

Towards a mechanistic understanding of lung loss

3.1 Abstract

Lungless salamanders develop an embryonic lung rudiment, but although some aspects of early lung development are conserved in lungless species the factors responsible for regression of the rudiment and adult lung loss are unknown. To understand the genetic changes driving lung loss I compared the transcriptome of the lung rudiment between lungless and lunged salamanders. I found that the expression of several genes and signaling pathways are significantly different in lungless salamander lung primordia. Among these, Tgfβ-signaling is strongly upregulated. Administration of exogenous Tgfβ in lunged salamander embryos results in impaired lung development, while reduction of Tgfβ signaling expands lung specification. These data support a role for Tgfβ upregulation in lung loss. Despite adult lunglessness, plethodontid salamanders conserve some aspects of early lung development,

indicating that they may be competent to develop lungs. By transplantation of mesoderm from lunged salamanders into lungless salamander embryos I demonstrate the formation of ectopic structures resembling lungs, suggesting that lung loss is due to loss of inductive cues from the pulmonary mesoderm. My study provides insight into the mechanism of lung loss in plethodontid salamanders. Evolutionary changes in the pulmonary gene regulatory network may result in lung loss, while avoiding detrimental effects stemming from coding mutations within pleiotropic lung-patterning genes.

3.2 Introduction

"Why lose the lungs?" Scientists have pondered this question since the discovery of lungless salamanders in the late 1800s (Lonnberg, 1899). A satisfactory answer to this question is still lacking more than 100 years later. A no less interesting, and perhaps more tractable, question is: "How have salamanders lost their lungs?" For instance, what are the developmental mechanisms responsible for the lack of lungs in adults? Progress on these questions is hampered by the fact that plethodontid salamanders are not model organisms and few experimental genetic or genomic resources exist for them. Nonetheless, lung loss represents a major evolutionary transition with corresponding and pronounced changes to organismal physiology and ecology. A better understanding of how lung loss occurs will offer a window into the genetics of macroevolutionary transitions.

Organ loss is often associated with major evolutionary transitions, such as limb loss in snakes or tooth loss in birds. A frequent impediment to studying such losses is the lack of embryonic rudiments of lost organs, forcing reliance on genomic signatures of loss and candi-

date gene approaches. In contrast, studying organ loss when a transient embryonic rudiment is present offers a distinct advantage: the gene expression pattern within the transient structure can be analyzed directly. Therefore, embryonic rudiments are not only among the most telling morphological indications of past evolutionary history (Darwin, 1859), but they can be leveraged to reveal the mechanisms underlying evolutionary organ loss.

I present evidence that embryonic plethodontid salamanders form a pulmonary rudiment (Chapter 2). The lungs are correctly specified in the lungless species *Plethodon cinereus*; a diverticulum from the foregut endoderm emerges and begins to express markers of pulmonary identity. The lung rudiment in plethodontids regresses by apoptosis and this degeneration results in the absence of lungs in adults. Despite the remarkable conservation of at least some aspects of the early lung developmental program in plethodontids, unknown genetic differences between lunged and lungless salamanders result in lung loss. Comparing the pulmonary gene regulatory networks of plethodontids and lunged salamanders may enable the identification of genes involved in lung loss.

The process of lung development is remarkably conserved across tetrapods (Rankin and Zorn, 2014; Rankin et al., 2015). Lung specification, outgrowth, and differentiation are governed by a well-characterized gene regulatory network dependent on signaling interactions between epithelium and mesenchyme (Table 3.1) (Alescio and Cassini, 1962; Morrisey and Hogan, 2010; Rankin and Zorn, 2014; Spooner and Wessells, 1970). Among the earliest embryonic events necessary for lung development is establishing the future pulmonary domain of the foregut. Fibroblast growth factors (Fgfs) secreted from the cardiac mesoderm pattern the foregut in a concentration-dependent fashion: high levels of Fgfs help specify pulmonary domains within the foregut, while lower levels (farther posterior from the Fgf source) activate

liver developmental programs (Serls et al., 2005). Fgfs continue to play roles in later lung development, as discussed below.

Retinoic acid (RA) plays a pleiotropic role in both early and late lung development. RA promotes canonical Wnt signaling through the suppression of Wnt inhibitor Dickkopf1 and functions to repress Transforming growth factor-β (Tgfβ) signaling (Chen et al., 2007, 2010). RA also acts indirectly on *wnt* expression, by activating factors that in turn activate *wnt* expression, such as *tbx5* (Ryckebusch et al., 2008). Tbx5 can activate Wnt signaling and *fgf10* expression (Arora et al., 2012). Therefore, one of the main early roles of RA is a permissive one: creating an embryonic region where Fgfs and Wnt can be expressed and initiate lung development (Rankin et al., 2012).

The Tgfβ superfamily encompasses several signaling pathways, such as bone morphogenetic protein (Bmp), Tgfβ, Nodal, Gdf and Activin signaling (Massagué and Gomis, 2006). Signaling network topology is fairly conserved between these different ligand classes, such as their shared utilization of dual receptors and Smad proteins for signal transduction, though ligand function and expression pattern is highly diverse. All Tgfβ superfamily receptors function as heteromeres composed of Type I and Type II receptors (de Caestecker, 2004). Ligand affinity varies, but generally Tgfβ binds the type II receptor TgfβrII and then recruits type I receptor Alk5. Activin binds to the type II receptor Acvr2 before recruiting Alk4. Bmp ligands Bmp4 and Bmp2 tend to first bind to type I receptors (Alk3 and Alk6) before recruiting type II receptors such as Bmpr2 (de Caestecker, 2004). However, ligands within the Tgfβ superfamily often overlap in use of receptors, providing some flexibility and complexity.

Ligand binding results in phosphorylation of Smad proteins (homologs of mothers against

decapentaplegic in *Drosophila*). Smads are usually classified into three groups: a receptor-regulated group (R-Smads; Smads 1,2,3,5,8), a co-mediator Smad group (co-Smads; Smad and Smad4), and the inhibitory Smads (I-Smads; Smad6 and Smad7). Bmp signaling results in phosphorylation of the r-Smads 1, 5, and 8, while Tgfβ signaling results in Smad2 and Smad3 phosphorylation (Dijke and Hill, 2004). All r-Smads interact with Smad4 and can activate transcription within this heteromeric complex, provided they translocate to the nucleus (Kitisin et al., 2007).

Among Tgfβ superfamily ligands, Tgfβ and Gdf8 are unique in that they are synthesized in an inactive form. Synthesized Tgfβ peptide self-dimerizes and then is cleaved to form the small latent complex made up of the C-terminal mature peptide and the N-terminal latency-associated propeptide (LAP). Typically, the inactive small latent complex is bound by one of the latent Tgfβ binding proteins (Ltbps), forming the large latent complex. This complex of Tgfβ, LAP and Ltbp is then secreted from the cell and deposited in the extracellular matrix (ECM) anchored by Ltbp (Hayashi and Sakai, 2012; Munger et al., 1997). Accumulated Tgfβ in the ECM then can be rapidly and specifically released to affect target cells (Mazzieri et al. 2005; Munger et al. 1997). Ltbp is crucial for most Tgfβ signaling; reduced levels of Ltbp results in decreased Tgfβ activity (Koli et al., 2004; Yoshinaga et al., 2008).

Tgfβ, when ectopically administered or upregulated, represses lung development through the repression of *fgf10* expression and repression of *nkx2.1* expression, resulting in the formation of a developmentally arrested endodermal diverticulum but no lung formation (Chen et al., 2007). Tgfβ also functions to repress later lung development when ectopically expressed in developing alveoli (Zhou et al., 1996) or when administered in culture (Serra et al., 1994).

The balance of canonical Wnt signaling is critically important for the development of

the lungs. In particular, Wnt2, Wnt2b and Wnt7b govern lung specification and development (Goss et al., 2009; Rajagopal et al., 2008). Canonical Wnt signaling depends on the nuclear accumulation of β-catenin, which recruits transcription factors to activate target genes and its own negative feedback loop. In the absence of bound Wnt ligand, β-catenin is phosphorylated by the β-catenin destruction complex, then subsequently ubiquitinated and proteolytically degraded. The heteromeric β-catenin destruction complex is composed of Axin, adenomatous polyposis coli (Apc), Gsk-3β, and Ck1α. Wnt ligand binds to heterodimer receptors composed of a Frizzled protein and a low-density lipoprotein receptor related protein (Lrp5/6 in vertebrates). Binding of Wnt results in the inactivation of the destruction complex, leading to accumulation of β-catenin and activation of target gene expression (reviewed by Volckaert and De Langhe, 2015).

Wnt/β-catenin signaling via Wnt2/2b is necessary and sufficient for lung specification (Goss et al., 2009; Harris-Johnson et al., 2009). While single mutants of Wnt2 have lung hypoplasia, the combined effect of Wnt2 and Wnt2b knockout is complete lung agenesis (Goss et al., 2009). Similar to Wnt2 mutants, Wnt7b mutants also possess hypoplastic lungs, though lung anatomy is mostly intact, indicating a role for Wnt7b in epithelial and mesenchymal development and proliferation (Rajagopal et al., 2008). The effects of Wnt2/2b are mediated by β-catenin, as evidenced by endoderm-specific conditional knockouts (Goss et al., 2009; Harris-Johnson et al., 2009). Stabilization of β-catenin (activation of Wnt signaling) in the endoderm results in expansion of pulmonary fate into the esophagus and the stomach (Goss et al., 2009). Wnt is also required in pulmonary proximal-distal development and differentiation later in development (Shu et al., 2005).

Lung and tracheal development share many developmental mechanisms, but phenotypes

from several knockout models are confined to one structure or another (Table 3.1). For instance, Fgf10 knockouts appear to develop a partial trachea but no lungs, while Bmp or Shh signaling deficiency results in the formation of lungs but no trachea. Fgf10 expression is required for the formation of the lung buds (Sekine et al., 1999) and this role appears to be conserved in frogs (Rankin et al., 2012). The expression pattern of Fgf10 in mesenchymal foci underlying the site of lung bud emergence and Fgf10 orthology to the well-characterized *Drosophila* tracheal development gene branchless is suggestive of a role directing proximal-distal outgrowth and branching morphogenesis (Metzger and Krasnow, 1999; Metzger et al., 2008). However, recent work points to Fgf10 as a permissive signal in lung development, where focal expression domains are not required for outgrowth (Volckaert et al., 2013).

Although the evolution of lunglessness marks a major morphological and physiological transition in amphibians, nothing is known about the genes that are responsible for lung regression. In plethodontid salamanders, the lung begins to develop embryonically during the same developmental window and at the same axial location as in lunged salamanders (Chapter 2). I observed expression of *wnt2b* surrounding the presumptive plethodontid lung primordium. Wnt2b acts to specify the lungs in mammals and frogs (Rankin et al., 2012). Later in development, the lung diverticulum expresses the pulmonary marker *nkx2.1* and the marker of distal pulmonary identity, *sox9* (Chapter 2). Expression of Sox9 throughout the lung rudiment is evidence that the tissue has taken on an alveolar-, or distal-type identity.

While candidate genes can reveal much about the process of development and regression of the lungs in plethodontids, unbiased sequencing approaches have the potential to reveal in greater detail the molecular interactions and global expression pattern of genes within the lung rudiment. Provided that lung loss occurs through differential gene regulation and

not wholesale gene loss or coding mutations, transcriptomics may reveal the mechanism underlying lung loss.

To determine what genes drive the loss of lungs in plethodontid salamanders I sequenced the transcriptome of the lung rudiment in one species of lungless salamander (*Plethodon cinereus*) and compared these transcriptome libraries to comparable samples from the lunged species *Ambystoma mexicanum*. I find that numerous genes implicated in lung development are differentially expressed in *P. cinereus* relative to *A. mexicanum*. Key differences include upregulation of Wnt signaling and upregulation of Tgfβ signaling. Increased Tgfβ signaling as well as decreased Shh signaling and levels of *foxa1* may result in the observed downregulation of the critical pulmonary gene *nkx2.1*, potentially resulting in lung loss. I find that administration of ectopic Tgfβ to lunged salamanders represses pulmonary development, while inhibition of Tgfβ signaling expands respiratory specification. Finally, epithelial-mesenchymal signaling interactions are implicated in lung loss by the probable rescue of lung development in a lungless species by transplantation of splanchnic mesenchyme derived from a lunged donor species.

3.3 Materials and methods

3.3.1 Animal husbandry

All plethodontid embryos were collected under Massachusetts Division of Fisheries and Wildlife permits and local permits, where applicable: DFW permit numbers 181.10SCRA (2010), 080.11SCRA (2011), 080.11SCRA (2012), 027.13SCRA (2013), 083.14SCRA (2014),

Table 3.1: Lung, trachea, and esophagus presence (+) or absence (–) in mouse knockout, conditional knockout and downregulation studies.

Gene	Expression	Lung	Trachea	Esophagus	References
Wnt2/2b (compound)	Mesenchyme	–	–	+	Goss et al. 2009
β-catenin (Conditional knockout using Shh-Cre)	Epithelium	–	–	+	Goss et al. 2009, Harris-Johnson et al. 2009
Tankyrase (through XAV939 treatment)	Epithelium	–	–	+	Rankin et al. 2012
Fgf10	Mesenchyme at site of branch outgrowth	– [a]	+	+	Sekine et al. 1999, Min et al. 1998
Fgfr2-IIIb	Epithelium	– [e]	– [g]	+	De Moerlooze et al. 2000
Shh	Epithelium	+ [b,c]	–	+	Litingtung et al. 1998
Hhip	Mesenchyme	+ [b,c]	+	+	Chuang and McMahon 1999, Chuang et al. 2003
Gli3 (single)	Mesenchyme	+[b,c]	+	+	Grindley et al. 1997
Gli2 (single)	Mesenchyme	+[b,c]	+[b]	+[b]	Motoyama et al. 1998
Gli2/3 (compound)[f]	Mesenchyme	–	–	–	Motoyama et al. 1998
Nkx2.1	Epithelium	+[h]	+[b]	+	Minoo et al. 1999, Kimura et al. 1996, 1999
Sox2	Epithelium	+	+	–	Que et al. 2007
Bmp4 (Conditional using Foxg1-Cre)	Mesenchyme	+[b]	–	+	Li et al. 2008
Bmpr1a/b	Epithelium	+	–	+	Domyan et al. 2011
Osr1/2	Epithelium	–	–	+	Rankin et al. 2012
Noggin	Mesenchyme	+[b]	+	–	Que et al. 2006
RA signaling (reduced through acute retinoid deprivation)		–	– [d]	– [d]	Dickman et al. 1997

[a]Small primary buds do form (Gail Martin, personal communication).
[b]Reduced size.
[c]Impaired branching.
[d]Laryngotracheal groove forms, but tracheoesophageal septation fails.
[e]Lung formation, then regression. Presence of a band of remaining mesenchyme, which regresses by apoptosis.
[f]Compound Gli2/3 mutants display different lung abnormalities depending on which gene is absent or reduced. Double-knockout mice completely lack expression in lungs, esophagus, and trachea, as well as other anterior foregut-derived structures.
[g]A "tracheal bud" is present, but it is mainly composed of mesenchyme and appears to be highly apoptotic at E10.5.
[h]Little to no proximal-distal differentiation of the lung.

and 022.15SCRA (2015); Cape Cod National Seashore permit number CACO-00214 (*Hemidactylium scutatum*). Locality data is provided in Appendix A, Table A.1. *Plethodon cinereus* embryos were typically found underneath moss atop fallen hemlock logs or inside rotting hemlock logs. *Hemidactylium scutatum* embryos were found within clumps of sphagnum moss overlying fresh waterways.

Embryos were kept at 15–17°C on filter paper moistened with 0.1X MMR (Marc's Modified Ringer solution: 0.01 M NaCl, 0.2 mM KCl, 0.1 mM $MgSO_4$, 0.2 mM $CaCl_2$, 0.5 mM HEPES pH 7.4) and 100 µg/ml gentamicin (Sigma, St. Louis, MO) or fully immersed in the same solution. Fungal infections were treated by dechorionating embryos and raising them in solution. Embryos were monitored daily for fungal infections and to check their developmental stage. Staging was performed as described for *Plethodon cinereus* (Kerney 2011) and for *Hemidactylium scutatum* (Hurney et al. 2015).

Ambystoma mexicanum embryos were obtained from the *Ambystoma* Genetic Stock Center, University of Kentucky. Embryos were reared at 17°C immersed in 20% Holtfreter solution and staged according to Bordzilovskaya et al. (1989).

3.3.2 RNA isolation, library preparation and sequencing

See Appendix B.

3.3.3 Bioinformatics

See Appendix B.

3.3.4 Explants

Ambystoma mexicanum explants were generated by transversely cutting stage-34 embryos anterior to the outflow tract of the heart and posterior to the posterior terminus of the liver. Surgery was performed under anesthetic (MS-222, Sigma, St. Louis, MO) in sterile 0.1x MMR. Explants were maintained in 50% L-15 medium (Sigma) containing 500 µg/ml gentamicin in 12-well tissue culture plates at 17°C. 50 mg/ml Tgfβ1 (Peprotech #100-21, Rocky Hill, NJ) or 50 mg/ml BSA (Sigma) were added to the solution. Solutions were changed every 48 hr. Explant health was assessed by looking for heartbeats. Heartbeats were maintained throughout the 10-day experiment until stage 42. End staging was assessed from limb bud morphology. After fixation in MEMFA (0.1 M MOPS (pH 7.4), 2 mM EGTA, 1 mM $MgSO_4$, 3.7% formaldehyde), explants were examined for SPC expression using wholemount *in situ* hybridization.

3.3.5 Chemical treatments

Chemical antagonist treatments were carried out on dechorionated embryos in 12-well tissue culture plates. The Alk5 antagonist A-83-01 (Tocris #2939, Bristol, UK) was utilized at a final concentration of 10 µM (for *A. mexicanum*) or 20 µM (for *P. cinereus*), diluted in sterile 0.1X MMR with 100 µg/ml gentamicin (Sigma). A dimethyl sulfoxide (DMSO) control with 2 µl/ml DMSO was run in parallel. Full solution changes were completed every 48 hr. For *P. cinereus*, embryos were treated from late stage 19 until stage 23 (a period of 26 days). After fixation, some embryos were stained using wholemount *in situ* hybridization for *sox9* expression, while others were histologically sectioned and stained. For *A. mexicanum*,

embryos were treated from stage 33 through 43 (10 days), with solution refreshed every 48 hr.

The Tankyrase inhibitor (hence Wnt/β-catenin inhibitor) XAV939 (Sigma) was administered to *A. mexicanum* embryos at 50 µM concentration from stage 31 through 43 (11 days) alongside a DMSO control. XAV939 treatment resulted in several gross phenotypic defects including repression of limb development, shorter gills than controls, and slight cranial malformations. XAV939 solutions were changed every 48 hr until fixation. Treated embryos were examined using histology and SPC *in situ* hybridization.

Additional chemical treatments yielded no lung phenotypes in axolotl or in plethodontids. These included: (1) Treatment with all-*trans* retinoic acid at 0.01 µM and 0.05 µM from stages 33 to 41 in *A. mexicanum* or treatment with 0.1 µM and 1 µM all-trans RA from stages 20 through 23 in *P. cinereus*. (2) Treatment with the competitive substrate for aldehyde dehydrogenase (Aldh) N,N-diethylamniobenzaldehyde (DEAB; Sigma) at 0.05 µM (from stages 33 to 41 in *A. mexicanum*) and 10 µM (from stages 26 to 41 in *A. mexicanum*) (Chute et al., 2006; Koppaka et al., 2012; Russo et al., 1988). (3) Treatment with 20 µM LDN193189 (Sigma), an antagonist of Alk2 and Alk3, from stages 33 to 41 in *A. mexicanum* and various stages in plethodontids. LDN193189 treatment resulted in digit malformations and extreme overgrowth of tail fin epithelium, likely due to repression of BMP-mediated interdigital apoptosis and tail fin apoptosis. (4) Treatment with the activator of BMP signaling 4'-Hydroxychalcone (VWR, Atlanta, GA) (Vrijens et al., 2013) in *A. mexicanum*, *H. scutatum*, and *Eurycea bislineata*. No gross phenotypes were observed.

3.3.6 *In situ* hybridization

DIG-labeled mRNA *in situ* hybridization probes were generated using the clones described in Chapter 1 (*P. cinereus* Sox9) and Chapter 4 (*A. mexicanum* SPC). Embryos were fixed overnight in MEMFA at 4°C, dehydrated and stored in 70 or 100% MeOH at −20°C. Whole-mount mRNA *in situ* hybridization (ISH) was performed by rehydrating samples, then samples were treated with between 5 and 10 µg/ml proteinase K for 30 to 60 min., washed with PBTw (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 0.2% Tween-20), post-fixed in 4% paraformaldehyde, washed with PBTw, and pre-hybridized in hybridization buffer for 2 hr at 65°C (hybridization buffer: 50% formamide, 5x SSC, 0.1 mg/ml heparin, 1x Denhardt's, 0.01% CHAPS, 0.2 mg/ml tRNA, 0.1% Tween-20; all solutions RNase-free). DIG-labeled riboprobes were diluted approximately 1:40 in hybridization buffer then denatured at 85°C for 10 min, before adding to specimens. Hybridization was carried out overnight at 65°C. Posthybridization washes were performed with a solution of 50% formamide, 5x SSC, and 0.2% Tween-20 at 65°C for 8 changes of 30 min each. Specimens were washed with maleic acid buffer plus 0.2% Tween-20 (MABT) prior to blocking and antibody incubation. Antibody block solution included 20% heat-inactivated sheep serum, 2% blocking reagent (Roche, Penzberg, Germany) in MABT. Samples were incubated with 1:2500 anti-DIG-AP Fab fragments (Roche) diluted in blocking solution overnight at 4°C. Extensive washes with MABT were performed prior to color development using BM-Purple (Roche) or NBT/BCIP (Sigma). Color development occurred over several hours and was stopped by administration of 1 mM EDTA in PBS, followed by postfixing in 4% PFA. In the case of Sox9 *in situ* hybridization and for analysis of SPC expression in XAV939-treated

embryos, embryos were then embedded for cryosectioning at 14–16 µm thickness. For examination of wholemount SPC expression, embryos were dehydrated into 100% methanol, then cleared and imaged in a 1:2 benzyl alcohol to benzyl benzoate solution (Sigma).

3.3.7 Bead implantation

Affi-gel blue beads at 100–200 mesh size (Bio-Rad, Hercules, CA) were washed three times with sterile phosphate-buffered saline using centrifuge filter columns. A small amount of washed beads were placed in a 5-µg/ml solution of recombinant Tgfβ1 (Peprotech) or BSA (Sigma) and soaked overnight at 4°C. The following day, embryos were dechorionated, placed into a large droplet of 3% methylcellulose in 0.1x MMR (Sigma), then anesthetized with MS-222 (Sigma). Using sterilized freshly sharpened tungsten needles, a small incision was made ventral to the limb bud and dorsal to the pericardium. A bead was inserted through this incision and manipulated until it was lodged between the pericardium and the foregut endoderm. Embryos were allowed to heal for approximately 5 min and then placed in 0.1X MMR with 100-µg/ml gentamicin in a 12-well dish. Survivorship was near 100%. Lung growth was assayed by SPC wholemount *in situ* hybridization.

3.3.8 Heterospecific transplantation

Stage-matched *Hemidactylium scutatum* and ubiquitously expressing GFP+ *Ambystoma mexicanum* individuals at 3–4 somites stage (stage 18 in *H. scutatum*—Hurney et al., 2014; stage 19 in *A. mexicanum*—Bordzilovskaya et al., 1989) were placed alongside each other on an agarose plate in 100% Holtfreter solution. Lateral plate mesoderm from *A. mexicanum*

embryos was removed by peeling back the ectoderm using freshly sharpened tungsten needles and then making a rectangular incision through the lateral plate. A comparable piece of lateral plate was excised from *H. scutatum* and replaced with GFP+ *A. mexicanum* lateral plate. Embryos were allowed to heal for approximately 60 min, then transferred to individual wells of a 12-well dish and maintained in 0.1X MMR plus 100-µg/ml gentamicin. Mock transplants were performed in parallel: the ectoderm was removed from *H. scutatum* embryos, and then the LPM was excised and then transplanted back in place. Photographs were taken 1–2 days post-transplantation and then at fixation at stage 28 or later (30 days post-transplantation) (Hurney et al., 2015). Individuals were cryosectioned at 14–16 µm thickness and then processed for immunohistochemistry (Piekarski et al., 2014). Briefly, sections were blocked and then incubated with an anti-GFP primary antibody (Abcam ab290, Cambridge, MA), followed by labeling with Alexa-Fluor 488 secondary antibody (Thermo Fisher #Z25302, Grand Island, NY) and nuclear staining with DAPI (Sigma). Failure to obtain a reliable antibody against pulmonary markers prevented molecular analysis, but graft recipients were assayed for morphology of the foregut.

3.3.9 Histology and imaging

Histology was performed as described in Chapter 2. All imaging was with performed with a Leica DMRE microscope (Wetzlar, Germany) for sections or a Leica MZFLIII for whole-mounts. Images were acquired with a QImaging Retiga 2000r camera and a QImaging RGB slider (Model: RGB-HM-S-IR) (Surrey, Canada) and Volocity 6.0 software (PerkinElmer, Waltham, MA). Adjustment of exposure, levels and color balance were performed in Adobe

Photoshop CS5.

3.4 Results

3.4.1 Lung primordia transcriptomes

With the goal of understanding the mechanism underlying the arrest and regression of the lung rudiment in lungless salamanders, I sequenced the transcriptome of the lung rudiment from *Plethodon cinereus* at two time points: during early lung rudiment development (stage 19; "Pc19") and just prior to morphological regression (stage 21; "Pc21"). Comparable transcriptome libraries were generated for lunged *Ambystoma mexicanum* embryos at stage 40 ("Am40"), after the emergence of the laryngotracheal tube and lung buds. Am40 is a roughly comparable stage to Pc19. Three libraries were generated at each time point. I generated an additional library from the lungs of a sub-adult *A. mexicanum* ("Am53") in order to better assemble pulmonary markers and potential markers not expressed embryonically. Embryonic RNA was isolated by microdissection of unfixed embedded frozen tissue under magnification on dry ice. The endoderm of the lung primordia as well as surrounding mesenchyme were removed using fine glass needles; extracted RNA was prepared for Illumina sequencing as described in Appendix B.

De novo assembly of *P. cinereus* and *A. mexicanum* transcriptomes was completed using Trinity (Grabherr et al., 2011; Haas et al., 2013). Orthologs between species were identified using reciprocal best hit BLAST. Principal components analysis (PCA) using the top 500 differentially expressed genes revealed that the first principal component axis separating

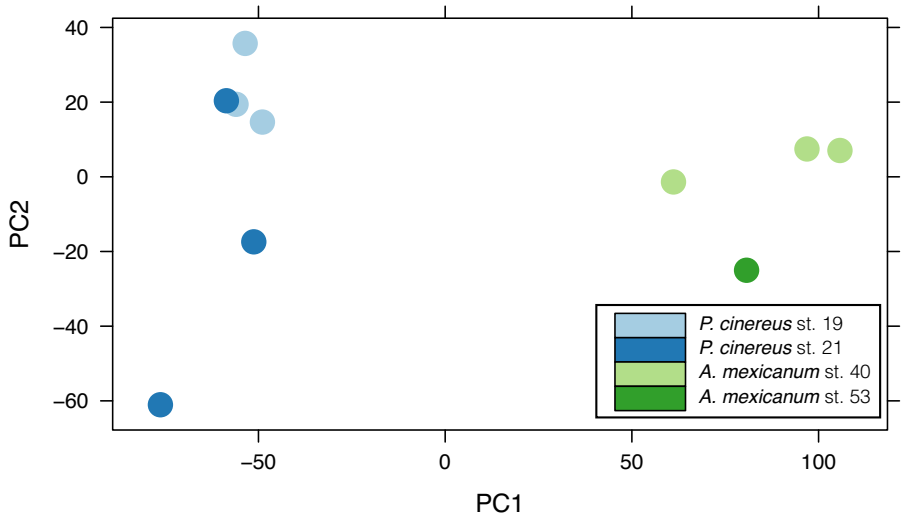


Figure 3.1: Principal-components (PC) analysis using the top 500 differentially expressed genes between lung primordia transcriptome libraries. Libraries separate by species along PC axis 1 (PC1); PC2 corresponds approximately to developmental stage. Analysis completed using the plotPCA function within DESeq (Anders and Huber, 2010).

libraries corresponded to species. The second axis corresponds to developmental stage (Fig. 3.1). There exists slightly higher within group heterogeneity in *P. cinereus* stage 21 libraries (Pc21). Hierarchical clustering based on the top 100 differentially expressed transcripts correctly splits the groups by stage and species, except in the case of *P. cinereus* at stage 21, which is paraphyletically clustered; a result of expression divergence within library Pc21_2 (Fig. 3.2). This library (Pc21_2) still clusters with its conspecific libraries (Fig. 3.1) and displays a very similar expression profile to other *P. cinereus* libraries in the heatmap (Fig. 3.2). Therefore, as judged by PCA and hierarchical clustering, the sequenced transcriptomes for the lung primordia from both species are internally consistent with no substantial biases.

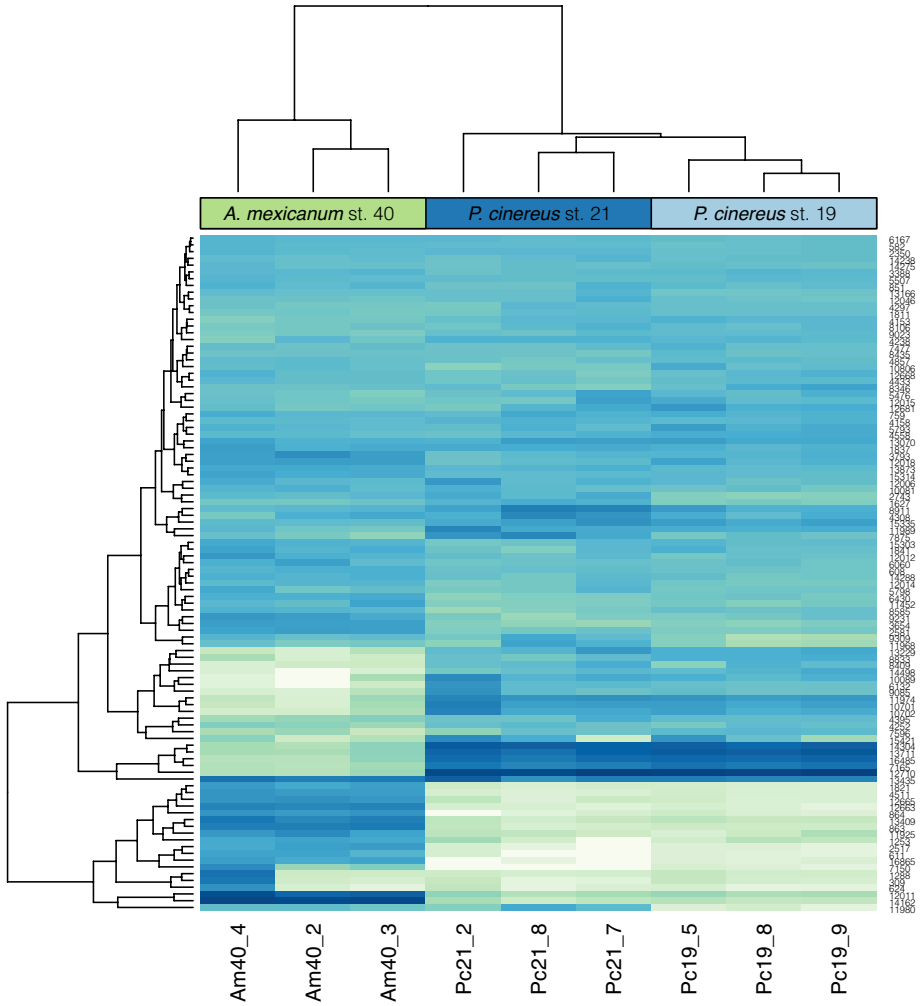


Figure 3.2: Heatmap of the top 100 differentially expressed genes within lung primordia transcriptomes from lungless and lunged salamanders. Heatmap shows the variance-stabilized counts in decreasing order by row mean. Hierarchical clustering (at the top) splits libraries appropriately by stage and species, except for library Pc21_2, which is divergent from the other Pc21 libraries.

3.4.2 Differential expression between lung primordia of *Ambystoma mexicanum* and *Plethodon cinereus*

I assessed differential expression between stages and species using DESeq (Anders and Huber, 2010). The primary comparison that I employed was between Pc19 and Am40 (Fig. 3.3).

Under this comparison, both stages are morphologically comparable and the lung primordia have reached similar stages of development (Chapter 2). Additional comparisons included expression divergence between Pc19 and Pc21 to highlight genes with differential expression during lung rudiment emergence and during the onset of rudiment regression. Lastly, Pc21 was compared to Am40. This latter comparison is completed with the assumption that the expression pattern within the lunged salamander lung primordia remains relatively constant through later stages of embryonic development. Given this assumption, more emphasis is placed on differential expression between the highly comparable stages Pc19 and Am40, but differential expression between Pc21 and Am40 is presented as well (Fig. 3.3).

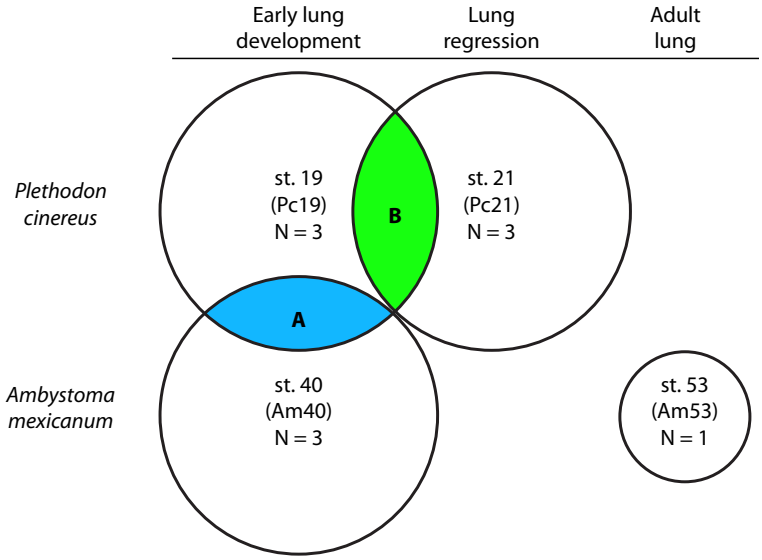


Figure 3.3: Schematic Venn diagram illustrating the sampling across stages and the primary differential expression comparisons employed. (A) Comparison between *Plethodon cinereus* and *Ambystoma mexicanum* at early lung development stages. (B) Comparison of expression between early lung development and lung regression within *P. cinereus*. The single stage-53 *A. mexicanum* lung library was sequenced to aid assembly of lung transcripts, but it was not used for differential expression analysis.

Differentially expressed sequences were first manually screened for potentially informa-

tive hits. Tables 3.2–3.7 show manually curated lists of differentially expressed sequences known to be involved in lung development, pulmonary physiology, intercellular signaling, and apoptosis. Genes upregulated in Pc19 relative to Am40 include *etv4*, a member of the Fgf signaling cascade, and *hhip*, a negative regulator of Shh signaling. Also upregulated were markers of Wnt signaling, including *aida*, *axin2*, Adenomatous polyposis coli protein (*apc*) and *yap1* (Jho et al., 2002; Konsavage et al., 2012; Li et al., 2013). Upregulation of Tgfβ signaling superfamily members was also observed (discussed below). Downregulated genes in Pc19 relative to Am40 include *fgf2* and *fgfr3*, and Pulmonary surfactant proteins A, B, and C (Table 3.3). Also significantly downregulated is the homeobox gene *nkx2.1*, a critical regulator of lung development (Tables 3.3, 3.5).

Comparing stages Pc19 to Pc21 there is a marked downregulation of *hhip* (Shh negative regulator) and the Wnt effector β-catenin at the later stage in *Plethodon*. The apoptotic regulator *bcl2* is downregulated at Pc21 relative to Pc19, a signature of apoptosis (Table 3.7). Differentially expressed sequences between Pc21 and Am40 largely overlap with the differentially expressed sequences between Pc19 and Am40. *Tbx5*, a gene with pleiotropic roles in lung, limb and heart development is significantly downregulated in Pc21 relative to Am40. However, *tbx5* is downregulated at Pc19 as evidenced by a two-sample t-test ($p = 0.033$), but not significantly downregulated according to DESeq (Table 3.5).

3.4.3 Altered lung gene regulatory network in *Plethodon cinereus*

Because the gene regulatory network (GRN) underlying lung development is both well-characterized and highly conserved (Rankin and Zorn, 2014; Volckaert and De Langhe, 2015),

I mapped differential expression data onto the known genetic interactions governing early lung development (Fig. 3.4). There are several key deviations to the GRN of *P. cinereus* GRN relative to *A. mexicanum*. As noted above, key indicators of Wnt signaling activity are significantly upregulated in *P. cinereus*, including β-catenin, *apc* and *axin2*. Wnt2 and Wnt2b are critical for specifying the pulmonary domain within the foregut endoderm. Although there is no significant difference in expression level of *wnt2b* between *P. cinereus* and *A. mexicanum*, *wnt2b* expression is approximately two fold higher in *P. cinereus* (*wnt2* was not assembled). Tankyrase (*tnks*), a positive regulator of Wnt signaling, is strongly upregulated in Pc19 compared to Am40, adding support to this hypothesis (Table 3.2). leading to the conclusion that *P. cinereus* displays higher levels of Wnt signaling than *A. mexicanum* (at Pc19) (Table 3.2). At Pc21, though still upregulated in comparison to *A. mexicanum*, β-catenin expression significantly decreases compared to Pc19, indicating a drop in Wnt signaling in the *P. cinereus* lung rudiment just before lung regression (Table 3.7).

Early cell fate decisions and lung morphogenesis are governed by the establishment of dorsoventral polarity in the foregut tube, marked by expression domains of Nkx2.1 (ventral) and Sox2 (dorsal) (Domyan et al., 2011; Que et al., 2007). Significantly reduced levels of *nkx2.1* in *P. cinereus* may result in a number of downstream effects and has the potential to lead to lung loss. Nkx2.1 expression can be reduced by two general mechanisms: increased suppression by *nkx2.1* inhibitors or reduced activation by *nkx2.1* activators. Sox2 inhibits *nkx2.1* expression (Que et al., 2007). Sox2 expression is higher in *P. cinereus* than in *A. mexicanum*, but not significantly so. However, RNA extraction targeted ventral endoderm and mesoderm, therefore *sox2* expression levels are likely underestimates. Bmp signaling via Bmpr1a and Bmpr1b (Alk3 and Alk6) represses *sox2* expression (Domyan et al., 2011).

Bmpr1b is upregulated in *P. cinereus*, but it is unclear if upregulation of the receptor indicates increases in Bmp signaling. Expression of Bmp ligands is not significantly different in any condition, but Bmp-binding endothelial regulator protein (*bmper*), a known inhibitor of Bmp signaling, is strongly downregulated in *P. cinereus* at both Pc19 and Pc21 (Figs. 3.5, 3.5). Hence there are some indications that Bmp signaling in general is upregulated in *P. cinereus*, which presumably would result in decreased expression of *sox2* in the dorsal foregut endoderm and subsequent increased expression of *nkx2.1* (Rankin et al., 2012). Tgfβ inhibits *nkx2.1* expression (Li et al., 2013) and Tgfβ signaling appears to be generally elevated in *P. cinereus* (discussed below), potentially contributing to significantly lower expression of *nkx2.1* in this species.

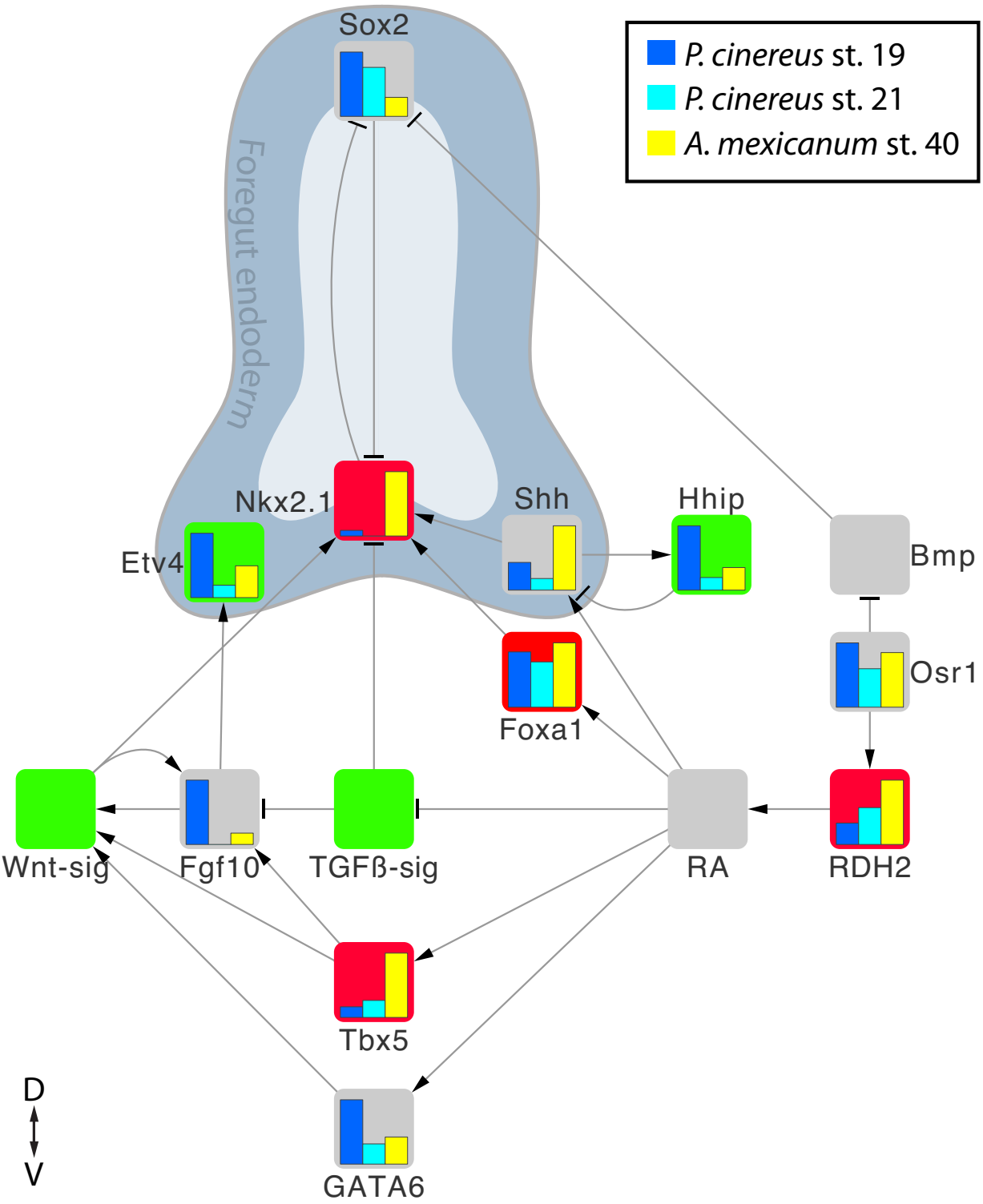
Known activators of *nkx2.1* expression include Forkhead proteins Foxa1 and Foxa2 (Ikeda et al., 1995; Minoo et al., 2007), Wnt signaling (via β-catenin) and Shh signaling (Rankin and Zorn 2014). *Foxa1* expression is significantly reduced at both Pc19 and Pc21 relative to Am40. Pc19 and Pc21 likely display decreased Shh signaling relative to Am40. The expression of the downstream effector of Shh signaling, *gli1* (Grindley et al., 1997), is significantly reduced in both *P. cinereus* stages (Tables 3.3, 3.5). In addition, the negative regulator of Shh signaling, *hhip*, is upregulated (Chuang and McMahon, 1999). However, Shh ligand expression is only marginally lower in Pc19 relative to Am40 ($p = 0.06$). The combined downregulation of Nkx2.1 activators (*foxa1* and Shh signaling) and the upregulation of inhibitory Tgfβ signaling (and potential upregulation of inhibitory *sox2*) may result in decreased levels of *nkx2.1* expression in *P. cinereus* and result in lung loss.

Retinoic acid plays a number of pleiotropic roles in lung development (Chen et al., 2007, 2010; Rankin et al., 2012). Retinoic acid (RA) production is likely reduced in *P. cinereus*

due to the downregulation of the retinol dehydrogenase *rdh2* (*sdr16c5*), which is responsible for catalyzing the first and rate-limiting step of RA production, the conversion of retinol to retinaldehyde. Reduction in RA synthesis may be responsible for the downregulation of *foxa1* and *tbx5*. In addition, since RA represses Tgfβ signaling, reduced RA may result in observed upregulation in Tgfβ signaling (discussed below). However, other retinol dehydrogenases are not differentially expressed in *P. cinereus*, and it is possible that these enzymes compensate for reduced expression of *rdh2*. Therefore, RA is marked in grey on the GRN, likely lower in *P. cinereus*, but with insufficient evidence to draw any strong conclusions.

Figure 3.4: Differential expression within the lung gene regulatory network of lungless *Plethodon cinereus* relative to lunged *Ambystoma mexicanum*. Lung rudiment transcriptomes were analyzed for differential expression using DESeq (Anders and Huber, 2010). Lung primordia expression values for stage-19 *P. cinereus* (Pc19), stage-21 *P. cinereus* (Pc21) and stage-40 *A. mexicanum* (Am40) are mapped on each node within the gene regulatory network. Network topology corresponds to the topology proposed by Rankin and Zorn (2014), with any additions discussed in the text of the chapter. Genes with expected foregut expression are arranged along the dorsoventral axis of a schematic transverse section through the foregut. All other genes are presumably expressed in the mesoderm, with the exception of Tgfβ signaling (Tgfβ-sig) and Wnt signaling (Wnt-sig). These two broad groupings indicate general patterns of upregulated expression that are explained further in the text. Hence, members of these signaling pathways may be expressed mesodermally or endodermally. Nodes are marked green if the gene is significantly upregulated in Pc19 relative to Am40 and red if the gene is significantly downregulated in Pc19 relative to Am40. One exception is *tbx5*, which is significantly downregulated at Pc21 but is only marginally significant at Pc19. Grey nodes are either not assembled or not differentially expressed. Bar charts within each node correspond to the average expression levels within each species and stage; colors are defined in the key at upper right.

Figure 3.4, continued



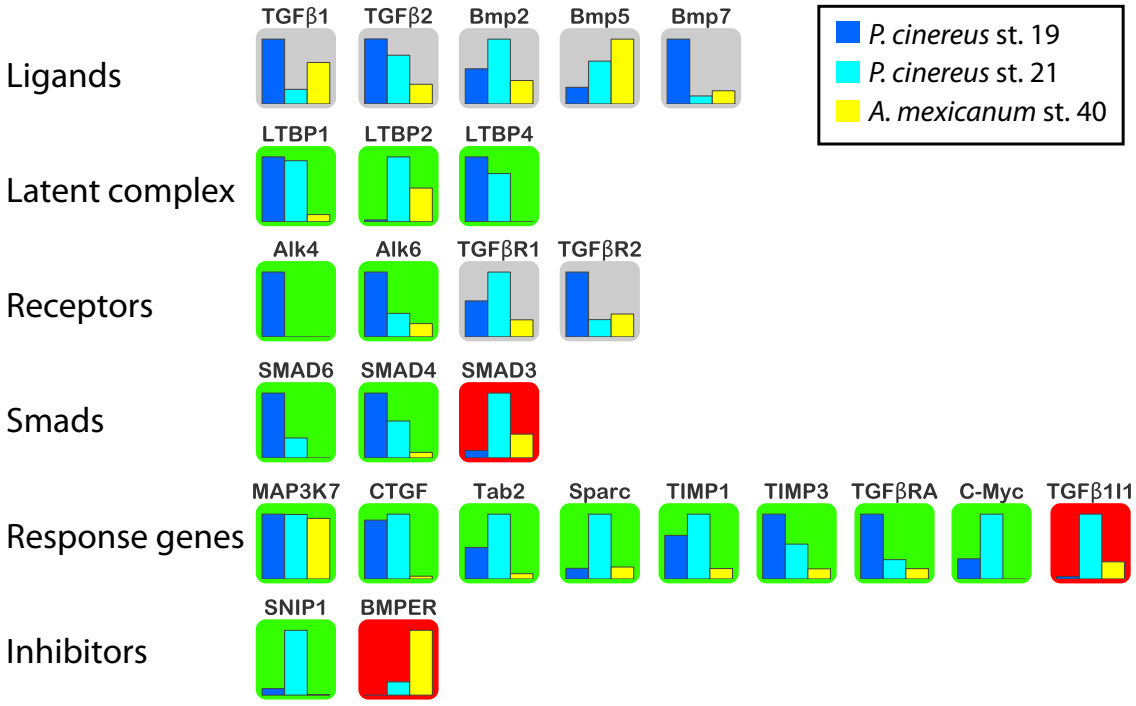


Figure 3.5: Differential expression of Tgfβ superfamily members in lung primordia between *Plethodon cinereus* and *Ambystoma mexicanum*. Tgfβ superfamily signaling pathway members are partitioned into ligands, latent complex constituents, receptors, Smads, response genes and inhibitors. Differential expression of each gene was determined using DESeq. Green indicates genes with significantly higher expression in *P. cinereus*. Red indicates significant downregulation relative to *A. mexicanum*. Grey means that there was no significant difference. Average expression values for each time point are charted in each box; colors correspond to the key at upper right. *ltbp2* and *sparc* were not significantly differentially expressed between Pc19 and Am40, but they were significantly higher in Pc21 relative to Am40.

3.4.4 Gene ontology enrichment

Gene ontology (GO) term enrichment analysis was used to examine broad categories of differentially regulated genes (Fig. 3.6). The top overrepresented GO category in genes upregulated in Pc19 relative to Am40 was Tgfβ-activated receptor activity (GO: 0005024; including *smad4, map3k7, ski-like, ltbp1, apc, smad6, acvr1b, snip1* and *bmpr1b*). Also overrepresented in this set of upregulated genes were transmembrane receptor protein kinase activity (GO: 0019199, including most of the same genes as GO: 0005024) and apoptotic mitochondrial changes (e.g., GO:0010623, including *bcl2, fis1, m3k1*) (Fig. 3.6A). Within the genes downregulated in Pc19 relative to Am40 are a number of overrepresented GO categories, such as extracellular matrix structural constituent, cell adhesion, biological adhesion, integrin complex and vascular endothelial growth factor receptor signaling. Several GO categories related to apoptosis are overrepresented in the genes upregulated in Pc21 relative to Pc19, presaging the regression of the lung rudiment. These GO categories include B cell apoptotic process, lymphocyte and leukocyte apoptotic processes, and intrinsic apoptotic signaling pathway (Fig. 3.6B).

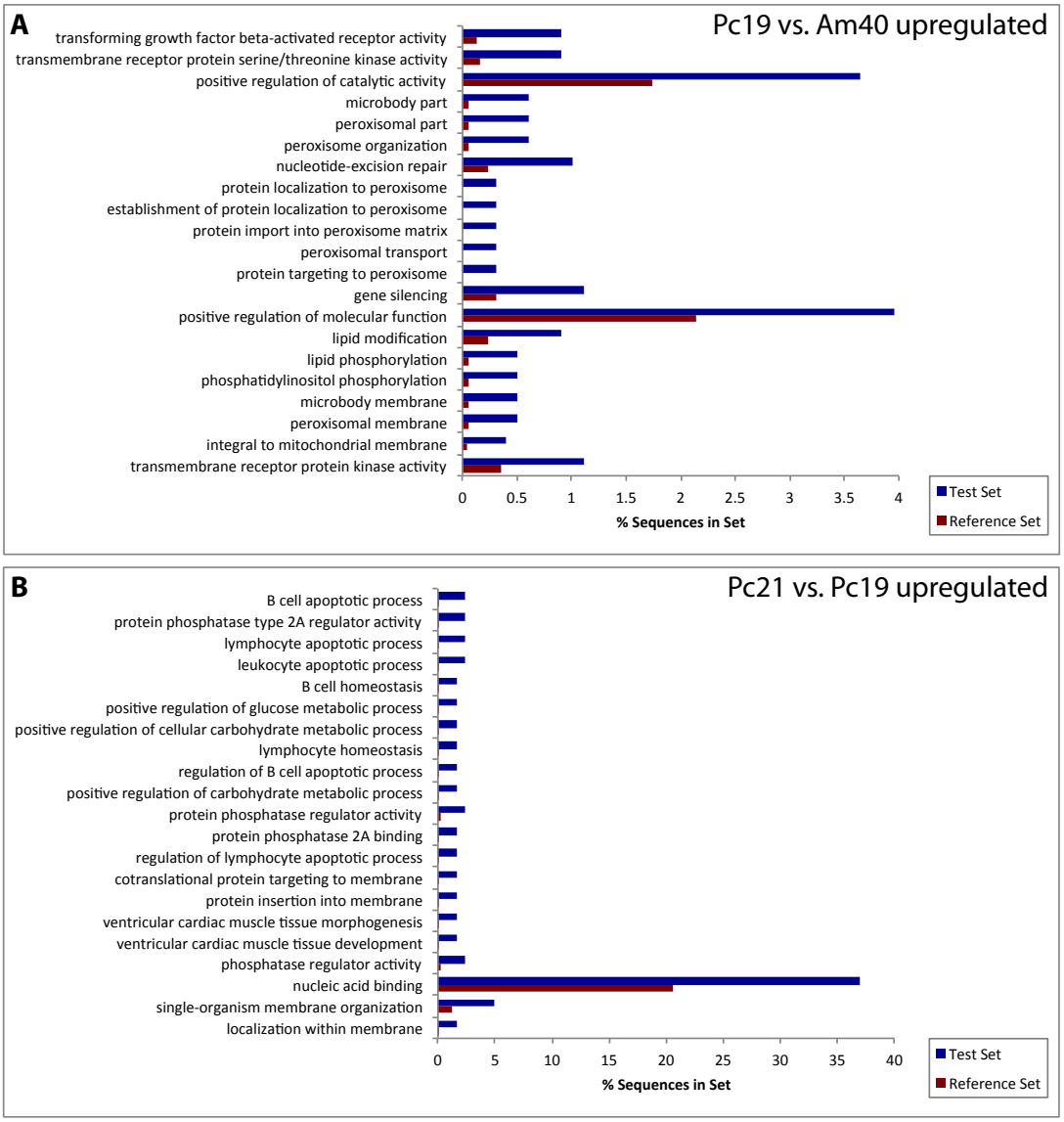


Figure 3.6: Gene ontology term enrichment within sets of differentially expressed genes. (A) Top overrepresented gene ontology (GO) terms within the genes significantly upregulated in Pc19 relative to Am40, as ranked by p-value. DESeq was used to determine differential expression. GO overrepresentation was determined using Fisher's Exact Test within Blast2GO. The values represent the percent of genes assigned to a particular GO category present either within the test set (differentially expressed genes) or within the reference set (a set of all expressed genes). (B) Top overrepresented gene ontology (GO) terms within the genes significantly upregulated in Pc21 relative to Pc19, as ranked by p-value.

3.4.5 Upregulation of Tgfβ superfamily signaling in *Plethodon cinereus*

A number of members of the Tgfβ superfamily are upregulated in *P. cinereus* relative to *A. mexicanum* (Fig. 3.5; Tables 3.2, 3.4). These genes include receptors *bmpr1b* (*alk6*) and *acvr1b* (*alk4*), which are Type I receptors for Bmp and Activin, respectively. Latent Tgfβ-binding proteins *ltbp1* and *ltbp4* are expressed at significantly higher levels in *P. cinereus* at Pc19. The same *ltbp* genes are expressed at significantly higher levels in Pc21, with the addition of *ltbp2*. A number of indicators of active Tgfβ signaling have been identified (Benson et al., 2002; Chen et al., 2007; Duncan et al., 1999; Verrecchia et al., 2001). Indicators of Tgfβ signaling including Tgfβ receptor associated protein (*tgfbra*), Connective tissue growth factor (*ctgf*), Tgfβ-activated kinase 1 (*tab2*), *sparc, timp1, timp3, c-myc* and *map3k7* (also known as *tak1*) are significantly upregulated in *P. cinereus* (Fig. 3.5). Smad homologs 4 and 6 are strongly upregulated in *P. cinereus*. Smad4 is a co-Smad, which forms a heteromere with active receptor activated Smads (Shi and Massague, 2003). While a large number of indicators of active Tgfβ superfamily signaling are overexpressed in *P. cinereus*, upregulation of inhibitory *smad6*, Smad nuclear interacting protein 1 (*snip1*) and downregulation of Tgfβ1i1 and *smad3* at Pc19 relative to Am40 may be evidence of reduced Tgfβ superfamily signaling. Due to overlap in Tgfβ superfamily family signaling pathways it is difficult to conclusively demonstrate upregulation of one pathway (e.g., Tgfβ vs. Bmp vs. Activin). However, upregulation of key Tgfβ signaling readouts argue strongly for upregulated Tgfβ signaling in *P. cinereus*, offering a potential proximate mechanism underlying lung loss (Fig. 3.4).

3.4.6 Tgfβ restricts lung specification and development in *Ambystoma mexicanum*

To test the hypothesis that Tgfβ signaling is involved in lung loss, I manipulated Tgfβ signaling during lung development in the lunged species *A. mexicanum*. Under the hypothesis that increased Tgfβ signaling is responsible for lung loss, I expect that increasing Tgfβ will negatively impact lung development in *A. mexicanum*. Initially, Tgfβ1 protein was administered to the lung region by implantation of Affi-gel blue beads soaked in recombinant human Tgfβ1 protein. No effects on lung development were observed (data not shown). The retention and secretion of compounds from beads is time limited, and lung development occurs over the course of many days. Therefore, any phenotype reliant on sustained ectopic expression of Tgfβ1 would not be observed using bead implantation. To circumvent this problem, I created an explant system in which thoracic segments of *A. mexicanum* embryos could be cultured under constant exposure to exogenous Tgfβ1 in solution. Embryos were cleaved just anterior to the outflow tract of the heart and just posterior to the liver, resulting in an open thoracic segment. Thoracic explants were cultured in 50% L-15 medium at 17°C in the presence of Tgfβ1 or bovine serum albumin (BSA) as a negative control. Administration of 50 mg/ml Tgfβ1 from stage 34 (prior to lung emergence) through stage 42 resulted in incomplete development of the trachea in 89% of assayed explants (N = 9; Fig. 3.7B). Lung buds were still present, but appeared to emerge directly from the foregut. Control individuals developed full trachea and lung buds (100%, N = 8; Fig. 3.7A). These data imply that ectopic Tgfβ represses tracheal formation.

Although excess Tgfβ appears to repress tracheal and/or lung development (present

study; Chen et al., 2007), little is known about the effects of reduced Tgfβ signaling on lung development. I therefore used the chemical antagonist A-83-01 to inhibit Tgfβ signaling. A-83-01 potently inhibits the function of Alk5 (TgfβR1), but it also inhibits Alk7 (Nodal receptor) and Alk4 (Activin receptor) (Tojo et al., 2005). Treatment of *A. mexicanum* embryos with 10 µm A-83-01 from stages 33 through 43 results in the ectopic expression of Surfactant protein C (SPC), a specific marker of the lungs in *A. mexicanum* (Fig. 3.8B). In 100% of treated individuals (N = 4) SPC expression was dramatically expanded into the dorsal and ventral pharynx. Since SPC is a marker of lung identity, reduced Tgfβ signaling results in expansion of the lung domain, supporting a role for Tgfβ in repressing lung development. Lung development still proceeded normally in treated individuals. Expansion of SPC expression was not observed in the case of treatment with LDN193189, an inhibitor of Alk2 and Alk3 (thus preferentially inhibiting BMP signaling; data not shown).

