Summer 2013

Differential Expressions of Nodal-Signal Transducers and Global Transcriptional Repression Commit Vegetal Cells in Eleutherodactylus Coqui to Form Nutritional Endoderm

Suman Chatterjee

DIFFERENTIAL EXPRESSION OF NODAL-SIGNAL TRANSDUCERS AND GLOBAL TRANSCRIPTIONAL REPRESSION COMMIT VEGETAL CELLS IN *ELEUTHERODACTYLUS COQUI* TO FORM NUTRITIONAL ENDODERM

A Dissertation

Submitted to the Bayer School of Natural and Environmental Sciences

Duquesne University

In partial fulfillment of the requirements for

the degree of Doctor of Philosophy

By

Suman Chatterjee

August 2013

DIFFERENTIAL EXPRESSION OF NODAL-SIGNAL TRANSDUCERS AND GLOBAL TRANSCRIPTIONAL REPRESSION COMMIT VEGETAL CELLS IN *ELEUTHERODACTYLUS COQUI* TO FORM NUTRITIONAL ENDODERM

By

Suman Chatterjee

Approved April 11, 2013

Richard P. Elinson Professor of Biological Sciences (Committee Chair)	Philip E. Auron Professor of Biological Sciences (Committee Member)
John A. Pollock Associate Professor of Biological Sciences (Committee Member)	Charles A. Ettensohn Professor of Biological Sciences Carnegie Mellon University, Pittsburgh (Committee Member)
David W. Seybert Dean, Bayer School of Natural & Environmental Sciences Professor of Chemistry and Biochemistry	Joseph R. McCormick Chair, Department Biological Sciences Associate Professor of Biological Sciences

ABSTRACT

DIFFERENTIAL EXPRESSION OF NODAL-SIGNAL TRANSDUCERS AND GLOBAL TRANSCRIPTIONAL REPRESSION COMMIT VEGETAL CELLS IN *ELEUTHERODACTYLUS COQUI* TO FORM NUTRITIONAL ENDODERM

By

Suman Chatterjee

August 2013

Dissertation supervised by Richard P. Elinson, Ph.D.

The vegetal core cells of a *Xenopus laevis* embryo commit to mesendoderm via the Nodal-signaling pathway. In *Eleutherodactylus coqui*, a direct developing frog, mesendoderm is specified at the marginal zone of the early gastrulae and vegetal core cells transform into nutritional endoderm. It is a novel tissue consisting of transient, yolk rich cells that provide nutrition, but do not differentiate into adult tissues. We hypothesized that a disruption of Nodal-signaling is responsible for committing vegetal core cells to nutritional endoderm. I report a dual regulation involved in the generation of nutritional endoderm. First, differential expression of Nodal-signaling components like Smad2 and Smad4 was observed during early gastrulation between cells in the marginal zone and in the vegetal core. Although *EcSmad2* RNA, as well as both native and the active forms of EcSmad2, were detected in the vegetal core by qPCR and western blot

respectively, western analysis revealed that Smad4 isoforms were expressed at a low level during early gastrulation. Immunostaining showed that only 12% and 50% of vegetal core cells were positive for nuclear Smad2 and Smad4 signals, respectively, compared to 100% in marginal zone cells. These differential expressions may indicate a signaling blockade in vegetal core cells. Second, I found global transcription repression in vegetal core cells by immunostaining. At late blastula, both the marginal zone and vegetal core cell were transcriptionally silent. At the onset of gastrulation, marginal zone cells, but not vegetal core ones, became transcriptionally active. This indicates the occurrence of a mid-blastula transition in the marginal zone by early gastrulation. Global transcriptional repression prevails in the vegetal core through development. A combination of differential Nodal-signaling and global transcriptional repression in vegetal core cells may account for its lack of differentiation.

DEDICATION

This dissertation is dedicated to my dad, late Mr. Sanjoy Chatterjee, who passed away on March 14, 2010. He was the first person in our business family to act against the wave and inspire me to dream of higher education abroad. Despite of several difficulties he provided me with the opportunity of an education from one of the best institutions in USA and supported my every decision till death. It is unfortunate that he did not live long enough to see this day of my life, but I am sure his well wishes are always there for me. I am and will be greatly indebted to him for everything he had done for me…

ACKNOWLEDGEMENT

"…The answer my friend is blowin' in the wind

The answer is blowin' in the wind."

- Bob Dylan, "Blowin' in the Wind"

I was 15, when I first listened to this song, written and sang by one of the great poets of all time and since then it has become a bit of wisdom I have followed most of my life. This song has been a tremendous inspiration to me in all the difficult times, when I looked for an answer. I realized the actual power of this song, when I came to USA to pursue my Ph.D. It helped me to become a stronger and an independent person during my initial days in USA, away from my family and home. Although it is true that the answers are always in the wind, nobody possesses the default vision to find the answer. There are several people in my life for last four years, who guided me in the right directions to find those answers I was looking for.

I would like to express my deepest gratitude towards Dr. Elinson for accepting me in his lab and for providing with extraordinary guidance. I enjoyed every moment I spent with him talking not only about research, but also literature, music and films. I am sincerely and heartily grateful to him, for the support and ideas he provided me with throughout my dissertation writing. It would have not been possible without his help. My gratitude is extended to my committee members Dr. Auron, Dr. Pollock and Dr. Ettensohn (CMU, Pittsburgh, PA). I would also like to show my appreciation to them for

all those valuable comments and advices they provided to shape up and orient my research projects in the right direction.

I owe a debt of gratitude to Kim Nath, our lab technician and Uma Karadge, the last graduate student left in the lab to face my questions on a variety of subjects. I would like to thank Dr. Selcer for always providing me with extraordinary ideas for my protein works. Credit should be given to my close friend in the department, Juraj, for helping me out with my qPCR experiments. I would like to thank Dr. Welsh, our A&P advisor for always being a friend, philosopher and guide. He helped me a lot to become an expert in a field alien to my endeavor.

It is with immense gratitude that I acknowledge the support and help I received throughout my graduate life at the Department of Biological Sciences. People like Dr. McCormick, Dr. Auron, Dr. Pollock, Dr. Seaman, Heather Costello, Phil Hoschar, Lalitha Rajakumar are only few of them to name. I am also deeply thankful to Judy Quinque and Pam Ferchak from the Biology office, without them life could have been miserable.

Finally, I owe sincere and earnest thankfulness to my wife, Swati for being there for me, when I needed her most. I wish to thank her parents, Mr. and Mrs. Banerjee, for their trust and faith they put on me. I have no words to express my gratitude to my mom, Madhumita Chatterjee and my brother, Samrat for the immense understanding and constant support they have shown. I owe you the most.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

ALK ⇒ Activin receptor-like kinase

ARE ⇒ Activin response element

A/V ⇒ Animal-Vegetal

BMP ⇒ Bone morphogenetic protein

CT ⇒ Threshold cycle

CTD ⇒ C-terminal domain

CBB ⇒ Coomassie Brilliant Blue

ΔΔCT ⇒ Comparative CT

DE ⇒ Definitive Endoderm

DEPC ⇒ Diethylpyrocarbonate

DI ⇒ Deionized

D/V ⇒ Dorsal-Ventral

HB ⇒ Homogenization buffer

NLS ⇒ Nuclear localization signal

NES ⇒ Nuclear export signal

MBT ⇒ Mid-blastula transition

MH1 ⇒ MAD Homology 1

MH2 ⇒ MAD Homology 2

MZ ⇒ Marginal Zone

NE ⇒ Nutritional Endoderm

NF ⇒ Nieuwkoop and Faber Staging Prefix

ORF ⇒ Open Reading Frame

PGC ⇒ Primordial Germ Cell

RNAPII ⇒ RNA Polymerase II

TS ⇒ Townsend and Stewart Staging Prefix

VC ⇒ Vegetal Core

CHAPTER ONE: INTRODUCTION

Complete or holoblastic cleavage similar to that in the extant, biphasic, amphibian model system *Xenopus laevis,* is considered primitive. It is found in invertebrates like annelids and mollusks, in deuterostome outgroups of craniates, namely sea urchins, tunicates, amphioxus, and in placental mammals. Transition to incomplete, meroblastic cleavage happened five times in the whole vertebrate phylogenetic distribution, including lineages leading to the development of hagfish, sharks and other cartilaginous fishes, teleosts, coelacanth, and amniotes (Collazo et al., 1994; Chea et al., 2005; Elinson, 2009; Takeuchi et al., 2009). In the course of evolution of terrestrial vertebrates, the two major groups, the amphibians and the amniotes, acquired fundamental differences in their early development. Characterized by the presence of extraembryonic membranes, gastrulation through a primitive streak, and meroblastic cleavage, amniote embryos are likely derived from embryos like those of extant amphibians (Collazo et al., 1994; Elinson, 2009). Whereas the entire amphibian egg develops into embryonic tissues, both extraembryonic as well as embryonic tissues develop from amniote eggs. The amniote gastrula is different from amphibians, as the embryo is composed of several flat layers requiring body folds after gastrulation to produce the characteristic three-tubed chordate embryo (Elinson, 2009).

An increase in the amount of yolk in eggs could account for these differences, which arose 360 million years ago (Elinson and Beckham, 2002; Elinson, 2009). Increase in yolk makes holoblastic cleavage difficult and favors the appearance of meroblastic cleavage in the course of evolution. To investigate the impact of increased yolk on

embryonic patterning, we analyzed the development of large eggs from *Eleutherodactylus coqui*.

E. coqui, the direct developing Puerto Rican tree frog, possesses a novel tissue, the nutritional endoderm. Nutritional endoderm (NE) arises in addition to the three germ layers endoderm, mesoderm and ectoderm, during early embryonic development (Buchholz et al., 2007). NE provides nutrition to the developing froglets, but does not differentiate into any organs in adults. Direct development is the process of formation of adult morphology as a direct consequence of embryogenesis, without the free-swimming larval stage of the more familiar biphasic amphibian. Although the fundamental mechanisms of direct development remain relatively unexplored, it is a common alternative life history strategy in recent amphibians (Jennings and Hanken, 1998).

For my Ph.D. dissertation, I investigated the molecular events, including the roles of some important signaling molecules underlying the generation of NE. In addition, I address the mechanism of its origin in *E. coqui* development.

I. Direct development and *Eleutherodactylus coqui*:

A. Indirect development: A common mode of anuran development.

Amphibians are a diverse group of animals including frogs and toads, which belong to the order Anura. Anurans usually exhibit indirect development, characterized by an intermediate aquatic larva between the embryo and the terrestrial adult. These free-swimming larvae, known as tadpoles, are drastically different from their adults in terms of morphology as well as life style. The process of transformation of tadpoles into adults is metamorphosis, which involves profound morphological, physiological, biochemical,

ethological and environmental changes (Dent, 1968; Just et al., 1981). This indirect development is very common in anurans and extensively documented in the life history of the model organism system of African clawed frog, *X. laevis*.

B. Direct development: An evolutionary significant adaptation:

Direct development, on the other hand, is the alternative strategy of development, where adults develop directly from the embryo skipping the intermediate free-swimming larva. The term "direct-development" has been used in many contexts. Duellman and Trueb (1986) suggested a direct correlation between the adoption of a terrestrial life style and the need of direct development. Absence of tadpoles in the life history of direct developers also raises the question of whether they undergo thyroid hormone-dependent metamorphosis (Callery and Elinson, 2000a). A more ecological definition was provided by Wake (1980), in which he included animals, which hatch as non-feeding tadpoles and undergo metamorphosis once their endogenous supply of yolk is exhausted. *Nectophrynoides malcolmi*, an anuran, would be an example of this type of development. Although it is considered a direct developer, a relatively non-motile tadpole form exists in its life history, which does not explore for food.

The most radical example of direct development is exhibited by the genus *Eleutherodactylus* (Salthe and Mecham, 1974). With only a little hint of their metamorphic ancestry in their life history (Townsend and Stewart, 1985), the embryogenesis of these anurans is highly modified. The extraordinary success of such a developmental strategy is well advocated by the presence of more than 450 species within the genus, which makes *Eleutherodactylus* the most species-rich vertebrate genus

(Hanken, 1992). Recent studies indicated the presence of 882 species of New World frogs that belong to the subfamily Eleutherodactylinae of the family Leptodactylidae. All these species breed on land and show direct development (Hedges et al., 2008). They lay eggs on land, where the eggs hatch into froglets. The geographical distribution of these frogs range from southern parts of the United States to northern Argentina, but are most diverse in Central America, the West Indies, and South America (Frost, 2007; Hedges et al., 2008; Heinicke *et al.* 2007).

C. *Eleutherodactylus coqui,* the direct developing, Puerto Rican tree frog.

The *Eleutherodactylid* frogs are K-strategists, characterized by high parental care investment per clutch. A clutch usually consists of 30 embryos. The embryos are large and filled with yolk to compensate for the lack of feeding tadpoles (Townsend and Stewart, 1985; Elinson, 1987a, b). *E. coqui* has been studied in detail (Elinson, 1990; Elinson and Beckham, 2002). Fertilization is internal and the embryos, after being laid, are brooded by the male until hatching (Elinson, 1990). The diameter of an *E. coqui* egg is 3.5 mm, whereas that of a *X. laevis* egg is 1.3 mm. This difference in the size indicates a volume ratio of 20:1 between the two species (Elinson, 1987). This large maternal storage of nutrients in the form of yolk favors the prolonged intra-oval development and production of young froglets directly. This common arboreal tree frog is endemic to the Caribbean island of Puerto Rico, where it is also considered as a "national symbol" (Joglar, 1998; Velo-Antón et al., 2007). In the late 1980s, *E. coqui* was accidentally introduced to Hawaii (Kraus et al., 1999), where it is considered a pest as a consequence

of affecting the state's floriculture and nursery industries and property values (Beard, 2007).

D. Characteristics of direct development in *E. coqui*

Many structural features that characterize an aquatic embryo and a tadpole are rudimentary or lacking in the life history of direct developers (Elinson, 2001 & 2013). *E. coqui* embryos showcase a fairly unremarkable development until the late neural plate stage, which is marked by the appearance of both fore- and hind-limb buds (Elinson and Fang, 1998; Townsend and Stewart, 1985). Unusual characteristics of the direct developing *E. coqui* are: (1) early and simultaneous appearance of limbs, (2) no horny mouth parts or adhesive organ, (3) great reduction in size of gills, (4) absence of a coiled intestine, (5) presence of a large yolk reserve, (6) development of a large membranous but non-muscular tail, (7) presence of egg tooth which is different from a tadpole's keratinous teeth and beak (Townsend and Stewart, 1985), (8) absence of a hatching gland as they use their egg tooth to break open the jelly capsule for hatching, (9) incomplete development of operculum (Callery and Elinson, 2000a), (10) absence of cement glands, which are required by the non-motile aquatic embryos to attach themselves to a substratum (Fang and Elinson, 1996, 1999; Elinson, 2013), and (11) absence of lateral line organs, sense organs used by aquatic organisms like tadpoles to detect water currents (Schlosser et al., 1999; Elinson, 2013).

The large quantities of yolk serve as a source of nutrients and minerals like calcium, magnesium, and phosphorous to build the body of growing froglets (Packard et al., 1996). Utilization of yolk during direct development involves a rapid decrease in the

dry yolk mass with a proportional gain in weight of the embryonic carcass, highlighting the importance of yolk in direct development (Packard et al., 1996). Although the increase in yolk makes holoblastic cleavage difficult and favors the appearance of meroblastic cleavages in the course of evolution, *E. coqui* embryos undergo holoblastic cleavage and the yolk is compartmentalized into individual cells as NE (Elinson, 1987; Buchholz et al., 2007; Elinson, 2009). The NE attaches to the gut, encased within a vascular intestinal sac and persists as a nutrient source even after hatching (Lynn, 1942; Valett and Jameson, 1961; Buchholz et al., 2007; Singamsetty and Elinson, 2010). Development of froglets from embryos in *E. coqui* requires thyroid hormone, like the indirect developers. This period of their life history is called "cryptic metamorphosis" (Callery and Elinson, 2000b; Elinson, 2013).

E. Thyroid-dependent metamorphosis in *E. coqui*:

Thyroid hormone plays crucial roles in the "cryptic metamorphosis" of *E. coqui*. When thyroid hormone synthesis is blocked by addition of methimazole, an anti-thyroid drug, development arrests at TS12 (staging according to Townsend and Stewart, 1985). Such arrested TS12 embryos have a frog-like head, eyes, legs, and general body morphology, but lack complete formation of skin, body musculature, jaw cartilage, and digestive tract (Callery and Elinson, 2000b; Singamsetty and Elinson, 2010; Elinson, 2013). Such developmentally arrested embryos do not hatch and stay inside their fertilization membrane and jelly coats. All of these methimazole effects are reversible upon addition of thyroid hormone (Callery and Elinson, 2000b; Singamsetty and Elinson, 2010). *E. coqui* is characterized by the presence of a highly vascularized tail, used for

respiration rather than swimming. As swimming is not part of the life style, the tail loses most of the musculature (Elinson, 2013). This tail is reabsorbed by the action of thyroid hormone after hatching around TS15 (Callery and Elinson, 2000b; Elinson, 2013). Over the course of embryogenesis, utilization of yolk and disappearance of NE are regulated by thyroid hormone (Singamsetty and Elinson, 2010).

II. Early embryonic development in *X. laevis*:

Over the last couple of decades, application of advanced molecular tools to the examination of development in the model organism system, *X. laevis*, has increased our knowledge of processes involved in generation of an embryo from an egg and formation of an adult from an embryo.

A. Origin of three germ layers:

Gastrulating metazoan embryos generate three germ layers - endoderm, mesoderm and ectoderm, from which organs arise. Unlike ectoderm or mesoderm, the induction and differentiation of endodermal cells and formation of the respective organs from endoderm was poorly understood. In last decade, however, genetic and molecular approaches involving zebrafish mutants and *X. laevis* and knockout mice have advanced our understanding of endoderm development and differentiation.

In amphibian oocytes, establishment of animal-vegetal (A/V) axis and formation of germ plasm are crucial events. Upon fertilization, the dorsal-ventral (D/V) polarity is developed as a second axis. Temporal and spatial superimposition of these two axes then serves as the blue print for the basic body plan (Harland and Gerhart, 1997). The cells of

the three prospective germ layers become committed to their individual fates at gastrulation (Horb and Slack, 2001; Takeuchi et al., 2009). The ectoderm is formed from the pigmented animal hemisphere and gives rise to epidermis and nervous system. Endoderm, the innermost layer of the three, is derived from the yolky vegetal hemisphere and becomes committed to form the gut and associated organs, including the epithelial lining of the liver, gall bladder, pancreas and respiratory system (Dale and Slack, 1987; Stainier, 2002; Fukuda and Kikuchi, 2005). Finally, mesoderm originates as an annulus at the equator (Clements et al., 1999) and forms blood, heart, muscle, kidneys and skeleton. Mesoderm also differentiates into the smooth muscle and connective tissues of the gut (Chalmers and Slack, 1998). All of these cell fate determinations are specified by the combined activity of different signaling pathways and interplay among essential molecular determinants. One crucial pathway is the Nodal-signaling pathway, which plays pivotal roles in endoderm/ mesoderm induction and left – right axis determination (Schier, 2009).

B. Early embryogenesis:

In *X. laevis*, fertilization is followed by twelve rapid and synchronous divisions leading to the formation of a blastula (Heasman, 2006). The first cell cycle lasts 90-minutes, whereas each of the next 11 divisions continues for 20 to 30 minutes. These rapid cell cycles consist of only DNA replication and mitosis with no gap phases. These early embryonic cells are transcriptionally quiescent, and depend solely upon maternal supplied stockpiles of proteins and RNAs for survival (Newport and Kirschner, 1982; Kisielewska and Blow, 2012). These divisions lead to the formation of an embryonic

mass of 4000 cells, enclosing a fluid filled blastocoel cavity (Heasman, 2006). After the twelfth division, the cell cycle gradually slows down and lasts about 50 minutes. The divisions become asynchronous and are characterized by appearance of gap phases. The cell-cycle changes from checkpoint unregulated to checkpoint regulated ones. This transition was first described by Signoret and Lefresne as "transition blastuleenne" (Korzh, 2009; Signoret and Lefresne, 1971). Later, Gerhart (1980) referred to it as the mid-blastula transition (MBT), a critical event, which marks the onset of embryonic gene expression (Gerhart, 1980; Heasman, 2006; Shiokawa, 2012).

MBT is also referred to as MZT for Maternal-to-Zygotic transition (Tadros and Lipshitz, 2009). MBT is extensively characterized in various model organisms such as *Caenorhabditis elegans* (nematode), *Drosophila melanogaster* (insect), *Strongylocentrotus purpuratus* (echinoderm), *Danio rerio* (fish), and *X. laevis* (amphibian). MBT is also evident in plants (Baroux et al., 2008). Morphologically, the mid-blastula embryo consists of three regions, the animal cap forming the roof of the blastocoel, the equatorial or marginal zone forming the walls of the blastocoel, and the vegetal mass, which constitutes the floor of the blastocoel. At a molecular level, two processes characterize MBT: elimination of a subset of maternal mRNAs and beginning of zygotic genome transcription.

Until the beginning of MBT, all the RNAs and proteins are maternally derived. Experimental analysis in zebrafish and *X. laevis* has indicated that at MBT the nucleocytoplasmic (N/C) ratio increases, which in turn promotes maternal supply of replication factors (Shimuta et al., 2002; Gotoh et al., 2011). This results in stalled replication forks, which activates a physiological checkpoint preventing cells from

entering into mitosis (Shimuta et al., 2002; Gotoh et al., 2011). In addition to the beginning of zygotic gene expression at MBT, some cells acquire mobility ultimately leading to gastrulation, which is marked by decreased mitoses and formation of the dorsal lip of the blastopore in the 15th cycle (Shimuta et al., 2002; Heasman, 2006; Gotoh et al., 2011).

The dorsal blastopore lip is the site where the Spemann organizer gene network is established. The organizer plays important roles in patterning the three germ layers. Analogous to the node in amniotes, Spemann organizer causes the neuralization of the ectoderm, dorsalization of the mesoderm and anteriorization of the endoderm (Gilbert, 2000). As a result of gastrulation, the embryonic ball converts into three layers, forming the definitive D/V and A/P axes.

Experimental analysis of formation, specification, differentiation and interaction of three germ layers in *X. laevis* has provided evidence that the vegetal hemisphere of the embryo gives rise to the endoderm (Dale and Slack, 1987; Horb and Slack, 2001, Heasman, 2006). Complex morphogenetic events during gastrulation transform the vegetal region into the archenteron, the tubular primitive gut. The cells from the original vegetal surface are internalized to cover the floor of the archenteric cavity, whereas the surface cells from the animal border of the blastopore lip of the gastrula and constitute the roof of archenteron. Finally, the archenteron cavity is closed and replaced by the definitive lumen of the gut (Chalmers and Slack, 2000).

III. Nodal-signaling underlying endoderm/mesoderm specification:

A complicated and interconnected network of different signaling pathways guides development of an early vertebrate embryo. One such pathway is the Nodal-signaling pathway, which determines many important fates during vertebrate embryogenesis. Ectopic expression of various members of the TGF-β family, like Activin and Nodal proteins such as Xnr-1, Xnr-2 and Derriere, induce the animal cap cells of early *X. laevis* embryos to express mesendoderm specific genes (Clements, 2001). When Nodal expression is blocked in the vegetal half of early embryos by the use of dominant negative forms of Nodal (Osada and Wright, 1999), Activin receptors (Clements et al., 1999; Yasuo and Lemaire, 1999; Chang and Hemmati-Brivanlou, 2000), and Derriere (Sun et al., 1999), it reduces or totally abrogates expression of mesendoderm specific genes (Stainier, 2002). These results strongly suggest an essential role of TGF-β signaling in *X. laevis* endoderm/mesoderm formation during early embryogenesis.

Like most TGF-β signals, Nodal ligands activate serine/threonine kinase receptors, which in turn phosphorylate Smad proteins to regulate endoderm/mesoderm specific genes (Schier, 2003; Shen, 2007; Wu and Hill, 2009) (Fig. 1). The Nodal signal is received by type I and II Activin receptor duo and co-receptors EGF-CFC. A core Nodal-signaling pathway consists of several steps. Receptor complex activation leads to the phosphorylation and activation of cytoplasmic Smad2/3 proteins, which in association with Smad4 translocate into the nucleus to bind with additional transcription factors to regulate endoderm/mesoderm specific genes (Schier, 2009) (Fig. 1). Several extracellular factors, like modifying enzymes, inducers, and antagonists as well as intracellular ones like transcriptional cofactors, proteins involved in receptor trafficking, and miRNAs,

regulate this pathway. In the following paragraphs, different components of the signaling cascade are described in detail.

A. The ligands:

Nodals, in addition to endoderm/mesoderm induction, are involved in establishing development of the left-right axis (Burdine and Schier, 2000; Capdevila et al., 2000; Hamada et al., 2002; Mercola and Levin, 2001). The Nodal-signaling pathway has been genetically and biochemically dissected in several vertebrate model systems including zebrafish, *X. laevis*, and mouse (Stainier, 2002; Schier, 2009). Activation of *Nodal* genes is spatially restricted. Nodal ligands can act locally as well as morphogens, acting at a distance in a concentration-gradient dependent manner (Chen and Schier, 2001; Green 2002, Gurdon and Bourillot 2001; Meno et al., 2001). Adjacent to the presumptive mesoderm in the vegetal half of the *X. laevis* embryos, *Nodal* genes are expressed, overlapping with the endoderm precursors. Higher levels of Nodal signaling in and close to the source produce endoderm but lower levels lead to the development of mesoderm in the neighboring cells (Schier, 2003; Schier and Talbot, 2005; Shen, 2007).

The Nodal proteins of the TGF-β superfamily are abundant only in chordates. So far the group of proteins include one Nodal in mouse, three proteins - Cyclops, Squint and South-paw in zebrafish, and Xnr1, 2, 4, 5, 6, and Derriere in *X. laevis* (Whitman, 2001; Stainier, 2002; Schier, 2003, 2009). Unlike Wnt and Notch pathways, Nodal signaling is not found in *D. melanogaster* or *C. elegans*, which suggests more specific roles for this pathway during vertebrate development than these other highly conserved ones (Schier, 2003). Experiments in *X. laevis* and other vertebrate systems have revealed

that Nodals possess similar biochemical activities, and their functional specificity is determined by complex regulation of *Nodal* gene expression. Additional complexity in ligand functioning could result from heterodimerization with other TGF-β family members, which in turn could lead to cooperative interactions or mutual inhibition (Eimon and Harland, 2002).

Nodal proteins are translated as proproteins consisting of a prodomain and a ligand domain. In response to specific signals, subtilisin-like proprotein convertases Spc1 and Spc4 proteolytically cleave the proproteins at R-X-(K/R)-R and R-X-X-R consensus sequences and produce the active Nodal ligand (Fig. 1) (Beck et al., 2002; Schier and Shen 2000; Schier 2009). While these convertases are involved in the processing of the proproteins, latent TGF-β binding proteins (LTBP) contribute to the maturation and presentation of the signals by binding to the peptide. Altmann et al. (2002) illustrated the role of *X. laevis* LTBP-1 in potentiation of Nodal and Activin signaling. Proteins like Lefty and Cerberus act as extracellular inhibitors and can restrict Nodal signaling spatially and temporally by binding to Nodal itself (Adamson et al., 2001; Branford and Yost, 2002; Chen and Schier, 2002; Feldman et al., 2002, Meno et al., 2001).

B. The Nodal receptors:

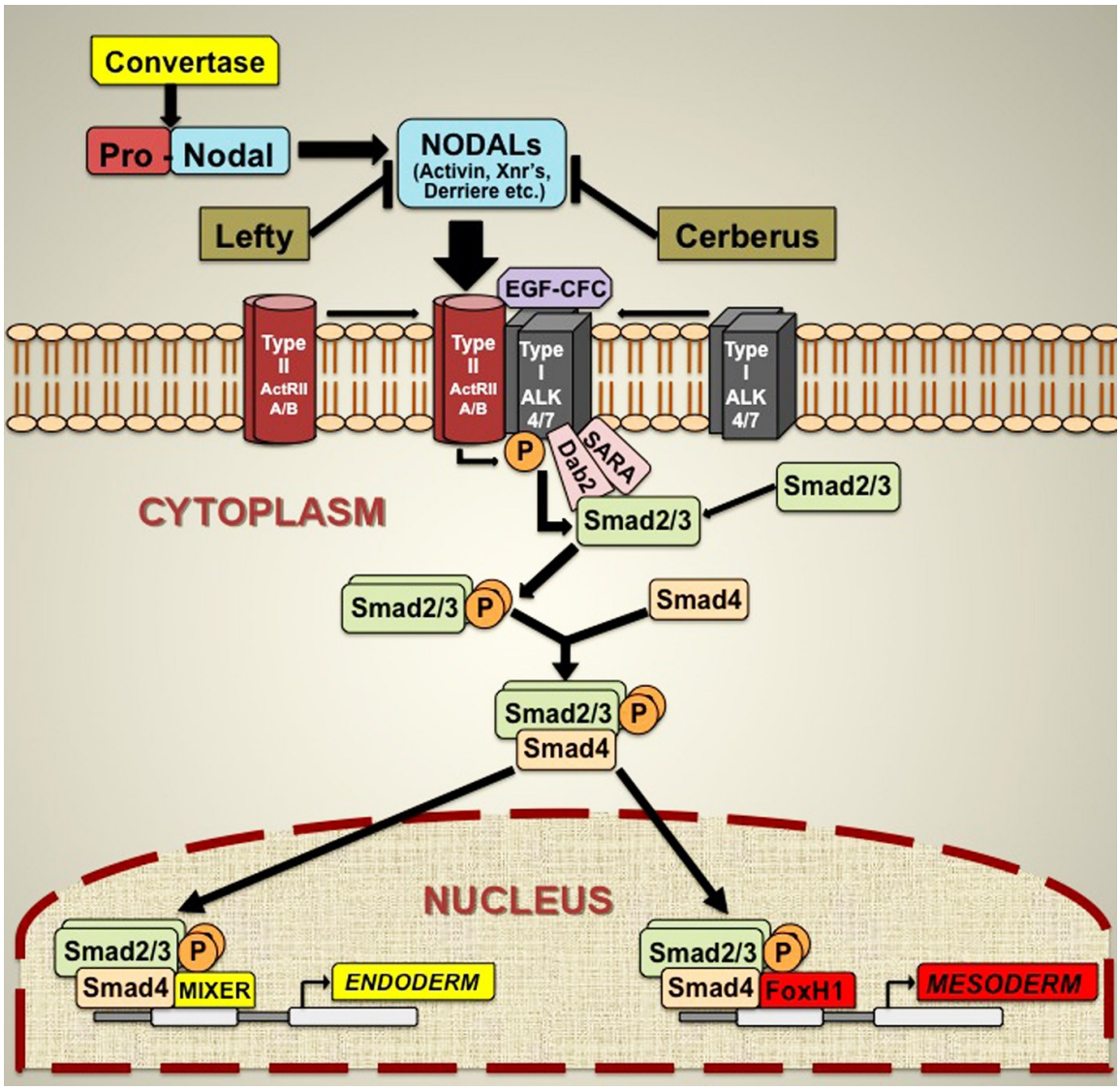
Receptors for the TGF-β superfamily ligands are composed of a cysteine-rich extracellular domain, a single pass transmembrane domain, and a highly conserved intracellular serine/ threonine kinase domain (Kimelman, 2006). Different Activin isoforms and the other members of the TGF-β superfamily of ligands turn on signaling by binding to a heteromeric complex between a type I and a type II serine-threonine kinase

receptors. There are seven type I and five type II receptors including those for Activins. These Activin receptors are represented by Activin receptor-like kinase 4 (ALK4), a type I receptor, and Activin receptor type IIA (*Acvr2a*, or ActRIIA) and Activin receptor type IIB (*Acvr2b*, or ActRIIB), the type II receptors (Mathews and Vale, 1991; Attisano and Wrana, 2002; Reissman et al., 2001; Yeo and Whitman, 2001; Yan et al., 2002; Shi and Massagué, 2003). A second type I receptor, ALK2 has been reported, but it is not involved in transduction of the signals (Willis et al., 1996; ten Dijke et al., 1994; Tsuchida et al., 1993). Another example of type I receptor is ALK7, which is instrumental in propagating signals from Activin B specifically, as shown in murine pancreatic β cell line (Tsuchida et al., 2004). An example of *X. laevis* type I receptor is XALK4, which is maternally contributed to the embryo. Overexpression of *XALK4* induces mesoderm formation in *X. laevis* embryos and explants (Hemmati-Brivanlou et al., 1992; Mathews et al., 1992).

Activin type II receptors are highly conserved, showing ~98% sequence homology among human, mouse and rat. ActRIIA is an important member in conveying Nodal signals downstream. The open reading frame (ORF) of *X. laevis* ActRIIA homolog is 1545 bp (Xenbase: XB-GENE-865037), which produces a peptide of 514 amino acids with a predicted molecular weight of 57.9 kDa.

Fig. 1: Nodal signaling pathway involved in endoderm/mesoderm specification. In the presence of proper cues, convertase proteins cleave Nodal pro-proteins to generate active Nodal ligands. Availability of active ligands brings preformed dimers of TGFβ receptors type I and type II together on the cell surface. Upon type I receptor phosphorylation by type II, an active ligand-receptor heteromeric complex is formed. Anchoring proteins like SARA, and an adaptor protein, like Dab2, recruit inactive Smad2/3 molecules from the cytoplasm to the active receptor complex, where active type I receptor phosphorylates the C-terminal domain of Smad2/3. Active Smad2/3 dislodges from the complex and forms a complex with the co-Smad protein Smad4. Active Smad2/3/4 complexes translocate inside the nucleus and bind specific transcription factors, already recruited on the genes. The formation of such multi-protein transcriptional activation complexes leads to the expression of endoderm/mesoderm specific target genes. In the case of endoderm specification, the active Smad2/3/4 complex usually binds a Mixer-like transcription factor. For mesoderm specification, it is FoxH1, which accommodates Smad2/3/4. (Scheme modified after Schier, 2003)

Fig. 1:


Convertase
Pro - Nodal
NODALs
(Activin, Xnr's, Derriere etc.)
Lefty
Cerberus
EGF-CFC
Type II ActRII A/B
Type II ActRII A/B
Type I ALK 4/7
Type I ALK 4/7
P
Dab2
SARA
CYTOPLASM
Smad2/3
Smad2/3
Smad2/3 P
Smad4
Smad2/3 P
Smad4
NUCLEUS
Smad2/3 P
Smad4
MIXER
ENDODERM
Smad2/3 P
Smad4
FoxH1
MESODERM

C. Active receptor complex formation and signal transduction:

The signals from Nodal ligands bring together pre-formed dimers of type II receptors and dimers of type I receptors to form the heterotetrameric active receptor complex (Fig. 1). It is the type II receptor, not type I, which binds the ligands. The type I receptor binds to the pre-formed type II-ligand complex and initiates signal transduction. Overexpression studies in *X. laevis* have suggested the involvement of Activin receptor ActRIB (ALK4), ActRIIA and ActRIIB in Nodal signaling (Reissman et al., 2001; Yeo and Whitman, 2001). The type II receptor is constitutively active, and within the ligand-bound receptor complex, it phosphorylates and activates type I receptor at several serine and threonine residues in a glycine and serine (GS) rich juxtamembrane domain (Kimelman, 2006). This GS-domain is highly conserved among all the members of the type I receptor family members. Phosphorylation of this GS-domain leads to the recruitment of receptor-regulated Smads from the cytoplasm.