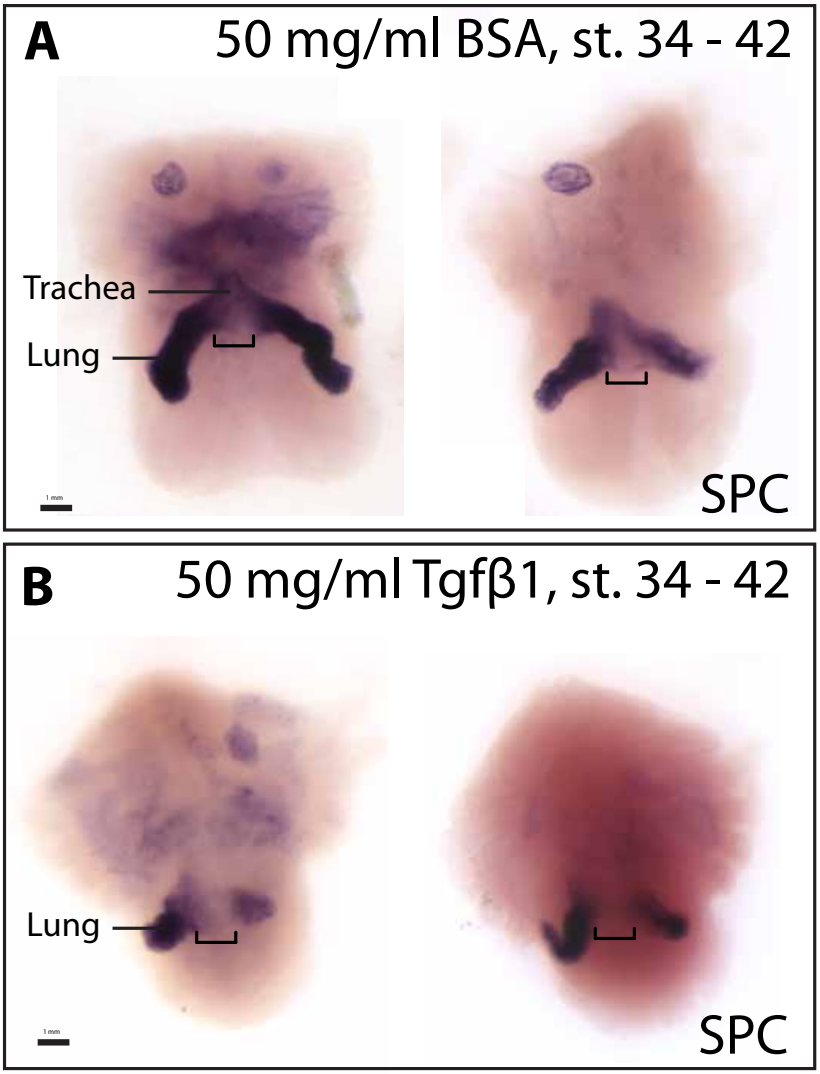


Figure 3.7: Treatment of *Ambystoma mexicanum* thorax explants with Tgfβ1 represses tracheal development. Explants were cultured for 10 d (from stages 34 through 42) in BSA (control) or recombinant human Tgfβ1 protein at 50 mg/ml. They were then stained for presence of the lung and trachea using *in situ* hybridization against SPC. Ventral views. (A) All control explants develop full trachea and lungs (100%; N = 8). (B) All but one explant treated with Tgfβ1 fail to develop trachea, as indicated by lack of median expression of SPC (89%; N = 9). Lungs still appear to develop normally.

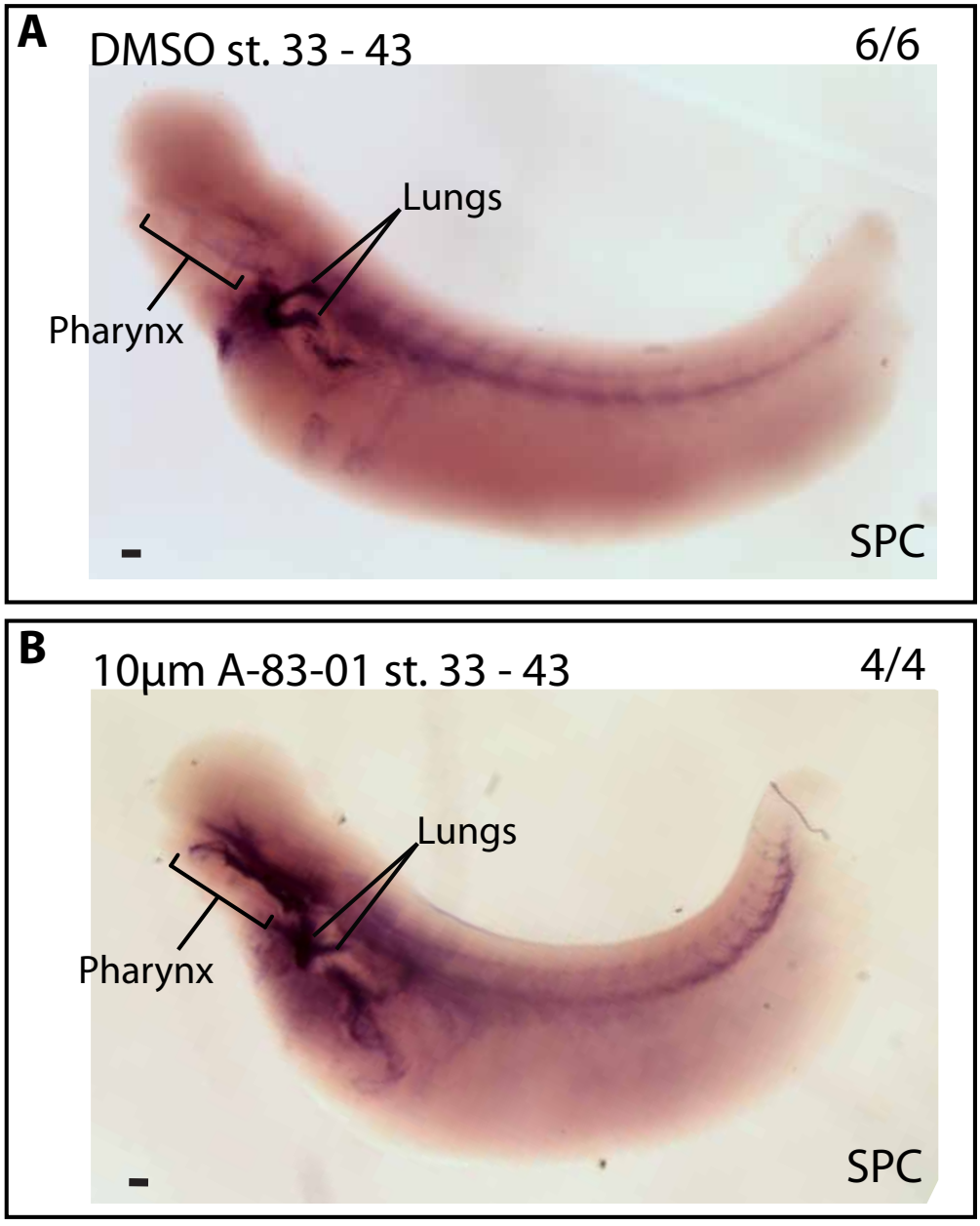


Figure 3.8: Treatment of *Ambystoma mexicanum* embryos with the Tgfβ signaling antagonist A-83-01 expands lung specification into the mouth and pharynx. (A) DMSO treatment from stages 33 through 43 results in normal expression of SPC in the lungs and trachea. (B) Embryos treated with 10 µM A-83-01 from stages 33 through 43 develop lungs and trachea normally, but they also display expanded expression of SPC into the pharynx and oral cavity. Scale bars: 100 µm.

3.4.7 The role of Tgfβ in lung loss in *Plethodon cinereus*

Given the hypothesized role of upregulated Tgfβ signaling in lung loss, I used the chemical antagonist A-83-01 to inhibit Tgfβ signaling in *P. cinereus*. Treatment of *P. cinereus* embryos with A-83-01 began at stage 19 and terminated at stage 23, by which time regression of the lung rudiment is complete. Treated individuals were slightly smaller than controls but had no gross morphological defects and appeared to be at the same stage judging by their cranial and digit morphology (Kerney, 2011). Treatment with 20 µm A-83-01 resulted in persistence of the lung rudiment in all treated individuals examined (N = 5). The lung rudiment was distinguishable by histology as a long diverticulum, which terminated anterior to the liver with an enlarged lumen (Fig. 3.9C,D). The retained rudiment expresses the pulmonary marker *sox9*, as demonstrated by *in situ* hybridization (Fig. 3.9F). Three of five (60%) negative controls treated with DMSO lacked lung rudiments. However, two control individuals appeared to retain a partial lung rudiment, though the structure was collapsed and smaller than treated individuals and did not strongly express *sox9* (Fig. 3.9B). The retention of the lung rudiment in these two control individuals points to a non-specific role of A-83-01 in preventing lung rudiment regression: the inhibitor may cause a developmental delay that results in later regression. Attempts were made to control for developmental delay by closely examining embryos prior to fixation. However, some heterochronic shifts may have gone unnoticed. Hence, A-83-01 treatment of *P. cinereus* embryos results in the retention of the lung rudiment, but this retention is possibly an off-target effect of a general developmental delay.

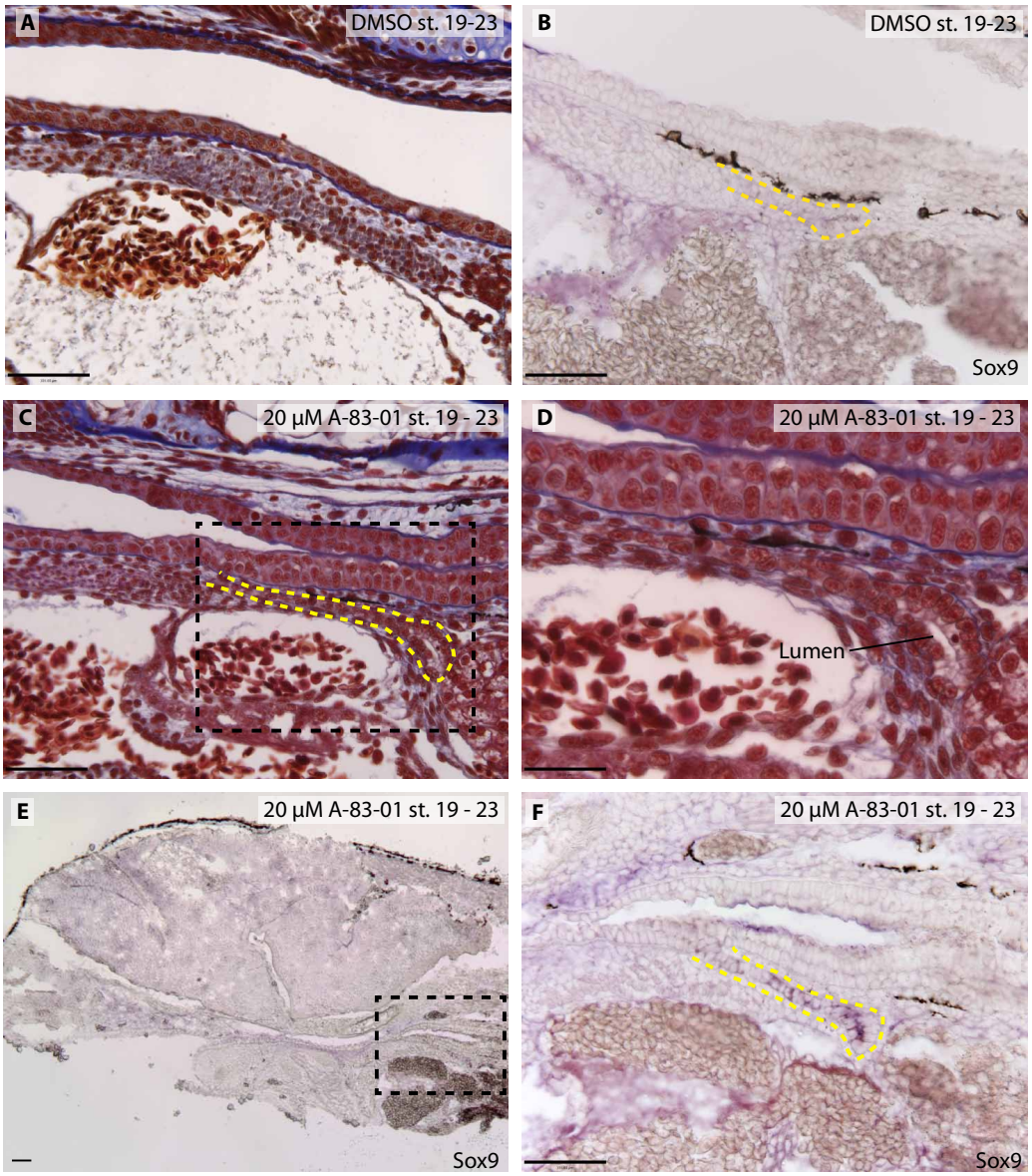


Figure 3.9: Treatment of *Plethodon cinereus* embryos with the Tgfβ signaling antagonist A-83-01 results in retention of the lung rudiment. Sagittal sections, anterior is to the left. (A, B) The lung rudiment has regressed (A) or nearly regressed at stage 23 in control embryos treated with DMSO from stages 19 through 23. (B) The remnant of the lung rudiment (outlined in a dashed yellow line) does not express Sox9. (C–F) Treatment of *P. cinereus* embryos with the Tgfβ signaling antagonist A-83-01 from stages 19 through 23 results in retention of the lung rudiment and formation of a large luminal space. (C) Histological section through the lung rudiment (dashed yellow line). (D) Enlargement of the boxed region in (C). (E) Section through a treated embryo stained for *sox9* expression with *in situ* hybridization. (F) Enlargement of the boxed region in (E). Sox9 is expressed in the elongated lung rudiment. Scale bars: 100 µm

3.4.8 Wnt signaling is critical for lung development in *Ambystoma mexicanum*

Wnt signaling plays a prominent role in lung specification (Goss et al., 2009; Harris-Johnson et al., 2009) and maintenance (Rankin et al., 2012). Canonical Wnt signaling results in the accumulation of nuclear-localized β-catenin, which would otherwise be targeted for degradation by the β-catenin destruction complex made up of proteins such as Axin, Apc and Gsk3. Tankyrase (Tnks) stimulates the degradation of Axin and therefore is permissive towards Wnt signaling. Inhibition of Tankyrase using the small molecule XAV939 results in reduced Wnt signaling (Huang et al., 2009). Treatment of *A. mexicanum* embryos with XAV939 from stage 31 until stage 43 resulted in complete lung loss (50%) or left lung reduction (50%) (Fig. 3.10C). Lung loss accompanies general endodermal defects including dense and disorganized endoderm and a small heart (Fig. 3.10D). These data demonstrate the conserved role that Wnt signaling plays in salamander lung development (Fig. 3.10).

3.4.9 Lunged salamander mesenchyme may induce lung development in lungless salamanders

Reciprocal signaling interactions between the foregut endoderm and the underlying mesoderm drive lung specification and morphogenesis (McCulley et al., 2015). I hypothesized that aberrant signaling between these tissues is responsible for lung loss. To test this hypothesis, I attempted to restore lung development in the lungless plethodontid species *Hemidactylium scutatum* by heterospecific transplantation of pulmonary mesoderm from *Ambystoma mexicanum*. *Hemidactylium scutatum* was selected because it is possible to obtain these embryos

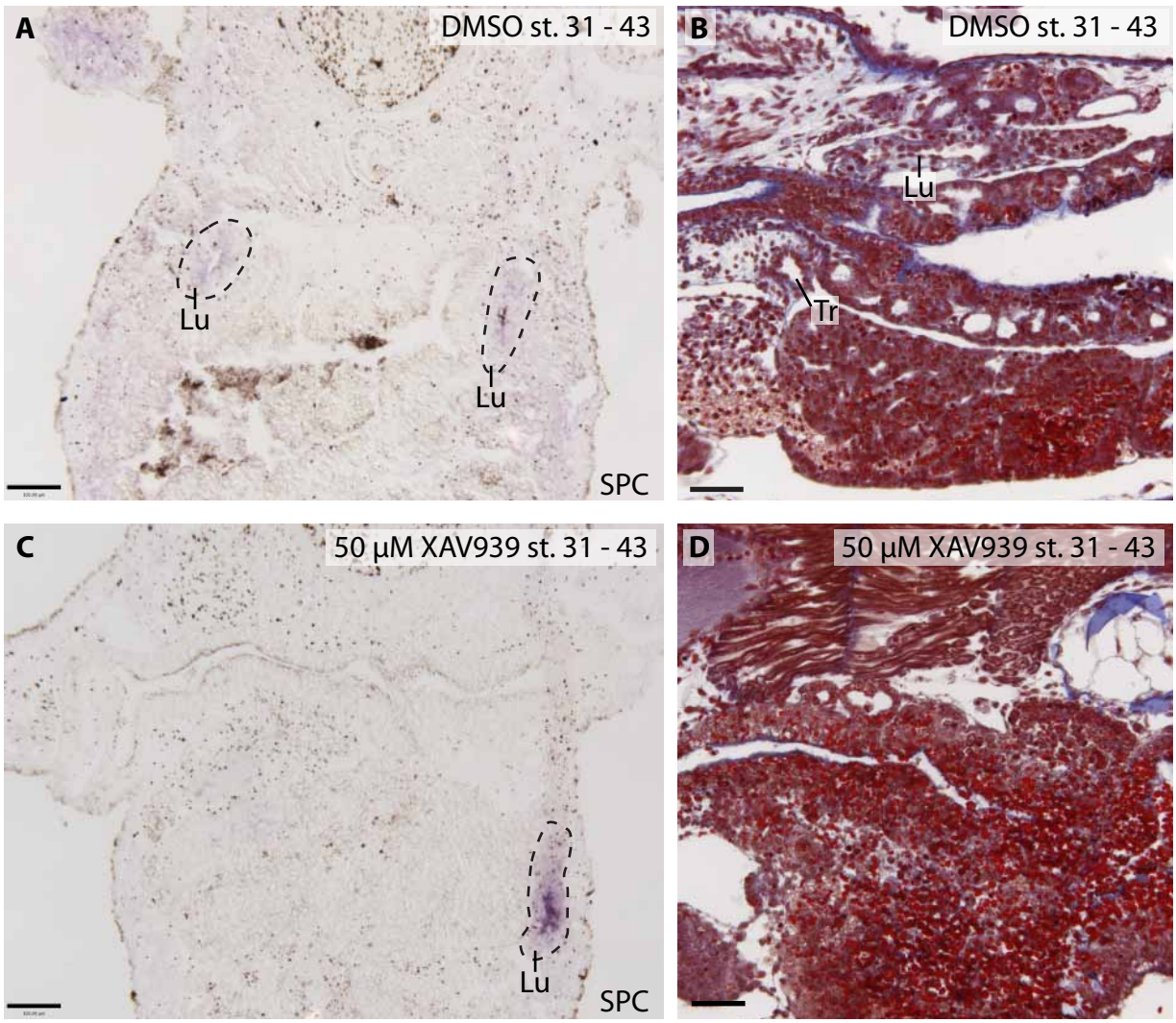


Figure 3.10: Treatment of *Ambystoma mexicanum* embryos with the Wnt signaling inhibitor XAV939 inhibits lung development. (A) Paired lung buds are visible in a DMSO control at stage 43, as revealed by SPC *in situ* hybridization (dashed lines). Transverse section. (B) Tracheal (Tr) and lung (Lu) development are normal in a DMSO control individual. Sagittal section. (C) Development of the left lung is inhibited following treatment with 50 µM XAV939. The right lung is still present. Transverse section. (D) Endodermal tissue is disorganized in a treated individual. Sagittal section. Scale bars: 100 µm

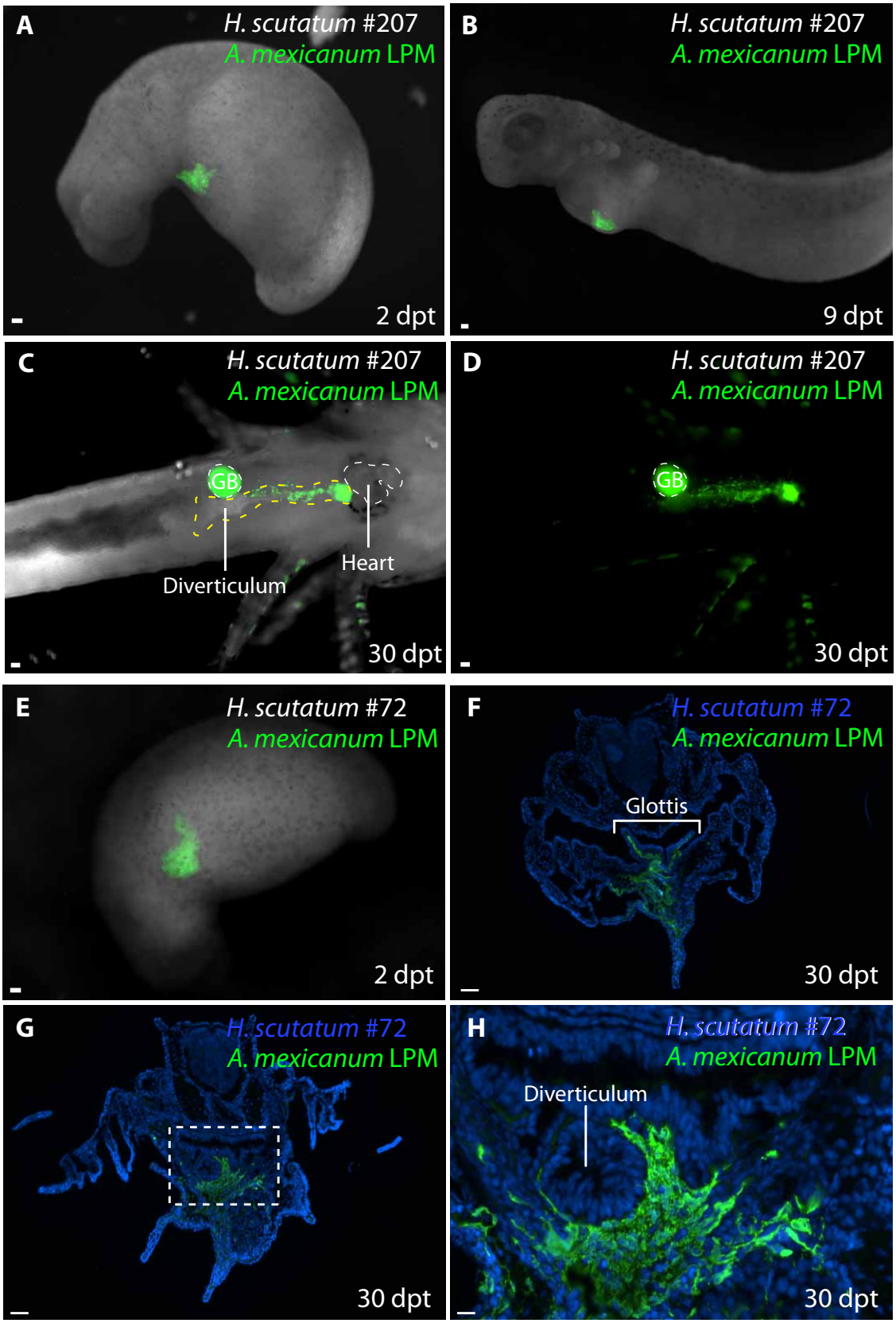
in large numbers and their mesolecithal biphasic development closely resembles *A. mexicanum*, in contrast to the direct-developer *P. cinereus*.

Homotypic grafts of GFP+ *A. mexicanum* lateral plate mesoderm were transplanted into

lungless *H. scutatum* hosts at the 3- or 4-somite stage. Embryos were allowed to develop for approximately 30 d. Out of 80 heterospecific transplantations from 2014, 49 embryos survived. Of these 49 embryos, nine (18%) displayed externally visible evidence of pulmonary development. In these embryos an ectopic endodermal diverticulum emerging from dorsal to the heart was present and surrounded by GFP+ splanchnic mesoderm derived from *A. mexicanum* (Fig. 3.11). The diverticulum was composed of light colored, spongy tissue. Distally, the unicameral diverticulum was located ventral to the liver. In all cases this tube was unpaired, perhaps reflecting the receipt of unilateral grafts. Several surviving embryos from the comparable experiment in 2013 were sectioned and stained for GFP presence (Fig. 3.11E–H). In these embryos, the endodermal diverticulum is clearly visible and surrounded by *A. mexicanum*-derived splanchnic mesoderm. One embryo possessed a glottis-like structure proximally (Fig. 3.11F). Mock transplant controls did not develop lung-like structures. Transplant recipients could not be examined for molecular indications of lung development due to the absence of *in situ* probes for *H. scutatum* and the absence of antibodies that mark pulmonary endoderm in salamanders. Despite this limitation, heterospecific transplantation offers some evidence that lung loss is the result of the lack of inductive or maintenance signals from the splanchnic mesoderm.

Figure 3.11: Transplantation of lunged *Ambystoma mexicanum* lateral plate mesoderm into lungless *Hemidactylium scutatum* embryos generates a putative lung. (A–D) Transplant recipient #207 2 days post-transplant (dpt), 9 dpt and 30 dpt. *Ambystoma mexicanum* lateral plate mesoderm is GFP positive and can be traced as it contributes to the splanchnic mesoderm surrounding the foregut. (C, D) Ventral view of the larva at 30 dpt illustrates the formation of an ectopic median diverticulum surrounded by *A. mexicanum* splanchnic mesoderm. The diverticulum emerges from dorsal to the heart and terminates ventral to the liver. The gall bladder (GB) is autofluorescent, as are some pigment cells in the skin. (E–H) Transplant recipient #72. (F) Transverse section through the pharynx illustrates the formation of a glottis, supported by *A. mexicanum* splanchnic mesoderm. (G) A transverse section farther posterior shows the ectopic diverticulum, which emerges from the foregut. (H) Enlargement of boxed region in (G). Scale bars: 100 µm, except (H): 50 µm

Figure 3.11, continued



3.5 Discussion

Examples of organ loss are rampant across metazoans, including numerous losses of wings, limbs, intromittent organs, eyes and teeth (Fong et al., 1995). Among tetrapods, lung loss is confined to four or more lineages of amphibians (Wilkinson and Nussbaum, 1997), but the homologous swim bladder has been lost over 30 times in teleost fishes and presumably at the base of Chondrichthyes (Liem, 1988; McCune and Carlson, 2004). Theoretically, organ loss could be the result of multiple evolutionary processes, which fit into three broad categories: (1) adaptive processes, (2) neutral processes, and (3) pleiotropic outputs. Adaptive processes could include energetic tradeoffs in organ investment and positively selected functional changes to organs resulting in vestigialization or loss. In plethodontids, most hypotheses regarding lung loss invoke adaptive explanations such as selection for reduced buoyancy (Wilder and Dunn, 1920) or selection for improved terrestrial feeding capability (Reagan and Verrell, 1991). Neutral processes may result in organ loss if maintenance of organ development is under relaxed stabilizing selection given the environment in which a population lives, resulting in genetic drift (for example, pigment loss in caves). In salamanders, it is possible that novel innovations in respiratory physiology (see Chapter 4), environmental changes, or small body size in general resulted in less reliance on pulmonary respiration and the accumulation of mutations in lung-patterning genes. Lastly, pleiotropy can result in organ loss. Selection for one constructive trait may cause changes to traits linked by shared developmental mechanisms, including shared utilization of patterning genes. Eye loss in cavefish may in fact be a result of pleiotropy and not neutral processes. Selection for increased taste bud density (better sensory function) via expanded Shh expression may

come at the expense of eye development (Jeffery, 2009).

In contrast to the ready availability of hypotheses regarding the basis of organ loss, it has been difficult to conclude much about the ultimate (selective) and proximate (mechanistic) causes of organ loss. One example where the mechanistic basis is well understood is in the loss of the intromittent organ in birds. Bmp4 upregulation has been implicated in apoptosis of the developing genital tubercle (Herrera et al., 2013). It is difficult to determine whether adaptive, neutral or pleiotropic processes have resulted in lunglessness. While ultimate causes of organ loss are difficult to retroactively address, my study represents an attempt to better understand the proximate causes of lung loss. Future work on pulmonary genes in plethodontids may reveal some information regarding the ultimate factors involved in lung loss, particularly if signatures of positive selection or accumulation of deleterious mutations are genomically identified.

Nearly all genes that have been previously implicated in lung development are present in *Plethodon cinereus*. Bmp4 is one exception. It was not assembled in either *P. cinereus* or *A. mexicanum*. This may indicate that Bmp4 does not play a role in lung development in salamanders. The expression of nearly all lung-patterning genes implies that lung loss is not due to loss of a master regulator of lung development, but instead is likely due to differences in expression at key network nodes. Neutral processes may result in the accumulation of mutations in lung patterning genes. However, because most of these signaling pathways (Fgf, Shh, Wnt, Tgfβ, Bmp) play roles across multiple tissue types I expect that stabilizing selection eliminates many deleterious coding mutations. Indeed, pleiotropy of these genes may be one reason for retention of the lung rudiment over millions of years since lungs were lost from adults (Chapter 2). On the other hand, gene regulatory mutations may be

well tolerated. For instance, discrete Shh enhancers drive expression at particular locations along the anterior-posterior axis of the endoderm (Sagai et al., 2009) and mutations to such tissue specific enhancers may result in down- or upregulation of gene expression in lungless salamanders without impacting the development of other organs.

One of the primary benefits of studying lung loss in plethodontids is the ability to directly sequence the transcriptome of the organ rudiment. However, extrapolating observed gene expression differences to mechanisms of lung loss requires two assumptions: (1) genetic changes responsible for lung loss in plethodontids are regulatory and not coding, and (2) differential expression between lunged and lungless salamanders is causative of lung loss and does not simply reflect historical artifacts of speciation, or genetic drift following the loss of lungs by a different mechanism. In reference to the first assumption, there are several lines of evidence based on this study that regulatory mutations are responsible for lung loss. The formation of a lung rudiment and the expression of markers of lung differentiation indicate conservation of many key mechanisms that underlie lung formation in lungless salamanders (Chapter 2). Late stages of lung development in mammals are dependent on maintenance of the same signals responsible for early lung bud emergence and for the most part not deployment of particular late-patterning genes (Morrisey and Hogan, 2010). Therefore, if coding mutations abrogate lung maintenance then these mutations should theoretically affect early lung primordium emergence, as well. This is not the case in *P. cinereus*. In fact, nearly all characterized markers of lung differentiation are expressed in *P. cinereus* (Chapter 2). It is also possible that an active process promoting regression is responsible for lung loss. Upregulation of particular endogenous lung developmental pathways, such as Tgfβ, can result in apoptosis (discussed below).

The second assumption of the present study is that gene expression differences underlie lung loss and do not simply reflect phylogenetic distance between the focal species or are the result of genetic drift following lung loss. It is difficult to directly address the validity of this assumption without additional sequenced pulmonary transcriptomes. If lung primordia transcriptomes from phylogenetically diverse lunged and lungless species display similar differences in expression, then this would provide support for this assumption.

3.5.1 The role of Tgfβ signaling in lung loss

I present evidence that gene regulatory networks underlying lung development are divergent between the lungless salamander *Plethodon cinereus* and the lunged species *Ambystoma mexicanum*. I show evidence that Tgfβ signaling, in particular, is strongly upregulated in *P. cinereus*. In addition, *P. cinereus* displays increased canonical Wnt signaling relative to *A. mexicanum*. Downregulated genes in *P. cinereus*, which include *nkx2.1*, *tbx5*, *gli1* and *foxa1*, are known to play essential roles in lung development.

Excess Tgfβ signaling may play a causative role in lung loss. Expression of Latent Tgfβ binding proteins 1 and 4 is markedly higher in *P. cinereus*. In addition, *ltbp2* expression is significantly elevated at Pc21. Ltpbs play a key yet underappreciated role in Tgfβ signaling and function to regulate Tgfβ signaling in the lung (Annes et al., 2003; Zhou et al., 2009). Mice hypomorphic for *ltbp4* develop pulmonary emphysema (Sterner-Kock et al., 2002). Receptors for Bmp and Activin signaling were upregulated in *P. cinereus*, but, as with upregulation of *smad* expression, stronger expression does not necessarily entail more active signaling. Two frequently used methods to analyze upregulation of Tgfβ superfamily

signaling are immunohistochemical examination of phosphorylated Smads particular to each pathway and the change in expression of downstream indicator genes. Phospho-Smad staining was not attempted here, but examination of a number of well-characterized indicators such as *ctgf, tab2* and *map3k7* provides strong evidence for upregulation of Tgfβ signaling in *P. cinereus* (Fig. 3.5; Benson et al., 2002; Chen et al., 2007; Duncan et al., 1999; Verrecchia et al., 2001).

I hypothesize that upregulated Tgfβ signaling in *P. cinereus* is responsible for lung loss. In support of this hypothesis, several studies from mice illustrate the inhibitory effects of Tgfβ on lung development. Treatment of embryonic mouse foreguts with exogenous Tgfβ1 results in the formation of a small endodermal pouch, but no lung (Chen et al., 2007). The effect of Tgfβ1 on early lung development in foregut explants is dose dependent: at higher levels there is no evidence of lung formation at all (Chen et al., 2007). Excess Tgfβ signaling is inhibitory at later lung development stages, including suppressing lung branching and growth in culture (Serra et al., 1994) and alveolar development in embryos by promotion of apoptosis (Neptune et al., 2003). Transgenic ectopic expression of Tgfβ1 using a lung-specific promoter leads to lung developmental arrest and prevents cellular differentiation (Zhou et al., 1996).

To test whether Tgfβ signaling upregulation may inhibit lung development in salamanders, I increased Tgfβ signaling in lunged *A. mexicanum* by administration of exogenous Tgfβ1 protein to explanted thoracic segments. Tgfβ1 treatment results in repression of tracheal development, as indicated by loss of median SPC expression (Fig. 3.7). Lung buds still form, but they appear to emerge directly from the foregut, akin to type III tracheal agenesis (Floyd et al., 1962) and similar in appearance to mouse conditional mutants of Bmp4 (Li et

al., 2008b), conditional mutants of Bmpr1a/b (Domyan et al., 2011), and embryos treated with Bmp-signaling inhibitors (Rankin et al., 2012). This phenotype also resembles *shh* -/- mice, which have defects in tracheo-esophageal septation (Litingtung et al., 1998). Notably, the expression pattern of Sox9 throughout the laryngotracheal tube in *P. cinereus* indicates loss of tracheal development in plethodontids, but maintenance of lung development, which corresponds to the observed loss of the trachea in the explant system.

Tracheal agenesis phenotypes are likely mediated by imbalance of *nkx2.1* and *sox2* expression early in lung development. Sox2 and Nkx2.1 are mutually antagonistic, and *sox2* expression is also repressed by Bmp (Domyan et al., 2011; Li et al., 2008b). Loss of Bmp signaling causes expansion of *sox2* expression into the ventral foregut and loss of tracheal specification (Domyan et al., 2011). Expression differences between *P. cinereus* and *A. mexicanum* were analyzed after lung specification and therefore may not be directly comparable to published reports. Nevertheless, the expression of *nkx2.1* in *P. cinereus* is significantly lower than in *A. mexicanum* (Table 3.3). *Sox2* expression is elevated in *P. cinereus*, and although this upregulation is not statistically significant, *sox2* expression levels likely are underestimates because the ventral endoderm was targeted for sequencing (Fig. 3.4). Hence, the balance of tracheal and esophageal specification could be perturbed in *P. cinereus*. Bmp signaling does not appear to be reduced in *P. cinereus*, pointing to reduction of Nkx2.1 inductive factors such as Shh signaling and Foxa1 as a proximate cause of the observed downregulation of *nkx2.1*, and/or an increase in factors other than Sox2 that repress *nkx2.1* expression such as Tgfβ (Billmyre et al., 2015).

Increased Tgfβ signaling results in downregulation of *nkx2.1* and *fgf10* expression (Chen et al., 2010; Li et al., 2008a; Napolitano et al., 2000; Xing et al., 2008). The direct relationship

between *nkx2.1* expression and Tgfβ signaling likely explains observed loss of the trachea in Tgfβ-treated *A. mexicanum* (Fig. 3.7). In contrast, canonical Wnt activates *nkx2.1* expression through TCF/LEF cis-regulatory enhancers (Gilbert-Sirieix et al., 2011; Goss et al., 2009; Harris-Johnson et al., 2009). Elevation of Wnt signaling is observed in *P. cinereus*. However, the suppression of *nkx2.1* expression by Tgfβ is functionally dominant over any activating effect of Wnt (Li et al., 2013), and this dominance of Tgfβ may result in *nkx2.1* downregulation in *P. cinereus*.

Although Tgfβ has a clear role in lung specification and early lung development, the effects of Tgfβ downregulation at these stages not been carefully characterized in model organisms. Chen et al. (2007) used a Tgfβ antibody to reduce Tgfβ signaling in E8.5 mouse foregut explants and did not find an inhibitory effect on initial lung outgrowth, but they did not examine expression patterns of pulmonary markers. Tgfβ knockouts have relatively mild lung phenotypes in mice (Dünker and Krieglstein, 2000; Kaartinen et al., 1995). However, embryonic lung development and the expression of pulmonary markers have not been characterized in knockout mutants. I inhibited Tgfβ signaling in *A. mexicanum* by administering the Alk5 inhibitor A-83-01 during lung specification and early lung development stages (Fig. 3.8). Treatment with this Tgfβ signaling inhibitor results in expansion of lung specification into the pharynx and oral epithelium, as evidenced by expression of SPC in these tissues (Fig. 3.8). There are no apparent lung defects in these treated animals. Therefore, Tgfβ likely represses lung development in salamanders by restricting lung specification to a narrow axial window. This proposed repressive role matches the well-established role of Tgfβ in mammals documented in overexpression experiments and discussed above. Expansion of SPC expression into the pharynx and anterior foregut is a phenotype observed in several ex-

perimental manipulations in both mice and frogs. For instance, activation of Wnt signaling and increased RA signaling both result in expanded lung specification (Goss et al., 2009; Harris-Johnson et al., 2009; Rankin et al., 2012). Each pathway likely acts in a parallel fashion to the downregulation of Tgfβ by upregulating *nkx2.1* expression (Fig. 3.4).

I hypothesized that if increased Tgfβ signaling is responsible for lung loss in plethodontids, then inhibition of Tgfβ signaling would result in retention, rather than regression, of the lung rudiment. To inhibit Tgfβ signaling, I treated *P. cinereus* embryos with A-83-01 from stage 19 until stage 23. Stage 19 is after the emergence of the lung rudiment. Hence, treatment targeted Tgfβ signaling during lung growth and maintenance, rather than specification. Stage 23 is after the lung rudiment has fully regressed (Chapter 2). A-83-01 treatment resulted in retention of the lung rudiment in all individuals examined (N = 5). The lung rudiment could be resolved histologically as a long diverticulum, which terminated distally near the liver with an enlarged distal lumen. The diverticulum robustly expresses the distal pulmonary marker *sox9* (Fig. 3.9F). The distal terminus of the lung rudiment extended posteriorly farther than in individuals examined in Chapter 2, indicating that inhibition of Tgfβ signaling permits more extensive lung development. However, a similar structure, though slightly smaller and with low *sox9* expression, was observed in two of five controls. The presence of the lung rudiment in these control embryos casts doubt on the efficacy of the A-83-01 treatment. It is possible that A-83-01 treatment causes a general developmental delay, which results in lung rudiment retention. A-83-01 treated individuals were grossly normal and apparently had attained the same embryonic stage as DMSO control individuals based on digit and cranial phenotypes (Kerney, 2011). However, treated embryos appeared slightly smaller than controls. Optimally, the experiment should be repeated over a longer time course to

guarantee that the retained rudiment is not the result of developmental delay.

Given the significant upregulation of Tgfβ signaling in *P. cinereus* and the demonstrated role of Tgfβ in lung development, I propose that Tgfβ is involved in lung loss via downregulation of *nkx2.1* expression. Nkx2.1 is a critical node in the lung gene regulatory network and I have demonstrated expression differences in multiple genes that may result in *nkx2.1* downregulation, including reduced *foxa1* expression and Shh signaling. Tgfβ signaling, Shh signaling and *foxa1* expression are all regulated by retinoic acid (RA) signaling (Chen et al., 2007, 2010; Rankin and Zorn, 2014).

3.5.2 Retinoic acid signaling in lung loss

Retinoic acid plays pleiotropic roles in lung development by balancing Wnt and Tgfβ signaling, among other functions (Chen et al., 2010; Malpel et al., 2000). Decreased RA signaling would be predicted to increase Tgfβ signaling while decreasing *foxa1* and *shh* expression, which is precisely the situation observed in *P. cinereus* (Chen et al., 2010; Volckaert and De Langhe, 2015). RA levels were not directly assessed in lunged or lungless salamanders. However, it is likely that RA levels are lower in *P. cinereus* due to significant downregulation of the retinol dehydrogenase *rdh2*. Although lungs are specified and begin to develop in *Xenopus laevis* following inhibition of RA signaling, lung development is arrested (Rankin et al., 2012). Lung development arrests during the formation of the laryngotracheal tube in retinoid-deficient mouse embryos (Dickman et al., 1997). Slightly different results are obtained in RA-deficient mouse foregut explants: lungs do not form at all when RA signaling is inhibited prior to lung specification (Chen et al., 2007). Therefore, the gene expression

pattern and stage at lung developmental arrest in plethodontids appear to phenocopy the most frequent defects observed in RA-deficient embryos. However, treatment of *P. cinereus* embryos with 0.01 µM and 0.05 µM all-*trans*-RA from stages 20 through 23 failed to prevent lung regression (data not shown). Treatment of *A. mexicanum* embryos with DEAB, which acts as a competitive substrate for aldehyde dehydrogenase (Aldh), failed to repress lung development (data not shown). However, DEAB may fail to fully inhibit RA synthesis and does not serve to block RA signaling via stored RA (Koppaka et al., 2012). Future studies should attempt to inhibit other points in the RA-signaling pathway, such as signaling via retinoic acid receptors (e.g., BMS493; Rankin et al., 2012).

3.5.3 Increased Wnt signaling in lungless salamanders

As discussed above, multiple studies support a strong dependence on canonical Wnt signaling for initiation of lung development (Chen et al., 2010; Goss et al., 2009; Harris-Johnson et al., 2009; Rankin et al., 2012). Treatment of *A. mexicanum* embryos with the Tankyrase inhibitor XAV939 to reduce canonical Wnt signaling (Huang et al., 2009) resulted in partial or complete lung agenesis, confirming that salamanders are reliant on Wnt for lung specification and development (Fig. 3.10). Canonical Wnt signaling also plays a later role in distal lung development by promoting the formation of alveolar type II cells, while repressing proximal lung development (Shu et al., 2005). Key indicators of Wnt signaling, such as *axin2*, *apc*, *β-catenin, yap1* and *tnks*, were upregulated in Pc19 relative to Am40 (Jho et al., 2002; Konsavage et al., 2012; Li et al., 2013). Among these markers, *axin2* upregulation is considered the most reliable indicator of active canonical Wnt activation (Clevers, 2006; Jho

et al., 2002). Wnt signaling elevation accords with the robust expression of wnt2b observed in *P. cinereus* by *in situ* hybridization (Chapter 2).

Modulation of Wnt signaling has divergent effects depending on whether it occurs before or after lung specification. β-catenin stabilization with the Gsk3 inhibitor BIO after gastrulation and during somitogenesis (prior to lung specification) promotes midgut and hindgut fate at the expense of anterior endoderm derivatives such as the lung (Scott Rankin, personal communication). Stabilization of β-catenin just prior to lung specification results in expansion of lung specification into the pharynx, esophagus and foregut (Goss et al., 2009; Rankin et al., 2012). Lung phenotypes following BIO treatment in *Xenopus laevis* have not been fully characterized; lung buds are still present, but possibly malformed (Rankin et al., 2012; Scott Rankin, personal communication). Later periods of Wnt signaling primarily serve to promote distal alveolar epithelial identity (Shu et al., 2005). However, overactivation of Wnt signaling at these later stages inhibits lung differentiation (Okubo and Hogan, 2004).

Elevated Wnt signaling in *P. cinereus* may therefore have multiple effects depending on which stage it commences. For instance, overactivation of Wnt signaling following lung specification may result in distalization of the lung rudiment (Shu et al., 2005) or inhibition of lung differentiation (Okubo and Hogan, 2004). The expression of *sox9* throughout the lung rudiment (Chapter 2) argues that the lung rudiment has distal lung identity. Distal identity could result from either the failure of proximal structures to develop or the activation of a distal signaling cascade, possibly due to increased canonical Wnt signaling.

In addition to a general elevation in Wnt signaling in *P. cinereus* relative to *A. mexicanum*, Wnt signaling likely decreases between Pc19 and Pc21, as indicated by significantly reduced transcription of β-catenin as well as reduced expression (but not significantly) of

axin2, *apc*, *aida* and *yap1* at this later stage. Reduced Wnt signaling in the lung following lung outgrowth results in defects in lung morphogenesis and cellular differentiation but not in lung regression (Shu et al., 2005). Therefore, it is unlikely that decreased Wnt signaling before lung regression is directly responsible for lung loss.

3.5.4 A role for Fgf signaling in lung loss

Fgf signaling has multiple roles in the developing lung, including the control of lung outgrowth, branching (in mammals) and cell differentiation (Volckaert and De Langhe, 2015). Of the Fgf ligands found in the lung and lung mesenchyme, Fgf10 is the only essential protein for lung bud induction (Min et al., 1998). Mesenchymal Fgf10 initiates lung outgrowth (Min et al., 1998; De Moerlooze et al., 2000; Sekine et al., 1999; Volckaert et al., 2013). *Fgf10* is not differentially expressed between lunged and lungless transcriptomes at any stage, but the Fgf signaling indicator *etv4* (Liu et al., 2003) is significantly upregulated at Pc19 relative to Am40. However, *etv4* expression at Pc21 has diminished to levels below Am40, while *fgf10* expression was not observed at all (Fig. 3.4). This parallel reduction of *etv4* and *fgf10* expression indicates reduced Fgf signaling preceding lung regression. Several upstream regulators of *fgf10* expression have been identified, including Tgfβ, Tbx5 and Wnt. Tgfβ negatively regulates *fgf10* expression (Chen et al., 2010), hence upregulation of Tgfβ signaling may be responsible for a drop in Fgf signaling. Tbx5 is a positive regulator of *fgf10* expression (Arora et al., 2012), and the decreased expression of *tbx5* could also be responsible for reduction of *fgf10* and *etv4* transcripts preceding lung regression.

Fgf2 and *fgfr3* are significantly downregulated in lungless salamanders relative to lunged.

Both of these Fgf signaling-pathway members function in late lung development, e.g., alveolar development, but not in early lung formation (Morrisey and Hogan, 2010).

3.5.5 The reversibility of lung loss

Plethodontids are a diverse and ancient family of salamanders. Given the conserved expression of many members of the lung GRN as well as the formation of the rudimentary lung, it is intriguing to consider why complete lung development has not been regained in any plethodontid species. Dollo's law on the irreversibility of evolution loosely states that lost organs never re-evolve in their same form (Gould, 1970). There exist several notable instances of evolution in reverse such as the re-evolution of teeth following their loss (Kurten, 1963; Lynch and Wagner, 2010; Wiens, 2011). While these are well-supported instances of trait re-evolution, it remains to be seen whether any of these examples violate the strict definition of Dollo's Law: structures cannot re-evolve in the same form as a lost structure. Modern conceptions of reverse evolution place more emphasis on re-evolved organ function rather than whether conserved mechanisms pattern ancestral and re-gained organ development.

Marshall et al. (1994) calculate that reverse evolution is improbable after 10 million years of evolution, based on the mutation rates of silenced genes. However, their calculation assumes that genes are no longer being expressed and do not have pleiotropic roles in other developmental pathways. In contrast, a prime example of reversibility in evolution is the re-evolution of mandibular teeth in marsupial frogs (*Gastrotheca*) over 200 million years after their original loss (Wiens, 2011). In this instance, reverse evolution of mandibular teeth may be facilitated by the fact that teeth in the upper jaw still form and are maintained within

Gastrotheca. Reverse evolution is hypothetically more likely when homologous structures are conserved in other anatomical areas (such as conserved upper jaw teeth in frogs), or for meristic traits such as nipple number or caudal vertebrae number (Galis et al., 2010). Conversely, reverse evolution is constrained when organs develop within more sensitive early embryonic periods (Galis et al., 2010). Constraint in regain also may be due the improbability of acquiring suites of mutations to completely restore not only ancestral morphology, but also ancestral function (Aigler et al., 2014).

Provided that the mechanism of organ loss is understood, changing a single causative genetic pathway may or may not result in organ re-evolution (Aigler et al., 2014). Tissue recombination experiments in birds indicate that if competence to respond to inductive signals for tooth formation is maintained, then the presence of a functional inductive tissue can partially induce trait development following loss (Mitsiadis et al., 2003). However, fully differentiated teeth cannot re-evolve in chickens due to the loss of genes for enamel and dentine formation (Sire et al., 2008). In lung loss, markers of lung differentiation are present and expressed, indicating that continued lung development could theoretically result in functional, differentiated lungs.

Pulmonary mesenchyme induces and maintains lung development (McCulley et al., 2015). I found that transplantation of lunged salamander mesenchyme into lungless salamander embryos appears to restore lung development into post-embryonic stages of development (Fig. 3.11). This implies that the foregut endoderm remains competent to respond to pulmonary inductive cues despite the ancient loss of lungs. The foregut could maintain such competence provided that lung-patterning genes are not subjected to deleterious mutations. As discussed, pleiotropic roles of many of these genes likely prevent pseudogenization and

gene loss.

In spite of hypothesized competence to respond to inductive cues, fully formed lungs have not been regained in plethodontids. Bridgham and colleagues (2009) demonstrated constraint in evolution of one form of a glucocorticoid receptor to the sequence of a closely related ancestral protein. In this case, reverse evolution of glucocorticoid receptors is constrained not only by the necessary mutations that would need to occur to functional sites, but also by additional mutations, which would render the protein unstable if changes to functional sites should occur (Bridgham et al., 2009). Complex gene regulatory networks may evolve in a similar fashion: restoring ancestral developmental potential may require toleration of reduced overall fitness until critical permissive mutations are fixed.

3.5.6 Summary

On the face of it, evolutionary lung loss seems improbable. Not only are there potential drawbacks of lungless respiration, but also lungs develop within a highly integrated developmental network and their formation is governed by the expression and function of genes with highly pleiotropic roles in the development of organs such as the limbs, gut, heart and thyroid. Additionally, lungs influence the morphogenesis of other tissues such as the heart (Chapter 5; Hoffmann et al., 2009). By comparing the pulmonary gene regulatory networks between lunged and lungless species, I identify how differential expression at key developmental nodes has the potential to result in lung loss. I suggest that high Tgfβ signaling plays a causative role in lung loss. Tgfβ signaling integrates and controls several aspects of lung development, including the expression of *fgf10* and *nkx2.1*, which are both required for

lung development. Experimental up- and downregulation of Tgfβ signaling in the lunged salamander *A. mexicanum* represses and expands lung development, respectively. Downregulation of Tgfβ signaling in lungless *P. cinereus* results in the retention of the lung rudiment, but it remains to be determined whether this phenotype is mediated by changes in Tgfβ signaling or a developmental delay induced by administration of the pharmacological inhibitor. Promisingly, lunged salamander splanchnic mesoderm appears to restore lung development in lungless embryos and larvae.

While candidate gene approaches have illuminated some causative mechanisms underlying organ loss (e.g., Herrera et al., 2013), the scope of candidate gene approaches is necessarily narrow. Transcriptomic analysis of organ loss allows not only the determination of proximate factors immediately upstream of organ loss, but also reveals a comprehensive picture of the overarching genetic interactions that result in observed expression differences. Lungless salamanders are a unique system in that the pulmonary homolog can be directly sequenced. Future work focusing on sequencing lung transcriptomes from additional lunged and lungless species will help to confirm the mechanistic basis of salamander lung loss.

Table 3.2: Selected genes upregulated in the lung primordium of *Plethodon cinereus* at stage 19 relative to the lung primordium of *Ambystoma mexicanum* at stage 40. List is ordered by fold change. Abbreviations: *Amex*, *Ambystoma mexicanum*; *Pcin*, *Plethodon cinereus*

Unique ID	p-value	Adjusted p-value	Base Mean *Pcin*	Base Mean *Amex*	Log_2(fold change)	*Amex* Annotation	Short Gene ID	*Amex* contig ID	*Pcin* contig ID
12807	2.84E-03	2.72E-02	44.697	0	-Inf	Apoptosis facilitator Bcl-2-like protein 14 OS=Homo sapiens GN=BCL2L14	BCL2L14	c52403_g1	c503422_g2
15340	1.52E-03	1.65E-02	8.86	0	-Inf	Latent-transforming growth factor beta-binding protein 4 OS=Mus musculus GN=Ltbp4	Ltbp4	c820703_g1	c453336_g1
12991	1.41E-02	9.63E-02	8.81	0	-Inf	Activin receptor type-1B OS=Mus musculus GN=Acvr1b	Acvr1b	c538429_g1	c396492_g2
8591	3.66E-03	3.33E-02	8.447	0	-Inf	Mucin-5B OS=Homo sapiens GN=MUC5B	MUC5B	c416812_g2	c484185_g1
13557	2.24E-04	3.17E-03	5.994	0	-Inf	Mothers against decapentaplegic homolog 6 OS=Homo sapiens GN=SMAD6	SMAD6	c607771_g1	c225433_g1
2675	8.53E-03	6.54E-02	118.865	1.176	-6.659	Homeobox protein SIX5 OS=Mus musculus GN=Six5	Six5	c289509_g1	c511723_g2
7498	7.72E-07	1.92E-05	45.288	0.608	-6.218	Roundabout homolog 4 OS=Homo sapiens GN=ROBO4	ROBO4	c409824_g1	c513249_g3
16450	1.59E-05	3.03E-04	15.821	0.324	-5.608	Smad nuclear-interacting protein 1 OS=Homo sapiens GN=SNIP1	SNIP1	c953923_g1	c476584_g1
7200	3.78E-11	1.81E-09	200.008	4.508	-5.471	Cathepsin B OS=Gallus gallus GN=CTSB	CTSB	c407304_g2	c425315_g1
2948	2.72E-03	2.63E-02	26.307	0.649	-5.342	Transforming growth factor-beta receptor-associated protein 1 OS=Bos taurus GN=TGFBRAP1	TGFBRAP1	c304506_g1	c500055_g2
2276	1.16E-22	1.36E-20	168.473	4.588	-5.198	Yorkie homolog OS=Homo sapiens GN=YAP1	YAP1	c265761_g1	c497216_g1
15384	6.84E-11	3.17E-09	48.563	2.439	-4.316	Death-associated protein-like 1-A OS=Xenopus laevis GN=dapl1-a	dapl1-a	c824363_g1	c502619_g1
7004	7.40E-05	1.32E-03	14.077	0.732	-4.266	Metalloproteinase inhibitor 3 OS=Xenopus laevis GN=timp3	timp3	c405688_g1	c507827_g1
9080	5.50E-14	3.52E-12	110.345	5.921	-4.22	Retinoic acid-induced protein 1 OS=Homo sapiens GN=RAI1	RAI1	c419531_g3	c478558_g2
14443	1.29E-13	8.09E-12	34.712	2.074	-4.065	Apoptosis regulator R1 (Fragment) OS=Xenopus laevis		c705076_g1	c371261_g2
11546	6.44E-03	5.23E-02	290.927	20.683	-3.814	Tankyrase-1 OS=Homo sapiens GN=TNKS PE=1 SV=2	TNKS	c431729_g4	c499188_g1

Table 3.2 Continued

Unique ID	p-value	Adjusted p-value	Base Mean *Pcin*	Base Mean *Amex*	Log_2(fold change)	*Amex* Annotation	Short Gene ID	*Amex* contig ID	*Pcin* contig ID
10965	4.56E-03	3.97E-02	77.654	5.808	-3.741	Foxf1	FOXF1	c426096_g5	c507771_g4
7619	1.41E-05	2.71E-04	207.96	20.921	-3.313	Aldehyde dehydrogenase family 16 member A1 OS=Bos taurus GN=ALDH16A1	ALDH16A1	c410675_g1	c498246_g1
1820	6.09E-11	2.85E-09	98.453	10.14	-3.279	Axin interactor, dorsalization-associated protein OS=Mus musculus GN=Aida	Aida	c241825_g1	c414678_g1
11751	3.02E-03	2.84E-02	231.76	25.3	-3.195	Neurogenic locus notch homolog protein 1 OS=Xenopus laevis GN=notch1	notch1	c432841_g8	c487673_g3
926	2.41E-12	1.32E-10	20.336	2.434	-3.063	Bcl-2-like protein 11 OS=Homo sapiens GN=BCL2L11	BCL2L11	c15934_g1	c437088_g1
10942	2.97E-04	4.04E-03	38.938	4.682	-3.056	Latent-transforming growth factor beta-binding protein 1 OS=Homo sapiens GN=LTBP1 PE=1 SV=4	LTBP1	c428727_g3	c503464_g1
1644	4.43E-06	9.50E-05	67.655	8.847	-2.935	Mothers against decapentaplegic homolog 4 OS=Homo sapiens GN=SMAD4	SMAD4	c229865_g1	c484062_g1
884	2.11E-03	2.13E-02	29.936	3.95	-2.922	Wilms tumor protein homolog (Fragment) OS=Alligator mississippiensis GN=WT1	WT1	c156884_g1	c455377_g1
3212	6.30E-07	1.84E-05	57.099	7.634	-2.903	Metalloproteinase inhibitor 1 OS=Equus caballus GN=TIMP1	TIMP1	c323539_g1	c485911_g1
1446	1.56E-04	2.30E-03	39.73	5.517	-2.848	Msx2-interacting protein OS=Mus musculus GN=Spen	Spen	c219545_g1	c491201_g1
6361	5.44E-05	9.02E-04	17.113	2.685	-2.672	ETS translocation variant 4 OS=Homo sapiens GN=ETV4	ETV4	c399238_g1	c475887_g1
7796	1.66E-05	3.15E-04	26.443	4.383	-2.593	HHIP-like protein 1 OS=Mus musculus GN=Hhipl1	Hhipl1	c412098_g1	c483806_g1
2079	1.25E-02	8.80E-02	10.624	1.954	-2.443	Vascular endothelial growth factor A OS=Coturnix coturnix japonica GN=VEGFA	VEGFA	c255440_g1	c432942_g1
10213	2.25E-05	4.08E-04	99.612	18.467	-2.431	Axin2 OS=Xenopus laevis	Axin2	c425052_g1	c514573_g1
8941	4.50E-06	9.65E-05	63.199	12.878	-2.295	Slit homolog 3 protein OS=Mus musculus GN=Slit3	Slit3	c418800_g3	c519641_g3
10742	1.23E-04	1.86E-03	97.814	21.663	-2.175	Adenomatous polyposis coli protein OS=Homo sapiens GN=APC	APC	c427544_g5	c509536_g2

Table 3.2 Continued

Unique ID	p-value	Adjusted p-value	Base Mean *Pcin*	Base Mean *Amex*	Log_2(fold change)	*Amex* Annotation	Short Gene ID	*Amex* contig ID	*Pcin* contig ID
10880	1.22E-05	2.39E-04	57.792	13.001	-2.152	Bone morphogenetic protein receptor type-1B OS=Gallus gallus GN=BMPR1B	BMPR1B	c428450_g6	c507672_g2
3900	1.68E-03	1.78E-02	111.153	35.707	-1.638	Programmed cell death protein 6 OS=Mus musculus GN=Pdcd6	Pdcd6	c349281_g1	c228248_g1
5877	3.69E-09	1.38E-07	277.641	91.085	-1.608	Catenin beta OS=Xenopus laevis GN=ctnnb1	ctnnb1	c393131_g1	c370305_g1
5389	6.85E-05	1.11E-03	73.051	28.45	-1.36	Mitogen-activated protein kinase kinase kinase 7 OS=Rattus norvegicus GN=Map3k7	Map3k7	c386114_g1	c502993_g2
8875	1.26E-04	1.90E-03	281.212	111.414	-1.336	Cell cycle and apoptosis regulator protein 2 OS=Mus musculus GN=Ccar2	Ccar2	c418344_g2	c485901_g1
6789	4.68E-04	5.98E-03	10.089	4.055	-1.315	Protein delta homolog 1 OS=Homo sapiens GN=DLK1	DLK1	c403733_g3	c460194_g1

Table 3.3: Selected genes downregulated in the lung primordium of *Plethodon cinereus* at stage 19 relative to the lung primordium of *Ambystoma mexicanum* at stage 40. List is ordered by fold change. Abbreviations: *Amex*, *Ambystoma mexicanum*; *Pcin*, *Plethodon cinereus*

Unique ID	p-value	Adjusted p-value	Base Mean *Pcin*	Base Mean *Amex*	Log_2(fold change)	*Amex* Annotation	Short Gene ID	*Amex* contig ID	*Pcin* contig ID
7559	2.11E-04	3.00E-03	0	8.031	Inf	Fibroblast growth factor 2 OS=Monodelphis domestica GN=FGF2	FGF2	c410188_g1	c105053_g1
10576	7.45E-03	5.88E-02	0	9.283	Inf	Epidermal growth factor-like protein 6 OS=Xenopus laevis GN=egfl6	egfl6	c426701_g1	c78937_g1
8951	5.43E-09	1.98E-07	0	27.85	Inf	BMP-binding endothelial regulator protein OS=Homo sapiens GN=BMPER	BMPER	c418859_g6	c756295_g1
9014	1.12E-05	2.20E-04	0	23.685	Inf	Zinc finger protein GLI1 OS=Homo sapiens GN=GLI1 PE=1 SV=1	GLI1	c419261_g3	c332734_g1
1406	8.93E-65	3.93E-62	0.107	438.829	11.995	Pulmonary surfactant-associated protein C OS=Neovison vison GN=SFTPC	SFTPC	c21543_g1	c4587_g1
9160	1.01E-09	4.06E-08	0.566	166.801	8.203	Mucin-1 OS=Mesocricetus auratus GN=MUC1	MUC1	c419880_g1	c515092_g3
10517	4.17E-08	1.33E-06	0.605	54.54	6.493	Forkhead box protein O4 OS=Mus musculus GN=Foxo4	Foxo4	c426408_g4	c506303_g3
8885	6.50E-03	5.27E-02	2.435	145.154	5.897	Gastrokine-2 OS=Rattus norvegicus GN=Gkn2	Gkn2	c418412_g1	c428764_g1
12120	2.77E-05	4.92E-04	3.518	192.952	5.778	Homeobox protein Nkx-2.1 OS= Thamnophis sirtalis GN=Nkx2-1	Nkx2-1	c441853_g1	c830881_g1
12095	4.79E-03	4.11E-02	0.391	20.038	5.681	Msx2-interacting protein OS=Mus musculus GN=Spen	Spen	c439640_g1	c491202_g1
5513	1.30E-189	3.11E-186	3.541	171.485	5.598	Pulmonary surfactant-associated protein B OS=Rattus norvegicus GN=Sftpb	Sftpb	c387854_g2	c367617_g1
6204	2.95E-07	7.95E-06	2.399	93.316	5.282	Platelet-derived growth factor subunit A OS=Xenopus laevis GN=pdgfa	pdgfa	c397635_g1	c70951_g1
6526	1.26E-02	8.83E-02	1.13	30.41	4.75	Tafazzin OS=Macaca mulatta GN=TAZ	TAZ	c400643_g1	c334758_g2
10910	1.21E-02	8.55E-02	1.56	40.065	4.682	Latent-transforming growth factor beta-binding protein 1 OS=Mus musculus GN=Ltbp1	Ltbp1	c428598_g1	c490111_g7

Table 3.3 Continued

Unique ID	p-value	Adjusted p-value	Base Mean *Pcin*	Base Mean *Amex*	Log_2(fold change)	*Amex* Annotation	Short Gene ID	*Amex* contig ID	*Pcin* contig ID
6921	2.18E-10	9.65E-09	2.213	51.429	4.539	Epidermal retinol dehydrogenase 2 OS=Mus musculus GN=Sdr16c5	Sdr16c5	c405025_g2	c429356_g1
10282	3.34E-05	5.82E-04	5.704	129.211	4.502	Mothers against decapentaplegic homolog 3 OS=Gallus gallus GN=SMAD3	SMAD3	c425326_g1	c493317_g3
8085	1.38E-03	1.52E-02	1.075	21.922	4.35	Transforming growth factor beta-1-induced transcript 1 protein OS=Rattus norvegicus GN=Tgfb1i1	Tgfb1i1	c413915_g3	c263362_g1
15298	2.61E-82	2.08E-79	6.238	90.471	3.858	Pulmonary surfactant-associated protein A1 OS=Homo sapiens GN=SFTPA1	SFTPA1	c818299_g1	c1013790_g1
12847	4.29E-03	3.78E-02	0.283	3.692	3.705	Growth arrest-specific protein 6 OS=Mus musculus GN=Gas6	Gas6	c527118_g1	c946240_g1
8185	8.92E-05	1.39E-03	3.659	35.073	3.261	BCL-6 corepressor OS=Mus musculus GN=Bcor	Bcor	c414538_g1	c507964_g4
10877	5.29E-14	3.41E-12	59.572	555.897	3.222	Fibroblast growth factor receptor 3 OS=Pleurodeles waltl GN=FGFR3	FGFR3	c428443_g3	c506457_g1
7215	8.74E-06	1.77E-04	1.94	15.83	3.029	Apoptosis-inducing factor 3 OS=Mus musculus GN=Aifm3	Aifm3	c407397_g1	c238845_g1
8186	5.05E-03	4.29E-02	3.948	31.245	2.984	BCL-6 corepressor OS=Homo sapiens GN=BCOR	BCOR	c414538_g2	c507964_g2
3768	1.10E-30	1.85E-28	38.07	223.016	2.55	Insulin-like growth factor 1 receptor OS=Xenopus laevis GN=igf1r	igf1r	c343998_g1	c491456_g6
9406	1.25E-03	1.39E-02	216.547	1256.774	2.537	Mucin-5B OS=Homo sapiens GN=MUC5B	MUC5B	c421084_g1	c464689_g3
11521	1.04E-02	7.56E-02	17.634	101.254	2.522	Apoptosis inhibitor 5 OS=Gallus gallus GN=API5	API5	c431608_g4	c475130_g1
7277	4.40E-04	5.66E-03	16.767	75.861	2.178	Forkhead box protein N3 OS=Sus scrofa GN=FOXN3	FOXN3	c408062_g1	c494038_g1
1928	6.22E-03	5.10E-02	2.265	10.2	2.171	Platelet-derived growth factor C OS=Homo sapiens GN=PDGFC	PDGFC	c246572_g1	c409282_g2
4822	1.37E-02	9.43E-02	1.68	7.031	2.065	Slit homolog 2 protein (Fragment) OS=Rattus norvegicus GN=Slit2	Slit2	c374490_g2	c412607_g1
11390	1.09E-02	7.88E-02	35.264	123.389	1.807	Forkhead box protein A1-A OS=Xenopus laevis GN=foxa1-a	foxa1-a	c430948_g4	c512274_g3

Table 3.4: Selected genes upregulated in the lung primordium of *Plethodon cinereus* at stage 21 relative to the lung primordium of *Ambystoma mexicanum* at stage 40. List is ordered by fold change. Abbreviations: *Amex*, *Ambystoma mexicanum*; *Pcin*, *Plethodon cinereus*

Unique ID	p-value	Adjusted p-value	Base Mean *Pcin*	Base Mean *Amex*	Log_2(fold change)	*Amex* Annotation	Short Gene ID	*Amex* contig ID	*Pcin* contig ID
4286	3.772E-06	9.259E-05	152.899	0.000	-Inf	Latent-transforming growth factor beta-binding protein 4 OS=Mus musculus GN=Ltbp4	Ltbp4	c362489_g3	c154514_g1
12807	1.134E-02	8.554E-02	31.547	0.000	-Inf	Apoptosis facilitator Bcl-2-like protein 14 OS=Homo sapiens GN=BCL2L14	BCL2L14	c52403_g1	c503422_g2
694	5.698E-03	5.016E-02	25.824	0.000	-Inf	Pulmonary surfactant-associated protein D OS=Rattus norvegicus GN=Sftpd	Sftpd	c141841_g1	c513945_g2
7200	7.395E-10	3.439E-08	183.422	4.508	-5.346	Cathepsin B OS=Gallus gallus GN=CTSB	CTSB	c407304_g2	c425315_g1
16450	5.545E-04	7.469E-03	12.543	0.324	-5.273	Smad nuclear-interacting protein 1 OS=Homo sapiens GN=SNIP1	SNIP1	c953923_g1	c476584_g1
15384	1.034E-15	8.785E-14	87.541	2.439	-5.166	Death-associated protein-like 1-A OS=Xenopus laevis GN=dapl1-a	dapl1-a	c824363_g1	c502619_g1
2276	3.087E-20	3.828E-18	138.192	4.588	-4.913	Yorkie homolog OS=Homo sapiens GN=YAP1	YAP1	c265761_g1	c497216_g1
7875	3.184E-04	4.615E-03	12805.763	552.086	-4.536	Collagen, type 1, alpha 2 precursor OS=Xenopus tropicalis	COL1A2	c412651_g2	c489973_g1
7498	1.796E-03	1.991E-02	13.411	0.608	-4.462	Roundabout homolog 4 OS=Homo sapiens GN=ROBO4	ROBO4	c409824_g1	c513249_g3
7004	7.398E-05	1.321E-03	14.077	0.732	-4.266	Metalloproteinase inhibitor 3 OS=Xenopus laevis GN=timp3	timp3	c405688_g1	c507827_g1
1634	4.215E-03	3.943E-02	6.002	0.407	-3.881	Connective tissue growth factor OS=Bos taurus GN=CTGF	CTGF	c229364_g1	c422226_g1
9080	1.004E-12	6.418E-11	86.261	5.921	-3.865	Retinoic acid-induced protein 1 OS=Homo sapiens GN=RAI1	RAI1	c419531_g3	c478558_g2
4916	1.382E-02	9.950E-02	29.625	2.109	-3.812	Cathepsin L1 OS=Chlorocebus aethiops GN=CTSL	CTSL	c376980_g1	c491890_g1
10985	7.035E-03	5.871E-02	107.136	8.035	-3.737	TGF-beta-activated kinase 1 and MAP3K7-binding protein 2 OS=Rattus norvegicus GN=Tab2	Tab2	c429049_g2	c497290_g3
16760	2.567E-04	3.840E-03	12.661	0.973	-3.702	Latent-transforming growth factor beta-binding protein 2 OS=Mus musculus GN=Ltbp2	Ltbp2	c980457_g1	c381716_g1

Table 3.4 Continued

Unique ID	p-value	Adjusted p-value	Base Mean *Pcin*	Base Mean *Amex*	Log_2(fold change)	*Amex* Annotation	Short Gene ID	*Amex* contig ID	*Pcin* contig ID
10942	1.171E-04	1.981E-03	43.895	4.682	-3.229	Latent-transforming growth factor beta-binding protein 1 OS=Homo sapiens GN=LTBP1	LTBP1	c428727_g3	c503464_g1
1820	1.287E-10	6.567E-09	81.422	10.140	-3.005	Axin interactor, dorsalization-associated protein OS=Mus musculus GN=Aida	Aida	c241825_g1	c414678_g1
7103	6.048E-03	5.268E-02	2.591	0.324	-2.998	Retinol dehydrogenase 7 OS=Rattus norvegicus GN=Rdh7	Rdh7	c406462_g1	c486414_g1
3212	6.299E-07	1.840E-05	57.099	7.634	-2.903	Metalloproteinase inhibitor 1 OS=Equus caballus GN=TIMP1	TIMP1	c323539_g1	c485911_g1
1644	1.525E-04	2.484E-03	47.113	8.847	-2.413	Mothers against decapentaplegic homolog 4 OS=Homo sapiens GN=SMAD4	SMAD4	c229865_g1	c484062_g1
8941	4.649E-05	8.824E-04	65.991	12.878	-2.357	Slit homolog 3 protein OS=Mus musculus GN=Slit3	Slit3	c418800_g3	c519641_g3
8983	1.858E-04	2.938E-03	67.367	13.183	-2.353	Mitogen-activated protein kinase kinase kinase 1 OS=Homo sapiens GN=MAP3K1	MAP3K1	c419087_g2	c512953_g7
14443	1.280E-04	2.141E-03	9.421	2.074	-2.183	Apoptosis regulator R1 (Fragment) OS=Xenopus laevis		c705076_g1	c371261_g2
1627	1.250E-20	1.610E-18	2780.205	629.369	-2.143	SPARC OS=Coturnix coturnix japonica GN=SPARC PE=2 SV=1	SPARC	c229138_g1	c484311_g1
926	6.446E-03	5.512E-02	10.489	2.434	-2.108	Bcl-2-like protein 11 OS=Homo sapiens GN=BCL2L11	BCL2L11	c15934_g1	c437088_g1
1446	1.250E-02	9.184E-02	17.827	5.517	-1.692	Msx2-interacting protein OS=Mus musculus GN=Spen	Spen	c219545_g1	c491201_g1
9084	1.794E-03	1.991E-02	190.568	59.733	-1.674	Apoptotic chromatin condensation inducer in the nucleus OS=Mus musculus GN=Acin1	Acin1	c419547_g1	c518836_g6
3900	1.069E-03	1.299E-02	104.857	35.707	-1.554	Programmed cell death protein 6 OS=Mus musculus GN=Pdcd6	Pdcd6	c349281_g1	c228248_g1
11261	9.335E-04	1.159E-02	174.170	63.426	-1.457	Programmed cell death 6-interacting protein OS=Xenopus laevis GN=pdcd6ip	pdcd6ip	c430312_g4	c441256_g1
10742	1.228E-02	9.084E-02	53.539	21.663	-1.305	Adenomatous polyposis coli protein OS=Homo sapiens GN=APC	APC	c427544_g5	c509536_g2

Table 3.4 Continued

Unique ID	p-value	Adjusted p-value	Base Mean *Pcin*	Base Mean *Amex*	Log_2(fold change)	*Amex* Annotation	Short Gene ID	*Amex* contig ID	*Pcin* contig ID
5389	7.025E-04	9.145E-03	66.698	28.450	-1.229	Mitogen-activated protein kinase kinase kinase 7 OS=Rattus norvegicus GN=Map3k7	Map3k7	c386114_g1	c502993_g2
10880	9.924E-03	7.706E-02	27.688	13.001	-1.091	Bone morphogenetic protein receptor type-1B OS=Gallus gallus GN=BMPR1B	BMPR1B	c428450_g6	c507672_g2
5877	1.043E-02	7.986E-02	177.859	91.085	-0.965	Catenin beta OS=Xenopus laevis GN=ctnnb1	ctnnb1	c393131_g1	c370305_g1

Table 3.5: Selected genes downregulated in the lung primordium of *Plethodon cinereus* at stage 21 relative to the lung primordium of *Ambystoma mexicanum* at stage 40. List is ordered by fold change. Abbreviations: *Amex*, *Ambystoma mexicanum*; *Pcin*, *Plethodon cinereus*

Unique ID	p-value	Adjusted p-value	Base Mean *Pcin*	Base Mean *Amex*	Log_2(fold change)	*Amex* Annotation	Short Gene ID	*Amex* contig ID	*Pcin* contig ID
12847	1.244E-02	9.156E-02	0.000	3.692	Inf	Growth arrest-specific protein 6 OS=Mus musculus GN=Gas6	Gas6	c527118_g1	c946240_g1
5646	4.876E-05	9.193E-04	0.000	12.498	Inf	Protein sprouty homolog 1 OS=Homo sapiens GN=SPRY1	SPRY1	c390017_g3	c738250_g1
6526	1.205E-03	1.443E-02	0.000	30.410	Inf	Tafazzin OS=Macaca mulatta GN=TAZ	TAZ	c400643_g1	c334758_g2
10517	4.133E-08	1.504E-06	0.000	54.540	Inf	Forkhead box protein O4 OS=Mus musculus GN=Foxo4	Foxo4	c426408_g4	c506303_g3
12120	9.594E-07	2.700E-05	0.000	192.952	Inf	Homeobox protein Nkx-2.1 OS= Thamnophis sirtalis GN=Nkx2-1	Nkx2-1	c441853_g1	c830881_g1
1406	4.420E-51	1.897E-48	0.000	438.829	Inf	Pulmonary surfactant-associated protein C OS=Neovison vison GN=SFTPC	SFTPC	c21543_g1	c4587_g1
6204	5.968E-06	1.407E-04	1.512	93.316	5.948	Platelet-derived growth factor subunit A OS=Xenopus laevis GN=pdgfa	pdgfa	c397635_g1	c70951_g1
5134	1.361E-04	2.254E-03	0.216	10.141	5.553	Wilms tumor protein 1-interacting protein homolog OS=Xenopus laevis GN=wtip	wtip	c381261_g2	c338370_g1
7547	1.262E-02	9.251E-02	3.195	94.076	4.880	Retinol dehydrogenase 10 OS=Xenopus tropicalis GN=rdh10	rdh10	c410133_g1	c499654_g1
5386	7.075E-03	5.893E-02	7.037	128.357	4.189	Glycogen synthase kinase-3 alpha OS=Homo sapiens GN=GSK3A	GSK3A	c386102_g2	c243342_g1
9160	2.981E-03	2.998E-02	10.149	166.801	4.039	Mucin-1 OS=Mesocricetus auratus GN=MUC1	MUC1	c419880_g1	c515092_g3
3849	1.371E-02	9.891E-02	1.080	16.942	3.972	Vang-like protein 2-A OS=Xenopus laevis GN=vangl2-a	vangl2-a	c346731_g1	c325189_g1
8951	6.763E-05	1.228E-03	1.899	27.850	3.874	BMP-binding endothelial regulator protein OS=Homo sapiens GN=BMPER	BMPER	c418859_g6	c756295_g1
15298	8.356E-58	4.372E-55	6.798	90.471	3.734	Pulmonary surfactant-associated protein A1 OS=Homo sapiens GN=SFTPA1	SFTPA1	c818299_g1	c1013790_g1

Table 3.5 Continued

Unique ID	p-value	Adjusted p-value	Base Mean *Pcin*	Base Mean *Amex*	Log_2(fold change)	*Amex* Annotation	Short Gene ID	*Amex* contig ID	*Pcin* contig ID
9782	4.259E-03	3.976E-02	0.432	4.871	3.496	Vascular endothelial growth factor receptor 1 OS=Gallus gallus GN=FLT1	FLT1	c422886_g1	c231021_g1
10877	2.070E-13	1.386E-11	51.008	555.897	3.446	Fibroblast growth factor receptor 3 OS=Pleurodeles waltl GN=FGFR3	FGFR3	c428443_g3	c506457_g1
8185	3.764E-03	3.611E-02	3.899	35.073	3.169	BCL-6 corepressor OS=Mus musculus GN=Bcor	Bcor	c414538_g1	c507964_g4
6921	1.201E-04	2.023E-03	7.758	51.429	2.729	Epidermal retinol dehydrogenase 2 OS=Mus musculus GN=Sdr16c5	Sdr16c5	c405025_g2	c429356_g1
7215	2.601E-03	2.699E-02	2.604	15.830	2.604	Apoptosis-inducing factor 3 OS=Mus musculus GN=Aifm3	Aifm3	c407397_g1	c238845_g1
11390	3.972E-03	3.762E-02	20.680	123.389	2.577	Forkhead box protein A1-A OS=Xenopus laevis GN=foxa1-a	foxa1-a	c430948_g4	c512274_g3
10206	1.264E-06	3.446E-05	6.707	39.367	2.553	Cathepsin D OS=Chionodraco hamatus GN=ctsd	ctsd	c425030_g3	c509674_g3
9014	4.400E-03	4.085E-02	4.402	23.685	2.428	Zinc finger protein GLI1 OS=Homo sapiens GN=GLI1	GLI1	c419261_g3	c332734_g1
5905	3.795E-03	3.631E-02	4.014	19.919	2.311	Retinoic acid-induced protein 1 OS=Homo sapiens GN=RAI1	RAI1	c393583_g1	c469672_g3
10892	1.186E-02	8.857E-02	8.859	39.241	2.147	T-box transcription factor TBX5 OS=Gallus gallus GN=TBX5	TBX5	c428542_g2	c408294_g1
4985	6.954E-03	5.821E-02	105.273	393.992	1.904	Msx2-interacting protein OS=Homo sapiens GN=SPEN	SPEN	c378409_g1	c509930_g2

Table 3.6: Selected genes upregulated in the lung primordium of *Plethodon cinereus* at stage 21 relative to *P. cinereus* at stage 19. List is ordered by fold change.

Unique ID	p-value	Adjusted p-value	baseMean Pc19	baseMean Pc21	Log_2(fold change)	Annotation	Short Gene ID	*Plethodon* contig ID
7559	6.778E-02	1.291E-03	0.000	2.591	Inf	Fibroblast growth factor 2 OS=Monodelphis domestica GN=FGF2	FGF2	c105053_g1
7467	3.948E-02	5.947E-04	0.000	4.107	Inf	Pro-cathepsin H OS=Sus scrofa GN=CTSH	CTSH	c406924_g2
16760	8.534E-04	5.385E-06	0.215	12.661	5.880	Latent-transforming growth factor beta-binding protein 2 OS=Mus musculus GN=Ltbp2	Ltbp2	c381716_g1
9973	7.348E-03	7.107E-05	1.505	53.484	5.152	Platelet endothelial cell adhesion molecule OS=Sus scrofa GN=PECAM1	PECAM1	c488672_g1

Table 3.7: Selected genes downregulated in the lung primordium of *Plethodon cinereus* at stage 21 relative to *P. cinereus* at stage 19. List is ordered by fold change.

Unique ID	p-value	Adjusted p-value	baseMean Pc19	baseMean Pc21	Log_2(fold change)	Annotation	Short Gene ID	*Plethodon* contig ID
5711	6.631E-02	1.243E-03	10.640	0.000	-Inf	Cell division cycle protein 20 homolog B OS=Homo sapiens GN=CDC20B	CDC20B	c278195_g1
9782	1.420E-03	9.466E-06	9.402	0.432	-4.444	Vascular endothelial growth factor receptor 1 OS=Gallus gallus GN=FLT1	FLT1	c231021_g1
8046	4.467E-02	7.323E-04	21.831	2.164	-3.335	Bcl-2-associated transcription factor 1 OS=Mus musculus GN=Bclaf1	Bclaf1	c490882_g3
7796	2.882E-02	3.926E-04	26.443	5.742	-2.203	HHIP-like protein 1 OS=Mus musculus GN=Hhipl1	Hhipl1	c483806_g1
14443	6.479E-02	1.173E-03	34.712	9.421	-1.882	Apoptosis regulator R1 (Fragment) OS=Xenopus laevis	2	c371261_g2
9471	1.603E-03	1.097E-05	512.343	151.092	-1.762	Insulin-like growth factor 2 mRNA-binding protein 1 OS=Gallus gallus GN=IGF2BP1	IGF2BP1	c489495_g1
926	1.682E-02	1.934E-04	20.336	10.489	-0.955	Bcl-2-like protein 11 OS=Homo sapiens GN=BCL2L11	BCL2L11	c437088_g1
5877	6.499E-02	1.184E-03	277.641	177.859	-0.642	Catenin beta OS=Xenopus laevis GN=ctnnb1	ctnnb1	c370305_g1

3.6 References

Aigler, S.R., Jandzik, D., Hatta, K., Uesugi, K., and Stock, D.W. (2014). Selection and constraint underlie irreversibility of tooth loss in cypriniform fishes. Proc. Natl. Acad. Sci. U. S. A. 111, 7707–7712.

Alescio, T., and Cassini, A. (1962). Induction *in vitro* of tracheal buds by pulmonary mesenchyme grafted on tracheal epithelium. J. Exp. Zool. 150, 83–94.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol 11, R106.

Annes, J.P., Munger, J.S., and Rifkin, D.B. (2003). Making sense of latent TGF-beta activation. J. Cell Sci. 116, 217–224.

Arora, R., Metzger, R.J., and Papaioannou, V.E. (2012). Multiple roles and interactions of Tbx4 and Tbx5 in development of the respiratory system. PLoS Genet. 8, e1002866.

Benson, M., Carlsson, B., Carlsson, L.M.S., Mostad, P., Svensson, P.-A., and Cardell, L.-O. (2002). DNA microarray analysis of Transforming Growth Factor-β and related transcripts in nasal biopsies from patients with allergic rhinitis. Cytokine 18, 20–25.

Billmyre, K.K., Hutson, M., and Klingensmith, J. (2015). One shall become two: Separation of the esophagus and trachea from the common foregut tube. Dev. Dyn. 244, 277–288.

Bordzilovskaya, N., Dettlaff, T., Duhon, S., and Malacinski, G. (1989). Developmental-stage series of axolotl embryos. In Developmental Biology of the Axolotl, J. Armstrong, and G. Malacinski, eds. (Oxford: Oxford University Press), pp. 201–219.

Bridgham, J.T., Ortlund, E.A., and Thornton, J.W. (2009). An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. Nature 461, 515–519.

de Caestecker, M. (2004). The transforming growth factor-β superfamily of receptors. Cytokine Growth Factor Rev. 15, 1–11.

Chen, F., Desai, T.J., Qian, J., Niederreither, K., Lü, J., and Cardoso, W. V (2007). Inhibition of Tgfβ signaling by endogenous retinoic acid is essential for primary lung bud induction. Development 134, 2969–2979.

Chen, F., Cao, Y., Qian, J., Shao, F., Niederreither, K., and Cardoso, W. V (2010). A retinoic acid–dependent network in the foregut controls formation of the mouse lung primordium. J. Clin. Invest. 120, 2040–2048.

Chuang, P., and McMahon, A.P. (1999). Vertebrate Hedgehog signaling modulated by induction of a Hedgehog-binding protein. Nature 397, 617–622.

Chute, J.P., Muramoto, G.G., Whitesides, J., Colvin, M., Safi, R., Chao, N.J., and McDonnell, D.P. (2006). Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells. Proc. Natl. Acad. Sci. U. S. A. 103, 11707–11712.

Clevers, H. (2006). Wnt/beta-catenin signaling in development and disease. Cell 127, 469–480.

Darwin, C. (1859). On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life (London: John Murray).

Dickman, E.D., Thaller, C., and Smith, S.M. (1997). Temporally-regulated retinoic acid depletion produces specific neural crest, ocular and nervous system defects. Development 124, 3111–3121.

Dijke, P. ten, and Hill, C.S. (2004). New insights into TGF-β–Smad signalling. Trends Biochem. Sci. 29, 265–273.

Domyan, E.T., Ferretti, E., Throckmorton, K., Mishina, Y., Nicolis, S.K., and Sun, X. (2011). Signaling through BMP receptors promotes respiratory identity in the foregut via repression of Sox2. Development 138, 971–981.

Duncan, M.R., Frazier, K.S., Abramson, S., Williams, S., Klapper, H., Huang, X., and Grotendorse, G.R. (1999). Connective tissue growth factor mediates transforming growth factor-β-induced collagen synthesis: down-regulation by cAMP. Faseb J 13, 1774–1786.

Dünker, N., and Krieglstein, K. (2000). Targeted mutations of transforming growth factor-beta genes reveal important roles in mouse development and adult homeostasis. Eur. J. Biochem. 267, 6982–6988.

Floyd, J., Campbell Jr, D.C., and Dominy, D.E. (1962). Agenesis of the trachea. Am. Rev. Respir. Dis. 86, 557–560.

Fong, D.W., Kane, T.C., and Culver, D.C. (1995). Vestigialization and loss of nonfunctional characters. Annu. Rev. Ecol. Syst. 26, 249–268.

Galis, F., Arntzen, J.W., and Lande, R. (2010). Dollo's law and the irreversibility of digit loss in *Bachia*. Evolution 64, 2466–2476; discussion 2477–2485.

Gilbert-Sirieix, M., Makoukji, J., Kimura, S., Talbot, M., Caillou, B., Massaad, C., and Massaad-Massade, L. (2011). Wnt/Beta-Catenin signaling pathway is a direct enhancer of thyroid transcription Factor-1 in human papillary thyroid carcinoma cells. PLoS One 6, 1–9.

Goss, A.M., Tian, Y., Tsukiyama, T., Cohen, E.D., Zhou, D., Lu, M.M., Yamaguchi, T.P., and Morrisey, E.E. (2009). Wnt2/2b and beta-catenin signaling are necessary and sufficient to specify lung progenitors in the foregut. Dev. Cell 17, 290–298.

Gould, S.J. (1970). Dollo on Dollo's law: irreversibility and the status of evolutionary laws. J. Hist. Biol. 3, 189–212.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D. a, Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.

Grindley, J.C., Bellusci, S., Perkins, D., and Hogan, B.L. (1997). Evidence for the involvement of the Gli gene family in embryonic mouse lung development. Dev. Biol. 188, 337–348.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494–1512.

Harris-Johnson, K.S., Domyan, E.T., Vezina, C.M., and Sun, X. (2009). Beta-catenin promotes respiratory progenitor identity in mouse foregut. Proc. Natl. Acad. Sci. U. S. A. 106, 16287–16292.

Hayashi, H., and Sakai, T. (2012). Biological significance of local TGF-β activation in liver diseases. Front. Physiol. 3, 1–11.

Herrera, A.M., Shuster, S.G., Perriton, C.L., and Cohn, M.J. (2013). Developmental basis of phallus reduction during bird evolution. Curr. Biol. 23, 1065–1074.

Hoffmann, A.D., Peterson, M.A., Friedland-Little, J.M., Anderson, S.A., and Moskowitz, I.P. (2009). Sonic hedgehog is required in pulmonary endoderm for atrial septation. Development 136, 1761–1770.

Huang, S.-M.A, Mishina, Y.M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G.A, Charlat, O., Wiellette, E., Zhang, Y., Wiessner, S., et al. (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. Nature 461, 614–620.

Hurney, C.A., Babcock, S.K., Shook, D.R., Pelletier, T.M., Turner, S.D., Maturo, J., Cogbill, S., Snow, M.C., and Kinch, K. (2015). Normal table of embryonic development in the four-toed salamander, *Hemidactylium scutatum*. Mech. Dev. 136, 99–110.

Ikeda, K., Clark, J.C., Shaw-White, J.R., Stahlman, M.T., Boutell, C.J., and Whitsett, J. a (1995). Gene structure and expression of human thyroid transcription factor-1 in respiratory epithelial cells. J. Biol. Chem. 270, 8108–8114.

Jeffery, W.R. (2009). Regressive evolution in *Astyanax* cavefish. Annu. Rev. Genet. 43, 25–47.

Jho, E.-H., Zhang, T., Domon, C., Joo, C.-K., Freund, J.-N., and Costantini, F. (2002). Wnt/Beta-Catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of

the signaling pathway. Mol. Cell. Biol. 22, 1172–1183.

Kaartinen, V., Voncken, J.W., Shuler, C., Warburton, D., Bu, D., Heisterkamp, N., and Groffen, J. (1995). Abnormal lung development and cleft palate in mice lacking TGF-beta3 indicates defects of epithelial-mesenchymal interaction. Nat. Genet. 11, 415–421.

Kerney, R. (2011). Embryonic staging table for a direct-developing salamander, *Plethodon cinereus* (Plethodontidae). Anat. Rec. 294, 1796–1808.

Kitisin, K., Saha, T., Blake, T., Golestaneh, N., Deng, M., Kim, C., Tang, Y., Shetty, K., Mishra, B., and Mishra, L. (2007). TGF-beta signaling in development. Sci. STKE 2007, 1–5.

Koli, K., Wempe, F., Sterner-Kock, A., Kantola, A., Komor, M., Hofmann, W.-K., von Melchner, H., and Keski-Oja, J. (2004). Disruption of LTBP-4 function reduces TGF-beta activation and enhances BMP-4 signaling in the lung. J. Cell Biol. 167, 123–133.

Konsavage, W.M., Kyler, S.L., Rennoll, S.A., Jin, G., and Yochum, G.S. (2012). Wnt/β-Catenin signaling regulates Yes-associated Protein (YAP) gene expression in colorectal carcinoma cells. J. Biol. Chem. 287, 11730–11739.

Koppaka, V., Thompson, D.C., Chen, Y., Ellermann, M., Nicolaou, K.C., Juvonen, R.O., Petersen, D., Deitrich, R.A, Hurley, T.D., and Vasiliou, V. (2012). Aldehyde dehydrogenase inhibitors: a comprehensive review of the pharmacology, mechanism of action, substrate specificity, and clinical application. Pharmacol. Rev. 64, 520–539.

Kurten, B. (1963). Return of a lost structure in the evolution of the felid dentition. Comment. Biol. Soc. Sci. Fenn. 26, 1–12.

Li, C., Li, A., Xing, Y., Li, M., Chan, B., Ouyang, R., Taketo, M.M., Kucherlapati, R., Borok, Z., and Minoo, P. (2013). Apc deficiency alters pulmonary epithelial cell fate and inhibits Nkx2.1 via triggering TGF-beta signaling. Dev. Biol. 378, 13–24.

Li, M., Li, C., Liu, Y., Xing, Y., Hu, L., Borok, Z., Kwong, K.Y.-C., and Minoo, P. (2008a). Mesodermal deletion of transforming growth factor-beta receptor II disrupts lung epithelial morphogenesis: cross-talk between TGF-beta and Sonic hedgehog pathways. J. Biol. Chem. 283, 36257–36264.

Li, Y., Gordon, J., Manley, N.R., Litingtung, Y., and Chiang, C. (2008b). Bmp4 is required for tracheal formation: a novel mouse model for tracheal agenesis. Dev. Biol. 322, 145–155.

Liem, K.F. (1988). Form and function of lungs: The evolution of air breathing mechanisms. Integr. Comp. Biol. 28, 739–759.

Litingtung, Y., Lei, L., Westphal, H., and Chiang, C. (1998). Sonic hedgehog is essential to foregut development. Nat. Genet. 20, 58–61.

Liu, Y., Jiang, H., Crawford, H.C., and Hogan, B.L.M. (2003). Role for ETS domain transcription factors Pea3/Erm in mouse lung development. Dev. Biol. 261, 10–24.

Lonnberg, E. (1899). Salamanders with and without lungs. Zool. Anz. 22, 545–548.

Lynch, V.J., and Wagner, G.P. (2010). Did egg-laying boas break Dollo's law? Phylogenetic evidence for reversal to oviparity in sand boas (Eryx: Boidae). Evolution 64, 207–216.

Malpel, S., Mendelsohn, C., and Cardoso, W. V (2000). Regulation of retinoic acid signaling during lung morphogenesis. Development 127, 3057–3067.

Marshall, C.R., Raff, E.C., and Raff, R.A. (1994). Dollo's Law and the death and resurrection of genes. Proc. Natl. Acad. Sci. U. S. A. 91, 12283–12287.

Massagué, J., and Gomis, R.R. (2006). The logic of TGFβ signaling. FEBS Lett. 580, 2811–2820.

McCulley, D., Wienhold, M., and Sun, X. (2015). The pulmonary mesenchyme directs lung development. Curr. Opin. Genet. Dev. 32, 98–105.

McCune, A.R., and Carlson, R.L. (2004). Twenty ways to lose your bladder: common natural mutants in zebrafish and widespread convergence of swim bladder loss among teleost fishes. Evol. Dev. 6, 246–259.

Metzger, R., and Krasnow, M. (1999). Genetic control of branching morphogenesis. Science 284, 1635–1639.

Metzger, R.J., Klein, O.D., Martin, G.R., and Krasnow, M.A. (2008). The branching programme of mouse lung development. Nature 453, 745–750.

Min, H., Danilenko, D.M., Scully, S. a., Bolon, B., Ring, B.D., Tarpley, J.E., DeRose, M., and Simonet, W.S. (1998). Fgf-10 is required for both limb and lung development and exhibits striking functional similarity to *Drosophila* branchless. Genes Dev. 12, 3156–3161.

Minoo, P., Hu, L., Xing, Y., Zhu, N.L., Chen, H., Li, M., Borok, Z., and Li, C. (2007). Physical and functional interactions between homeodomain NKX2.1 and winged helix/forkhead FOXA1 in lung epithelial cells. Mol. Cell. Biol. 27, 2155–2165.

Mitsiadis, T.A., Chéraud, Y., Sharpe, P., and Fontaine-Pérus, J. (2003). Development of teeth in chick embryos after mouse neural crest transplantations. Proc. Natl. Acad. Sci. U.S.A. 100, 6541–6545.

De Moerlooze, L., Spencer-Dene, B., Revest, J.M., Hajihosseini, M., Rosewell, I., and Dickson, C. (2000). An important role for the IIIb isoform of fibroblast growth factor receptor 2 (FGFR2) in mesenchymal-epithelial signalling during mouse organogenesis. Development 127, 483–492.

Morrisey, E.E., and Hogan, B.L.M. (2010). Preparing for the first breath: genetic and

cellular mechanisms in lung development. Dev. Cell 18, 8–23.

Munger, J.S., Harpel, J.G., Gleizes, P.E., Mazzieri, R., Nunes, I., and Rifkin, D.B. (1997). Latent transforming growth factor-beta: structural features and mechanisms of activation. Kidney Int. 51, 1376–1382.

Napolitano, G., Montani, V., Giuliani, C., Di Vincenzo, S., Bucci, I., Todisco, V., Laglia, G., Coppa, A., Singer, D.S., Nakazato, M., et al. (2000). Transforming growth factor-beta1 down-regulation of major histocompatibility complex class I in thyrocytes: coordinate regulation of two separate elements by thyroid-specific as well as ubiquitous transcription factors. Mol. Endocrinol. 14, 486–505.

Neptune, E.R., Frischmeyer, P. A, Arking, D.E., Myers, L., Bunton, T.E., Gayraud, B., Ramirez, F., Sakai, L.Y., and Dietz, H.C. (2003). Dysregulation of TGF-beta activation contributes to pathogenesis in Marfan syndrome. Nat. Genet. 33, 407–411.

Okubo, T., and Hogan, B.L.M. (2004). Hyperactive Wnt signaling changes the developmental potential of embryonic lung endoderm. J. Biol. 3, 11.

Piekarski, N., Gross, J.B., and Hanken, J. (2014). Evolutionary innovation and conservation in the embryonic derivation of the vertebrate skull. Nat. Commun. 5, 5661.

Que, J., Okubo, T., Goldenring, J.R., Nam, K.-T., Kurotani, R., Morrisey, E.E., Taranova, O., Pevny, L.H., and Hogan, B.L.M. (2007). Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. Development 134, 2521–2531.

Rajagopal, J., Carroll, T.J., Guseh, J.S., Bores, S.A., Blank, L.J., Anderson, W.J., Yu, J., Zhou, Q., Melton, D.A., and McMahon, A.P. (2008). Wnt7b stimulates embryonic lung growth by coordinately increasing the replication of epithelium and mesenchyme. Development 135, 1625–1634.

Rankin, S.A., and Zorn, A.M. (2014). Gene regulatory networks governing lung specification. J. Cell. Biochem. 115, 1343–1350.

Rankin, S.A., Gallas, A.L., Neto, A., Gómez-Skarmeta, J.L., and Zorn, A.M. (2012). Suppression of Bmp4 signaling by the zinc-finger repressors Osr1 and Osr2 is required for Wnt/β-catenin-mediated lung specification in *Xenopus*. Development 139, 3010–3020.

Rankin, S.A., Thi Tran, H., Wlizla, M., Mancini, P., Shifley, E.T., Bloor, S.D., Han, L., Vleminckx, K., Wert, S.E., and Zorn, A.M. (2015). A molecular atlas of *Xenopus* respiratory system development. Dev. Dyn. 244, 69–85.

Reagan, N.L., and Verrell, P.A. (1991). The evolution of plethodontid salamanders: did terrestrial mating facilitate lunglessness? Am. Nat. 138, 1307–1313.

Russo, J.E., Hauquitz, D., and Hilton, J. (1988). Inhibition of mouse cytoslic aldehyde dehydrogenase by 4-(diethylamino)benzaldehyde. Biochem. Pharmacol. 37, 1639–1642.

Ryckebusch, L., Wang, Z., Bertrand, N., Lin, S.-C., Chi, X., Schwartz, R., Zaffran, S., and Niederreither, K. (2008). Retinoic acid deficiency alters second heart field formation. Proc. Natl. Acad. Sci. U.S.A. 105, 2913–2918.

Sagai, T., Amano, T., Tamura, M., Mizushina, Y., Sumiyama, K., and Shiroishi, T. (2009). A cluster of three long-range enhancers directs regional Shh expression in the epithelial linings. Development 136, 1665–1674.

Sekine, K., Ohuchi, H., Fujiwara, M., Yamasaki, M., Yoshizawa, T., Sato, T., Yagishita, N., Matsui, D., Koga, Y., Itoh, N., et al. (1999). Fgf10 is essential for limb and lung formation. Nat. Genet. 21, 138–141.

Serls, A.E., Doherty, S., Parvatiyar, P., Wells, J.M., and Deutsch, G.H. (2005). Different thresholds of fibroblast growth factors pattern the ventral foregut into liver and lung. Development 132, 35–47.

Serra, R., Pelton, R.W., and Moses, H.L. (1994). TGF beta 1 inhibits branching morphogenesis and N-myc expression in lung bud organ cultures. Development 120, 2153–2161.

Shi, Y., and Massague, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. Cell 113, 685–700.

Shu, W., Guttentag, S., Wang, Z., Andl, T., Ballard, P., Lu, M.M., Piccolo, S., Birchmeier, W., Whitsett, J.A., Millar, S.E., et al. (2005). Wnt/beta-catenin signaling acts upstream of N-myc, BMP4, and FGF signaling to regulate proximal-distal patterning in the lung. Dev. Biol. 283, 226–239.

Sire, J.-Y., Delgado, S.C., and Girondot, M. (2008). Hen's teeth with enamel cap: from dream to impossibility. BMC Evol. Biol. 8, 246.

Spooner, B.S., and Wessells, N.K. (1970). Mammalian lung development: interactions in primordium formation and bronchial morphogenesis. J. Exp. Zool. 175, 455–466.

Sterner-Kock, a, Sterner-Kock, a, Thorey, I.S., Thorey, I.S., Koli, K., Koli, K., Wempe, F., Wempe, F., Otte, J., Otte, J., et al. (2002). Disruption of the gene encoding the latent transforming growth factor-beta binding protein 4 (LTBP-4) causes abnormal lung development, cardiomyopathy, and colorectal cancer. Genes Dev 16, 2264–2273.

Tojo, M., Hamashima, Y., Hanyu, A., Kajimoto, T., Saitoh, M., Miyazono, K., Node, M., and Imamura, T. (2005). The ALK-5 inhibitor A-83-01 inhibits Smad signaling and epithelial-to-mesenchymal transition by transforming growth factor-beta. Cancer Sci. 96, 791–800.

Verrecchia, F., Chu, M.L., and Mauviel, A. (2001). Identification of novel TGF-beta/Smad gene targets in dermal fibroblasts using a combined cDNA microarray/promoter transactivation approach. J. Biol. Chem. 276, 17058–17062.

Volckaert, T., and De Langhe, S.P. (2015). Wnt and FGF mediated epithelial-mesenchymal crosstalk during lung development. Dev. Dyn. 244, 342–366.

Volckaert, T., Campbell, A., Dill, E., Li, C., Minoo, P., and De Langhe, S. (2013). Localized Fgf10 expression is not required for lung branching morphogenesis but prevents differentiation of epithelial progenitors. Development 140, 3731–3742.

Vrijens, K., Lin, W., Cui, J., Farmer, D., Low, J., Pronier, E., Zeng, F.-Y., Shelat, A.A, Guy, K., Taylor, M.R., et al. (2013). Identification of small molecule activators of BMP signaling. PLoS One 8, e59045.

Wiens, J.J. (2011). Re-evolution of lost mandibular teeth in frogs after more than 200 million years, and re-evaluating Dollo's law. Evolution 65, 1283–1296.

Wilder, I.W., and Dunn, E. (1920). The correlation of lunglessness in salamanders with a mountain brook habitat. Copeia 1920, 63–68.

Wilkinson, M., and Nussbaum, R.A. (1997). Comparative morphology and evolution of the lungless caecilian *Atretochoana eiselti* (Taylor) (Amphibia: Gymnophiona: Typhlonectidae). Biol. J. Linn. Soc. 62, 39–109.

Xing, Y., Li, C., Hu, L., Tiozzo, C., Li, M., Chai, Y., Bellusci, S., Anderson, S., and Minoo, P. (2008). Mechanisms of TGF-beta inhibition of LUNG endodermal morphogenesis: The role of TgfBRII, Smads, Nkx2.1 and Pten. Dev. Biol. 320, 340–350.

Yoshinaga, K., Obata, H., Jurukovski, V., Mazzieri, R., Chen, Y., Zilberberg, L., Huso, D., Melamed, J., Prijatelj, P., Todorovic, V., et al. (2008). Perturbation of transforming growth factor (TGF)-beta1 association with latent TGF-beta binding protein yields inflammation and tumors. Proc. Natl. Acad. Sci. U. S. A. 105, 18758–18763.

Zhou, L., Dey, C.R., Wert, S.E., and Whitsett, J.A. (1996). Arrested lung morphogenesis in transgenic mice bearing an SP-C-TGF-B1 chimeric gene. Dev. Biol. 175, 227–238.

Zhou, Y., Koli, K., Hagood, J.S., Miao, M., Mavalli, M., Rifkin, D.B., and Murphy-Ullrich, J.E. (2009). Latent transforming growth factor-beta-binding protein-4 regulates transforming growth factor-beta1 bioavailability for activation by fibrogenic lung fibroblasts in response to bleomycin. Am. J. Pathol. 174, 21–33.

Chapter 4

Neofunctionalization of a novel lung gene paralog may facilitate respiration in lungless salamanders

Collaborators: Jorge A. Dorantes, James Hanken

4.1 Abstract

Most terrestrial vertebrates use lungs to breathe. The few exceptions include lungless amphibians (hundreds of species of salamanders, but only two other amphibians), which breathe entirely through the skin and lining of the mouth. How lungless amphibians are able to meet metabolic demands is a topic of considerable interest. Lunglessness places theoretical limits on thermal tolerance and body size, but lungless salamanders paradoxically live across diverse thermal environments and reach relatively large body sizes. Lungless salamanders

display an increase in vascularization of extrapulmonary tissues (the skin and buccal cavity) compared to lunged species, but the possibility remains that molecular and physiological adaptations also contribute to the greater extrapulmonary respiratory capacity of lungless salamanders. I identify an undescribed gene present in salamanders, which likely arose from gene duplication of the critical respiratory gene, surfactant protein C (SPC). This new paralog is expressed specifically in the lung in lunged salamander, which is similar to the SPC expression pattern. In a lungless species, the new paralog is expressed in the skin and buccal cavity, which are sites of respiratory activity in this species. These expression differences, combined with structural modeling predictions and examination of respiratory tissue ultrastructure suggest that the newly discovered paralog of SPC is involved in extrapulmonary respiration in lungless salamanders. Increased capability for cutaneous and buccopharyngeal respiration may allow lungless salamanders to inhabit diverse ecological niches and reach relatively large body sizes.

4.2 Introduction

Most amphibians inhabit two worlds and are challenged with the task of respiration both in the water and on the land. To further complicate the issue, amphibians perform gas exchange across four respiratory sites—the lungs, the gills, the integument (skin) and the buccopharyngeal mucosa (oral and pharyngeal cavities)—which are utilized to varying extents depending on the species and its life history stage. The integument of amphibians is a particularly important site of gas exchange (Gatz et al., 1974a; Whitford and Hutchison, 1963, 1965). For instance, in adult salamanders 50% or more of oxygen uptake is through

the skin (Whitford and Hutchison, 1965). Lability in sites of gas exchange is critical for amphibians, especially for metamorphic species, which face different respiratory demands in air and water. However, little is known about the molecular mechanisms that enable the respiratory transition from water breathing to air breathing. The mechanism of the transition to aerial respiration is even more enigmatic in animals that rely purely on non-pulmonary sites of respiration, such as the lungless salamanders.

In the field of amphibian respiratory physiology few species have elicited as much attention as those with reduced lungs (Hutchison et al., 1976) or those that are completely lungless such as the salamanders in the family Plethodontidae (Feder, 1983; Feder and Burggren, 1985; Gatz et al., 1974a; Whitford and Hutchison, 1965). The plethodontid salamanders are a monophyletic family, which account for over two thirds of all salamander species (Wake and Hanken, 1996). All species in this family lack lungs as adults. Respiration instead takes place over the integument (skin), buccopharyngeal mucosa (oral cavity), and (in some species or at some life history stages) through the gills. Lunglessness is not unique to plethodontids and has evolved several times in amphibians (Wake and Donnelly, 2010). The adaptive significance of lung loss is unknown, though hypotheses include the reduction of buoyancy (Bruce et al., 1994; Dunn, 1926; Wilder and Dunn, 1920), transition from aquatic to terrestrial mating (Reagan and Verrell, 1991), or a reduction in the width of the head, which may have rendered buccal pumping inefficient (Ruben and Boucot, 1989).

All amphibians employ the skin and buccopharyngeal membranes in respiration. However, it is apparent that plethodontids are adapted for highly efficient gas exchange at these sites. Early researchers initially proposed that the pharynx of plethodontids functions as a "pharyngo-esophageal lung" because of its high vascularity (Barrows, 1900; Czopek, 1961,

1962, 1965; Seelye, 1906; Wilder, 1901). It is likely that other physiological differences in addition to increased vascularization account for the highly efficient extrapulmonary (skin and mouth) respiration observed in plethodontids. In order to study the molecular mechanisms underlying the adaptation to lunglessness, I investigated the expression of crucial pulmonary surfactant-associated proteins in plethodontid salamanders.

Proper pulmonary function depends on the presence of pulmonary surfactant, a complex and evolutionarily variable mix of phospholipids and proteins that aid mucous spreading and lung compliance (Daniels et al., 1995; Orgeig et al., 2007; Pattle and Hopkinson, 1963; Whitsett and Weaver, 2002). Non-mammalian vertebrates are less reliant on pulmonary surfactant for lung compliance and instead are hypothesized to employ it to prevent collapsed lungs or swim bladders from self-adhering or to facilitate mucous clearance (Daniels et al., 1995, 1998). Phospholipids make up about 90% of pulmonary surfactant and phosphatidylcholine is the dominant lipid present, with about half of it present as dipalmitoylphosphatidylcholine (DPPC; Veldhuizen et al., 1998). Pulmonary surfactant proteins make up the other 10% of pulmonary surfactant.

Pulmonary surfactant-associated protein C (SPC) is a hydrophobic protein, which localizes to the air-liquid interface of the lung's lumen and helps to reduce surface tension by aiding the adsorption and distribution of lipids within pulmonary surfactant (Johansson, 1998; Possmayer et al., 2001; Weaver and Conkright, 2001; Whitsett and Weaver, 2002). SPC also helps regulate production and turnover of phosphatidylcholine (Rice et al., 1989; Weaver and Whitsett, 1991). One underappreciated role of SPC is to limit the depth of the liquid layer lining the lung mucous, known as the hypophase or alveolar surface liquid (Guyton et al., 1984). Thickness of a mucous layer greatly impacts gas exchange between the

environment and the blood supply (Ultsch and Gros, 1979). Additionally, oxygen uptake is positively influenced by the presence of surfactant in the hypophase, likely due to convective effects that facilitate mixing of oxygen and mucous (Sosnowski et al., 1998). SPC can decrease hypophase depth by regulating fluid balance and surface tension, thereby influencing oxygen uptake by reducing the distance between the environment and blood supply.

The expression pattern of SPC is highly conserved across tetrapods: all tetrapods express SPC specifically in the lungs (Bourbon and Chailley-Heu, 2001; Fisher et al., 1989; Glasser et al., 2000; Hyatt et al., 2007; Rankin et al., 2015; Weaver and Whitsett, 1991; Wert et al., 1993; Wohlford-Lenane et al., 1992; Yin et al., 2010). In mammals, SPC is confined to alveolar type II cells. In anamniotes, SPC is expressed throughout the lung (Hyatt et al., 2007; Rankin et al., 2012, 2015). Alveolar type I and type II cells are not readily distinguishable in anamniotes (Goniakowska-Witalińska, 1980) and it is possible that amphibians and fishes have a single type of pulmonary epithelial cell, which performs both secretory and gas exchange functions (see Chapter 6).

While looking for salamander orthologs of SPC, I discovered a previously undescribed gene with high sequence similarity to SPC. I identified three species of salamanders, one with lungs (*Ambystoma mexicanum*) and two lungless (*Desmognathus fuscus* and *Plethodon cinereus*), which express this undescribed gene in addition to SPC. This new gene, SPC-like, is likely a paralog of SPC.

Based on sequence conservation and phylogenetics, SPC duplication likely occurred a single time prior to about 150 Ma, predating the divergence of ambystomatids and plethodontids. SPC-like expression, as assayed by *in situ* hybridization, closely matches the expression pattern of SPC in the lunged species *A. mexicanum*: the gene is expressed specifically in the

lungs, but at a lower expression level than SPC. Unexpectedly, in the lungless salamander *Desmognathus fuscus*, *in situ* hybridization reveals SPC-like expression in the integument and the buccopharyngeal mucosa. The site of expression of SPC-like corresponds with life history stage: embryonically, SPC-like is found in the skin. Expression is maintained within the skin during larval stages. At metamorphosis, SPC-like expression shifts to the mouth and remains buccopharyngeal into adult life. I propose that SPC-like is neofunctionalized for extrapulmonary respiration in this species of lungless salamander. Neofunctionalization is the process by which one daughter gene of a gene duplication event evolves novel function while the other copy retains ancestral function (He and Zhang 2005). Neofunctionalization of SPC-like may enable efficient lungless respiration and may account for success of this physiological strategy across many aquatic and terrestrial habitats.

4.3 Materials and methods

4.3.1 Animal husbandry

Desmognathus fuscus (Northern dusky salamander) embryos were collected under Massachusetts DFW collection permits 080.11SCRA (2012), 027.13SCRA (2013), 083.14SCRA (2014), and 022.15SCRA (2015) and appropriate local permits from the following GPS coordinates: 42.483111, -72.761263 and 42.450922, -71.913009. Adults were collected from 42.671606, -71.776156. Embryos were maintained at 15°C until hatching in 0.1x Marc's Modified Ringer solution (MMR; 0.01M NaCl, 0.2mM KCl, 0.1mM $MgSO_4$, 0.2mM $CaCl_2$, 0.5mM HEPES pH 7.4). Following hatching, larvae were fed Artemia spp. and maintained

Table 4.1: Primers used to clone SPC and SPC-like.

Gene	Species	Forward	Reverse
SPC	*A. mexicanum*	5'-CAC ACA GAR AMG ATT TTC CAG ATG-3'	5'-CGT CTT GTC CAT TTT TGT KAB GTA GCA-3'
SPC-like	*A. mexicanum*	5'-AAG ATG GAA ACC GGC AGC AAG C-3'	5'-CGT CTT GTC CAT TTT TGT KAB GTA GCA-3'
SPC-like	*D. fuscus*	5'-AAG ATG GAA ACC GGC AGC AAG C-3'	5'-AGT ATT GGA AGC GGT CTG GGT G-3'
SPC-like	*P. cinereus*	5'-AAG ATG GAA ACC GGC AGC AAG C-3'	5'-GGT GTA GTC ATA GAC CAC-3'

at 17–20°C until metamorphosis at approximately 7 months post-hatching. Older larvae were hand-fed blood worms. Embryos and larvae were fixed at intermediate stages from embryogenesis until 3–5 days post-metamorphosis in MEMFA (0.1M MOPS pH 7.4, 2mM EGTA, 1mM $MgSO_4$, 3.7% formaldehyde) at 4°C overnight, then dehydrated and stored at −20°C in 100% methanol. Adults were fixed directly after collection in a similar manner to embryos.

Ambystoma mexicanum (Mexican axolotl) embryos were obtained from *Ambystoma* Genetic Stock Center, University of Kentucky, and maintained in 20% Holtfreter solution at 17°C. After hatching larvae were raised similarly to *Desmognathus fuscus* larvae. Fixation was performed as described above.

4.3.2 PCR

Embryonic cDNA from *Desmognathus fuscus*, *Plethodon cinereus*, and *Ambystoma mexicanum* was used to clone SPC-like. SPC was cloned from *A. mexicanum* and attempts were made to clone SPC from plethodontids. Degenerate and non-degenerate PCR primers were used (Table 4.1).

Primers were designed based on alignment of SPC sequences from *Desmognathus fuscus*, *Xenopus laevis, X. tropicalis, Anolis carolinensis, Neovison vison, Bos taurus, Monodelphis*

domesticus, and *Homo sapiens*. All sequences were obtained in GenBank, with the exception of the *D. fuscus* sequence, which was kindly provided by Dr. David Weisrock.

4.3.3 SPC phylogeny

SPC sequences were identified through TBLASTN and TBLASTX searches of sequenced transcriptomes from this study (*Plethodon cinereus* and *Ambystoma mexicanum*) as well as from *Rana (Lithobates) pipiens* (Christenson et al., 2014), *Hynobius chinensis* (Che et al., 2014) *Ambystoma tigrinum* (Doyle et al., 2013) and *Desmognathus fuscus* (David Weisrock, personal communication). For transcriptome assembly methods, see Chapter 3 and Appendix B. SPC sequences from annotated genomes were also taken from NCBI and ENSEMBL (Table 4.2). Outgroups were selected based on previous phylogenies of SPC (Hughes 2007, Hedlund et al. 2009). Predicted amino acid sequences were generated from all nucleotide sequences. Multiple sequence alignment was performed in using PRANK (Löytynoja and Goldman, 2010) and resulting nexus files were exported to Mesquite to check alignments. Subsequently, 95% maximum clade credibility gene trees were reconstructed in MrBayes (v3.2.5; Ronquist et al., 2012) using Markov chain Monte Carlo analysis with 500,000 generations sampling every 100 generations and a burn in of 25%. Convergence of the posterior probabilities was assessed by examining output statistics, including the potential scale reduction factor, which equaled or exceeded 1.000.

Multiple sequence alignment for the text figure was completed using the MSA package in R.

Table 4.2: Accession numbers and contig IDs for SPC, SPC-like and outgroups. Multiple contigs or genes occasionally match SPC

Species	Gene	Accession/Contig ID	Citation or source
Ambystoma mexicanum	SPC	c21543_g1_i2	Present study
Ambystoma mexicanum	SPC-like	c421069_g1_i1	Present study
Ambystoma mexicanum	SPC	Cloned from embryo	Present Study
Plethodon cinereus	SPC	c4587_g1_i1	Present Study
Plethodon cinereus	SPC-like	Cloned from embryo	Present Study
Desmognathus fuscus	SPC-like	isotig01894, isotig01892	Courtesy of David Weisrock
Ambystoma tigrinum	SPC	Contig5897.1, Contig19649.1	Doyle et al. 2013
Hynobius chinensis	SPC	GAQK01000789.1	Che et al. 2014
Xenopus laevis	SPC	NP_001090190.1	
Xenopus tropicalis	SPC	NP_001122132.1	
Rana pipiens	SPC	Transcript _409015_10368	Christenson et al. 2014
Anolis carolinensis	SPC	XP_003228279.1	
Neovison vison	SPC	CAA79577.1	
Bos taurus	SPC	AAI23537.1	
Homo sapiens	SPC	AAH05913.1	
Monodelphis domestica	SPC	XP_001373402.1	
Chrysemys picta belli	SPC	XP_005285292.1	
Chelonia mydas	SPC	XP_007066382.1	
Alligator sinensis	SPC	XP_006028986.1	
Python bivattatus	SPC	XP_007427166.1	
Latimeria chalumnae	Unchar. protein	ENSLACG 00000001041.1	LatCha1 Assembly GCA 000225785.1
Danio rerio	Predicted BRICHOS domain containing protein 5	XP_001332252.1	
Xenopus tropicalis	Lectin 1	NP_001090798.1	
Xenopus tropicalis	Tenomodulin	XP_004919251.1	

4.3.4 *In situ* hybridization

Embryos were fixed overnight in 4% paraformaldehyde (PFA) or MEMFA at 4°C, dehydrated and stored in 70% or 100% MeOH at −20°C. Wholemount mRNA *in situ* hybridization (ISH) was performed by rehydrating samples, then samples were treated with between 5 and 10 µg/ml proteinase K for 30 to 60 min., washed with PBTw (137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO4, 0.2% Tween-20), post-fixed in 4% PFA, washed with PBTw, and pre-hybridized in hybridization buffer for 2 hr at 65°C (hybridization buffer: 50% formamide, 5x SSC, 0.1 mg/ml heparin, 1x Denhardt's, 0.01% CHAPS, 0.2 mg/ml tRNA, 0.1% Tween-20; all solutions RNase-free). DIG-labeled riboprobes were diluted approximately 1:40 in hybridization buffer then denatured at 85°C for 10 min, before adding to specimens. Hybridization was carried out overnight at 65°C. Posthybridization washes were performed with a solution of 50% formamide, 5x SSC, and 0.2% Tween-20 at 65°C for 8 changes of 30 min each. Specimens were washed with maleic acid buffer plus 0.2% Tween-20 (MABT) prior to blocking and antibody incubation. Antibody block solution included 20% heat-inactivated sheep serum, 2% blocking reagent (Roche, Penzberg, Germany) in MABT. Samples were incubated with 1:2500 anti-DIG-AP Fab fragments (Roche) diluted in blocking solution overnight at 4°C. Extensive washes with MABT were performed prior to color development using BM-Purple (Roche) or NBT/BCIP (Sigma, St. Louis, MO). Color development occurred over several hours. Embryos were then embedded for cryosectioning at 14–16 µm thickness.

4.3.5 Structure models

An experimentally determined protein data bank (PDB) model for porcine SPC was downloaded from the Research Collaboratory for Structural Bioinformatics PDB (Johansson et al., 1994). The secondary and supersecondary structure of *Desmognathus fuscus* SPC-like was predicted using Quark *Ab initio* Protein Structure Prediction (Xu and Zhang, 2012). The N- and C-terminal extent of the *Desmognathus* SPC-like sequence was chosen based on alignment with mature forms of SPC found in mammals.

A structure model for *D. fuscus* SPC-like was also predicted in SWISS-MODEL using the Johansson et al. (1994) model as a template. The PDB files for each SPC-like model were imported into PyMol (Schrödinger) and aligned with the SPC model to graphically illustrate structural similarities.

4.3.6 Transmission electron microscopy

Two 24-mm *D. fuscus* larvae were euthanized and decapitated. Specimens were then dissected in fixative (2.5% glutaraldehyde, 2% paraformaldehyde in 0.1 M HEPES; the aldehydes were free of alcohol stabilizers). I divided the head into three 1-mm sagittal sections. In addition, I euthanized an 18-cm adult *Ambystoma mexicanum* and dissected out the lungs and a section of gular skin from the ventral chin. These samples were trimmed to 1 mm thick pieces in fixative.

The samples were left in fresh fixative for three days and then washed two times quickly with 0.1 M HEPES and three times for 5 min with milli-Q H_2O (mqH_2O). Next, samples were fixed for 24 hr at 4°C in aqueous 1% osmium tetroxide, followed by five washes in mqH_2O for 5

min each. Subsequently, specimens were stained with 2% uranyl acetate (EMS, Hatfield, PA) overnight at 4°C, then washed two times for 5 min with mqH_2O. Specimens were dehydrated with 5-min washes of 50%, 70%, and 95% ethanol, followed by three 10-min washes with 100% ethanol, then two quick rinses with propylene oxide (PO). Specimens were embedded in Embed 812 resin (EMS) formulated to medium hardness by rinsing 30 min each in 1:1 PO to Embed 812, 1:2 PO to Embed 812, then 60 min in 1:4 PO to Embed 812. Specimens were then transferred to 100% Embed 812 and incubated overnight at room temperature, followed by two subsequent changes of Embed 812, over a total embedding time of 48 hr. Samples were then positioned in molds and placed at 60°C for three days to polymerize.