Members of the extracellular GPI-anchored protein family EGF-CFP are required as co-receptors in the type I/II receptor coupling mechanisms (Schier and Shen, 2000; Dorey and Hill, 2002; Onuma et al., 2006; Schier, 2009). The EGF-CFP family includes one-eyed pinhead of zebrafish, FRL-1 of frogs, chick CFC, and mouse and human Cripto and Criptic (Shen and Shier, 2000). Although the actual mechanism is unknown, it is believed that the coreceptors change the confirmation of ALK4 to allow interaction with Nodals. An alternative hypothesis also exists, which states that EGF-CFC proteins provide an additional interaction surface for Nodals or can change the confirmation of Nodals to allow binding with the Activin receptors (Schier, 2003). Activins themselves can activate the receptors without EGF-CFC proteins, which raises questions regarding

the requirement of the EGF-CFC proteins (Gritsman et al., 1999; Schier and Shen, 2000; Whitman, 2001). Studies in *X. laevis* showed that the EGF-CFC protein, FRL-1 is involved in MAPK signaling. FRL-1 behaves like a neural inducer by activating MAPK and inhibiting BMP pathway (Kinoshita et al., 1995; Yabe et al., 2003).

Nodal receptor complex activation leads to the phosphorylation of regulatory Smads in the cytoplasm. Phosphorylation of Smad2 and possibly of Smad3 activates the Smads. Activated Smads bind their partner Smad4, forming a multi-protein complex to translocate into the nucleus. Inside the nucleus, this complex of proteins act as transcription factors to activate the transcription of distinct but partially overlapping sets of Nodal downstream genes (Kumar et al., 2001; Lee et al., 2001; Yeo and Whitman, 2001; Ross and Hill, 2008; Schier, 2009). In *X. laevis*, the nuclear accumulation of the Smad complexes occurs only after MBT (Saka et al., 2007).

D. Downstream targets:

In the case of endoderm/mesoderm specification in *X. laevis*, the most widely studied transcription factor that associates with phospho-Smad2/3 is Forkhead protein FoxH1, initially known as Fast1, which regulates the downstream targets of Nodals (Whitman 2001). FoxH1/Fast was identified in *X. laevis*, where it bound an Activin response element (ARE) in the promoter region of the mesendodermal gene *mix2* (Chen et al., 1996; Watanabe and Whitman, 1999). Enhancer elements adjacent to *lim1*, *goosecoid*, *lefty2*, *Nodal*, and *pitx2* mediate Nodal signaling via FoxH1-phospho-Smad2 interaction. In *X. laevis* and zebrafish, FoxH1 is required for the formation of mesoderm, but not endoderm (Watanabe and Whitman, 1999; Pogoda et al., 2000; Sirotkin et al.,

2000). This result suggests presence of additional transcription factors that recruit Smad2/3 for endoderm specification (Stainier, 2002).

There are certain members in the Mix family of paired-class homeodomain transcription factors, which also bind phospho-Smad2 (PSmad2). Some important members sharing this function are Mm1 in mouse (Pearce and Evans, 1999; Robb et al., 2000), Bonnie and clyde or Bon (Alexander et al., 1999; Kikuchi et al., 2000), Mtx1 and 2 (Hirata et al., 2000) in zebrafish, and seven proteins in *X. laevis* (Rosa et al., 1989; Mead et al., 1996; Vize et al., 1996; Ecochard et al., 1998; Henry and Melton, 1998; Lemaire et al., 1998; Tada et al., 1998). All of these proteins bind active Smad2 via a short motif present not only in them, but interestingly also in FoxH1 (Germain et al., 2000).

The endodermally expressed gene *Mix1* in *X. laevis* was initially identified through its responsiveness to mesodermal inducers (Rosa, 1989) and implicated in endoderm specification (Lemaire et al., 1998). Henry and Melton (1998) established that *X. laevis Mixer* is exclusively expressed in the endodermal precursors during early gastrulation. Ectopic expression of *Mixer* gene led to the activation of endoderm specific genes in animal cap cells. Experiments with zebrafish Mix family member *bon* revealed that *bon* expression happens before the onset of *Sox17* expression, and *bon* mutants fail to develop 90% of their endoderm (Alexander et al., 1999). In *X. laevis*, FoxH1 probably also regulates the expression of *Mixer*, an idea supported by the presence of a FoxH1-binding site in the *Mixer* gene promoter. *Smad2* mutant cells extensively colonized ectodermal and mesodermal lineages, but failed to form definitive endoderm during gastrulation. These findings identify Smad2 as the key nodal effector for endoderm and

mesoderm specification. Smad3 could compensate during mesoderm formation, but only when Smad2 is absent (Tremblay et al., 2000; Stainier, 2002).

E. Regulators of Nodal expression:

A turning point in the molecular understanding of *X. laevis* endoderm specification was the discovery that VegT, a T-box transcription factor, is both necessary and sufficient for endoderm formation (Lustig et al., 1996; Stennard et al., 1996; Zhang and King, 1996; Horb and Thomsen, 1997; Zhang et al., 1998; Xanthos et al., 2001). The striking evidence was provided by Zhang et al. (1998), showing that depletion of the maternal supply of the *VegT* RNA blocked endoderm specification leading to a disruption of germ layer patterning in the *X. laevis* blastula. This study also illustrated for the first time, that asymmetric expression of maternal *VegT* in the vegetal half of the *X. laevis* oocyte is required for both mesoderm and endoderm specification (D'Souza et al., 2003; Stennard et al., 1996; Xanthos et al., 2001; Zhang et al., 1998).

VegT regulates endoderm formation in several ways. It activates the expression of several TGF-β family members like *Xnr1*, *Xnr2*, *Xnr4* and *Derriere* (Kofron et al., 1999; Hyde and Old, 2000). It also directly regulates the expression of *Bix1* and *Bix4*, members of *Mix* type homeobox gene family (Tada et al., 1998; Casey et al., 1999). VegT becomes an essential controller of Nodal signaling, which in addition to the maintenance of its own expression, activates the zygotic expression of *VegT*. These events lead to the establishment of a positive feedback loop for endoderm specification in vegetal cells (Clements et al., 2001; Clements and Woodland, 2003) as well as a key synergy between

VegT and Nodal actions in endoderm and mesoderm specification (Clements and Woodland, 2003).

Another important VegT target is the ultimate endodermal determinant gene *Sox17*. Sox17 is a transcriptional regulator containing high mobility group (HMG) DNA-binding domain, which was first implicated in endoderm specification in *X. laevis* (Hudson et al., 1997). When *Sox17* was over-expressed in animal cap explants, it promoted endodermal gene expression. Blocking Sox17 activity in cells fated to become endoderm, promoted those cells to become mesoderm or ectoderm (Clements and Woodlands, 2000). Sinner et al. (2004) showed that Sox17 is involved in the activation of β-catenin, also required for transcription of endodermal target genes in *X. laevis*. Sox17 can act as an antagonist towards Wnt signaling (Zorn et al., 1999). Studies in mouse submucosal gland cell lines showed inhibitory effects of *Sox17* expression on Wnt3/β-catenin-mediated activation of *Lef-1* promoter (Liu et al., 2010). Sox17 is an essential component of the transcriptional network that defines the extraembryonic endoderm (Niakan et al., 2010). According to reports in *X. laevis*, Sox17 is directly involved in the regulation of many important genes such as *Hnf1β*, *Foxa2*, *Foxa4*, *Edd*, and *Zfp202* (Dickinson et al., 2005; Howard et al., 2007; Hudson et al., 1997; Sinner et al., 2004, 2006; Patterson et al., 2008).

IV. Smad2: The key effector of Nodal signaling.

During early vertebrate embryogenesis, Smad2 signaling in response to the signals from TGF-β ligand including Nodals, Activins, and Vg1-related proteins, is important and essential for germ layer specification and axial patterning (Ho et al., 2010).

Experiments in *D. melanogaster* and *C. elegans* first identified Smads as the downstream targets of the type I receptor. Following these findings, use of degenerate PCR and independent functional expression screens in *X. laevis* identified vertebrate Smad homologs of *D. melanogaster Mad* and *C. elegans Sma*. A total of nine vertebrate homologs have been identified, and the number is still increasing. In *X. laevis,* the expression of either full length or the C-terminal half of Smad2 had the mesoderm inducing effects of Activin on animal blastomeres (Baker and Harland 1996; Graff et al., 1996; Whitman 1998). In mice, mutation in *Smad2* gene abrogated all the derivatives of embryonic germ layers (Nomura and Li, 1998; Weinstein et al., 1998; Heyer et al., 1999). Chimeric studies in *Smad*2 mutant cells also showed these cells were able to colonize ectodermal and mesodermal lineages, but failed to colonize the definitive endodermal lineage during gastrulation (Tremblay et al., 2000). These results suggest that Smad2 is the key Nodal effector, necessary for endoderm specification.

A. Types:

There are three distinct categories of Smads. The *Pathway-specific or Receptor-regulated Smads (R-Smads)* such as Smad1, Smad2, Smad3, Smad5 and Smad8, are directly phosphorylated by active type I receptor. When over-expressed, they are able to phenocopy a signaling pathway. R-Smads 1, 5, and 8 are the principal substrates for BMP and anti-Mullerian receptors, whereas Smad2 and Smad3 serve as transducers of signals from TGF-β, Activin, and Nodal receptors (Massagué et al., 2005). The *Co-Smad,* Smad4, that is shared among multiple signaling pathways, and oligomerizes with activated R-Smads. The *Inhibitory Smads (I-Smads),* like Smad6 and Smad7, inhibit

TGF-β superfamily signaling. I-Smad inhibition is accomplished by competing with R-Smad for receptor binding and marking the receptor for degradation upon binding (Moustakas et al., 2001). There is a high level of sequence conservation among the different vertebrate homologs.

B. Domain structure:

Smad2 is composed of two highly conserved domains, the N-terminal Mad homology 1 (MH1) and the C-terminal Mad homology 2 (MH2) domains (Fig. 2). A less conserved, linker region separates these two domains. Functional studies involving point and deletion mutational analysis together with X-ray crystallography of MH1 and MH2 domains revealed insights into the principal interactions of Smad proteins with other proteins and with DNA (Shi and Massagué, 2003; Massagué 2005). These structural studies indicate an association between MH1 and MH2 in the inactive state, but upon receptor activation, the structure opens up to initiate the effector function of each domain (Baker and Harland 1996; Hata et al., 1997).

Sequence and structural analysis indicates a strong similarity between MH1 domain and the diverse His-Me (histidine-metal-ion) finger family of endonucleases. This similarity suggests that the MH1 domain might have evolved from an ancient enzymatic domain, which lost its catalytic function but retained the DNA binding power (Grishin, 2001). The MH1 domain is a DNA binding module, which is stabilized by a closely attached zinc atom. MH1 promotes two important functions, regulation of nuclear import and transcription via binding with DNA and other nuclear proteins. The DNA

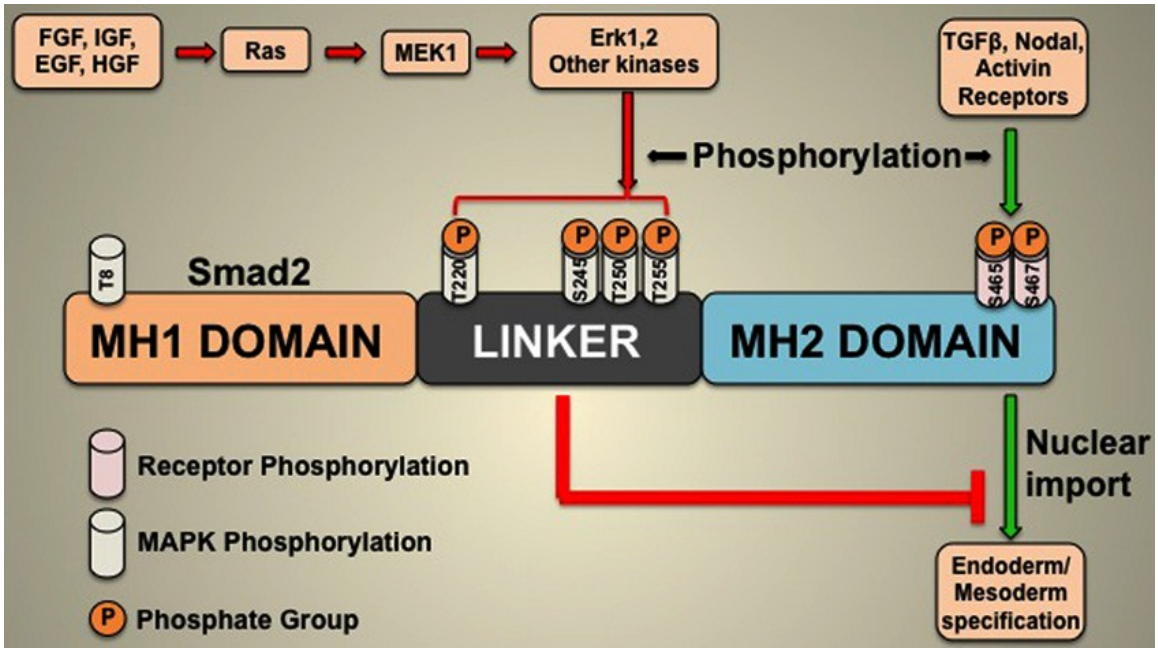
binding feature of MH1 is achieved by a β-hairpin structure, conserved in all R-Smads and Smad4.

The MH2 domain, a highly conserved domain among all known Smads, is the most versatile protein-interacting module in any signaling cascade. It contains multiple α-helices and loops surrounding a β-sandwich (Shi, 2001). The MH2 domain structure resembles that of a forkhead-associated (FHA) domain, a phospho-peptide binding domain commonly found in many transcription and signaling factors (Li et al., 2000). A conserved C-terminal Ser-Ser-X-Ser (SSXS) motif is present at the end of the MH2 domain. Receptor-mediated phosphorylation of the last two serine residues is crucial for activation (Fig. 2). Functions of the MH2 domain include oligomerization with different or the same types of Smads, recognition by type I receptors, and interaction with cytoplasmic adaptors and several other transcription factors. On the surface of the MH2 domain, a set of repeating hydrophobic patches is present that mediates Smad2's interaction with other proteins like cytoplasmic retention proteins, nucleoporins and DNA binding cofactors (Massagué et al., 2005).

The linker region, which lies between MH1 and MH2, serves as a binding site for proteins like Smurf (Smad ubiquitination-related factor) ubiquitin ligase. Moreover, this region provides sites for phosphorylation by several important classes of protein kinases like MAPK (Fig. 2). In the case of Smad4, the linker region also contains a nuclear export signal (NES).

Fig. 2: Smad2 phosphorylation. Smad2 protein consists of MH1 and MH2 domains with a linker region in the middle. The linker region is responsible for connecting Smad2 signaling to Ras/MAPK pathway. Phosphorylation of Smad2 at S465 and S467 by Nodal, Activin or TGFβ receptor(s) leads to its activation. This in turn specifies the endodermal and mesodermal cell fates in *X. laevis* development. Growth factors activate the Ras/MAPK pathway, which in turn induces MAPK. Four specific sites (T220, S245, S250 and S255) in the linker region are phosphorylated by MAPK, which acts in opposition to the TGFβ receptor mediated Smad2 activation. (Scheme modified from Kretzschmar et al., 1997; Massagué, 2003; Ross and Hill, 2008)

Fig. 2:


FGF, IGF, EGF, HGF
Ras
MEK1
Erk1,2 Other kinases
TGFβ, Nodal, Activin Receptors
Phosphorylation
T8
Smad2
T220
S245
T250
T255
S465
S467
MH1 DOMAIN
LINKER
MH2 DOMAIN
Nuclear import
Receptor Phosphorylation
MAPK Phosphorylation
Phosphate Group
Endoderm/ Mesoderm specification

C. Activation:

Smad2 is phosphorylated by type I receptor at its highly conserved C-terminal $S^{465/467}$ residues (Fig. 2). These two conserved serines together with a third, non-phosphorylated serine, S^{464} constitute an evolutionary conserved SSXS motif in all R-Smads (Abdollah et al., 1997; Souchelnytskyi et al., 1997). S^{465} and S^{467} residues were identified as the two major ligand-receptor regulated phosphorylation sites by radiochemical sequencing and phospho-amino acid analysis of phosphorylated peptides of Smad2 (Souchelnytskyi et al., 1997). The phosphorylation of these sites happens sequentially following an obligate order. S^{467} residue is first recognized by the type I receptor kinase, which then displays a new recognition sequence leading to the phosphorylation of S^{465}. When Souchelnytskyi et al. (1997) replaced S^{467} with the phosphomimetic aspartic acid residue (S→A) *in vitro*, phosphorylation of S^{465} was not restrained. This result suggests that presence of negative charge at S^{467} is enough to induce the second phosphorylation at S^{465}. Mutations at these two C-terminal serine residues ($S^{465/467}$) as well as at all three serine residues ($S^{464/465/467}$) of Smad2 in independent studies, blocked nodal-signaling (Macías-Silva et al., 1996; Souchelnytskyi et al. 1997).

Substrate specification is an important event at the onset of Nodal signaling. The L45 loop structure in type I receptor and L3 loop in Smad2 or Smad3 are involved in substrate specificity. Such specific structural interactions make TGF-β and Activin receptors specifically phosphorylate the R-Smads like Smad2/3, whereas Smads1/5/8 are phosphorylated by bone morphogenetic protein or BMP-pathway receptors (Chen et al., 1998). Unphosphorylated Smad2/3 proteins usually reside in the cytoplasm as monomers.

Phosphorylation of the SSXS motif leads to homo-oligomerization, which in turn quickly forms hetero-oligomers along with the co-Smad protein, Smad4 (Correia et al., 2001; Kawabata et al., 1998). Inactive Smad2 proteins in the cytoplasm are intrinsically auto-inhibited via intramolecular interaction between the MH1 and the MH2 domains (Hata et al., 1997), and the presence of a unique loop structure in the MH2 domain of Smad4 protein inhibits its oligomerization in absence of any signals (Tada et al., 1999). Receptor mediated phosphorylation induces conformational changes, and that is sufficient to remove the auto-inhibitions.

D. Shuttling between cytoplasm and nucleus:

The proper anchoring of Smad2 proteins to the receptor complex involves a cytoplasmic protein called Smad Anchor for Receptor Activation (SARA) and the adaptor molecule Dab2 that specifically forms a bridge between the inactive Smad2 and the receptor complex (Fig. 1). This is an essential scaffolding step for proper activation of the Smad2 protein. This stable interaction in turn inhibits the nuclear import of Smad2 (Tsukazaki et al., 1998; Xu et al., 2000; Ross and Hill, 2008).

Type I receptor phosphorylation of Smad2 disrupts the association between SARA and Smad2, resulting in the release Smad2 from the scaffolding complex. This event leads to the subsequent exposure of the nuclear localization signal (NLS) on Smad2 MH2 domain (Xu et al., 2000) and makes Smad2 available to interact with Smad4 (Shi and Massagué, 2003). Receptor-mediated phosphorylation of R-Smads is also reported to augment its affinity for Smad4 (Shi and Massagué, 2003).

All R-Smads, as well as co-Smads, such as human Smad4 and *X. laevis* Smad4α are cytoplasmic, whereas *X. laevis* Smad4β along with I-Smads reside inside the nucleus (Howell et al., 1999; Shi and Massagué, 2003). In the case of *X. laevis*, the hetero-oligomerization is preferred between Smad2 and Smad4α. The MH1 domain of Smad3 contains a lysine rich motif, which acts as NLS (Xiao et al., 2000; Xiao et al., 2001). In the case of Smad3, C-terminal phosphorylation leads to conformational changes exposing the NLS, which is then bound by importinβ1 and undergoes Ran-dependent nuclear translocation (Xiao et al., 2000). In the case of Smad2, the nuclear import is NLS-independent, and it occurs in a cytosolic-factor-independent manner, involving the MH2 domain (Xu et al., 2000). The difference in the import mechanisms between Smad2 and Smad3 is due to the presence of the unique exon 3 in Smad2 MH1 domain (Kurisaki et al., 2001). Inside the nucleus, Smad3 and Smad4 can bind directly, but with a lower affinity to the Smad-binding elements (SBEs), which have a minimal motif of 5'-CAGAC-3'. Such binding is mediated by a conserved β-hairpin loop and an additional α-helix 2 structures present in the MH1 domain (Kusanagi et al., 2001). Smad2, in contrast, lacks the DNA binding ability to SBE because of its unique exon 3 encoded sequence in MH1 (Yagi et al., 1999).

An alternative mechanism of Smad2/3 nuclear import and export is based on the finding that R-Smad MH2 domain can interact directly with the components of nuclear pore complex, especially CAN/Nup214 and Nup153. Direct contact between the nuclear pore complex and Smad2/3 allows the R-Smads to undergo constant shuttling (Liu et al., 1997; Inman et al., 2002; Xu et al., 2002).

E. Downstream targets:

Experiments in *X. laevis* embryos have identified two classes of transcription factors that are involved in the recruitment of phosphorylated Smad2-Smad4 complexes to the promoters of Nodal signaling target genes. The transcription factor FoxH1a and FoxH1b can recruit the complex to ARE within the promoters of genes like *Mix.2* (Chen, 1996; Howell, 2002, Ross and Hill, 2008). *Mixer*, *Milk* (*Bix2*) and *Bix3*, members of the *X. laevis Mix* transcription factor family, can also recruit the complex to the distal elements within the promoter region of mesoderm specific genes like *goosecoid* (Germain et al., 2000; Randal et al., 2002). Although these different transcription factors interact with activated Smad2 via a common proline-rich motif called Smad interaction motif (SIM), FoxH1 factors possess the Fast/FoxH1 motif (FM), which helps them to interact specifically with Smad2 (Randall et al., 2004). At the end of its performance as a transcription factor, Smad2 dissociates from the complex followed by dephosphorylation. Finally the inactive Smad2 is exported back to the cytoplasm (Derynck and Zhang, 2003).

F. Linker region acting as a bridge between Nodal signaling and the Mitogen-Activated Protein Kinase pathway:

Phosphorylation of Smad2 by mitogen-activated protein kinases (MAPKs) can also have an effect on TGF-β signaling. MAPK phosphorylation sites are usually abundant in the linker region between the MH1 and the MH2 domain of Smad2/3 and Smad4, with some lying in the MH1 domain (Fig. 2). Smad2/3 and Smad4 are usually phosphorylated at proline directed serine/threonine residues by MAPKs Erk1 and Erk2

and also by stress-activated MAP kinase p38 and JNK (Javelaud and Mauviel, 2006). These MAPK induced phosphorylations of Smad2 can have different effects. Phosphorylation of T8, T220, S245, S250 and S255 on Smad2 linker region and T179, S204, S208, and S213 on Smad3 linker region by mitogens such as EGF, hepatocyte growth factor (HGF), MEK kinase 1 (MEKK1), JNK stimulate TGF-β signaling, showing a synergy between the TGF-β and MAPK signaling (Funaba et al., 2002; LaBonne and Whitman 1994). Smad2-dependent formation of mesoderm in *X. laevis* coincides with the regions of peak MAPK activation (Schohl and Fagotto, 2002). In contrast, oncogenic Ras, which stimulates Erk1 and Erk2, inhibits TGF-β signaling. In this case, Smad2/3 linker phosphorylation by MAPK blocks their transcriptional activity and nuclear accumulation (Kretzschmar et al., 1999) (Fig. 2). Finally, such phosphorylation of Smad2/3 leads to the proteasome regulated degradation.

V. Smad2Δexon3, an alternatively spliced variant of Smad2:

Although there is 91% amino acid sequence identity between Smad2 and Smad3, their biological activities are different. Unlike Smad3 and Smad4, Smad2 cannot bind directly to DNA. Kurokawa et al. (1998) showed that Smad3 can bind the Evi-1 transcriptional regulator, but Smad2 cannot. Although both Smad2 and Smad3 are essential signal transducers of the TGF-β pathway and TGF-β is more potent than Activin. Studies in HaCaT cells revealed that TGF-β can phosphorylate both Smad2 and Smad3, but Activin can only phosphorylate Smad3 (Shimizu et al., 1998). These functional differences were explained by the presence of an alternatively spliced variant

of Smad2 protein, called Smad2Δexon3, which is present in various cell types including *X. laevis* (Yagi et al., 1999; Faure et al., 2000).

A 30 amino acid sequence is present in the Smad2 MH1 domain, which is not present in that of Smad3. There are 11 exons that compose the *Smad2* gene ORF. Exon 3 is translated into these 30 amino acids, which are in the middle of MH1 domain. Although the functional superiority of Smad2Δexon3 over wild type Smad2 in terms of DNA binding has been experimentally demonstrated, the overall transcriptional activation by the two isoforms was similar. Although the presence of an extra 30 amino acids inhibits Smad2's direct DNA binding via the MH1 domain, it cannot block activated Smad2 from its indirect binding to DNA. DNA-binding factors like FAST1 or FAST2 interact and bind the MH2 domain of Smad2. Functionally, Smad2Δexon3 is equivalent to Smad3 and may substitute for it in certain cell types lacking Smad3 expression (Yagi et al., 1999).

VI. Smad4, the co-Smad:

Smad4, a member of common-mediator Smad (Co-Smad) family of proteins, binds receptor-activated Phospho-Smad2 (PSmad2) followed by nuclear translocation. This complex mediates the endoderm/mesoderm specific transcriptional regulation in *X. laevis*. Smad4 was originally identified as the product of *DPC* tumor suppressor gene in pancreatic cancer cells (Hahn et al., 1996). Involvement of Smad4 in TGFβ signaling is not only suggested in mammals (Lagna et al., 1996; Zhang et al., 1996) but also held as an important mediator of Activin- and BMP-signaling in *X. laevis* embryos (Lagna et al.,

1996; Zhang et al., 1996). Smad4 interacts with activated R-Smads via its C-terminal domain (Hata et al., 1997; Wu et al., 1997).

In *X. laevis*, two isoforms of Smad4, α and β, have been identified (LeSueus and Graff, 1999; Masuyama et al., 1999; Howell et al., 1999; Hill, 2001). *XSmad4α* is the potential *X. laevis* ortholog of *hSmad4*. Its ORF is 1650 bp in length producing a 549 amino acid protein with a predicted molecular weight of 59.8 kDa. *XSmad4*β, also known as *Smad10*, is not an alternatively spliced isoform, but a product of a different gene (Watanabe et al., 2000). It has a 1683 bp ORF, which produces a 560 amino acid protein with a predicted molecular weight of 61.2 kDa.

According to Howell et al. (1999), these two isoforms show strikingly different temporal expression patterns in early *X. laevis* embryos. Adult tissues were also characterized by their expression in different ratios, which suggests isoform specific roles. From MBT until the onset of midgastrula stages, expression of XSmad4β is much higher than that of XSmad4α, whereas at the midgastrula stages, the reverse expression pattern for the two isoforms has been observed (Howell et al., 1999). XSmad4β protein expression predominates during the early embryonic development in *X. laevis* and was part of two different activin induced Smad/transcription factor complexes, which bind ARE on *Mix2* gene. Moreover, TGF-β family members regulating transcription differ between these two time frames (Howell et al., 1999). Therefore, it is predicted that during early *X. laevis* embryonic development, maternal Activin-like and/or BMP signaling targets a set of zygotic gene expression via XSmad4β. At midgastrulation, when other TGF-β members like *derriere* (Sun et al., 1999), *Xnr-1*, *Xnr-2*, and *Xnr-3* (Jones et al.,

1995; Smith et al., 1995) and *BMP4* (Nishimatsu et al., 1992) are expressed, the pathway is mediated by XSmad4α.

Although the nuclear import of R-Smads does not need Smad4 (Liu et al., 1997), it co-translocates with R-Smads (Hoodless et al., 1999; Liu et al., 1997; Watanabe et al., 2000). Unlike the ligand-dependent import of R-Smads, Smad4 shuttles continuously between the cytoplasm and the nucleus because of the combined effects of the NLS in the MH1 domain and an NES in its linker region (Inman et al., 2002). Surprisingly, XSmad4β, not XSmad4α, lacks the NES. Therefore XSmad4β is almost exclusive to the nucleus, whereas XSmad4α resides in the cytoplasm in the absence of any signal due to the presence of NES (Masuyama et al., 1999). Also, XSmad4β and not XSmad4α, is phosphorylated, but the purpose is unknown (Masuyama et al., 1999).

Pierreus et al. (2000) and Watanabe et al. (2000) provided evidence that Smad4 undergoes nucleocytoplasmic shuttling on its own without involving R-Smads. The common Smad4 (Smad4α in *X. laevis*) possesses multiple motifs in its MH1 domain in addition to the NES in the linker region, which helps Smad4 in nuclear translocation without Smad2/3 (Xiao et al., 2003).

Smad2-independent nuclear translocation of Smad4 has been reported in several instances, especially when it serves as a co-Smad in multiple signaling pathways. In response to BMP4 signals, Smad4 binds with activated Smad1, not Smad2, and undergoes nuclear translocation (Lagna et al., 1996). There are also several instances, where Smad4 translocates inside the nucleus without being in a complex with active Smad2 or Smad3 in a TGF-β-inducible manner. SMIF, a Smad4-interacting protein

forms a TGF-β/BMP4-inducible complex with Smad4 and translocates inside the nucleus in a Smad4 dependent manner (Bai et al., 2002).

VII. Mesendoderm specification in *E. coqui*: A comparative analysis.

E. coqui excludes the feeding tadpole from its ontogeny and the young froglet emerges directly from the embryo (Callery et al., 2001). To compensate absence of a larval phase, the eggs are supplied with enough yolk to support the embryo to reach their froglet stage (Fang et al., 2000; Elinson and Beckham, 2002). As a consequence, the first and foremost difference in the early development between *E. coqui* and *X. laevis*, a metamorphosing frog, is the huge size of the egg in *E. coqui* (Elinson, 1987). The presence of a large amount of yolk shifted the early embryogenic events towards the animal pole of the zygote.

A. Early events:

Unlike *X. laevis*, the first horizontal cleavage in *E. coqui* occurs at the 16-cell stage (Elinson and Beckham, 2002). This cleavage produces eight small animal blastomeres, which represent barely 1% of the total embryo volume (Ninomiya et al., 2001). Although relatively small in size, these blastomeres contribute to both the mesoderm and ectoderm. Due to the presence of so much yolk, the cleavage furrows extend very slowly resulting in large, incompletely divided, vegetal yolky cells. The appearance of the blastopore lip equatorially, much closer to the animal pole than that in *X. laevis*, marks the onset of gastrulation (Beckham et al., 2003). The mesendoderm inducing activity, presumably under the influence of TGF-β signaling is localized

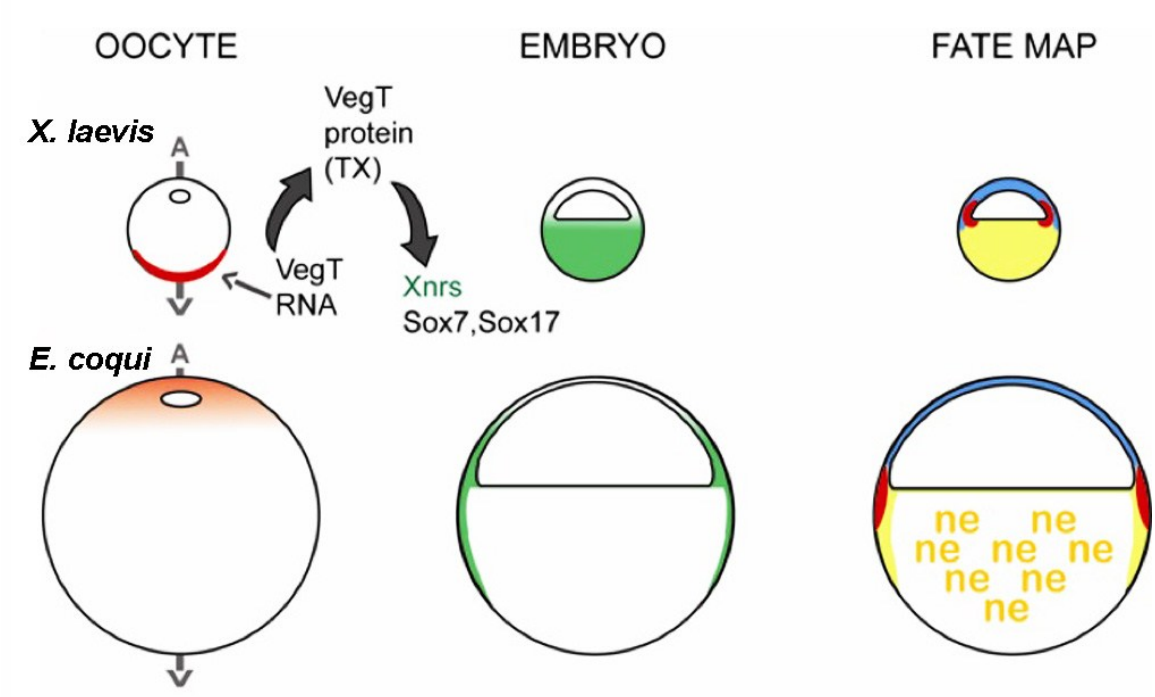
equatorially and sub-equatorially in the early gastrula and restricted to the superficial marginal zone cells (Ninomiya et al., 2001). Consequently, most of the vegetal region of the early gastrula lacks mesendoderm-inducing activity (Fig. 3) (Ninomiya et al., 2001). Unlike *E. coqui*, the mesoderm-inducing activity is expressed throughout the vegetal region of a *X. laevis* gastrula.

B. Mesendoderm inducers - VegT, Vg1 and Sox17:

The key initiator of both mesoderm and endoderm in *X. laevis* is the T-box transcription factor VegT. Spatial restriction of maternal *VegT* and *Vg1* RNAs to the vegetal cortex in the oocyte leads to the mesendoderm induction in vegetal half (Fig. 3). In *X. laevis*, VegT is necessary and sufficient to regulate endoderm formation by controlling *Nodal* gene expression and by directly activating transcription of endodermal determinants (Zhang et al., 1998; Clements et al., 2001; Xanthos et al., 2001). Zhang et al. (1998) and Xanthos et al. (2001) provided evidence that depletion of maternal *VegT* RNA causes lack of endodermal markers and mesoderm inducing signals in early embryos. This implicated the essential role of VegT in endoderm and mesoderm specification in *X. laevis* early embryos (Zhang et al., 1998; Xanthos et al., 2001). VegT is the important regulator of several TGF-β signals including Derriere and Xnrs 1, 2, 4, 5, and 6 (Clements et al., 1999, 2003; Kofron et al., 1999; Rex et al., 2002; Sun et al., 1999; Takashi et al., 2000). This interplay between VegT and Nodals is absolutely necessary for mesoderm induction (Hemmati-Brivanlou and Melton, 1992: Piccolo et al., 1999) and sustenance of endoderm specific gene expression (Zhang et al., 1998; Clements et al., 1999, 2003; Yasuo and Lemaire, 1999; Chang and Hemmati-Brivanlou, 2000; Engleka et

Fig. 3: Comparison of fate maps between *X. laevis* and *E. coqui* (in sagittal views, drawn to scale). In *X. laevis, VegT* and *Vg1* RNAs (red), localized to the oocyte vegetal (V) cortex, lead to Nodal signaling (green) in the vegetal half. This leads to endoderm (yellow) and mesoderm (red) in the fate map. In *E. coqui*, *EcVegT and EcVg1* RNA location is near the animal pole and Nodal signaling is hypothesized to be restricted to the peripheral marginal and sub- marginal zones. Absence of VegT activity and Nodal signaling, in the vegetal core is hypothesized to lead to nutritional endoderm (ne) development. (Modified after Elinson and del Pino, 2011, with permission)

Fig. 3:


OOCYTE
EMBRYO
FATE MAP
X. laevis
A
VegT
protein
(TX)
VegT
RNA
Xnrs
Sox7,Sox17
E. coqui
A
ne ne
ne ne ne
ne ne
ne

al., 2001; Xanthos et al., 2001). Clements et al. (2003) suggested a strong synergism between VegT and TGF-β signals, which is used by VegT to augment the cell's sensitivity towards TGFβ signals. VegT induces several other downstream targets, which are implicated in the specification of endoderm. These targets include Mix/Bix/Mixer family members, GATAs 4, 5, and 6, and Xsox17α and β (Hudson, 1997; Tada et al., 1998; Casey et al., 1999; Clements et al., 1999, 2003; Weber et al., 2000; Xanthos et al., 2001).

In the *X. laevis* embryo, a TGF-β superfamily member, *Vg1*, is expressed as maternal RNA in the vegetal hemisphere cells, which produce endoderm and induce mesoderm formation (Rebaglaiti et al., 1985; Weeks and Melton, 1987). Mutant Vg1 ligand studies by Joseph and Melton (1998) provided evidence that blocking Vg1 signaling results in embryos lacking dorsal mesoderm and axial structures. A large amount of Vg1 precursor proteins are synthesized and localized in the vegetal half of the *X. laevis* oocyte, but are only selective processed after fertilization to produce the mature protein in response to cortical rotation event (Thomsen and Melton, 1993). This small amount of potent mature Vg1 protein then initiates dorsal mesoderm formation and the turning on of several endodermal marker genes (Gamer and Wright, 1995; Henry et al., 1996). One such gene is *Xlhbox8*, which encodes a transcription factor expressed in vegetal core (VC) cells and later in pancreas (Gamer and Wright, 1995; Henry et al., 1996; Wright et al., 1988).

In *X. laevis* oocytes, *VegT* regulates the expression of the endoderm specific transcription factors like Sox7 (Zhang et al., 2005) and Sox17 (Yasuo & Lemaire 1999; Clements et al., 1999; Clements & Woodland 2003; Howard et al., 2007; Zhang et al,

2007). Sox7 and Sox17 are F-type Soxs, a group of Sox proteins based on degree of sequence similarities within their HMG domains (Zhang et al, 2007). In *X. laevis*, *Sox7* mRNA is maternally contributed and localizes to the vegetal half of the early embryo (Zhang et al, 2007). Zhang et al. (2005a) showed VegT regulates the expression of Sox7 and blocking *Sox7* expression leads to the inhibition of VegT's effects in animal cap. Zhang et al. (2005a, 2005b) also illustrated the direct involvement of Sox7 in regulation of expression of mesoderm-inducing Nodal-genes such as *Xnr1*, *Xnr2*, *Xnr4*, *Xnr5*, and *Xnr6*.

The other F-type Sox, Sox17, is involved exclusively in endodermal specification. Hudson et al. (1997) showed that two transcription factors, Sox17α and Sox17β, are expressed in the whole vegetal half fated to form endoderm in the *X. laevis* embryo. Nodal-signaling via intermediates like Mixer-like and GATA proteins in *X. laevis* activates Sox17 expression (Engleka et al., 2001; Shivdasani, 2002; Clements and Woodland, 2003; Loose and Patient, 2004; Zhang and Klymkowsky, 2007). Unlike Sox7, Sox17 activates only *Xnr4* (Sinner et al., 2004; Zhang et al., 2005b; Zhang and Klymkowsky, 2007). It also cooperates with β-catenin in initiation of transcription of endoderm specific target genes in *X. laevis* (Sinner et al., 2004).

In animal cap explants, over-expression of *Sox17* augmented endodermal gene expression (Stainier, 2002). Disruption of *Sox17* expression in vegetal cells fated to become endoderm, forced the cells to enter mesodermal or ectodermal lineages (Clements and Woodland, 2000; Stainier, 2002). According to Engleka et al. (2001), both VegT and Nodal are required for *Sox17α* expression at gastrula, but only VegT is necessary to initiate *Sox17α* expression at MBT. They also reported that misexpression of

Sox17α in the marginal zone of the early embryo led to the inhibition of expression of mesodermal genes. This result suggests incompatibility of Sox17α with mesoderm specification. Sox17α also blocks expression of mesodermal markers like *Xbra* and *MyoD* (Engleka et al., 2001). These results corroborate the idea that activation of Sox17α by VegT in the vegetal hemisphere of *X. laevis* oocyte (Fig. 3) defines the endodermal domain, and the response of Sox17α to Nodal signal is biased against mesoderm specification (Engleka et al., 2001).

EcVegT and *EcVg1* cDNAs, the *E. coqui* orthologs of *X. laevis VegT* and *Vg1*, were cloned in our laboratory. In situ hybridization in full-grown *E. coqui* oocytes detected presence of the *EcVegT* and *EcVg1* RNAs near the animal pole (Beckham et al., 2003) in contrast to *X. laevis*, where *VegT* and *Vg1* RNAs are at the vegetal cortex (Zhang and King, 1996; Melton, 1987). The results in *E. coqui* were confirmed by RT-PCR analysis (Beckham et al., 2003). In situ hybridization in *E. coqui* early embryos, detected zygotic *EcVegT* expression in the marginal zone (MZ) during early gastrulation (Beckham et al., 2003). The expression of *EcVegT* in the marginal zone of the embryo became stronger by mid-gastrulation (Beckham et al., 2003).

Although, in situ hybridization data on *EcSox17* RNA met the expectation of being present only in the marginal zone of the *E. coqui* early embryos, RT-PCR results revealed the presence of the RNA in both the MZ and the VC (Buchholz et al., 2007). The maternal contribution of *EcSox17* in *E. coqui* could account for the presence of *EcSox17* RNA in the VC. *Sox17* RNA is not present in *X. laevis* eggs (Hudson et al., 1997). The function of *EcSox17* RNA in vegetal cells remains elusive. As a consequence, *EcSox17* expression data could not be used to distinguish NE from DE.

Recombinant analysis by Ninomiya et al. (2001) revealed the presence of strong mesoderm inducing activity in the outermost cells of the MZ of *E. coqui* early embryos. This study did not detect mesoderm-inducing activity in the blastocoel floor and VC of the embryo. Expression of *Brachyury* (*EcBra*), the *E. coqui* homolog of *X. laevis Xbra*, an early mesoderm marker, was detected around the entire MZ (Ninomiya et al., 2001).

Based on these data, the *E. coqui* fate map (Fig. 3) was generated (Elinson and del Pino, 2012). Unlike *X. laevis*, *EcVegT* and *EcVeg1* RNA localization towards the animal pole leads to mesendoderm specification towards the marginal zone in the fate map and absence of any mesendoderm inducing activity in the VC cells from late blastula and early gastrula (Elinson and del Pino, 2012).

C. Germ plasm in *E. coqui*:

In spite of the deviations described above, localization of germ plasm to the vegetal cortex is conserved in both *E. coqui* and *X. laevis*. Germ cells arise early in frog development due to this cytoplasmic localization. Frog germ plasm contains mitochondria, an electron dense nuage material, and many germ plasm specific RNAs. Germ plasm RNAs in *X. laevis* include *dazl*, *nanos1*, *Xpat*, and *DEADSouth* (King et al., 2005). These RNAs are localized at the vegetal cortex during oogenesis. After fertilization, they form islands of germ plasm, and during cleavage, these islands segregate to a small number of cells. Such cells are called primordial germ cells (PGCs), and they migrate from the endoderm to the genital ridges of the future gonads.

In *E. coqui*, islands of germ plasm were identified by $DiOC_6$ staining and by *Ecdazl* and *EcDEADSouth* RNA localization (Elinson et al., 2011). The islands cover a

large part of the vegetal cortex of the cleaving embryo. These distributions of RNAs, although similar to that in *X. laevis*, are comparatively more extensive.

During gastrulation in *X. laevis*, cells containing germ plasm maintain totipotency despite the presence of maternal *VegT* RNA. They are prevented from acquiring an endodermal fate, and finally give rise to the PGCs (Venkatarama et al., 2010). The fate of the germ plasm containing cells is determined by a global transcriptional repression of those specific blastomeres. The other blastomeres, which are not transcriptionally repressed, become committed to the endodermal fate. Both the endodermal and PGC lineages originate from the vegetal cytoplasm in *X. laevis*, but the latter do not respond to VegT (Venkatarama et al., 2010).

VIII. Nutritional endoderm, the novel tissue in *E. coqui*:

E. coqui eggs, with elevated levels of yolk, have a large vegetal region devoid of molecular determinants (Ninomiya et al., 2001; Beckham et al., 2003). The absence of molecular determinants in the VC is responsible to form the novel tissue NE in this direct developing frog (Buchholz et al., 2007). Presence of NE in *E. coqui* dramatically changes the fate map relative to the *X. laevis* fate map (Fig. 3).

NE consists of transient, yolk rich cells that provide nutrition to the growing embryo but do not differentiate into any definitive endodermal tissues. NE is grossly a mass of large yolky cells attached to the differentiated intestinal tube. These cells disappear from the differentiated intestine after the yolk is utilized (Buchholz et al., 2007). There are several lines of evidence that support the idea that NE performs only a nutritional role. Lack of endoderm specific molecular determinants like VegT from NE

provides evidence that NE does not behave like regular endodermal cells. When the gut of a newly hatched froglet was examined, there were two types of grossly different tissues. One of them is the translucent tissue, which forms the differentiating stomach and intestine, and the other type is the yolky NE (Buchholz et al., 2007). Histological examination emphasized the sharp distinction between the two tissues. Fate mapping, performed by injecting FDA into the large vegetal cells of cleaving embryos, resulted into a strong labeling of only the large yolky cells, not the differentiated gut tissue (Buchholz et al., 2007). Instead of differentiating, the mass of labeled yolky cells decreased in number, and label remained only in the mesonephros and cloaca. The labeling of mesonephros indicated that the FDA was being cleared from blood, which suggests FDA uptake into circulation when the labeled cells died. This result suggested death as the fate of these large yolky cells. In attempt to show death in NE cell, Buchholz et al. (2007) performed TUNEL staining, but failed to detect labeled cell nuclei. Also, nuclear histological changes associated with apoptosis in *X. laevis* such as nuclear fragmentation or chromatin condensation (Ishizuya-Oka and Ueda, 1996), were not observed in *E. coqui* (Buchholz et al., 2007).

It is evident that NE lacks the signaling pathways and events related to mesendoderm specification. Whether NE is present in indirect developers like *X. laevis*, *Rana,* or other frogs with tadpoles, has not been determined. Detection of such cells, which would be only a minor component of the yolk rich endoderm, would require a more detailed fate mapping than has been done so far (Buchholz et al., 2007; Elinson, 2008).

IX. Global transcriptional status:

NE cells disappear in the course of development and do not contribute to any differentiated tissues in *E. coqui* froglets. In contrast to MZ cells, VC cells during early *E. coqui* embryogenesis lack expressions of molecular determinants of mesendoderm inducers such as *EcVegT*, *EcVg1*, and mesodermal marker like *EcBra*. These results could account for the lack of mesendodermal development for the VC, but the presence of *EcSox17* RNA in VC cells suggests a way to bypass the usual endodermal signaling routes. It is expected that the presence of transcription factors in *E. coqui* early embryos should also perform their anticipated role, but that depends on the functional status of the most important transcribing enzyme, RNA polymerase II.

A. Regulation of RNAP II CTD-phosphorylation in transcription:

RNA polymerase II (RNAP II) is a eukaryotic enzyme that catalyzes transcription of mRNA, snRNA, and microRNA. In other words, the most important enzyme for gene expression is RNAP II. RNAP II is a 550 kDa multiprotein complex, with 12 subunits. It interacts with a wide range of transcription factors that are required for it to bind to its promoters and begin transcription. The C-terminal domain (CTD) of RNAP II is an unusual extension, which is attached to the C-terminus of the largest subunit and serves as a flexible binding scaffold for transcription factors (Phatnani and Greenleaf, 2006). The transcription factors to be bound depend on the phosphorylation status of the CTD repeats. This repeat domain is evolutionarily conserved. In fungi, plants, and animals, the CTD contains 25 to 52 tandem copies of the consensus heptapeptide repeat $Y_1\mathbf{S_2}P_3T_4\mathbf{S_5}P_6S_7$ (Corden, 1990; Phatnani and Greenleaf, 2006; Venkatarama et al.,

2010). Regulation of phosphorylation of this heptapeptide repeat determines the transcriptional status of a cell (Venkatarama et al., 2010). Unphosphorylated RNAP II is recruited to a promoter to form the initiation complex (Fig. 4). Within the pre-initiation complex, each of the CTD-serine 5 (P-Ser5)'s is phosphorylated by the cyclin dependent kinase CDK7 (Price, 2000; Phatnani and Greenleaf, 2006; Sutherland and Bickmore, 2009; Venkatarama et al., 2010). P-Ser5 marks transcription initiation. This is followed by transcription elongation, which occurs only when the CTD-serine 2 (P-Ser2)'s are phosphorylated by cyclin-dependent kinase CDK9 (Fig. 4) (Price, 2000; Phatnani and Greenleaf, 2006; Sutherland and Bickmore, 2009; Venkatarama et al., 2010).

Based on the reports of Phatnani and Greenleaf (2006) and Sutherland and Bickmore (2009), a working model for transcription initiation and elongation correlated with CTD phosphorylation has been proposed (Fig. 4). According to the model, the RNAP II CTD repeat $Y_1\mathbf{S_2}P_3T_4\mathbf{S_5}P_6S_7$ exists in four phosphorylation states along the body of a gene during expression. First, in the pre-initiation complex at the gene promoter, the CTD repeat is hypo-phosphorylated (non-P), and the enzyme is associated with a set of general factors and mediators. Second, the Ser5 residue of the repeat is phosphorylated (Ser5P) by CDK7 or by the kinase activity of the general transcription factor (GTF), TFIIH. This event leads to transcription initiation at the 5' end of the gene, accompanied by the binding of a 5'end-processing factor like capping enzyme. Third, an elongation-phase kinase like CDK9 phosphorylates Ser2 exclusively to create a dually phosphorylated state of CTD (Ser2P,5P), which is essential for elongation. Finally, towards the 3'end of the gene, Ser2 residues dominate, while Ser5 positions are mostly

dephosphorylated by protein phosphatase (PPase) (Phatnani and Greenleaf, 2006; Sutherland and Bickmore, 2009).

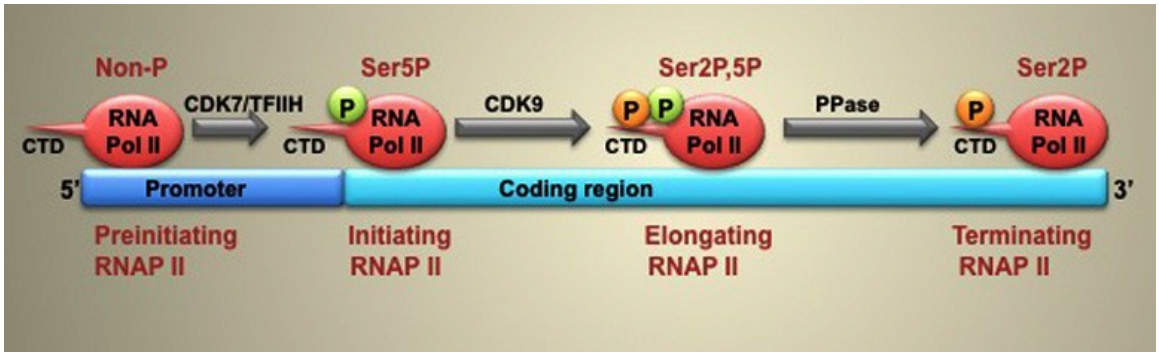
B. Global transcriptional status as a fate determinant:

In *D. melanogaster* and *C. elegans*, global repression of transcription is a well-known phenomenon, preventing the germ cell lineage from acquiring the somatic fate (Seydoux et al., 1996; Van Doren et al., 1998; Strome and Lehmann, 2007). Use of specific monoclonal antibodies H14 and H5 against P-Ser5 and P-Ser2, respectively, revealed that global gene expression is very low or absent in the germline, whereas the neighboring somatic cells showed global gene expression (Seydoux and Dunn, 1997; Ghosh and Seydoux, 2008). In these invertebrate model systems, the RNAPII activity is blocked in the germline temporarily via regulation of the phosphorylation state of the RNAPII CTD. In *C. elegans*, involvement of protein factors like OMA-1/OMA-2 (Guven-Ozkan et al., 2008) and PIE-1 (Seydoux et al., 1996) leads to the complete inhibition of CDK7 and partial blockage to CDK9 functioning transcriptional repression targeting both initiation and elongation steps (Ghosh and Seydoux, 2008). In *D. melanogaster*, repression of CTD-Ser2 phosphorylation is inhibited by expression of *polar granule component* (*pgc*). Polar Granule Component is reported to interact with CDK9 and prevent it from entering the transcription site (Hanyu-Nakamura et al., 2008). In *pgc*-null mutants stocks of *D. melanogaster*, poles cells become degraded before reaching the gonad (Hanyu-Nakamura et al., 2008; Timinszky et al., 2008).

Global transcriptional repression also occurs in the germ line precursor cells in the ascidian *Halocynthia roretzi* (Tomioka et al., 2002). Recent reports in mouse strongly

Fig. 4: A working model for the existence of four phosphorylation states of RNAP II CTD. The RNAP II CTD repeat $Y_1\mathbf{S_2}P_3T_4\mathbf{S_5}P_6S_7$ is hypo-phosphorylated (non-P) in the pre-initiation complex at the promoter of a gene. Ser5 phosphorylation (Ser5P) by CDK7 or TFIIH is responsible for transcription initiation. CDK9 phosphorylates Ser2 exclusively and a dually phosphorylated state of CTD (Ser2P,5P) leads to elongation. At the 3'end of the gene, Ser2P remains active, Ser5P is de-phosphorylated by PPase. (Based on Phatnani and Greenleaf, 2006; Sutherland and Bickmore, 2009).

Fig. 4:


Non-P
Ser5P
Ser2P,5P
Ser2P
CDK7/TFIIH
CDK9
PPase
CTD
RNA Pol II
P
5'
Promoter
Coding region
3'
Preinitiating RNAP II
Initiating RNAP II
Elongating RNAP II
Terminating RNAP II

suggest the presence of global transcriptional repression as a means to protect PGC identity during migration in mammalian germline (Seki et al., 2007).

In *X. laevis*, a similar event has been documented, where PGCs are formed from cells containing germ plasm islands at the vegetal cortex, while the neighboring regions of the vegetal half become committed to endodermal fates (Venkatarama et al., 2010). Palancade et al. (2001) reported the existence of several distinct cycles of RNAP II CTD phosphorylation, which correlate with active transcription in *X. laevis*. At fertilization, the RNAP II CTD is almost completely de-phosphorylated with only a few phosphorylated forms being present, as detected by the monoclonal antibody (H14 and H5) staining. This correlates to a very low level of overall transcription. Twelve divisions later, at MBT, when zygotic gene expression kicks in, RNAP II CTD undergoes hyper-phosphorylation to a maximum level for both P-Ser5 and P-Ser2 (Venkatarama et al., 2010).

In *X. laevis* PGCs, although *VegT* RNA is present, its downstream targets are not expressed due to the global transcriptional block at MBT. The neighboring somatic cells at the same time showed the presence of hyper-phosphorylated RNAP II CTD and become committed to an endodermal fate. Such transcriptional blockage existed in PGCs for at least ten hours until neural tube has been specified (Venkatarama et al., 2010). Although transcriptional blockage in PGCs was accompanied by differences in the amounts of histone linker proteins and DNA methylation, no significant changes in chromatin remodeling were observed that could account for the global transcriptional repression (Venkatarama et al., 2010). Therefore, regulation via RNAP II CTD phosphorylation appears to be pivotal in maintaining the germline in both invertebrates

and vertebrates. The evolutionary conservation of this phenomenon makes it worthwhile to investigate whether such event also occur in *E. coqui* during early embryogenesis.

The presence of germ plasm has been documented in the vegetal half of *E. coqui* embryos (Elinson et al., 2011). Germ plasm markers such as *Ecdazl* and *EcDEADSouth* were detected towards the vegetal surface of *E. coqui* early embryos. Although, no correlation has been established between the vegetal location of germ plasm islands and generation of NE, it is enough to speculate that a global transcriptional repression event also could occur in *E. coqui* vegetal half. Such a repression could provide an explanation for why NE does not behave like definitive endoderm, despite the presence of endodermal determinants like EcSox17.

X. Current investigation:

Direct development in *E. coqui* is accompanied by the emergence of NE in addition to the three germ layers. Unlike *X. laevis*, NE develops from the VC of the embryo, whereas the neighboring MZ cells become committed to the mesendoderm fate. In this dissertation, I investigate the probable causes underlying the origin of NE by asking whether NE develops because of the absence of the molecular determinants of Nodal signaling in the VC in contrast to MZ. More specifically, I characterize the expression of the Nodal signal transducer, Smad2 during early embryogenesis. I also address the functional aspect of this tissue, which provides only nutrients to the growing embryos, but does not differentiate. Lack of differentiation raises the question of the functional inertness of these cells, which could be a consequence of the nature of its gene expression. My study includes the following aims:

(1) *Cloning the EcSmad2 gene*: The role of Smad2 in germ layer specification has been well documented in *X. laevis* and other vertebrate species, but nothing is known for direct developing amphibians. Smad2 is not only an essential component of Nodal signaling, but it also serves as a hub connecting many important signaling pathways. As the genome of *E. coqui* is not sequenced, I clone the Smad2 cDNA from *E. coqui* to examine its expression.

(2) *Establishing temporal and spatial expression patterns for EcSmad2*: The only temporal expression profile of Smad2 protein in embryos is available in *X. laevis*. Spatial and temporal RNA expression patterns have not yet been reported in any frog. It will be important to find out whether *EcSmad2* RNA is maternally contributed and whether it is present throughout all of the early developmental stages. It is possible that *EcSmad2* expression is spatially restricted to the MZ only. Altogether, it will be helpful to generate a complete profile of temporal and spatial *EcSmad2* expression.

(3) *Phospho-EcSmad2, EcSmad4 and ActRIIA protein expression in the prospective definitive endoderm vs. prospective nutritional endoderm:* The VC of *E. coqui* does not develop into definitive endoderm; rather it develops into the NE. If nodal signaling does not occur in the vegetal core of *E. coqui* blastulae and gastrulae, then Smad2 should be phosphorylated in the superficial MZ, but not in the VC. To test this hypothesis, I will analyze protein expression of both native and active forms of EcSmad2. Although a lot of work has been done on *X. laevis* Smad2 protein expression, a complete profile for temporal and spatial expression of Smad2 and PSmad2 does not exist. Creating such a profile is instrumental for our understanding of Nodal signaling in frogs, especially the direct developing ones. I will look at the expression profiles of the co-

Smad, Smad4, and one of the activin receptors ActRIIA. Although some work has been done with the receptor at the protein level, Smad4 protein expression has not been examined. Looking at all of these components of Nodal signaling in the prospective definitive endoderm vs. prospective NE will furnish us with new ideas about *E. coqui* early embryogenesis.

(4) *Detection of cellular location of PSmad2, Smad4 and receptor ActRIIA in prospective definitive endoderm vs. prospective NE:* The presence of an active form of protein does not necessarily indicate the pathway is active, especially when VC is not differentiating. It will be important to find out whether the active Smad2 or Smad4 proteins are cytoplasmic or are inside the nucleus. Identification of cellular locations of Smad2 and these other components of the pathway will provide crucial information to determine whether the pathway is active in NE.

(5) *Determination of transcriptional status of NE cells:* Finally, determination of transcriptional status of NE cells will help us understand better the nature of this tissue. Determining the transcriptional status may provide us with an explanation for why the VC cells of early embryos do not undergo differentiation. I will explore differences in transcriptional activity between those VC cells and those from the MZ.

CHAPTER TWO: MATERIALS AND METHODS

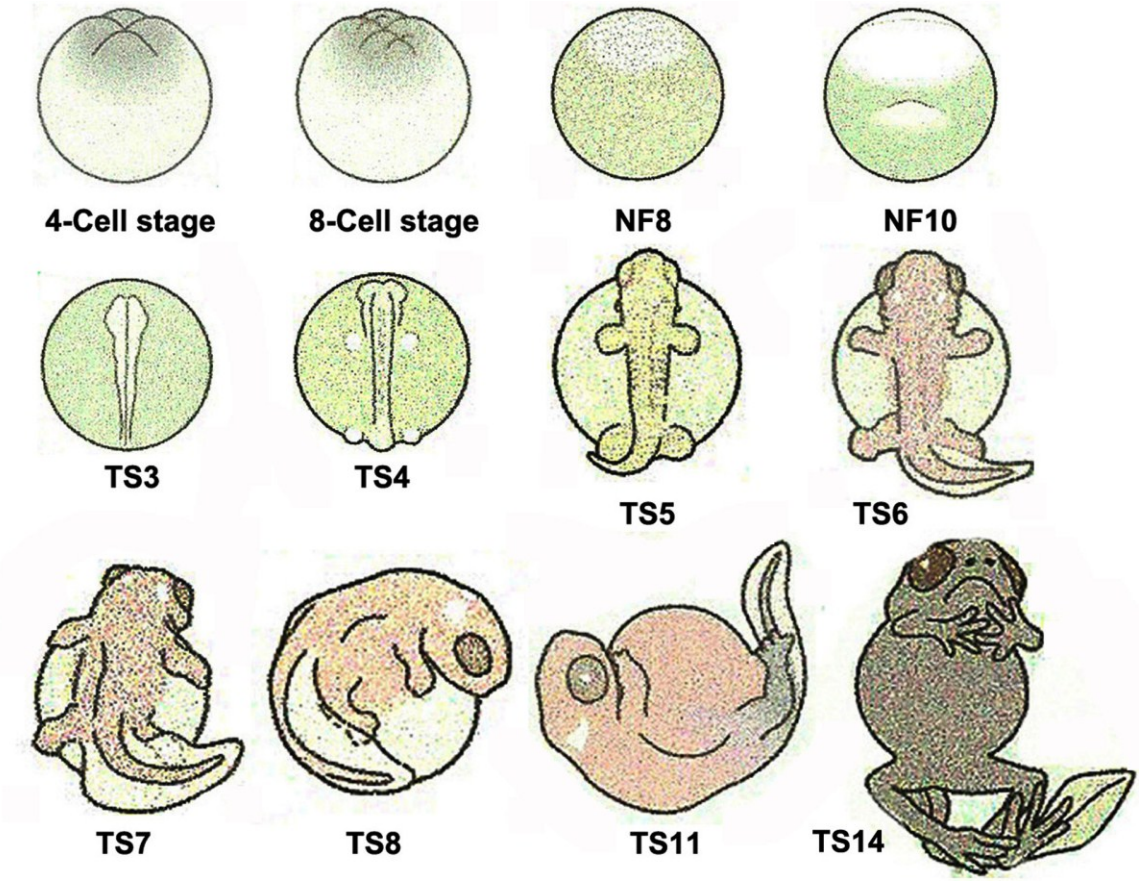
I. Animals, oocytes, and embryos:

Adult *E. coqui* frogs were captured on the Big Island of Hawaii following Injurious Wildlife Export permits from the Department of Land and Natural Resources of Hawaii. Male and female frogs were kept as pairs as a reproductive colony at Duquesne University. Adults and embryos were maintained following the protocols approved by the Institutional Animal Care and Use Committee (IACUC). Female frogs in each pair lay fertilized eggs in clutches, which are then guarded by the male frogs.

Clutches of embryos were collected and kept in a plastic petri dish with a filter paper soaked in 20% Steinberg's solution, made by dilution of a 200% Steinberg's solution. The 200% Steinberg's solution was made by adding 100 mls of two separate 20X stock solutions - A and B (Stock A = 1.16 M NaCl, 13.41 mM KCl, 16.59 mM $MgSO_4.7H_2O$, 6.7 mM $Ca(NO)_3$; Stock B = 10 mM Tris.HCl, pH 7.4) to a total 1000 ml volume with deionized water (dIH_2O). Embryos were gently separated from each other using watchmaker's forceps in order to ensure proper development. *E. coqui* embryos are staged according to the table by Townsend and Stewart (1985). Each stage is given a TS number starting from TS3 to TS15, with TS15 denoting the hatching froglets. Embryos from stages earlier than TS3 (neural tube formation) were staged according to the Nieuwkoop and Faber (1994) stages for *X. laevis*. Embryos with a thick animal cap, small blastocoel, and absence of a dorsal lip are considered to be at NF8. NF10 is characterized by the presence of a very thin animal cap, a large blastocoel, and the appearance of the dorsal lip (Fig. 5).

Fig. 5: Stages of *E. coqui* development. Cleavage is represented by 4-cell and 8-cells stages. NF8 represents a pre-gastrulation stage and NF10 is gastrulation. Eight post-gastrulation stages are TS3, TS4, TS5, TS6, TS7, and TS8, which were used for various experiments. Figures taken from Winter Edition of Biohistory Research Hall, 2012 (www.brh.co.jp)

Fig 5:


4-Cell stage
8-Cell stage
NF8
NF10
TS3
TS4
TS5
TS6
TS7
TS8
TS11
TS14

Ovarian oocyte collection was performed by removal of ovary from a sacrificed, reproductively retired female frog. A female frog was anesthetized in 0.1% MS222 (Tricaine methane sulfonate), pH 7.4 and killed by decapitation. Ovaries were dissected out and placed in 200% Steinberg's solution until ready to use. Full-grown oocytes were defolliculated with watchmaker's forceps after incubating for one hour in Ca^{2+} free 200% Steinberg's solution (116 mM NaCl, 1.34 mM KCl, 0.8 mM $MgSO_4$, 9.2 mM Tris, pH 7.4) with 1 mM EGTA (Beckham et al., 2003).

II. Dejellying and removal of fertilization membrane:

In order to perform experiments using molecular techniques, embryos at desired stages were transferred to a new petri dish. They were flooded with 20% Steinberg's solution to cause swelling of the outer jelly layer. Fine forceps (No. 5, Dumont Electronic quality, Switzerland) were used to remove the outer and middle jelly layers. Pre-gastrula stage embryos were subjected to additional treatments to remove the inner jelly layer and the fertilization membrane. First, they were incubated in 3% cysteine, pH 8 [L-(+)-Cysteine, Hydrochloride, Monohydrate; J.T.Baker] for 8 – 10 minutes with gentle swirling to remove the inner jelly layer. After thoroughly rinsing three to four times with copious amounts of 20% Steinberg's solution, embryos were incubated in Hennen's solution for 45 – 50 seconds to weaken the fertilization membrane. Hennen's solution consists of 2 % cysteine, 0.2% papain (Sigma) and 0.2% α-chymotrypsin (Type II, from bovine pancreas), pH 8.0 (Hennen, 1973). Embryos were immediately washed three to four times with excess amounts of 20% Steinberg's solution and transferred to a new petri dish containing 1% Bovine Serum Albumin (BSA, Fraction V, Heat Shock

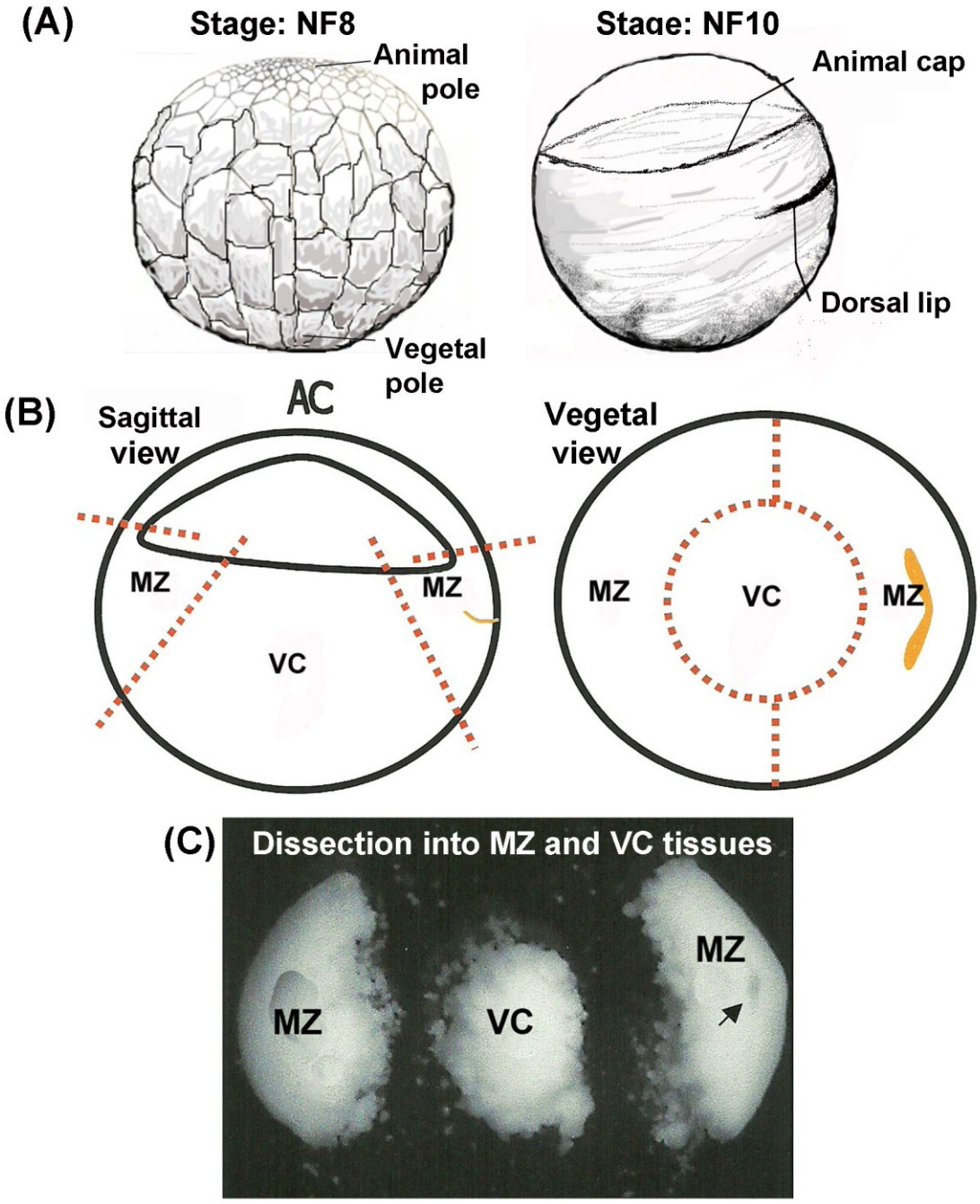
Isolation, OmniPur) in 100% Steinberg's solution. In this final medium, the weakened fertilization membrane was carefully removed with fine forceps under a dissecting microscope.

III. Dissection of embryos:

After dejellying and removing the fertilization membrane as described above, embryos were dissected in 100% Steinberg's solution with 1% BSA. Without fertilization membranes, pre-gastrulation and gastrulation stage embryos flattened and appeared like a flatbread. At these early stages, embryos were physically manipulated with the help of a loop made from human hair attached to the end of glass Pasteur pipette. First, the embryos were positioned with their animal pole facing upward. At NF8, it was easy to detect vegetal surface by the presence of bigger cells. A fine incision was made with forceps in a circle around the animal cap and the cap was removed. After this, the embryo was turned upside down bringing the vegetal pole up. Fine incisions were made using a hair loop to separate out the thin, peripheral MZ tissue (prospective definitive endoderm or DE) consisting of both dorsal and ventral sides from the larger inner VC tissue (prospective NE) (Fig. 6). Finally the MZ and the VC were separated into different microfuge tubes for further use. For NF 10 embryos, the presence of a more translucent animal cap and the appearance of the blastopore lip made the dissection of prospective DE and NE much easier. For post-gastrulation stages when the embryonic regions have already developed, dissections became much less complicated. For all post-gastrulation stages the whole embryo was dissected into the embryonic tissues and the NE as shown in Fig. 6.