Sectioning was performed on a Leica UCT ultramicrotome, using glass knives for trimming blocks and generating semi-thin (1 µm) sections and a DiATOME diamond knife for generating thin sections at approximately 60-100 nm thickness (target thickness = 80 nm). Sections were flattened with chloroform vapor, transferred onto precoated Formvar/carbon 200 mesh copper grids (#01803F, Ted Pella, Redding, CA), and dried on filter paper.

Grids were imaged with an FEI Tecnai G2 series F20 transmission electron microscope (Hillsboro, OR) at 80 kV using a Gatan CCD camera and Gatan Digital Micrograph Software (Pleasanton, CA).

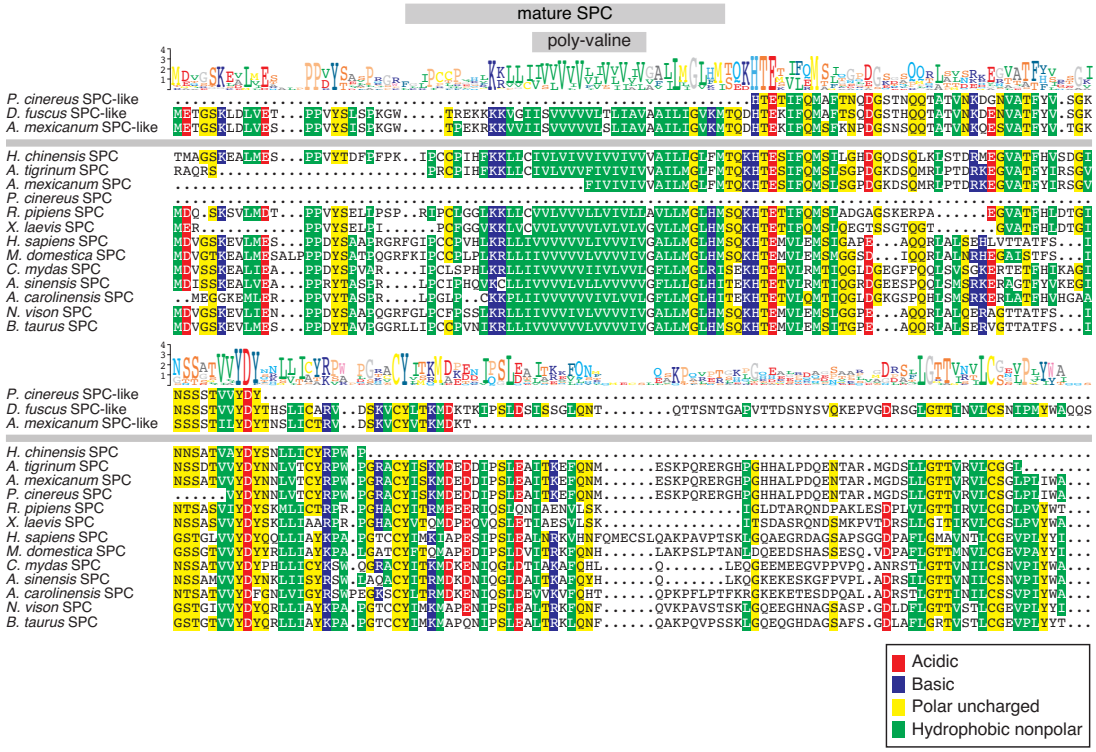


Figure 4.1: Amino acid alignment of SPC and SPC-like sequences. SPC and SPC-like sequences were aligned by MUSCLE. Hydropathy of residues is displayed using a color scheme according to the key. Sequence logo (top) indicates sequence conservation. The extent of the mature peptide sequence of human SPC and the poly-valine domain are marked above. SPC and SPC-like sequences are fairly conserved across the mature peptide region. The alignment and sequence logo were produced with the MSA package for R. Genera are as follows (from top to bottom): *Plethodon, Desmognathus, Ambystoma, Hynobius, Ambystoma, Plethodon, Rana, Xenopus, Homo, Monodelphis, Chelonia, Alligator, Anolis, Neovison,* and *Bos.*

4.4 Results

4.4.1 Discovery of a new pulmonary surfactant protein

Due to specificity of SPC expression to the lungs and its conserved role across tetrapods I decided to examine the expression of SPC in lungless salamanders compared to lunged salamanders. I attempted to clone SPC from cDNA synthesized from total RNA extracted from whole embryonic homogenates of *Ambystoma mexicanum*, *Plethodon cinereus* and *Desmog-*

nathus fuscus. I cloned a fragment from *Ambystoma mexicanum* and translated the cDNA into a predicted peptide sequence. The *A. mexicanum* cloned fragment matched SPC by an NCBI TBLASTN search of the nucleotide collection database (NR). All 100 hits were surfactant protein C and the top hit was *Chrysemys picta belli* SPC (44% identity, 67% positives, e-value = 2e-36). In addition I cloned a second sequence from *A. mexicanum*, which was divergent in nucleotide sequence from the initial *A. mexicanum* SPC sequence, yet still only matched SPC within the NCBI NR database (top match: *Anolis carolinensis*; 41% identity, 53% positives, e-value = 3e-13). I term this form "SPC-like" while the form with higher sequence identity to amniote and frog SPC is termed "SPC." The presence of two SPC sequences indicates that there are two forms of SPC expressed in *A. mexicanum*. SPC is alternatively spliced in mammals (Connelly and Possmayer, 1992; Glasser et al., 1988). However, the two forms of SPC found in *A. mexicanum* do not correspond to splice variants as their sequences markedly diverge within exonic regions, but not according to putative splice boundaries (Fig. 4.1).

In attempting to clone SPC from the lungless salamanders *Desmognathus fuscus* and *Plethodon cinereus* I amplified sequences where SPC was the only candidate match in the hit table using NCBI TBLASTN against the NR database. However, in terms of sequence similarity, the *P. cinereus* and *D. fuscus* sequences matched *A. mexicanum* SPC-like sequence more than the *A. mexicanum* SPC sequence. For instance, the *D. fuscus* cloned sequence was 75% identical to SPC-like of *A. mexicanum* (86% positive, e-value= 4e-73) while only matching the *A. mexicanum* SPC sequence with 43% identity (60% positives, e-value=1e-26). Using transcriptome sequencing, I identified a SPC sequence from *P. cinereus* with high sequence identity to SPC in *A. mexicanum* and amniotes (Fig. 4.1).

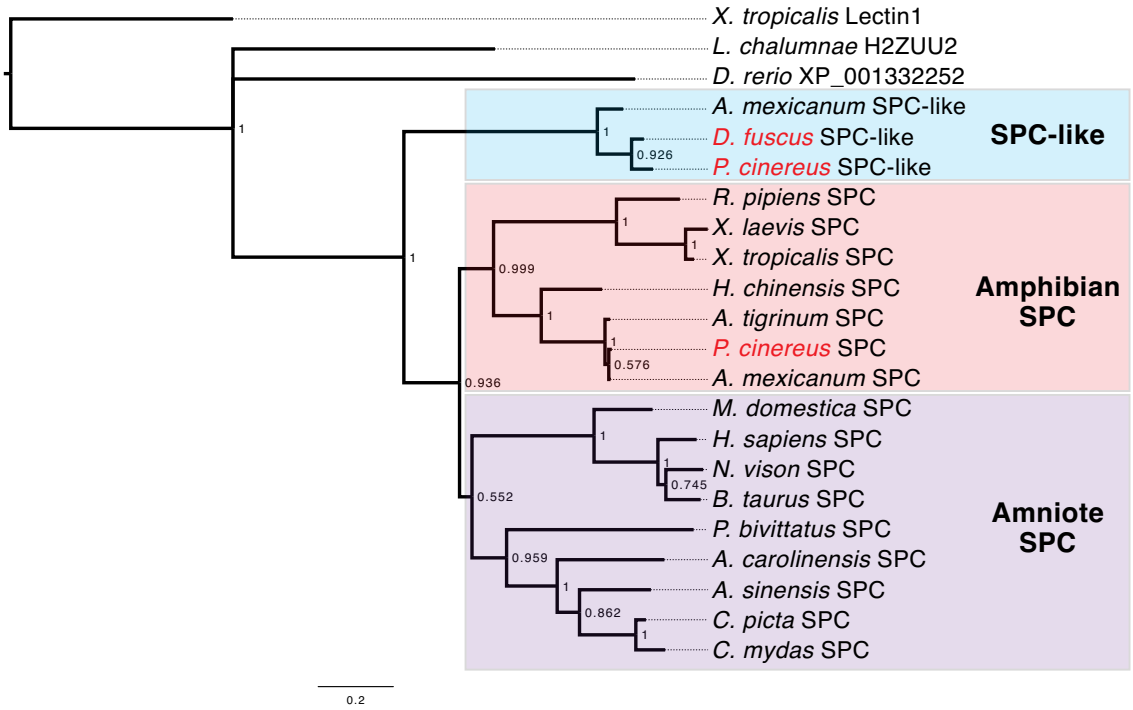


Figure 4.2: Bayesian phylogeny of SPC sequences reveals the presence of SPC-like transcripts in salamanders. The tree represents a 95% maximum clade credibility phylogeny based on predicted amino acid sequences of SPC and related nucleotide sequences. Full species names and accession numbers are listed in Table 4.2. Node values are posterior probabilities. Scale indicates 0.2 expected changes per site. Red taxa are lungless (plethodontid) salamanders.

Additional sequences of salamander SPC were obtained from sequenced transcriptomes (Table 4.2) and SPC sequences were aligned using PRANK (Löytynoja and Goldman, 2010), also incorporating uncharacterized proteins from *Danio rerio* and *Latimeria chalumnae* that were top matches to SPC from BLAST searches against their respective genomes. In addition, I aligned BRICHOS domain-containing proteins from the chondromodulin family, of which SPC is a member, to the SPC and SPC-like sequences from NCBI, cloning, and transcriptomes. The Bayesian tree topology supports a relationship of SPC-like sister to all other tetrapod SPC sequences, with coelacanth and zebrafish BRICHOS domain proteins sister to the SPC + SPC-like group (Fig. 4.2). Because both SPC and SPC-like are present in

at least two salamander species (*A. mexicanum* and *P. cinereus*), this tree topology implies a duplication of an ancestral BRICHOS domain protein to form paralogs SPC and SPC-like on the tetrapod lineage after the last common ancestor of coelacanth and tetrapods. However, this interpretation of the phylogeny requires multiple losses of SPC-like in frogs, amniotes and possibly some salamanders. A more parsimonious explanation is that a SPC duplication event within salamanders resulted in the formation of two paralogs. One paralog, "SPC," maintained sequence conservation to other tetrapod SPC sequences. The other paralog, "SPC-like," diverged in sequence due to selection or drift. High sequence divergence likely creates a phylogenetic artifact due to low taxon sampling within amphibians resulting in SPC-like falling out sister to other tetrapod SPC sequences. With limited available amphibian SPC and SPC-like sequences and no genomes available for salamanders it is difficult to rule out a tetrapod duplication of SPC followed by multiple losses. However, the most parsimonious explanation for the observed sequence divergence is gene duplication within salamanders prior to the last common ancestor of ambystomatids and plethodontids. In either case, SPC-like represents an undescribed paralog closely related to highly conserved lung-specific SPC.

4.4.2 Structure models of SPC-like

SPC is proteolytically processed to form an approximately 35 amino acid-long mature peptide, which takes on an alpha-helical conformation in mammals (Johansson, 1998; Johansson et al., 1994). Mature SPC contains a highly conserved poly-valine region, which contributes to its extremely non-polar nature and hydrophobicity (Fig. 4.1). SPC forms an alpha helix,

which encompasses residues 9–34 of the mature peptide. Two models were generated for *Desmognathus fuscus* SPC-like secondary structure. SWISS-MODEL (Arnold et al., 2006) prediction using resolved SPC as a template (Johansson et al., 1994) results in near perfect alignment of SPC and SPC-like (Fig. 4.3A,D,E). However, the quality measure for this structure is low. *Ab initio* prediction, using Quark (Xu and Zhang, 2012), yields a different structure of *D. fuscus* SPC-like: the predicted structure includes two alpha helices separated by a three-residue loop, representing a possible helix hairpin. The smaller, N-terminal alpha helix spans residues 7–12, while the larger predicted alpha helix spans residues 16–34 (Fig. 4.3B,C). SPC-like does not conserve the cysteine residues found at positions 5 and 6 in the mature form. These residues are palmitoylated in mammals leading to hydrophobicity of the N-terminal (non-helical) portion of the mature peptide (Curstedt et al., 1990; Weaver and Whitsett, 1991). Structure models for amphibian and human SPC reconstructed using Quark *Ab initio* structure prediction (Xu and Zhang, 2012) matched the experimentally validated SPC structure from Johansson et al. (1994) (data not shown), giving confidence that Quark predictions are capable of accurately predicting secondary structure for this protein.

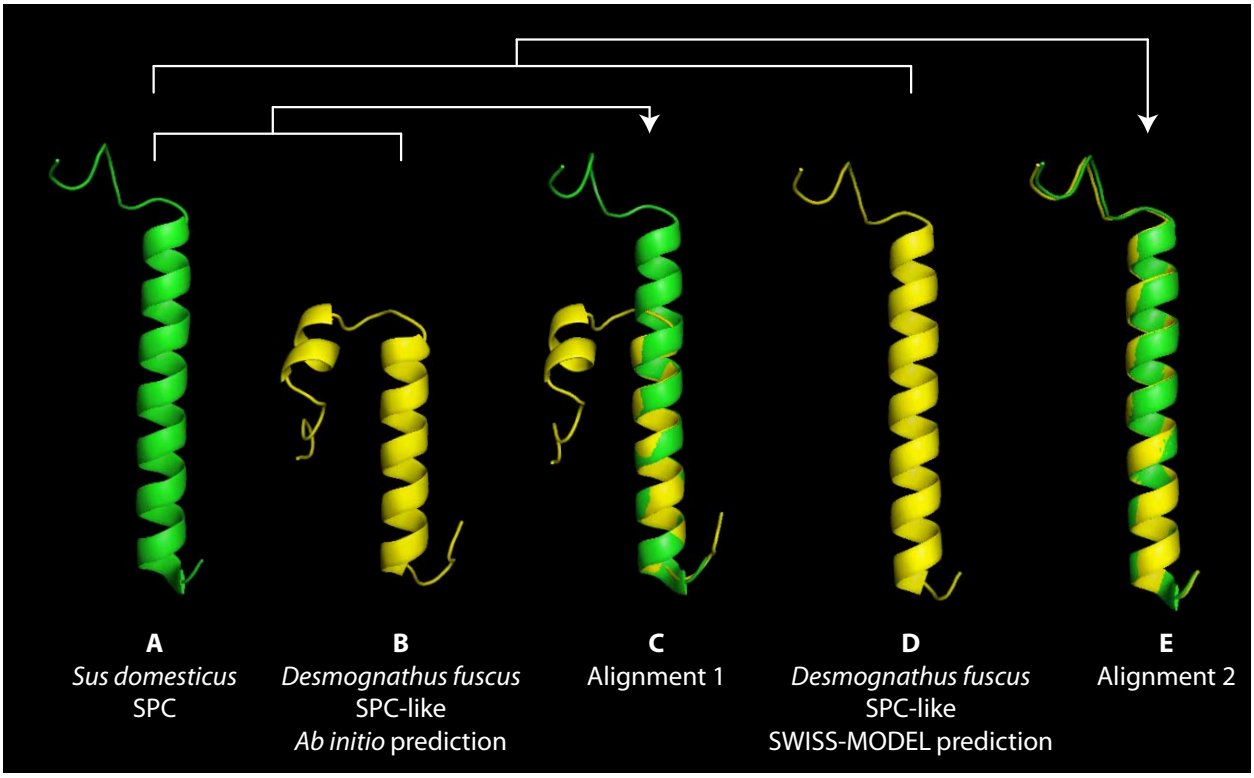


Figure 4.3: SPC and SPC-like secondary structure. SPC secondary structure is compared to the predicted secondary structure of SPC-like. (A) SPC secondary structure prediction for *Sus domesticus* (porcine) SPC was experimentally obtained by Johansson et al. (1994). *Desmognathus fuscus* SPC-like secondary structure was predicted in two ways: (B) *Ab initio* structure prediction (Xu and Zheng 2012) and (D) by SWISS-MODEL prediction using (A) as a template (Arnold et al. 2006). Resulting PDB files were visualized and aligned in PyMol using the alignment function (C, E).

4.4.3 Expression pattern of SPC in the lunged salamander *Ambystoma mexicanum*

I cloned a fragment of SPC from *Ambystoma mexicanum* and examined its expression by wholemount mRNA *in situ* hybridization. SPC is expressed specifically in the lungs in *A. mexicanum*. Detectable expression of SPC begins at stage 38, just following the formation of the laryngotracheal groove (Fig. 4.4A). Expression strengthens and by stage 40 SPC is present throughout the lungs and trachea. This expression pattern is maintained throughout embryonic, juvenile and adult life in *A. mexicanum* (Fig. 4.4B-E).

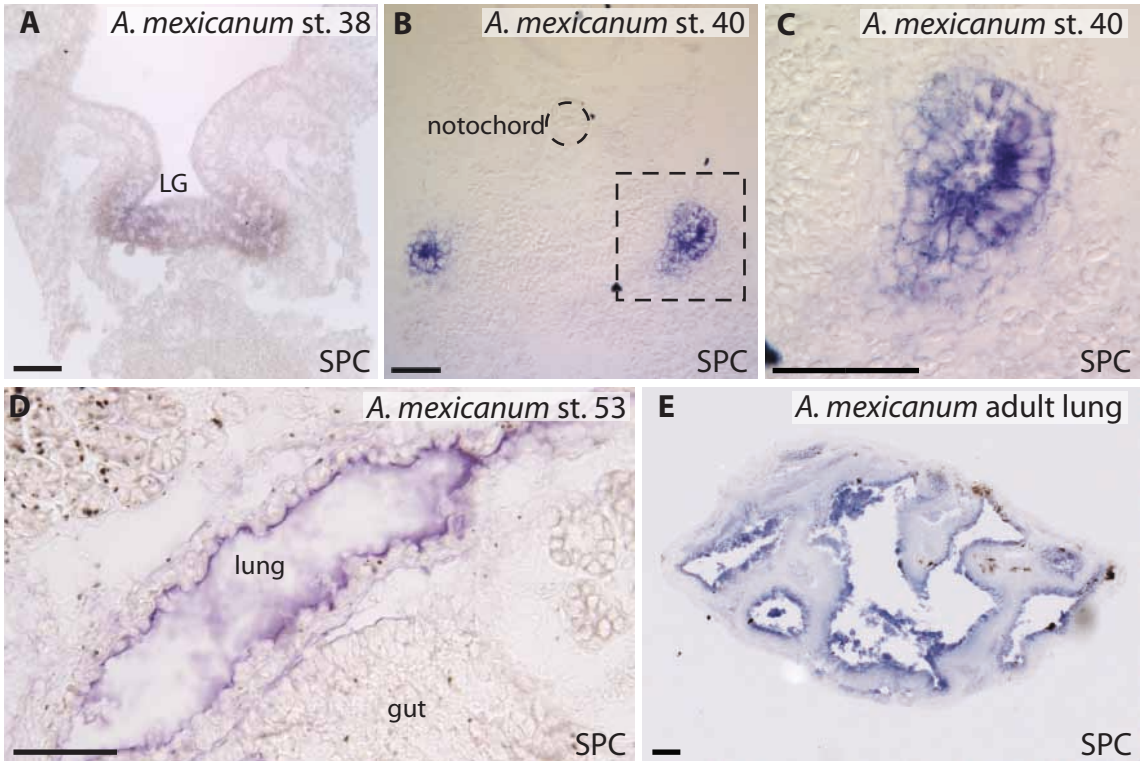


Figure 4.4: SPC expression in lunged *Ambystoma mexicanum*. Antisense wholemount *in situ* hybridization for SPC in the lunged salamander, *A. mexicanum*. (A-C) Transverse sections through embryonic *A. mexicanum* reveal SPC expression in the laryngotracheal groove (LG; A) and lung buds (B, C). (D) Sagittal section through a post-hatching *A. mexicanum* shows SPC expression within the lung. Anterior is to the left. (E) Transverse section through an adult lung with expression of SPC along the lumen. Scale bars: 100 µm.

4.4.4 Expression pattern of SPC-like in the lunged salamander *Ambystoma mexicanum*

The SPC paralog SPC-like is expressed in the lungs of *A. mexicanum*. However, expression of SPC-like is temporally and quantitatively different from SPC expression in this species. SPC-like expression could not be detected by transcriptome sequencing of the lung rudiment at stage 40. Expression in the lungs was not detectable by ISH at embryonic stages. During post-embryonic stages, SPC-like is expressed specifically in the lung (Fig. 4.5). SPC-like and SPC expression levels were quantified in the lungs of a single post-embryonic (stage 53) axolotl larvae by transcriptome sequencing and SPC expression is approximately 87 times higher than SPC-like (SPC-like = 17.8 FPKM; SPC = 1541.1 FPKM). SPC and SPC-like *in situ* hybridization at stage 53 developed in parallel revealed lower color development in SPC-like relative to SPC, supporting a lower expression level of this transcript (Figs. 4.4D, 4.5C).

4.4.5 Expression of SPC in the lungless salamander *Plethodon cinereus*

Plethodon cinereus expresses SPC within the lung rudiment, as assayed by transcriptome sequencing (Chapter 3). However, SPC expression was only identified in one individual at stage 19 and expression level was extremely low, at 0.44 fragments per kilobase of transcript per million mapped reads (FPKM).

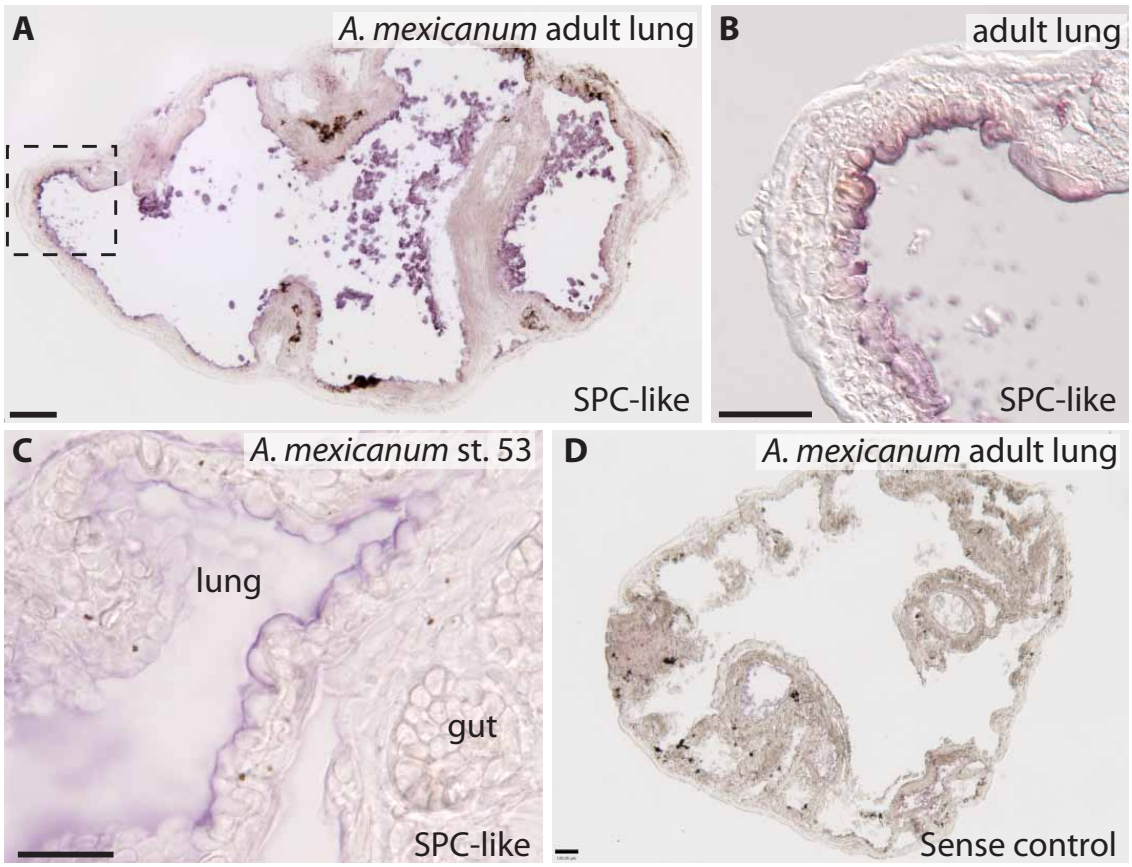


Figure 4.5: SPC-like expression in *Ambystoma mexicanum*. Wholemount *in situ* hybridization for SPC-like in the lunged salamander *A. mexicanum*. (A, B, D) Transverse sections of adult lungs stained with an SPC-like antisense probe (A, B) or sense control (D). (C) Sagittal section through a post-hatching *A. mexicanum* illustrates expression of SPC-like in the lung. Anterior is to the left. Scale bars: 100 µm.

4.4.6 Expression pattern of SPC-like in the lungless salamander *Desmognathus fuscus*

The *Desmognathus fuscus* ortholog of SPC-like displays a dynamic expression pattern over ontogeny. Embryonically, SPC-like is strongly expressed throughout much of the integument, with reduced staining on the dorsal (internal) surface of the operculum (gill covering) and in the limbs (Fig. 4.6). By raising a developmental series of *D. fuscus* I captured the expression pattern of SPC-like over larval development (approximately 7 months) into

adulthood. Integumentary expression of SPC-like is maintained in young larval stages and closely matches the expression pattern of embryos (Fig. 4.7). Expression is not observed on the gills. Immediately prior to metamorphosis, the expression of SPC-like begins to change: expression in the integument begins to diminish, while expression expands to the buccopharyngeal mucosa (oral epithelium) (Fig. 4.8). At this late pre-metamorphic stage, SPC-like expression in the integument is patchy. Remaining SPC-like-expressing cells in the integument appear to have been pushed apically, taking on an irregular morphology (Fig. 4.8B,D). Immediately following metamorphosis, SPC-like expression is completely absent from the integument and present only in the buccopharyngeal mucosa (Fig. 4.9). In adults, SPC-like is expressed in the buccal cavity and extends into the pharynx (Fig. 4.10). Adult expression is tightly confined to oral mucosa, with a strict boundary of expression at the marginal teeth. Expression is not observed along the vomers (near the internal nares) or in the center of the tongue. However, expression is strong at the margin of the tongue (Fig. 4.10).

SPC-like is also expressed in *Plethodon cinereus* as evidenced by tissue-specific RT-PCR (Fig. 4.11) and cloning. SPC-like is present in embryonic homogenates and appears to be expressed in the adult pharynx and foregut (Fig. 4.11E). In situ hybridization revealed possible expression within the embryonic pharynx and foregut (Fig. 4.11A,B) and possible adult expression in the submucosal layer of the pharynx (Fig. 4.11C). However, staining was weak and inconsistent over multiple trials, indicating that expression is either low, or that the *in situ* probe does not work adequately. Therefore, these expression patterns should be treated as tentative.

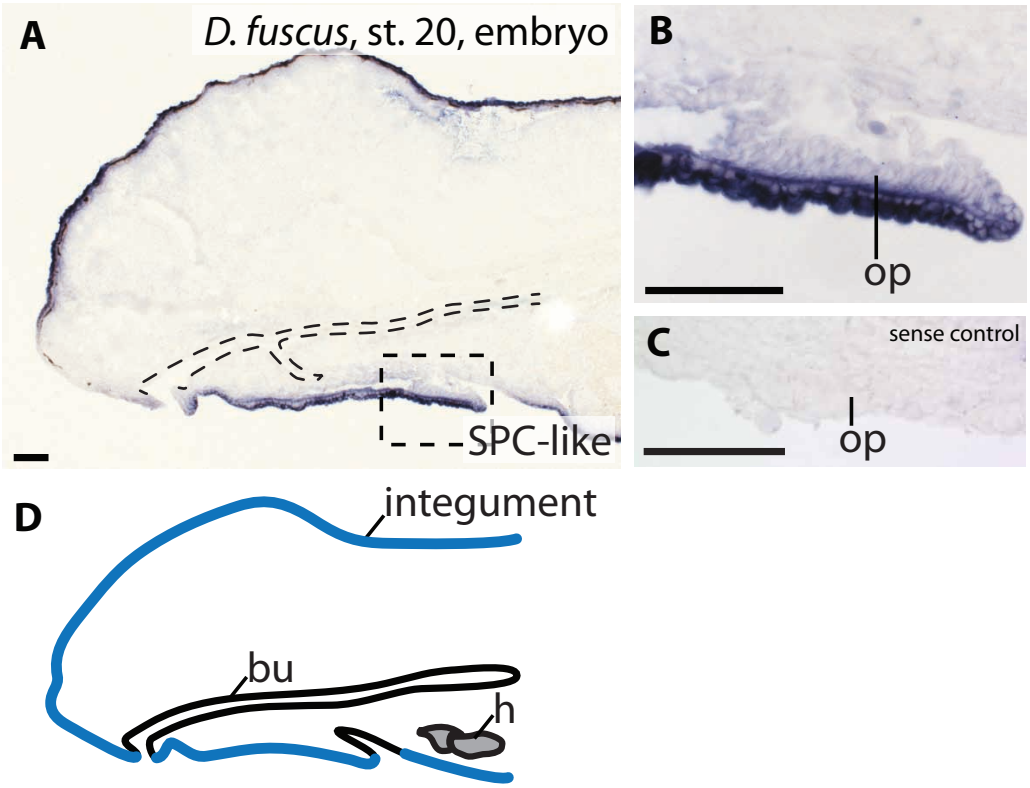


Figure 4.6: SPC-like expression in embryonic *Desmognathus fuscus*. Wholemount *in situ* hybridization for SPC-like in an embryonic lungless *D. fuscus* specimen. (A) SPC-like is expressed throughout much of the integument. Sagittal section; anterior is to the left. The buccopharyngeal mucosa is outlined with a dashed line. (B) Enlargement of boxed region in (A), corresponding to the opercular region (op), showing strong expression ventrally, but no expression dorsally. (C) Sense control of the opercular region displays no staining for SPC-like. (D) Expression summary. Blue lines indicate the expression pattern of SPC-like in a schematic sagittal section. The integument, heart (h) and buccopharyngeal mucosa (bu) are labeled. Scale bars: 100 µm.

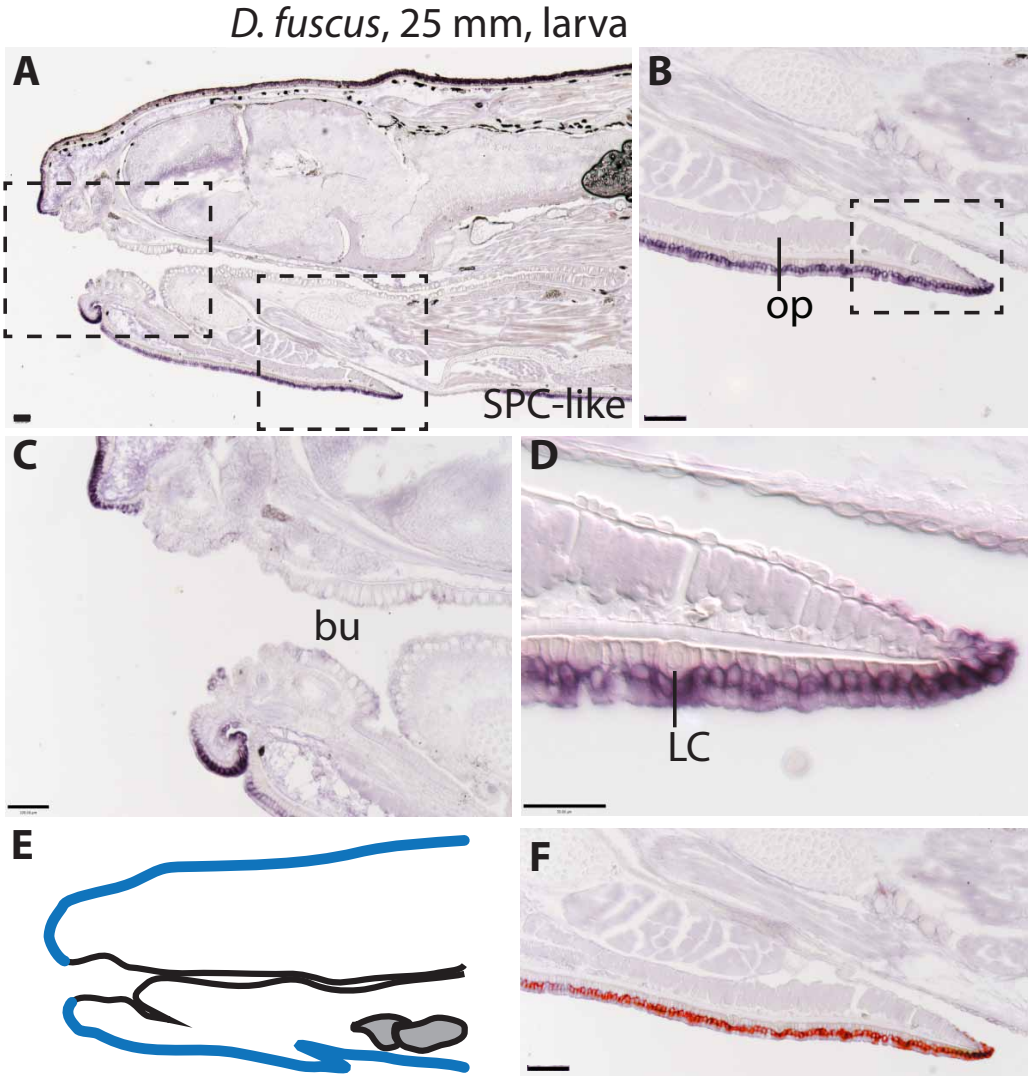


Figure 4.7: SPC-like expression early larval *Desmognathus fuscus*. (A-D) Integumental expression of SPC-like in the lungless species *D. fuscus* during an early larval stage (25 mm total length). Sagittal section; anterior is to the left. (B,C) Are enlargements of the boxed regions in (A). (B) Prominent expression in the skin covering the operculum (op) is maintained. (C) There is no expression inside of the buccopharyngeal cavity (bu). (D) Enlargement of the boxed region in (B), showing SPC-like expression in the superficial layer of the epidermis, and not in deep cells such as Leydig cells (LC; see Fig. 4.12). (E) Expression summary. Blue lines indicate expression pattern of SPC-like in a schematic sagittal section. (F) Replace color manipulation of (B) to highlight expression pattern in red. Scale bars: 100 µm, except (D): 50 µm.

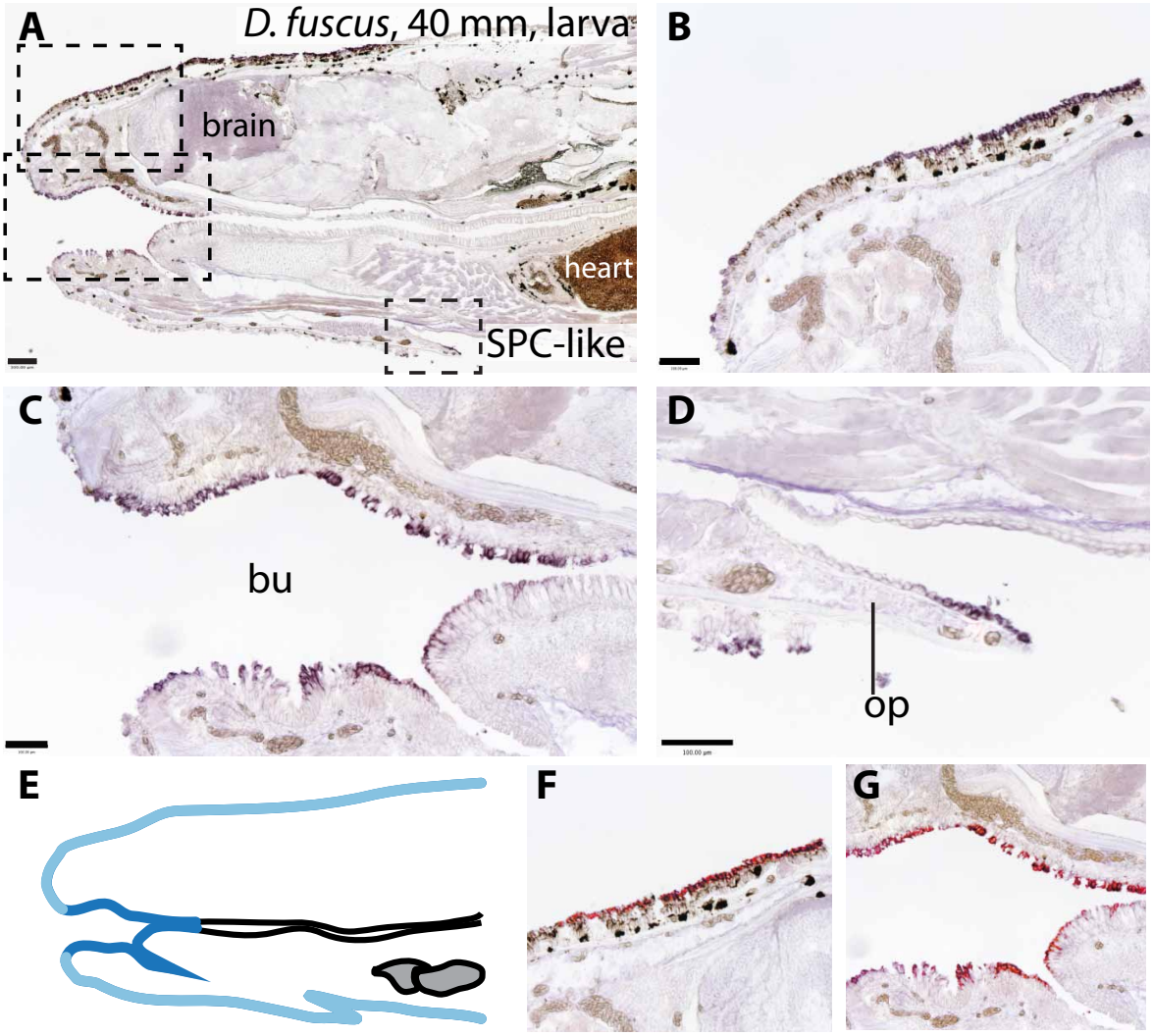


Figure 4.8: SPC-like expression in late larval *Desmognathus fuscus*. (A-D) Expression of SPC-like in a pre-metamorphic *D. fuscus* at 40 mm total length. Sagittal section; anterior is to the left. Boxed regions in (A) are enlarged in (B-D). (B, D) Expression of SPC-like in the integument has begun to diminish and become patchy. (C) Expression has shifted to the buccopharyngeal cavity (bu). (E) Expression summary. Blue lines indicate expression pattern and expression level in a schematic sagittal section. Light blue indicates lower expression than dark blue. (F,G) Replace color manipulation of panels (B,C) to illustrate expression pattern in red. Scale bars: 100 µm.

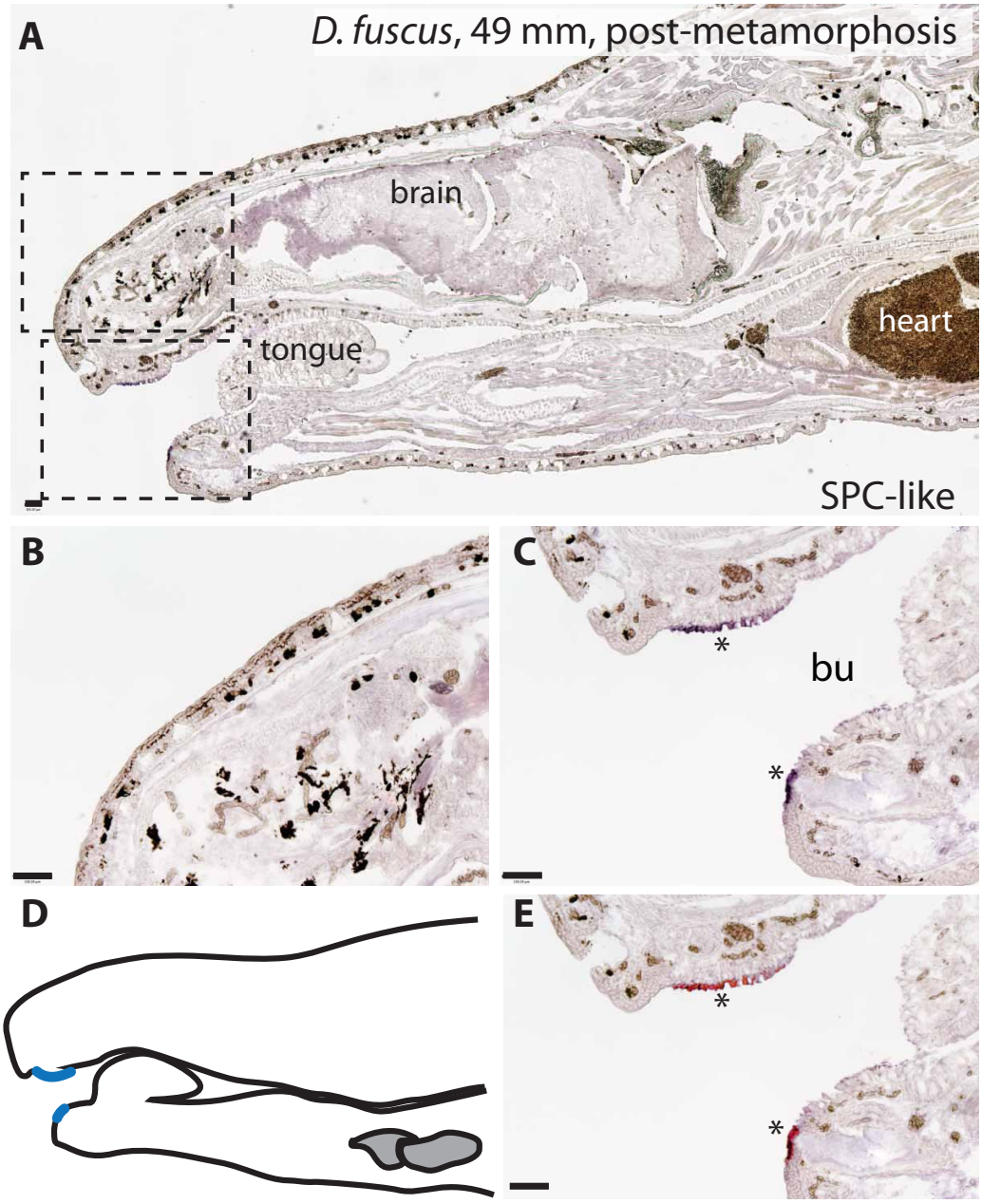


Figure 4.9: SPC-like expression in a recently metamorphosed *Desmognathus fuscus*. (A-C) Expression of SPC-like in a post-metamorphic *D. fuscus* at 49 mm total length. Sagittal section; anterior is to the left. Boxed regions in (A) are enlarged in (B,C). (B) Expression in the integument is absent. (C) Expression remains in a small portion of the buccopharyngeal cavity (bu), indicated by the asterisks. (D) Expression summary. Blue lines indicate expression pattern in this schematic sagittal section. (F) Replace color manipulation of (C) to illustrate expression pattern in red. Scale bars: 100 µm.

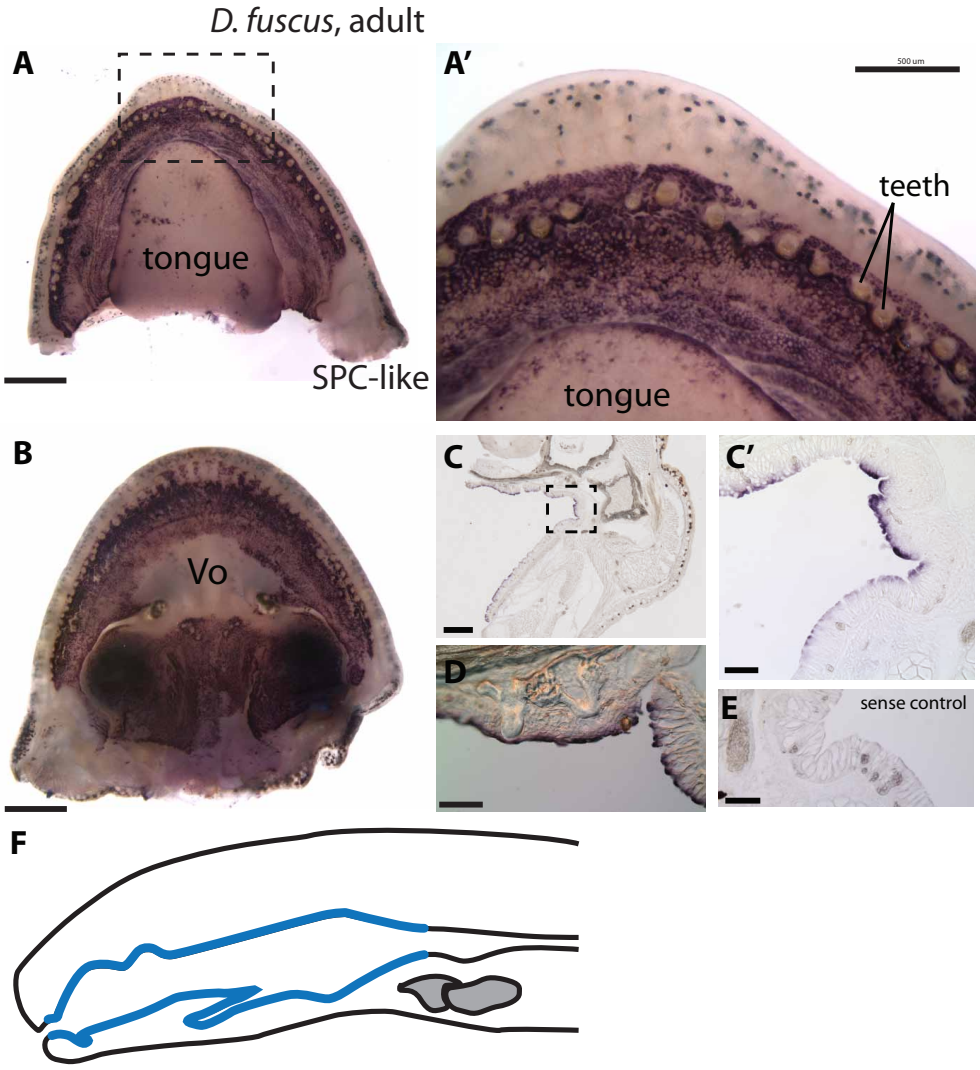


Figure 4.10: SPC-like expression in adult *Desmognathus fuscus*. Expression of SPC-like is assessed by wholemount *in situ* hybridization of the buccal cavity (A, B) and pharynx (C-E) of an adult *D. fuscus*. (A) The lower jaw and tongue seen in dorsal view. (A') Enlargement of (A) showing the extent of SPC-like expression, which does not extend exterior to the oral mucosa. Expression is also present along the margin of the tongue. (B) The dorsal buccal cavity, with SPC-like expression present over much of the surface, with the exception of the epithelium corresponding to the approximate extent of the vomers (Vo). (C, C', D) Transverse sections through the pharynx showing SPC-like expression along the pharyngeal mucosa. (C') Enlargement of boxed region in (C). (D) Differential interference contrast micrograph illustrating cellular morphology along the membrane. The pharynx was bisected mid-sagittally and the opposite half was hybridized in parallel with a sense control probe (E). (F) Expression summary. Blue lines indicate expression pattern in this schematic sagittal section. Scale bars: (A, B) 1 mm, (A', C) 500 µm, (C', D, E) 100 µm.

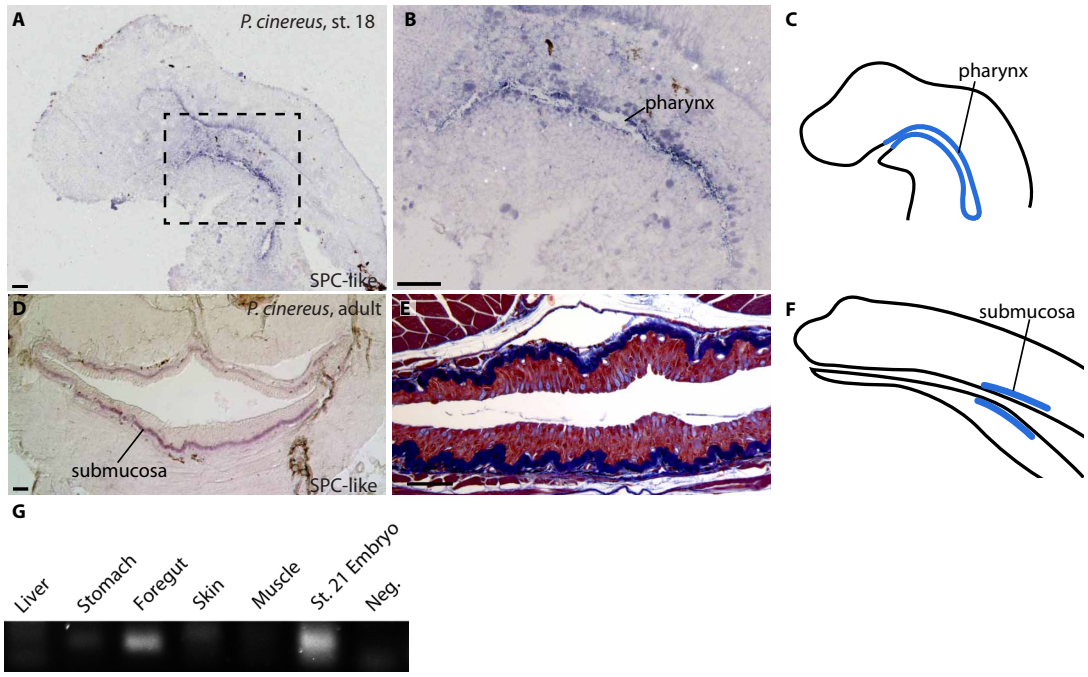


Figure 4.11: SPC-like expression in *Plethodon cinereus*. (A,B) Embryonic expression pattern of SPC-like. Sagittal section; anterior is to the left. Boxed region in (A) is enlarged in (B). (C) Schematic embryonic expression pattern in a sagittal section. Blue lines indicate expression of SPC-like. (D) Transverse section through the adult pharynx reveals potential SPC-like expression in the submucosal layer. (E) Transverse histological section of adult *P. cinereus* shows pharyngeal tissue layers. (F) Schematic adult expression pattern in a sagittal section. (G) RT-PCR on mRNA isolated from adult tissues, as well as whole embryonic mRNA reveals expression of SPC-like in the pharynx and foregut (labeled "foregut") of adults, and generally in embryos. Scale bars: 100 µm.

4.4.7 Ontogenetic changes to the integument in *Desmognathus fuscus*

Changes to the integument during lungless salamander metamorphosis have been documented extensively (Seelye, 1906; Wilder, 1925). The cessation of integumentary expression of SPC-like corresponds with several described metamorphic transitions including the degeneration of putative larval secretory cells (termed "Leydig cells," but not to be confused with testicular Leydig cells) and the process of molting (Wilder, 1925). Though not discussed by Wilder (1925), anatomical changes to the oral epithelium of *D. fuscus* at metamorphosis parallel changes to the integument such as increased stratification and potential molting or sloughing of tissue (Fig. 4.8). Histological sectioning confirms profound remodeling of the integument during metamorphosis. Prior to metamorphosis, the integument is composed of four layers: the deep basal lamina, a basal cell layer, an inner cell layer, and a keratinized superficial layer (Fig. 4.12A,B). The integument thickens and becomes more complex at metamorphosis, including the formation of large acinous glands (Figs. 4.8B, 4.12C).

4.4.8 Pronounced secretory activity from the larval skin of *Desmognathus fuscus*

I characterized the ultrastructure of the larval *D. fuscus* integument using transmission electron microscopy (TEM). Moving from deep to superficial, the integument is composed of a basal lamina (BL), containing collagen, which is linked to the basement membrane (Fig. 4.13B,C). The deep layer mostly contains cells with a fairly homogenous cytoplasm (Fig. 4.13B). Leydig cells are dispersed among these basally located cells (Figs. 4.13D, 4.14A).

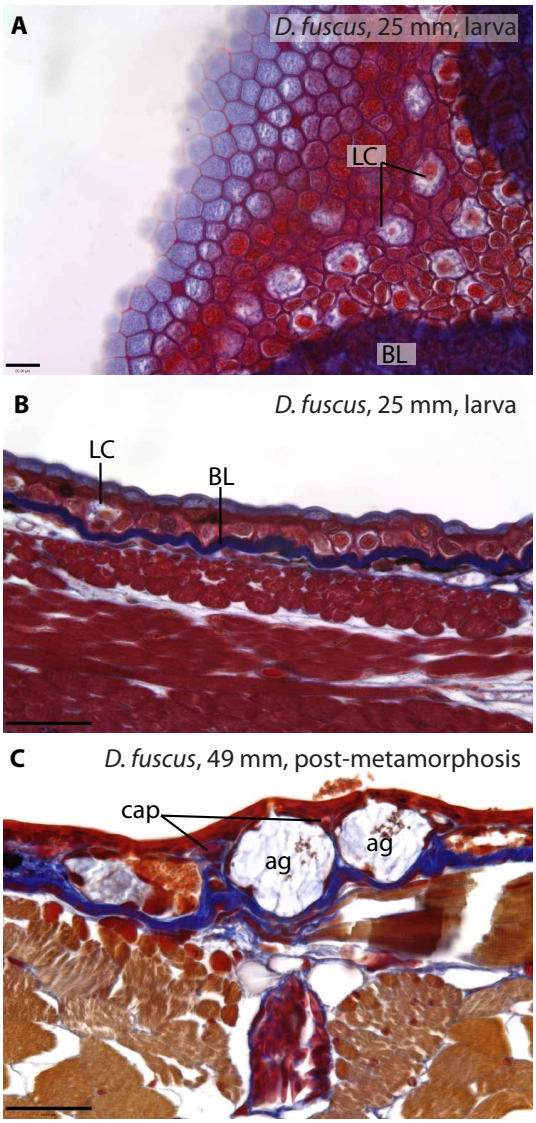


Figure 4.12: Integument histology in *Desmognathus fuscus* before and after metamorphosis. (A) Tangential section through the gular region of the larval integument shows the various layers of the skin from left to right: flattened cuticle-like keratinized layer, an inner cell layer, large cuboidal Leydig cells (LC), intermingled with capillaries and other supporting cells, purple: basal lamina (BL). (B) Sagittal section taken from the abdominal region of a larva. (C) Transverse section through the integument of a recently metamorphosed specimen showing the acinous glands (ag) and capillaries (cap). Scale bars: (A) 20 µm, (B,C) 50 µm.

In the central layer of the integument are cells with pronounced secretory activities, which are heavily vacuolated at their apical extent. The most superficial layer of the integument

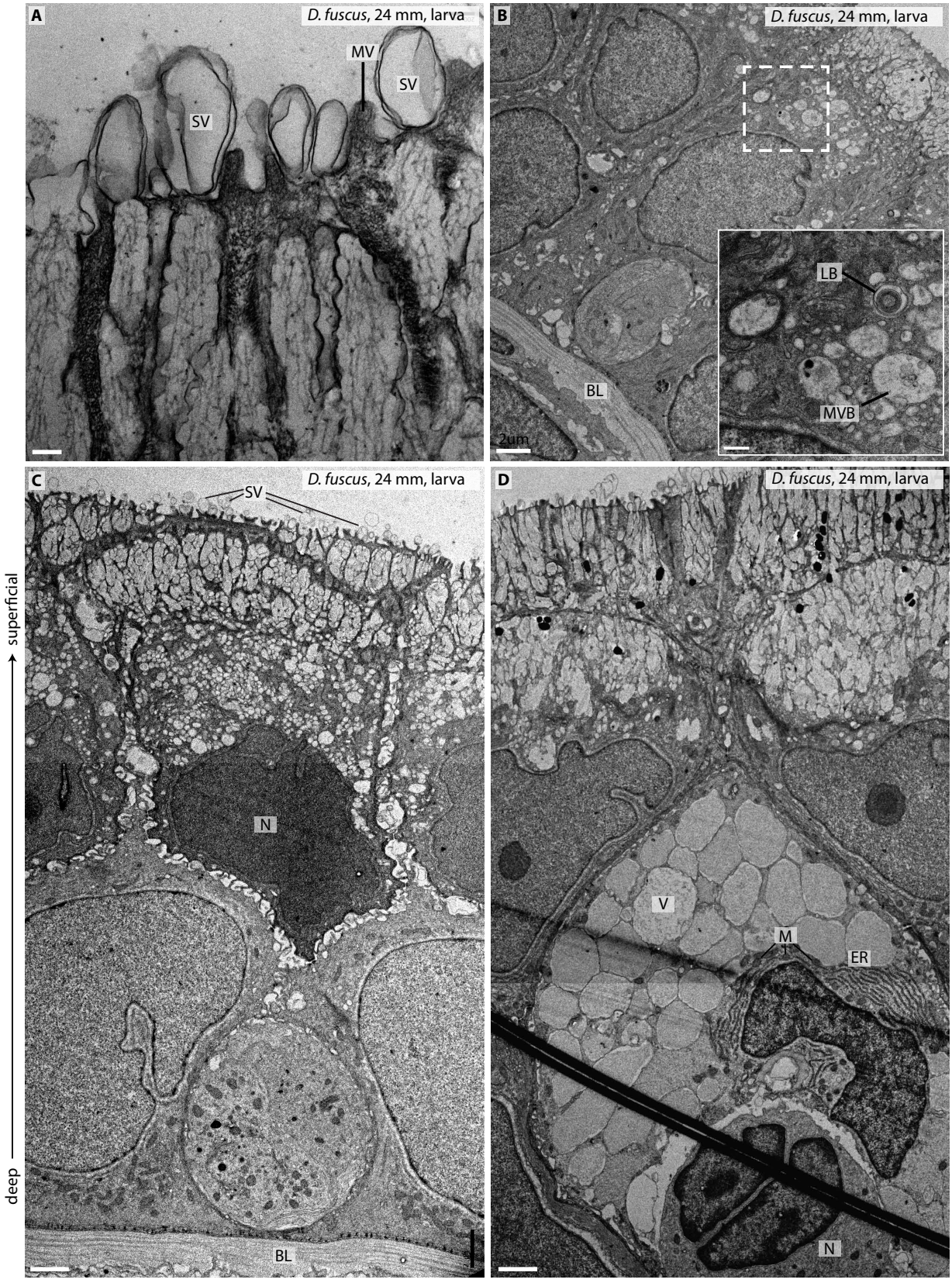
appears distinct from the inner cells due to a dark and consistent division between this layer and the underlying cells, but no clear membrane separates the two (Fig. 4.13C,D). This anucleated superficial layer likely represents the stratum corneum, a protective layer composed of keratinized cells. The stratum corneum displays an extremely high level of secretory activity evidenced by the near universal distribution of bilamellar secretory vesicles along its superficial surface (Fig. 4.13A,C). These secretory vesicles are similar in appearance to those secreted from the alveolar epithelial cells of the lung in *A. mexicanum* (Fig. 4.15F).

In the lung, pulmonary surfactant forms characteristic lamellar bodies (Schmitz and Müller, 1991). Preliminary characterization of the integument revealed several lamellar body-like structures (Figs. 4.13B, 4.14). Lamellar bodies in the skin suggest that a surfactant-like substance is being produced. Further work needs to be done to determine the distribution of lamellar bodies within the larval *D. fuscus* integument.

The alveolar surface of the *A. mexicanum* lung also shows indications of high secretory activity, including the secretion of lamellar body contents (Fig. 4.15E) and more simple bilamellar secretory vesicles, as discussed above (Fig. 4.15F). The integument of *A. mexicanum* does not appear to have high secretory activity. Several multivesicular bodies were identified on the external surface of the skin and internally in extracellular spaces. However, no secretory vesicles or active secretion was observed (Fig. 4.16).

Figure 4.13: Pronounced secretory activity in larval *Desmognathus fuscus* integument. (A-C) Transmission electron microscopic examination of the cranial integument of a 24-mm *D. fuscus* larva. Thin ($\approx$ 90 nm) sagittal sections through the epidermis, with the superficial (apical) surface pointing upwards, or to the upper right (B). (A) The apical surface of the integument is covered with secretory vesicles (SV), which emerge from columnar vacuolar structures. The apical surface is also interspersed with microvilli (MV). (B) The layers of the integument. The inset figure is an enlargement of the boxed region in (B), showing a potential lamellar body (LB) as well as a multivesicular body (MVB). (C) A low magnification view of the integument, showing discernable layers in the superficial cells. (D) A Leydig cell is filled with vacuoles mostly containing a floccular material. Abbreviations: BL, basal lamina; ER, endoplasmic reticulum; LB, lamellar body; M, mitochondria; MVB, multivesicular body; MV, microvilli; N, nucleus; SV, secretory vesicle; V, vacuole. Scale bars (A) 200 nm, (B,C,D) 2 µm, (B inset) 500 nm.

Figure 4.13, continued



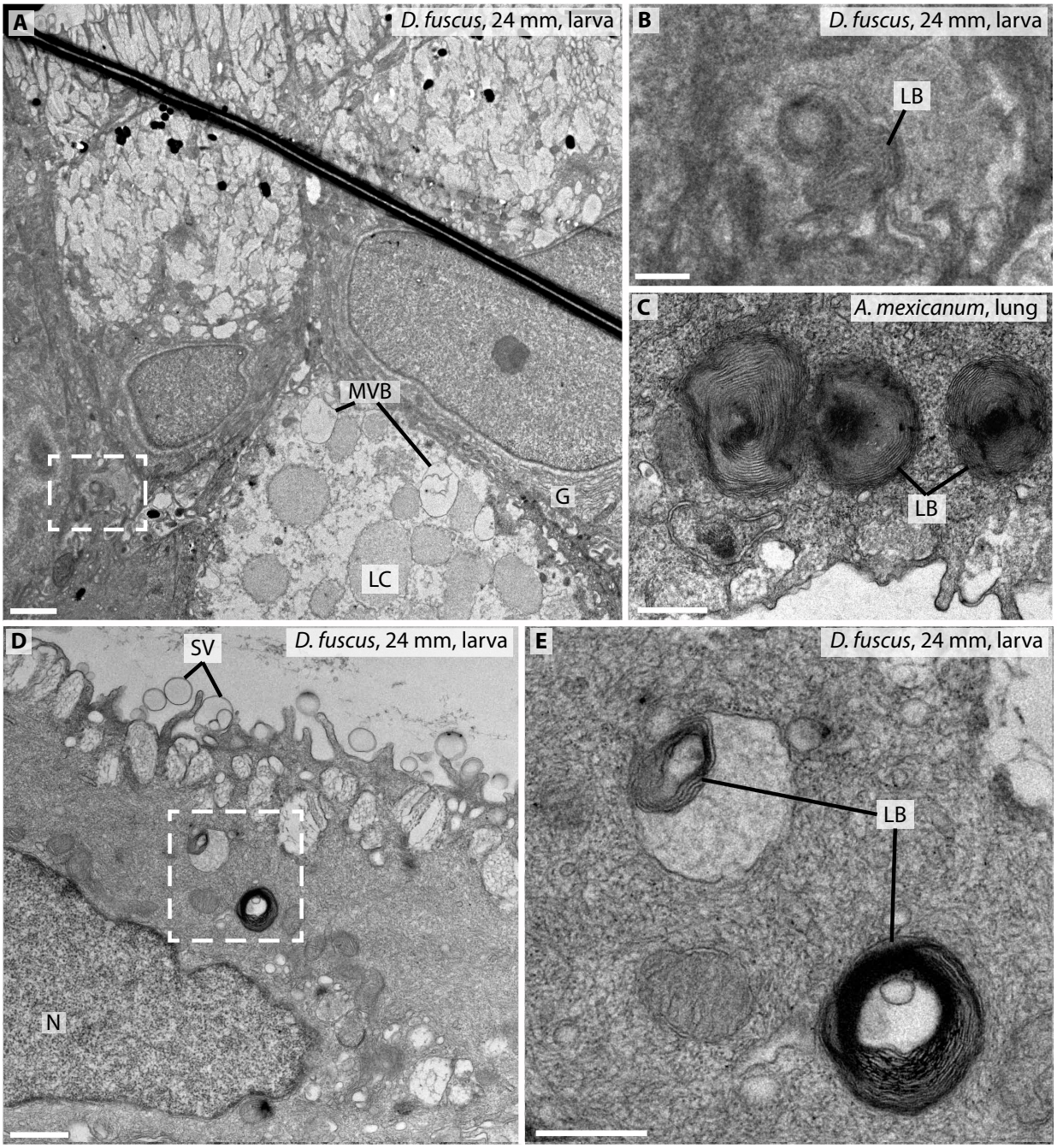
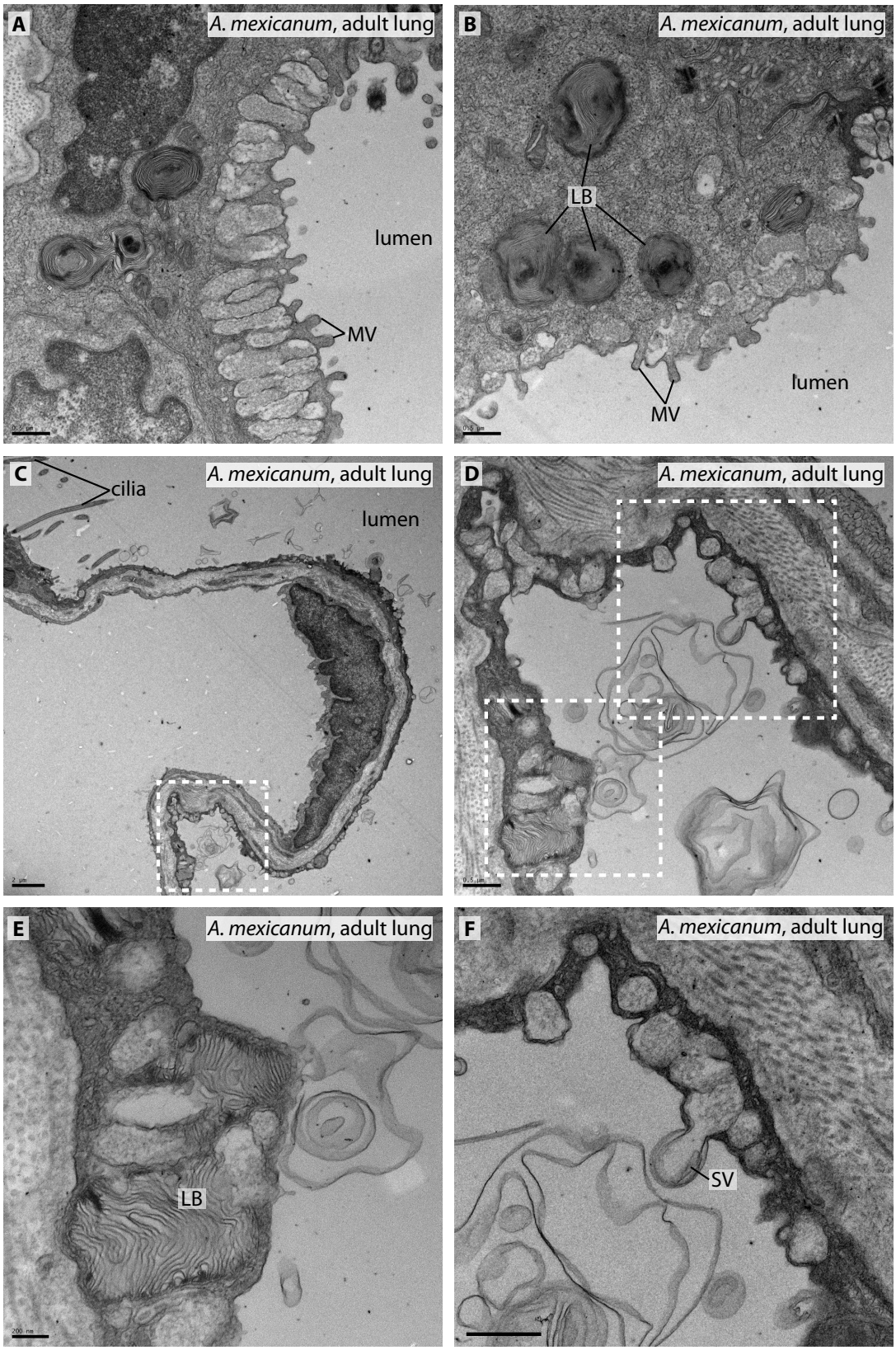


Figure 4.14: Potential lamellar bodies in *Desmognathus fuscus* integument. (A, B, D, E) Thin (≈ 90 nm) sections through the cranial integument of a larval *D. fuscus*. (A) Multiple cell types are visible in the epidermis, including a vacuolated Leydig cell (LC) as well as more superficial secretory cells. (B) Enlargement of boxed region in (A), showing a lamellar body-like structure (LB). (C) Section through an adult lung of *Ambystoma mexicanum* showing lamellar bodies (LB). See Fig. 4.16 for more information. (D, E) Lamellar bodies are visible in the *D. fuscus* integument. (E) is an enlargement of the boxed region in (D). Abbreviations: LB, lamellar body; LC, Leydig cell; MVB, multivesicular body; N, nucleus; SV, secretory vesicle. Scale bars: (A) 2 µm, (B,C,E) 0.5 µm, (D) 1 µm.

Figure 4.15: Lamellar bodies and secretory vesicles in the lungs of *Ambystoma mexicanum*. (A-F) Thin (≈70-80 nm) section through the distal portion of the *A. mexicanum* adult lung. (A,B) Lamellar bodies (LB) are visible within the cuboidal regions of the alveolar epithelial cells. (C-F) The squamous cytoplasmic processes of an alveolar epithelial cell. (C) The cytoplasmic process is visible as a dark band flanking the lumen. (D) Enlargement of the boxed region in (C), showing secretory activity in the lung. (E) Enlargement of the lower boxed region in (D) illustrates probable surfactant secretion by lamellar bodies. (F) Enlargement of the upper boxed region in (D) showing secretory vesicles similar to those seen on the skin of *Desmognathus fuscus* (Fig. 4.14A). Abbreviations: LB, lamellar body; MV, microvilli; SV, secretory vesicle; Scale bars: (A,B,D,F) 0.5 µm, (C) 2 µm, (E) 200 nm.

Figure 4.15, continued



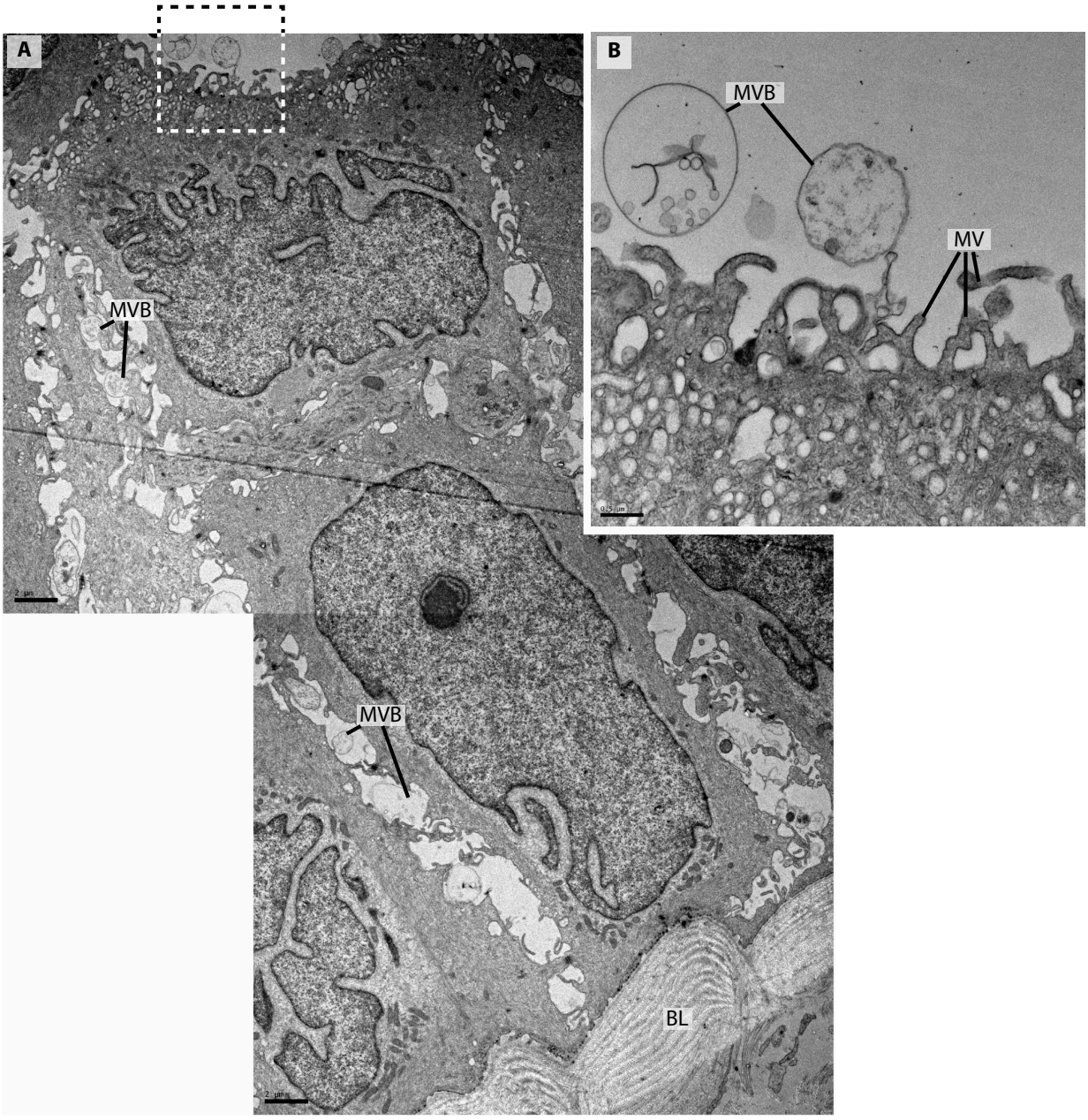


Figure 4.16: The integument of *Ambystoma mexicanum* does not play a pronounced secretory role. (A) The skin of a sexually mature *A. mexicanum*, from a thin (≈ 70-80 nm) section of the gular skin in transverse orientation. Multivesicular bodies (MVB) can be observed in the extracellular spaces. (B) Enlargement of the boxed region in (A). MVB are present external to the skin, but no active secretion of vesicles is visible. Abbreviations: BL, basal lamina; MV, microvilli; MBV, multivesicular bodies. Scale bars: (A) 2 µm, (B) 0.5 µm.

4.5 Discussion

I have discovered an undescribed gene with close sequence similarity to pulmonary surfactant protein C (SPC), which I name SPC-like. To date, SPC-like is found only in salamanders and not in any sequenced vertebrate genomes, suggesting that it arose from a gene duplication of SPC within salamanders. SPC expression is confined to the lungs in all vertebrates studied. SPC-like expression in a salamander with lungs, *Ambystoma mexicanum*, matches this conserved expression pattern. In contrast, SPC-like expression in the lungless salamander species *Desmognathus fuscus* is extrapulmonary: in the integument (skin) or the buccopharyngeal mucosa (mouth) depending on life history stage. During embryonic and larval life, SPC-like is expressed strongly in the integument. Prior to metamorphosis, expression expands to the buccopharyngeal cavity. Integumentary expression diminishes just prior to metamorphosis and is absent in recently metamorphosed animals and adults (Fig. 4.17). SPC-like most likely conserves some or all of the characteristic alpha-helical secondary structure of SPC (Fig. 4.3). These results, combined with histology, ultrastructure, and published observations on lungless salamander physiology (discussed below), suggest that following gene duplication SPC-like may have become neofunctionalized for extrapulmonary respiration in lungless salamanders.

4.5.1 SPC expression in tetrapods

Pulmonary surfactant is a complex and evolutionarily variable mixture of phospholipids and proteins (Daniels et al., 1994, 1995; Pattle and Hopkinson, 1963). The primary role of surfactant in mammals is to increase pulmonary compliance, but non-mammalian vertebrates

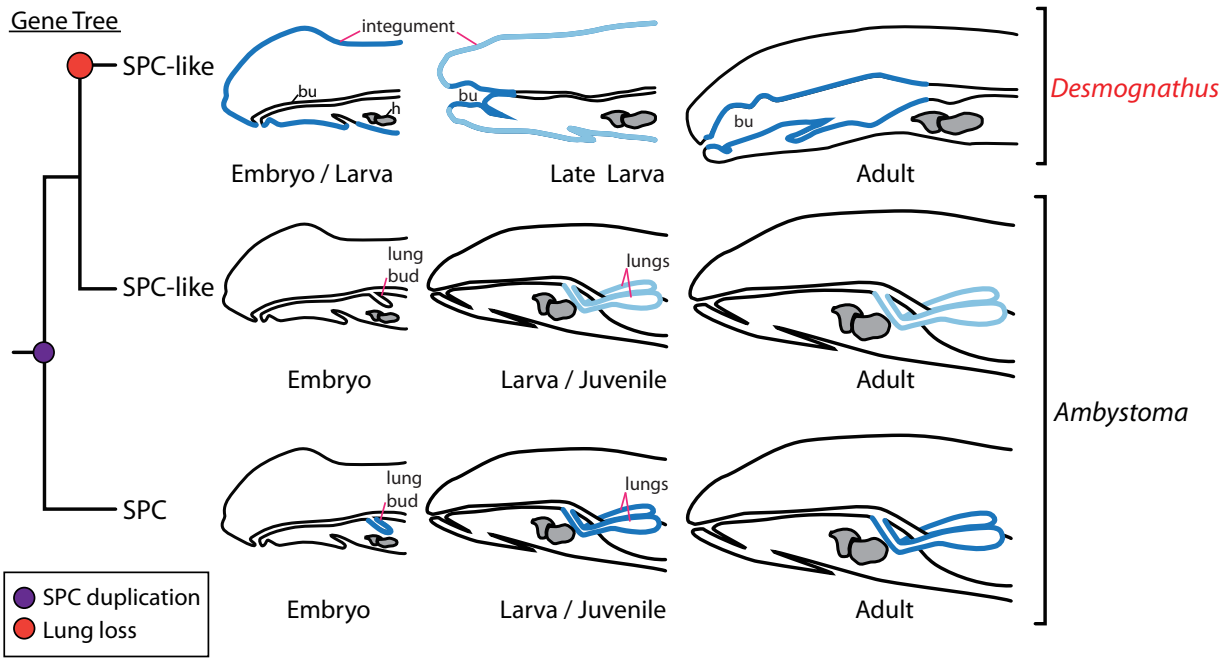


Figure 4.17: Summary of SPC and SPC-like expression patterns over ontogeny. Schematic sagittal sections showing SPC and SPC-like expression as blue outlines. Dark blue indicates a higher expression level than light blue. Abbreviations: bu, buccopharyngeal cavity; h, heart.

such as amphibians, reptiles, and lungfish purportedly use surfactant to prevent adherence of the collapsed lung to itself and to facilitate mucocilliary clearance of particles (Daniels et al., 1995, 2004).

A preponderance of studies on model species from mammals to amphibians report SPC specifically localized to the lungs (Bourbon and Chailley-Heu, 2001; Fisher et al., 1989; Glasser et al., 2000; Hyatt et al., 2007; Weaver and Whitsett, 1991; Wert et al., 1993; Wohlford-Lenane et al., 1992; Yin et al., 2010). However, four reports cite expression of SPC outside of the lungs in humans: Mo et al. (2006) report SPC in human fetal and adult skin. However, this study relies on immunohistochemistry using an antibody that, in my experience, yields spurious labeling, and on RT-PCR data without sequencing follow

up, which is subject to contamination and mispriming. Bräuer et al. (2009) report SPC expression in submandibular and parotid glands using RT-PCR and western immunoblots. These results are more credible because western blots reveal a protein of the expected size of the SPC pro-protein and RT-PCR of focal cDNA is performed alongside a positive control (although sequencing is still not employed). In a related study, Schicht et al. (2014) report the presence of SPC in saliva from human patients based on western blots and ELISA. These authors do not identify the antibody they used to detect SPC, and their isolated band is at 16 kDa, while SPC proprotein is 21 kDa and mature SPC 3.7 kDa (Beers and Mulugeta, 2005), casting doubt on these findings. Additionally, saliva may be subject to contamination by surfactant produced in the lungs. Schob et al. (2013) report SPC expression in central nervous system tissue and cerebrospinal fluid by RT-PCR but fail to rule out genomic DNA contamination. The western blots from Schob et al. (2013) also fail to demonstrate a band at the correct size for SPC, given the reported antibody that they employed. The authors also express confusion about how mRNA for SPC could be present in cerebrospinal fluid. Therefore, there are several weaknesses of all recent studies that cite extrapulmonary expression of SPC in humans, including possible contamination. Numerous studies in mammals and frogs, including ISH and reporter knockins, have failed to demonstrate extrapulmonary expression of SPC. For instance, a mouse line with Cre recombinase knocked in to the SPC locus fails to label cells outside of the lung when crossed to a reporter line (Barkauskas et al., 2013 and Christine Barkauskas, personal communication). It remains possible that humans and perhaps other animals endogenously express SPC outside of the lungs. Optimally, ISH and transcriptome sequencing should be used to confirm the human results presented above.

In mammals, Nkx2.1 is the principal regulator of SPC expression. Nkx2.1 acts synergis-

tically with a number of transcription factors, including Nuclear Factor I, Erm, Gata6, and TAZ to induce SPC expression specifically in type II alveolar epithelial cells (Bachurski et al., 2003; Kelly et al., 1996; Lin et al., 2006; Liu et al., 2002; Park et al., 2004). While SPC is normally expressed specifically in the lungs in lunged animals, experimental manipulation can affect localization of SPC transcripts. By increasing Wnt signaling using BIO or by the downregulation of Bmp4, which negatively regulates Wnt signaling, Rankin et al. (2012) observed expansion of SPC expression domains into the mouth and pharynx in *Xenopus laevis*. Bmp signaling downregulation results in both expanded SPC expression and lung loss, thus phenocopying the expression pattern and lung loss of adult *D. fuscus*.

SPC is has no close homologs (Hedlund et al., 2009; Hughes, 2007). The closest related gene is likely BRICHOS domain-containing protein 5 (Hedlund et al., 2009). SPC-like shows considerably higher sequence similarity to SPC than to BRICHOS domain-containing protein 5. BRICHOS domain-containing protein 5 sequences fall sister to SPC and SPC-like (Fig. 4.2 and data not shown). The two genes with closest sequence similarity to SPC in *Danio rerio* and *Latimeria chalumnae* likely both represent BRICHOS domain-containing protein 5. SPC is proposed to be absent in birds, indicating that it has been lost in association with the loss of blind ending alveoli and the presumed decreased selection for lung compliance (Hughes, 2007). That SPC is absent in birds, but maintained in plethodontids points to a functional role for this protein in the absence of lungs.

4.5.2 The discovery of SPC-like

I have identified a previously unknown gene, SPC-like, which is found so far only in three species of salamanders, *Ambystoma mexicanum* (Ambystomatidae), *Plethodon cinereus* and *Desmognathus fuscus* (Plethodontidae). These species share a last common ancestor (LCA) approximately 150 Ma (Shen et al., 2016). The ambystomatids and plethodontids phylogenetically bracket several families of salamanders and I expect that SPC-like is also found within these families. The fact that the SPC-like gene tree does not match the species tree given the proposed gene duplication of the ancestral SPC sequence is not problematic (Fig. 4.2). Instead, it points to a more ancient duplication of the ancestral SPC sequence than at the LCA of ambystomatids and plethodontids. Although gene evolution and speciation are coupled, neutral and adaptive processes result in sequence divergence that does not always recapitulate species divergence. Considerable coding mutations to SPC-like have occurred relative to SPC, and the resulting gene tree reflects this history.

Models differ as to the predicted secondary structure of SPC-like. Using the resolved structure as a template, SWISS-MODEL (Arnold et al., 2006) predicts a nearly identical secondary structure of SPC-like to SPC (Fig. 4.3E). This model is of low overall quality and should be treated with reservations. However, quality predictions for membrane proteins in SWISS-MODEL are generally inaccurate. The secondary structure of mature SPC-like by *Ab initio* predictive methods (Xu and Zhang, 2012) is similar to SPC at its C-terminal half, including the conservation of the extremely hydrophobic poly-valine domain (forming the C-terminal helix) (Johansson, 1998; Johansson et al., 1994). However, N-terminal sequence divergence results in a predicted helix hairpin. If this secondary structure model is cor-

rect, then the membrane-spanning capability of SPC-like may be divergent from SPC. Some mammalian SPC is palmitoylated at two N-terminal cysteine residues (5 and 6), leading to hydrophobicity of this non-helical domain (Curstedt et al., 1990). These residues are not conserved in SPC-like, which is further evidence that the N-terminal region of SPC-like is functionally divergent from SPC. However, these residues are not highly conserved across tetrapod SPC sequences, indicating that palmitoylation does not have an essential functional role (Fig. 4.1).

The mechanisms of surfactant and surfactant protein action are critical to understand from a clinical perspective. Respiratory distress syndrome (RDS), which is caused by insufficient production of pulmonary surfactant, results in collapsed lungs and difficulty breathing in infants, among other symptoms (Ma and Ma, 2012; Whitsett et al., 2015). Treatment for RDS involves the administration of surfactant. The addition of surfactant proteins to synthetic phospholipid surfactant improves clinical outcomes (Ma and Ma, 2012). SPC-like may have important clinical applications for the treatment of RDS. Future functional testing can confirm whether SPC-like performs similarly or better than SPC in aiding lipid adsorption or preventing pneumonitis and RDS.

4.5.3 Respiration in *Desmognathus fuscus*

Gas exchange requires a permeable surface exposed to water or air, but permeability in air results in susceptibility to dehydration. In biphasic lungless salamanders (those with aquatic and terrestrial stages), respiratory and hydric demands may shift dramatically at metamorphosis and the transition to terrestrial life. For instance, water loss through the skin becomes

a critical concern on land but reduction of skin permeability negatively influences cutaneous gas exchange. *Desmognathus fuscus* are biphasic, semi-aquatic lungless salamanders. Since the skin is their main respiratory source, post-metamorphic lungless salamanders are acutely challenged with balancing competing demands of fluid retention and efficient respiration because they have few means to actively regulate the amount of gas exchange and fluid loss over the skin. More flexibility is observed, however, in the buccopharyngeal mucosa, which can be ventilated on demand. Theoretically then, in an environment where fluid retention is not a limiting factor, larval salamanders should perform more cutaneous than buccopharyngeal respiration. At metamorphosis, this respiratory pattern would be expected to shift, resulting in decreased cutaneous respiration and increased buccopharyngeal respiration. This hypothesis is supported by a positive correlation between degree of terrestriality and degree of buccopharyngeal respiration (Whitford and Hutchison, 1965).

Buccopharyngeal respiration was recognized as a potential mode of compensation for lung loss upon the discovery of lungless salamanders (Wilder, 1896). Camerano (1896) experimentally tested the role of buccopharyngeal respiration in the lungless salamander *Hydromantes italicus*. He found that prevention of buccopharyngeal respiration (in adults, presumably) resulted in death and concluded that cutaneous respiration alone was insufficient for gas exchange in this species. He subsequently proposed that buccopharyngeal respiration predominates in lungless species (Camerano, 1896). Bethge (1898) discovered that the pulmonary artery, which normally delivers blood to the lungs, anastomoses with vessels in the pharynx and esophagus in the plethodontid *H. italicus*. Subsequently, H.H. Wilder and his student Anne Seelye (neé Barrows) morphologically characterized of the lining of the mouth of *Desmognathus fuscus* and proposed that this species breathes through

a "pharyngo-oesphageal lung" (Barrows, 1900; Seelye, 1906; Wilder, 1901). Elkan (1955, 1958) studied the histology of the skin and buccopharyngeal mucosa of lunged and lungless species. Elkan claimed that only two species of lungless salamanders that he examined, the semi-aquatic species *Eurycea bislineata* and *Pseudotriton ruber*, possessed extensive vascularization of the buccopharyngeal cavity, while he observed no pronounced vascularization in *D. fuscus*, terrestrial plethodontids and lunged salamanders. Czopek (1961, 1962, 1965) rigorously quantified the degree of integumentary and buccopharyngeal vascularization in amphibians. Czopek did not perform any statistical comparisons, but re-analysis of his data on salamander vascularization reveals that plethodontids, when normalized by body mass, have significantly higher cutaneous capillary length and buccal capillary length than lunged salamanders (both $p < 0.005$ by Welch's two-sample T-test). Many plethodontids form diverticula from their buccopharyngeal capillaries, which may serve to increase surface area (Czopek 1965). Approximately 10% of all respiratory capillaries in *D. fuscus* are found within the mouth. This proportion is among the highest for plethodontid salamanders (contradicting Elkan's qualitative assessments) and considerably higher than the mean proportion for lunged salamanders, which is about 2.48% (Czopek, 1965).

Camerano's (1896) assertion that buccopharyngeal respiration plays a larger role than cutaneous respiration in plethodontids has not been borne out by subsequent morphological or physiological studies. However, buccopharyngeal respiration occurs to a much greater extent in plethodontids than in lunged salamanders (Whitford and Hutchison 1965). *Desmognathus quadramaculatus* use 85% cutaneous respiration, compared to 76% in terrestrial *Plethodon glutinosus* and 50% in a lunged ambystomatid (Whitford and Hutchison 1965). Although it is difficult to uncouple buccopharyngeal and pulmonary respiration in lunged amphibians,

Whitford and Hutchison (1965) estimate that only approximately 2% of oxygen uptake is buccopharyngeal in lunged rough-skinned newts. This is markedly less than the 15–24% buccopharyngeal respiration in plethodontids.

There is evidence that plethodontids increase rates of buccopharyngeal ventilation (or "gular pumping") under hypoxia, high temperature, or activity (Czopek, 1965; Sheafor et al., 2000; Whitford and Hutchison, 1963, 1965). Therefore, it is possible that plethodontid buccopharyngeal membranes are employed as an adaptive respiratory surface to meet metabolic demands. However, based solely on histological comparisons, Elkan (1955, 1958) argues that buccopharyngeal ventilation is olfactory and does not play a role in respiration. Gatz et al. (1974b) also argue that gular pumping is not respiratory. Selection for efficient buccopharyngeal respiration may have played a major role in the adaptive radiation of terrestrial plethodontids (Whitford and Hutchison 1965). Cutaneous and buccopharyngeal respiration frees plethodontids from the constraints of pulmonary ventilation by buccal pumping, allowing them to occupy numerous habitats and employ ballistic tongue projection (Gatz et al., 1974a; Lombard and Wake, 1976, 1986).

4.5.4 A proposed role of SPC-like in plethodontids

The most commonly cited role for SPC in mammals is the reduction of surface tension at the air-liquid interface in the lung in order to prevent lung collapse and aid compliance. As discussed above, authors have proposed that surfactant acts as an anti-glue in amphibians to prevent the collapsed lung from sticking to itself (Daniels and Orgeig, 2001; Daniels et al., 1995). Provided that SPC-like performs similar functions, it is an open question as to what

purpose SPC-like serves in salamanders. Integumentary and buccopharyngeal expression of SPC-like raises additional functional questions: if not acting as an anti-glue, or aiding compliance, what could SPC-like be doing? I propose that SPC-like functions to reduce the thickness of the mucous layer.

Passive respiration over tissue (such as cutaneous respiration) is a function of several primary variables including the diffusivity of the tissue, the surface area of the tissue over which gas exchange occurs, the partial pressure difference of gasses across the tissue, and the thickness of the barrier between the underlying blood and the environment (Feder and Burggren, 1985). To modify gas exchange properties across a tissue, one or more of these variables must be altered. The partial pressure gradient is determined by the environment. Diffusivity is determined by tissue properties and is not easily altered. Remaining variables are surface area and barrier thickness. For cutaneous respiration, surface area can be modified at the ontogenetic level, such as in *Telmatobius culeus*, *Cryptobranchus alleghaniensis*, or potentially *Trichobatrachus robustus*, which possess additional skin surface and behavioral modifications to ventilate this tissue (Hutchison et al., 1976; Noble, 1925). Effective surface area can also be modified ontogenetically by changes to capillary density (Czopek, 1965). For instance, the amphibious fish *Kryptolebius marmoratus* may undergo vasculogenesis in the oral cavity following transition to air breathing, as evidenced by the onset of expression of CD31 (platelet endothelial cell adhesion molecule) within this tissue (Turko et al., 2014). On a shorter time scale, changing the degree of capillary perfusion within a tissue can actively modify the surface area of gas exchange, and there is evidence that this occurs in many species (Feder and Burggren, 1985).

Barrier thickness is mainly a function of the distance between the environment and the

blood supply. This distance may change over ontogeny with the thickening of the skin at metamorphosis. However, barrier thickness also depends on the thickness of the mucous layer between the environment and respiratory tissue. Mucous has a low diffusivity—about 30% lower than water (Ultsch and Gros, 1979). In the carp (*Cyprinus carpio*), a mucous layer of 5 µm would result in an 81% increase in resistance to diffusion of oxygen and this resistance scales linearly with increasing thickness. Increased resistance is due in part to the low diffusivity of mucous and in part to the creation of a non-convective boundary layer around the tissue (Ultsch and Gros 1979). Reduction of surface tension by pulmonary surfactant helps to maintain a thin layer of airway surface liquid (Daniels et al., 1998; Guyton et al., 1984) and increases convection within the mucous, resulting in better uptake of oxygen (Sosnowski et al., 1998). SPC-like may help perform one or both of these functions in order to reduce the effective barrier thickness or diffusivity of mucous. Additionally, SPC plays a role in regulating the production of DPPC, which is the main component of surfactant, thereby indirectly influencing mucous layer thickness (Rice et al., 1989). All of these potential reductions of mucous layer thickness would favorably alter gas diffusion properties across extrapulmonary tissue. SPC-like may act to decrease mucous thickness and increase efficiency of cutaneous and buccopharyngeal respiration.

4.5.5 The integument of plethodontid salamanders

Considerable attention has been placed on the morphology of the integument in lungless salamanders because of its function in respiration and its ability to protect terrestrial species from desiccation (raising the vital limit) (Littleford et al., 1947). In a broad study of plethod-

ontid integument, McMahon and Pav (1982) concluded that there were no apparent histological differences in the integument in terrestrial versus semiaquatic plethodontid species that would account for different vital limits. Instead, they suggested that physiological factors, not morphology, likely account for different vital limits. SPC-like could be one such factor.

SPC-like expression shifts at metamorphosis in *D. fuscus* coincident with changes in the morphology of the integument. Wilder (1925) carefully documented the metamorphic changes to the integument in amphibians. In the lungless species *Eurycea bislineata*, the larval skin is approximately two cell layers thick, well vascularized and possesses a type of non-glandular yet potentially secretory cells, which Wilder calls "Leydig" cells. (These Leydig cells are not to be confused with the Leydig cells in the testes.) Towards the beginning of metamorphosis, basally located cells begin to mitotically divide, increasing the thickness of the integument. The Leydig cells become more columnar, with nuclei located basally, and begin to degenerate (Wilder, 1925). The mucous (acinous) glands begin to form late in the pre-metamorphic period (Fig. 4.12). Metamorphosis involves three molting periods when various layers of the integument are sloughed off.

Although the loss of integumentary SPC-like expression correlates with the degeneration and eventual shedding of the Leydig cells, there is evidence to suggest that Leydig cells do not express SPC-like. *Desmognathus fuscus* larval epidermis is similar in gross structure to that reported for *E. bislineata* (Wilder, 1925) and other salamanders (Warburg and Lewinson, 1977). It is composed of roughly three layers: the deep layer is apposed to the collagenous basal lamina, the inner layer apical to these basal cells, and an anucleated keratinized superficial layer. Leydig cells are located in the basal layer (Figs. 4.12, 4.13, 4.14), whereas expression of SPC-like occurs in the superficial layer (Fig. 4.7). Putative Leydig cells in sec-

tioned material do not appear to express SPC-like (Fig. 4.7D). Instead, SPC-like expression appears strongly expressed in the inner layer and weakly expressed in the superficial layer of epidermis (Fig. 4.7). Ultrastructurally, cells in the inner layer possess an extensive Golgi apparatus and many apical vacuoles (Figs. 4.13, 4.14). The superficial layer corresponds to the stratum corneum, a keratinized protective layer, which appears anucleated but still active in secretion.

The larval *D. fuscus* stratum corneum appears to actively secrete bilamellar secretory vesicles along nearly the entirety of the surface of the integument (Fig. 4.13A). These bilamellar vesicles resemble those found in the *A. mexicanum* lung (Fig. 4.15F). The bilamellar vesicles also resemble known surfactant-containing vesicles reported in the literature. For instance, fractionated pulmonary lavage contains ultrastructurally similar structures within a fraction containing multilamellar bodies and tubular myelin (Magoon et al., 1983; Schmitz and Müller, 1991). Other studies have reported simple bilamellar bodies in the lungs of lizards (Wetzstein et al., 1980). In contrast, I find no similar structures in the skin of the lunged species *A. mexicanum* (Fig. 4.16) and I cannot find any report of these secretory vesicles in or on the integument of other species of salamander (Fox, 1984; Kelly, 1966a, 1966b; Lewinson et al., 1987; Ohmura and Wakahara, 1998; Warburg and Lewinson, 1977).

I also found evidence for multilamellar bodies in the larval integument of *D. fuscus* (Figs. 4.13B, 4.14). If future detailed characterization of these structures confirms morphological similarity to pulmonary lamellar bodies, then this will provide more evidence that *D. fuscus* secretes surfactant, along with SPC-like, from its skin.

4.5.6 Multimodal respiration and the adaptive radiation of plethodontid salamanders

My data support an ontogenetic shift in expression of SPC-like in *Desmognathus fuscus*, which correlates with the transition from aquatic respiration to aerial respiration. Gatz et al. (1974a) propose that adaptations for high rates of cutaneous respiration in *D. fuscus* (and plethodontid salamanders in general) make them highly suited for respiration in water and in air, potentially contributing to the evolutionary success of this lineage. Until now, researchers have sought morphological adaptations to explain high rates of buccopharyngeal and cutaneous respiration in plethodontids. I propose that these morphological differences work in concert with molecular changes such as the neofunctionalization of SPC-like to enable efficient extrapulmonary respiration. A better understanding of the evolution and function of SPC-like will contribute to a more complete picture of the evolution of lung loss and the evolution of major physiological transitions in general. My results suggest a mechanism by which plethodontids have become the largest and most widespread radiation of salamanders on the planet, despite the theoretical limitations on thermal tolerance and body size imposed by lunglessness.

4.6 References

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22, 195–201.

Bachurski, C.J., Yang, G.H., Currier, T.A., Gronostajski, R.M., and Hong, D. (2003). Nuclear factor I/thyroid transcription factor 1 interactions modulate surfactant protein C

transcription. Mol. Cell. Biol. 23, 9014–9024.

Barkauskas, C.E., Cronce, M.J., Rackley, C.R., Bowie, E.J., Keene, D.R., Stripp, B.R., Randell, S.H., Noble, P.W., and Hogan, B.L.M. (2013). Type 2 alveolar cells are stem cells in adult lung. J. Clin. Invest. 123, 3025–3036.

Barrows, A.I. (1900). Respiration of *Desmognathus*. Anat. Anz. 18, 461–464.

Beers, M.F., and Mulugeta, S. (2005). Surfactant protein C biosynthesis and its emerging role in conformational lung disease. Annu. Rev. Physiol. 67, 663–696.

Bethge, E. (1898). Das blutgefälssystem von *Salamandra maculata*, *Triton taeniatus* und *Spelerpes fuscus*; mit betrachtungen über den ort der athmung bein lungenlosen *Spelerpes fuscus*. Z. Wiss. Zool. 63, 680–706.

Bourbon, J.R., and Chailley-Heu, B. (2001). Surfactant proteins in the digestive tract, mesentery, and other organs: evolutionary significance. Comp. Biochem. Physiol. A, Mol. Integr. Physiol. 129, 151–161.

Bräuer, L., Möschter, S., Beileke, S., Jäger, K., Garreis, F., and Paulsen, F.P. (2009). Human parotid and submandibular glands express and secrete surfactant proteins A, B, C and D. Histochem. Cell Biol. 132, 331–338.

Bruce, R.C., Beachy, C.K., Lenzo, P.G., Pronych, S.P., and Wassersug, R.J. (1994). Effects of lung reduction on rheotactic performance in amphibian larvae. J. Exp. Zool. 268, 377–380.

Camerano, L. (1896). Recherches anatomo-physiologiques sur les salamandres normalement privees de poumons. Arch. Ital. Biol. 21, 387–395.

Che, R., Sun, Y., Wang, R., and Xu, T. (2014). Transcriptomic analysis of endangered Chinese salamander: Identification of immune, sex and reproduction-related genes and genetic markers. PLoS One 9, e87940.

Christenson, M.K., Trease, A.J., Potluri, L.-P., Jezewski, A.J., Vincent, M., Knight, L.A., Kolok, A.S., Davis, P.H., Davis, V.M., Knight, L.A., et al. (2014). *De novo* assembly and analysis of the Northern leopard frog *Rana pipiens* transcriptome. J. Genomics 2, 141–149.

Connelly, I., and Possmayer, F. (1992). cDNA sequence and alternative mRNA splicing of surfactant-associated protein C (SP-C) in rabbit lung. Biochim Biophys Acta 1127, 199–207.

Curstedt, T., Johansson, J., Persson, P., Eklund, a, Robertson, B., Löwenadler, B., and Jörnvall, H. (1990). Hydrophobic surfactant-associated polypeptides: SP-C is a lipopeptide with two palmitoylated cysteine residues, whereas SP-B lacks covalently linked fatty acyl groups. Proc. Natl. Acad. Sci. U. S. A. 87, 2985–2989.

Czopek, J. (1961). Vascularization of respiratory surfaces in some Plethodontidae. Zool. Pol. 11, 131–148.

Czopek, J. (1962). Vascularization of respiratory surfaces in some caudata. Copeia 1962, 576–587.

Czopek, J. (1965). Quantitative studies on the morphology of respiratory surfaces in amphibians. Acta Anat. (Basel). 62, 296–323.

Daniels, C.B., and Orgeig, S. (2001). The comparative biology of pulmonary surfactant: past, present and future. Comp. Biochem. Physiol. Part A, Mol. Integr. Physiol. 129, 9–36.

Daniels, C.B., Orgeig, S., Wilsen, J., and Nicholas, T.E. (1994). Pulmonary-type surfactants in the lungs of terrestrial and aquatic amphibians. Respir. Physiol. 95, 249–258.

Daniels, C.B., Orgeig, S., and Smits, A.W. (1995). The evolution of the vertebrate pulmonary surfactant system. Physiol. Zool. 68, 539–566.

Daniels, C.B., Lopatko, O. V, and Orgeig, S. (1998). Evolution of surface activity related functions of vertebrate pulmonary surfactant. Clin. Exp. Pharmacol. Physiol. 25, 716–721.

Daniels, C.B., Orgeig, S., Sullivan, L.C., Ling, N., Bennett, M.B., Schürch, S., Val, A.L., and Brauner, C.J. (2004). The origin and evolution of the surfactant system in fish: insights into the evolution of lungs and swim bladders. Physiol. Biochem. Zool. 77, 732–749.

Doyle, J.M., Siegmund, G., Ruhl, J.D., Eo, S.H., Hale, M.C., Marra, N.J., Waser, P.M., and Dewoody, J.A. (2013). Microsatellite analyses across three diverse vertebrate transcriptomes. Genome 56, 407–414.

Dunn, E. (1926). The Salamanders of the Family Plethodontidae (Northampton: Smith College).

Elkan, E. (1955). The buccal and pharyngeal mucous membranes in urodeles. Proc. Zool. Soc. London 125, 685–711.

Elkan, E. (1958). Further contributions on the buccal and pharyngeal mucous membranes in urodeles. Proc. Zool. Soc. London 131, 335–355.

Feder, M.E. (1983). Effect of hypoxia and body size on the energy metabolism of lungless tadpoles, *Bufo woodhousei*, and air-breathing anuran larvae. J. Exp. Zool. 40, 127–19.

Feder, M.E., and Burggren, W.W. (1985). Cutaneous gas exchange in vertebrates: design, patterns, control and implications. Biol. Rev. Camb. Philos. Soc. 60, 1–45.

Fisher, J.H., Shannon, J.M., Hofmann, T., and Mason, R.J. (1989). Nucleotide and

deduced amino acid sequence of the hydrophobic surfactant protein SP-C from rat: expression in alveolar type II cells and homology with SP-C from other species. Biochim. Biophys. Acta 995, 225–230.

Fox, H. (1984). The skin of Amphibia: epidermis. In Biology of the Integument II: Vertebrates, J. Bereiter-Han, A. Matoltsy, and K. Richards, eds. (Berlin: Springer-Verlag), pp. 78–110.

Gatz, R.N., Crawford, E.C., and Piiper, J. (1974a). Respiratory properties of the blood of a lungless and gill-less salamander, *Desmognathus fuscus*. Respir. Physiol. 20, 33–41.

Gatz, R.N., Crawford, E.C., and Piiper, J. (1974b). Metabolic and heart rate response of the plethodontid salamander *Desmognathus fuscus* to hypoxia. Respir. Physiol. 20, 43–49.

Glasser, S.W., Korfhagen, T.R., Perme, C.M., Pilot-Matias, T.J., Kister, S.E., and Whitsett, J.A. (1988). Two SP-C genes encoding human pulmonary surfactant proteolipid. J. Biol. Chem. 263, 10326–10331.

Glasser, S.W., Burhans, M.S., Eszterhas, S.K., Bruno, M.D., and Korfhagen, T.R. (2000). Human SP-C gene sequences that confer lung epithelium-specific expression in transgenic mice. Am. J. Physiol. Lung Cell. Mol. Physiol. 278, L933–L945.

Goniakowska-Witalińska (1980). Ultrastructural and morphometric changes in the lung of newt, *Triturus cristatus carnifex* Laur. during ontogeny. J. Anat. 130, 571–583.

Guyton, A.C., Moffatt, D.S., and Adiar, T.H. (1984). Role of alveolar surface tension in transepithelial movement of fluid. In Pulmonary Surfactant, B. Robertson, L. Van Golde, and J. Batenburg, eds. (Amsterdam: Elsevier Science Publishers), pp. 171–185.

Hedlund, J., Johansson, J., and Persson, B. (2009). BRICHOS - a superfamily of multidomain proteins with diverse functions. BMC Res. Notes 2, 180.

Hughes, A.L. (2007). Evolution of the lung surfactant proteins in birds and mammals. Immunogenetics 59, 565–572.

Hutchison, V.H., Haines, H.B., and Engbretson, G. (1976). Aquatic life at high altitude: respiratory adaptations in the Lake Titicaca frog, *Telmatobius culeus*. Respir. Physiol. 27, 115–129.

Hyatt, B.A., Resnik, E.R., Johnson, N.S., Lohr, J.L., and Cornfield, D.N. (2007). Lung specific developmental expression of the *Xenopus laevis* surfactant protein C and B genes. Gene Expr. Patterns 7, 8–14.

Johansson, J. (1998). Structure and properties of surfactant protein C. Biochim. Biophys. Acta 1408, 161–172.

Johansson, J., Szyperski, T., Curstedt, T., and Wüthrich, K. (1994). The NMR structure

of the pulmonary surfactant-associated polypeptide SP-C in an apolar solvent contains a valyl-rich alpha-helix. Biochemistry 33, 6015–6023.

Kelly, D.E. (1966a). The Leydig cell in larval amphibian epidermis. Fine structure and function. Anat. Rec. 154, 685–699.

Kelly, D.E. (1966b). Fine structure of desmosomes, hemidesmosomes, and an adepidermal globular layer in developing newt epidermis. J. Cell Biol. 28, 51–72.

Kelly, S.E., Bachurski, C.J., Burhans, M.S., and Glasser, S.W. (1996). Transcription of the lung-specific Surfactant protein C gene is mediated by Thyroid transcription factor 1. J. Biol. Chem. 271, 6881–6888.

Lewinson, D., Rosenberg, M., and Warburg, M.R. (1987). Ultrastructural and ultracytochemical studies of the gill epithelium in the larvae of *Salamandra salamandra* (Amphibia, Urodela). Zoomorphology 107, 17–25.

Lin, S., Perl, A., and Shannon, J.M. (2006). Erm/Thyroid transcription factor 1 interactions modulate surfactant protein C transcription. J. Biol. Chem. 281, 16716–16726.

Littleford, R.A., Keller, W.F., and Phillips, N.E. (1947). Studies on the vital limits of water loss in the plethodont salamanders. Ecology 28, 440–447.

Liu, C., Glasser, S.W., Wan, H., and Whitsett, J.A. (2002). GATA-6 and Thyroid transcription tactor-1 directly interact and regulate Surfactant protein-C gene expression. J. Biol. Chem. 277, 4519–4525.

Lombard, R.E., and Wake, D.B. (1976). Tongue evolution in the lungless salamanders, family Plethodontidae I. Introduction, theory and a general model of dynamics. J. Morphol. 148, 265–286.

Lombard, R.E., and Wake, D.B. (1986). Tongue evolution in the lungless salamanders, family Plethodontidae IV. Phylogeny of plethodontid salamanders and the evolution of feeding dynamics. Syst. Zool. 35, 532–551.

Löytynoja, A., and Goldman, N. (2010). webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. BMC Bioinformatics 11, 579.

Ma, C.C.-H., and Ma, S. (2012). The role of surfactant in respiratory distress syndrome. Open Respir. Med. J. 6, 44–53.

Magoon, M.W., Wright, J.R., Baritussio, A., Williams, M.C., Goerke, J., Benson, B.J., Hamilton, R.L., and Clements, J.A. (1983). Subfractionation of lung surfactant. Implications for metabolism and surface activity. Biochim Biophys Acta 750, 18–31.

McMahon, S.H., and Pav, D.I. (1982). Histological examination of the integument of fifteen species of plethodontid salamanders (Amphibia: Urodela: Plethodontidae). J. Tenn.

Acad. Sci. 57, 78–81.

Mo, Y.K., Kankavi, O., Masci, P.P., Mellick, G.D., Whitehouse, M.W., Boyle, G.M., Parsons, P.G., Roberts, M.S., and Cross, S.E. (2006). Surfactant protein expression in human skin: evidence and implications. J. Invest. Dermatol. 127, 381–386.

Noble, G.K. (1925). The integumentary, pulmonary, and cardiac modifications correlated with increased cutaneous respiration in the amphibia: A solution of the "hairy frog" problem. J. Morphol. 40, 341–416.

Ohmura, H., and Wakahara, M. (1998). Transformation of skin from larval to adult types in normally metamorphosing and metamorphosis-arrested salamander, *Hynobius retardatus*. Differentiation 63, 237–246.

Orgeig, S., Bernhard, W., Biswas, S.C., Daniels, C.B., Hall, S.B., Hetz, S.K., Lang, C.J., Maina, J.N., Panda, A.K., Perez-Gil, J., et al. (2007). The anatomy, physics, and physiology of gas exchange surfaces: is there a universal function for pulmonary surfactant in animal respiratory structures? Integr. Comp. Biol. 47, 610–627.

Park, K.-S., Whitsett, J.A., Di Palma, T., Hong, J.-H., Yaffe, M.B., and Zannini, M. (2004). TAZ interacts with TTF-1 and regulates expression of surfactant protein-C. J. Biol. Chem. 279, 17384–17390.

Pattle, R., and Hopkinson, D. (1963). Lung lining in bird, reptile and amphibian. Nature 200, 894.

Possmayer, F., Nag, K., Rodriguez, K., Qanbar, R., and Schürch, S. (2001). Surface activity in vitro: role of surfactant proteins. Comp. Biochem. Physiol. Part A 129, 209–220.

Rankin, S.A., Gallas, A.L., Neto, A., Gómez-Skarmeta, J.L., and Zorn, A.M. (2012). Suppression of Bmp4 signaling by the zinc-finger repressors Osr1 and Osr2 is required for Wnt/β-catenin-mediated lung specification in *Xenopus*. Development 139, 3010–3020.

Rankin, S.A., Thi Tran, H., Wlizla, M., Mancini, P., Shifley, E.T., Bloor, S.D., Han, L., Vleminckx, K., Wert, S.E., and Zorn, A.M. (2015). A molecular atlas of *Xenopus* respiratory system development. Dev. Dyn. 244, 69–85.

Reagan, N.L., and Verrell, P.A. (1991). The evolution of plethodontid salamanders: did terrestrial mating facilitate lunglessness? Am. Nat. 138, 1307–1313.

Rice, W.R., Sarin, V.K., Fox, J.L., Baatz, J., Wert, S., and Whitsett, J.A. (1989). Surfactant peptides stimulate uptake of phosphatidylcholine by isolated cells. Biochim. Biophys. Acta 1006, 237–245.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: Efficient Bayesian

phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.

Ruben, J., and Boucot, A. (1989). The origin of the lungless salamanders (Amphibia: Plethodontidae). Am. Nat. 134, 161–169.

Schicht, M., Stengl, C., Sel, S., Heinemann, F., Götz, W., Petschelt, A., Pelka, M., Scholz, M., Rausch, F., Paulsen, F., et al. (2015). The distribution of human surfactant proteins within the oral cavity and their role during infectious diseases of the gingiva. Ann. Anat. - Anat. Anzeiger 199, 92–97.

Schmitz, G., and Müller, G. (1991). Structure and function of lamellar bodies, lipid-protein complexes involved in storage and secretion of cellular lipids. J. Lipid Res. 32, 1539–1570.

Schob, S., Schicht, M., Sel, S., Stiller, D., Kekulé, A., Paulsen, F., Maronde, E., and Bräuer, L. (2013). The detection of surfactant proteins A, B, C and D in the human brain and their regulation in cerebral infarction, autoimmune conditions and infections of the CNS. PLoS One 8, e74412.

Seelye, A. (1906). Circulatory and respiratory systems of *Desmognathus fusca*. Proc. Bost. Soc. Nat. Hist. 32, 335–357.

Sheafor, E.A., Wood, S.C., and Tattersall, G.J. (2000). The effect of graded hypoxia on the metabolic rate and buccal activity of a lungless salamander (*Desmognathus fuscus*). J. Exp. Biol. 203, 3785–3793.

Shen, X.-X., Liang, D., Chen, M.-Y., Mao, R.-L., Wake, D.B., and Zhang, P. (2016). Phylogeny, time and biogeography of plethodontids. Syst. Biol. 65, 66–81.

Sosnowski, T., Gradon, L., Skoczek, M., and Drozdiel, H. (1998). Experimental evaluation of importance of the pulmonary surfactant for oxygen transfer rate in human lungs. Int. J. Occup. Saf. Ergon. 4, 391–409.

Turko, A.J., Robertson, C.E., Bianchini, K., Freeman, M., and Wright, P. a (2014). The amphibious fish *Kryptolebias marmoratus* uses different strategies to maintain oxygen delivery during aquatic hypoxia and air exposure. J. Exp. Biol. 217, 3988–3995.

Ultsch, G.R., and Gros, G. (1979). Mucus as a diffusion barrier to oxygen: Possible role in O_2 uptake at low pH in carp (*Cyprinus carpio*) gills. Comp. Biochem. Physiol. Part A Physiol. 62, 685–689.

Veldhuizen, R., Nag, K., Orgeig, S., and Possmayer, F. (1998). The role of lipids in pulmonary surfactant. Biochim. Biophys. Acta 1408, 90–108.

Wake, D.B., and Hanken, J. (1996). Direct development in the lungless salamanders: what are the consequences for developmental biology, evolution and phylogenesis? Int. J. Dev. Biol. 40, 859–869.

Wake, M.H., and Donnelly, M.A. (2010). A new lungless caecilian (Amphibia: Gymnophiona) from Guyana. Proc. Biol. Sci. 277, 915–922.

Warburg, M.R., and Lewinson, D. (1977). Ultrastructure of epidermis of *Salamandra salamandra* followed throughout ontogenesis. Cell Tissue Res. 181, 369–393.

Weaver, T.E., and Conkright, J.J. (2001). Functions of surfactant proteins B and C. Annu. Rev. Physiol. 63, 555–578.

Weaver, T.E., and Whitsett, J.A (1991). Function and regulation of expression of pulmonary surfactant-associated proteins. Biochem. J. 273, 249–264.

Wert, S.E., Glasser, S.W., Korfhagen, T.R., and Whitsett, J.A. (1993). Transcriptional elements from the human SP-C gene direct expression in the primordial respiratory epithelium of transgenic mice. Dev. Biol. 156, 426–443.

Wetzstein, H., Stratton, C., and Olson, N. (1980). The stereoscanning electron microscopy and ultrastructural histochemistry of the avian and reptilian surfactant systems: Indian dove, desert spiny and Taiwan golden skink lizards. Anat. Rec. 197, 63–73.

Whitford, W., and Hutchison, V. (1963). Cutaneous and pulmonary gas exchange in the spotted salamander, *Ambystoma maculatum*. Biol. Bull. 124, 344–354.

Whitford, W.G., and Hutchison, V.H. (1965). Gas exchange in salamanders. Physiol. Zool. 38, 228–242.

Whitsett, J.A., and Weaver, T.E. (2002). Hydrophobic surfactant proteins in lung function and disease. N. Engl. J. Med. 347, 2141–2148.

Whitsett, J.A., Wert, S.E., and Weaver, T.E. (2015). Diseases of pulmonary surfactant homeostasis. Annu. Rev. Pathol. Mech. Dis. 10, 371–393.

Wilder, H.H. (1896). Lungless salamanders. Anat. Anz. 12, 182–192.

Wilder, H.H. (1901). The pharyngo-oesophageal lung of *Desmognathus*. Am. Nat. 35, 183–186.

Wilder, I.W. (1925). The Morphology of Amphibian Metamorphosis (Northampton: Smith College).

Wilder, I.W., and Dunn, E. (1920). The correlation of lunglessness in salamanders with a mountain brook habitat. Copeia 1920, 63–68.

Wohlford-Lenane, C.L., Durham, P.L., and Snyder, J.M. (1992). Localization of Surfactant-associated protein C (SP-C) mRNA in fetal rabbit lung tissue by *in situ* hybridization. Am. J. Respir. Cell Mol. Biol. 6, 225–234.

Xu, D., and Zhang, Y. (2012). *Ab initio* protein structure assembly using continuous

structure fragments and optimized knowledge-based force field. Proteins 80, 1715–1735.

Yin, A., Winata, C.L., Korzh, S., Korzh, V., and Gong, Z. (2010). Expression of components of Wnt and Hedgehog pathways in different tissue layers during lung development in *Xenopus laevis*. Gene Expr. Patterns 10, 338–344.

Chapter 5

Convergent evolution of the heart in lungless salamanders

Co-author for the publication: James Hanken

5.1 Abstract

Nearly two thirds of the approximately 700 species of living salamanders are lungless. These species breathe entirely through the skin and buccopharyngeal mucosa. Lung loss dramatically impacts the configuration of the circulatory system, but the effects of evolutionary lung loss on cardiac morphology have long been controversial. For example, the atrial septum presumably would have little function in lungless salamanders due to the absence of pulmonary veins and the presence of a single source of mixed blood flowing into the heart, but whether lungless salamanders even possess an atrial septum and whether the sinoatrial aperture is located in the left or right atrium are unresolved; authors have stated opposing claims since

the late 1800s. Here, we use µ-CT imaging, gross dissection and histological reconstruction to compare cardiac morphology among lungless plethodontid salamanders (Plethodontidae), salamanders with lungs, and the convergently lungless species *Onychodactylus japonicus* (Hynobiidae). Plethodontid salamanders have partial atrial septa and incomplete separation of the atrium into left and right halves. Partial septation is also seen in *O. japonicus*. The partial septum in plethodontids can make it appear that the sinoatrial aperture is in the left atrium, but this interpretation is incorrect. Outgroup comparisons demonstrate that the aperture is located in a posterodorsal extension of the right atrium into the left side of the heart. Hence, lungless salamanders from two lineages convergently evolve similar morphology of the atrial septum. Convergent losses of the atrial septum may have a similar developmental basis. In mammals, the lungs induce formation of the atrial septum by secreting morphogens to neighboring mesenchyme. Candidate genes involved in atrial septum development are expressed at significantly lower levels in lungless salamander second heart field mesoderm relative to a lunged salamander, offering a potential mechanism of atrial septum reduction. We hypothesize that the lungs induce atrial septum development in amphibians in a similar fashion to mammals, and that atrial septum reduction in lungless salamanders is a direct result of lunglessness.

5.2 Introduction

Integrated organ systems are hallmarks of bilaterian life. However, such integration may also constrain potentially adaptive evolutionary changes (Gould, 1980). Saltatory changes to single organs, such as organ loss, may be maladaptive within interconnected organ systems

because of the need for the system as a whole to function in order to maintain organismal fitness. These constraints can be circumvented by pleiotropy, whereby single genes govern development of multiple organs, or circumvented by evolved molecular interactions between organ systems that ensure matching morphologies develop in tandem. An example of the latter may be found in the cardiopulmonary system. In vertebrates, the heart and lungs function as an integrated system. Complete cardiac septation, including the development of ventricular and atrial septa, has evolved convergently in association with the evolution of endothermy in order to achieve separation of oxygenated and deoxygenated blood and efficient gas exchange (Jensen et al., 2013; Koshiba-Takeuchi et al., 2009). Molecular signals secreted from the developing lungs mediate the development of the atrial septum. Indeed, the lungs may provide cues to the heart that enable the functional integration of both organs (Hoffmann et al., 2009, 2014). This interaction provides one paradigm for understanding how the lungs and heart are capable of correlated development, and thus correlated evolution.

Complete cardiac separation is found only within mammals and certain reptiles. Amphibians typically have a three-chambered heart, with a septate atrium and an undivided ventricle (Noble, 1925). Yet, and despite the lack of ventricular septation, deoxygenated and oxygenated blood mostly remain separate as they pass through the heart (Haberich, 1965; Johansen, 1962; Noble, 1925). Atrial septum development is best understood in mammals, in part due to the high prevalence of human atrial septal defects (ASDs). Patent foramen ovale, for example, may afflict as many as 25% of the adult population (Taussig, 1960). Congenital ASDs, while less frequent, are still among the most common birth defects (Bjornard et al., 2013; Parker et al., 2010). This has prompted several recent investigations into both the development of the atrial septum (AS) in mammals (Goddeeris et al., 2008; Hoffmann

et al., 2009, 2014; Xie et al., 2012) and the genetic basis of certain AS malformations, such as those associated with Holt-Oram Syndrome (Basson et al., 1997) caused by mutations in the T-box protein Tbx5 (Li et al., 1997). Atrial septum development in mammals depends on the expression of the Forkhead box transcription factors Foxf1a and Foxf2 (Hoffmann et al., 2014). Sonic hedgehog (Shh) secreted by the developing lung activates the Shh-signaling pathway in second heart field mesoderm, which lies ventral to the pulmonary endoderm. The Shh pathway transcription factor Gli1 and Tbx5 bind to a cis-regulatory element for Foxf1a and synergistically activate Foxf1a expression. Foxf2 expression similarly depends on Shh and Tbx5, but the cis-regulatory interactions are unknown (Hoffmann et al., 2014). Although mammalian atrial septum development is now well characterized, an understanding of AS development and morphology in other taxa is lacking.