Fig. 6: Dissection of blastula (NF8) and early gastrula (NF10) into MZ and VC. (A) Diagrams of NF8 and NF10 stages. (B) Diagram of the sagittal and the vegetal sections of NF8/NF10 embryos with the red dotted line indicating the positions of MZ and VC tissues and paths of dissections. (C) Photograph of the three pieces of dissected tissues corresponding to the vegetal view diagram. [Taken from Buchholz et al., 2007, with permission.]

Fig 6:


(A)
Stage: NF8
Animal pole
Vegetal pole
Stage: NF10
Animal cap
Dorsal lip
(B)
Sagittal view
AC
MZ
MZ
VC
Vegetal view
MZ
VC
MZ
(C)
Dissection into MZ and VC tissues
MZ
VC
MZ

IV. Dissociation and isolation of individual cells from different tissues:

Dissociation of MZ and VC cells was required for the immunostaining experiments. Once the MZ and VC were dissected, they were transferred to separate small petri dishes containing Ca^{2+}-Mg^{2+}-free modified 100% Steinberg's solution [68 mM NaCl, 1.34 mM KCl, 4.05 mM Na_2HPO_4, 0.73 mM KH_2PO_4, 1.0 mM EGTA.Na_2 with final pH 7.4] (modified from Ninomiya et al., 2001). Incubation for 10 – 15 minutes in this solution dissociated cells from their respective tissues, after which cells were washed in 1X PBS and were ready to use.

This solution only worked well for embryos from NF8, NF10, and TS3 stages. For more mature embryos with more definitive and rigid body structures, the dissociation medium was not enough even after incubating for more than an hour. For dissociation of NE cells from embryos at TS6, TS8, TS11, and TS14 stages, dissected NE tissues were incubated in Ca^{2+}-free 200% Steinberg's solution with 1.0 mM EGTA [116.4 mM NaCl, 1.342 mM KCl, 0.812 mM $MgSO_4.7H_2O$, 4.05 mM Na_2HPO_4, 0.73 mM KH_2PO_4, pH 7.4, along with 1.0 mM final concentration of EGTA (pH 7.0)] (modified from Shibuya and Masui, 1988) for at least an hour.

V. Trizol RNA Extraction:

RNAs were isolated from whole embryos as well as from dissected prospective DE and prospective NE following the Trizol RNA extraction protocol standardized in our laboratory. Five dejellied whole embryos with their fertilization membranes or five pieces of individual tissues from dissected embryos were transferred into a 1.5 ml microfuge tube. A clean glass Pasteur pipette was used to remove excess buffer solution. Microfuge

pestles were cleaned with RNAse Zap solution (Ambion, Cat. # 9780. 9782) and rinsed two to three times with Diethylpyrocarbonate (DEPC, Sigma) treated deionized water (dIH_2O). DEPC-dIH_2O was made by addition of 1 ml of DEPC to 1000 ml of dIH_2O, which was autoclaved after overnight incubation at room temperature. 400 µl of ice-cold Trizol reagent (Invitrogen) was added to each tube containing whole embryos and 200 µl to the tubes containing dissected tissues. Embryos or tissues were homogenized using RNAse free pestles attached to an electrical homogenizer. All the samples were homogenized vigorously for two to three minutes followed by addition of 400 or 200 µl of Trizol to each tube, depending on the sample type, and mixed thoroughly. After five minutes incubation at room temperature, 400 or 200 µl chloroform was added to respective tubes. Tubes were rigorously shaken for 15 to 20 seconds. Mixtures were incubated for three minutes at room temperature and centrifuged at 16,100 xg for 10 minutes. As a result of centrifugation, all RNAs were present in an aqueous phase at the top of each tube, and the aqueous phase was carefully transferred to new 1.5 ml microfuge tubes without disturbing the interphase. Depending on the sample type, 500 or 250 µl of isopropanol was added to each tube and swirled gently. Samples were incubated at room temperature for 10 minutes. Tubes were centrifuged for 10 minutes at 16,100 xg. All supernatants were carefully discarded, keeping the pellets at the bottom of each tube. One ml of 70% ethanol was added to each pellet and vortexed for 30 seconds for proper washing. Tubes were spun at 16,100 xg for 10 minutes. Supernatants were discarded and pellets were air dried by laying the tubes with the cap open on the tabletop for 20 – 25 minutes at room temperature. Dried pellets were dissolved in 25 µl of DEPC-

dIH_2O. 25 µl of 8M LiCl was added to each tube along with 1 µl of glycogen, and the tubes were stored at -20^oC overnight.

The next day, the tubes were centrifuged for 15 minutes at 16,100 xg at 4^oC. Supernatants were discarded, and pellets were washed with 500 µl of 70% ethanol by gently vortexing. Tubes were centrifuged again for 10 minutes 16,100 xg at 4^oC. After discarding the supernatants from each tube, pellets were air dried and dissolved in 30 µl of DEPC-dIH_2O by incubating tubes at 60^oC for 5 minutes. This was followed by gentle vortexing of the tubes and brief spinning. DNAse treatment was performed by adding 7.5 µl of DEPC-dIH_2O, 10.0 µl of DNAse reaction buffer (Promega), 2.0 µl of RNAse free DNAse (Promega) and 0.5 µl of RNAse inhibitor (Takara) to each tube, mixing gently, and incubating the tubes at 37^oC for 45 minutes. After incubation, tubes were centrifuged briefly, and 150 µl of DEPC-dIH_2O + 200 µl of PCI (Phenol: Chloroform: Isoamyl alcohol:: 25: 24: 1) was added to each tube and vortexed rigorously for 30 seconds. Tubes were centrifuged for 5 minutes at 16,100 rcf at room temperature. The upper aqueous layers (~180 µl) were transferred very carefully to new 1.5 microfuge tubes without disturbing the interface of the two layers, which consists of digested genomic DNA and proteins. 18 µl of 3M Sodium Acetate pH 6.0, and 500 µl of ice-cold ethanol were added to each tube and stored at – 20^0C overnight.

The following day, the tubes were centrifuged for 15 minutes at 15,700 xg at room temperature. Supernatants were discarded and pellets were washed with 70% ethanol and air dried as before. Finally pellets were dissolved in 25 µl of DEPC-dIH_2O. One µl of RNA was mixed with 99 µl of DEPC-dIH_2O and quantified at 260 nm using a UV spectrophotometer (Beckman DU 530).

VI. Reverse transcription to produce cDNA:

A total of 2-5 µg RNA was used as a template to synthesize the first strand of cDNA by reverse transcription. A 50 µl reaction was set up in a 1.5 ml microfuge tube on ice by adding the following reagents in order: 10 µl 5X RT Buffer (Promega), 2.5 µl 100 mM dNTPs (USB), 5 µl of 500 nM/µl 12-18 mer Oligo (dT) (Promega), 2-5 µg of RNA in a volume of 10 µl, and 2 µl RNAse Inhibitor (NEB). The final volume of the mix was adjusted to 47.5 µl by adding DEPC-dIH_2O. Reactions were initiated by adding 2.5 µl of M-MLV-RT enzyme (500 units, Promega) to each reaction. Reagents were mixed well by pipetting up and down and finally by brief spinning in a tabletop micro-centrifuge. Tubes were incubated in a 37^0C water bath for a minimum of 2 hours. Each reaction was accompanied with a negative RT reaction, which contained all the mentioned reagents except for M-MLV-RT enzyme. Once, the incubation was over, cDNA samples were purified by running the reaction mixtures through the spin-50 mini-columns (USA Scientific). Retrieved purified cDNA samples were stored at – 20^0C until ready to use.

VII. Cloning the *EcSmad2* gene:

The complete ORF of *EcSmad2*, most of the 5'UTR and the complete 3'UTR ending with Poly(A) tail were cloned in four steps:

A. Degenerate PCR cloning of the *EcSmad2* ORF:

Three sets of forward and reverse degenerate primers were designed based on Clustal analysis (Clustal w1.81) of *Smad2* ORF sequences of *X. laevis*, *X. tropicalis*, *Mus musculus* and *Homo sapiens* (Fig. 7).

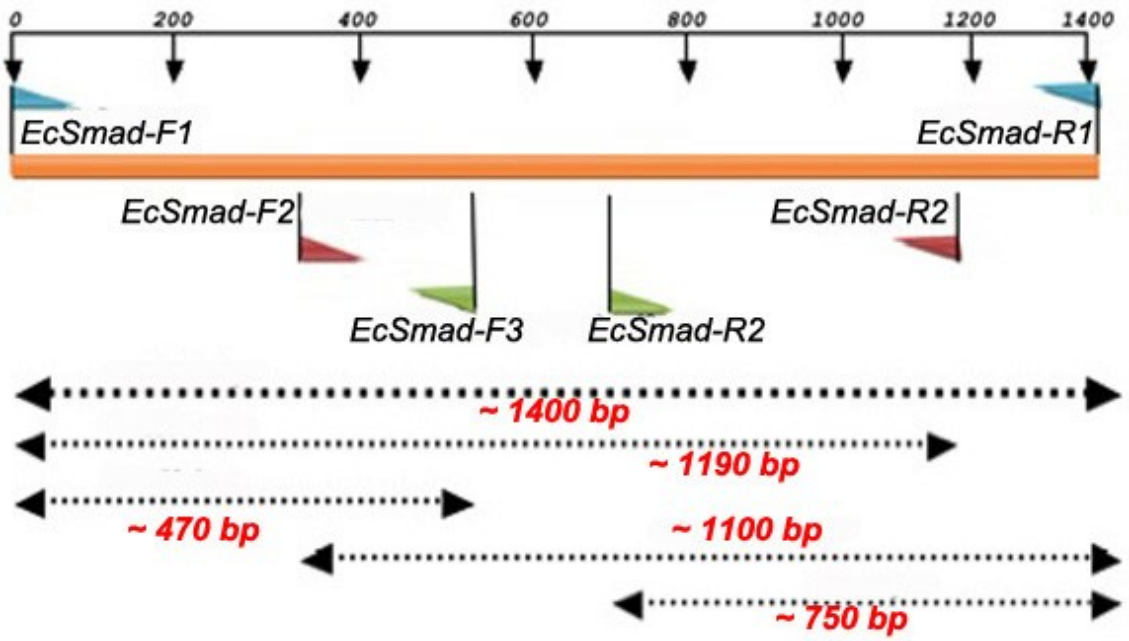
Fig. 7: Degenerate primers for cloning *EcSmad2*. (A) Three forward and three reverse degenerate primers. Red letters represent the degenerate code. (B) Positions of the three pairs of degenerate PCR primers on the hypothetical EcSmad2 ORF, based on sequence comparison with other species. The expected product sizes are: Reaction # 1: F1-R1 = ~1400 bp, Reaction # 2: F1-R2 = ~1190 bp, Reaction # 3: F1-R3 = ~470 bp, Reaction # 4: F2-R1 = ~1100 bp, Reaction # 5: F3-R1 = ~750 bp,

Fig. 7 (A):

Primer Name	Sequence	Length (bp)	Tm (50 mM NaCl)
EcSmad2 -F1	5′- ATGTCSTCSATYTTGCCATTYACBCC -3′	26	60.6^{0}C
EcSmad2 -F2	5′- ACAGGCCTTTACAGCTTCTCTGAACA -3′	26	60.0^{0}C
EcSmad2 -F3	5′- CCTGGATATATYAGTGAAGATGGAGA -3′	26	54.4^{0}C
EcSmad2 -R1	5′- TTABGACATGCTTGAGCADCGSACTG -3′	26	61.2^{0}C
EcSmad2 -R2	5′- GACTGAGCSAGRAGAGCAGCAAAYTC -3′	26	61.3^{0}C
EcSmad2 -R3	5′- ACTTCATCYTTTTTMAGRTTAAAAGC -3′	26	51.7^{0}C

[Degenerate Code: B = C, G, or T, D = A, G or T, M = A or C, R = A or G, S = C or G, Y = C or T; F = forward and R = reverse]

Fig. 7 (B):



Gradient PCR reactions were set up for each of the above combination of primers in a TGradient Thermocycler (Whatman, Biometra). Each 20 µl reaction mix consisted of 2 µl 10X PCR Reaction Buffer (Invitrogen, Carlsbad, CA), 0.6 µl of 50 mM $MgCl_2$ (Invitrogen, Carlsbad, CA), 0.4 µl of 10mM dNTPs (Fisher), 0.1 µl of Taq DNA Polymerase (Invitrogen, Carlsbad, CA), 1.0 µl of template cDNA, 1.0 µl of each of the forward and reverse primers, and 13.9 µl of DEPC-dIH_2O. The cDNA from *E. coqui* embryos of neurula stage (TS3) was used as template in each reaction. A doublet of bands was obtained for reactions 4 and 5 at expected sizes using the following program:

94^0C – 3 minutes

94^0C – 45 seconds

52^0C – 45 seconds } 29 Cycles

72^0C – 1 minute 30 seconds

72^0C – 10 minutes

4^0C – Over night hold

The amplified PCR products from reactions 4 (F2-R1) and 5 (F3-R1) were run at 100 V for 45 minutes on a 1.5% Agarose gel with 2 µl of 5 mg/ml Ethidium Bromide (EtBr). A Kodak imaging system was used to illuminate the gel with UV light to take a photograph followed by the excision of the desired sized bands using a sharp razor blade. DNA fragments from the excised gel pieces were extracted using GeneJET^{TM}Gel Extraction Kit (Fermentas, Cat # K0691). DNA was eluted from the column in 50 µl DEPC-dIH_2O, and the concentration of DNA was measured using a UV spectrophotometer (Beckman DU 530).

The F2-R1 and F3-R1 DNA fragments were ligated into a pGEM vector using Promega Ligation pGEM-Teasy Vector System (Cat # A3600). The ligated vectors were transformed into JM 109 High Efficiency Competent cells following Promega's transformation protocol. Finally, 100 µl from each transformation reaction culture was plated on LB/Ampicillin/IPTG/X-Gal plates. Twenty colonies from each plate were screened for the right inserts by PCR using nested primers. Two such positive colonies from each type were selected and grown overnight in 500 µl LB medium. The following day, 50 µl of those overnight cultures were used to inoculate 5 ml LB medium and incubated overnight in a 37^0C shaker. The following day, plasmid DNA was extracted using GeneJET^{TM}Plasmid Miniprep Kit (Fermentas, Cat # K0502). A PCR reaction with the same set of nested primers was done again to verify the presence of the right clone in those DNA samples. The clones were sequenced at the DNA Sequencing Facility, University of Pittsburgh, Pittsburgh, PA using ABI 3100 Sequencer (Applied Biosystems Inc., Foster City, CA).

B. Cloning the *EcSmad2* 3'UTR from an *E. coqui* Ovarian cDNA Library:

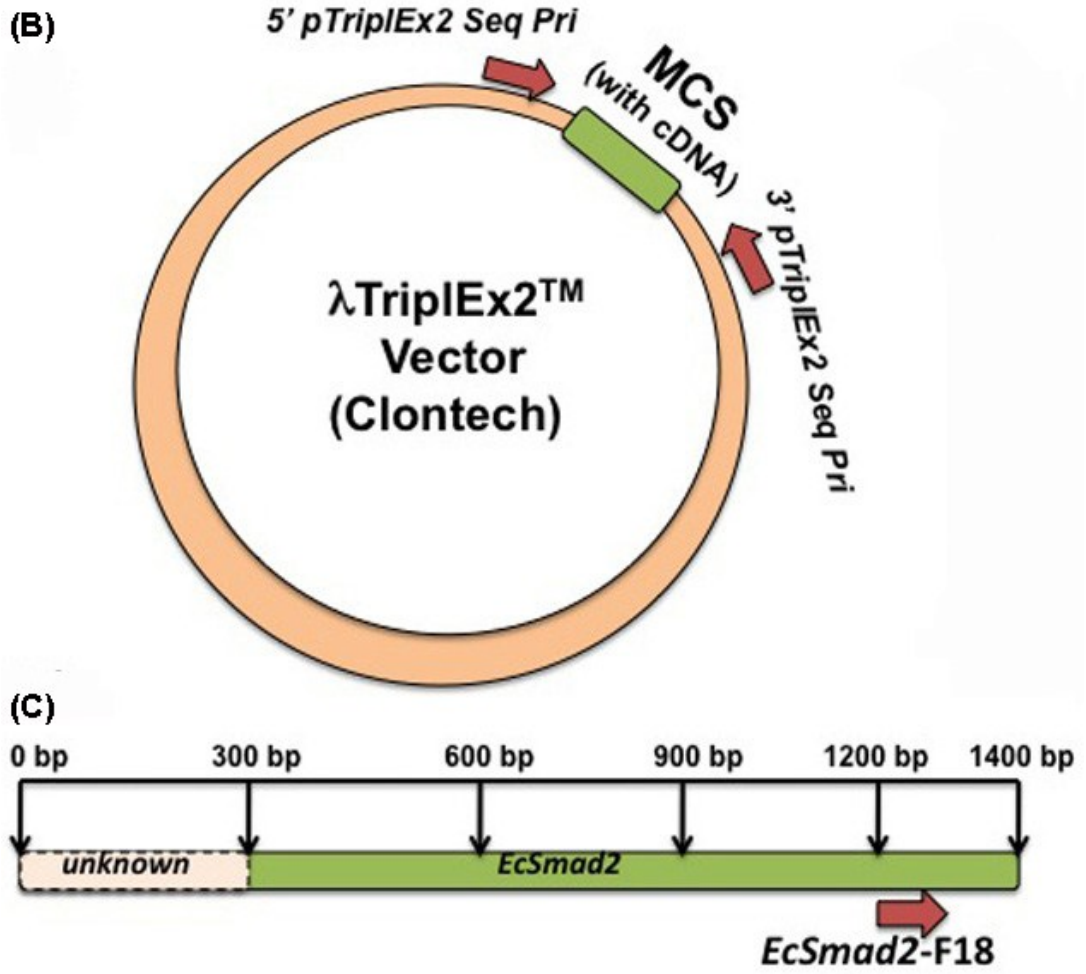
In order to obtain the 3'UTR region of *EcSmad2*, I used PCR on Custom SMART^{TM} cDNA Library in λTriplEx2^{TM} (Clontech, Cat # CS1023u). The RNA source for this library was *E. coqui* ovary. Random fragments of *E. coqui* ovarian cDNAs were used as inserts. Based on the previous sequencing results, an exact forward primer for the *EcSmad2* ORF was designed towards the 3' end (Fig. 8). As I did not know the orientation of the insert, I designed two primers for the vector sequence – forward

Fig. 8: Primers for cloning the 3'UTR. (A) Exact *EcSmad2* and λTriplEx2 vector primers. (B) Diagram showing positions and directions of 5' pTriplEx2 Seq Pri and 3' pTriplEx2 Seq Pri on the λTriplEx2 vector. MCS (green bar) stands for the multiple cloning site, which contains complete or incomplete fragments of *E. coqui* genes. (C) Diagram showing position and direction of *EcSmad2*-F18 forward primer.

Fig. 8:

(A)

Primer Name	Sequence	Length (bp)	Tm (50 mM NaCl)
EcSmad2-F18	5′-GGGCTTTGAAGCAGTTTACCAGTTAACG-3′	28	60.0^{0}C
5' pTriplEx2 Seq Pri	5′-GCCAAGCTCCGAGATCTGGACGAGC-3′	25	65.0^{0}C
3' pTriplEx2 Seq Pri	5′-GAATTGTAATACGACTCACTATAGGGCGAA-3′	30	57.8^{0}C



(5’ pTriplEx2 Seq Pri) and reverse (3’ pTriplEx2 Seq Pri) flanking the Multiple Cloning Site (MCS).

PCR reactions were set up using *EcSmad2*-F18 with 5’ pTriplEx2 Seq Pri and *EcSmad2*-F18 with 3’ pTriplEx2 Seq Pri in 50 µl total volume as follow: 6.5 µl 10X PCR Reaction Buffer A (Fisher), 1.0 µl of 10mM dNTPs (Fisher), 0.5 µl of Taq DNA Polymerase (Fisher), 5.0 µl of ovarian library cDNA as template, 4.0 µl of each of the forward and reverse primers, and 29.0 µl of DEPC-dIH_2O. Reaction with *EcSmad2*-F18 and 3’ pTriplEx2 Seq Pri produced a band using the following program:

94^0C – 3 minutes

94^0C – 45 seconds

56^0C – 45 seconds } 29 Cycles

72^0C – 1 minute 30 seconds

72^0C – 10 minutes

4^0C – Over night hold

The amplified PCR product obtained only from reaction with *EcSmad2*-F18 and 3’ pTriplEx2 Seq Pri, was gel extracted, ligated into a cloning vector, transformed, screened and sequenced as described in section VII.A.

C. Cloning the 5’ UTR and initial segment of *EcSmad2* ORF using 5’RACE:

The initial 5’ end of the ORF and the whole 5’ UTR was cloned using the 5’RACE PCR Kit (Invitrogen, Carlsbad, CA; Cat # 18374-058). Invitrogen’s protocol for 5’ RACE system for Rapid Amplification of cDNA Ends, Version 2.0 was followed with some minor modifications. According to the protocol, gene specific primers (GSP) were

designed for *EcSmad2* – GSP1, GSP2 and GSP3. These were reverse primers and running towards the beginning of the ORF (Fig 9).

A PCR of dC-tailed cDNA was done using the following recipe: 5.0 µl 10X PCR Buffer, 3.0 µl of 25mM $MgCl_2$, 2.0 µl of 10mM dNTPs mix, 3.0 µl of nested GSP2 primer (10 µM solution), 3.0 µl of Abridged Anchor primer (10 µM solution), 10.0 µl of dC-tailed cDNA as template, 23.0 µl of DEPC-dIH_2O and 0.5 µl of Taq DNA Polymerase (5 units/µl). Product was obtained using the following program:

94^0C – 5 minutes

94^0C – 1 minute

55^0C – 1 minute 30 seconds } 29 Cycles

72^0C – 2 minutes

72^0C – 10 minutes

4^0C – Over night hold

The product was further verified to contain *EcSmad2* sequence by doing PCR with nested GSP3 reverse and GSP4 forward primers. The 5'RACE product was gel extracted, ligated into cloning vector, transformed, screened and sequenced as described in section VII.A. Finally, when the full ORF sequence was known, one forward and one reverse exact primer were designed (Table 1) to amplify the full ORF sequence using PCR with high fidelity Taq Polymerase, Pfu Ultra fusion Hs DNA polymerase (Stratagene, Cat # 600670) as follow: 5.0 µl 10X Pfu Ultra Rxn Buffer, 2.0 µl of 10mM dNTPs mix, 3.0 µl of EcSmad2 Full ORF FP (10 µM solution), 3.0 µl of EcSmad2 Full ORF RP (10 µM solution), 4.0 µl of TS3 whole embryo cDNA as template, 32.0 µl of DEPC-dIH_2O and 1.0 µl of Pfu Ultra fusion Hs DNA polymerase.

Fig. 9: Primers for 5'RACE to clone 5' UTR and the initial segment of the ORF. (A) Primers used in 5'RACE. (B) Positions and directions of primers. GSP1 and GSP2 are designed based on the known EcSmad2 sequence (green bar). GSP3 and GSP4 are designed based on the preliminary sequencing data on the 5'RACE product.

Fig. 9:

(A)

Primer Name	Sequence	Length (bp)	Tm (50 mM NaCl)
***EcSmad2*-RP16 GSP1**	5′-GCATATTCACAATTTTCAAT -3′	**20**	**44.5^{0}C**
***EcSmad2*-RP17 GSP2**	5′-CCAGAGACGACAGTAGATAACGTGCGGTAA-3′	**30**	**62.2^{0}C**
***EcSmad2*-RP18 GSP3**	5′-GGCGGCCATCTAGAGACCTGGTTTGTTCA-3′	**29**	**65.0^{0}C**
***EcSmad2*-FP19 GSP4**	5′-GACAGGCCAGCTGGACGAACTCG-3′	**23**	**64.0^{0}C**

(B)

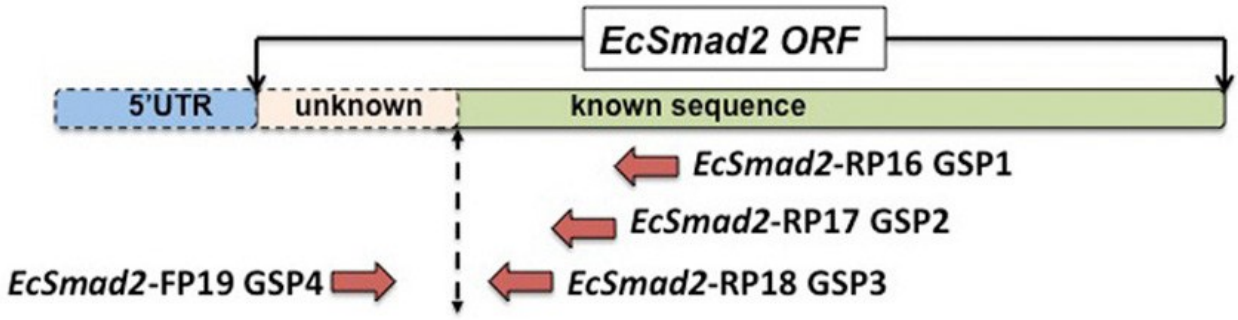


Table 1: *EcSmad2* Full ORF amplification primers. *EcSmad2* full ORF forward and reverse primers were used to amplify the full ORF for both the isoforms of the gene.

Table 2: qPCR primers. One pair of EcSmad2 specific primers, *EcSmad2*-qPCR FP1 and *EcSmad2*-qPCR RP1 was used for all the qPCR experiments. *EcL8* F4 and *EcL8* R5 primers were used as endogenous controls.

Table 1:

Primer Name	Sequence	Length (bp)	Tm (50 mM NaCl)
EcSmad2 **-Full ORF FP**	5′-**ATGTCCTCCATACTGCCCTTTACACC** -3′	**26**	**60.2^{0}C**
EcSmad2 **-Full ORF RP**	5′-**TTAGGACATGCTTGAGCAGCGGA**-3′	**23**	**60.7^{0}C**

Table 2:

Primer Name	Sequence	Length (bp)	Tm (50 mM NaCl)
EcSmad2 **-qPCR FP1**	5′-**CCCTTTACACCTCCCGTTGTGAAACGT** -3′	**27**	**62.5^{0}C**
EcSmad2 **-qPCR RP1**	5′-**CCTTCTCGAGTTCGTCCAGCTGGC**-3′	**24**	**63.4^{0}C**
***EcL8*-F4**	5′-**GAAGGTCATCTCTTCTGCAAACAGAGC** -3′	**27**	**61.0^{0}C**
***EcL8*-R5**	5′-**TAAGACCAACTTTGCGACCAGCTGG**-3′	**25**	**63.0^{0}C**

Following program was used:

95^0C – 3 minutes

95^0C – 30 seconds

55^0C – 45 seconds } 34 Cycles

72^0C – 1 minute 15 seconds

72^0C – 5 minutes

4^0C – Over night hold

Products were gel extracted, ligated into cloning vector, transformed and screened as previously described. Two clones for each of the isoforms of *EcSmad2* (Full ORF and Δexon3) were stored at – 80^0C as freezer stocks.

VIII. Quantitative PCR (qPCR) Analysis:

The detection of temporal and spatial regulation of *EcSmad2* expression was done using ΔΔCt method of qPCR. I designed *EcSmad2* ORF specific forward and reverse primers (Table 2), which would produce a small band of around 175 bp, ideal for qPCR analysis. Specific primers were used for the ribosomal protein coding gene *EcL8*, which served as the endogenous control.

For temporal expression, cDNAs for qPCR were synthesized using 5 μg of RNAs extracted from whole *E. coqui* oocytes and embryos at various stages of development – NF8, NF10, TS3, TS4, TS5, TS6, TS7, and TS8. Both *EcL8* and *EcSmad2* primer efficiencies were first tested as described by Singamsetty (2009). In order to ensure good quality of cDNAs, a regular reverse transcriptase PCR (RT-PCR) was performed before every qPCR run.

To do a relative quantification, reactions were set up in MicroAmp fast optical 48 or 96 well plates (Applied Biosystems, Foster City, CA). The total volume for each reaction was 20 µl. First, a master mix for each gene was made in a 1.5 ml sterile microfuge tube. Each reaction mix contained 1.5 µl of cDNA template (150 – 200 ng), 0.6 µl of each of the forward and reverse primers, 7.3 µl of DEPC-dIH_2O and 10.0 µl of 2X Maxima^{TM} SYBR Green/ROX qPCR Master Mix (Fermentas, cat # K0221). Care was taken at every step of assembling the reactions and pipetting mixes into the wells. Every reaction was represented in triplicates on the reaction plates. Plates were then sealed with MicroAmp 48 or 96-well optical adhesive films (Applied Biosystems, Foster City, CA).

The reactions were run on StepOne Real Time PCR System (Applied Biosystems, Foster City, CA). A ΔΔCt method of qPCR program was set to run, and a melt curve was obtained for amplification phase each time. *EcSmad2* was amplified for 40 cycles with the following parameters:

Initial denaturation: @ 94^0C for 10 minutes

Denaturing: @ 94^0C for 50 seconds

Annealing: @ 56^0C for 50 seconds

Extension: @ 72^0C for 1 minute 15 seconds

The experiments were analyzed using StepOne V2.0. The data was exported as an Excel spread sheet and processed in Microsoft Excel 7.0. Final RQ values for each set were then plotted to show the relative *EcSmad2* expression.

For spatial expression, cDNAs for qPCR were synthesized using 5 µg of RNAs extracted from dissected DE and NE tissues of embryos at various stages of development. I followed same protocol as described for temporal expression.

IX. In vitro transcription of capped mRNA and microinjection into embryos:

Ambion mMESSAGE mMACHINE^{TM} Kit FOR SP6 (Ambion, Cat # 1340) was used to synthesize capped mRNA of *X. laevis* EGFP tagged *Smad2* cDNA (XeSmad2-EGFP). The construct was a gift from Dr. Chenbei Chang. The plasmid vector was CS105 pDH105 and the linearizing enzyme was *Asc*I. The Ambion instruction manual was followed with a minor modification. The transcription reaction was carried out at 37^0C for three hours instead of one. Transcribed capped RNAs were treated with DNaseI and recovered using phenol: chloroform extraction and isopropanol precipitation as per the instructions provided with the kit. All RNA samples were quantified by UV absorbance at 260 nm and stored at – 20^0C until needed.

X. laevis and *E. coqui* embryos were cultured in 5% Ficoll (Fisher, Cat # BP525-100) dissolved in 20% Steinberg's solution throughout the microinjection procedure. For practice purposes, *X. laevis* embryos to be injected were placed in a drop of the medium on a RNAse free silane-prep glass slide (Sigma). In the case of the bigger *E. coqui* embryos, they were placed on a square piece of Nitex Nylon screen cloth no. 2380. The square screen cloth with the embryos on it was kept immersed in the microinjection medium in a 60x15 mm plastic petri dish (VWR). All the petri dishes and the screen cloths were cleaned with RNAse away (Molecular Bioproducts) prior to use. A needle puller (KOPF Needle/ Pipette puller, Model 730) was used to make glass needles and the tips were trimmed with forceps as needed. Small drops of RNA to be injected were placed on a piece of parafilm and drawn through suction into the glass needle filled with light mineral oil (Fisher, Cat # 0121-1) with the help of Drummond Nanoject variable automatic injector.

While observing under a dissecting scope, a hair loop and forceps were used to position each *E. coqui* embryo at the 4-8 cell stage with its animal pole up. RNAs were injected as a single shot of 4.6 nl, 9.2 nl or 13.8 nl at an angle to the undivided vegetal half of the embryo.

For each microinjection experiments, at least six embryos were kept as un-injected controls. All other sets of embryos were injected with either 9.2 nl of DEPC-H_2O, or *EGFP-Smad2* mRNA. In case of the *EGFP-Smad2*, three separate concentrations were used: 10 pg, 100 pg and 900 pg.

Once injected, the embryos were transferred to a fresh petri dish containing 5% Ficoll in 20% Steinberg's solution and kept in that medium at least until the next day. Embryos were transferred to 20% Steinberg's solution, and were cultured until they reached NF10, when they were dissected into the prospective DE and NE tissues. Tissues were dissociated into individual cells in the dissociation medium as described in section IV.

X. Whole protein isolation, SDS-PAGE and Western Blot Analysis:

The whole protein preparation from yolky *E. coqui* embryos was a difficult task due to the huge amounts of lipids and yolk platelets in them and the lack of existing protocols. In order to establish a standard protocol to isolate protein from *E. coqui* embryos, several recipes for homogenization buffers (HB) were tried. The standard protein isolation protocol (HB1), used for *X. laevis* embryos (Birsoy et al., 2005), did not work well on *E. coqui* embryos. I altered the recipe 15 times, leading to a series of buffer recipes HB1 to HB16. HB16, which was modified after Callery et al. (1996), eliminated

interference by yolk platelets. Proteins extracted using HB16 produced very nice bands when run on SDS-PAGE. In Table 3, I have listed a few of the HB recipes in order to have a better understanding of the modifications of the original buffer recipe.

Eight to ten *E. coqui* whole embryos were homogenized in 150 µl of HB16 for 30 seconds on ice using a Kontes Pellet Pestle Cordless Motor (Kimble Chase Kontes, Cat # K749540-0000). Blue polypropylene, autoclavable Kontes disposable pellet pestles (Sigma-Aldrich, Cat # Z359947-100EA) were used for grinding the tissues in autoclaved 1.5 ml microfuge tubes until the solution turned milky white. In the case of MZ or VC from early stages and DE or NE from post-gastrulation stages, 15-16 pieces of dissected tissues were homogenized in the same volume of HB16 for the same time span.

Extracts were centrifuged at 16,100 xg for 15 minutes on a tabletop microcentrifuge at 4^0C. After spinning, the yolky, white masses settled to the bottom of the tubes. A clear supernatant appeared on top with a very thin, white, almost elastic layer on the surface. The transparent supernatants were carefully separated to new 1.5 ml microfuge tubes. If any contamination with the white mass was observed in any of the new tubes, they were subjected to another 5 minutes spin at the same speed in order to get rid of contaminants. In each extraction, almost 120 µl of transparent supernatant were obtained. Twenty µl of each supernatant was then used with BCA Protein Assay Reagent kit (Thermo Scientific, Cat # 23225) following the manufacturer's protocol to measure the protein concentration. Meanwhile, equal volume of freshly made 2X Laemmli sample buffer (BioRad) was mixed thoroughly with the transparent protein supernatants and boiled for 5-6 minutes [1 ml 2X Laemmli sample buffer = 950 µl Laemmli Sample buffer (BioRad, Cat # 161-0737) + 50 µl of 2-Mercaptoethanol (BioRad, Cat # 161-0710)].

Finally, tubes were spun briefly at room temperature (RT) and stored either on ice or at – 20^0C until further use. Once the concentrations for each of the protein preparations in an experiment were determined, they were equalized to the lowest one by adding 1X Laemmli sample buffer.