Though amphibians descended from ancestors with lungs, lungless species or clades have originated at least four times independently across all living amphibian orders (Bickford et al., 2008; Noble, 1925; Nussbaum and Wilkinson, 1995; Wilder, 1896). Lung loss is a remarkable example of morphological convergence. Little attention has been paid to the corresponding morphology of the circulatory system in lungless species, and the entirety of this work has focused on only a single lineage of lungless salamanders, the family Plethodontidae. Adult plethodontids are completely lungless salamanders that respire cutaneously and through the buccopharyngeal mucosa (Whitford and Hutchison, 1965). Despite of several studies of this one group, the degree of atrial septation in plethodontid salamanders has long been controversial. Gross dissection of plethodontid hearts has yielded a number of contradictory hypotheses regarding atrial septum morphology (Fig. 5.1).

The morphology of the plethodontid AS, and even its presence or absence, has been con-

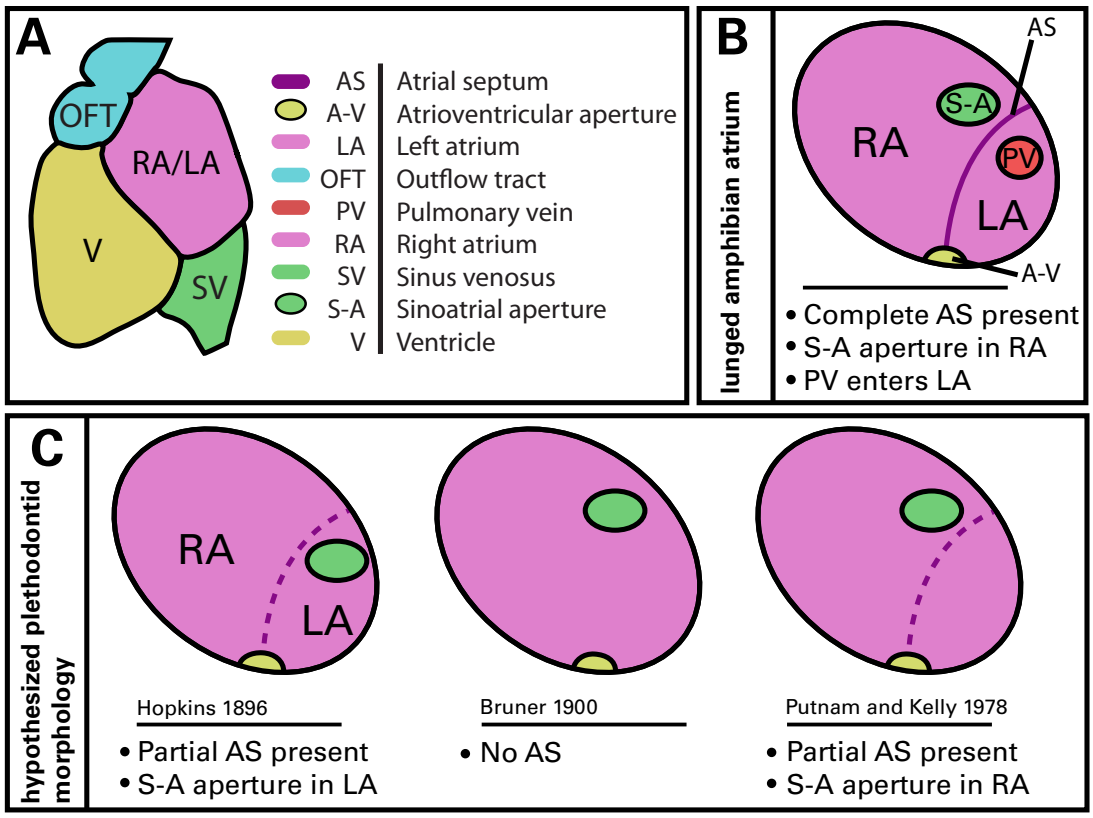


Figure 5.1: Previous hypotheses on the morphology of the plethodontid atrium. (A) Abbreviations, color schema and general heart morphology from a ventral perspective. (B) The generalized morphology of the atrium of an amphibian with lungs. (C) Three hypotheses on the morphology of the atrium in plethodontid (lungless) salamanders. The hypotheses differ regarding the location of the sinoatrial aperture and whether an atrial septum is present.

tentious since the end of the 19th century. The atrial septum is thin and delicate, which may contribute to the difficulty in resolving contradictory hypotheses. In addition, each previous study used different fixation techniques and examined a different plethodontid species, sometimes without comparison to a lunged outgroup (Bruner, 1900; Hopkins, 1896; Noble, 1925; Putnam and Kelly, 1978). No attempt has been made to study this structure using even standard histology, let alone more advanced visualization methods such as computed tomography (CT). Approaches that incorporate novel morphological techniques and broad

comparative data sets may help to reconcile the conflicting hypotheses on the plethodontid AS and yield greater understanding of the evolution of integrated organ systems in general.

Among the few non-plethodontid lungless amphibians is the hynobiid salamander genus *Onychodactylus* from mainland Asia and Japan (Yoshikawa and Matsui, 2014; Yoshikawa et al., 2008). The morphology of the heart in *Onychodactylus* has not been examined, despite its relevance to the unresolved hypotheses regarding of the morphology of the heart of other lungless amphibians. Exploring the mechanisms and consequences of lung loss in *Onychodactylus* will enable a better understanding of the basis of convergent evolution between lungless clades.

Here we employ several techniques to examine AS morphology across several lungless and lunged salamander species (Table 5.1). Histological sectioning and three-dimensional reconstruction are combined with gross dissection and contrast-stained μ-CT. Unlike lunged salamanders, which have a two-part AS, plethodontids instead possess a partial AS due to absence of one of the two primary components found in lunged salamanders. The partial septum of plethodontids would be unable to maintain separation of blood from the left and right atria. The atrial morphology of convergently lungless salamander *Onychodactylus japonicus* is similar in that of plethodontids and has presumably evolved independently. Finally, genes associated with atrial septum development are downregulated in plethodontid salamander embryos relative to lunged salamanders, suggesting a mechanism by which atrial septum reduction occurs in plethodontids.

5.3 Materials and methods

5.3.1 Animal collection and husbandry

Ambystoma mexicanum (axolotl) embryos were procured from the Ambystoma Genetic Stock Center, University of Kentucky, and maintained in 20% Holtfreter solution at 17°C. All plethodontid embryos were collected under Massachusetts DFW permits and local permits, where applicable. DFW permit numbers: 181.10SCRA (2010), 080.11SCRA (2011), 080.11SCRA (2012), 027.13SCRA (2013), 083.14SCRA (2014), and 022.15SCRA (2015). Locality data is provided in Appendix A (Table A.1). *Hemidactylium scutatum* embryos were collected at Cape Cod National Seashore (permit CACO-2012-SCI-0008) and maintained in the lab on filter paper moistened with 0.1x MMR with 100 µg/ml gentamicin at 15°C. At hatching, they were transferred to 20% Holtfreter solution and fed Artemia spp. until fixation with Bouin fixative. *Plethodon cinereus* embryos were field collected and maintained in the lab similarly to *H. scutatum*, except they were fixed at hatching or before in neutral-buffered formalin (NBF) or Bouin fixative. *Desmognathus fuscus* embryos were field collected under the aforementioned Massachusetts DFW permits and raised similarly to *H. scutatum* embryos and larvae until fixation.

Animals were staged according to (Bordzilovskaya et al., 1989; Kerney, 2011; Nye et al., 2003).

5.3.2 Museum specimens and gross dissection

Specimens examined are listed in Table 5.1. Dissections were carefully performed with a scalpel, fine scissors and #5 forceps in order to not tear the delicate atrial septum. Congealed

Table 5.1: List of specimens and stages examined.

Species	Stage	Specimen No.	Techniques[1]
Lunged species			
Ambystoma mexicanum	44, 52, 57, adult	N/A	H, µ, D
Amphiuma tridactylium	Adult	MCZ A-115797	D
Cynops ensicauda	Adult	MCZ 26607	D
Lungless species			
Desmognathus fuscus	Larvae	N/A2	H
Desmognathus quadramaculatus	Adult	MCZ 117899	D
Hemidactylium scutatum	Embryo, late larva, adult	N/A2	H, µ
Plethodon glutinosus	Adult	MCZ Z-27269	D
Plethodon cinereus	Embryo, hatchling, adult	N/A2	H, µ
Pseudotriton montanus	Adult	MCZ 5739	D
Onychodactylus japonicus	Late larva	MCZ A-119652	µ

[1] D, dissection; H, histology; µ, X-ray µ-CT.

blood was removed from the atria manually and by rinsing with 70% ethanol.

5.3.3 Micro-computed tomography

For X-ray µ-CT, specimens were fixed in 10% NBF, washed with 1x phosphate-buffered saline and then dehydrated in 70% ethanol. Specimens were soaked in 2.5% phosphomolybdic acid (Sigma, St. Louis, MO) in 70% ethanol for 1–2 wk, washed briefly with 70% ethanol, and embedded in 0.8% agarose to stabilize the specimen and prevent desiccation. Samples were mounted as described in Metscher (2011). 360-degree µ-CT scans at 0.1 degree increments were run at 80 kV, 70 µA, and 800 ms exposure time at pixel dimensions of 5.3–9.9 µm on a SkyScan 1173 benchtop scanner (Bruker µ-CT, Kontich, Belgium). Reconstruction was performed with the native reconstruction software using ring artifact reduction and

post-alignment corrections where appropriate (NRecon, Bruker µ-CT, Kontich, Belgium). Reconstructed TIFF files were visualized in Amira 6.0 (FEI, Hillsboro, OR) and segmented using the segmentation editor. Separate label fields were utilized for the septal and endocast components so their opacity could be adjusted independently.

5.3.4 Histological reconstruction

Specimens were fixed in Bouin solution for 24–48 hr prior to washing, dehydrating and embedding in paraplast (McCormick Scientific, Wetzlar, Germany). Serial sections (7 or 8 µm) were stained using the Mallory trichrome method (Presnell et al., 1997) with the following modifications: a 10-min stain in Mayer hematoxylin followed by a 10-min wash in running dH_2O preceded 30 sec in 1% acid fuchsin, several rinses with dH_2O, 5 min in 1% phosphomolybdic acid, then 3 min in a modified Mallory II (1% orange G, 1% aniline blue, and 2% oxalic acid) followed by dehydration and mounting. Sections were imaged using a Leica DMRE microscope (Wetzlar, Germany) equipped with a QImaging Retiga 2000r camera and a QImaging RGB slider (Surrey, Canada) and Volocity 6.0 software (PerkinElmer, Waltham, MA). Monochrome images were imported into Amira and aligned using the Align Slices module. Endocasts of cardiac chambers were made using segmentation tools in Amira, as described above. Reconstruction was only performed on specimens with nearly complete or complete serial sections. In the event that a section was damaged or distorted, labels were interpolated from the flanking slides.

5.3.5 Gene expression analysis

Transcriptome sequencing is described in Chapter 3 and Appendix B.

5.4 Results

5.4.1 Lunged salamanders develop a two-part atrial septum

In all lunged salamanders we examined, the atrial septum comprises two perpendicularly oriented interconnected endothelial tissues, AS1 and AS2. AS1 is a thin sheet oriented anteroposteriorly; it stretches from the center of the atrioventricular aperture (A-V) to the anterior wall of the atrium. It is emarginated posterodorsally, forming an arch through which blood flows from the sinoatrial aperture (S-A) into the right atrium (RA; Fig. 5.2E, F). AS2 is oriented dorsoventrally; it attaches to the left side of AS1 and stretches to the left wall of the atrium (Fig. 5.2E). Removal of the ventral atrial wall reveals AS2, and further excision of AS2 reveals the S-A posterodorsal to it (Fig. 5.2C, D). The pulmonary vein (PV) wraps around the sinus venosus (SV) dorsally to enter the left atrium (LA) at its anterodorsal surface, anteroventral to AS2 (Fig. 5.2B). A medial orifice in AS2 allows blood to pass through the A-V valve. This orifice is small in adult *Cynops ensicauda* but much larger in juvenile *Ambystoma mexicanum* (Fig. 5.2C, E).

The AS1 in *A. mexicanum* develops late and slowly. At the post-embryonic stage 52 (2-3 weeks post-hatching), AS1 is extending ventrally but has reached only halfway across the atrium. By stage 57 (sub adult), the separation of left and right atria is nearly complete (Figs. 5.3, 5.4); AS1 is not yet complete, but the S-A is fully covered by AS2 (Figs. 5.2,

5.3, 5.4). Adult *A. mexicanum* have a complete AS1 as determined by gross dissection (not shown); blood from the PV passes ventral to AS2. Prior to complete development of AS1, oxygenated blood from the PV could potentially mix with deoxygenated blood from the RA before passing through the A-V. However, when viewed from the anterior, the AS2 completely covers the S-A and thus would likely direct blood from the SV into the RA and maintain oxygenated blood from the PV in the LA (Fig. 5.2). Adult *A. mexicanum* possess a complete AS1, thus enabling complete separation of the two blood streams.

To evaluate the condition of the atrial septum in an appropriate outgroup to plethodontid salamanders, we dissected a specimen of *Amphiuma tridactylium*. Molecular phylogenies consistently place the family Amphiumidae, to which this species belongs, as the sister group to Plethodontidae (Pyron and Wiens, 2011). Our dissection agrees with previously published findings (Johansen, 1963). *Amphiuma tridactylium* has a large LA separated from the RA by a robust, imperforate atrial septal complex (data not shown).

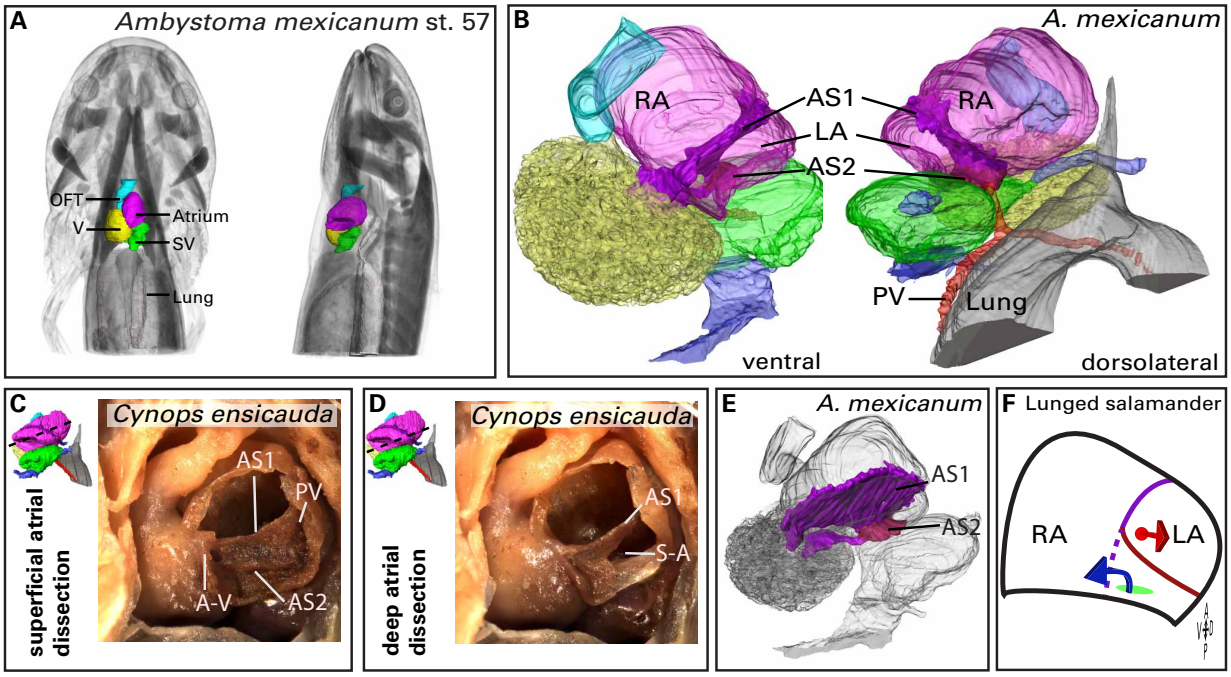
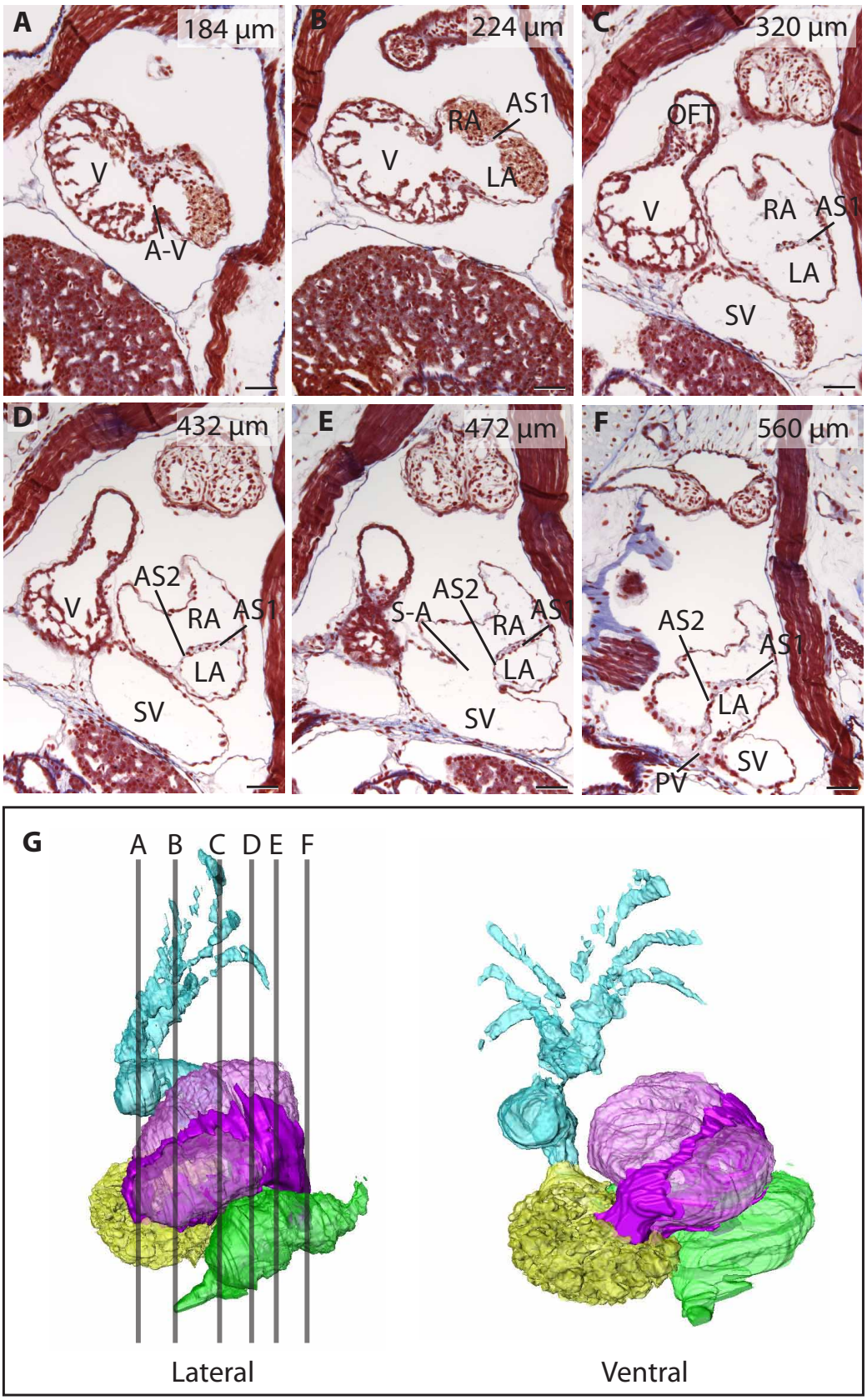


Figure 5.2: Atrial morphology in lunged salamanders. (A) Contrast μ-CT of a stage 57 (juvenile) axolotl, *Ambystoma mexicanum*, depicted in ventral (left) and lateral perspectives. (B) Reconstruction of the *A. mexicanum* heart at stage 57 generated from 8-μm histological sections. (C) Superficial dissection of the atrium in the swordtail newt, *Cynops ensicauda*. The left atrial chamber is bordered dorsally by AS2 and anteriorly by AS1. Ventral view with the ventral atrial wall removed. Dashed line in the colored schematic illustrates the approximate angle and depth of dissection. (D) Deep dissection of the specimen in (C). AS2 has been partially removed, revealing the S-A dorsal to it. (E) Components of the atrial septum are highlighted in the st. 57 *A. mexicanum* heart. (F) Schematic frontal section of a lunged salamander atrium, showing the location of the pulmonary vein (red), S-A (green) and associated blood streams.

Figure 5.3: Histological sections through the heart of the lunged salamander *Ambystoma mexicanum* at stage 52. (A–F) The distance of each frontal section from the ventral surface of the heart is designated in the upper right. Anterior is to the top. (G) 3-D reconstruction of the heart in lateral and ventral perspectives. The approximate position of each section is indicated by a vertical line. Abbreviations as in Figure 5.1. Scale bars, 100 µm

Figure 5.3, continued



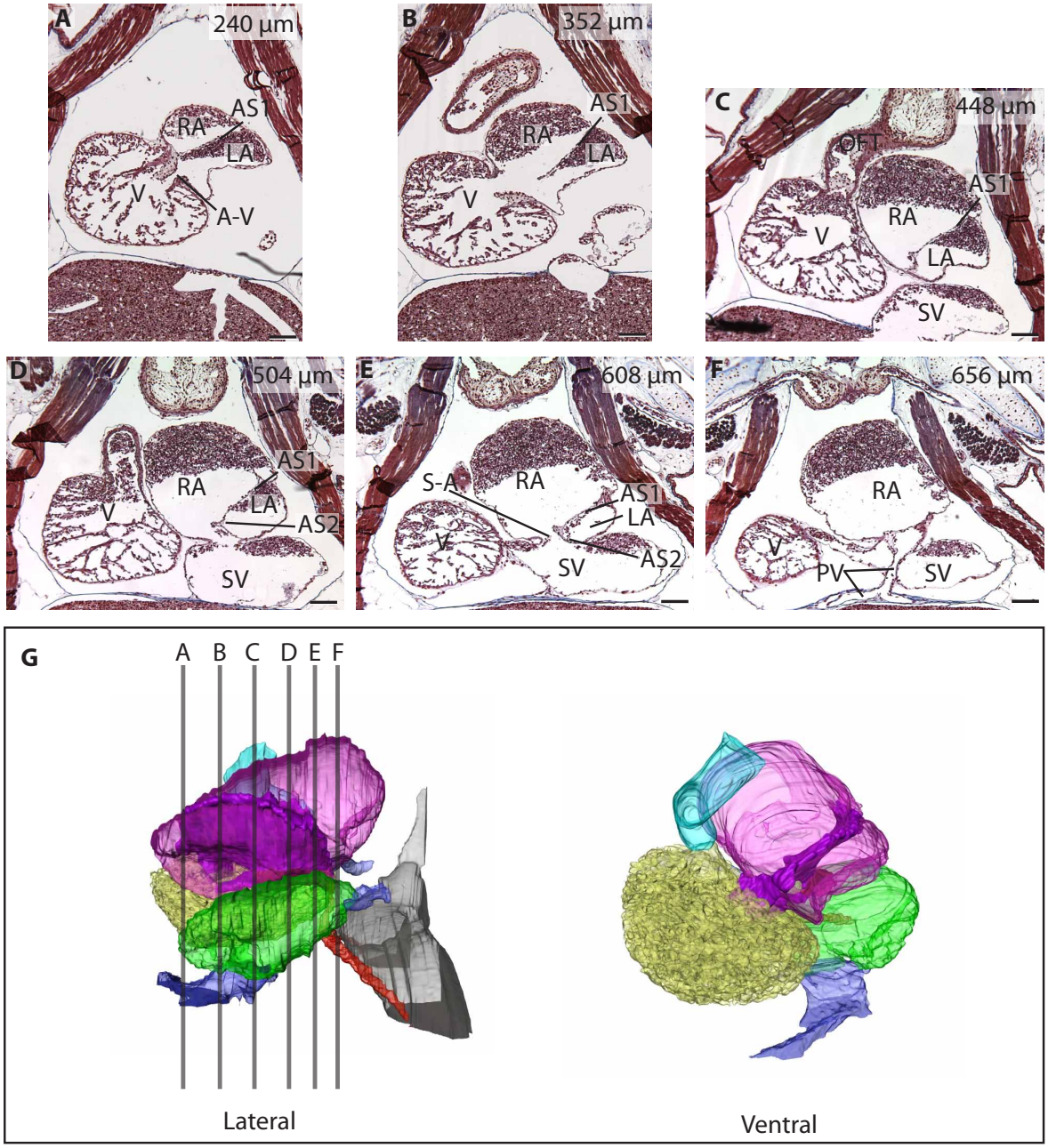


Figure 5.4: Histological sections through the heart of the lunged salamander *Ambystoma mexicanum* at stage 57. (A–F) The distance of each frontal section from the ventral surface of the heart is designated in the upper right. Anterior is to the top. (G) 3-D reconstruction of the heart in lateral and ventral perspectives. The approximate position of each section is indicated by a vertical line. Abbreviations as in Figure 5.1. Scale bars, 200 µm.

5.4.2 Plethodontid salamanders lack AS2

Several species of plethodontid salamanders were dissected, sectioned histologically, or subjected to contrast-stained μ-CT imaging (Table 5.1). The atrial chamber of plethodontids broadly resembles that of salamanders with lungs: the AS1 is present and oriented anteroposteriorly, it lacks perforations, and its dorsal border is emarginate (Figs. 5.5, 5.6).

Plethodontids, however, lack AS2 (Fig. 5.5B, C, E). Since AS2 essentially forms the dorsal wall of the LA in lunged salamanders, the absence of this structure results in the complete confluence of left and right atria. Plethodontids also lack pulmonary veins. All inflow to the heart is through the left and right common cardinal veins (ducts of Cuvier) and the posterior cardinal vein, which empty into the SV. Blood enters the atrium solely through the S-A from which it would presumably flow to either side of AS1 before passing through the A-V during atrial systole.

Previous authors disagree on the position of the S-A in plethodontids (Fig. 5.1). Our gross dissection reveals that the S-A is to the left side of AS1 when the atrium is viewed ventrally (Fig. 5.5C). However, the LA in lunged salamanders is bordered by the AS1 dextrally and the AS2 dorsally. Thus, if the AS2 were intact, then the S-A would fall within the right atrium (Figs. 5.7, 5.9).

As a result of these derived features, plethodontids have incomplete atrial septation. AS1 is present and morphologically similar to AS1 in lunged salamanders in that it occupies the same position in the heart and is not fenestrated. AS1 may function as a sinoatrial valve.

Either of two ontogenetic scenarios could account for the observed morphology of adult plethodontid hearts. First, the atrial septum may develop fully and subsequently regress.

Alternatively, the atrial septum may never develop fully. Our data support the latter scenario. By reconstructing hearts from plethodontids at stages ranging from hatchling to late larva to adult, we find no sign of a complete atrial septum at any point in development (Fig. 5.5).

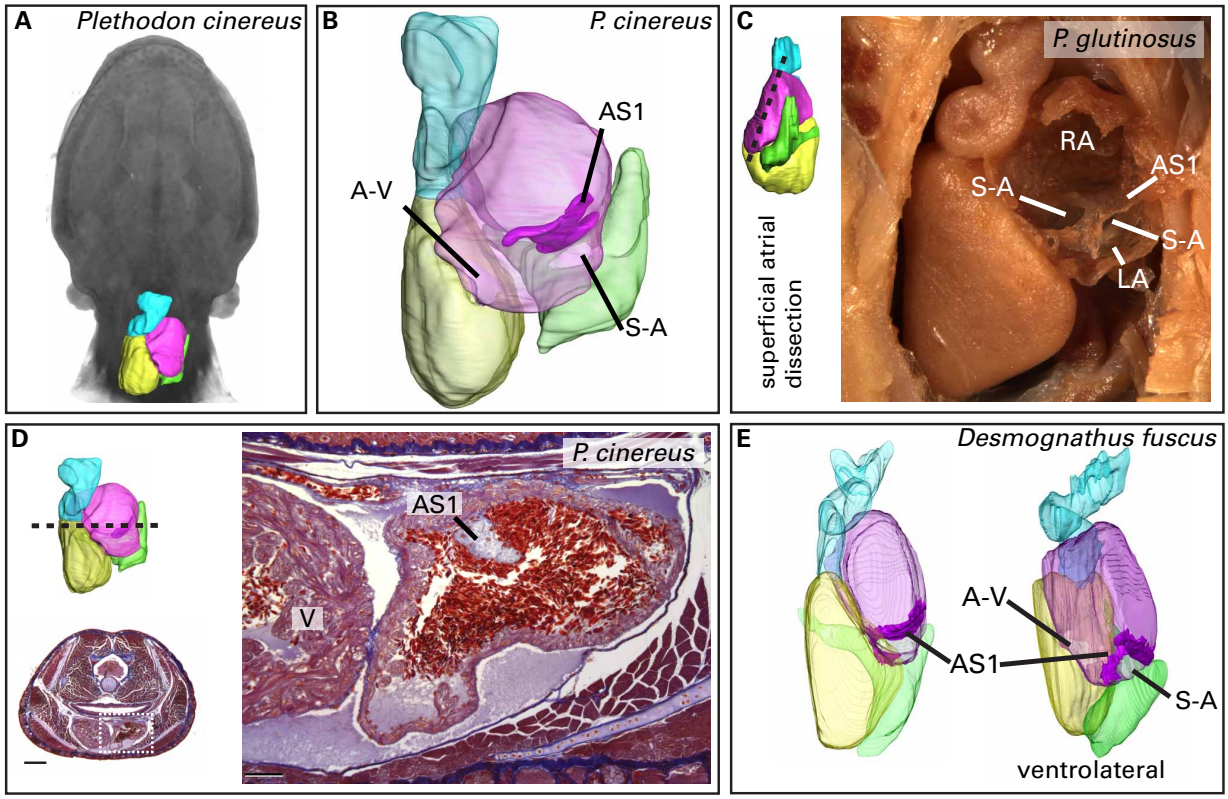
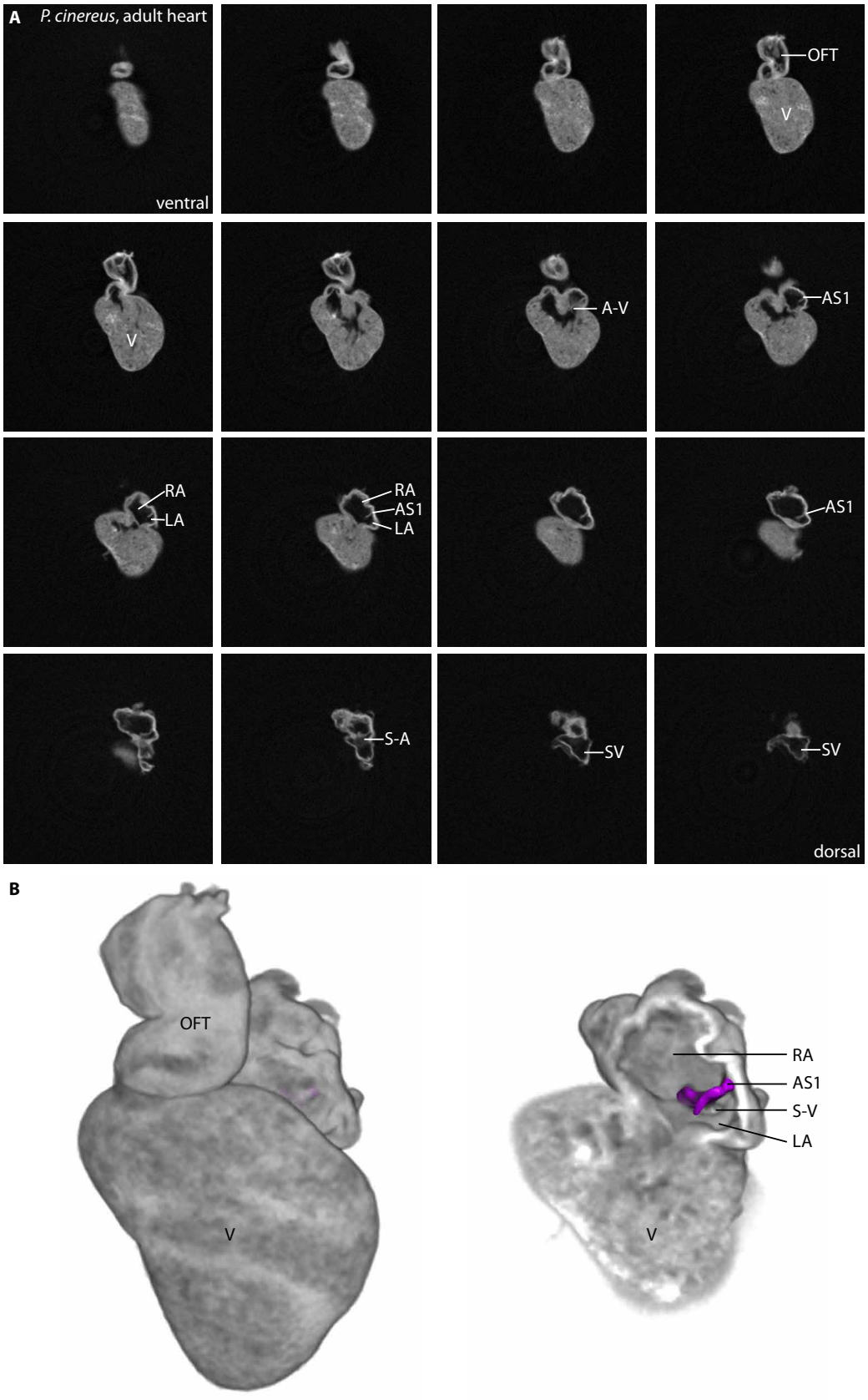


Figure 5.5: Atrial morphology in plethodontid salamanders. (A) Contrast µ-CT of a *Plethodon cinereus* embryo at hatching (st. 24; Kerney 2011). Heart chambers have been segmented and colored. (B) Stage-24 *P. cinereus* heart magnified to show the atrial septum. Only AS1 is present. (C) Atrium of an adult *Plethodon glutinosus*. The plane and angle of dissection are indicated by the colored reconstruction. Ventral view. (D) Transverse section through the atrial septum of an adult *P. cinereus*. The approximate plane is indicated in the schematic on the left. Scale bars 500 µm and 100 µm for the left and right images, respectively. (E) Reconstruction of the heart of a *Desmognathus fuscus* larva in ventral and left ventrolateral views.

Figure 5.6: Micro-CT sections through an adult heart of *Plethodon cinereus*. (A) Frontal sections at 57 µm increments from ventral to dorsal illustrate the components of the *P. cinereus* heart. The sections are arranged from left to right, top to bottom. (B) Volume renderings of the heart from a ventral perspective. Cutaway (right) reveals AS1, segmented in purple. Abbreviations as in Fig. 5.1.

Figure 5.6, continued



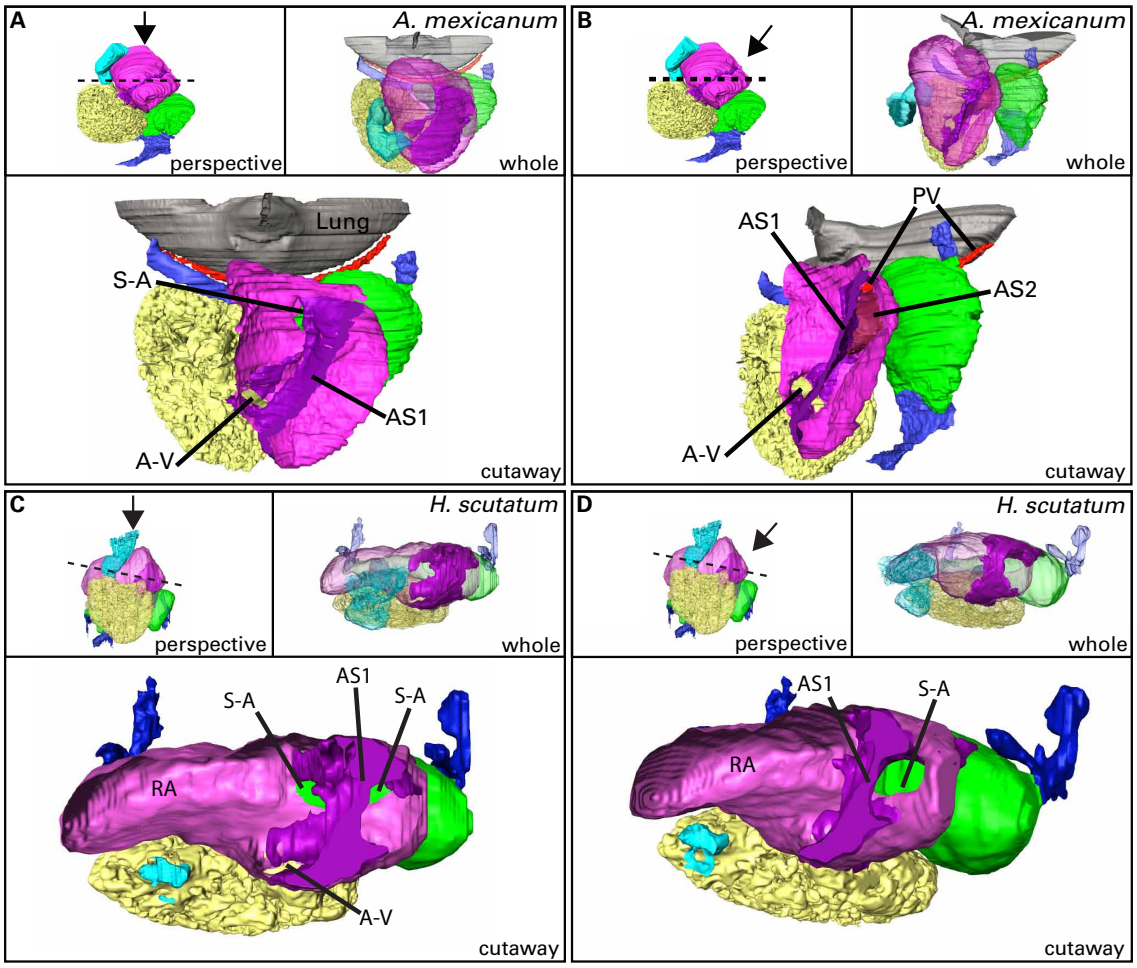


Figure 5.7: Comparative morphology of the posterior atrial wall. (A-D) Insets reflect the angle of cutaway and viewing, as well as a whole transparent heart viewed from the same angle as the cutaway. (A) Cutaway into the atria of a stage-57 *Ambystoma mexicanum* (lunged) heart from an anterior perspective. The S-A is visible. (B) The heart in (A) viewed from a left lateral angle reveals AS2 and the pulmonary vein inlet. The S-A is not visible because it is covered by AS2. (C) Cutaway into the atria of a *Hemidactylium scutatum* (lungless) heart from an anterior perspective. The S-A is visible on both sides of AS1. This heart is reconstructed from a 16.5 mm (total length) larva. (D) The heart in (C) viewed from a left lateral angle reveals the lack of AS2 and, consequently, the ability to see the S-A through the presumptive LA.

5.4.3 *Onychodactylus japonicus* lacks AS2

The heart from a single larval specimen of *Onychodactylus japonicus* (6.5 cm total length, 3.5 cm snout-vent length) was examined using contrast μ-CT (Fig. 5.8). AS1 is clearly visible as a region within the atria devoid of blood (Fig. 5.8). In contrast to plethodontids, where a dorsal emargination of AS1 would cause blood to flow first dorsally and then posteriorly towards the A-V, in *O. japonicus* AS1 is emarginated posteroventrally. Blood likely flows from the S-A posteroventrally around the free margin of AS1 and then through the A-V. The lack of dorsal emargination and few points of attachment to the atrial walls indicate that AS1 functions mainly as an S-A valve. There is no AS2, and the left atrium is extremely small.

5.4.4 Atrial septum patterning genes are downregulated in *Plethodon cinereus* embryos

Atrial septum development in mammals depends on Shh signaling between the developing lung and second heart field mesoderm, combined with the mesodermal expression of Tbx5 (Hoffmann et al., 2014). We sequenced the transcriptome of the pulmonary endoderm and underlying second heart field mesoderm in the lungless salamander *P. cinereus* and the lunged species *A. mexicanum* (Chapter 3). Gene expression in stage 19 *P. cinereus* embryos and stage 40 *A. mexicanum* embryos was compared, because these stages are developmentally comparable and comparable to the stage when atrial septal progenitors are specified in mammals. Known atrial septum-patterning genes are significantly downregulated in lungless *P. cinereus* embryos relative to *A. mexicanum* (Fig. 5.10). This includes a significant

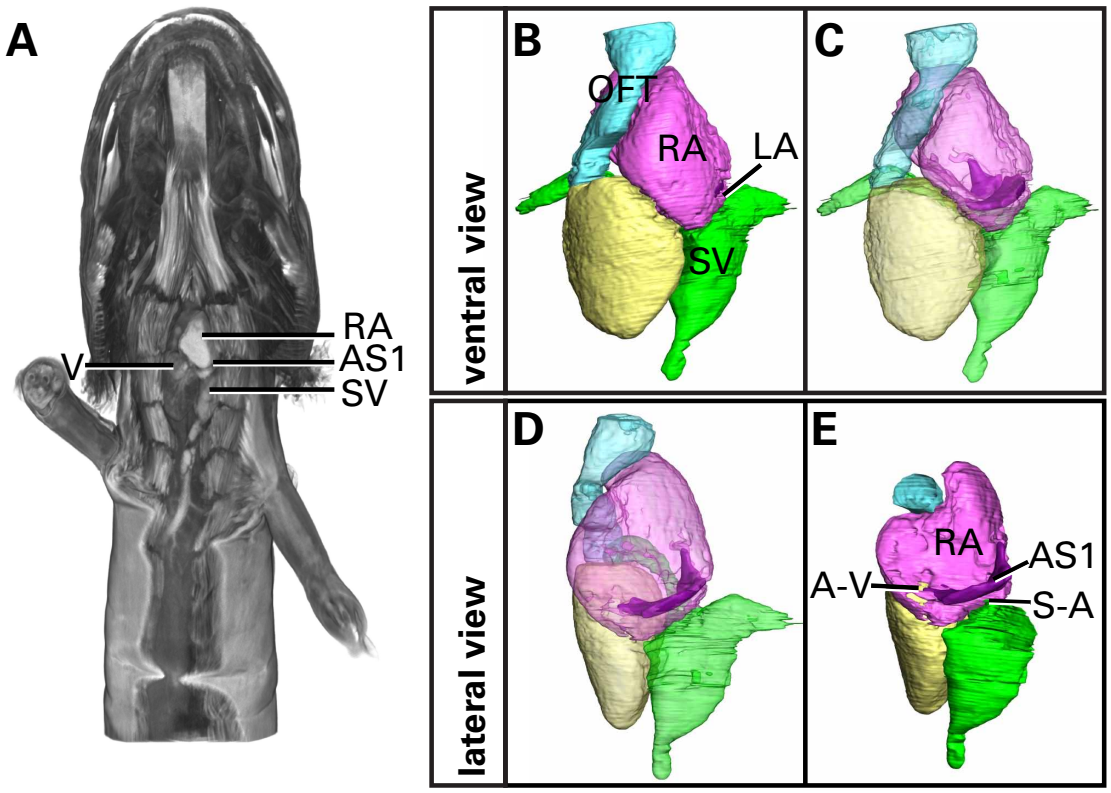


Figure 5.8: Atrial morphology in the convergently lungless species *Onychodactylus japonicus*. (A) Frontal orthoslice through the contrast-stained µ-CT scan at the level of the heart. Atrial chambers are radiopaque due to the dense mass of blood contained within them. The AS1 is visible as a less-opaque structure positioned far towards the left edge of the atrium. Ventral view of the segmented heart rendered opaque (B) and transparent (C) to view the atrial chambers and the AS1. Lateral views of the *O. japonicus* heart rendered transparent (D) and opaque with a cutaway (E). Cutaway view reveals the S-A and A-V.

downregulation of Gli1 and Tbx5 and a marginally significant downregulation of Foxf2 (Fig. 5.10). Foxf1a expression was not significantly different between species.

5.5 Discussion

By examining cardiac morphology in a diverse sample of plethodontid salamanders, lunged salamanders, and the convergently lungless species *Onychodactylus japonicus*, we demon-

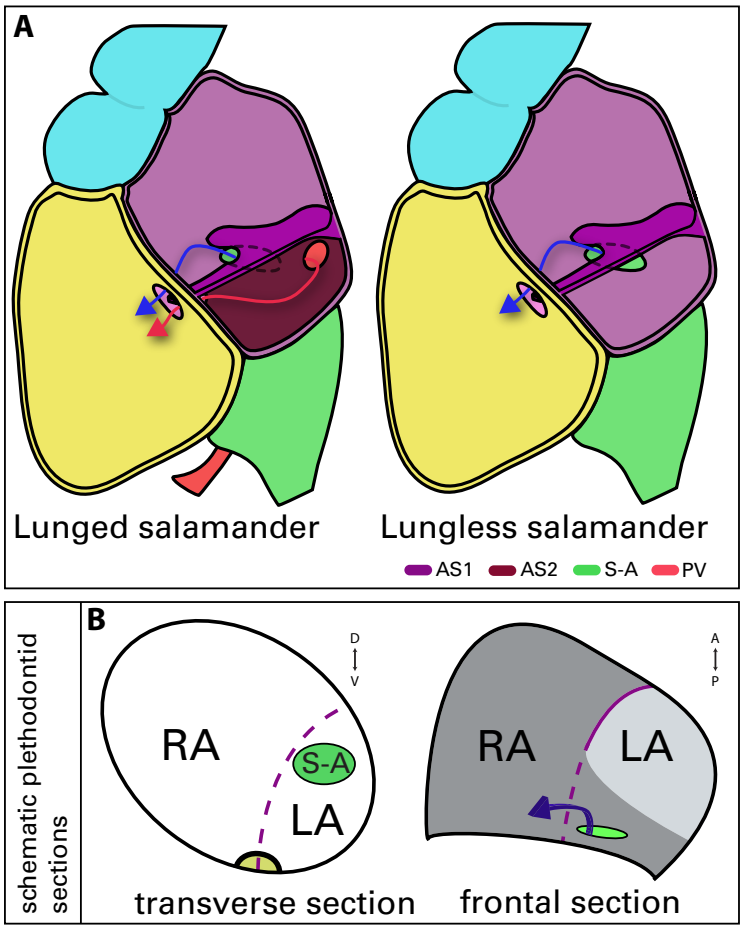


Figure 5.9: Atrial morphology in salamanders. (A) Lunged salamanders have two sources of blood to the atrium: the SV (via S-A) and the pulmonary vein. Blood from S-A passes dorsal to AS2 and then posterodorsal through the emarginated portion of AS1 and through A-V (blue arrow). Blood from the pulmonary vein passes ventral to AS2 before flowing through A-V (red arrow). Lungless salamanders lack pulmonary veins and AS2, so blood from SV could presumably flow freely between chambers. (B) Schematic sections of plethodontid atria. When viewed from the anterior, the S-A appears displaced to the left side of AS1, a condition that some previous authors took to indicate that S-A had shifted to the left atrium. When AS1 is interpreted as just a component of the bipartite atrial septal complex, it is clear that the RA (shaded region) extends posterodorsal to the left atrium. The lack of AS2 has caused the assignment of this region to the LA. Comparison to lunged outgroups (e.g., Fig. 5.2F) is necessary for making this assessment.

strate that two lineages of lungless salamanders have independently evolved similar atrial morphology, including the loss of a significant component of the atrial septum (AS). Our work is the first description of the atrial morphology of *O. japonicus*. Previous studies of

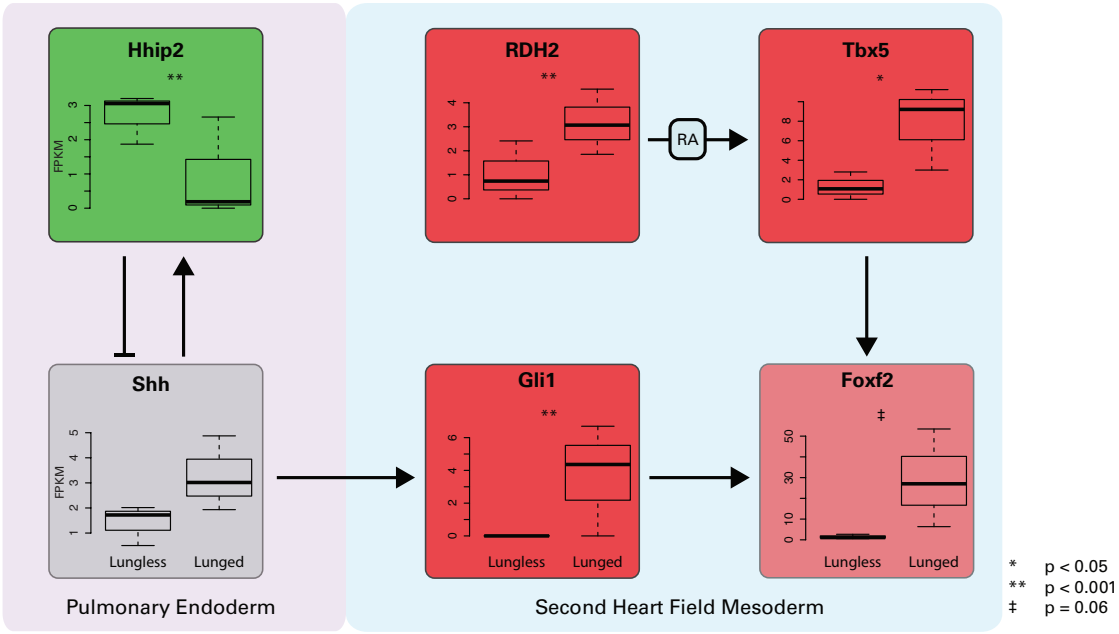


Figure 5.10: Candidate genes involved atrial septum development are downregulated in *Plethodon cinereus* embryos relative to *Ambystoma mexicanum*. The expression of several genes that have been previously implicated in atrial septum development is significantly different lungless *P. cinereus* embryos and lunged *A. mexicanum* embryos. Transcriptomes of pulmonary endoderm and underlying second heart field mesoderm were sequenced from stage 19 *P. cinereus* and stage 40 *A. mexicanum*. Expression values are normalized to fragments per kilobase of transcript per million mapped reads (FPKM).

plethodontid heart morphology left several questions unresolved. In particular, the position of the sinoatrial aperture (S-A) and the presence or absence of the atrial septum has been debated for over 100 years (Fig. 5.1). Recent advances in computer-assisted, 3-dimensional reconstruction from histological or tomographic datasets facilitates study of the complex hearts of salamanders.

Given complete atrial septation in amphiumid, ambystomatid and salamandrid salamanders (present study; Johansen 1963), combined with published data on septation in frogs and caecilians (de Bakker et al., 2015; Noble, 1925), we infer the basal condition for salamanders to be fully septate. Hence, the morphology of lungless salamanders represents an

evolutionary loss.

5.5.1 Why lose the septum?

Although some species of plethodontids retain the pulmonary arch, adult lungless amphibians lack pulmonary veins (McMullen, 1938). Consequently, all blood flowing into the heart originates from the left and right common cardinal veins (ducts of Cuvier) and the posterior cardinal vein via the sinus venosus (Barrows, 1900; McMullen, 1938; Seelye, 1906). Since this blood is from a single mixed source, lungless amphibians do not require separation of blood in the heart. In fact, the presence of an AS might be disadvantageous in this situation, as it could lead to pressure differentials between atrial chambers.

In humans, anomalies of pulmonary venous return are developmental defects, which result in anastomosis of the pulmonary veins to the right inflow tract. These misconnections may cause extreme pressure differentials between RA and LA (El-Said et al., 1972; Taussig, 1960); patients display high mortality unless they possess compensatory defects in the AS or undergo surgical intervention. The physiological disadvantage of left-right pressure differential resulting from the lack of pulmonary veins might have selected for atrial septum reduction during the evolution of lungless salamanders. If so, then lungless salamanders have evolved a condition similar to a human atrial septal defect in response to the unique configuration of their cardiac inflow. Salamanders may offer a beneficial model for the study of atrial septal defects or anomalous pulmonary venous return.

5.5.2 Atrial septum controversy

Atrial morphology in plethodontids has been contentious, with no agreement regarding even basic features such as the presence or absence of a septum, the position of the S-A and the degree of septal fenestration (Noble, 1925; Putnam and Kelly, 1978; Fig. 5.1). In the first published work on plethodontid hearts, Hopkins (1896) declared that plethodontids have a partial atrial septum and that the S-A has shifted from the RA, where it is located in all other amphibians, to the LA. Subsequent authors disagree regarding the position of the S-A and the degree of septation (Fig. 5.1; Bruner, 1900; Noble, 1925; Putnam and Kelly, 1978). The lack of agreement could be due to the different species that the authors examined or the different analytic techniques employed. Alternatively, it may be due to the failure of all studies but one (Putnam and Kelly, 1978) to directly compare plethodontids to lunged outgroups.

Hopkins (1896) asserted that the S-A lies within the left atrium, but he failed to recognize the two-part nature of the atrial septum, which is apparent when one compares plethodontids to lunged outgroups (Figs. 5.1, 5.9). The S-A is left of AS1, but in lunged salamanders the RA extends dorsal to the LA and the loss of AS2 makes it appear that the S-A is in the LA. Putnam and Kelly (1978) were largely correct in their interpretation of atrial morphology, but they failed to recognize the artifactual nature of Hopkins' account, which led them to incorrectly depict the position of the S-A relative to AS1. Bruner (1900) did observe tissue within the atrium, but he interpreted this tissue to represent the sinoatrial valve. Cords (1923) supported Bruner's view. Our data are consistent with a functional role for the partial septum as a sinoatrial valve. However, we did not observe a separate sinoatrial valve

in lunged salamanders. The AS in lunged salamanders is a complex structure of several parts, which may function both as a sinoatrial valve and as an interatrial partition (Putnam and Kelly, 1978). Functional studies would be required to confirm this.

Several early authors declared that salamanders have a highly fenestrated atrial septum (reviewed in Noble 1925). However, Noble (1925) did not find this to be the case in salamanders with mostly pulmonary respiration. He suggested instead that fenestration is a feature of species that predominately use non-pulmonary respiration, such as *Rhyacotriton olympicus* and *Cryptobranchus alleganiensis* (Noble, 1925; Putnam and Parkerson Jr, 1985). Interestingly, *R. olympicus*, which possess vestigial lungs, retain a complete but fenestrated AS. A number of caecilians with reduced reliance on pulmonary respiration also show partial reduction or fenestration of the atrial septum (Wilkinson and Nussbaum, 1997). We did not observe AS fenestration in lunged or lungless specimens. AS reduction and AS fenestration likely function similarly by accounting for reduced or absent pulmonary return with appropriate shunting of blood between atria (Lawson, 1966).

5.5.3 Convergent reduction of the atrial septum

Our data support at least two independent evolutionary losses of AS2 in salamanders, one in the completely lungless family Plethodontidae and a second in the completely lungless genus *Onychodactylus*. *Salamandrina perspicillata*, a salamandrid species with reduced lungs, also may have lost AS2, as illustrated in one published drawing: Cords (1923) depicts a septum covering the S-A when viewed anteriorly. This septum is likely the AS1, which has been reflected posteriorly. If our interpretation is correct, then the AS2 has been lost

independently at least three times in salamanders, indicating the functional importance of septum reduction during the evolution of lunglessness and/or the existence of a shared developmental mechanism that links lung loss/reduction and septal reduction.

Cords (1923) closes her manuscript with a prerogative for further research on the atrial septum in salamanders. She asks whether the AS develops fully and subsequently regresses in lungless or lung-reduced salamanders, or if it instead fails to develop fully. Our data, drawn from multiple ontogenetic stages of plethodontid salamanders, supports the latter proposition.

5.5.4 Atrial septum development

The AS complex is slow to develop in *A. mexicanum*, a lunged species. At hatching, AS2 is present but AS1 is small. During larval stages, AS1 grows to fully partition the atrium (Figs. 5.3, 5.4). In lungless plethodontids, AS1 is present at hatching but shows no sign of continued development during larval or juvenile stages. In the frog *Xenopus laevis*, the atrial septum is one of the last heart structures to develop (Mohun et al., 2000). Previous authors note the origin of the atrial septum from the dorsal wall of the atrium; the AS then extends towards the atrioventricular opening, eventually separating the two atria (Mohun et al., 2000; Simons, 1957). Our results on AS growth are in concordance with these findings (Figs. 5.3, 5.4), but we did not examine the earliest stages of *A. mexicanum* to determine the precise origin of the AS.

Recent developmental-genetic studies highlight the importance of signaling interactions and genetic pleiotropy between the lungs and heart during organogenesis. From the char-

acterization of human birth defects, there is increasing awareness that many critical heart development genes play pleiotropic roles in the development of other organs (reviewed in Li et al., 1997). For instance, Holt-Oram syndrome, which is due to a mutation in the T-box transcription factor Tbx5, is characterized by atrial septal defects, upper limb malformation and occasional right lung agenesis (Basson et al., 1997; Li et al., 1997; Tseng et al., 2007). Tbx5 is known to regulate Fgf10 expression in the developing limb (Agarwal et al., 2003) and is important for lung branching (Cebra-Thomas et al., 2003). Congenital ASDs also include cor triloculare biventriculare, or complete absence of the atrial septum (Sangam et al., 2011), which, based on associated malformations that include polydactyly and cleft palate, is likely due to a defect in the Shh pathway (personal observation). Ellis-Van Creveld syndrome, which is caused by a mutation in a positive mediator of Shh signaling, results in ASDs (Baujat and Le Merrer, 2007). Mouse mutants for the important pulmonary gene Wnt2 also lack atrial septa (Tian et al., 2010). Finally, crucial mediators of both lung and AS development include members of the Bmp pathway such as Alk3, Bmp4 and Osr1 (Briggs et al., 2013; Jiao et al., 2003; Zhou et al., 2015). Osr1 mutations result in ASDs and lung phenotypes (Wang et al., 2005; Xie et al., 2012), and knockdown of Osr1 and Osr2 results in lung loss (Rankin et al., 2012). All these Bmp signaling-associated genes play essential roles in both heart and lung morphogenesis. Interestingly, several defects associated with mutations of these genes have orthologous phenotypes to those displayed in lungless salamander hearts (viz., left atrium hypoplasia, atrial septal defects and lack of pulmonary circulation). In a rare instance of human lung loss, an ASD is also present (Devi and More, 1966).

Research on heart septation in mice implicates pulmonary and pharyngeal endoderm in producing morphogens that are directly responsible for development of the atrial septum and

the outflow tract, respectively (Goddeeris et al., 2007, 2008; Hoffmann et al., 2009, 2014). Shh secreted from the tracheal pulmonary endoderm to the second heart field mesoderm potentiates a population of cells that migrate into the heart and form the atrial septum. Shh is required to specify these cells but is not required for their migration or for septal morphogenesis. Subsequent work pinpoints activation of Foxf genes by Gli1 and Tbx5 as critical for pulmonary-induced septal morphogenesis (Hoffmann et al., 2014). The above results imply a direct molecular link between lung and heart morphogenesis during the evolution of cardio-pulmonary respiration.

We found significantly decreased expression of Gli1 and Tbx5 between lunged and lungless embryos in the second heart field mesoderm, suggesting a mechanism by which atrial septum reduction occurs in lungless salamanders (Fig. 5.10). Lung loss may further decrease levels of Shh signaling later in plethodontid embryonic development. Provided that the mechanism of atrial septum development is conserved between amphibians and mammals, these data help confirm a direct link between lung loss and atrial septum reduction.

5.5.5 Conclusion

Convergent evolution is a fascinating yet often vexing phenomenon in systematic and morphological research. There are several developmental mechanisms by which convergent morphologies might arise (Manceau et al., 2010; Stern, 2013). These include independent mutations of a single gene, mutations of multiple genes within a pathway, selection on standing genetic variation, or introgression of loci, among other mechanisms (Stern, 2013). Certain convergent traits may also emerge as pleiotropic effects of selection on molecularly or de-

velopmentally linked phenotypes. The heart and lungs function as an integrated system; functional mismatch between them would have disastrous consequences for organismal physiology. We propose that the molecular crosstalk between the developing lungs and heart may underlie convergent reduction of the atrial septum across species of lungless salamanders. Our study supports a process by which pleiotropic interactions between developing heart and lung pattern the co-evolution of the functionally constrained cardiopulmonary system.

5.6 References

Agarwal, P., Wylie, J.N., Galceran, J., Arkhitko, O., Li, C., Deng, C., Grosschedl, R., and Bruneau, B.G. (2003). Tbx5 is essential for forelimb bud initiation following patterning of the limb field in the mouse embryo. Development 130, 623–633.

de Bakker, D.M., Wilkinson, M., and Jensen, B. (2015). Extreme variation in the atrial septation of caecilians (Amphibia: Gymnophiona). J. Anat. 226, 1–12.

Barrows, A.I. (1900). Respiration of *Desmognathus*. Anat. Anz. 18, 461–464.

Basson, C.T., Bachinsky, D.R., Lin, R.C., Levi, T., Elkins, J.A., Soults, J., Grayzel, D., Kroumpouzou, E., Traill, T.A., Leblanc-Straceski, J., et al. (1997). Mutations in human cause limb and cardiac malformation in Holt-Oram syndrome. Nat. Genet. 15, 30–35.

Baujat, G., and Le Merrer, M. (2007). Ellis-Van Creveld syndrome. Orphanet J. Rare Dis. 2, 27.

Bickford, D., Iskandar, D., and Barlian, A. (2008). A lungless frog discovered on Borneo. Curr. Biol. 18, 374–375.

Bjornard, K., Riehle-Colarusso, T., Gilboa, S.M., and Correa, A. (2013). Patterns in the prevalence of congenital heart defects, metropolitan Atlanta, 1978 to 2005. Birth Defects Res. Part A - Clin. Mol. Teratol. 97, 87–94.

Bordzilovskaya, N., Dettlaff, T., Duhon, S., and Malacinski, G. (1989). Developmental-stage series of axolotl embryos. In Developmental Biology of the Axolotl, J. Armstrong, and G. Malacinski, eds. (Oxford: Oxford University Press), pp. 201–219.

Briggs, L.E., Phelps, A.L., Brown, E., Kakarla, J., Anderson, R.H., van den Hoff, M.J.B.,

and Wessels, A. (2013). Expression of the BMP receptor Alk3 in the second heart field is essential for development of the dorsal mesenchymal protrusion and atrioventricular septation. Circ. Res. 112, 1420–1432.

Bruner, H.L. (1900). On the heart of lungless salamanders. J. Morphol. 16, 323–336.

Cebra-Thomas, J.A., Bromer, J., Gardner, R., Lam, G.K., Sheipe, H., and Gilbert, S.F. (2003). T-box gene products are required for mesenchymal induction of epithelial branching in the embryonic mouse lung. Dev. Dyn. 226, 82–90.

Cords, E. (1923). Das herz eines lungenlosen salamanders (*Salamandrina perspicillata*). Anat. Anz. 57, 205–213.

Devi, B., and More, J.R.S. (1966). Total tracheopulmonary agenesis associated with asplenia, agenesis of umbilical artery and other anomalies. Acta Paediatr. Scand. 55, 107–116.

El-Said, G., Mullins, C.E., and Mcnamara, D.G. (1972). Management of total anomalous pulmonary venous return. Circulation 45, 1240–1250.

Goddeeris, M.M., Schwartz, R., Klingensmith, J., and Meyers, E.N. (2007). Independent requirements for hedgehog signaling by both the anterior heart field and neural crest cells for outflow tract development. Development 134, 1593–1604.

Goddeeris, M.M., Rho, S., Petiet, A., Davenport, C.L., Johnson, G.A., Meyers, E.N., and Klingensmith, J. (2008). Intracardiac septation requires hedgehog-dependent cellular contributions from outside the heart. Development 135, 1887–1895.

Gould, S.J. (1980). The evolutionary biology of constraint. Daedalus 109, 39–52.

Haberich, F. (1965). The functional separation of venous and arterial blood in the univentricular frog heart. Ann. N. Y. Acad. Sci. 127, 459–476.

Hoffmann, A.D., Peterson, M.A., Friedland-Little, J.M., Anderson, S.A., and Moskowitz, I.P. (2009). Sonic hedgehog is required in pulmonary endoderm for atrial septation. Development 136, 1761–1770.

Hoffmann, A.D., Yang, X.H., Burnicka-Turek, O., Bosman, J.D., Ren, X., Steimle, J.D., Vokes, S.A., McMahon, A.P., Kalinichenko, V. V., and Moskowitz, I.P. (2014). Foxf genes integrate Tbx5 and Hedgehog pathways in the second heart field for cardiac septation. PLoS Genet. 10, e1004604.

Hopkins, G. (1896). The heart of some lungless salamanders. Am. Nat. 30, 829–833.

Jensen, B., Wang, T., Christoffels, V.M., and Moorman, A.F.M. (2013). Evolution and development of the building plan of the vertebrate heart. Biochim. Biophys. Acta - Mol. Cell Res. 1833, 783–794.

Jiao, K., Kulessa, H., Tompkins, K., Zhou, Y., Batts, L., Baldwin, H.S., and Hogan, B.L.M. (2003). An essential role of Bmp4 in the atrioventricular septation of the mouse heart. Genes Dev. 17, 2362–2367.

Johansen, K. (1962). Double circulation in the amphibian *Amphiuma tridactylum*. Nature 194, 991–992.

Johansen, K. (1963). Cardiovascular dynamics in *Amphiuma tridactylum*. PhD Diss. Univ. Oslo.

Kerney, R. (2011). Embryonic staging table for a direct-developing salamander, *Plethodon cinereus* (Plethodontidae). Anat. Rec. 294, 1796–1808.

Koshiba-Takeuchi, K., Mori, A.D., Kaynak, B.L., Cebra-Thomas, J., Sukonnik, T., Georges, R.O., Latham, S., Beck, L., Henkelman, R.M., Black, B.L., et al. (2009). Reptilian heart development and the molecular basis of cardiac chamber evolution. Nature 461, 95–98.

Lawson, R. (1966). The anatomy of the heart of *Hypogeophis rostratus* (Amphibia, Apoda) and its possible mode of action. J. Zool. 149, 320–336.

Li, Q.Y., Newbury-Ecob, R.A., Terrett, J.A., Wilson, D.I., Curtis, A.R., Yi, C.H., Gebuhr, T., Bullen, P.J., Robson, S.C., Strachan, T., et al. (1997). Holt-Oram syndrome is caused by mutations in TBX5, a member of the brachyury (T) gene family. Nat. Genet. 15, 21–29.

Manceau, M., Domingues, V.S., Linnen, C.R., Rosenblum, E.B., and Hoekstra, H.E. (2010). Convergence in pigmentation at multiple levels: mutations, genes and function. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365, 2439–2450.

McMullen, E.C. (1938). The morphology of the aortic arches in four genera of plethodontid salamanders. J. Morphol. 62, 559–597.

Mohun, T.J., Leong, L.M., Weninger, W.J., and Sparrow, D.B. (2000). The morphology of heart development in *Xenopus laevis*. Dev. Biol. 218, 74–88.

Noble, G.K. (1925). The integumentary, pulmonary, and cardiac modifications correlated with increased cutaneous respiration in the amphibia: A solution of the "hairy frog" problem. J. Morphol. 40, 341–416.

Nussbaum, R.A., and Wilkinson, M. (1995). A new genus of lungless tetrapod: A radically divergent caecilian (Amphibia: Gymnophiona). Proc. R. Soc. London. Ser. B Biol. Sci. 261, 331–335.

Nye, H.L.D., Cameron, J.A., Chernoff, E.A.G., and Stocum, D.L. (2003). Extending the table of stages of normal development of the axolotl: limb development. Dev. Dyn. 226, 555–560.

Parker, S.E., Mai, C.T., Canfield, M. a, Rickard, R., Wang, Y., Meyer, R.E., Anderson, P., Mason, C.A., Collins, J.S., Kirby, R.S., et al. (2010). Updated National Birth Prevalence estimates for selected birth defects in the United States, 2004-2006. Birth Defects Res. A. Clin. Mol. Teratol. 88, 1008–1016.

Presnell, J.K., Schreibman, M.P., and Humason, G.L. (1997). Humason's Animal Tissue Techniques (Baltimore: Johns Hopkins University Press).

Putnam, J., and Kelly, D. (1978). A new interpretation of interatrial septation in the lungless salamander, *Plethodon glutinosus*. Copeia 1978, 251–254.

Putnam, J., and Parkerson Jr, J. (1985). Anatomy of the heart of the Amphibia II. *Cryptobranchus alleganiensis*. Herpetologica 41, 287–298.

Pyron, R.A., and Wiens, J.J. (2011). A large-scale phylogeny of Amphibia including over 2800 species, and a revised classification of extant frogs, salamanders, and caecilians. Mol. Phylogenet. Evol. 61, 543–583.

Rankin, S.A., Gallas, A.L., Neto, A., Gómez-Skarmeta, J.L., and Zorn, A.M. (2012). Suppression of Bmp4 signaling by the zinc-finger repressors Osr1 and Osr2 is required for Wnt/β-catenin-mediated lung specification in *Xenopus*. Development 139, 3010–3020.

Sangam, M.R., Devi, S.S.S., Krupadanam, K., and Anasuya, K. (2011). Cor triloculare biventriculare with left superior vena cava. Folia Morphol. (Warsz). 70, 135–138.

Seelye, A. (1906). Circulatory and respiratory systems of *Desmognathus fusca*. Proc. Bost. Soc. Nat. Hist. 32, 335–357.

Simons, J. (1957). The pulmonary return as an agent in the final development of the atrium in *Rana temporaria*. J. Embryol. Exp. Morphol. 5, 250–255.

Stern, D.L. (2013). The genetic causes of convergent evolution. Nat. Rev. Genet. 14, 751–764.

Taussig, H. (1960). Congenital Malformations of the Heart (Cambridge: Harvard University Press).

Tian, Y., Yuan, L., Goss, A.M., Wang, T., Yang, J., Lepore, J.J., Zhou, D., Schwartz, R.J., Patel, V., Cohen, E.D., et al. (2010). Characterization and *in vivo* pharmacological rescue of a Wnt2-Gata6 pathway required for cardiac inflow tract development. Dev. Cell 18, 275–287.

Tseng, Y.R., Su, Y.N., Lu, F.L., Jeng, S.F., Hsieh, W.S., Chen, C.Y., Chou, H.C., and Peng, S.S.F. (2007). Holt–Oram syndrome with right lung agenesis caused by a *de novo* mutation in the TBX5 gene. Am. J. Med. Genet. Part A 143, 1012–1014.

Wang, Q., Lan, Y., Cho, E.-S., Maltby, K.M., and Jiang, R. (2005). Odd-skipped related 1

(Odd 1) is an essential regulator of heart and urogenital development. Dev. Biol. 288, 582–594.

Whitford, W.G., and Hutchison, V.H. (1965). Gas exchange in salamanders. Physiol. Zool. 38, 228–242.

Wilder, H.H. (1896). Lungless salamanders. Anat. Anz. 12, 182–192.

Wilkinson, M., and Nussbaum, R.A. (1997). Comparative morphology and evolution of the lungless caecilian *Atretochoana eiselti* (Taylor) (Amphibia: Gymnophiona: Typhlonectidae). Biol. J. Linn. Soc. 62, 39–109.

Xie, L., Hoffmann, A.D., Burnicka-Turek, O., Friedland-Little, J.M., Zhang, K., and Moskowitz, I.P. (2012). Tbx5-hedgehog molecular networks are essential in the second heart field for atrial septation. Dev. Cell 23, 280–291.

Yoshikawa, N., and Matsui, M. (2014). A new salamander of the genus *Onychodactylus* from Tsukuba Mountains, eastern Honshu, Japan (Amphibia, Caudata, Hynobiidae). Zootaxa 3866, 53–78.

Yoshikawa, N., Matsui, M., Nishikawa, K., Kim, J., and Kryukov, A. (2008). Phylogenetic relationships and biogeography of the Japanese clawed salamander, *Onychodactylus japonicus* (Amphibia: Caudata: Hynobiidae), and its congener inferred from the mitochondrial cytochrome b gene. Mol. Phylogenet. Evol. 49, 249–259.

Zhou, L., Liu, J., Olson, P., Zhang, K., Wynne, J., and Xie, L. (2015). Tbx5 and Osr1 interact to regulate posterior second heart field cell cycle progression for cardiac septation . J. Mol. Cell. Cardiol. 85 , 1–12.

Chapter 6

Discussion

Lung loss is not an isolated phenotype that can be analyzed independent of its physiological, developmental and evolutionary context. Regardless of the selective advantage of lung loss, which is unresolved, proto-plethodontid salamanders overcame a number of practical constraints in order to evolve lung loss. Among these is the considerable amount of pleiotropy exhibited by critical lung patterning genes. Coding mutations to many of these genes likely entail the malformation of other organs such as appendages, endodermal organs and the heart. The formation of the lung rudiment may be evidence that lung gene function is maintained, despite millions of years of potential genetic drift since functional adult lungs became lost (Chapter 2). It is likely that cis-regulatory evolution of the lung gene regulatory network mitigated the negative effects of coding mutations to these critical developmental genes (Chapter 3). The fact that the lung gene regulatory network topology is essentially intact (albeit with significant differences in expression levels) argues that stabilizing selection has maintained lung developmental modules. Constraint is also evident physiologically. The evolution of lunglessness requires profound transitions in the respiratory and circulatory

system. Lungless salamanders evolved new approaches to cutaneous and buccopharyngeal respiration, potentially facilitated by the neofunctionalization of a pulmonary gene paralog (Chapter 4). Lastly, lung loss affects the configuration of the circulatory system, and several lungless lineages independently evolved similar cardiac morphology to account for these changes (Chapter 5). An integrative and multidisciplinary approach is required to understand the evolutionary and developmental context for lung loss.

6.1 Summary of major findings

In Chapter 2, I demonstrate that several species of lungless salamanders develop a transient lung rudiment. Lungs are correctly specified in the lungless species *Plethodon cinereus*, as evidenced by the expression of Wnt2b in the presumptive pulmonary mesenchyme. A laryngotracheal tube develops in multiple species of lungless salamander. This laryngotracheal tube, also referred to as a lung rudiment, expresses markers of lung differentiation including Sox9, Nkx2.1, and pulmonary surfactant proteins A, B, C, and D. The lung rudiment then regresses by apoptosis, accounting for adult lunglessness in all species of Plethodontidae.

The development of an incipient lung indicates that some aspects of lung development are conserved in plethodontids. In particular, lung specification, early outgrowth and cellular differentiation are intact. However, lung development ceases in plethodontid embryos and the lungs regress. The expression pattern of certain markers hints at the tissue identity of the lung rudiment. In particular, Sox9 expression through the lung rudiment signals that the tissue has an alveolar, or distal, identity. This implies that proximal portions of the lung, such as the trachea, fail to form. However, the unpaired nature of the laryngotracheal tube

does not support this interpretation.

In Chapter 3, transcriptome sequencing reveals differences in the gene expression profiles of the lung primordia between lungless (*P. cinereus*) and lunged (*A. mexicanum*) salamanders in several important respects, including in the expression of indicators of Tgfβ signaling. I demonstrate that the development of the trachea in *A. mexicanum* is inhibited by ectopic Tgfβ protein. In contrast, lung specification expands when Tgfβ signaling is inhibited in *A. mexicanum*. The lung rudiment is retained when Tgfβ signaling is inhibited in *P. cinereus*, but the phenotype may be the result of developmental delay. I also show that introduction of lunged salamander splanchnic mesoderm to lungless salamanders by heterospecific transplantation may restore lung development.

Transcriptome sequencing offers a comprehensive picture of the differences in gene expression between the lung primordia of lunged and lungless salamanders. While the expression of several genes or genetic pathways is significantly downregulated or upregulated in lungless salamanders, I focused on the role of Tgfβ signaling in lung loss. Tgfβ signaling has many roles in the development of the lungs (Fig. 3.4), including the suppression of Nkx2.1 expression and Fgf10 expression, both of which are necessary for complete lung development. The lack of trachea in explants treated with ectopic Tgfβ protein implies that overactive Tgfβ signaling may repress proximal lung development.

In Chapter 4, I report the discovery of a paralog of the gene SPC, which I name SPC-like. I find that SPC gene duplication likely occurred in salamanders. SPC-like has a conserved three-dimensional structure relative to SPC, indicating conserved function in surface-tension reduction. SPC-like is expressed in *A. mexicanum* solely in the lung, which is similar to the SPC expression pattern in this species. In contrast, SPC-like is expressed in the integument in

embryonic and larval specimens of the lungless salamander species *Desmognathus fuscus*. In *D. fuscus*, SPC-like expression shifts to the buccopharyngeal mucosa (oral cavity) just prior to metamorphosis and expression is maintained in this tissue into adulthood. I examine the integument of *D. fuscus* by histology and transmission electron microscopy and find evidence for elevated secretory activity of the larval integument in *D. fuscus*. I also find structures that resemble lamellar bodies in the integument, which are associated with surfactant trafficking and exocytosis in the lung.

The results from Chapter 4 indicate that, following its formation from gene duplication of an ancestral SPC sequence, SPC-like has experienced distinct patterns of regulatory evolution in different lineages of salamanders. In the lunged species *A. mexicanum*, SPC-like demonstrates little divergence from the SPC expression pattern. In lungless *D. fuscus*, substantial shifts in the expression pattern of SPC-like indicate pronounced gene regulatory changes. Expression has shifted from the pulmonary endoderm to the embryonic and larval integument and after metamorphosis to the adult buccopharyngeal membranes.

SPC plays a major role in pulmonary surfactant function and may help reduce mucous layer thickness, resulting in more efficient gas exchange (Guyton et al., 1984; Sosnowski et al., 1998; Ultsch and Gros, 1979). In mammals, pulmonary surfactant increases lung compliance and prevents alveolar collapse. However, the function of pulmonary surfactant in amphibians is unresolved because amphibians have intrinsically compliant lungs and their large alveolar spaces are not prone to collapse. Previous authors have attempted to explain surfactant presence in amphibians as important for preventing the sides of the lungs from self-adhering (Daniels et al., 1998). While theoretically possible, there is little experimental support for this "anti-glue" hypothesis (Daniels et al., 1995). I propose a different hypothesis:

surfactant reduces the thickness of the mucous layer, which directly facilitates gas exchange. This hypothesis accounts for both the presence of pulmonary surfactant in amphibian lungs and the conserved expression of pulmonary surfactant proteins in plethodontids. SPC-like may help to maintain a thin mucous layer over the extrapulmonary respiratory tissues of *D. fuscus*.

This proposed functional role for SPC-like in combination with the correlation of SPC-like expression with the sites over which gas exchange is being performed suggest that SPC-like is neofunctionalized in lungless salamanders for extrapulmonary respiration. Neofunctionalization occurs when one gene paralog assumes a function distinct from the ancestral function, which is maintained by the other paralog (He and Zhang, 2005).

In Chapter 5, I characterize atrial septum morphology and development across a diverse sample of lunged and lungless salamanders. I find that lunged salamanders possess a two-part atrial septum, whereas plethodontid salamanders lack a portion of the atrial septum. I show that partial loss of the atrial septum has also occurred in the convergently lungless salamander species *Onychodactylus japonicus*. Transcriptomics reveals that genes that are known to regulate the formation of the atrial septum in mammals are significantly downregulated in the lungless species *P. cinereus* relative to *A. mexicanum*.

These data indicate that previous hypotheses regarding the loss of complete atrial septation in lungless or lung-reduced salamanders may be incorrect. Previous authors have proposed that cutaneous respiration makes mixed systemic blood high in oxygen, resulting in relaxed selective pressure to retain atrial septation (Johansen and Hanson, 1968). I propose an alternative hypothesis for observed atrial septum reduction in salamanders: complete atrial septation is maladaptive when pulmonary return is reduced or absent due to a blood

pressure differential between atrial chambers. The strong correlation between lung reduction/loss and atrial septum reduction points to convergent evolution mediated by positive selection against complete atrial septation. Atrial septum reduction may be mechanistically achieved through developmental crosstalk between the heart and lungs (Chapter 5). Lungless salamanders may serve as beneficial models for the study of human cardiopulmonary defects including atrial septal defects and anomalous pulmonary venous return (Fig. 6.1).

6.2 Future directions

There are several additional studies that would complement my exploration of lung loss in salamanders. Determining which genes are directly responsible for lung loss may require additional sequencing and experimental efforts. For instance, sequencing the lung primordia transcriptomes from two or more additional lunged and lungless salamanders may help to determine whether the observed changes in gene expression patterns contribute to lung loss. Particularly, it will be important to demonstrate consistent gene expression within lunged samples and within lungless samples. Additional experiments to change levels of Tgfβ signaling in lunged and lungless species may add support to my hypothesis that elevated Tgfβ signaling is responsible for lung loss. These experiments may include further treatments on lungless salamanders or gene-editing experiments in *A. mexicanum*. In addition, future work should interrogate the function of other genetic pathways in lung loss, including Shh signaling and RA signaling.

While heterospecific transplantation results are promising, this study requires a better method to prove that pulmonary induction occurs. This may involve generating a set of *in*

situ hybridization probes for *Hemidactylium scutatum*, or finding or generating a crossreactive antibody to the lungs of salamanders, potentially against Nkx2.1.

Lastly, additional descriptive data would be beneficial. For instance, despite sectioning several lungless *Eurycea bislineata* embryos, I did not identify the laryngotracheal tube. Adding species to my study of lung rudiment development in plethodontids will help to demonstrate whether the formation of the laryngotracheal tube is common to all plethodontids or only a subset of species. If possible, convergently lungless embryos from the genus *Onychodactylus* should be analyzed to determine whether lunglessness in these species is due to regression of an embryonic rudiment, providing documentation of the developmental basis for this convergent trait.

To better understand the evolution and function of SPC-like, I propose several future experiments. First, the origin of the putative SPC gene duplication needs to be determined. The novel SPC paralog, SPC-like, is expressed in *D. fuscus*, *P. cinereus* and *A. mexicanum*, which have an ancient last common ancestor. Therefore, phylogenetic bracketing implies that SPC-like is present in many other species of salamanders, but has not yet been sequenced. In future work, I intend to sequence SPC and SPC-like from a broad array of amphibians in order to pinpoint when the duplication event occurred. I will then examine the tissue specific expression pattern of SPC-like in additional species to determine whether divergence in expression pattern is unique to lungless salamanders. The molecular evolution of SPC-like is unknown. I intend to look for signatures of positive selection on SPC-like using phylogenetic analysis by maximum likelihood (PAML). If possible, cis-regulatory sequences can also be examined for signatures of selection in lungless salamanders.

While SPC is expressed in *P. cinereus* in the lung rudiment, it is expressed at very low

levels. It remains to be demonstrated if *D. fuscus* expresses SPC in addition to SPC-like. If both are expressed, then the expression pattern of SPC needs to be investigated. Initial efforts to clone SPC have resulted in cloning of SPC-like.

It is unknown whether SPC-like plays a role in respiration in lungless salamanders. I have identified several potential functional approaches that can be used to address this question. *In vitro* experiments utilizing captive bubble surfactometry may reveal if SPC-like reduces surface tension or aids phospholipid adsorption to the air-liquid interface in a similar fashion to SPC (Schürch et al., 1998). However, *in vitro* approaches depend on the ability to isolate recombinant SPC-like (Hafner et al., 1999). SPC-like can be ectopically expressed in lunged salamanders to determine if SPC-like aids cutaneous respiration. Ectopic expression can be induced using an MMLV or similar retroviral system (Whited et al., 2013). Cutaneous respiration of treated and control individuals can then be measured in a simple two-chambered closed system respirometer, which decouples pulmonary and cutaneous respiration (Whitford and Hutchison, 1965). An alternative to ectopic expression using a retroviral system is to perform cutaneous respirometry on lunged salamanders with and without the addition of clinically available synthetic pulmonary surfactant to their aquatic media. If adding surfactant aids cutaneous respiration, then this points to a role of SPC-like in cutaneous gas exchange.

An ideal functional experiment would be to abrogate the expression of SPC-like in *D. fuscus* using the CRISPR-Cas9 system and test for respiratory function. However, this approach relies on the collection of early *D. fuscus* embryos, which would be challenging.

Lastly, it is possible to investigate whether SPC-like functionally substitutes for SPC by using an SPC-deficient mouse model. Humans and mice with SPC mutations develop

emphysema-like morphology and pneumonitis (Beers and Mulugeta, 2005; Glasser et al., 2003). By generating a transgenic mouse with SPC-like targeted to the SPC locus and then crossing this mouse line into SPC knockout mice, the functional role of SPC-like can be better determined. SPC-like knockin heterozygote mice can be bred to heterozygous SPC knockout mice. Resulting SPC-like (+) / SPC (–) mice can then be backcrossed to SPC +/– heterozygotes. Resulting SPC-like (+) / SPC (–) progeny can be compared to littermate controls of genotypes SPC –/– and SPC +/– and characterized by histology.

The study of atrial septum loss in plethodontid salamanders may have biomedical applications because plethodontids phenocopy atrial septal defects and the condition anomalous pulmonary venous return (Fig. 6.1). For a more robust molecular analysis, *in situ* hybridization should be employed to confirm putative differences in gene expression between lunged and lungless species. Subsequently, genome editing approaches can be employed to reduce levels of Foxf genes, Tbx5, and Gli1 expression in *A. mexicanum*, to see whether these genes play a conserved role in salamander atrial septum development (Hoffmann et al., 2014). Finally, apoptosis was observed in the second heart field mesoderm preceding laryngotracheal tube regression (Chapter 2). Further study is necessary to determine whether apoptosis of progenitor cells is responsible for atrial septum reduction, or whether atrial septum reduction is due to other factors such as altered progenitor cell specification.

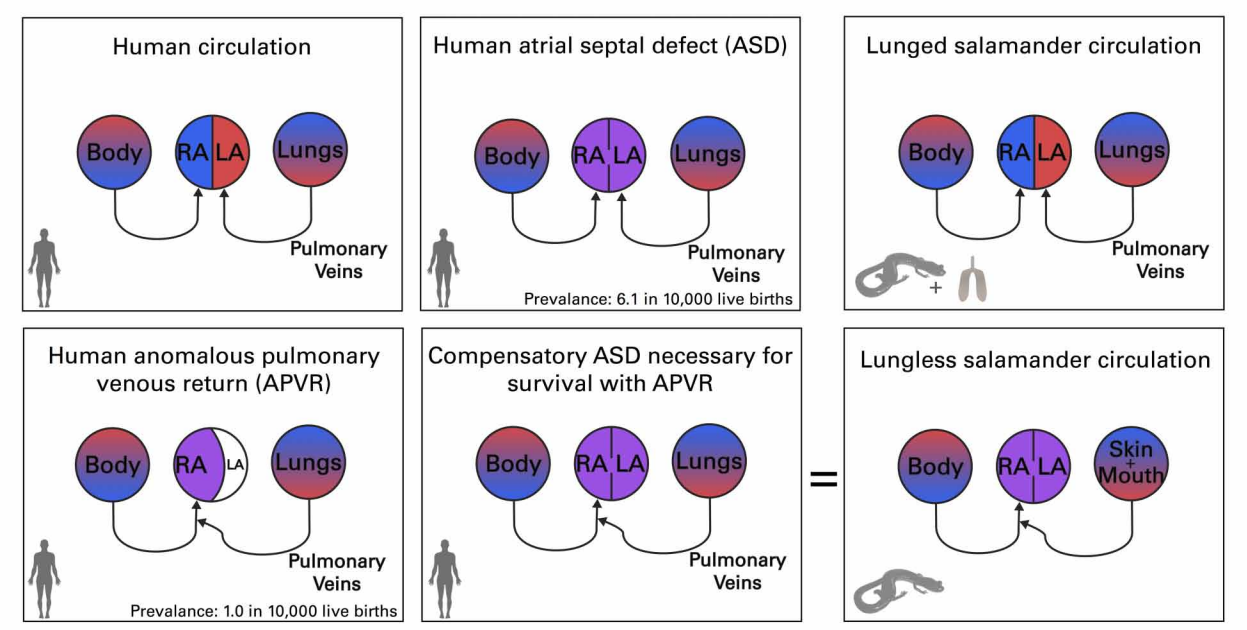


Figure 6.1: Lungless salamanders phenocopy human cardiac birth defects. Congenital heart defects are among the most common human birth defects (Bjornard et al., 2013). Interestingly, the configuration of the circulatory system of lungless salamanders morphologically resembles anomalous pulmonary venous return (APVR) with an atrial septal defect (ASD). Patients with APVR experience deleterious pressure differentials between atrial chambers. Partial septum loss may prevent pressure differentials in lungless salamanders. Abbreviations: LA, left atrium; RA, right atrium.

6.3 What can amphibians can tell us about mammalian lung development?

In my dissertation I use a comparative approach to understand the process of lung development in salamanders and the types of genetic changes that may result in lung loss. Amphibians are optimal models for the study of lung development given that lung loss in lunged species is not lethal, which allows for experimental approaches not possible in mammals (Garber, 1930). The study of amphibian lung development may reveal novel insights into the evolution and development of the lungs. For instance, my explant studies help

confirm that the lungs and trachea develop as modules and suggest that the early formation of the laryngotracheal groove may be independent from the formation of the alveolar buds (Domyan et al., 2011).

Future transcriptomic approaches to the study of lung development in non-model species will likely yield invaluable information regarding the formation and evolution of the lung. Mammalian alveoli are composed of two principal epithelial cell types: type I and type II alveolar epithelial cells (AT1 and AT2). AT1 cells are characterized by a flattened morphology, suggestive of a role in gas exchange, while AT2 cells play a role in secretion of pulmonary surfactant. Ultrastructurally, AT2 cells have numerous organelles and highly distinctive lamellar bodies, which function to store and exocytose pulmonary surfactant fluid (Schmitz and Müller, 1991). AT1 and AT2 cells arise from a bipotential progenitor (BP) cell (Desai et al., 2014; Treutlein et al., 2014). BP cells co-express sets of molecular markers found in either AT1 or AT2 cells, and the upregulation or downregulation of certain genes within this set of markers has been precisely charted at the single-cell level during the short period of differentiation into AT1 or AT2 cells at high temporal resolution (Treutlein et al., 2014).

Whether amphibians possess differentiated AT1 and AT2 cells or a single cell type combining properties of both is unresolved. Some researchers have described differentiated AT1 and AT2 cells in frogs and salamanders (Okada et al., 1965a). However, most papers cite the presence of only a single type of alveolar epithelial cell (AEC), which has both lamellar bodies and thin cytoplasmic projections, indicating that these cells function in both secretion and gas exchange (Dierichs and Dosche, 1982; Hightower et al., 1975; Meban, 1979; Pattle et al., 1977). My preliminary ultrastructural analysis in *A. mexicanum* appears to confirm

this dual-function hypothesis. These amphibian cells with AT1 and AT2 properties share ultrastructural similarities to the epithelial cells in the swim bladder in fish (Podkowa and Goniakowska-Witalińska, 1998) or those in the lungfish lung (Power et al., 1999). Differentiated AT1 and AT2 cells are present in reptiles and mammals (Meban, 1978; Okada et al., 1965b). Therefore, amphibians represent a critical node in understanding the evolution of differentiated alveolar epithelial cells.

The formation of differentiated cell types from a common progenitor cell is both a developmental and evolutionary process. It is likely that differentiated AT1 and AT2 cells not only differentiate from a genetically intermediate bipotential progenitor but also have evolved from a dual functioning AEC. Since genes involved in the developmental differentiation of AECs are well known, in a preliminary study I examined the presence of these genes across various animal genomes as well as the expression of these genes within the swimbladder or salamander lung (Fig. 6.2).

The evolution of differentiated AT1 and AT2 cells is likely associated with the evolution of new genes, changes in expression of existing genes, or changes in gene function. Both fishes and amphibians lack several genes associated with fully differentiated AT1 and AT2 cells (Fig. 6.2). It is also interesting to note genes present within amphibian and fish genomes that are not expressed within the lungs and swimbladders. Heterotopic expression of existing genes within the lungs of amniotes may be important for the evolution of AT1 and AT2 cells. Examples include expression of Hopx in AT1 cells and Lamp3 in AT2 cells (Fig. 6.2). Genes specific to amniote genomes, such as s100a6, may also be important for cellular differentiation. S100a6, for instance, plays a role in regulating cell shape (Breen and Tang, 2003), and the evolution and expression of this gene may be involved in the squamous

morphology of AT1 cells. Examination of cellular differentiation within an evolutionary and developmental context using transcriptomics may yield a more complete picture of cell type evolution. Future characterization of anamniote lung transcriptomes may reveal essential genes for the development and evolution of differentiated AT1 and AT2 cells.

6.4 Conclusion

One critique of the current state of research in the field of evolution and development, or evo-devo, is that much work is focused on trait loss as opposed to the evolution of novelty (Hoekstra and Coyne, 2007). I argue that the study of trait loss remains critical to understanding the developmental basis of phenotypic evolution and biodiversity. While trait losses themselves do not explain the evolution of new traits, losses are often associated with the origin of new clades of animals or transitions to new ecological niches. Furthermore, trait loss is often associated with a plethora of novel adaptations contingent upon trait loss, or novel features potentially resulting from adaptation in the absence of the original trait. For plethodontid salamanders, lung loss may have contributed to their ability to feed terrestrially, and hence their remarkable adaptive radiation (Lombard and Wake, 1986). I posit that lung loss may have been enabled by the evolution of novel mechanisms for extrapulmonary respiration. Alternatively, novel respiratory mechanisms may have evolved following lung loss (Chapter 4).

Rudimentary structures are often subject to great variability in morphology (Darwin, 1859, 1874). However, in my study of the lung rudiment in plethodontids I find that morphology and development of the laryngotracheal tube is conserved across species of lungless

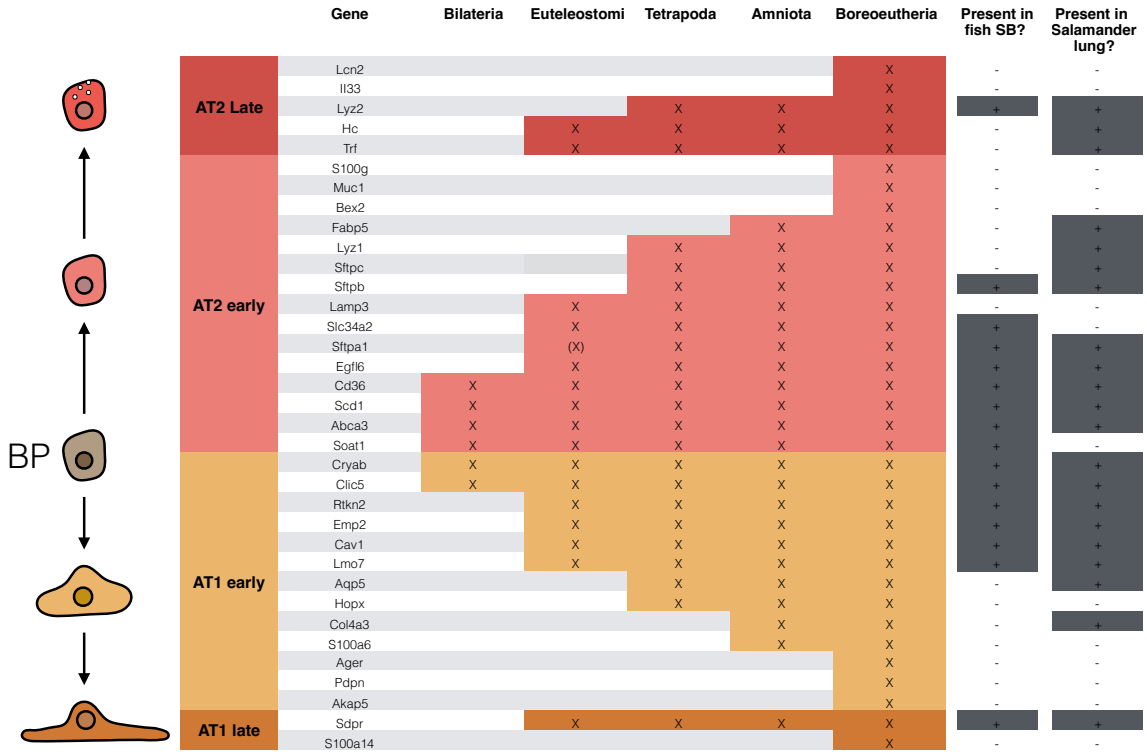


Figure 6.2: Evolutionary developmental differentiation of alveolar epithelial cells into alveolar type I (AT1) and alveolar type II (AT2) cells. The genes expressed during the transition from a mammalian bipotential progenitor (BP) cell to either AT1 or AT2 cells have been characterized by Treutlein et al. (2014). Using NCBI's Homologene database, I determined whether orthologs of these mammalian genes have been identified in other animal genomes. An "X" and a colored box indicate the genomic presence of a gene within a broad metazoan taxonomic grouping. Sftpa1 is likely present across Euteleostomi, though not represented in the Homologene database, therefore it is represented with and "(X)" (Daniels et al., 2004). I next determined whether each of these genes is expressed in swimbladders (SB) or lungs, using transcriptome data for swimbladders from *Danio rerio* and *Takifugu rubripes* and salamander lungs from *Ambystoma mexicanum* and *A. tigrinum* (Cui et al., 2014; Eo et al., 2012; Zheng et al., 2011; Present study, Chapter 3). The presence of these genes within the swimbladder and lung transcriptomes is indicated by a + sign and a grey box in the right two columns.

salamanders. The morphological and molecular conservation of early lung development in lungless salamanders indicates that pleiotropic constraints act to maintain these early morphogenetic processes over long evolutionary timescales (Wright, 1964). The concept of how

an organ continues to regress following loss of biological function has historically been slightly vexing. Darwin wrote, "There remains, however, this difficulty. After an organ has ceased being used, and has become in consequence much reduced, how can it be still further reduced in size until the merest vestige is left; and how can it be finally quite obliterated?" (Darwin, 1859). Many current researchers would argue that genetic drift could result in the complete loss of organ rudiments. Darwin (1859) posited that complete loss is favored if further reduction is selectively advantageous in terms of energy economy. In either case, the fact that lung development is not completely abrogated in plethodontids further argues for either a functional role of the structure itself, or stabilizing selection on the genes that pattern it due to their pleiotropic actions.

In the coming years, advances in genomics will greatly enhance our ability to understand the genetic basis of animal evolution. Rudiments will continue to play a key role in these studies because they offer a dynamic view of the evolution of development—both evidence of evolutionary history and snapshots of evolution's genetic underpinnings. The rudiment, though perhaps of little use to its possessor, has been and will continue to be of great utility for scientists. The study of rudiments stands little chance of atrophy.

6.5 References

Beers, M.F., and Mulugeta, S. (2005). Surfactant protein C biosynthesis and its emerging role in conformational lung disease. Annu. Rev. Physiol. 67, 663–696.

Bjornard, K., Riehle-Colarusso, T., Gilboa, S.M., and Correa, A. (2013). Patterns in the prevalence of congenital heart defects, metropolitan Atlanta, 1978 to 2005. Birth Defects Res. Part A - Clin. Mol. Teratol. 97, 87–94.

Breen, E.C., and Tang, K. (2003). Calcyclin (S100A6) regulates pulmonary fibroblast proliferation, morphology, and cytoskeletal organization in vitro. J. Cell. Biochem. 88, 848–854.

Cui, J., Liu, S., Zhang, B., Wang, H., Sun, H., Song, S., Qiu, X., Liu, Y., Wang, X., Jiang, Z., et al. (2014). Transciptome Analysis of the Gill and Swimbladder of *Takifugu rubripes* by RNA-Seq. PLoS One 9, e85505.

Daniels, C.B., Orgeig, S., and Smits, A.W. (1995). The evolution of the vertebrate pulmonary surfactant system. Physiol. Zool. 68, 539–566.

Daniels, C.B., Lopatko, O. V, and Orgeig, S. (1998). Evolution of surface activity related functions of vertebrate pulmonary surfactant. Clin. Exp. Pharmacol. Physiol. 25, 716–721.

Daniels, C.B., Orgeig, S., Sullivan, L.C., Ling, N., Bennett, M.B., Schürch, S., Val, A.L., and Brauner, C.J. (2004). The origin and evolution of the surfactant system in fish: insights into the evolution of lungs and swim bladders. Physiol. Biochem. Zool. 77, 732–749.

Darwin, C. (1859). On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life (London: John Murray).

Darwin, C. (1874). The Descent of Man, and Selection in Relation to Sex (New York: A.L. Burt).

Desai, T.J., Brownfield, D.G., and Krasnow, M.A. (2014). Alveolar progenitor and stem cells in lung development, renewal and cancer. Nature 507, 190–194.

Dierichs, R., and Dosche, C. (1982). The alveolar-lining layer in the lung of the axolotl, *Ambystoma mexicanum*. Cell Tissue Res. 222, 677–686.

Domyan, E.T., Ferretti, E., Throckmorton, K., Mishina, Y., Nicolis, S.K., and Sun, X. (2011). Signaling through BMP receptors promotes respiratory identity in the foregut via repression of Sox2. Development 138, 971–981.

Eo, S.H., Doyle, J.M., Hale, M.C., Marra, N.J., Ruhl, J.D., and DeWoody, J.A. (2012). Comparative transcriptomics and gene expression in larval tiger salamander (*Ambystoma tigrinum*) gill and lung tissues as revealed by pyrosequencing. Gene 492, 329–338.

Garber, S. (1930). Metamorphosis of the axolotl following lung extirpation. Physiol. Zool. 3, 373–378.

Glasser, S.W., Detmer, E.A., Ikegami, M., Na, C.-L., Stahlman, M.T., and Whitsett, J.A. (2003). Pneumonitis and emphysema in SP-C gene targeted mice. J. Biol. Chem. 278, 14291–14298.

Guyton, A.C., Moffatt, D.S., and Adiar, T.H. (1984). Role of alveolar surface tension in

transepithelial movement of fluid. In Pulmonary Surfactant, B. Robertson, L. Van Golde, and J. Batenburg, eds. (Amsterdam: Elsevier Science Publishers), pp. 171–185.

Hafner, D., Germann, P.G., Hauschke, D., and Kilian, U. (1999). Effects of early treatment with rSP-C surfactant on oxygenation and histology in rats with acute lung injury. Pulm. Pharmacol. Ther. 12, 193–201.

He, X., and Zhang, J. (2005). Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. Genetics 169, 1157–1164.

Hightower, J.A., Burke, J.D., and Haar, J.L. (1975). A light and electron microscopic study of the respiratory epithelium of the adult aquatic newt, *Notophthalmus viridescens*. Can. J. Zool. 53, 465–472.

Hoekstra, H.E., and Coyne, J.A. (2007). The locus of evolution: evo devo and the genetics of adaptation. Evolution 61, 995–1016.

Hoffmann, A.D., Yang, X.H., Burnicka-Turek, O., Bosman, J.D., Ren, X., Steimle, J.D., Vokes, S.A., McMahon, A.P., Kalinichenko, V. V., and Moskowitz, I.P. (2014). Foxf genes integrate Tbx5 and Hedgehog pathways in the second heart field for cardiac septation. PLoS Genet. 10, e1004604.

Johansen, K., and Hanson, D. (1968). Functional anatomy of the hearts of lungfishes and amphibians. Am. Zool. 8, 191–210.

Lombard, R.E., and Wake, D.B. (1986). Tongue evolution in the lungless salamanders, family Plethodontidae IV. Phylogeny of plethodontid salamanders and the evolution of feeding dynamics. Syst. Zool. 35, 532–551.

Meban, C. (1978). Functional anatomy of the lungs of the green lizard, *Lacerta viridis*. J. Anat. 125, 421–431.

Meban, C. (1979). An electron microscope study of the respiratory epithelium in the lungs of the fire salamander (*Salamandra salamandra*). J. Anat. 128, 215–224.

Okada, Y., Ishiko, S., Daido, S., Kim, J., and Ikeda, S. (1965a). Comparative morphology of the lung with special reference to alveolar epithelial cells. 1. Lung of the Amphibia. Acta Tuberc. Jpn. 14, 89–95.

Okada, Y., Ishiko, S., Daido, S., Kim, J., and Ikeda, S. (1965b). Comparative morphology of the lung with special reference to alveolar epithelial cells. 2. Lung of the Reptilia. Acta Tuberc. Jpn. 14, 89–95.

Pattle, R.E., Schock, C., Creasey, J.M., and Hughes, G. (1977). Surpellic films, lung surfactant, and their cellular origin in newt, caecilian, and frog. J. Zool. London 182, 125–136.

Podkowa, D., and Goniakowska-Witalińska, L. (1998). The structure of the airbladder of the catfish *Pangasius hypophthalmus* Roberts and Vidthayanon 1991 (previously *P. sutchi* Fowler 1937). Folia Biol. 46, 189–196.

Power, J.H., Doyle, I.R., Davidson, K., and Nicholas, T.E. (1999). Ultrastructural and protein analysis of surfactant in the Australian lungfish *Neoceratodus forsteri*: evidence for conservation of composition for 300 million years. J. Exp. Biol. 202, 2543–2550.

Schmitz, G., and Müller, G. (1991). Structure and function of lamellar bodies, lipid-protein complexes involved in storage and secretion of cellular lipids. J. Lipid Res. 32, 1539–1570.

Schürch, S., Green, F.H.Y., and Bachofen, H. (1998). Formation and structure of surface films: Captive bubble surfactometry. Biochim. Biophys. Acta 1408, 180–202.

Sosnowski, T., Gradon, L., Skoczek, M., and Drozdiel, H. (1998). Experimental evaluation of importance of the pulmonary surfactant for oxygen transfer rate in human lungs. Int. J. Occup. Saf. Ergon. 4, 391–409.

Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M.A, and Quake, S.R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature 509, 371–375.

Ultsch, G.R., and Gros, G. (1979). Mucus as a diffusion barrier to oxygen: Possible role in O_2 uptake at low pH in carp (*Cyprinus carpio*) gills. Comp. Biochem. Physiol. Part A Physiol. 62, 685–689.

Whited, J.L., Tsai, S.L., Beier, K.T., White, J.N., Piekarski, N., Hanken, J., Cepko, C.L., and Tabin, C.J. (2013). Pseudotyped retroviruses for infecting axolotl *in vivo* and *in vitro*. Development 140, 1137–1146.

Whitford, W.G., and Hutchison, V.H. (1965). Gas exchange in salamanders. Physiol. Zool. 38, 228–242.

Wright, S. (1964). Pleiotropy in the evolution of structural reduction and of dominance. Am. Nat. 98, 65–69.

Zheng, W., Wang, Z., Collins, J.E., Andrews, R.M., Stemple, D., and Gong, Z. (2011). Comparative transcriptome analyses indicate molecular homology of zebrafish swimbladder and mammalian lung. PLoS One 6, e24019.

Appendix A

Collection sites, primers, and plasmids

Table A.1: Locality for plethodontid salamander collection.

Species	Site Name	Approx. GPS Coordinates
Desmognathus fuscus	Ashfield, MA	42.483111, -72.761263
Desmognathus fuscus	Mass Audubon Wachusett Meadows Preserve	42.450922, -71.913009
Hemidactylium scutatum	Cape Cod National Seashore	41.9597778, -70.0567333
Plethodon cinereus	Mass Audubon Wachusett Meadows Preserve	42.450922, -71.913009
Plethodon cinereus	Mass Audubon Rutland Brook Sanctuary	42.467594, -72.157106
Plethodon cinereus	Harvard Forest	42.456331, -72.165636
Plethodon cinereus	Harvard Forest	42.507220, -72.216362
Plethodon cinereus	Willard Brook State Forest	42.671606, -71.776156
Plethodon cinereus	Estabrook Woods	42.490386, -71.348666

Table A.2: Primers used for cloning lung genes.

Gene Name	Species	Left primer	Right Primer
Wnt2b	*P. cinereus*	5'-GGG CGC ATA TGA ATT TGC A-3'	5'-CAG CTG CTA GGG TCT ACA GC-3'
Wnt2b	*A. mexicanum*	5'-CAT GCA GCC AAG GAG AGA TTA-3'	5'-GGA CTA CAG CTC CCA GAA TTG-3'
Tbx5	*P. cinereus*	5'-CAG CCA TGC CAG GAC GAT TA-3'	5'-CTG TGA GCT CAG GTT GGA GG-3'
Tbx5	*A. mexicanum*	5'-GCA CAC TTG ACC AGC TGT TG-3'	5'-ATA CGT CCA TCC GGA TTC GC-3'
Nkx2.1	*P. cinereus*	5'-CAG GCG AAC AAA CAG GCT TC-3'	5'-AGG CCC AGG TGT ATG AGC TA-3'
Nkx2.1	*A. mexicanum*	Nested PCR:	Nested PCR:
		5'-ACT CCT TTC TCA GTG TCT GAC-3'	5'-GTA GCG GTG GTT CTG GAA CC-3'
		followed by:	followed by:
		5'-GTC CCC TGG AGG AGA GCT AC-3'	5'-GTG AGG TGG ATC ATG CTG GCT A-3'
Sox2	*P. cinereus*	5'-GAA CGG CAG TCC CAC ATA CA-3'	5'-GAC CGT CAT CCT GAT TGC CA-3'
Sox2	*A. mexicanum*	5'-AAA CAG GAT GAA GAG CCG GG-3'	5'-TAG GAC TCC TGC GAA CCT CA-3'
Ptch1	*P. cinereus*	5'-CTC TTC CTG CTG AAC CCC TG-3'	5'-GTC GAC ACC TTG CAT CTG GA-3'
Axin2	*P. cinereus*	5'-TCC ACC TCC ATG ACA CGT TG-3'	5'-TCG TAA CGT GCT TGC AGG AT-3'
Axin2	*A. mexicanum*	5'-CCC TTC TGT GAC CTT GGG AC-3'	5'-CGA TCC ACA GTC CAT CCT CG-3'
APC	*P. cinereus*	5'-GGC AAA TAA GAG CCG CCA TG-3'	5'-ATC TGG TGT TCT GCT GGA GC-3'
APC	*A. mexicanum*	5'-CGT TGT ACC TTC CGT GGA CA-3'	5'-ACG AGC AAA GGC ATC CAA GA-3'
Acvr1b	*P. cinereus*	5'-TGC GGA CAT GCA TCT CCT TT-3'	5'-CGA ACA GGG ATC CGT ACT CG-3'
Acvr1b	*A. mexicanum*	5'-CGT GAT AGT CCG AGA CCA GC-3'	5'-TTT CTA CTG CCA GAG CAC GG-3'
Smad6	*P. cinereus*	5'-GGC TTG GAT TGG GTC TCT GT-3'	5'-GGC CTA GTC GTA CGC AAA GT-3'
Smad6	*A. mexicanum*	5'-TCA AAC ACT TTG GCA GAG TAG C-3'	5'-CTA CTG GGA GCA CCG AAC G-3'
SPC	*A. mexicanum*	5'-CAC ACA GAR AMG ATT TTC CAG ATG-3'	5'-CGT CTT GTC CAT TTT TGT KAB GTA GCA-3'
SPC-like	*A. mexicanum*	5'-AAG ATG GAA ACC GGC AGC AAG C-3'	5'-CGT CTT GTC CAT TTT TGT KAB GTA GCA-3'
SPC-like	*D. fuscus*	5-AAG ATG GAA ACC GGC AGC AAG C-3'	5'-AGT ATT GGA AGC GGT CTG GGT G-3'
SPC-like	*P. cinereus*	5'-AAG ATG GAA ACC GGC AGC AAG C-3'	5'-GGT GTA GTC ATA GAC CAC-3'

Table A.3: Plasmids obtained from other researchers.

Gene Name	Species	Plasmid information
Sox9	*A. mexicanum*	Courtesy of Tanaka lab (in pCRII, antisense: NotI/SP6; AmpR,)
Sox9	*P. cinereus*	Courtesy of R. Kerney (in pBSK, antisense: T7/BamHI; AmpR)
Fgf10	*A. mexicanum*	Courtesy of N. Frobisch (in pCRII-TOPO, antisense: SP6/ApaI; Kan/Amp)
Shh	*A. mexicanum*	Courtesy of N. Frobisch (in pCRII-TOPO, antisense: SP6/NotI; Kan/Amp)
Shh	*P. cinereus*	Courtesy of J. Gross (in pGEMT, antisense: Sp6/NcoI; AmpR)

Appendix B

Sequencing methods

B.1 RNA isolation, library preparation and sequencing

Specimens were anesthetized using MS-222 (Sigma, St. Louis, MO), then immersed in ice-cold 15% sucrose containing MS-222 for 15 min, then 30% ice-cold sucrose containing MS-222 for 15 min for cryopreservation. Specimens were embedded sagittally in OCT medium (Sakura, Torrence, CA) and frozen in a dry ice/ethanol bath then stored at -80°C. Total RNA was extracted by cryosectioning each specimen until an approximately mid-sagittal plane was reached. A WPI Kwik-Fil borosilicate glass capillary (#M1B150-4, Hertfordshire, UK) was briefly soaked in DRNAse Free (Argos Technologies, Elgin, IL) and dried. Using these RNase-free capillaries, needles were pulled on a World Precision Instruments PUL1 (Sarasota, FL) to the following specifications: external tip diameter, $\approx$ 0.14mm; external diameter (unpulled region), $\approx$ 1.15 mm; sidewall thickness (unpulled region), $\approx$ 0.26 mm; internal diameter (unpulled region), $\approx$ 0.62 mm; length of tapered region, $\approx$ 3.6mm. Sectioned specimens, still affixed to an aluminum chuck, were mounted in dry ice and placed under

a dissecting microscope. The lung bud region was dissected out by plunging the tip of the glass needle into the tissue just dorsal to the wall between the sinus venosus and the atrium. Endoderm and mesoderm were not separated. For the stage-53 *Ambystoma mexicanum* lung sample, the animal was euthanized with MS-222 in DEPC H_2O, rinsed with DEPC H_2O, and then dissected using a sterile tungsten needle to remove lungs, but not trachea. Tissue was expelled or placed into an Eppendorf tube containing lysis buffer and TCEP from a Machery-Nagel Nucleospin RNA XS kit (#740902.10, Düren, Germany). Total RNA was isolated according to the kit protocol, but carrier RNA was not used.

Total RNA samples were analyzed on an Agillent Tape Station 2200 for RNA integrity and concentration using High Sensitivity R6K ScreenTape (Agilent, Santa Clara, CA). Using the IntegenX PrepX RNA-Seq Library Kit, 5 µl of total RNA (approximately 5 ng) was aliquotted for Single Primer Isothermal Amplification (SPIA) (IntegenX, Pleasanton, CA; Kurn et al., 2005) on an Apollo 324 robotic sample preparation system (WaferGen Biosystems, Fremont, CA), closely following kit instructions. Agencourt Ampure XP beads were used for magnetic purification steps (Beckman Coulter, Indianapolis, IN). Beads were kept at room temperature for 15 min before starting block setup. Beads were added last to the block, after a 30-sec vortex to fully resuspend them.

Following SPIA, concentration of samples was assessed using a Qubit 1.0 fluorometer (Invitrogen/Thermo Fisher Scientific, Grand Island, NY), high sensitivity dsDNA reagents (Molecular Probes/Thermo Fisher Scientific, Grand Island, NY), and Qubit Assay Tubes (Invitrogen). Samples were diluted to 20 µg/ml then sheared on a Covaris S220 Focused ultrasonicator (Covaris, Woburn, MA) using a 72-sec protocol, and targeting 220/320bp output. TapeStation HS D1K tape (Agilent) was used to examine sheared DNA for optimal

size range. BIOO Scientific NEXTflex DNA barcodes (#514102, Austin, TX) were diluted to 5µm and used in the IntegenX PrepX DNA Library ILM prep kit (#P003279, Pleasanton, CA). Library prep was performed on the Apollo 324 using the kit manufacturer's precise instructions. 5-µl aliquots of ≈ 2µg/ml samples were then subjected to limited cycle PCR amplification using NEB Next Master Mix (#M0541S, NEB, Ipswich, MA) and NEXTflex Primer Mix (BIOO Scientific) using the following cycle: denaturation at 98°C for 120 sec, then 5 cycles of 98°C 30 sec, 6°C 30 sec, 72°C 60 sec. Final extension for 5 min at 72°C. The Agilent Apollo 324 was used for cleanup of PCR samples using the built in PCR cleanup protocol and Agencourt Ampure XP beads. Following this, libraries were analyzed with Qubit 1.0, TapeStation and qPCR to assess library concentration, size and quality. Samples were diluted to 0.29 nM concentration each, then pooled. An initial 2 x 150bp Illumina HiSeq 2500 Rapid Run lane was used to assess clustering and relative concentrations (Illumina, San Diego, CA). Initial clustering was ≈ 25% of optimal, resulting in 49 million reads. A second lane of 2 x 150 bp Illumina HiSeq 2500 Rapid Run sequencing was run with adjustments based on the first sequenced flow cell in order to optimize clustering. This second lane yielded 184.5 million reads that passed filter for a total of 232.4 million reads (Table B.1).

B.2 RNA sequencing analysis

An overview of the complete workflow is provided in Figure B.1. All analyses were performed on Harvard's Odyssey Cluster, using Simple Linux Utility for Resource Management (SLURM). *Ambystoma mexicanum* stage-40 libraries are named Am40_2, Am40_3, and Am40_4 and the stage-53 lung library named Am53_lung. *Plethodon cinereus* stage-19

Table B.1: Read data from Illumina HiSeq 2500 2 x 150bp RapidRun sequencing.