Depending on the protein concentration in an experiment, 75 – 100 μg protein were loaded per lane of a 10% Tris-HCl BioRad Ready Gel (10X Cat # 161-1155). PageRuler™ Prestained Proteins ladder (Thermo Scientific, prod # 26616) was used in all the gels. All protein gels were run using BioRad Mini PROTEAN® 3 Gel system following the manufacturer's manual (BioRad, Cat # 165-3301). A 10X stock solution for the Gel Running Buffer was made as follows: 30 g Tris (Base), 144 g Glycine and 10 g Sodium Dodecyl Sulphate (SDS) in 1000 ml of dIH_2O. A 1X working solution was freshly made for each run. A maximum of two gels were run at a time at 100V for 90 minutes. In all experiments, one gel was always run with all samples for Coomassie Brilliant Blue (CBB) staining to check the quality of the protein preparations. The CBB working solution was 0.02% CBB, 50% methanol, and 10% glacial acetic acid in dIH_2O. Staining the gels with freshly made solution was usually enough to visualize the band patterns on the gels within an hour. Gels were de-stained by incubating them overnight in a 5% methanol and 10% acetic acid destain solution.

Table. 3: Recipes for four homogenization buffers to extract protein. HB1 was based on the recipe provided by Birsoy et al. (2005) for extraction of proteins from *Xenopus* embryos. HB14 and HB15 were modifications of HB1, while HB16 was modified from the recipe described by Callery et al. (1996).

Table. 3:

HB1 (Birsoy et al., 2005)	
Stock solution	
1 M Tris (pH 8.0)	100 µl
100 mM EDTA	20 µl
100 mM EGTA	50 µl
25% NP-40 (Sigma)	20 µl
250 mM Na-β glycerophosphate	100 µl
500 mM NaF (Sigma, Cat # 919-25ML)	200 µl
100 mM Sodium Pyrophosphate	100 µl
10 nM CalyculinA (Sigma, Cat # C5552-10UG)	2 µl
PIC (Sigma, Cat # P-8340)	10 µl
PMSF (Sigma, Cat # 93482-50ML-F)	5 µl
dIH_2O	393 µl
	1000 µl
HB11	
Stock solution	
100 mM Tris (pH 7.4)	250 µl
100 mM EDTA	20 µl
100 mM $MgCl_2$	100 µl
10 nM CalyculinA (Sigma, Cat # C5552-10UG)	2 µl
PIC (Sigma, Cat # P-8340)	10 µl
PMSF (Sigma, Cat # 93482-50ML-F)	5 µl
dIH_2O	613 µl
	1000 µl

HB14	
Stock solution	
100 mM Tris (pH 7.4)	250 µl
100 mM EDTA	20 µl
100 mM EGTA	50 µl
250 mM Na-β glycerophosphate	100 µl
100 mM Sodium Pyrophosphate	100 µl
10 nM CalyculinA (Sigma, Cat # C5552-10UG)	2 µl
PIC (Sigma, Cat # P-8340)	10 µl
PMSF (Sigma, Cat # 93482-50ML-F)	5 µl
dIH_2O	463 µl
	1000 µl
HB16 (Modified after Callery et al., 1996)	
Stock solution	
100 mM Tris (pH 7.4)	300 µl
100 mM EDTA	40 µl
100 mM $MgCl_2$	100 µl
10 nM CalyculinA (Sigma, Cat # C5552-10UG)	10 µl
PIC (Sigma, Cat # P-8340)	20 µl
PMSF (Sigma, Cat # 93482-50ML-F)	20 µl
23.8 mM NaF (Sigma, Cat # 919-25ML)	210 µl
dIH_2O	300 µl
	1000 µl

Gels for western blotting were incubated in dIH_2O on a shaker for 5 minutes followed by a 30 minute equilibration in transfer buffer, along with other components of the transfer cassette sandwich (2X Fiber pads, 4X filter papers, 1X PVDF membrane). Immun-Blot PVDF membrane (BioRad, Cat #162-0177) was used for transfers and each time before equilibration in the transfer buffer, they were soaked in methanol for 30 seconds. The transfer sandwich was assembled in a large tray submerged under the cold transfer buffer following the BioRad instruction manual. Transfer buffer was: 3.0 g Tris (Base), 14.4 g Glycine and 200 ml Methanol in 1000 ml total volume. The transfer buffer was made fresh each time and stored in the cold room. Proteins were transferred at 150 V for 80 – 90 minutes in a tank on a magnetic stirrer with a stir bar inside the tank to circulate the heat well.

Once the transfer was finished, the transfer cassette was dislodged from the tank and disassembled in a tray under dIH_2O. The PVDF membranes were incubated in Ponceau S stain (0.1% (w/v) Ponceau S in 1% (v/v) acetic acid) for 30 seconds to check for the presence of protein bands on the membrane. After staining and visualization of bands, the membrane was transferred into water for two washes of 5 minutes each. Finally the membranes were removed and incubated in TBST (10.0 mM Tris pH 8, 0.15 M NaCl, 1% BSA, 0.05% Tween-20) for 20 minutes on a shaker. Membranes were blocked in 5% Non Fat Dry Milk (NFDM) in TBST for an hour and half or overnight. After blocking, membranes were incubated in primary antibody dilutions (Table 4) made in blocking medium, overnight at 4^0C on a shaker. The next morning, each membrane was washed in TBST 3x15 minutes. Only in case of anti-PSmad2 antibody, the membrane was incubated after washing in Peroxidase Suppressor (Thermo Scientific,

Prod # 35000) for 30 minutes to reduce background staining. This was followed by another 3X15 minutes wash in TBST. All membranes were incubated in their respective HRP-tagged secondary antibody, made in 2% NFDM in TBST for two hours at RT. Membranes were washed for 3X15 minutes in TBST. Finally each membrane was treated with SuperSignal® West Femto Maximum Sensitivity Substrate (Thermo Scientific, Prod # 34096). Membranes were incubated in the mixture for 5 minutes and wrapped in a plastic sheet protector, and the remaining solution was squeezed out.

A Typhoon 8600 Variable Mode Imager (Molecular Dynamics, Sunnyvale, CA) was used for imaging the western blot. The membrane was placed with its face down on the glass platen towards a corner to minimize the scanning time. Scan area was selected using Scanner Control software (Amersham Biosciences), and resolution was selected at high with mode of scanning selected at Chemiluminiscence. The rest of the parameters were left in default setting. Developed images were analyzed using Image Quant Version 5.0 (Molecular Dynamics) or Adobe Photoshop CS5 Extended software.

XI. Immunocytochemistry (ICC) and DAPI staining with dissociated cells:

Immunocytochemistry was performed on dissociated cells to detect the cellular location of selected proteins. After cells were completely dissociated, they were washed gently in 1XPBS on a horizontal shaker for 3x10 minutes at a low speed. Cells were fixed in 3.7% formaldehyde solution (J.T.Baker, Cat # 2106-11; Stock solution 37%, diluted in 1X PBS) for 35 minutes with gently shaking. After that, cells were washed in 1X PBS for 3x10 minutes followed by permeabilization. Cells were permeabilized in 0.3% TritonX-100 in 1X PBS (ICN Biomedicals, Inc., Cat # 07426) for 35 minutes with constant gentle

shaking. Cells were washed in 1X PBS for 3x10 minutes. Cells were then blocked in 10% Goat Serum (GIBCO, Cat # 16210-064) in 1X PBS for 90 minutes on a shaker.

Different primary antibodies were diluted in 10% Goat Serum in 1X PBS according to the desired concentration (Table 4). Cells were incubated in a dilution of specific primary antibody overnight at 4^0C with gentle shaking. The next morning, cells were washed in 1% Goat Serum in 1X PBS for 3x15 minutes, followed by incubation in the appropriate dilution of a specific secondary antibody (Table 4) for at least an hour at room temperature on a shaker. As all the secondary antibodies were tagged with a fluorescent label, the incubation containers, flat bottom12 Well Cell Culture plates (COSTAR, Coring, NY; Cat # 3513) were wrapped with aluminum foil to protect them from light. All subsequent steps were performed in aluminum wrapping.

Cells were washed again in 1% Goat Serum in 1X PBS for 3x15 minutes followed by DAPI staining. DAPI stock solution was made by dissolving 1 mg DAPI (Invitrogen^{TM}/ Molecular Probes, Cat # D1306) in 1 ml DEPC-dIH_2O and stored in aliquots in 1.5 ml microfuge tubes at – 20^0C. Cells were incubated in 1:500 to 1:1000 dilution of the DAPI stock solution, in 1% Goat Serum in 1X PBS for 20 to 35 minutes on a shaker, followed by washing in 1% Goat Serum in 1X PBS for 3x15 minutes. Finally, cells were treated with SlowFade® Antifade kit (Invitrogen^{TM}/ Molecular Probes, Cat # S2828). When ready, cells were mounted in antifade medium on glass slides and observed with a Nikon Microphot-SA microscope (Nikon Corporation, Tokyo, Japan). Photographs were taken after exciting various fields with bright or fluorescent light through FITC, Rhodamine or UV filters. A QED camera was used to capture images, and they were processed using QCapture software and Adobe Photoshop.

Table. 4: Primary and secondary antibodies used in immunocytochemistry (IC) and western blotting experiments.

Name of the Antibody	Primary or Secondary	Company and Cat #	Dilution used for IC	Dilution used for Western
1. Purified Mouse monoclonal Anti-Smad2/3*	Primary	BD Transduction Laboratories; 610842	1:50	1:300
2. Rabbit Polyclonal Anti-Phospho-Smad2 (Ser465/467) Antibody**	Primary	Cell Signaling; 3101	1:25	1:150
3. Rabbit Anti-Smad4 Polyclonal Antibody	Primary	Thermo Scientific; PA1-41292	1:50	1:150
4. Goat Anti-human Activin RIIA Antibody	Primary	R&D Systems, Inc.; AF340	1:30	1:150
5. RNA Polymerase II 8WG16 Mouse Monoclonal Antibody	Primary	Covance; MMS-126R	1:50	N/A
6. RNA Polymerase II H14 Mouse Monoclonal Antibody	Primary	Covance; MMS-134R	1:50	N/A
7. RNA Polymerase II H5 Mouse Monoclonal Antibody	Primary	Covance; MMS-129R	1:50	N/A

Name of the Antibody	Primary or Secondary	Company and Cat #	Dilution used for IC	Dilution used for Western
8. Mouse Monoclonal Alpha Tubulin Antibody (DM1A) – loading control	Primary	abcam; ab7291	N/A	1:40,000
Anti-Green fluorescent protein, Rabbit IgG fraction.	Primary	Invitrogen™/ Molecular Probes; A11122	1:200	N/A
9. Anti-GFP, rabbit polyclonal antibody, Alexa fluor® 488 conjugate		Invitrogen™/ Molecular Probes; A21311	1:250	N/A
10. Alexa fluor® 488 Goat Anti-Mouse IgG (H+L)	Secondary	Invitrogen™/ Molecular Probes; A11029	1:250	N/A
11. Alexa fluor® 488 Goat Anti-Rabbit IgG (H+L)	Secondary	Invitrogen™/ Molecular Probes; A11034	1:250	N/A
12. Alexa fluor® 488 Donkey Anti-Goat IgG (H+L)	Secondary	Invitrogen™/ Molecular Probes; A11055	1:250	N/A
13. Alexa fluor® 546 Goat Anti-Mouse IgG (H+L)	Secondary	Invitrogen™/ Molecular Probes; A11030	1:250	N/A

Name of the Antibody	Primary or Secondary	Company and Cat #	Dilution used for IC	Dilution used for Western
14. Oregon Green, Goat Anti-Rabbit IgG (H+L)	Secondary	Invitrogen™/ Molecular Probes; O11038	1:500	N/A
15. Goat Anti-Mouse IgG (Fab specific) PEROXIDASE CONJUGATE	Secondary	Sigma; A2304	N/A	1:500 (for Anti-Smad2) 1:40,000 (for Anti-Alpha Tubulin)
16. Anti-Rabbit IgG (Whole Molecule) PEROXIDASE CONJUGATE	Secondary	Sigma; A0545	N/A	1:200 (for Anti-PSmad2) 1:300 (for Anti-Smad4)
17. Rabbit Anti-Goat IgG-HRP	Secondary	Santa Cruz Biotechnology; sc-A1111	N/A	1:2000 for Anti-ActRIIA

* Purified mouse monoclonal Anti-Smad2/3 was raised against mouse Smad2 amino acid 142 – 263. There is a possibility that due to high sequence conservation between Smad2 and Smad3, this antibody could also detect Smad3 protein. Previous use of this antibody in experiments involving *X. laevis* (Faure et al., 2000; Schohl & Fagotto, 2002) addressed detection of Smad2 only, not Smad3. Moreover, in mesendoderm specification in X. laevis, Smad3 plays a lesser role. Due to these reasons, I will refer to the protein detected by this antibody in *E. coqui* as Smad2.

** Due to the similar sequence conservation issue, the rabbit polyclonal anti-Phospho-Smad2 (Ser465/467) antibody may also detect the equivalent phosphorylated sites in Smad3. Due to the reasons as stated above, my western blot experiments in *E. coqui* involving anti-Phospho-Smad2 antibody will address the detection of PSmad2 only, not PSmad3.

CHAPTER THREE: RESULTS

I. Cloning the *EcSmad2* gene:

The protein encoded by the *Smad* gene belongs to the SMAD family of proteins, similar to the gene products of the *D. melanogaster* gene 'mothers against decapentaplegic' (*Mad*) and the *C. elegans* gene *Sma*. Smad2, the essential signal transducer of Nodal-signaling pathway in vertebrate embryonic development, has been cloned in several vertebrate species as well as in *D. melanogaster* and *C. elegans*. Cloning *EcSmad2* would help to understand the spatial and temporal activation of Nodal signaling. Two isoforms of *X. laevis* Smad2 are present, playing a crucial role in transducing Nodal related TGF-β signals into the nucleus to specify endodermal and mesodermal fates (Faure *et al*., 2000). Smad2Δexon3, a shorter alternatively spliced variant of Smad2 differs from the full length Smad2 functionally (Dennler *et al*., 1998; Labbé *et al*., 1998; Yagi *et al*., 1999). I have successfully cloned both the full length ORF of *EcSmad2* and the *EcSmad2*Δ*exon3* splice variant along with their 5'-and 3'-UTRs.

A. Degenerate PCR:

Alignment of *Smad2* sequences revealed a high degree of conservation among vertebrate species. Based on this, the size of *EcSmad2* ORF was predicted to be ~1400 bp. To clone the *Smad2* gene from *E. coqui*, several degenerate PCR primer sets were designed based on a comparison of the *Smad2* cDNA sequences of *X. laevis*, mouse and human. A series of degenerate PCRs were done on the *E. coqui* TS3 cDNA templates to generate fragments of *EcSmad2*.

Among the five initial reactions, only two, reaction#4: *EcSmad2*-F2-*EcSmad2*-R1, and reaction#5: *EcSmad2*-F3- *EcSmad2*-R1, gave results. Both the products were towards the 3'end of the ORF. The *EcSmad2*-F2-*EcSmad2*-R1 band size was ~1100 bp, whereas *EcSmad2*-F3- *EcSmad2*-R1 was ~750 bp (Fig. 10). Surprisingly, in both cases, doublets of bands were found indicating the presence of both isoforms as seen in other vertebrate species (Fig. 10). As the product obtained in reaction#4 was bigger than that of reaction#5, further experiments were performed on those ~1100 bp bands. Cloning and sequence analysis using NCBI BLAST program confirmed the presence of *Smad2* sequences in those clones.

B. Cloning the 3'UTR:

The 3'UTR region of *EcSmad2* was cloned using PCR on Custom SMART^{TM} *E. coqui* ovarian cDNA Library in λTriplEx2^{TM} (Clontech, Cat # CS1023u) as described in VII.B (Fig. 8) of Materials and Methods. Two vector specific primers, 5' pTriplEx2 Seq Pri (forward) and 3' pTriplEx2 Seq Pri (reverse), flanking the Multiple Cloning Site (MCS) were separately used in combination with an exact *EcSmad2* forward, *EcSmad2*-F18 designed on the basis of the known sequence. Product was obtained from reaction with *EcSmad2*-F18 and 3' pTriplEx2 Seq Pri. Sequencing of the amplified PCR product revealed the presence of 205 bps after the stop codon "TAA" and ending with a poly(A)-tail. Clustal W analysis of the deduced *EcSmad2* ORF sequence along with the 3'UTR revealed 85% and 82% sequence homology with that of *X. laevis* and *X. tropicalis* respectively (Table 8). The fragment also had a polyadenylation signal "AATAAA", which marks the cleavage site for the RNA transcript ~30 bp past the signal (Fig 11).

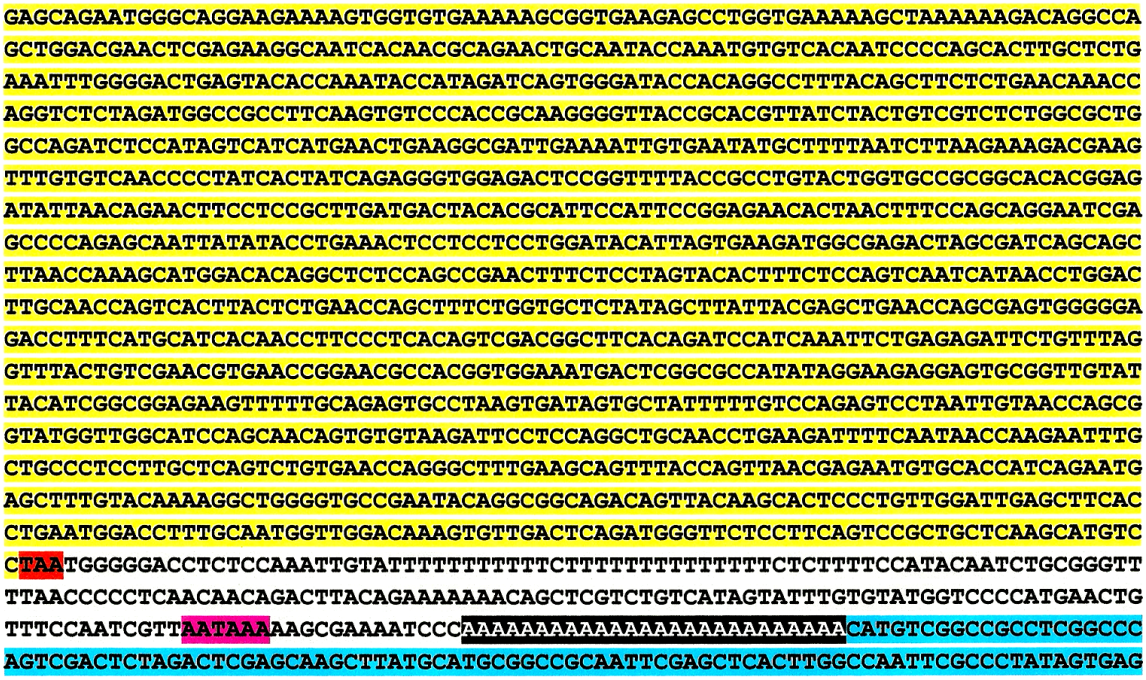
Fig. 10: PCR fragments of *EcSmad2* full length ORF and *EcSmad2*Δ*exon3* splice variant. Lane#1 is the 1Kb plus DNA ladder (usb), Lane#2 shows bands obtained for two isoforms from the *EcSmad2*-F2-*EcSmad2*-R1 reaction. In Lane#3 are the two products from *EcSmad2*-F3- *EcSmad2*-R1 reaction. Lane#4 is the negative control (- ve), where dIH_2O was used instead of a DNA template.

Fig. 10:


DNA Ladder
F2-R1
F3-R1
- ve control
1500 bp
1000 bp
500 bp
EcSmad2
EcSmad2Δexon3

Fig. 11: Incomplete nucleotide sequence of *EcSmad2* ORF along with full 3'UTR. Section with yellow background ( ) represents the portion of the ORF starting at ~300 bp from the 5' start site and ending in the stop codon TAA ( ). The grey section after the stop codon represents the 3'UTR of the gene, which contains the polyadenylation signal "AATAAA" ( ) and ends with poly(A)-tail ( ).

Fig 11:


GAGCAGAATGGGCAGGAAGAAAAGTGGTGTGAAAAAGCGGTGAAGAGCCTGGTGAAAAAGCTAAAAAAGACAGGCCA
GCTGGACGAACTCGAGAAGGCAATCACAACGCAGAACTGCAATACCAAATGTGTCACAATCCCCAGCACTTGCTCTG
AAATTTGGGGACTGAGTACACCAAATACCATAGATCAGTGGGATACCACAGGCCTTTACAGCTTCTCTGAACAAACC
AGGTCTCTAGATGGCCGCCTTCAAGTGTCCCACCGCAAGGGGTTACCGCACGTTATCTACTGTCGTCTCTGGCGCTG
GCCAGATCTCCATAGTCATCATGAACTGAAGGCGATTGAAAATTGTGAATATGCTTTTAATCTTAAGAAAGACGAAG
TTTGTGTCAACCCCTATCACTATCAGAGGGTGGAGACTCCGGTTTTACCGCCTGTACTGGTGCCGCGGCACACGGAG
ATATTAACAGAACTTCCTCCGCTTGATGACTACACGCATTCCATTCCGGAGAACACTAACTTTCCAGCAGGAATCGA
GCCCCAGAGCAATTATATACCTGAAACTCCTCCTCCTGGATACATTAGTGAAGATGGCGAGACTAGCGATCAGCAGC
TTAACCAAAGCATGGACACAGGCTCTCCAGCCGAACTTTCTCCTAGTACACTTTCTCCAGTCAATCATAACCTGGAC
TTGCAACCAGTCACTTACTCTGAACCAGCTTTCTGGTGCTCTATAGCTTATTACGAGCTGAACCAGCGAGTGGGGGA
GACCTTTCATGCATCACAACCTTCCCTCACAGTCGACGGCTTCACAGATCCATCAAATTCTGAGAGATTCTGTTTAG
GTTTACTGTCGAACGTGAACCGGAACGCCACGGTGGAAATGACTCGGCGCCATATAGGAAGAGGAGTGCGGTTGTAT
TACATCGGCGGAGAAGTTTTTGCAGAGTGCCTAAGTGATAGTGCTATTTTTGTCCAGAGTCCTAATTGTAACCAGCG
GTATGGTTGGCATCCAGCAACAGTGTGTAAGATTCCTCCAGGCTGCAACCTGAAGATTTTCAATAACCAAGAATTTG
CTGCCCTCCTTGCTCAGTCTGTGAACCAGGGCTTTGAAGCAGTTTACCAGTTAACGAGAATGTGCACCATCAGAATG
AGCTTTGTACAAAAGGCTGGGGTGCCGAATACAGGCGGCAGACAGTTACAAGCACTCCCTGTTGGATTGAGCTTCAC
CTGAATGGACCTTTGCAATGGTTGGACAAAGTGTTGACTCAGATGGGTTCTCCTTCAGTCCGCTGCTCAAGCATGTC
CTAATGGGGGACCTCTCCAAATTGTATTTTTTTTTTTCTTTTTTTTTTTTTTTCTCTTTTCCATACAATCTGCGGGTT
TTAACCCCCTCAACAACAGACTTACAGAAAAAAACAGCTCGTCTGTCATAGTATTTGTGTATGGTCCCCATGAACTG
TTTCCAATCGTTAATAAAAAGCGAAAATCCCAAAAAAAAAAAAAAAAAAAAAAAAACATGTCGGCCGCCTCGGCCC
AGTCGACTCTAGACTCGAGCAAGCTTATGCATGCGGCCGCAATTCGAGCTCACTTGGCCAATTCGCCCTATAGTGAG

In case of the *EcSmad2*, I found poly-adenylation site present 13 bp after the signal sequence in the 3'UTR. The sequence present after poly(A)-tail was of unknown identity.

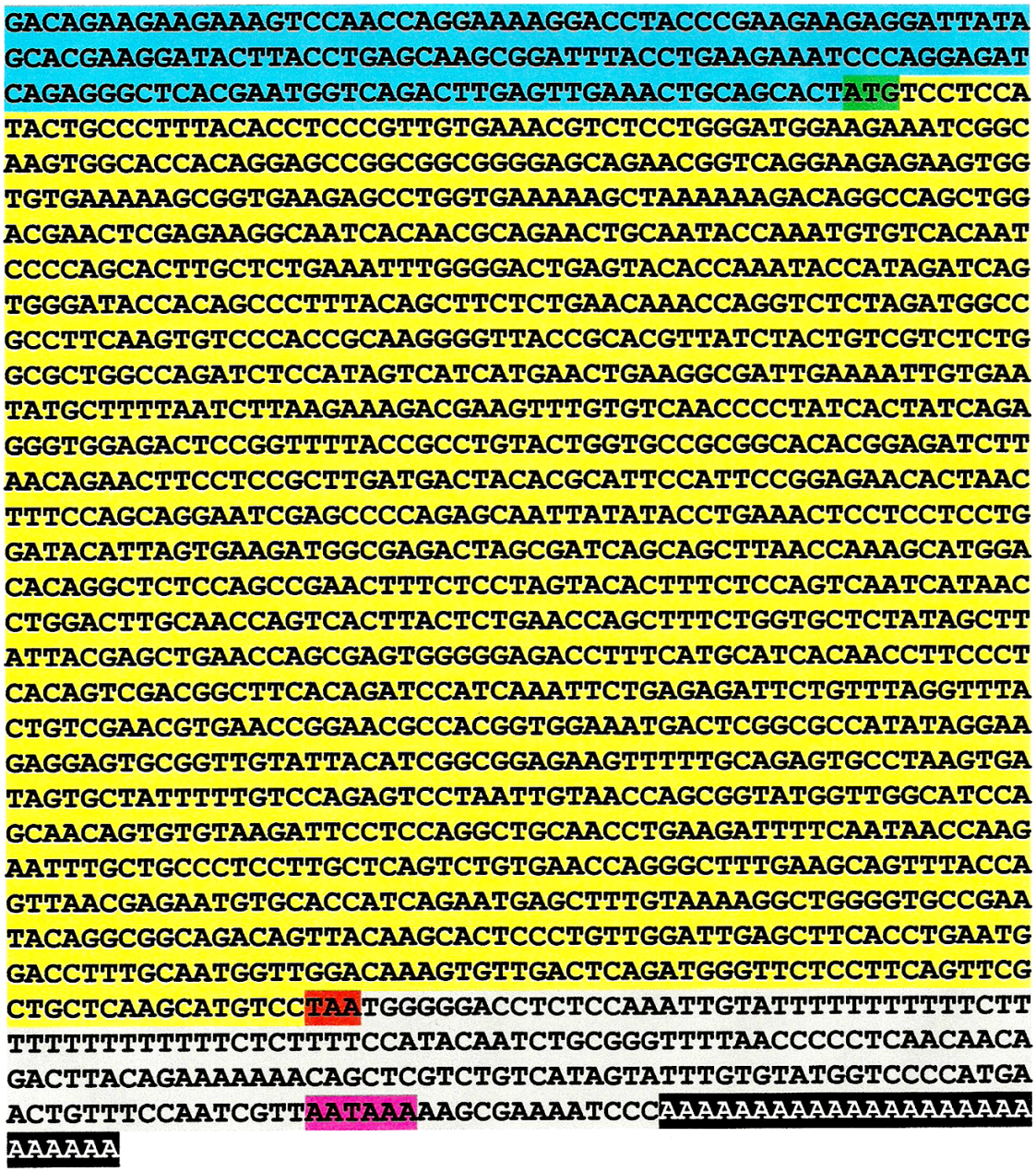
C. Cloning the 5'UTR and 5'-end of the ORF:

A 5'RACE PCR kit was used to amplify the initial missing segment of the *EcSmad2* ORF along with the 5'UTR, following the manufacturer's (Invitrogen) instructions. Two gene specific exact reverse primers - *EcSmad2*-RP16 GSP1 and *EcSmad2*-RP17 GSP2 were designed (Fig. 9). The RACE specific forward primers were provided with the kit. The primers were used on cDNA templates prepared from *E. coqui* TS3 embryos. The PCR product obtained was gel purified, cloned, and sequenced. An initial round of sequencing revealed the presence of 155 bp of 5'UTR along with the rest of the ORF, starting from 5' start site "ATG". The compiled sequence of the full length ORF along with its 5' and 3'-UTRs is shown in Fig. 12. Two new exact primers, *EcSmad2*-RP16 GSP3 and *EcSmad2*-RP17 GSP4, were designed to use in PCRs to screen transformant clones for *EcSmad2*.

Once the full length ORF sequence was deduced, two more primers (Table 1) were designed to PCR amplify the full ORF from TS3 cDNA template in order to clone and make freezer stocks. Both isoforms were recovered by cutting the gel, and they were separately cloned. Sequencing revealed the length of the full ORF to be 1404 bp. ApE 2.0.44 software was used to generate a peptide sequence of 467 amino acids with a predicted molecular weight of 52.4 kDa. The splice variant, *EcSmad2*Δ*exon3*, was recovered using the same primer set, and it had an ORF of 1314 bp. The predicted peptide was 437 amino acids, and the estimated molecular weight was 49 kDa.

Fig. 12: Full *EcSmad2* ORF along with its 5' and 3'-UTR. The first starting segment with BLUE background ( ) represents the 5'UTR. An ATG start codon ( ) initiates the ORF ( ), which ends with the TAA stop codon ( ). ( ) represents the 3'UTR which ends in poly(A)-tail ( ). A polyadenylation signal "AATAAA" ( ) lies inside the 3'UTR.

Fig 12:


GACAGAAGAAGAAAGTCCAACCAGGAAAAGGACCTACCCGAAGAAGAGGATTATA
GCACGAAGGATACTTACCTGAGCAAGCGGATTTACCTGAAGAAATCCCAGGAGAT
CAGAGGGCTCACGAATGGTCAGACTTGAGTTGAAACTGCAGCACTATGTCCTCCA
TACTGCCCTTTACACCTCCCGTTGTGAAACGTCTCCTGGGATGGAAGAAATCGGC
AAGTGGCACCACAGGAGCCGGCGGCGGGGAGCAGAACGGTCAGGAAGAGAAGTGG
TGTGAAAAAGCGGTGAAGAGCCTGGTGAAAAAGCTAAAAAAGACAGGCCAGCTGG
ACGAACTCGAGAAGGCAATCACAACGCAGAACTGCAATACCAAATGTGTCACAAT
CCCCAGCACTTGCTCTGAAATTTGGGGACTGAGTACACCAAATACCATAGATCAG
TGGGATACCACAGCCCTTTACAGCTTCTCTGAACAAACCAGGTCTCTAGATGGCC
GCCTTCAAGTGTCCCACCGCAAGGGGTTACCGCACGTTATCTACTGTCGTCTCTG
GCGCTGGCCAGATCTCCATAGTCATCATGAACTGAAGGCGATTGAAAATTGTGAA
TATGCTTTTAATCTTAAGAAAGACGAAGTTTGTGTCAACCCCTATCACTATCAGA
GGGTGGAGACTCCGGTTTTACCGCCTGTACTGGTGCCGCGGCACACGGAGATCTT
AACAGAACTTCCTCCGCTTGATGACTACACGCATTCCATTCCGGAGAACACTAAC
TTTCCAGCAGGAATCGAGCCCCAGAGCAATTATATACCTGAAACTCCTCCTCCTG
GATACATTAGTGAAGATGGCGAGACTAGCGATCAGCAGCTTAACCAAAGCATGGA
CACAGGCTCTCCAGCCGAACTTTCTCCTAGTACACTTTCTCCAGTCAATCATAAC
CTGGACTTGCAACCAGTCACTTACTCTGAACCAGCTTTCTGGTGCTCTATAGCTT
ATTACGAGCTGAACCAGCGAGTGGGGGAGACCTTTCATGCATCACAACCTTCCCT
CACAGTCGACGGCTTCACAGATCCATCAAATTCTGAGAGATTCTGTTTAGGTTTA
CTGTCGAACGTGAACCGGAACGCCACGGTGGAAATGACTCGGCGCCATATAGGAA
GAGGAGTGCGGTTGTATTACATCGGCGGAGAAGTTTTTGCAGAGTGCCTAAGTGA
TAGTGCTATTTTTGTCCAGAGTCCTAATTGTAACCAGCGGTATGGTTGGCATCCA
GCAACAGTGTGTAAGATTCCTCCAGGCTGCAACCTGAAGATTTTCAATAACCAAG
AATTTGCTGCCCTCCTTGCTCAGTCTGTGAACCAGGGCTTTGAAGCAGTTTACCA
GTTAACGAGAATGTGCACCATCAGAATGAGCTTTGTAAAAGGCTGGGGTGCCGAA
TACAGGCGGCAGACAGTTACAAGCACTCCCTGTTGGATTGAGCTTCACCTGAATG
GACCTTTGCAATGGTTGGACAAAGTGTTGACTCAGATGGGTTCTCCTTCAGTTCG
CTGCTCAAGCATGTCCTAATGGGGGACCTCTCCAAATTGTATTTTTTTTTTTCTT
TTTTTTTTTTTTCTCTTTTCCATACAATCTGCGGGTTTTAACCCCCTCAACAACA
GACTTACAGAAAAAAACAGCTCGTCTGTCATAGTATTTGTGTATGGTCCCCATGA
ACTGTTTCCAATCGTTAATAAAAAGCGAAAATCCCAAAAAAAAAAAAAAAAAAA
AAAAAA

Alignment of the two amino acid sequences is shown in Fig. 13, which also indicates the position of 90 bp exon 3. Two clones for each isoform (clone B and K for the full length form and clone 2 and 3 for the splice variant) were used to make freezer stocks and stored at – 80^0C. Alignment of *EcSmad2* coding sequences along with its 3' and 5'UTR showed 85% and 82% sequence conservation with those of *X. laevis* and *X. tropicalis* by Clustal W analysis (Table 5). When the full length *EcSmad2* ORF nucleotide (Table 6) sequences were aligned with those of *X. laevis* and *X. tropicalis,* mice and human using Clustal W, the result showed 86% nucleotide identity shared among the three frogs. *EcSmad2* nucleotide sequence was also 82% identical to that of both human and mice. The EcSmad2 predicted protein sequence alignment with Smad2 of different organisms revealed 99% conservation with those of *X. laevis* and *X. tropicalis,* and 98% with those of human, chicken and mouse (Table 7). Based on the cloning results, a schematic comparison of both isoforms of *EcSmad2* is shown in Fig. 14. The 99% conservation of Smad2 full-length peptide sequence among *E. coqui*, *X. laevis* and *X. tropicalis* is shown in Fig. 15. It shows the positions of the conserved MH1 and MH2 domains. The C-terminal S^{465} and S^{467}, phosphorylation of which are essential for activation of Smad2 under the influence of Nodal-signaling, are conserved among all three frogs.

Table 5: Clustal W alignments of the Smad2 ORF along with 3' and 5'UTRs from *E. coqui*, *X. laevis* and *X. tropicalis*.

Table 6: Clustal W alignments of *Smad2* ORF sequences from *E. coqui* and other vertebrates.

Table 7: Clustal W alignments of Smad2 predicted peptide sequences from *E. coqui* and other vertebrates.