Flow Cell 1

Adapter	Adapter Sequence	Sample	Reads (x10^6)	% of raw clusters per lane	% of >= Q30 Bases	Mean Quality Score
1	CGATGT	Pc19_5	6.202	6.49	93.86	35.58
2	TGACCA	Pc21_2	8.410	8.79	92.61	35.25
3	ACAGTG	Am40_2	9.550	9.99	94.81	35.85
4	GCCAAT	Am40_3	8.451	8.84	94.88	35.88
5	CAGATC	Pc19_8	14.748	15.42	93.97	35.57
6	CTTGTA	Pc19_9	8.198	8.57	94.57	35.79
7	ATCACG	Pc21_7	6.125	6.4	93.76	35.51
8	TTAGGC	Pc21_8	8.681	9.08	94.81	35.85
9	ACTTGA	Am40_4	10.544	11.03	94.36	35.72
10	GATCAG	Am53	8.095	8.47	94.4	35.7
NA		undetermined indices	6.621	6.92	90.12	34.08

Total Reads	Pass Filter Reads	% PF
49010104	47812292	95.6604

Flow Cell 2

Adapter	Adapter Sequence	Sample	Reads (x10^6)	% of raw clusters per lane	% of >= Q30 Bases	Mean Quality Score
1	CGATGT	Pc19_5	31.899	8.64	81.75	32.39
2	TGACCA	Pc21_2	32.672	8.85	79.76	31.87
3	ACAGTG	Am40_2	33.821	9.16	82.78	32.7
4	GCCAAT	Am40_3	32.108	8.7	82.7	32.68
5	CAGATC	Pc19_8	54.821	14.85	81.69	32.34
6	CTTGTA	Pc19_9	31.816	8.62	82.92	32.72
7	ATCACG	Pc21_7	23.553	6.38	82.09	32.44
8	TTAGGC	Pc21_8	30.772	8.34	82.31	32.61
9	ACTTGA	Am40_4	36.901	10	82.16	32.5
10	GATCAG	Am53	27.258	7.38	81.53	32.34
NA		undetermined indices	33.479	9.07	62.75	26.66

Total Reads	Pass Filter Reads	% PF
201665216	184551808	90.9296

Total Reads (Flow cells 1 and 2)	Total PF Reads (Flow cells 1 and 2)	% PF (Flow cells 1 and 2)
250675320	232364100	92.6952442

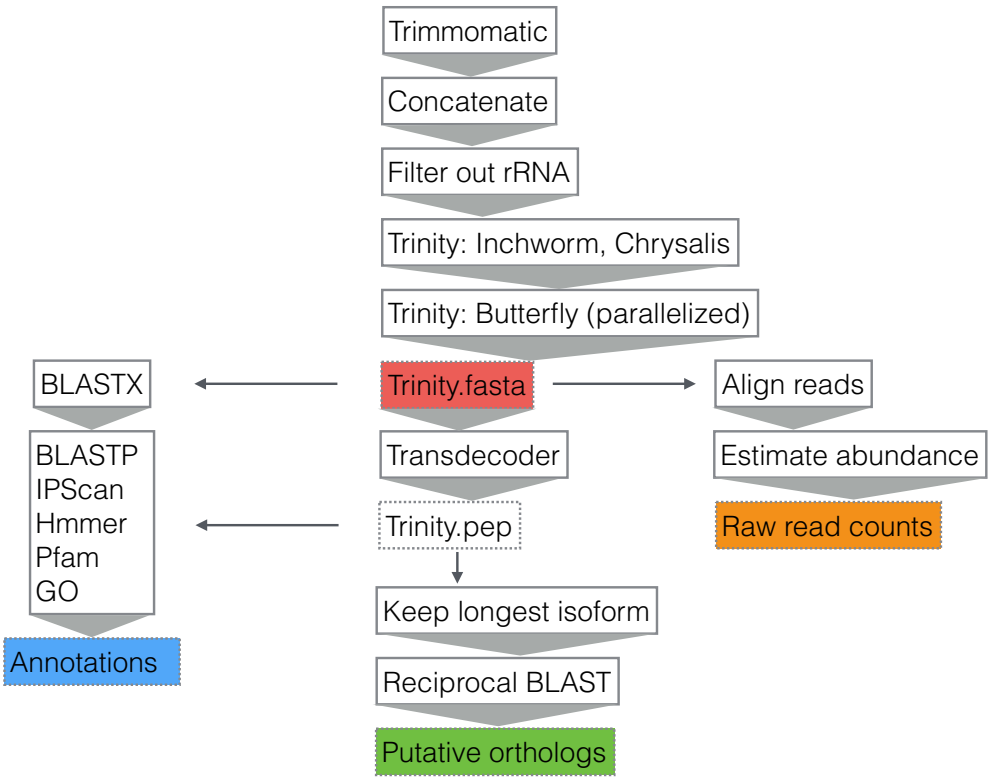


Figure B.1: Bioinformatic workflow for analysis of lung rudiment transcriptomes. Boxes with dotted outlines indicate databases, while solid outlines indicate processes. Colored databases were assembled into a relational database organized by orthology to assess differential expression between species.

libraries are named Pc19_5, Pc19_8, and Pc19_9. *Plethodon cinereus* stage-21 libraries are named Pc21_2, Pc21_7, and Pc21_8. Sequences were moved to a scratch drive and processed as follows:

1. Reads were trimmed with Trimmomatic (Bolger et al., 2014), using the following script, but changing the variable Am40_4 to the desired sequence name where appropriate. All 10 libraries run in parallel.

```
#!/bin/bash
```

```
#SBATCH --partition=serial_requeue
#SBATCH -t 60
#SBATCH -n 4
#SBATCH --mem=12000
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --output=40_4_trimmomatic_output.out

module load centos6/Trimmomatic-0.30

for SAMPLE in Am40_4; do
    echo ${SAMPLE}

PROJECT="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX"

java -jar $TRIMMOMATIC/trimmomatic-0.30.jar PE \
            -phred33 \
            $PROJECT/Sample_${SAMPLE}_*.R1.fastq.gz \
            $PROJECT/Sample_${SAMPLE}_*.R2.fastq.gz \
            $PROJECT/trimmed_fastq/${SAMPLE}.R1.paired.fastq.gz \
            $PROJECT/trimmed_fastq/${SAMPLE}.R1.unpaired.fastq.gz \
            $PROJECT/trimmed_fastq/${SAMPLE}.R2.paired.fastq.gz \
            $PROJECT/trimmed_fastq/${SAMPLE}.R2.unpaired.fastq.gz \
            ILLUMINACLIP:$TRIMMOMATIC/adapters/TruSeq3-PE.fa:2:30:10 \
            SLIDINGWINDOW:40:30 MINLEN:30
done
```

2. Reads from multiple flow cells were concatenated.

```
#!/bin/bash
#SBATCH --partition=serial_requeue
#SBATCH -t 600
#SBATCH -n 8
#SBATCH --mem=24000
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --job-name=concatenate
#SBATCH --output=concatentate.out

RUN1="/n/regal/mcz_users/zlewis/140523_D00365_0220_AH9N4DADXX/
    BclToFastq_Lane2_Indexlength6_Run1/Project_H9N4DADXX/trimmed_fastq"
RUN2="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX/trimmed_fastq"

cd $RUN2
cd ..
mkdir -p Run1+2_concat
```

```
for SAMPLE in Pc19_5 Pc19_8 Pc19_9 Pc21_2 Pc21_7 Pc21_8 Am40_2 Am40_3 Am40_4
    Am53_lung; do
     echo ${SAMPLE}
cd $RUN2
gunzip -c $RUN1/${SAMPLE}.R1.paired.fastq.gz $RUN2/${SAMPLE}.R1.paired.fastq.gz >
    ../Run1+2_concat/${SAMPLE}.R1.paired.concat.fastq
gunzip -c $RUN1/${SAMPLE}.R1.unpaired.fastq.gz $RUN2/${SAMPLE}.R1.unpaired.fastq.gz
     > ../Run1+2_concat/${SAMPLE}.R1.unpaired.concat.fastq
gunzip -c $RUN1/${SAMPLE}.R2.paired.fastq.gz $RUN2/${SAMPLE}.R2.paired.fastq.gz >
    ../Run1+2_concat/${SAMPLE}.R2.paired.concat.fastq
gunzip -c $RUN1/${SAMPLE}.R2.unpaired.fastq.gz $RUN2/${SAMPLE}.R2.unpaired.fastq.gz
     > ../Run1+2_concat/${SAMPLE}.R2.unpaired.concat.fastq

done
```

3. Reads from multiple libraries within each species were concatenated prior to *de novo* assembly.

```
#!/bin/bash
#SBATCH --partition=serial_requeue
#SBATCH -t 120
#SBATCH -n 8
#SBATCH --mem=24000
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --job-name=concatenate
#SBATCH --output=concatenate_samples.out

RUN2="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX"
CONCAT="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX/Run1+2_concat"

mkdir -p $RUN2/Pc.concat
mkdir -p $RUN2/Am.concat

cd $RUN2/Run1+2_concat

cat $CONCAT/Pc{19_5,19_8,19_9,21_2,21_7,21_8}.R1.paired.concat.fastq $CONCAT/Pc{19
    _5,19_8,19_9,21_2,21_7,21_8}.R1.unpaired.concat.fastq > $RUN2/Pc.concat/Pc.R1.p+
    s.fastq
cd $RUN2/Run1+2_concat
cat $CONCAT/Pc{19_5,19_8,19_9,21_2,21_7,21_8}.R2.paired.concat.fastq $CONCAT/Pc{19
    _5,19_8,19_9,21_2,21_7,21_8}.R2.unpaired.concat.fastq > $RUN2/Pc.concat/Pc.R2.p+
    s.fastq
cd $RUN2/Run1+2_concat
cat $CONCAT/Am{40_2,40_3,40_4,53_lung}.R1.paired.concat.fastq $CONCAT/Am{40_2,40_3
    ,40_4,53_lung}.R1.unpaired.concat.fastq > $RUN2/Am.concat/Am.R1.p+s.fastq
cd $RUN2/Run1+2_concat
```

```
cat $CONCAT/Am{40_2,40_3,40_4,53_lung}.R2.paired.concat.fastq $CONCAT/Am{40_2,40_3
    ,40_4,53_lung}.R2.unpaired.concat.fastq > $RUN2/Am.concat/Am.R2.p+s.fastq
```

4. Ribosomal rRNA reads were removed using Bowtie (Langmead et al., 2009). I constructed a BLAST database composed of deposited 12s, 16s, 18s, and 28s rRNA sequences on GenBank for both *Ambystoma mexicanum* and *Plethodon cinereus*. After the first lane of sequencing I also identified additional and/or longer rRNA sequences for each species and incorporated these into my BLAST database.

```
BLAST="/Users/Zack/ncbi-blast-2.2.29+/bin"

$BLAST/blastn -query Am_rRNA.fa -db Pc_Trinity5_Blast -out Pc_rRNA_hits.fa

$BLAST/blastn -query Am_rRNA.fa -db Am_Trinity4_Blast -out Am_rRNA_hits.fa
```

I assembled rRNA sequences into two FASTA files: Pc_rRNA_2.fa and Am_rRNA_2.fa. I placed these FASTA files into Pc.concat and Am.concat directories, respectively. In an interactive session from the directory above the species directory, I ran the following script, adapted from a script by Bob Freeman:

```
srun -n 1 --mem=2000 -p interact --pty bash
module load centos6/bowtie-1.0.0

for SPECIES in Pc Am; do
    echo ${SPECIES}
    cd /n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
        Project_H9YRCADXX/${SPECIES}.concat
    # make your output directories
    mkdir -p ./clean_noBlack/black_seqs
    mkdir -p ./clean_noBlack/clean_logs
    # build bowtie index
    bowtie-build ${SPECIES}_rRNA_2.fa ${SPECIES}_rRNA
    AXO_INDEX=${PWD}/${SPECIES}_rRNA
    sbatch -p serial_requeue -t 120 \
      -J blacklist_${SPECIES} \
      -n 4 -N 1 \
      --mem=8000 \
      --wrap="bowtie -p 4 -n 3 -q \
       --al ./clean_noBlack/black_seqs/${SPECIES}.fastq \
       --un ./clean_noBlack/${SPECIES}.clean.fastq \
```

```
    $AXO_INDEX \
    -1 ${SPECIES}.R1.p+s.fastq -2 ${SPECIES}.R2.p+s.fastq \
    > ./clean_noBlack/clean_logs/${SPECIES}.bowtie_out.txt"
done
```

All reads that did not match a known rRNA sequence were placed in the clean_noBlack directories. Resulting output indicates that, for instance, 7.74% of all *A. mexicanum* reads were rRNA. See below:

```
A. mexicanum:
# reads processed: 69635827
# reads with at least one reported alignment: 5389994 (7.74%)
# reads that failed to align: 64245833 (92.26%)
Reported 5389994 paired-end alignments to 1 output stream(s)

P. cinereus:
# reads processed: 105143462
# reads with at least one reported alignment: 2119068 (2.02%)
# reads that failed to align: 103024394 (97.98%)
Reported 2119068 paired-end alignments to 1 output stream(s)
```

5. Trinity *de novo* assembly, part 1 (Inchworm and Chrysalis) (Grabherr et al., 2011; Haas et al., 2013). This is a representative script for *Plethodon cinereus*, which was used to construct the final assembly (Pc_Trinity6 script #1). The Trinity job was stopped before butterfly and then resubmitted as a parallelized job using --grid_conf option. This parallelized job needs only a couple CPUs and low memory, so job submission parameters can be reset to total memory 10G, --JM memory to 8G, CPU to 1 and time-24:00. The script is the same for *A. mexicanum*, only substituting variable names.

```
#!/bin/bash
#SBATCH --partition=general
#SBATCH --time=7-0
#SBATCH --mem=210000
#SBATCH --nodes=1
#SBATCH --ntasks=8
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
```

```
#SBATCH --output=pc_trinity6.out
#SBATCH --job-name=pc_trinity6

PROJECT="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX"

mkdir -p $PROJECT/Pc_Trinity6_output

lfs setstripe -c 16 $PROJECT/Pc_Trinity6_output

module load centos6/samtools-0.1.19
module load centos6/bowtie2-2.1.0
module load centos6/trinityrnaseq_r20140413

Trinity   --seqType fq --JM 200G \
          --left $PROJECT/Pc.concat/clean_noBlack/Pc.clean_1.fastq \
          --right $PROJECT/Pc.concat/clean_noBlack/Pc.clean_2.fastq \
          --output $PROJECT/Pc_Trinity6_output \
          --min_kmer_cov 2 \
          --CPU 8 \
          --no_run_butterfly
```

6. Trinity *de novo* assembly, part 2 (Butterfly). (Pc_Trinity6 script #2.) I placed this grid configuration file in my working directory and saved it as gridconfig.txt. Again, the same script was used for *A. mexicanum*, only substituting variable names.

```
#-----------------------------------------------------------------
# grid type:
grid=SLURM
# template for a grid submission
cmd=sbatch -p serial_requeue --mem=10000 --time=2:00:00
# number of grid submissions to be maintained at steady state by the Trinity
    submission system
max_nodes=1000
# number of commands that are batched into a single grid submission job.
cmds_per_node=30
#-----------------------------------------------------------------
```

The Butterfly script is as follows:

```
#!/bin/bash
#SBATCH --partition=general
#SBATCH --time=7-0
#SBATCH --mem=10000
#SBATCH --nodes=1
#SBATCH --ntasks=1
#SBATCH --mail-type=END
```

```
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --output=am_trinity5_grid.out
#SBATCH --job-name=am_trinity5_grid

PROJECT="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX"

module load centos6/samtools-0.1.19
module load centos6/bowtie2-2.1.0
module load centos6/trinityrnaseq_r20140413

Trinity   --seqType fq --JM 8G \
          --left $PROJECT/Am.concat/clean_noBlack/Am.clean_1.fastq \
          --right $PROJECT/Am.concat/clean_noBlack/Am.clean_2.fastq \
          --output $PROJECT/Am_Trinity5_output \
          --grid_conf $PROJECT/gridconfig.txt \
          --CPU 1
```

Then I started the butterfly script as a job dependency based on completion of the first script.

```
sbatch --dependency=afterok:JOB_ID pc_trinity6_grid.sbatch
```

7. Resulting assembly statistics were compiled by running trinity stats (Grabherr et al., 2011).

```
################################
#Plethodon cinereus, Assembly 6
################################

################################
## Counts of transcripts, etc.
################################
Total trinity 'genes': 1145387
Total trinity transcripts: 1469859
Percent GC: 42.46

########################################
Stats based on ALL transcript contigs:
########################################

Contig N10: 1163
Contig N20: 823
Contig N30: 659
Contig N40: 550
Contig N50: 466
```

```
Median contig length: 335
Average contig: 426.49
Total assembled bases: 626885048

#######################################################
## Stats based on ONLY LONGEST ISOFORM per 'GENE':
#######################################################

Contig N10: 1065
Contig N20: 771
Contig N30: 623
Contig N40: 522
Contig N50: 443

Median contig length: 324
Average contig: 410.05
Total assembled bases: 469663025

###################################
###Ambystoma mexicanum, Assembly 6
###################################

#################################
## Counts of transcripts, etc.
#################################
Total trinity 'genes': 1071122
Total trinity transcripts: 1296815
Percent GC: 43.49

########################################
Stats based on ALL transcript contigs:
########################################

    Contig N10: 1115
    Contig N20: 808
    Contig N30: 656
    Contig N40: 552
    Contig N50: 469

    Median contig length: 336
    Average contig: 425.64
    Total assembled bases: 551972598

#####################################################
## Stats based on ONLY LONGEST ISOFORM per 'GENE':
#####################################################

    Contig N10: 1032
    Contig N20: 765
    Contig N30: 626
    Contig N40: 527
    Contig N50: 448
```

```
Median contig length: 326
Average contig: 410.90
Total assembled bases: 440125025
```

Table B.2: Final assembly statistics.

	Total Trinity 'genes'	Total Trinity transcripts	Contig N50*	Median contig length*	Average contig length*	Total assembled bases*
Plethodon cinereus assembly 6	1,145,387	1,469,859	466	335	426.49	626,885,048
Ambystoma mexicanum assembly 6	1,071,122	1,296,815	469	336	425.64	551,972,598

*Numbers based on all transcript contigs.

Interpretation of results: There are about three times as many transcripts as the previous assembly, which was constructed using the initial lane of sequencing data. I was expecting fewer, longer contigs based on better coverage. These additional contigs are likely basal transcription and make up most of the short and lower expressed sequences (Brian Haas, personal communication). To test the influence of read depth on assemblies I assembled transcriptomes at varying levels of completeness, as described below.

8. Assembly quality control. Trinity predicted a large number of contigs and genes for both species. To assess whether these predictions were a function of read depth (input reads) I performed independent assemblies on starting datasets of various completeness. I randomly subsetted the complete dataset of *Plethodon cinereus* reads at 10% increments from 10 to 100% and re-ran Trinity *de novo* assembly on each dataset. Resulting assemblies were analyzed for predicted numbers of genes, transcripts and average contig lengths (Fig. B.2).

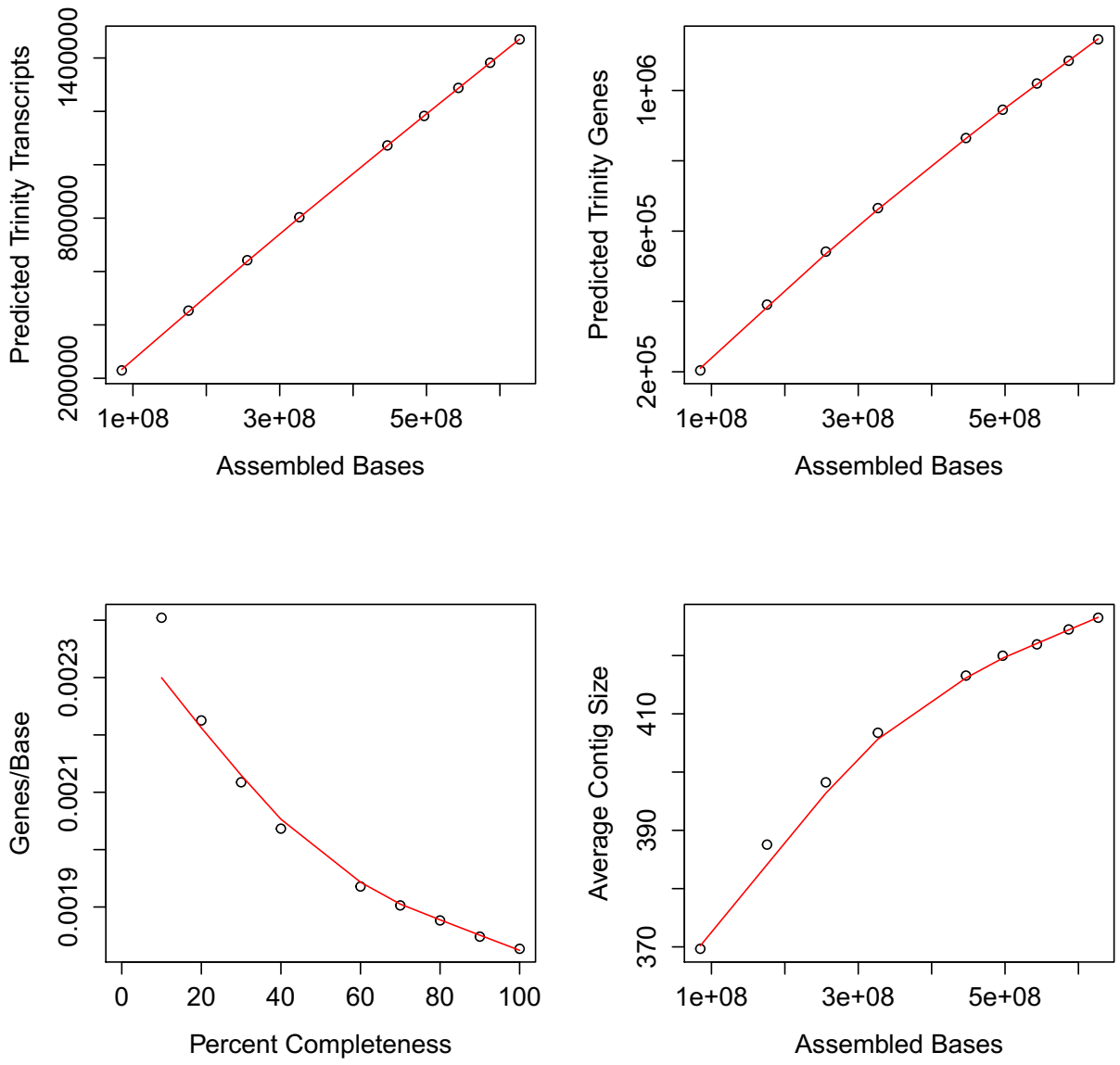


Figure B.2: Assembly statistics for *Plethodon cinereus* assemblies at various read depths. Random subsets of all *P. cinereus* reads were created at 10% intervals up to 100% ($\approx$ 626 million assembled bases). Libraries generated from these subsets were analyzed for number of predicted transcripts and genes, as well as for average contig size. The number of predicted transcripts or genes increases roughly linearly with library size. In contrast, the average contig size begins to plateau at high read depth, indicating that assemblies are nearing optimum. In addition, the assembled genes-per-base ratio begins to level off, indicating that adding additional bases (additional read depth) would negligibly increase transcriptome completeness.

9. I prepared my reference for alignment and mapping of reads:

```
#!/bin/bash
#SBATCH --partition=serial_requeue
#SBATCH -t 60
#SBATCH -n 2
#SBATCH --mem=8000
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --job-name=Pc_prepref

module load centos6/trinityrnaseq_r20140413
module load centos6/samtools-0.1.19
module load centos6/bowtie2-2.1.0
module load centos6/express-1.5.1
module load centos6/rsem-1.2.11

DIROUT="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX/Pc_Trinity6_output"

align_and_estimate_abundance.pl --transcripts $DIROUT/Trinity.fasta --est_method
    RSEM --aln_method bowtie --trinity_mode --prep_reference --output_dir $DIROUT
```

10. I next aligned reads and estimated abundance. I had already unzipped reads from each lane of sequencing and concatenated them together. They were in Run1+2_concat directory. For better speed, I disregarded the FOR loop and submit multiple scripts in parallel.

```
#!/bin/bash
#SBATCH --partition=serial_requeue
#SBATCH -t 300
#SBATCH -n 8
#SBATCH --mem=24000
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --job-name=Pc_align_est

module load centos6/trinityrnaseq_r20140413
module load centos6/samtools-0.1.19
module load centos6/bowtie2-2.1.0
module load centos6/express-1.5.1
module load centos6/rsem-1.2.11

cd /n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/Project_H9YRCADXX/
mkdir -p bowtie_rsem
```

```
for SAMPLE in Pc19_5 Pc19_8 Pc19_9 Pc21_2 Pc21_7 Pc21_8; do
    echo ${SAMPLE}
cd /n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/Project_H9YRCADXX/
    bowtie_rsem
mkdir -p ${SAMPLE}

DIRIN="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX/Pc_Trinity6_output"
READS="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX/Run1+2_concat"
DIROUT="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX/bowtie_rsem/${SAMPLE}"

align_and_estimate_abundance.pl \
     --transcripts $DIRIN/Trinity.fasta --seqType fq \
               --left $READS/${SAMPLE}.R1.paired.concat.fastq \
     --right $READS/${SAMPLE}.R2.paired.concat.fastq \
               --est_method RSEM --aln_method bowtie \
               --trinity_mode --output_dir $DIROUT \
     --output_prefix ${SAMPLE}

done
```

11. I looked at the number of expressed genes by opening an interactive node and running the count_features script within Trinity.

```
srun -n 2 --mem=8000 -p interact --pty bash

module load centos6/trinityrnaseq_r20140413

cd /n/sw/centos6/trinityrnaseq_r20140413/util/misc

srun count_features_given_MIN_FPKM_threshold.pl /n/regal/mcz_users/zlewis/
    BclToFastq_Lane1_Indexlength6_Run1/Project_H9YRCADXX/bowtie_rsem/Pc19_5/Pc19_5.
    genes.results >/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX/bowtie_rsem/Pc19_5/Pc19_5_cumul_counts.txt
```

An example using Pc19_8 library reveals the following (subset of the file):

```
neg_min_fpkm    num_features
-10    28892
-9     34033
-8     41177
-7     50326
-6     62788
-5     80509
-4     105563
```

```
-3      141193
-2      192463
-1      259068
0       1145387
```

Which indicates that 259,000+ genes are expressed over 1 FPKM (Fig. B.3). This implies that relatively stringent filtering can be used, perhaps even a cutoff of FPKM ≥10. I plotted the results for this one library in R using this code:

```
data = read.table("~/Documents/RStudio/RNA-seq/Pc19_8_cumul_counts.txt", header=T)
attach(data)
plot(neg_min_fpkm,num_features, xlim=c(-10,0), type='b')
```

12. To reduce the size of my library I filtered the fasta file by the FPKM values. First, I tried removing contigs with IsoPct < 1%, meaning that they represent at least 1% of the per-component expression value (essentially removing unsupported isoforms). Next, I filtered out transcripts below 1 FPKM. I also attempted filtering out transcripts below 5 FPKM. The first step involved copying isoform.results files to a new directory:

```
for SAMPLE in Pc19_5 Pc19_8 Pc19_9 Pc21_2 Pc21_7 Pc21_8; do
echo ${SAMPLE}
cp ${SAMPLE}/${SAMPLE}.isoforms.results ./isoforms.results_files/${SAMPLE}.isoforms
    .results
done

mkdir -p $PROJECT/Pc_Trinity6_filtered/removed_sequences
```

I then copied the perl script from the Trinity Utilities to Pc_Trinity6_filtered directory. I then ran the script. Below are the results set to filter at IsoPct >1%. The FPKM cutoff can be specified using the flag --fpkm_cutoff=5 with or without --isopct_cutoff=1.

```
#!/bin/bash
#SBATCH --partition=serial_requeue
#SBATCH -t 60
```

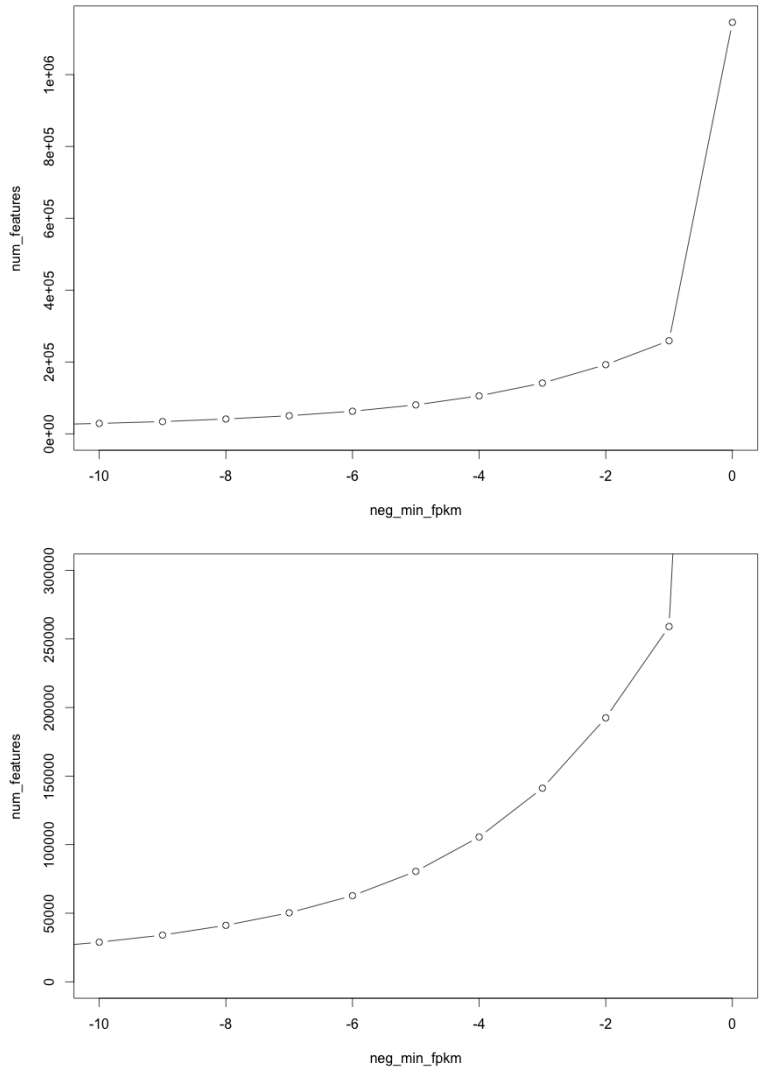


Figure B.3: The number of genes expressed at different FPKM thresholds. Top: Full chart. Bottom: The Y-axis scale has been adjusted to better illustrate the effects of a FPKM threshold greater than 1.

```
#SBATCH -n 8
#SBATCH -N 1
#SBATCH --mem=16000
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --job-name=filter_rsem
#SBATCH --output=filter_rsem_output

PROJECT="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX"
cd $PROJECT/Pc_Trinity6_filtered
```

Table B.3: Results of different filtering regimes on number of trinity products for *Plethodon cinereus* assembly 6.

	Total Trinity 'genes'	Total Trinity transcripts
Starting values	1,145,387	1,469,859
IsoPct >1	1,125,933	1,311,166
FPKM >1	863,282	1,019,404
FPKM >5; IsoPct >1	509,970	553,226

Table B.4: Results of filtering *Ambystoma mexicanum* assembly 6 for transcripts expressed over 1 FPKM.

	Total Trinity 'genes'	Total Trinity transcripts
Starting values	1,071,122	1,296,815
FPKM >1	851,037	964,863

```
module load centos6/BioPerl-1.6.1

perl filter_fasta_by_rsem_values.pl \
          -r $PROJECT/bowtie_rsem/isoforms.results_files/Pc19_5.isoforms.results \
          -r $PROJECT/bowtie_rsem/isoforms.results_files/Pc19_8.isoforms.results \
          -r $PROJECT/bowtie_rsem/isoforms.results_files/Pc19_9.isoforms.results \
          -r $PROJECT/bowtie_rsem/isoforms.results_files/Pc21_2.isoforms.results \
          -r $PROJECT/bowtie_rsem/isoforms.results_files/Pc21_7.isoforms.results \
          -r $PROJECT/bowtie_rsem/isoforms.results_files/Pc21_8.isoforms.results \
          --fasta=$PROJECT/Pc_Trinity6_output/Trinity.fasta \
          --output=$PROJECT/Pc_Trinity6_filtered/Pc_Trinity6_filtered.fasta \
          --filtered_output $PROJECT/Pc_Trinity6_filtered/removed_sequences/
              Pc_Trinity6_removed.fasta \
          --isopct_cutoff=1.00 \
          --log filter_fasta_by_rsem_values.log
```

IsoPct >1% did not affect transcriptome size markedly. Instead, decided to filter at ≥1 FPKM.

Bob Freeman's perl script fastq_stats.pl can be used to look at distribution of transcripts at different size bins.

https://github.com/macmanes/trinityrnaseq/blob/master/util/fastq_stats.pl

It is invoked as follows:

```
module load centos6/BioPerl-1.6.1

perl fastq_stats.pl -i Trinity.fasta -f
    "200,300,400,500,600,700,800,900,1000,1250,1500,2000,3000,4000,5000,7500,10000"

Resulting in:

Sequence stats:
Count1296815
Sum551972598
Mean425.637117090718

Min201
Max13867
Median336

Q0201
Q1250
Q2336
Q3501
Q413867

201(min)
<= 200.0:  0
<= 300.0: 536417
<= 400.0: 267203
<= 500.0: 168285
<= 600.0: 110811
<= 700.0: 71653
<= 800.0: 45792
<= 900.0: 29052
<= 1000.0: 18857
<= 1250.0: 24815
<= 1500.0: 10438
<= 2000.0:   8194
<= 3000.0:   4080
<= 4000.0:   797
<= 5000.0:   245
<= 7500.0:   157
<= 10000.0:  12
```

This script was employed to plot the distribution of contig sizes before and after filtering in R (Fig. B.4). This example shows *Ambystoma mexicanum*.

```
before <-c
    (536417,267203,168285,110811,71653,45792,29052,18857,24815,10438,8194,4080,797,245,157,12)

after <-c
    (297951,225605,149409,100203,64929,41664,25054,16803,21707,8964,6987,3499,710,218,143,10)

barplot(before, col="red", border="red",density=50, ylim=c(0,600000))
```

```
par(new=TRUE)
barplot(after, col="blue", border="blue",density=50, ylim=c(0,600000), xlab="Size (
    bp)", ylab=ŰCountsŰ, names.arg=c
    (300,400,500,600,700,800,900,1000,1250,1500,2000,3000,4000,5000,7500, 10000))
```

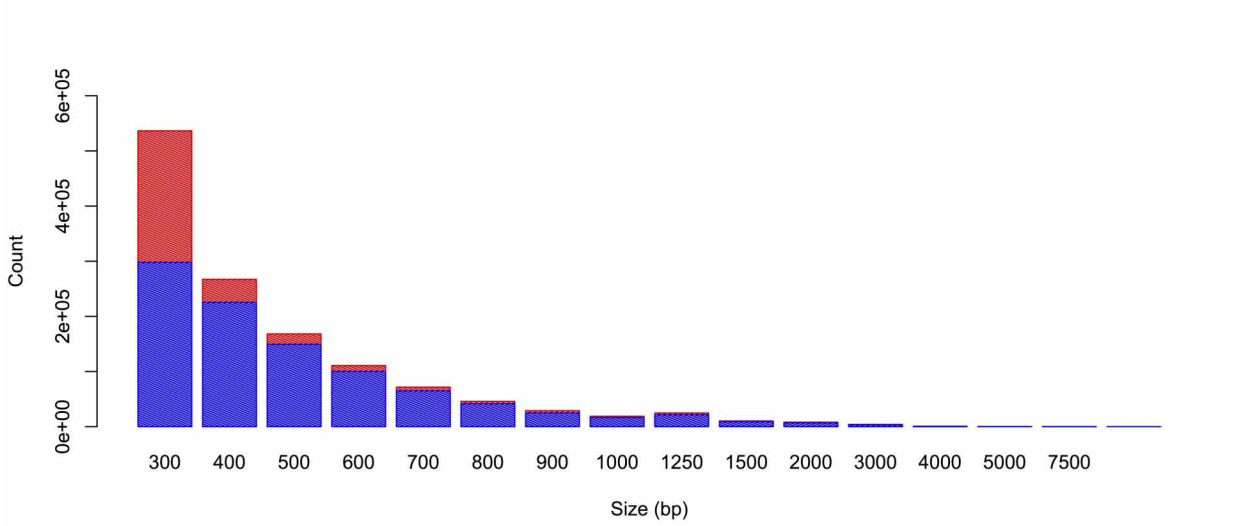


Figure B.4: *Ambystoma mexicanum* contig sizes before and after filtering by ≥1 FPKM. Before filtering=red; after filtering=blue. Note the marked reduction in contigs between 200 – 300 bp. Also note that bin size shifts after 1000 bp, resulting in a small hump.

13. Reads were re-mapped onto the reference. I used the same script as in step 9, but with assemblies that had been filtered by FPKM ≥1.

14. I translated my assembly using Transdecoder (Grabherr et al., 2011). I used grep determine what proportion of translations are complete (meaning that they have canonical start and stop codons) by searching for instances of "type:complete".

```
#!/bin/bash
#SBATCH --partition=serial_requeue
#SBATCH -t 300
#SBATCH -n 4
#SBATCH --mem=32000
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --job-name=Pc_transdecoder
```

```
module load centos6/TransDecoder_r20131117

PROJECT="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX"
WORKDIR="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX/Pc_Transdecoder"

TransDecoder    -t $PROJECT/Pc_Trinity6_filtered/Pc_Trinity6_filtered_fpkm1.fasta \
                  -m 50 \
                  --workdir $WORKDIR \
                  --retain_long_orfs 150
```

15. Initial annotation was performed using BLASTX and then BLASTP. Only the BLASTX script example is shown. In this example, I used a custom reference database composed of proteins from several chordates and assembled together by Bob Freeman.

```
#!/bin/bash
#SBATCH --partition=general
#SBATCH -t 3000
#SBATCH -n 16
#SBATCH -N 1
#SBATCH --mem=32000
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --job-name=Pc_blastx

module load centos6/ncbi-blast-2.2.29+

blastx -query Pc_Trinity6_filtered_fpkm1.fasta -db /n/regal/temp/
    model_chordate_proteins -num_threads 8 -max_target_seqs 1 -outfmt 6 > blastx.
    outfmt6blastx
```

16. In order to determine orthology between species, I needed to have only one translation per gene. To eliminate duplicates I selected only the longest translation in instances where there were two or more predicted translations per contig. I used a script from Bob Freeman to do this.

```
#!/usr/bin/perl -w

use strict;
use Bio::SeqIO;
```

```
use Bio::Seq;
use Bio::DB::Fasta;
use English   qw ( -no_match_vars );
use Text::Wrap;
$Text::Wrap::columns = 60;

local $OUTPUT_AUTOFLUSH             =      1;
my %sequences;
my %descriptions;

my $infile = shift || die "Error! 1st argument must be input fasta file!\n";

my $seqio = Bio::SeqIO->new( -file  => "<$infile",
                           -format   => 'fasta');

while (my $seqobj = $seqio->next_seq()) {
    #
    my $seq_id = $seqobj->display_id();
    my $sequence = $seqobj->seq();
    my $description = $seqobj->desc();

    # if entry not present, store sequence
    # if present, only store sequence if longer
    if (!defined $sequences{$seq_id}) {
        $sequences{$seq_id} = $sequence;
        $descriptions{$seq_id} = $description;
    } else {
        if (length ($sequence) > length ($sequences{$seq_id})) {;
            $sequences{$seq_id} = $sequence;
            $descriptions{$seq_id} = $description;
        }
    }
}
foreach my $key (keys %sequences) {
    print ">$key $descriptions{$key}\n";
    print wrap('', '', $sequences{$key});
    #print $sequences{$key};
    print "\n";
}
exit;

# END OF PROGRAM
```

This script was run as follows:

```
PROJECT="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX"
module load centos6/BioPerl-1.6.1

perl keep_longest_of_dupe_seqID_2.pl $PROJECT/keep_longest/
    Pc_Trinity6_filtered_fpkm1.fasta.transdecoder.fa > \
$PROJECT/keep_longest/Pc_Trinity6.fasta.transdecoder.longest.fa
```

17. To find orthologs between *A. mexicanum* and *P. cinereus* I used a reciprocal best hit BLAST approach (RBH). Each transcriptome was BLASTed against the other. Orthology was predicted for two contigs if the top BLAST hit from a species 1 vs. species 2 comparison was the same as the top BLAST hit from a species 2 vs. species 1 comparison.

```
################################
##Reciprocal BLAST:
################################

mkdir -p reciprocal_blast

module load centos6/ncbi-blast-2.2.29+

makeblastdb -in ../keep_longest/Pc_Trinity6.fasta.transdecoder.longest.fa -dbtype
    prot -out PcBlastDb -title ÙPcBlastDbÚ

makeblastdb -in ../keep_longest/Am_Trinity6.fasta.transdecoder.longest.fa -dbtype
    prot -out AmBlastDb -title ÙAmBlastDbÚ

##################
##Pc vs. Am:
##################

#!/bin/bash
#SBATCH --partition=general
#SBATCH -t 7-0
#SBATCH -n 16
#SBATCH -N 1
#SBATCH --mem=16000
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --job-name=Pc_vs_Am

PROJECT="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX"

module load centos6/ncbi-blast-2.2.29+

blastp -query $PROJECT/keep_longest/Pc_Trinity6.fasta.transdecoder.longest.fa -db
    AmBlastDb -num_threads 12 -outfmt 6 > 1v2.blastp

#Note: run job from within the blastdb folder

##################
##Am vs. Pc:
```

```
##################

#!/bin/bash
#SBATCH --partition=general
#SBATCH -t 7-0
#SBATCH -n 16
#SBATCH -N 1
#SBATCH --mem=16000
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --job-name=Am_vs_Pc

PROJECT="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX"

module load centos6/ncbi-blast-2.2.29+

blastp -query $PROJECT/keep_longest/Am_Trinity6.fasta.transdecoder.longest.fa -db
    PcBlastDb -num_threads 12 -outfmt 6 > 2v1.blastp
```

Reciprocal BLAST took a little over two days and generated huge datasets (3.5—4.5 gb). I next found the intersection of the two BLAST databases using a script designed by Bob Freeman and Leonid Peshkin (Wühr et al., 2014). I edited the script to stop at step 10.

```
module load centos6/BioPerl-1.6.1

perl genesym_assignment_parse_edited.pl -b ./1v2.blastp -c ./2v1.blastp -s trial1
    --Evalue 1e-5

wc RBH.tsv -l
16958 RBH.tsv
```

The reciprocal best-hit database (RBH) has 16,958 lines, corresponding to 16,958 putative orthologs between *A. mexicanum* and *P. cinereus*.

18. I computed expression values (abundance estimation) using native Trinity scripts (Grabherr et al., 2011).

```
#!/bin/bash
#SBATCH --partition=serial_requeue
```

```
#SBATCH -t 60
#SBATCH -n 4
#SBATCH --mem=12000
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --job-name=est_to_matrix

RSEM="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX/bowtie_rsem_filtered"

module load centos6/trinityrnaseq_r20140413
module load centos6/R-3.0.2

abundance_estimates_to_matrix.pl --est_method RSEM \
             --out_prefix Plethodon_matrix \
             $RSEM/Pc19_5/Pc19_5.genes.results \
             $RSEM/Pc19_8/Pc19_8.genes.results \
             $RSEM/Pc19_9/Pc19_9.genes.results \
             $RSEM/Pc21_2/Pc21_2.genes.results \
             $RSEM/Pc21_7/Pc21_7.genes.results \
             $RSEM/Pc21_8/Pc21_8.genes.results
```

```
#!/bin/bash
#SBATCH --partition=serial_requeue
#SBATCH -t 60
#SBATCH -n 4
#SBATCH --mem=12000
#SBATCH --mail-type=END
#SBATCH --mail-type=FAIL
#SBATCH --mail-user=zlewis@oeb.harvard.edu
#SBATCH --job-name=est_to_matrix_2

RSEM="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX/bowtie_rsem_filtered"

module load centos6/trinityrnaseq_r20140413
module load centos6/R-3.0.2

abundance_estimates_to_matrix.pl --est_method RSEM \
             --out_prefix Axolotl_matrix \
             $RSEM/Am40_2/Am40_2.genes.results \
             $RSEM/Am40_3/Am40_3.genes.results \
             $RSEM/Am40_4/Am40_4.genes.results \
             $RSEM/Am53_lung/Am53_lung.genes.results
```

19. I further annotated with BLAST, using the swissprot database. Here is an example of constructing a new swissprot BLAST database and running BLASTX and BLASTP against swissprot.

```
wget http://sourceforge.net/projects/trinotate/files/TRINOTATE_RESOURCES/20140708/
    uniprot_sprot.fasta.gz
gunzip uniprot_sprot.fasta.gz
module load centos6/ncbi-blast-2.2.29+
makeblastdb -in uniprot_sprot.fasta -dbtype prot
```

Here is an example BLAST command for *P. cinereus* BLASTP against swissprot:

```
PROJECT="/n/regal/mcz_users/zlewis/140523_D00365_0220_AH9N4DADXX/
    BclToFastq_Lane2_Indexlength6_Run1/Project_H9N4DADXX"
SWISSPROT="/n/regal/mcz_users/zlewis/swissprot"

module load centos6/ncbi-blast-2.2.29+

blastp -query $PROJECT/keep_longest/Pc_Trinity6.fasta.transdecoder.longest.fa \
     -db $SWISSPROT/uniprot_sprot.fasta -num_threads 20 -max_target_seqs 1 -outfmt
         5 > Pc_blastp_swissprot.xml
```

Note that BLASTX jobs need a significant amount of time, processors and memory to complete. For instance, BLASTX was run at 7 days, 24 processors and 40 GB of memory. Following job completion, XML files were parsed using the parse_multiblast_file.pl script from Wuhr et al. (2014).

```
screen -S parse
srun -n 4 --mem=12000 -p interact --pty bash
module load centos6/BioPerl-1.6.1
perl parse_multiblast_file_lite.pl -i Am_blastx_swissprot.xml -x

control-AD
```

20. I constructed a relational database combining multiple discrete databases in Filemaker Pro 13. All databases were exported as comma-separated value files and imported into Filemaker Pro as individual tables. Reciprocal best-hit BLAST results were used to assign orthologous contigs (isoforms have been collapsed) from each species a unique identifier. After linking orthologs by an identifier, annotations, peptide sequence, nucleotide sequence, and expression data could be related to the contig ID for each

ortholog (Fig. B.5). I next constructed an interface with which to interact with the database and perform basic graphing functions (Fig. B.6). New datasets can be easily exported from this relational database in order to perform cross-species differential expression analysis in R statistical software.

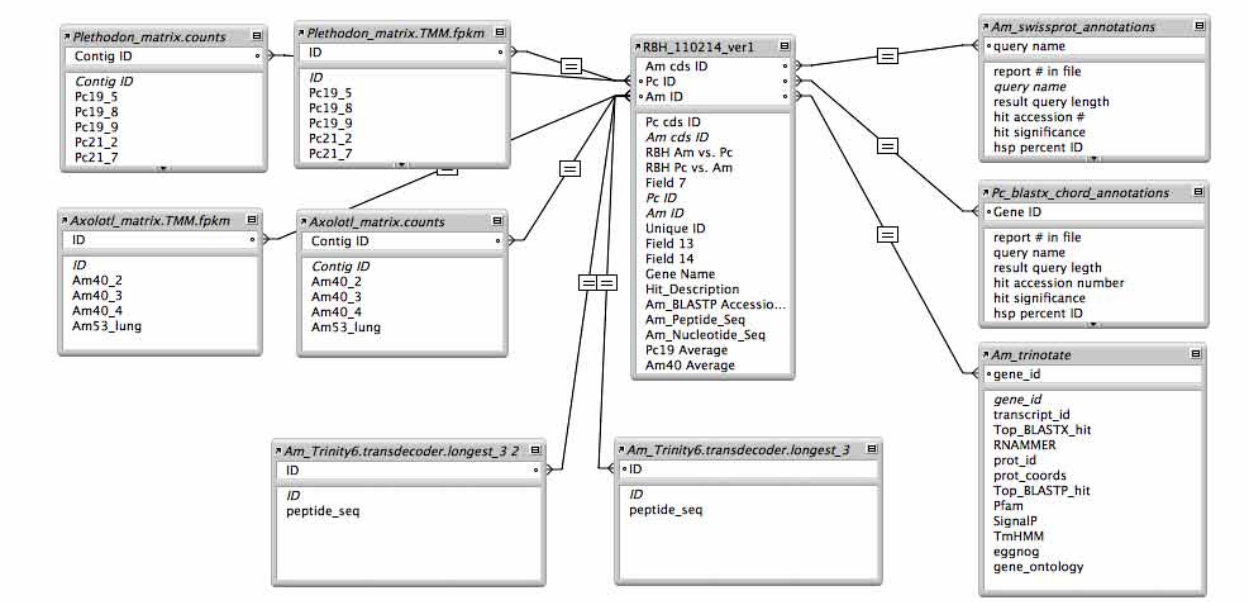


Figure B.5: Relational database structure in FileMaker Pro 13. The reciprocal best-hit blast determined orthologs (RBH) (in the center) are utilized to integrate other databases. Mapped reads (left), peptide sequences (center bottom) and annotations (right) are related through the RBH database.

21. Further annotation with Pfam and Interproscan. Pfam annotation was completed according to the instructions in (Haas et al., 2013). Interproscan was performed only

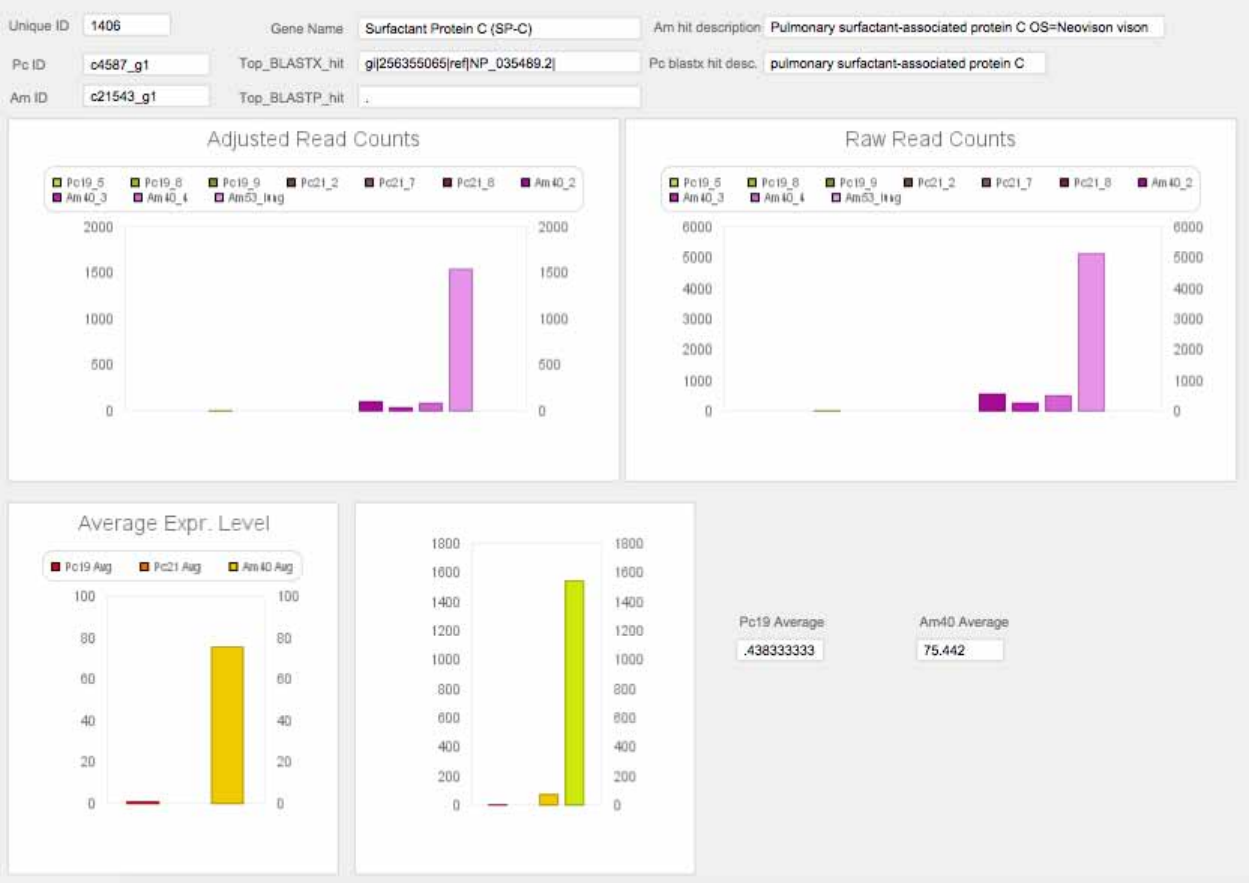


Figure B.6: Representative image of the FileMaker Pro interface. Genes can be searched for using annotations, unique identifiers, or contig IDs. Mapped reads are then displayed by library.

on peptides from *A. mexicanum* within the relational blast database. To convert tabular output from Filemaker Pro to a fasta file I wrote this script using SeqIO.

```
########################################
##saved following script as convert4.py:
########################################

from Bio import SeqIO

input_handle = open("Am_longest_fasta_seqs_in_RBH.tab", "rU")
output_handle = open("Am_longest_fasta_seqs_in_RBH.fa", "w")

sequences = SeqIO.parse(input_handle, "tab")
count = SeqIO.write(sequences, output_handle, "fasta")
```

```
output_handle.close()
input_handle.close()

print "Converted %i records" % count
```

I downloaded and extracted InterProScan standalone (Quevillon et al., 2005). I then adjusted the hyperlink in the configuration as specified. I next ran InterProScan with the following script:

```
#!/bin/bash
#SBATCH --partition=serial_requeue
#SBATCH -t 1200
#SBATCH -n 16
#SBATCH -N 1
#SBATCH --mem=32000
#SBATCH --job-name=ipscan
#SBATCH -o ./ips_RBH.out
#SBATCH -e ./ips_RBH.err

WORKDIR="/n/regal/mcz_users/zlewis/BclToFastq_Lane1_Indexlength6_Run1/
    Project_H9YRCADXX"
mkdir -p ./Am_RBH_IPS

srun interproscan.sh -b ./Am_RBH_IPS/Am_Trinity6_RBH_IPS -f tsv,xml,gff3,html,svg \
     --goterms -i $WORKDIR/keep_longest/Am_longest_fasta_seqs_in_RBH_2_noasterisk.fa
          -appl PfamA,PRINTS -iprlookup \
     -goterms --pathways -t p
```

InterProScan results were then loaded into Blast2GO along with the BLASTX XML results. Both annotations were merged into the database and were subsequently used in gene ontology (GO) analysis.

22. I analyzed differential expression with DESeq (Anders and Huber, 2010) in RStudio. Comments and instructions are in the R script, below.

```
#reading in data
count.data2<-read.table("~/Documents/RStudio/RNA-Seq/Combined_Counts_2.tab", header
    =TRUE, row.names=1, sep="\t")
head(count.data)

#rename columns
```

```
names(count.data2)[names(count.data2)=="X10.00"] <- "Pc19_8"
names(count.data2)[names(count.data2)=="X0.00"] <- "Pc19_9"
names(count.data2)[names(count.data2)=="X0.00.1"] <- "Pc21_2"
names(count.data2)[names(count.data2)=="X1.00"] <- "Pc21_7"
names(count.data2)[names(count.data2)=="X0.00.2"] <- "Pc21_8"
names(count.data2)[names(count.data2)=="X0.00.3"] <- "Am40_2"
names(count.data2)[names(count.data2)=="X4.00"] <- "Am40_3"
names(count.data2)[names(count.data2)=="X0.00.4"] <- "Am40_4"
names(count.data2)[names(count.data2)=="X1.00.1"] <- "Am53_lung"

#continue
data.design <- data.frame(row.names = colnames(count.data2), condition = c("1",
    "1", "1", "2", "2", "2", "3", "3", "3", "4"), libType = c("paired-end","paired-
    end","paired-end","paired-end","paired-end","paired-end","paired-end","paired-
    end","paired-end","paired-end"))
library(DESeq)
data <- as.matrix(count.data2)
storage.mode(data) = "integer"
conds <- factor( c("1", "1", "1", "2", "2", "2", "3", "3", "3", "4"))
cds <- newCountDataSet(data, conds)
head(cds)

#estimate the effective library size
cds <- estimateSizeFactors( cds )
sizeFactors(cds)

#The inference in DESeq relies on an estimation of the typical relationship between
     the dataÕs variance and their mean
#or, equivalently, between the dataÕs dispersion and their mean
#dispersion is the square of the coefficient of biological variation
cds <-estimateDispersions(cds)

#estimateDispersions fxn has several steps to fit a curve and give dispersions
#can examine the fit info using:
str( fitInfo(cds) )

#can plot dispersion estimates
plotDispEsts <- function( cds ) { plot(
  rowMeans( counts( cds, normalized=TRUE ) ), fitInfo(cds)$perGeneDispEsts,
  pch = '.', log="xy" )
  xg <- 10^seq( -.5, 5, length.out=300 )
  lines( xg, fitInfo(cds)$dispFun( xg ), col="red" )}
plotDispEsts( cds )

#Create a plotting function which takes your nbinomTest() object and the desired
    percent false discovery rate as arguments.
plotDE <- function( res, sig )
  plot(
    res$baseMean,
    res$log2FoldChange,
    log="x", pch=20, cex=.3,
    col = ifelse( res$padj < sig, "red", "black" ), xlab="Base Mean", ylab=
        expression(Log[2]~Fold~Change) )
```

```
#example standard test between time 1 and 2 in Plethodon:
res <- nbinomTest( cds, "1", "2" )
head(res)
plotDE(res, 0.1)

#filter for significant genes according to a FDR threshold
resSig <- res[ res$padj < 0.1, ]
head(resSig)

#look at the most strongly downregulated genes
head( resSig[ order( resSig$foldChange, -resSig$baseMean ), ] )

#write results to csv
write.csv( res, file="Pc19vsPc21_nbinom_test.csv" )

#example plot DE with significance cutoff of 0.05
plotDE(res2, 0.1)
plotDE(res3, 0.1)

#histogram of p-values to look at their distribution
hist(res2$pval, breaks=100, col="skyblue", border="slateblue", main="")
hist(res3$pval, breaks=100, col="skyblue", border="slateblue", main="")
#plotMA
plotMA(res)

#Multifactorial analysis
cdsFull <- newCountDataSet(data, data.design)
cdsFull <- estimateSizeFactors( cdsFull )
cdsFull <- estimateDispersions( cdsFull )
plotDispEsts( cdsFull )
fit0 <- fitNbinomGLMs( cdsFull, count ~ condition )
head(fit0)
select <- order(res$pval)[1:40]

# Making a heatmap
# Transform the count data
cdsBlind <- estimateDispersions( cds, method="blind" )
vsd <- getVarianceStabilizedData( cdsBlind )

#Order the genes by their significance values and display the top 1000 genes
select <- order(res$pval)[1:1000]

#Order the genes by their significance values and display the top 100 genes
select2 <- order(res$pval)[1:100]

#Set up the color pallet and plot the heat map
colors <- colorRampPalette(c("white","darkblue"))(100)
colors2 <- colorRampPalette(c("white","darkblue"))(1000)
heatmap( vsd[select,],col = colors, scale = "none")
heatmap( vsd[select,],col = colors2, scale = "none")

#pretty colors:
library("gplots")
library("RColorBrewer")
```

```
hmcol = colorRampPalette(brewer.pal(9, "GnBu"))(100)
heatmap( vsd[select,],col = hmcol, scale = "none")

#heatmap of the distance matrix reveals which samples are similar
cdsFullBlind <- estimateDispersions( cdsFull, method = "blind" )
vsdFull <- getVarianceStabilizedData( cdsFullBlind )
dists1 <- dist( t( vsdFull ) )
heatmap(as.matrix(dists1),symm=TRUE, scale="none", col=hmcol, margins=c(8,8),)

#Heatmap for 100 genes
heatmap( vsd[select2,],col = colors, scale = "none")

#Heatmap from rowMeans instead of res$pval
select4 <- order(rowMeans(counts(cds)), decreasing=TRUE)[1:100]
heatmap( vsd[select4,],col = hmcol, scale = "none", cexRow=0.50, cexCol=0.75)
#the above fxn stopped working, so switched to:
heatmap.2(exprs(vsd)[select4,], col = hmcol, trace="none", margin=c(7, 6))
#note: works when vsd <- getVarianceStabilizedData( cdsBlind )

#can do the same ordering trick on the res$pval heatmaps but the clustering changes
select5 <- order((res$pval), decreasing=TRUE)[1:500]
heatmap( vsd[select5,],col = hmcol, scale = "none", cexCol=0.8)

#plot the PCA; can specify number of the most variable genes to be used in
#calculating the PCA with ntop=x (default = 500)
vsd.2 = varianceStabilizingTransformation(cdsBlind)
plotPCA(vsd.2, intgroup=c("condition"))

###################################
#Used the general approach below to run analysis on all replicated samples (9 #
    libraries, not including A. mexicanum lung at stage 53, because it is not #
    replicated
###################################

count.data.subset <- count.data2[,1:9]
data.design.subset <- data.frame(row.names = colnames(count.data.subset), condition
     = c("1", "1", "1", "2", "2", "2", "3", "3", "3"), libType = c("paired-end","
    paired-end","paired-end","paired-end","paired-end","paired-end","paired-end","
    paired-end","paired-end"))
head(count.data.subset)
data.subset <- as.matrix(count.data.subset)
storage.mode(data.subset) = "integer"
conds.subset <- factor( c("1", "1", "1", "2", "2", "2", "3", "3", "3"))
cds.subset <- newCountDataSet(data.subset, conds.subset)
head(cds.subset)
cds.subset <- estimateSizeFactors( cds.subset )
cds.subset <-estimateDispersions(cds.subset)
plotDispEsts( cds.subset)
cdsFull.subset <- newCountDataSet(data.subset, data.design.subset)
cdsFull.subset <- estimateSizeFactors( cdsFull.subset )
cdsFull.subset <- estimateDispersions( cdsFull.subset )
plotDispEsts( cdsFull.subset )
fit7 <- fitNbinomGLMs( cdsFull.subset, count ~ condition )
head(fit7)
```

```
select <- order(res$pval)[1:100]
cdsBlind.subset <- estimateDispersions( cds.subset, method="blind" )
vsd.subset <- getVarianceStabilizedData( cdsBlind.subset )
heatmap( vsd.subset[select,],col = hmcol, scale = "none", cexRow=0.50, cexCol=0.75)

select8 <- order(rowMeans(counts(cds.subset)), decreasing=TRUE)[1:100]
heatmap( vsd.subset[select8,],col = hmcol, scale = "none", cexRow=0.50, cexCol
    =0.75)
#distance array
cdsFullBlind.subset <- estimateDispersions( cdsFull.subset, method = "blind" )
vsdFull.subset <- getVarianceStabilizedData( cdsFullBlind.subset )
dists2 <- dist( t( vsdFull.subset ) )
heatmap(as.matrix(dists2),symm=TRUE, scale="none", col=hmcol, margins=c(8,8),)

#For differential expression analysis I used an alternative method to estimate
    dispersions

#less conservative sharingMode in est dispersions:
cds <- newCountDataSet(data, conds)
cds <- estimateSizeFactors( cds )
cds.fit <-estimateDispersions(cds, sharingMode="gene-est-only")
plotDispEsts( cds.fit )

#test between Am40 and Pc19:
res.fit.2 <- nbinomTest( cds.fit, "1", "3" )
head(res.fit.2)
plotDE(res.fit.2, 0.1)

#filtered for significant genes according to a FDR threshold
resSig.fit <- res.fit[ res.fit$padj < 0.1, ]
head(resSig.fit)
#look at the most strongly downregulated genes
head( resSig.fit[ order( resSig.fit$foldChange, -resSig.fit$baseMean ), ] )

#look at the most strongly downregulated genes
head( resSig.fit.2[ order( resSig.fit.2$foldChange, -resSig.fit.2$baseMean ), ] )
write.csv( resSig.fit.2[ order( resSig.fit.2$foldChange, -resSig.fit.2$baseMean ),
    ], file="Pc19vsAm40_nbinom_test.csv" )
```

23. I also analyzed differential expression with EdgeR (Robinson et al., 2010). I followed EdgeR tutorial and performed similar contrasts to DESeq using the exact test and likelihood ratio test. Results were similar to DESeq and are not reported here.

24. Gene ontology (GO) analysis: Using the RBH dataset, I annotated each gene with its predicted peptide sequence and imported them to Blast2GO software. Annotations generated with BLASTX and InterProScan (see above) were loaded into Blast2GO

as well. Subsequent annotation within Blast2GO followed the standard procedure of merging annotations, mapping, running Blast2GO annotation. Enrichment analysis (using Fisher's Exact Test) revealed GO categories over or underrepresented across experimental conditions.

B.3 References

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol 11, R106.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D. a, Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494–1512.

Kurn, N., Chen, P., Heath, J.D., Kopf-Sill, A., Stephens, K.M., and Wang, S. (2005). Novel isothermal, linear nucleic acid amplification systems for highly multiplexed applications. Clin. Chem. 51, 1973–1981.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., and Lopez, R. (2005). InterProScan: Protein domains identifier. Nucleic Acids Res. 33, 116–120.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140.

Wühr, M., Freeman, R.M., Presler, M., Horb, M.E., Peshkin, L., Gygi, S.P., and Kirschner, M.W. (2014). Deep proteomics of the *Xenopus laevis* egg using an mRNA-derived

reference database. Curr. Biol. 24, 1467–1475.