Table 5:

SeqA	Name	Len(nt)	SeqB	Name	Len(nt)	Score
1	E_coqui	1520	2	X_laevis	1454	85
1	E_coqui	1520	3	X_tropicalis	1861	82
2	X_laevis	1454	3	X_tropicalis	1861	94

Table 6:

SeqA	Name	Len(nt)	SeqB	Name	Len(nt)	Score
1	H_sapiens	1404	2	M_musculus	1404	92
1	H_sapiens	1404	3	X_tropicalis	1394	83
1	H_sapiens	1404	4	X_laevis	1404	83
1	H_sapiens	1404	5	E_coqui	1404	82
2	M_musculus	1404	3	X_tropicalis	1394	83
2	M_musculus	1404	4	X_laevis	1404	82
2	M_musculus	1404	5	E_coqui	1404	82
3	X_tropicalis	1394	4	X_laevis	1404	95
3	X_tropicalis	1394	5	E_coqui	1404	86
4	X_laevis	1404	5	E_coqui	1404	86

Table 7:

SeqA	Name	Len(aa)	SeqB	Name	Len(aa)	Score
1	E_coqui	467	2	G_gallus	467	98
1	E_coqui	467	3	H_sapiens	467	98
1	E_coqui	467	4	M_musculus	467	98
1	E_coqui	467	5	X_tropicalis	467	99
1	E_coqui	467	6	X_laevis	467	99
2	G_gallus	467	3	H_sapiens	467	99
2	G_gallus	467	4	M_musculus	467	98
2	G_gallus	467	5	X_tropicalis	467	98
2	G_gallus	467	6	X_laevis	467	98
3	H_sapiens	467	4	M_musculus	467	99
3	H_sapiens	467	5	X_tropicalis	467	98
3	H_sapiens	467	6	X_laevis	467	98
4	M_musculus	467	5	X_tropicalis	467	98
4	M_musculus	467	6	X_laevis	467	97
5	X_tropicalis	467	6	X_laevis	467	99

Fig. 13: Predicted protein sequences of the two Smad2 isoforms. (A) Predicted amino acid sequences of the full-length ORF (clone B), and (B) *Smad2Δexon3* splice variant (clone 3) from their nucleotide sequences using ApE software. (C) Clustal W alignment of the two protein sequences showing the position of exon 3 (highlighted with yellow).

Fig. 13:

(A) Predicted amino acid sequence of Full ORF, clone B:

MSSILPFTPPVVKRLLGWKKSASGTTGAGGGEQNGQEEKWCEKAVKSLVKKLKKTGQLD
ELEKAITTQNCNTKCVTIPSTCSEIWGLSTPNTIDQWDTTALYSFSEQTRSLDGRLQVS
HRKGLPHVIYCRLWRWPDLHSHHELKAIENCEYAFNLKKDEVCVNPYHYQRVETPVLPP
VLVPRHTEILTELPPLDDYTHSIPENTNFPAGIEPQSNYIPETPPPGYISEDGETSDQQ
LNQSMDTGSPAELSPSTLSPVNHNLDLQPVTYSEPAFWCSIAYYELNQRVGETFHASQP
SLTVDGFTDPSNSERFCLGLLSNVNRNATVEMTRRHIGRGVRLYYIGGEVFAECLSDSA
IFVQSPNCNQRYGWHPATVCKIPPGCNLKIFNNQEFAALLAQSVNQGFEAVYQLTRMCT
IRMSFVKGWGAEYRRQTVTSTPCWIELHLNGPLQWLDKVLTQMGSPSVRCSSMS

(B) Predicted amino acid sequence of EcSmad2Δexon3, clone 3:

MSSILPFTPPVVKRLLGWKKSASGTTGAGGGEQNGQEEKWCEKAVKSLVKKLKKTGQLD
ELEKAITTQNCNTKCVTIPRSLDGRLQVSHRKGLPHVIYCRLWRWPDLHSHHELKAIEN
CEYAFNLKKDEVCVNPYHYQRVETPVLPPVLVPRHTEILTELPPLDDYTHSIPENTNFP
AGIEPQSNYIPETPPPGYISEDGETSDQQLNQSMDTGSPAELSPSTLSPVNHNLDLQPV
TYSEPAFWCSIAYYELNQRVGETFHASQPSLTVDGFTDPSNSERFCLGLLSNVNRNATV
EMTRRHIGRGVRLYYIGGEVFAECLSDSAIFVQSPNCNQRYGWHPATVCKIPPGCNLKI
FNNQEFAALLAQSVNQGFEAVYQLTRMCTIRMSFVKGWGAEYRRQTVTSTPCWIELHLN
GPLQWLDKVLTQMGSPSVRCSSMS

(C) CLUSTAL 2.0.12 multiple sequence alignment:

```
clone_B         MSSILPFTPPVVKRLLGWKKSASGTTGAGGGEQNGQEEKWCEKAVKSLVKKLKKTGQLDE  60
clone_3         MSSILPFTPPVVKRLLGWKKSASGTTGAGGGEQNGQEEKWCEKAVKSLVKKLKKTGQLDE  60
                ************************************************************

clone_B         LEKAITTQNCNTKCVTIPSTCSEIWGLSTPNTIDQWDTTALYSFSEQTRSLDGRLQVSHR 120
clone_3         LEKAITTQNCNTKCVTIP----------------------------RSLDGRLQVSHR  90
                ******************<---------- Exon 3 ---------->************

clone_B         KGLPHVIYCRLWRWPDLHSHHELKAIENCEYAFNLKKDEVCVNPYHYQRVETPVLPPVLV 180
clone_3         KGLPHVIYCRLWRWPDLHSHHELKAIENCEYAFNLKKDEVCVNPYHYQRVETPVLPPVLV 150
                ************************************************************

clone_B         PRHTEILTELPPLDDYTHSIPENTNFPAGIEPQSNYIPETPPPGYISEDGETSDQQLNQS 240
clone_3         PRHTEILTELPPLDDYTHSIPENTNFPAGIEPQSNYIPETPPPGYISEDGETSDQQLNQS 210
                ************************************************************

clone_B         MDTGSPAELSPSTLSPVNHNLDLQPVTYSEPAFWCSIAYYELNQRVGETFHASQPSLTVD 300
clone_3         MDTGSPAELSPSTLSPVNHNLDLQPVTYSEPAFWCSIAYYELNQRVGETFHASQPSLTVD 270
                ************************************************************

clone_B         GFTDPSNSERFCLGLLSNVNRNATVEMTRRHIGRGVRLYYIGGEVFAECLSDSAIFVQSP 360
clone_3         GFTDPSNSERFCLGLLSNVNRNATVEMTRRHIGRGVRLYYIGGEVFAECLSDSAIFVQSP 330
                ************************************************************

clone_B         NCNQRYGWHPATVCKIPPGCNLKIFNNQEFAALLAQSVNQGFEAVYQLTRMCTIRMSFVK 420
clone_3         NCNQRYGWHPATVCKIPPGCNLKIFNNQEFAALLAQSVNQGFEAVYQLTRMCTIRMSFVK 390
                ************************************************************

clone_B         GWGAEYRRQTVTSTPCWIELHLNGPLQWLDKVLTQMGSPSVRCSSMS 467
clone_3         GWGAEYRRQTVTSTPCWIELHLNGPLQWLDKVLTQMGSPSVRCSSMS 437
                *********************************************
```

Fig 14: Comparison of ORFs of *EcSmad2* and its splice variant *EcSmad2Δexon3*.

The lengths of the full length and the splice variant isoforms are 1404 bp and 1314 bp, respectively. The highly conserved MH1 and MH2 domains are separated by the less conserved linker region (light green box). The *EcSmad2Δexon3* ORF is lacking *exon* 3 (orange box) due to alternative splicing.

Fig. 14:

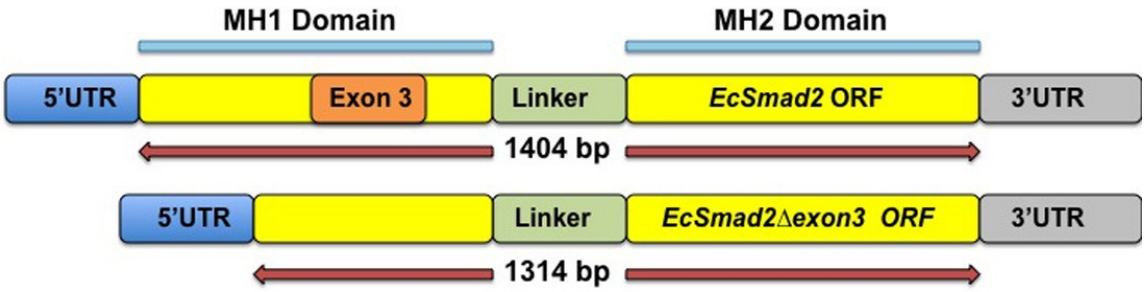
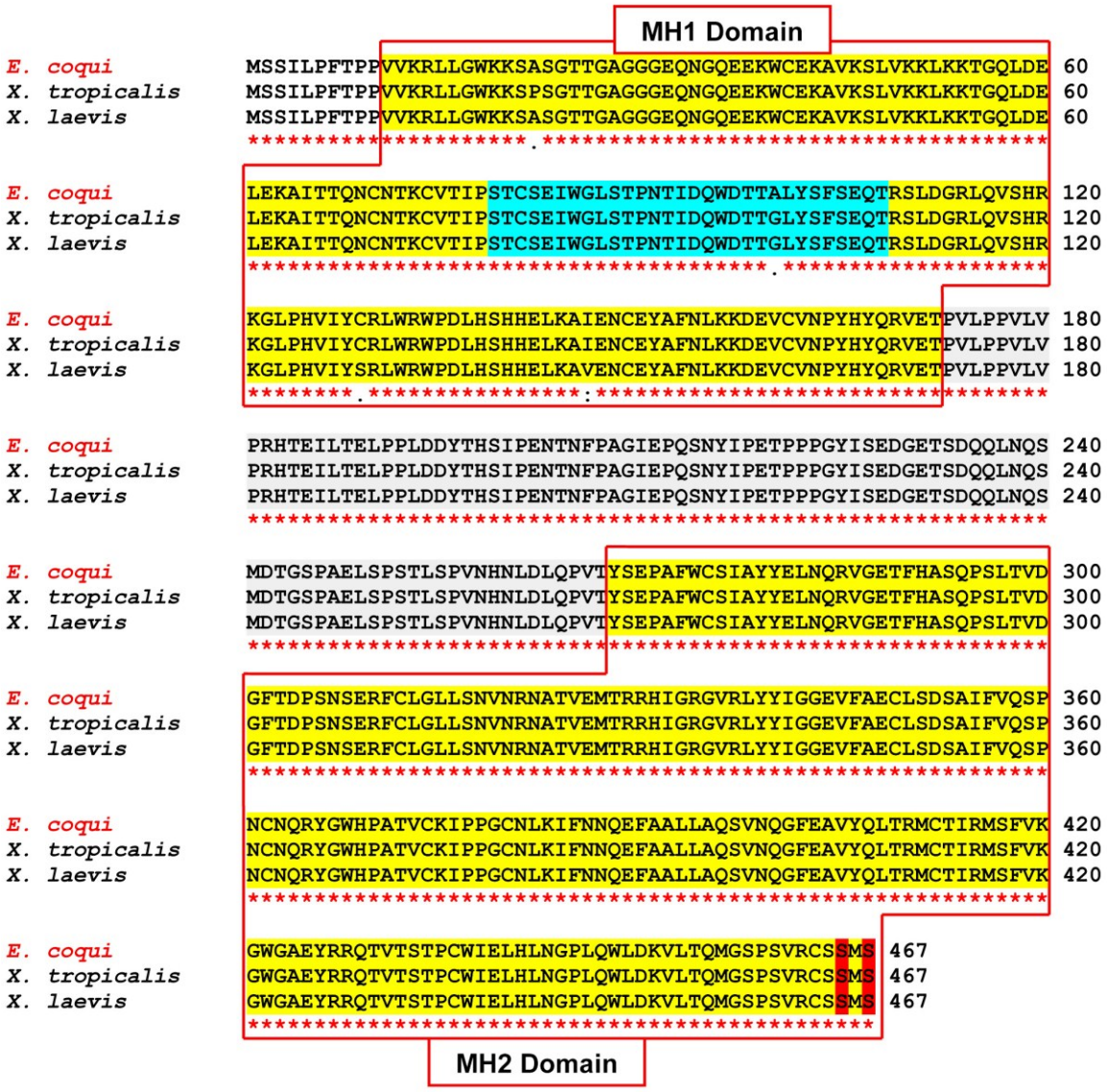

MH1 Domain
MH2 Domain
5'UTR
Exon 3
Linker
EcSmad2 ORF
3'UTR
1404 bp
5'UTR
Linker
EcSmad2Δexon3 ORF
3'UTR
1314 bp

Fig. 15: Conservation of Smad2 predicted peptide sequences from *E. coqui*, *X. tropicalis* and *X. laevis*. The first box with red boundary and yellow background, shows the complete conservation of the MH1 domain except for four amino acids. The MH1 domain contains exon 3 (blue background). This segment of amino acids is followed by a linker region, which is 100% conserved among the three species. Finally, the last segment of amino acids in the second box, with red boundary and yellow background, represents the MH2 domain, which is also 100% conserved among all three frogs. At the end of MH2 domain, there are two serine residues, S^{465} and S^{467}, phosphorylation of which are essential for activation of Smad2.

Fig. 15:


MH1 Domain
E. coqui MSSILPFTPPVVKRLLGWKKSASGTTGAGGGEQNGQEEKWCEKAVKSLVKKLKKTGQLDE 60
X. tropicalis MSSILPFTPPVVKRLLGWKKSPSGTTGAGGGEQNGQEEKWCEKAVKSLVKKLKKTGQLDE 60
X. laevis MSSILPFTPPVVKRLLGWKKSASGTTGAGGGEQNGQEEKWCEKAVKSLVKKLKKTGQLDE 60
E. coqui LEKAITTQNCNTKCVTIPSTCSEIWGLSTPNTIDQWDTTALYSFSEQTRSLDGRLQVSHR 120
X. tropicalis LEKAITTQNCNTKCVTIPSTCSEIWGLSTPNTIDQWDTTGLYSFSEQTRSLDGRLQVSHR 120
X. laevis LEKAITTQNCNTKCVTIPSTCSEIWGLSTPNTIDQWDTTGLYSFSEQTRSLDGRLQVSHR 120
E. coqui KGLPHVIYCRLWRWPDLHSHHELKAIENCEYAFNLKKDEVCVNPYHYQRVETPVLPPVLV 180
X. tropicalis KGLPHVIYCRLWRWPDLHSHHELKAIENCEYAFNLKKDEVCVNPYHYQRVETPVLPPVLV 180
X. laevis KGLPHVIYSRLWRWPDLHSHHELKAVENCEYAFNLKKDEVCVNPYHYQRVETPVLPPVLV 180
E. coqui PRHTEILTELPPLDDYTHSIPENTNFPAGIEPQSNYIPETPPPGYISEDGETSDQQLNQS 240
X. tropicalis PRHTEILTELPPLDDYTHSIPENTNFPAGIEPQSNYIPETPPPGYISEDGETSDQQLNQS 240
X. laevis PRHTEILTELPPLDDYTHSIPENTNFPAGIEPQSNYIPETPPPGYISEDGETSDQQLNQS 240
E. coqui MDTGSPAELSPSTLSPVNHNLDLQPVTYSEPAFWCSIAYYELNQRVGETFHASQPSLTVD 300
X. tropicalis MDTGSPAELSPSTLSPVNHNLDLQPVTYSEPAFWCSIAYYELNQRVGETFHASQPSLTVD 300
X. laevis MDTGSPAELSPSTLSPVNHNLDLQPVTYSEPAFWCSIAYYELNQRVGETFHASQPSLTVD 300
E. coqui GFTDPSNSERFCLGLLSNVNRNATVEMTRRHIGRGVRLYYIGGEVFAECLSDSAIFVQSP 360
X. tropicalis GFTDPSNSERFCLGLLSNVNRNATVEMTRRHIGRGVRLYYIGGEVFAECLSDSAIFVQSP 360
X. laevis GFTDPSNSERFCLGLLSNVNRNATVEMTRRHIGRGVRLYYIGGEVFAECLSDSAIFVQSP 360
E. coqui NCNQRYGWHPATVCKIPPGCNLKIFNNQEFAALLAQSVNQGFEAVYQLTRMCTIRMSFVK 420
X. tropicalis NCNQRYGWHPATVCKIPPGCNLKIFNNQEFAALLAQSVNQGFEAVYQLTRMCTIRMSFVK 420
X. laevis NCNQRYGWHPATVCKIPPGCNLKIFNNQEFAALLAQSVNQGFEAVYQLTRMCTIRMSFVK 420
E. coqui GWGAEYRRQTVTSTPCWIELHLNGPLQWLDKVLTQMGSPSVRCSSMS 467
X. tropicalis GWGAEYRRQTVTSTPCWIELHLNGPLQWLDKVLTQMGSPSVRCSSMS 467
X. laevis GWGAEYRRQTVTSTPCWIELHLNGPLQWLDKVLTQMGSPSVRCSSMS 467
MH2 Domain

II. Establishing temporal and spatial expression patterns for *EcSmad2* in *E. coqui* early embryo:

It was important to find out whether *EcSmad2* RNA is maternally contributed, whether it is expressed throughout all the early developmental stages, and whether expression is spatially restricted, such as to DE only. Based on my original hypothesis, it was possible that *EcSmad2* expression is spatially restricted to the MZ, which forms the DE. This idea was based on the fact that NE does not differentiate into any adult organs; it just provides nutrient to the growing embryos. Altogether, it was helpful to generate a complete profile of temporal and spatial regulation of *EcSmad2* gene expression.

The technique that I used for detecting *EcSmad2* expression was ΔΔCt method of Real-Time PCR (qPCR), a more quantitative method than RT-PCR. I designed *EcSmad2* specific primers (*EcSmad2*-qPCR FP1 and *EcSmad2*-qPCR RP1, Table 5) in the *EcSmad2* ORF. qPCR was performed on RNAs isolated from eggs as well as from different early embryonic stages like pre-gastrulation stage NF8, gastrulation stage NF10 and different post-gastrulation stages from TS3 to TS8. The ribosomal protein-coding gene *EcL8*, was previously shown to be expressed in both the MZ (or DE) and VC (or NE) tissues (Ninomiya et al., 2001), so it was chosen as an endogenous control. *EcL8* specific primers (*EcL8*-F4 and *EcL8*-R5, Table 5) were used in each experiment. Each qPCR experiment for temporal or spatial expression was performed three times independently with RNAs from different sources, and the overall results are discussed below.

A. Temporal Expression:

To detect temporal expression, I used RNAs isolated from whole embryos at each stage denoted as specific stage followed by "W". I made cDNAs using each RNA extract as template, and performed qPCR with them. Stages used in these experiments were: oocyte, NF8W, NF10W, TS3W, TS4W, TS5W, TS6W, TS7W and TS8W. The RQ value for oocyte was set to one, and the RQ values for other stages were expressed relative to oocyte. My results revealed a strong maternal contribution of *EcSmad2* transcript (Fig. 16). *EcSmad2* expression stays high at NF8 and NF10, followed by sharp decline in post-gastrulation stages.

B. Spatial Expression:

For spatial expression, RNAs were isolated from dissected embryonic tissues from each stage. For NF8 and NF10, RNAs were made from dissected MZ and VC tissues, as shown in Fig. 6 (denoted as specific stage followed by "MZ" or "VC"). For post-gastrulation stages from TS3 to TS8, the dissected tissues were denoted by the stage followed by "E" for embryo or "NE" for nutritional endoderm. After gastrulation, the embryos (E) have differentiated, and the NE has been formed from VC cells. The qPCR was performed with cDNAs as described before. The RQ values of MZ or E from each stage were set as one, and the RQ values for VC or NE are expressed relative to MZ or E. My results indicate higher relative expression of *EcSmad2* in VC tissues at both NF8 and NF10 (Fig. 17). This pattern changes after gastrulation is complete. In all post-gastrulation stages, a relatively lower level of *EcSmad2* RNA expression is present in NE.

Fig. 16: Temporal expression of *EcSmad2*. The ΔΔCt method of qPCR was used to establish the temporal expression pattern. Relative expression (RQ) at the oocyte stage was set to 1 to serve as reference, and the corresponding RQ values of all other stages are plotted relative to oocyte. *EcL8*, a ribosomal protein-coding gene was used as an endogenous control. Each green bar represents the mean of 3 independent experiments. Error bars show the standard deviations.

Fig. 16:

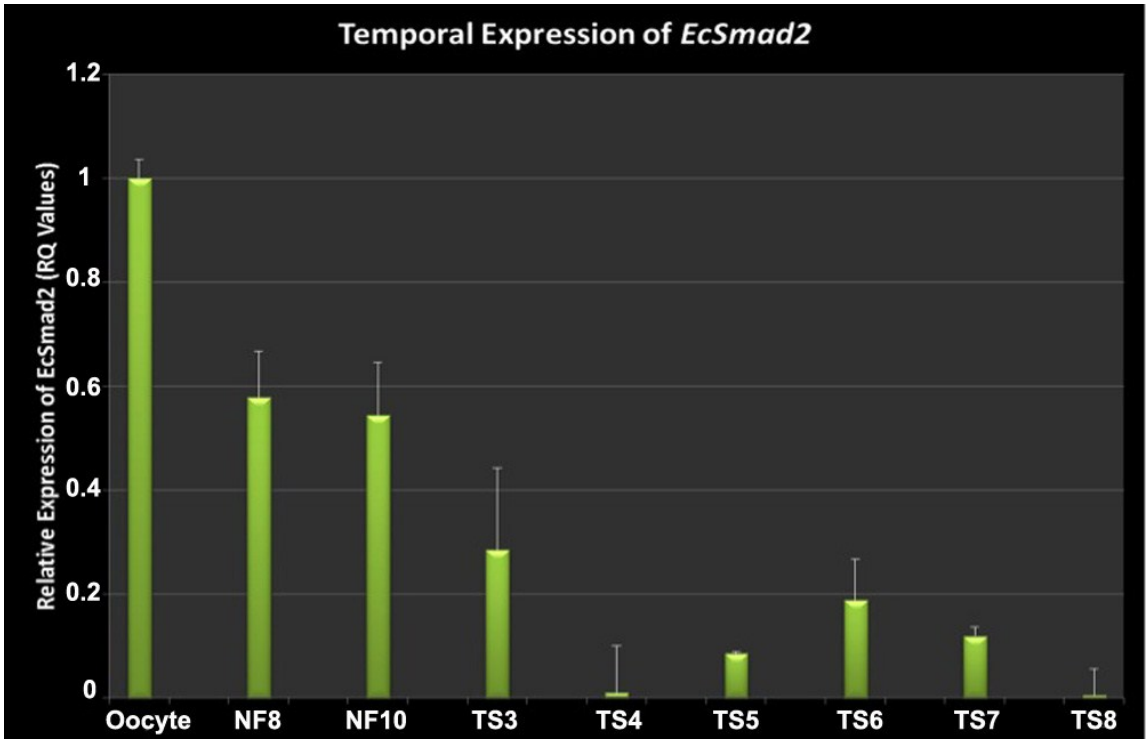
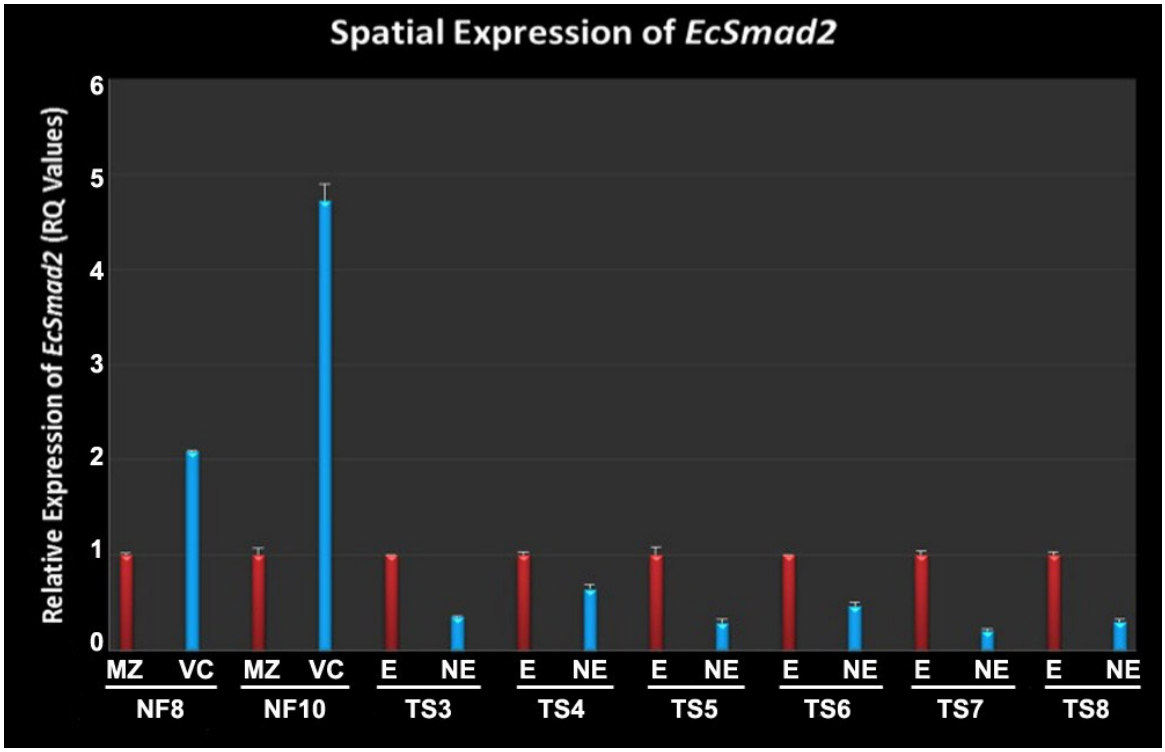

Temporal Expression of *EcSmad2*
Relative Expression of EcSmad2 (RQ Values)
1.2
1
0.8
0.6
0.4
0.2
0
Oocyte
NF8
NF10
TS3
TS4
TS5
TS6
TS7
TS8

Fig. 17: Spatial expression of *EcSmad2*. The pattern of *EcSmad2* expression was established by the ΔΔCt method of qPCR. Relative expression (RQ) of either MZ or E from each stage was set to 1 (red bars) to serve as a reference for the corresponding RQ values of the VC or NE samples (blue bars). *EcL8* serves as an endogenous control. Mean values of 3 independent experiments are used in the figure. The NF8 and NF10 stages showed higher levels of *EcSmad2* expression in VC, the prospective nutritional endodermal tissues. All the post-gastrulation stages tested showed comparatively lower levels of *EcSmad2* expression in nutritional endoderm.

Fig. 17:


Spatial Expression of EcSmad2
Relative Expression of EcSmad2 (RQ Values)
6
5
4
3
2
1
0
MZ VC MZ VC E NE E NE E NE E NE E NE E NE
NF8 NF10 TS3 TS4 TS5 TS6 TS7 TS8

III. Phospho-EcSmad2, EcSmad4 and ActRIIA in the prospective DE vs. prospective NE:

Nodal signaling leads to the phosphorylation of Smad2, which in turn binds with its partner Smad4 (Kumar et al., 2001; Lee et al., 2001; Yeo and Whitman, 2001; Ross and Hill, 2008). Nuclear translocation of this protein complex is an essential event leading to the specification of endoderm and mesoderm in early embryonic development (Lee et al., 2001; Yeo and Whitman, 2001; Ross and Hill, 2008). Although the qPCR results provide information for the presence of *EcSmad2* RNA in the prospective nutritional endoderm tissues of early embryonic stages, it was important to look at the expression of the native and active forms of Smad2 protein. A cell could express RNA, but the RNA could get degraded or the translation into its protein product or the activation of such protein could be blocked due to some regulatory mechanism. To shed light into the reason behind the absence of differentiation for NE, detection of both native and active forms of EcSmad2 becomes critical. Detection of native vs. PSmad2 in the prospective DE vs. prospective NE of early embryos was accomplished by western blot analysis using anti-Smad2/3 and anti-PSmad2 antibodies.

Before western blot analysis, the first major step was to standardize a technique to isolate protein samples from *E. coqui* embryos, which are packed with yolk platelets. Obtaining a standard high quality of protein preparation was a long process, which involved painstaking efforts. Sixteen different recipes of homogenization buffers (HB) in total were tested to isolate a sample of high quality, which looked acceptable on a SDS-PAGE gel. It began with HB1 (Table 7), which was based on a recipe mentioned in Birsoy et al. (2005) to isolate protein samples from *X. laevis* embryos. As the physical

characteristics and chemical composition of *E. coqui* embryos are widely different from those of *X. laevis* embryos, HB1 failed to produce a satisfactory protein sample (Fig. 18). Following several modifications of that original recipe, I ended up with HB16, which produced a protein sample that looked acceptable on a gel stained with CBB. Fig. 18 shows gels, where the protein samples were prepared using four of the HB recipes. Once the protein isolation technique was standardized, Western blot analysis was executed using two antibodies – anti-Smad2/3 and anti-PSmad2 to detect the level of expression of native and active Smad2, respectively.

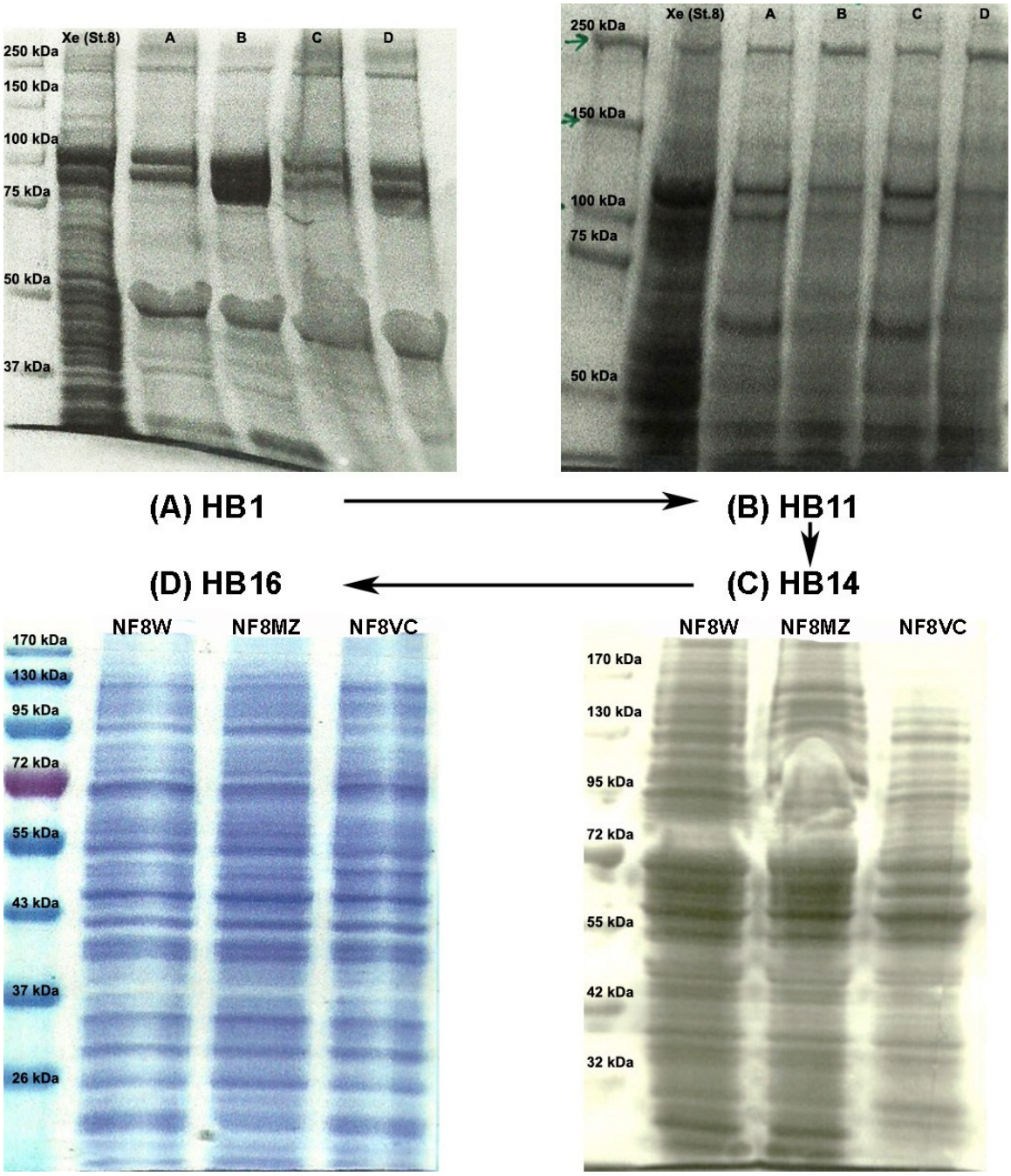
For the purpose of western analysis, the concentration of all protein samples was determined by the BCA method, and equal amounts were loaded in all lanes of a gel. In all experiments, one gel was run with the protein samples to check protein quality by CBB staining.

A. Detection of Smad2 and PSmad2:

The antibodies used for the experiments were purified mouse monoclonal anti-Smad2/3 (BD Transduction Laboratories; 610842) and rabbit polyclonal PSmad2 (Ser465/467) antibody (Cell Signaling; #3101) against native and the active phosphorylated forms of Smad2 protein, respectively. Although anti-Smad2/3 antibody was raised against mouse Smad2 amino acids 142 – 263, it was used successfully to detect *X. laevis* Smad2 protein on a western blot (Faure et al., 2000; Birsoy et al., 2005). The PSmad2 (Ser465/467) antibody was raised against a synthetic phospho-peptide (KLH-coupled), and detected endogenous levels of Smad2 only when dually phosphorylated at Ser465 and Ser467. This antibody was also successfully used on

Fig. 18: SDS-PAGE of proteins isolated with different homogenization buffers. (A) HB1: protein preparations were done using the HB1 recipe. Lane#1 – molecular weight (MW) markers; Lane#2 – protein preparation from *Xenopus laevis* embryos. Twenty NF8 embryos were homogenized in 200 µl of HB1; Lanes A-D – four protein preparations from *E. coqui* NF8, NF10, TS3 and TS4 embryos. The same numbers of embryos were homogenized in equal volumes of HB1 (see Materials and Methods). **(B) HB11**: Lane#1 – MW markers; Lane#2 – protein preparation from *X. laevis* NF8 embryos; Lanes A-D – four protein preparations from *E. coqui* NF8, NF10, TS3 and TS4 embryos, homogenized in equal volumes of HB11. **(C) HB14**: Lane#1 – MW markers; Lanes#2-4 – three individual protein preparations from *E. coqui* whole NF8 embryos (NF8W) and of dissected MZ and VC tissues (NF8MZ and NF8VC) made in HB14. **(D) HB16**: Lane#1 – MW markers; lanes#2-4 – a different set of protein preparations from whole and dissected MZ and VC tissues (NF8W, NF8MZ and NF8VC) made in HB16.

Fig. 18:


Xe (St.8)
A
B
C
D
250 kDa
150 kDa
100 kDa
75 kDa
50 kDa
37 kDa
(A) HB1
(B) HB11
(C) HB14
(D) HB16
NF8W
NF8MZ
NF8VC
170 kDa
130 kDa
95 kDa
72 kDa
55 kDa
43 kDa
42 kDa
37 kDa
32 kDa
26 kDa

protein samples prepared from *X. laevis* embryos (Faure et al., 2000; Birsoy et al., 2005). Based on my sequencing, the predicted EcSmad2 peptide sequence is a 99% match with those of *X. laevis* and *X. tropicalis*. Therefore there was a high chance that these would also work on the *E. coqui* protein samples on western blot.

The temporal expression of both the native and the active forms of EcSmad2 was detected in protein preparations made from oocytes and embryos at stages NF10, TS3, TS5 and TS7 (Fig 19). Both forms of EcSmad2 were present in all the stages tested. Moreover, the antibodies detected both isoforms of EcSmad2. Band intensity increased as the stages advanced (Fig. 19).

To detect the levels of native and PSmad2 in the prospective DE vs. prospective NE, western blotting was performed on protein preparations from dissected tissues. Fig. 20A shows the western blot for post-gastrulation stages. Native as well as active forms of Smad2 were detected in both embryonic and NE tissues. Levels of expressions in NE for both forms of the protein were lower than those in embryos. Fig. 20B represents a separate experiment, where the gastrulation stage NF10 and post-gastrulation stages TS3, TS5 and TS7 were compared. Both Smad2 and PSmad2 were detected in VC at NF10. Three independent experiments with protein samples prepared from different clutches of embryos at NF10 showed similar results, with the presence of both native and active forms of EcSmad2 in VC tissues.

B. Detection of Smad4:

NE does not differentiate into any adult tissue. Based on my Smad2 western analysis, both native and active forms of Smad2 are present in VC cells of the embryo.

Therefore it was necessary to look at the expression level of Smad4, the partner of PSmad2 in nuclear translocation of the protein complex (Kumar et al., 2001; Lee et al., 2001; Yeo and Whitman, 2001; Ross and Hill, 2008).

Although the *EcSmad4* gene has not been cloned, several commercial antibodies against Smad4 were available. One antibody was rabbit anti-Smad4 polyclonal antibody (Thermo Scientific; PA1-41292), which was raised against synthetic peptides corresponding to amino acid 186 – 199 and 509 – 523 of human Smad4. To my knowledge, this antibody has never been reported to be used for any frog samples, but based on sequence conservation of the immunogen it was predicted to cross react with *X. laevis*. There are two isoforms of Smad4, α and β, present in *X. laevis* (LeSueus and Graff, 1999; Masuyama et al., 1999; Howell et al., 1999; Hill, 2001). *X. laevis* Smad4 sequence revealed that Smad4β is larger than that of Smad4α. The predicted molecular weights for *X. laevis* Smad4α and Smad4β are 60.0 and 61.3 kDa, respectively.

Based on my experience with EcSmad2 and its sequence conservation with that of *X. laevis*, I used this antibody against *E. coqui* protein samples. The EcSmad4 western results are shown in Fig. 21. Anti-Smad4 polyclonal antibody detected two Smad4 isoforms. My western results, both temporal (Fig. 21A) and spatial (Fig. 21B), indicated the presence of two bands around 60 kDa. The bigger band (tentatively, Smad4β) showed a higher level of expression compared to the smaller isoform (tentatively Smad4α).

C. Detection of ActRIIA:

Activin RIIA is one of the important Nodal receptors, which plays essential role in the signaling pathway for endoderm-mesoderm specification in *X. laevis* (Reissman et

al., 2001; Yeo and Whitman 2001). Goat anti-human Activin RIIA antibody (R&D Systems, Inc.; AF340), raised against the recombinant human Activin RIIA (rhActivin RIIA) extracellular domain, is predicted to show less than 2% cross-reactivity with rhActivin RIIB or any other activin receptor. This antibody has been used successfully to detect the endogenous levels of the *X. laevis* homolog on a western blot (Martello et al., 2007). The predicted molecular weight of rhActivin RIIA is 58 kDa and that of the *X. laevis* homolog was 57.9 kDa (Xenbase: XB-GENE-865037), but Martello et al. (2007) did not provide any size information on their western. I made several attempts to use this antibody on westerns of *E. coqui* embryos, (Fig. 22). These always produced a prominent band around 50 kDa, instead of 60 kDa. In the absence of *EcActRIIA* sequence data, I decided not to proceed further.

IV. Cellular location of PSmad2, Smad4 and receptor ActRIIA in prospective DE vs. prospective NE:

Western blots with antibodies against Smad2, PSmad2, and Smad4 not only detected the presence of the respective proteins in all the stages of *E. coqui* development tested, they also indicated the presence of these TGF-β pathway components in both VC and NE. This data led me to investigate the cellular locations of these proteins. Given that PSmad2 was present in the VC of NF10 embryos (Fig. 20), it was important to see whether the active protein is nuclear or not. Immunostaining experiments were carried out on dissociated MZ and VC cells from NF10 embryos. For each of the experiments, separate sets of cells were treated with “no primary antibody, only secondary antibody” and “no

Fig. 19: Native and active forms of EcSmad2 in *E. coqui* development. Western blot of proteins isolated from *E. coqui* oocyte, NF10, TS3, TS5 and TS7. Whole embryos or oocytes were used to extract proteins. The amount of protein loaded in all the lanes was 75 μg. The top panel shows anti-Alpha Tubulin antibody staining, which served as a loading control. The middle panel represents expression of native EcSmad2 detected by purified mouse monoclonal anti-Smad2/3 (BD Transduction). The bottom panel shows the presence of active PSmad2 detected by rabbit polyclonal PSmad2 (Ser465/467) Antibody (Cell Signaling) in all stages tested. There was a general trend of increased band intensity as the development advanced. The experiment was performed three times with three independently prepared protein samples from different embryo clutches.

Fig. 19:

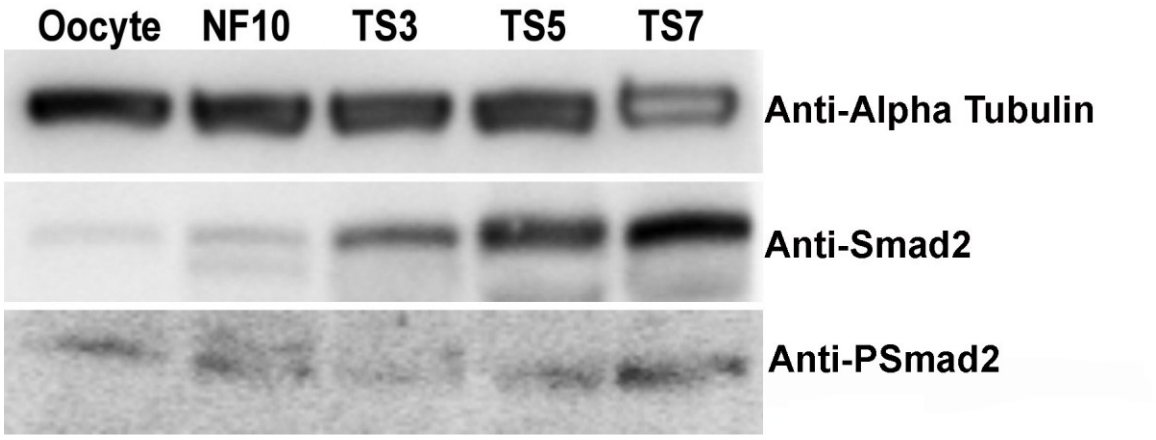
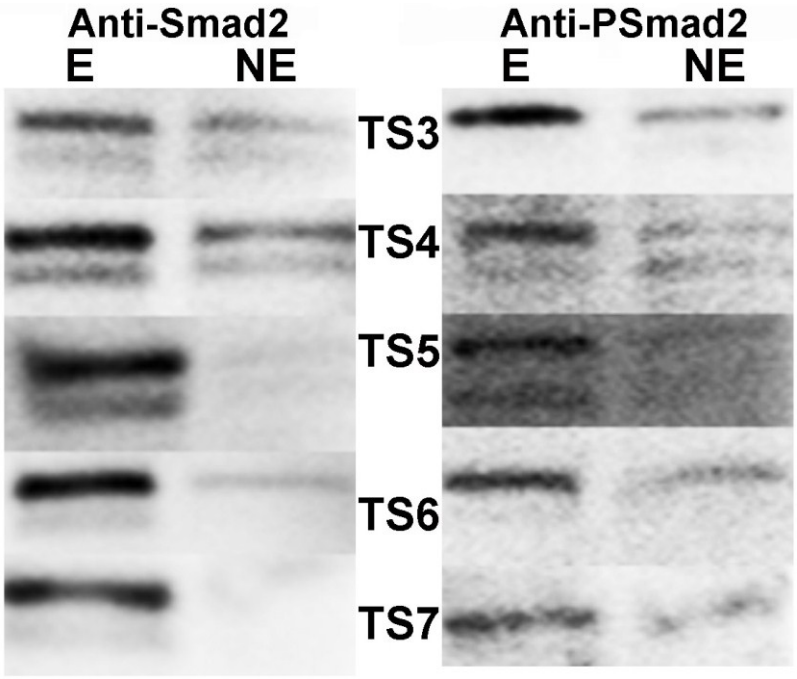

Oocyte
NF10
TS3
TS5
TS7
Anti-Alpha Tubulin
Anti-Smad2
Anti-PSmad2

Fig. 20: Spatial expression of Smad2 and PSmad2 in *E. coqui* development. (A) Western blot of dissected embryo (E) and NE tissues from post-gastrulation stages – TS3 to TS7. The left panel represents expression of native EcSmad2 detected by purified mouse monoclonal anti-Smad2/3 (BD Transduction). Right panel shows detection of active PSmad2 by rabbit polyclonal PSmad2 (Ser465/467) antibody (Cell Signaling). There are lower levels of Smad2 and PSmad2 in NE compared to E. **(B)** Western blot of dissected MZ and VC from gastrulation stage NF10 and E and NE from post-gastrulation stages – TS3, TS5 and TS7. The top panel shows the anti-Alpha Tubulin antibody staining serving as a loading control. Middle panel represents expression of native EcSmad2 detected by purified mouse monoclonal anti-Smad2/3 (BD Transduction). The bottom panel shows the presence of active PSmad2 being detected by rabbit polyclonal PSmad2 (Ser465/467) Antibody (Cell Signaling) in all stages tested. There was almost an equal amount of Smad2 and PSmad2 in MZ and VC from NF10. Spatial expression experiments were performed three times with three independently prepared protein samples from different embryo clutches. Amount of protein loaded in all the lanes was 100 μg.

Fig. 20:

(A)


Anti-Smad2
E
NE
Anti-PSmad2
E
NE
TS3
TS4
TS5
TS6
TS7

(B)

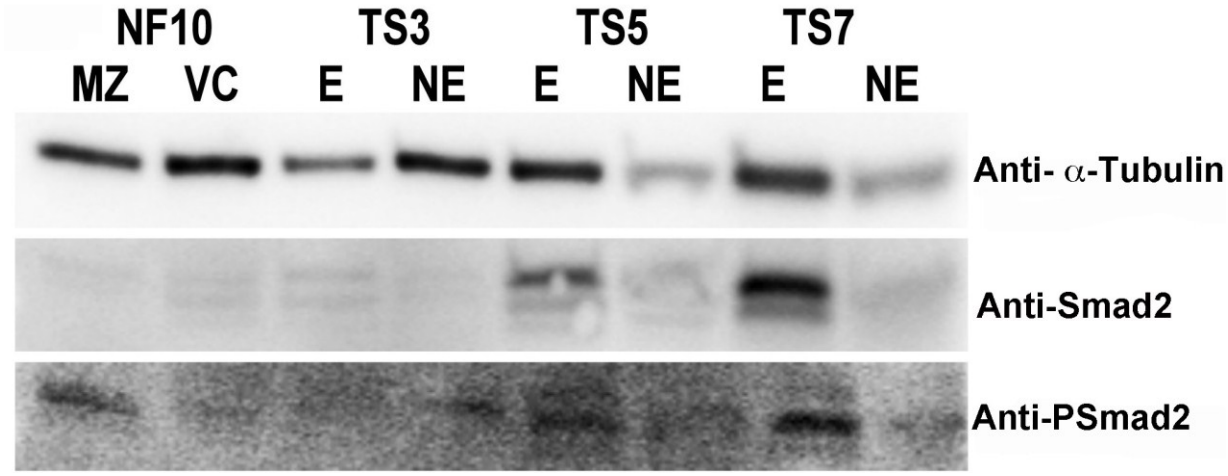

NF10
MZ
VC
TS3
E
NE
TS5
E
NE
TS7
E
NE
Anti- α-Tubulin
Anti-Smad2
Anti-PSmad2

Fig. 21: Western blot of EcSmad4. (A) Temporal expression of EcSmad4 isoforms in *E. coqui*. Top panel: α-Tubulin was used as a loading control for the experiment. **Bottom panel:** Two isoforms of EcSmad4 were detected (red arrow). 100 ug of whole protein preparations from 3 stages were run and probed with polyclonal rabbit anti-Smad4 antibody (Thermo Scientific). **(B) Spatial expression of EcSmad4 isoforms in *E. coqui*.** 100 μg of protein from NF10 whole embryos W, VC and MZ were run and probed with polyclonal rabbit anti-Smad4 antibody (Thermo Scientific). Two bands represent the two isoforms of the EcSmad4 protein (red arrow). Based on *X. laevis* sequence data, the band with higher molecular weight (MW) is tentatively Smad4β and the band with lower MW is tentatively Smad4α. For both temporal and spatial analysis purpose, each experiment was run three times with independently isolated protein extracts from different embryo clutches.

Fig. 21:

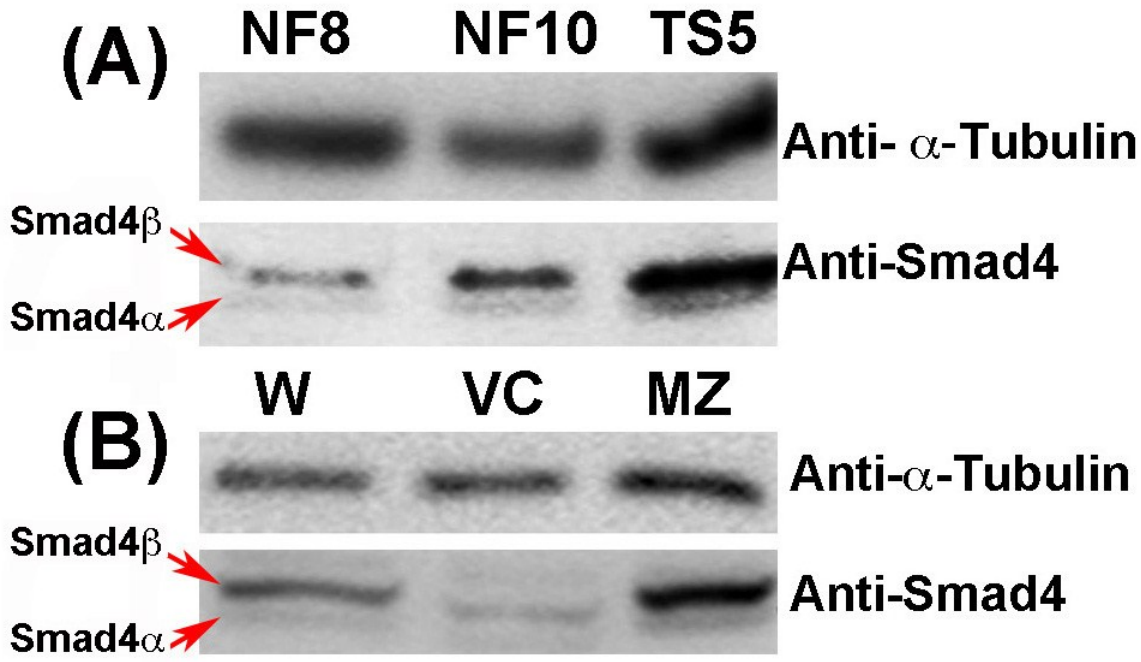
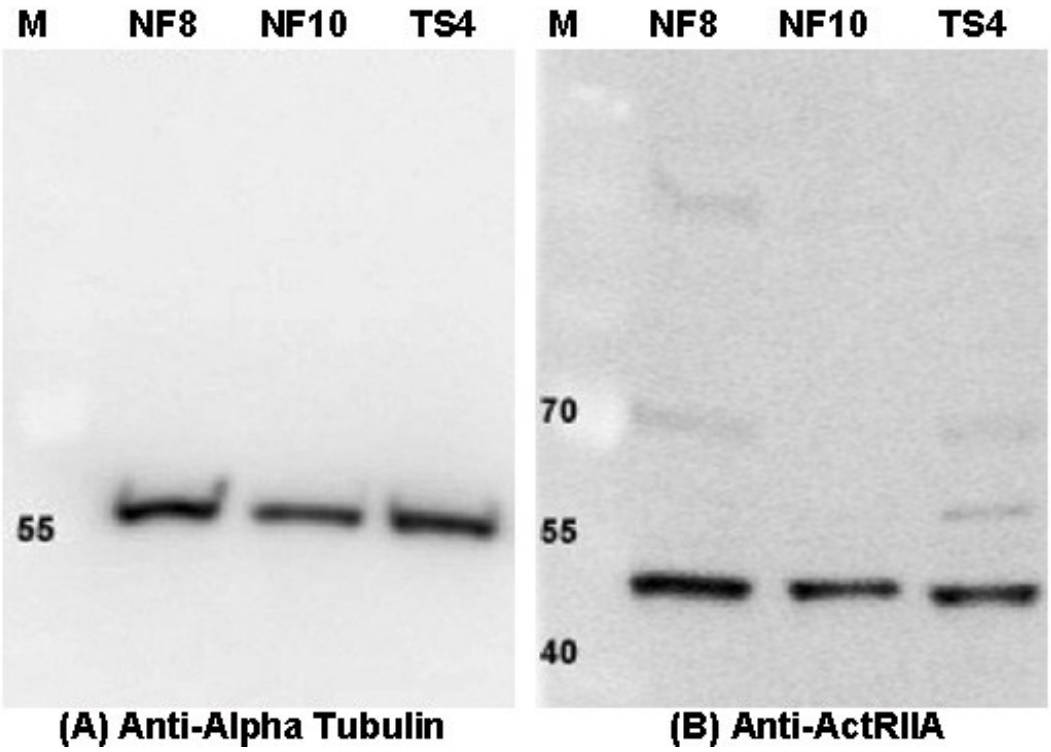

(A)
NF8
NF10
TS5
Anti- α-Tubulin
Smad4β
Smad4α
Anti-Smad4
(B)
W
VC
MZ
Anti-α-Tubulin
Smad4β
Smad4α
Anti-Smad4

Fig. 22: Western blot of ActRIIA. An Anti -ActRIIA antibody (Goat Anti-human Activin RIIA Antibody; R&D Systems) was tested on Western blots. (a) Alpha Tubulin band at 56 kDa serving as a loading control for the experiment. (b) Although the predicted size of ActRIIA in human and *X. laevis* is 58 – 60 kDa, this anti-ActRIIA antibody detected a band around 50 kDa in *E. coqui*. 100 ug of whole protein preparations from three stages were run and probed with the antibody. Each experiment was run three times with independently isolated protein extracts from different embryo clutches. (M is the protein marker lane)

Fig. 22:


M
NF8
NF10
TS4
55
(A) Anti-Alpha Tubulin
M
NF8
NF10
TS4
70
55
40
(B) Anti-ActRIIA

secondary antibody, only primary antibody", which served as negative controls with no signal being detected. Each experiment was repeated at least twice. In case of VC or NE, there were many anuclear cells, so only cells with single or multiple nuclei were scored.

A. Immunostaining for PSmad2:

During late blastula and early gastrula, 100% of the prospective DE cells from MZ show PSmad2 nuclear localization, but only 12% of VC cells have nuclear localization (Table 8, Fig. 23). The rest had a scattered cytoplasmic signal.

B. Immunostaining for Smad4:

Smad4 nuclear localization was evident in cells of both MZ and VC. About 50% of VC cells were positive for nuclear localization of Smad4, whereas 100% of the MZ cells were positive (Table 8, Fig. 24). There are two isoforms of Smad4, α and β, reported in *X. laevis*. The antibody used in immunostaining experiments was the same one used in western blotting, where it detected two bands at predicted molecular sizes. There was no way to know whether the nuclear Smad4 was α, or β.

C. Immunostaining for ActRIIA:

ActRIIA, a Nodal receptor, was found only in 17% of the VC cells and 20% of MZ cells (Table 8, Fig. 25). In both types of cells, signal was observed on the surface of the cells. This staining must be interpreted cautiously since the antibody may not have detected a protein of the correct molecular weight on a western blot (Fig. 22).

D. A second approach to detect nuclear accumulation of PSmad2:

So far, my immunostaining data strongly suggests that PSmad2 is absent from the nuclei of VC cells, but almost half of the population was stained positive for Smad4. Although Smad2-independent nuclear translocation of Smad4 has been reported (Bai *et. al*. 2002), absence of PSmad2 in the nuclei of VC cells favors the idea of lack of Nodal signaling in them. This lack could account for the lack of differentiation of the NE into endodermal tissue. As an alternative approach to further confirm PSmad2 immunostaining data, I injected *EGFP-Smad2* capped RNA into the vegetal half of the 8-32 cell stage *E. coqui* embryos.

The pCS105-EGFP-Smad2 construct was generated using *X. laevis Smad2* sequence and has been used in experiments involving *X. laevis* (Skirkanich et al., 2011). Six *E. coqui* embryos were kept as un-injected controls and six more embryos were injected with only DEPC-H_2O. For *EGFP-Smad2* mRNA, first 900 pg and 100 pg in 9.2 nl were injected. Embryos were cultured until NF10, when they were dissected into MZ and VC tissues. VC cells were dissociated, fixed and stained with anti-Green fluorescent protein antibody (Invitrogen^{TM}/ Molecular Probes; A11122) to detect nuclear localization of the EGFP tagged PSmad2 proteins. Cells were co-stained with DAPI to locate the nuclei. 100% of the VC cells showed nuclear location of EGFP overlapping with DAPI (Fig. 26A), but the control embryos injected with DEPC-H_2O failed to grow beyond neural tube formation (TS3). Development of uninjected embryos was normal.

In a further trial, 10 pg in 9.2 nl was injected as a tracer level. Embryos survived longer, but were developmentally looked behind those injected with DEPC-H_2O. When embryos injected with DEPC-H_2O were at TS11, experimental embryos, injected with 10

pg *EGFP-Smad2* mRNA, looked like TS8 (Fig. 26B). This tracer level of injected mRNA led to a nuclear EGFP signal in only 8% of the VC cells dissociated at NF10. One of those VC cells is shown in Fig. 26C-E.

V. Determination of transcriptional status of VC cells:

The presence of both PSmad2 and Smad4 in VC cells but no differentiation to adult tissues, leads to the question whether the VC cells are transcriptionally active. Depending on the needs of a cell, RNAPII activity is determined by the phosphorylation state of the CTD heptapeptide YSPTSPS (Corden, 1990; Phatnani and Greenleaf, 2006; Venkatarama et al., 2010). According to the model shown in Fig. 4, there are four phosphorylation states. The promoter regions are occupied normally by the unphosphorylated RNAPII. Ser 5 phosphorylation initiates transcription and Ser 2 phosphorylation leads to elongation.

To determine the transcriptional status of the prospective NE cells, VC and MZ cells were dissociated from embryos at NF8 and NF10 and NE cells were dissociated from embryos after gastrulation. Cells were stained with antibodies against RNAPII CTD-Ser 2 (H5), CTD-Ser 5 (H14) and unphosphorylated RNAPII large subunit (8WG16) as a positive control. For each of the experiments, separate sets of cells were treated with “no primary antibody, only secondary antibody” and “no secondary antibody, only primary antibody”, which served as negative controls. No signals were detected. Each experiment was repeated at least twice, and each time, a minimum of 50 cells was counted. In the case of VC or NE, anuclear cells were not considered.

Table 8: Summary of immunostaining results with anti-PSmad2, anti-Smad4 and anti-ActRIIA antibodies.

	MZ Cells	VC Cells
Nuclear Psmad2	100 % (N = 80)	12% (N = 150)
Nuclear Smad4	100% (N = 100)	50% (N = 150)
ActRIIA	20% (N = 100)	17% (N = 100)

Fig. 23: Nuclear localization of PSmad2 in MZ cells but not VC cells in *E. coqui* NF10 embryos. Dissociated cells isolated from dissected MZ and VC tissues were stained with DAPI (**b**, **e**, **h**, and **k**) to show the position of nuclei. The same cells were stained with rabbit polyclonal anti-PSmad2 (Ser465/467) antibody (Cell Signaling, #3101) to detect nuclear localization of PSmad2 (**c**, **f**, **i**, and **l**). **a**, **d**, **g**, and **j** represent bright field (BF) images of the same fields. All MZ cells showed nuclear signal for PSmad2 overlapping their respective DAPI signals, whereas most VC cell nuclei were PSmad2 negative. Scale bar represents 1 μm.

Fig. 23:

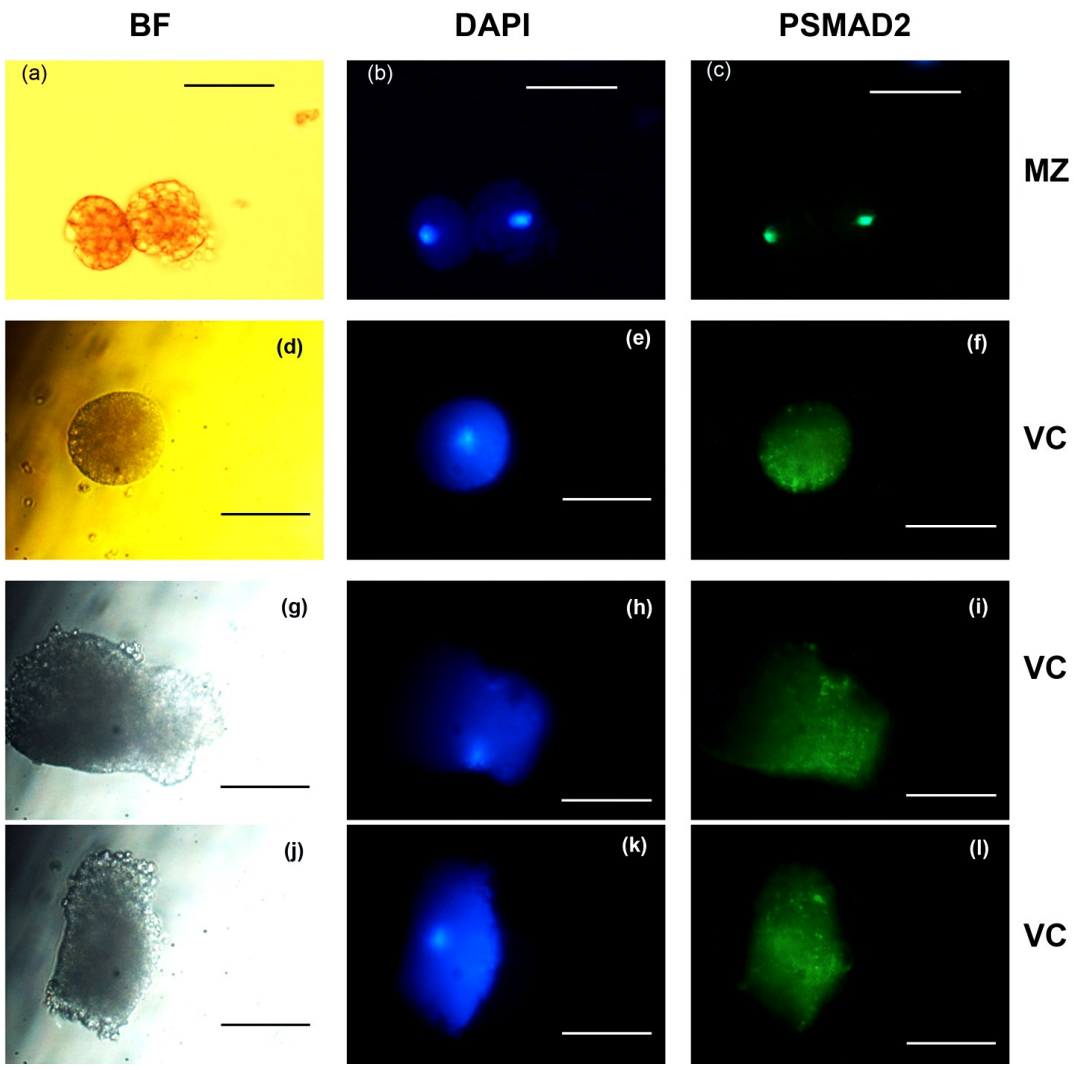


Fig. 24: Nuclear Smad4 signal in 50% of VC cells, but 100% of MZ cells at gastrulation. Dissociated MZ and VC cells dissected from NF10 embryos were stained with polyclonal rabbit anti-Smad4 antibody (Thermo Scientific, #PA1-41292) (**c**, **f**, **i**, **l** and **o**) and counter-stained with DAPI (**b**, **e**, **h**, **k** and **n**). **a**, **d**, **g**, **j** and **m** represents bright field (BF) images of the same fields. Scale bar represents 5 μm.

Fig. 24:

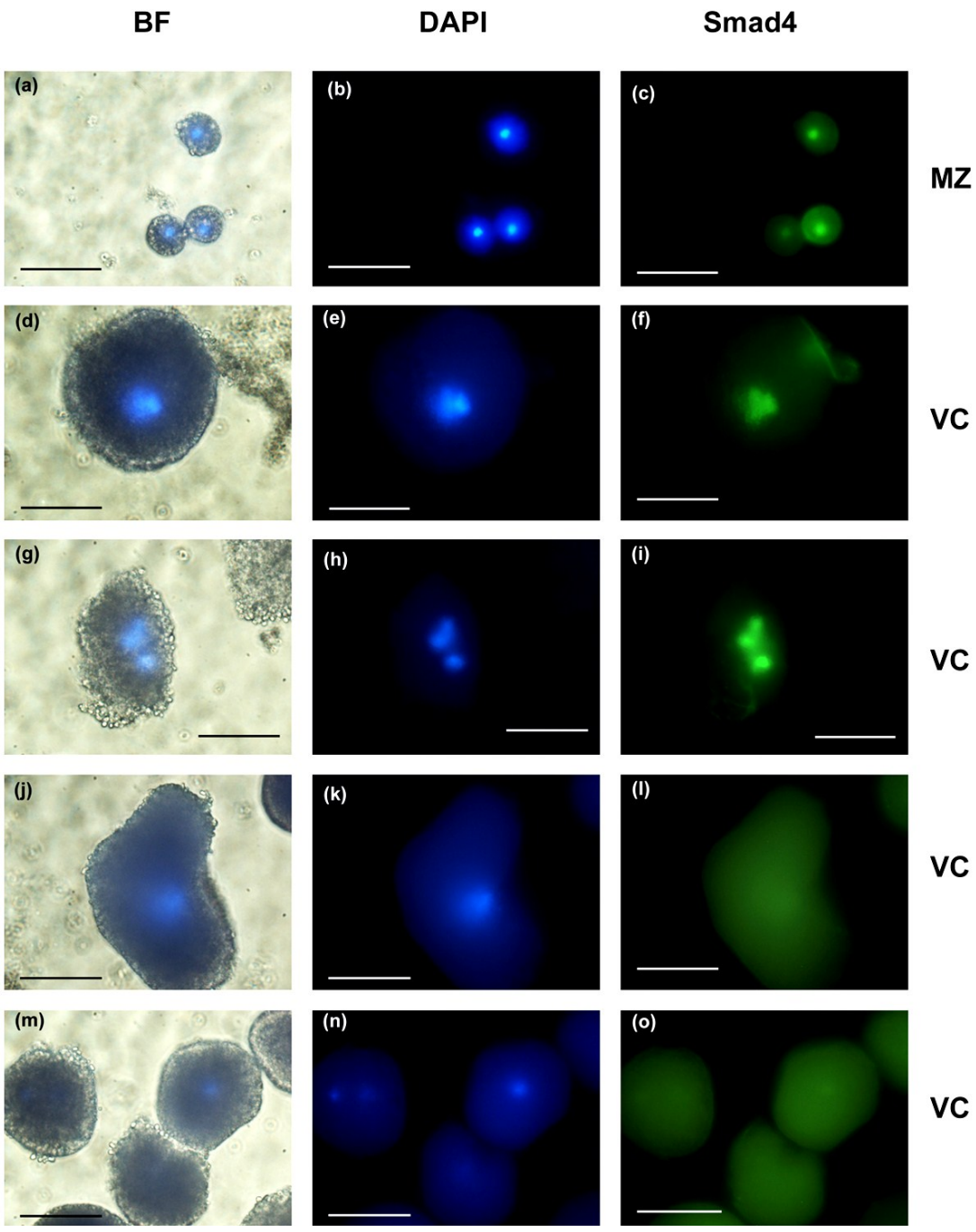


Fig. 25: Positive ActRIIA signal in 17% of the VC cells and 20% of MZ cells at gastrulation. Dissociated MZ and VC cells dissected from NF10 embryo were stained with polyclonal goat anti-human Activin RIIA antibody (R&D, #AF340) (**c**, **f**, **i**, and **l**) and counter-stained with DAPI (**b**, **e**, **h**, and **k**). **a**, **d**, **g**, and **j** represents bright field (BF) images of the same fields. Scale bar represents 5 μm.

Fig. 25:

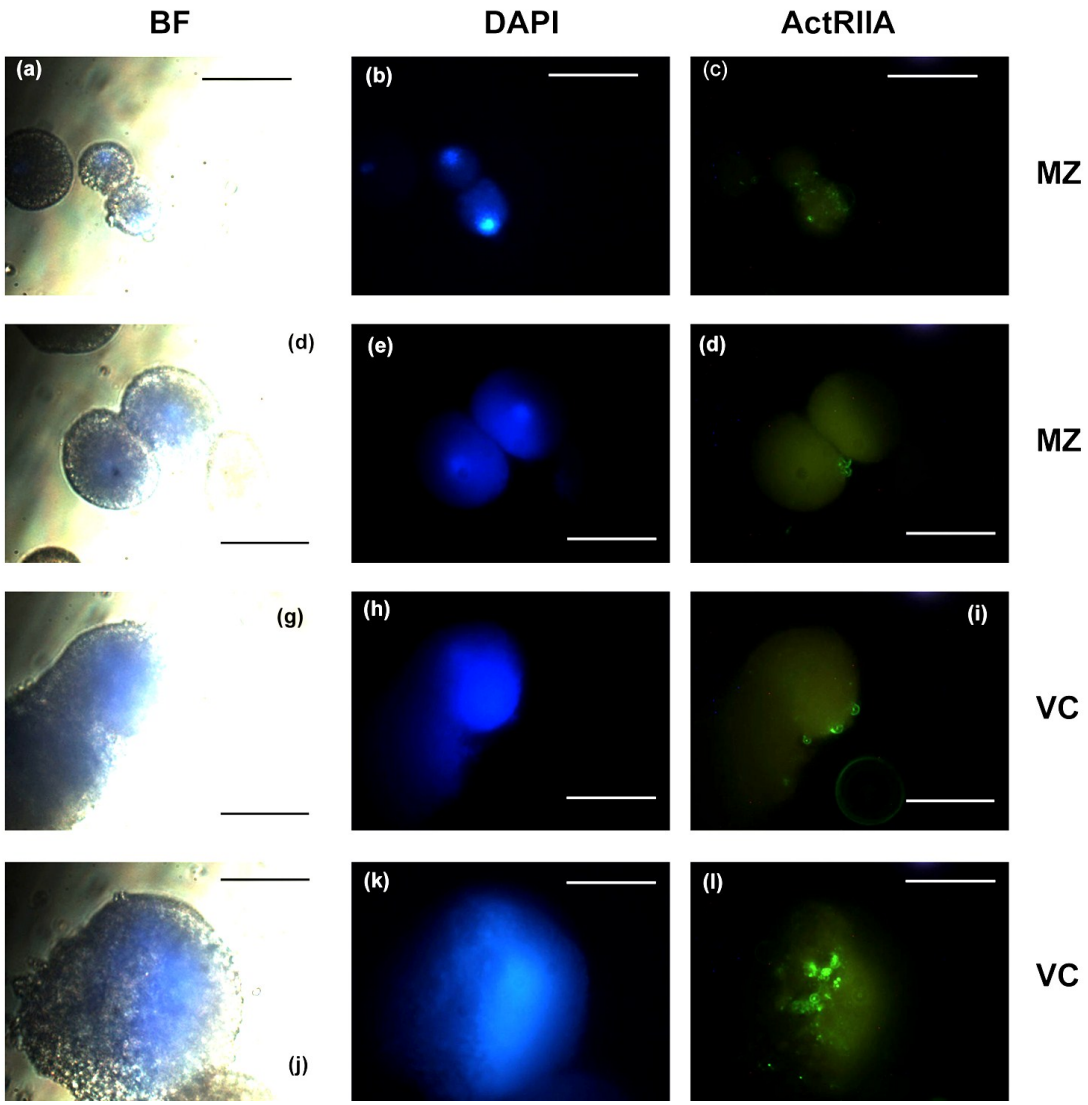
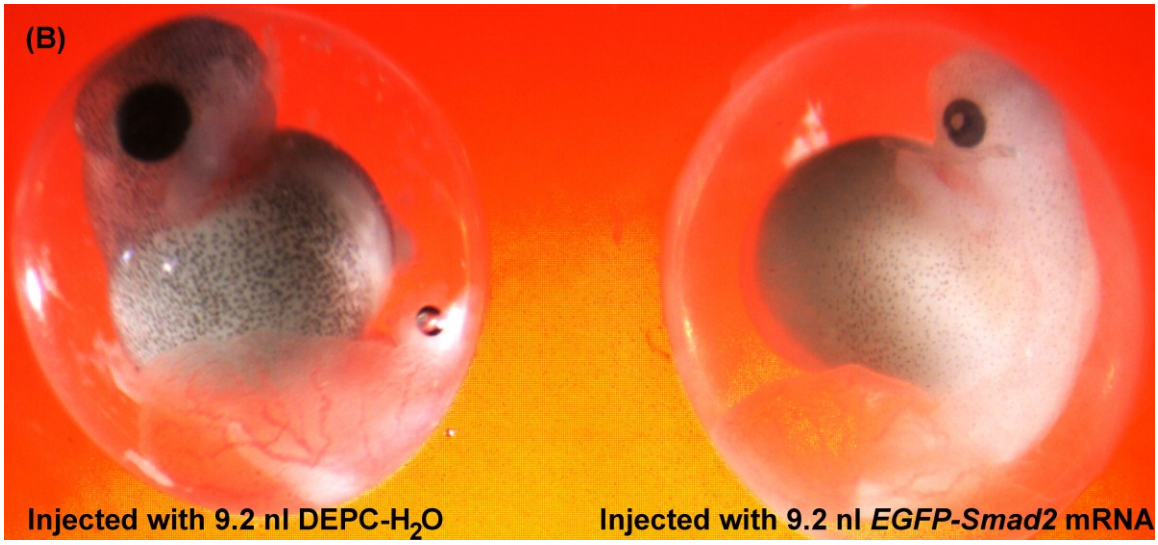


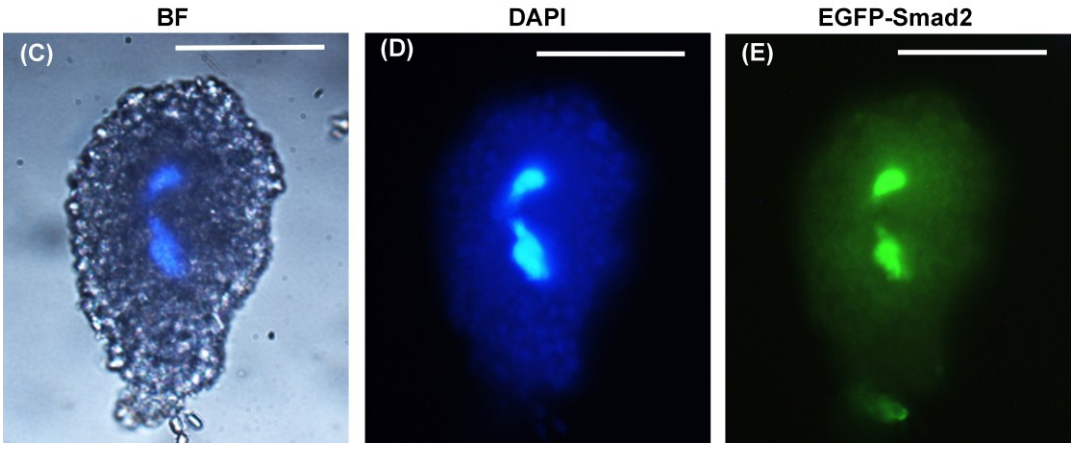
Fig. 26: Nuclear localization of EGFP-Smad2 in VC cells at NF10. **(A)** *E. coqui* embryos were injected with 900 and 100 pg *EGFP-Smad2* capped RNA at 8-32 cell stage. Although, 100% of the VC cells dissociated from NF10 embryos showed nuclear accumulation of the EGFP signal, none of the control embryos grew beyond neural tube formation. **(B)** *E. coqui* embryos were injected with 10 pg *EGFP-Smad2* capped RNA at 8-32 cell stage. Embryos were grown until NF10, when they were dissected and dissociated. Six uninjected embryos and six embryos injected with 9.2 nl DECP-H_2O were used as control. When embryos injected with DECP-H_2O reached TS11, embryos injected with 10 pg *EGFP-Smad2* RNA were developmentally behind, at around stage TS8 and did not survive beyond this point. Dissociated VC cells from NF10 injected embryos were fixed and stained. (**C)** represents a bright field (BF) image of one of the dissociated VC cell isolated from dissected VC tissues at NF10. VC cells were stained with DAPI to show the position of nuclei (**D**) to show the position of nuclei. The same cells were stained with anti-Green fluorescent protein antibody (Invitrogen^{TM}/ Molecular Probes; A11122) to detect nuclear localization of EGFP-Smad2. (**E**) represents one of 8% of VC cell nuclei, which were EGFP positive. No MZ cells showed nuclear signal for EGFP-Smad2. Scale bar represents 5 µm.

Fig. 26:

(A)

Volume injected	Concentration injected	Nuclear EGFP signal
9.2 nl	900 pg	100% (N = 150)
9.2 nl	100 pg	100% (N = 150)
9.2 nl	10 pg	8% (N = 200)





At NF8, both MZ and VC cells were positive for unphosphorylated RNA Pol II CTD (8WG16), but failed to show any signal for transcription initiation (H14) or transcription elongation (H5) (Table. 9, Fig. 27). At NF10, MZ and VC cells both showed positive signals for H14 antibody indicating the occurrence of transcription initiation. MZ cells, but not VC cells, were positive for H5, which marked transcription elongation (Table. 9, Fig. 28). This indicates that MZ cells were transcriptionally active, whereas VC cells were not.

At advanced post gastrulation stages, the embryo was already formed. It was hard to dissociate individual cells from the rigid embryo tissues. For all post-gastrulation stages, undissociated embryonic tissues were stained and were positive for all three antibodies used (not shown). NE cells dissociated from TS3 embryos were positive for 8WG16, but showed very faint or no signal for either H14 or H5, indicating a transcriptionally repressed state (Fig. 29). Similar signaling data was obtained for NE cells dissociated from all advanced stages including TS8 (Fig. 30), TS11 (Fig. 31) and TS14 (Fig. 32). These results suggest a global transcriptional suppression for nutritional endodermal cell throughout development. A summary of the 8WG16, H14 and H5 immunostaining results on cells from all the stages are represented in Table. 9.

X. laevis fertilization is followed by 12 rapid, synchronous cleavage divisions, which leads to the formation of blastula (Heasman, 2006). After the twelfth division, the cell cycle shows asynchronous divisions and the appearance of gap phases. This change in the cell cycle is referred to as MBT, a critical event, which marks the onset of embryonic gene expression (Gerhart, 1980; Heasman, 2006). In *E. coqui*, MZ cells were transcriptionally active at gastrulation (NF10) (Fig. 28) and remained active post-

gastrulation at TS3, TS8, TS11 and TS14. This was the first evidence for MBT in *E. coqui* and suggested that it occurred between NF8 and NF10. Unlike MZ, VC cells did not become transcriptionally active, before or after gastrulation.

Fig. 27: Neither transcription initiation nor elongation occurred in MZ and VC at NF8. Nuclei of both MZ and VC cells dissociated from NF8 were positive for unphosphorylated RNA Pol II large subunit. MZ and VC cells were stained with monoclonal mouse 8WG16 antibody (Covance, #MMS-126R) (**c**, **f**) and counter-stained with DAPI (**b**, **e**). **a** and **d** represent bright field (BF) images of the same fields. 8WG16 recognized unphosphorylated RNA Pol II large subunit and served as a positive control. MZ and VC cells were stained with monoclonal mouse H14 antibody (Covance, #MMS-134R) (**i**, **l**) and counter-stained with DAPI (**h**, **k**). **g** and **j** represent BF images of the same fields. H14 recognized phosphoserine-5 version of RNA Pol II, which indicated transcription initiation. Neither MZ nor VC nuclei were stained with H14. MZ and VC cells were stained with monoclonal mouse H5 antibody (Covance, #MMS-129R) (**o**, **r**) and counter-stained with DAPI (**n**, **q**). **m** and **p** represent bright field (BF) images of the same fields. H5 recognized phosphoserine-2 version of RNA Pol II, which is the mark for transcriptional elongation. Neither VC nor MZ nuclei stained with H5. Scale bar represents 5 μm.

Fig. 27:

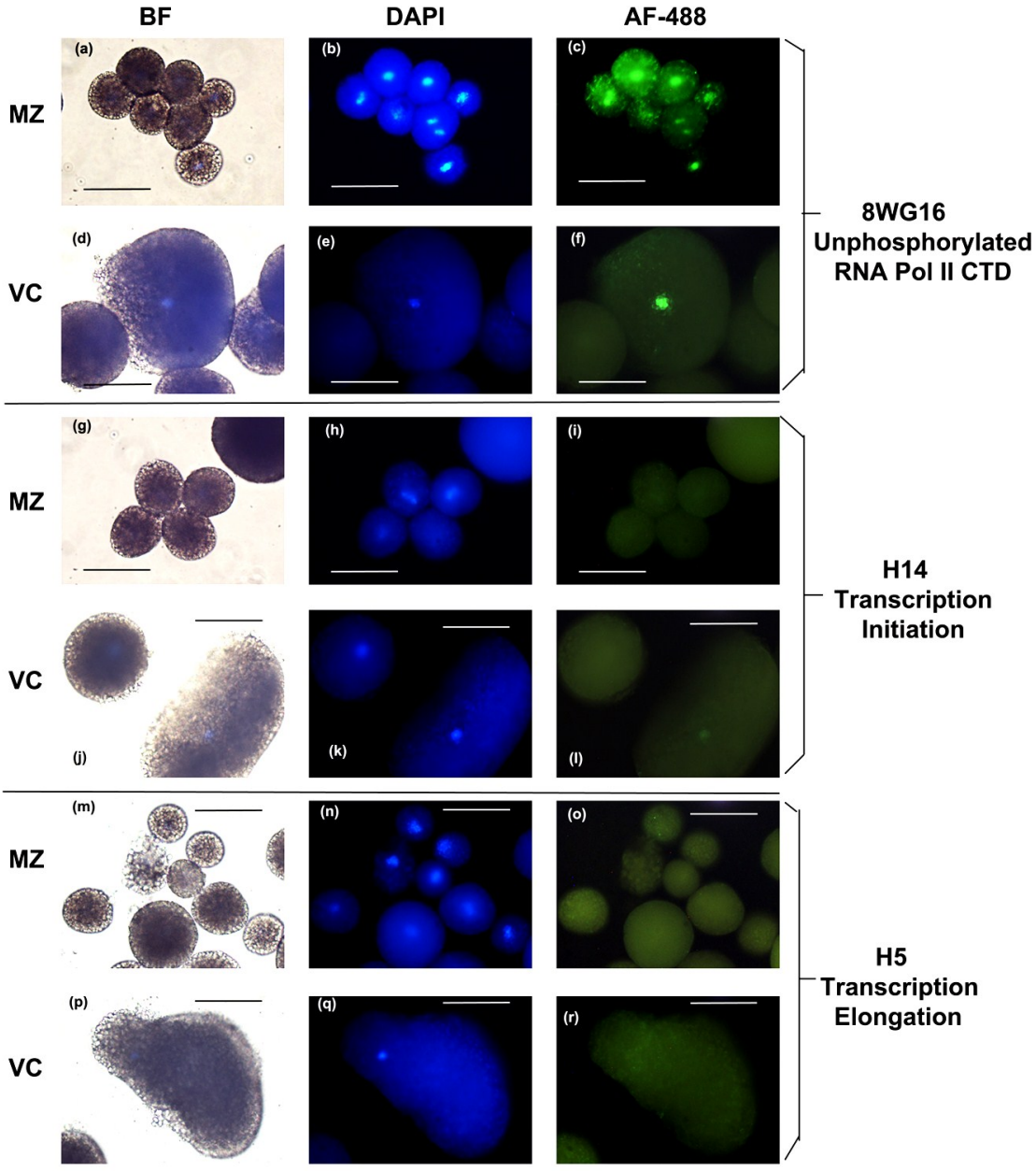

BF
DAPI
AF-488
MZ
VC
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
(m)
(n)
(o)
(p)
(q)
(r)
8WG16
Unphosphorylated
RNA Pol II CTD
H14
Transcription
Initiation
H5
Transcription
Elongation

Fig. 28: Both MZ and VC initiated transcription at NF10, but transcription elongation occurred only in MZ. MZ and VC cells dissociated from NF10 embryos were stained with monoclonal mouse 8WG16 antibody (Covance, #MMS-126R) (**c**, **f**) and counter-stained with DAPI (**b**, **e**). **a** and **d** represents bright field (BF) images of the same fields. Nuclei of both cell types were positive for 8WG16. MZ and VC cells were stained with monoclonal mouse H14 antibody (Covance, #MMS-134R) (**i**, **l**) and counter-stained with DAPI (**h**, **k**). **g** and **j** represents bright field (BF) images of the same fields. Nuclei of both MZ and VC cells were positively stained for H14 antibody. Both types of cells were stained with monoclonal mouse H5 antibody (Covance, #MMS-129R) (**o**, **r**) and counter-stained with DAPI (**n**, **q**). **m** and **p** represents bright field (BF) images of the same fields. Only MZ, and not VC nuclei were positive for H5. Scale bar represents 5 μm.

Fig. 28:

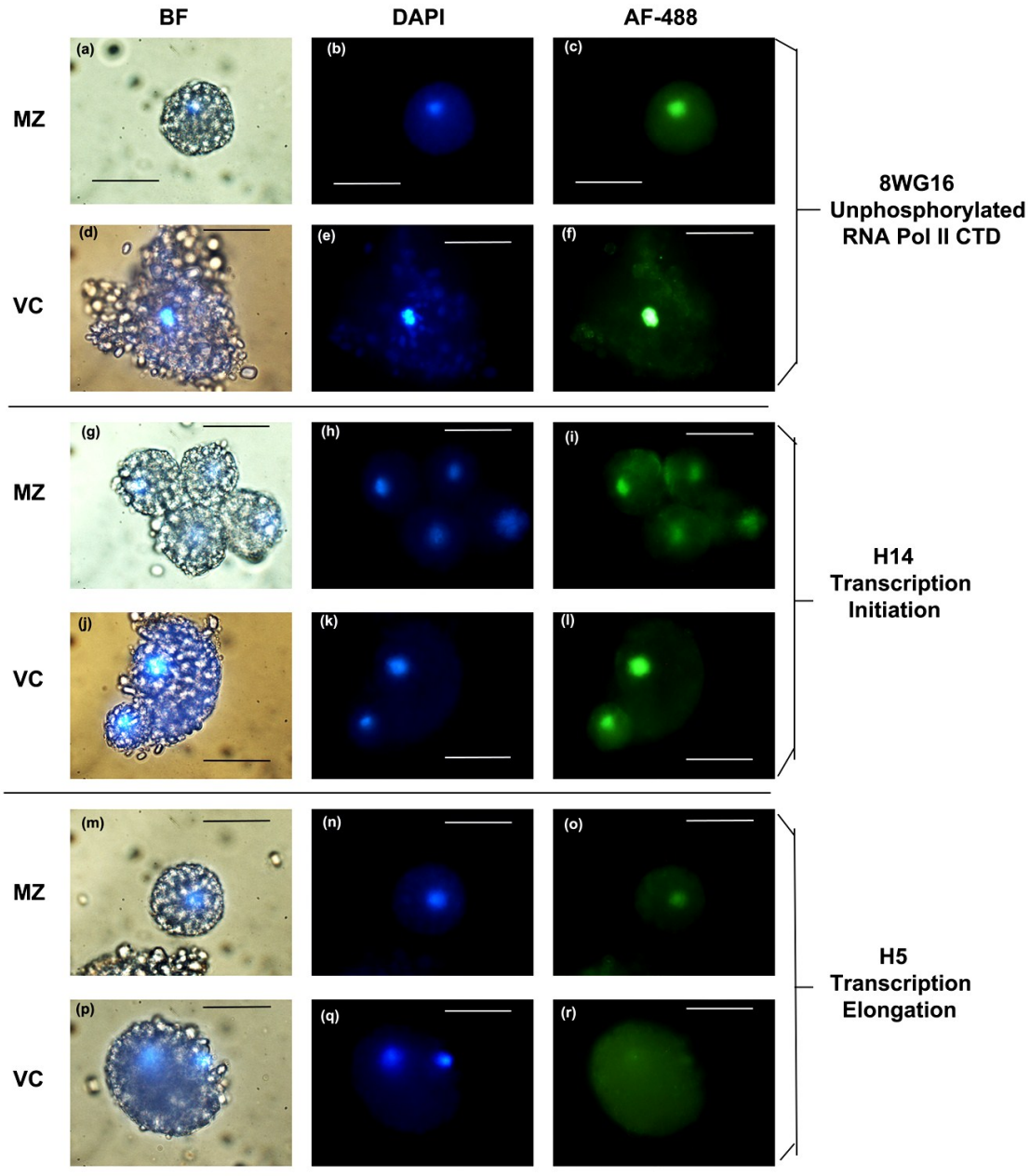

BF
DAPI
AF-488
MZ
VC
(a)
(b)
(c)
(d)
(e)
(f)
8WG16
Unphosphorylated
RNA Pol II CTD
MZ
VC
(g)
(h)
(i)
(j)
(k)
(l)
H14
Transcription
Initiation
MZ
VC
(m)
(n)
(o)
(p)
(q)
(r)
H5
Transcription
Elongation

Fig. 29: NE cells from TS3 were positive for 8WG16, but showed weak or no signal for initiation or elongation. NE cells dissociated from TS3 embryos were stained with 8WG16 (**c**) H14 (**f**) and H5 (**i**) antibodies and counter-stained with DAPI (**b, e**, and **h**). **a, d** and **g** represent bright field (BF) images of the same fields. Nuclei of NE cells were positive for 8WG16. Nuclear staining for both transcription initiation (H14) and elongation (H5) were either very weak or not present. Scale bar represents 5 μm.

Fig. 30: NE cells from TS8 were positive for 8WG16, but showed weak or no signal for initiation or elongation. NE cells dissociated from TS8 embryos were stained with 8WG16 (**c**) H14 (**f**) and H5 (**i**) antibodies and counter-stained with DAPI (**b, e** and **h**). **a, d** and **g** represent bright field (BF) images of the same fields. Nuclei of NE cells were positive for 8WG16. Nuclear staining for both transcription initiation (H14) and elongation (H5) were either very weak or not present. Scale bar represents 5 μm.

Fig. 29:

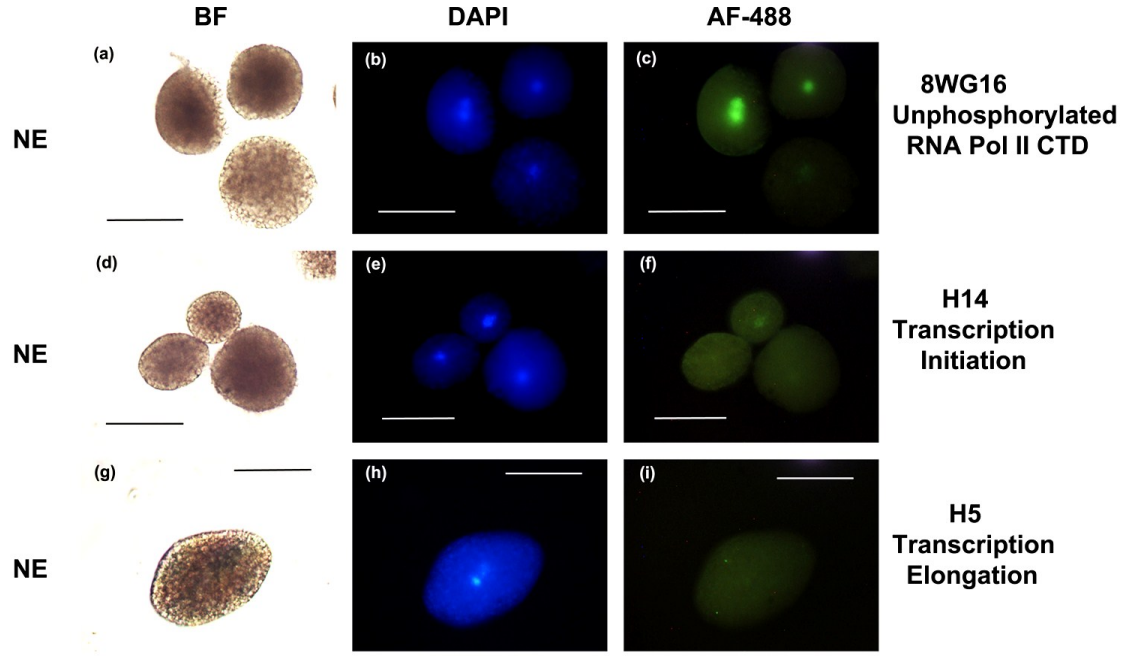

BF
DAPI
AF-488
(a)
(b)
(c)
NE
8WG16
Unphosphorylated
RNA Pol II CTD
(d)
(e)
(f)
NE
H14
Transcription
Initiation
(g)
(h)
(i)
NE
H5
Transcription
Elongation

Fig. 30:

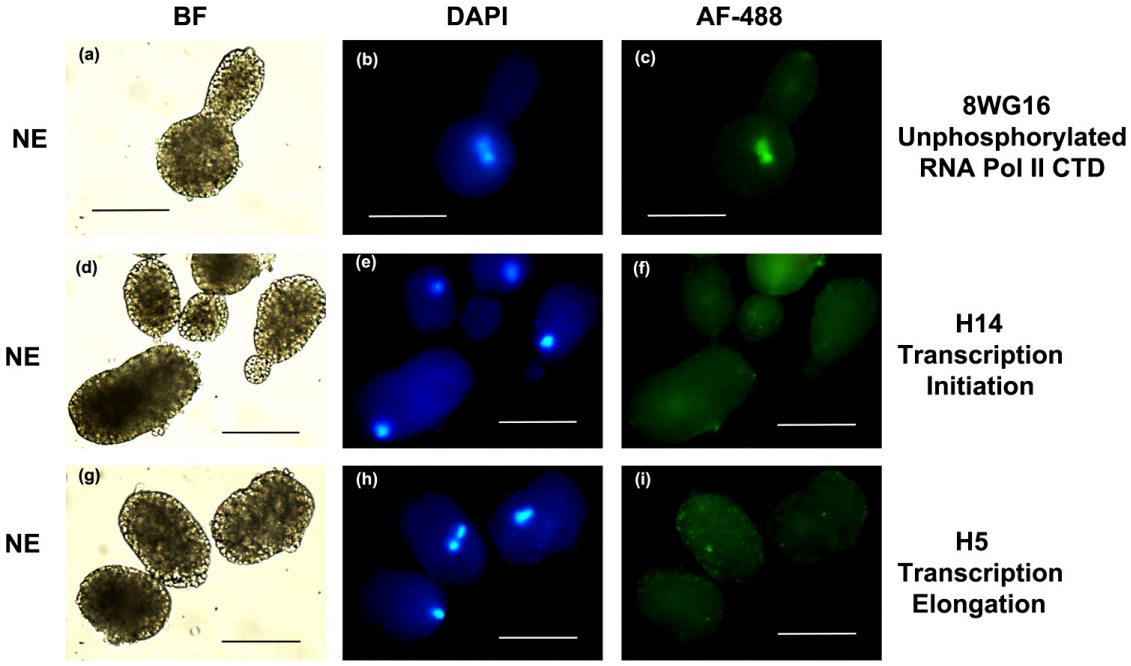

BF
DAPI
AF-488
(a)
(b)
(c)
NE
8WG16
Unphosphorylated
RNA Pol II CTD
(d)
(e)
(f)
NE
H14
Transcription
Initiation
(g)
(h)
(i)
NE
H5
Transcription
Elongation

Fig. 31: TS11 NE cells were positive for 8WG16, but showed weak or no signal for initiation or elongation. NE cells dissociated from TS11 embryos were stained with 8WG16 (**c**) H14 (**f**) and H5 (**i**) antibodies and counter-stained with DAPI (**b, e** and **h**). **a, d** and **g** represents bright field (BF) images of the same fields. Nuclei of NE cells were positive for 8WG16. Nuclear staining for both transcription initiation (H14) and elongation (H5) were either very weak or not present. Scale bar represents 5 μm.

Fig. 32: TS14 NE cells were positive for 8WG16, but weak or no signal for initiation or elongation. NE cells dissociated from TS14 embryos were stained with 8WG16 (**c**) H14 (**f**) and H5 (**i**) antibodies and counter-stained with DAPI (**b, e** and **h**). **a, d** and **g** represents bright field (BF) images of the same fields. Nuclei of NE cells were positive for 8WG16. Nuclear staining for both transcription initiation (H14) and elongation (H5) were either very weak or not present. Scale bar represents 5 μm.

Fig. 31:

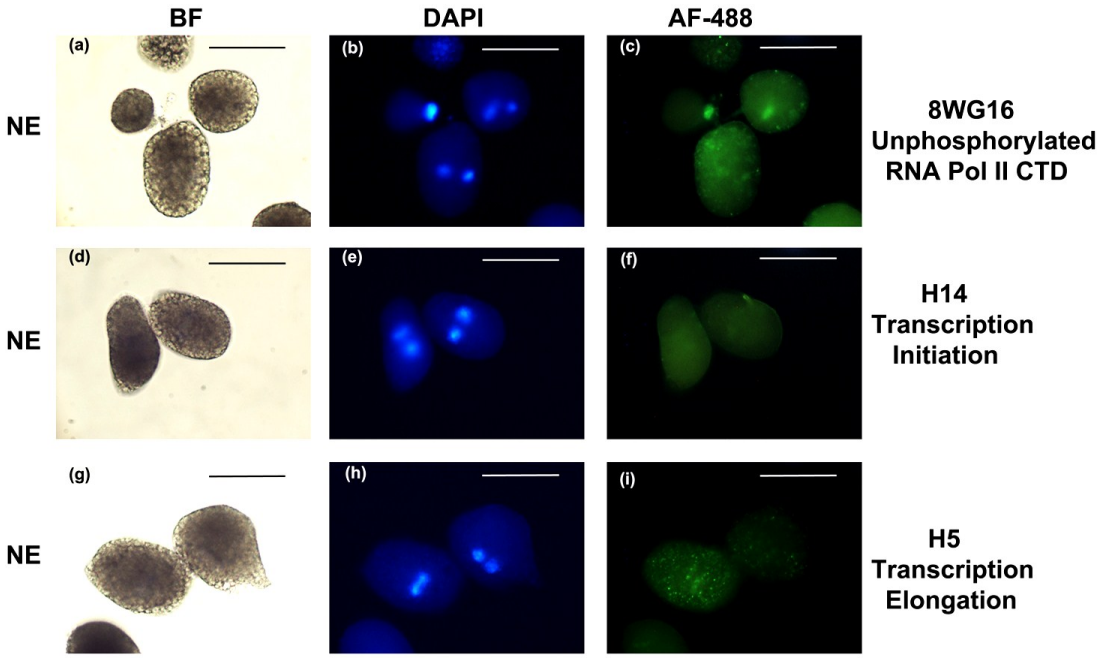

BF
DAPI
AF-488
NE
(a)
(b)
(c)
8WG16
Unphosphorylated
RNA Pol II CTD
NE
(d)
(e)
(f)
H14
Transcription
Initiation
NE
(g)
(h)
(i)
H5
Transcription
Elongation

Fig. 32:

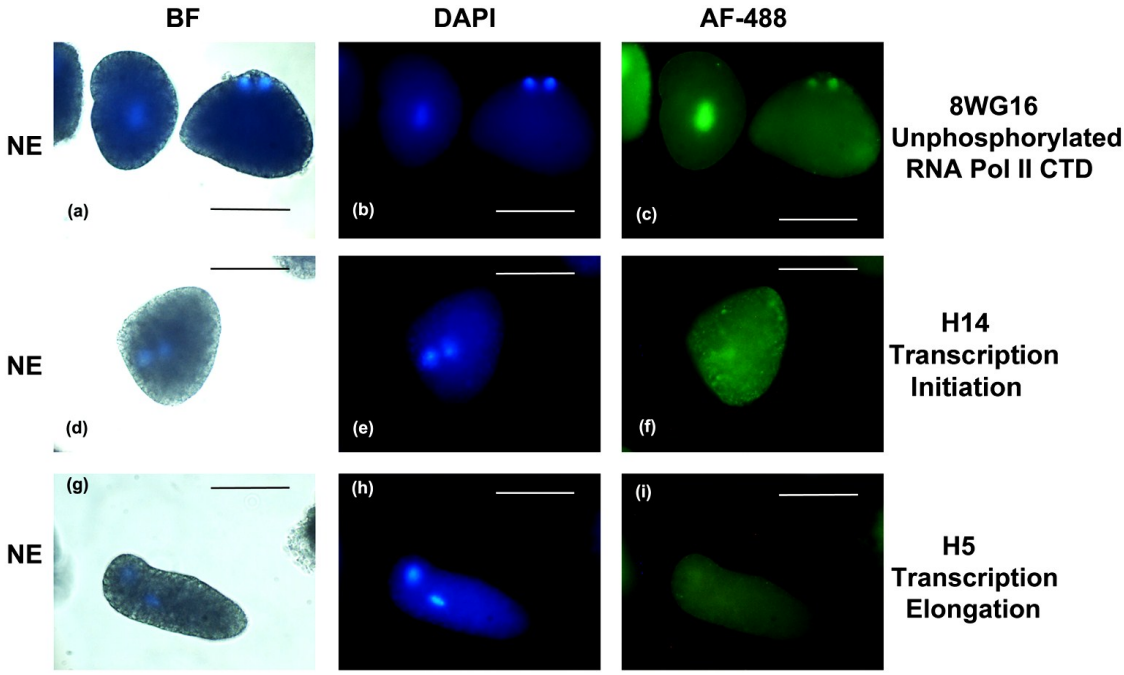

BF
DAPI
AF-488
NE
(a)
(b)
(c)
8WG16
Unphosphorylated
RNA Pol II CTD
NE
(d)
(e)
(f)
H14
Transcription
Initiation
NE
(g)
(h)
(i)
H5
Transcription
Elongation

Table 9. Summary of imunostaining with 8WG16, H14 and H5 on MZ and VC cells from NF8 and NF10, and NE cells from TS3, TS8, TS11 and TS14.

		8WG16 nuclear signal				H14 nuclear signal				H5 nuclear signal			
Stage	Tissue	No. of cells	No	Faint	Strong	No. of cells	No	Faint	Strong	No. of cells	No	Faint	Strong
NF8	MZ	100	2	6	92	100	100	0	0	100	100	0	0
	VC	200	10	7	183	150	150	0	0	150	150	0	0
NF10	MZ	80	0	3	77	150	0	6	144	150	0	2	148
	VC	200	7	0	193	200	187	13	0	200	195	5	0
TS3	NE	200	2	18	180	200	194	6	0	200	191	8	0
TS8	NE	200	6	8	186	200	198	2	0	200	195	5	0
TS11	NE	200	2	2	196	200	189	11	0	200	195	5	0
TS14	NE	200	0	2	198	200	185	15	0	200	186	14	0

CHAPTER FOUR: DISCUSSION

Activin and Nodal-related proteins of the TGF-β family of growth factors, which initiate the Nodal-signaling pathway, play central roles in endoderm/mesoderm specification in mouse, *X. laevis*, chick and zebrafish (Whitman, 2001; Shen 2007). In *X. laevis* embryos, maternal *VegT* and *Vg1* RNAs, localized in the vegetal blastomeres, regulate expression of Nodal ligands (Joseph and Melton, 1998; Heasman, 2006; White and Heasman, 2008). The whole vegetal half of *X. laevis* embryo initiates Nodal-signaling under the influence of maternal *VegT* expression (Zhang et al., 1998; Clements et al., 1999; Kofron et al., 1999; Xanthos et al., 2001), and these cells became committed to form the definitive endoderm. Maternal VegT also regulates expression of the endoderm specific transcription factor *Xsox17a* directly (Howard et al., 2007). VegT induced expression of *Xsox17a* at the midblastula transition is then maintained by Nodal signals (Engleka et al., 2001).

Based on expression profiles of *EcVegT* and *EcVg1* RNA (Beckham et al., 2003; Pérez et al., 2007), it was hypothesized for *E. coqui* that the early embryonic VC cells, which form the NE, would be devoid of the molecular determinants required for endoderm/mesoderm specification (Ninomiya et al., 2001). Absence of *EcVegT* RNA and mesoderm inducing activity in the VC of *E. coqui* embryos (Beckham et al., 2003; Ninomiya et al., 2001) implied an absence of endoderm/ mesoderm inducing activities in these cells. In contrast, Karadge (2012) indicated that VC cells from NF8 contained RNA for Activin ligands like *EcActivinB* and *EcDerriere*. Moreover, their expressions were upregulated at NF10, when gastrulation begins. Sun et al. (1999) reported that the *VegT*

expression in *X. laevis* regulates *Derriere* expression in the vegetal half of the embryo. One possible explanation for the finding in *E. coqui* is that the maternal supply of *EcDerriere* in the early embryo was sufficient to autoregulate its zygotic expression in VC cells at NF10 (Karadge, 2012). Karadge (2012) also suggested the presence of a weak mesoderm inducing activity in VC cells rather than a total absence from the VC, as reported by Ninomiya et al. (2001). The weak mesoderm inducing activity, which appeared later than in the MZ cells, apparently did not commit VC cells to differentiate as mesendoderm (Karadge, 2012).

For my work, I looked at the expression of Smad2, an essential transcription factor of the Nodal-signaling pathway. I showed that *EcSmad2* RNA is maternally supplied, and the level remained high until end of gastrulation. Moreover, during pre-gastrulation and early gastrulation, *EcSmad2* expression was greater in VC cells compared to MZ cells. Not only the RNA, but also both native and the active forms of EcSmad2 protein were expressed in VC and MZ cells. Immunostaining of NF10 embryos showed nuclear localization of PSmad2 in 100% of the MZ cells, whereas only 12% of VC cells were positive for nuclear PSmad2 signal. Nuclear Smad4 signal was detected in 100% MZ cells, but in 50% of the VC cells at NF10. I have summarized the similarities and differences between the early embryonic events in terms of Smad2 and Smad4 expressions between the *E. coqui* and *X. laevis* in Table 9. Based on my western and immunostaining results with EcSmad2 and EcSmad4 (Table 9), Nodal-signaling differs substantially between *X. laevis* and *E. coqui* VC cells.

I also investigated the nature of the VC with respect to its transcriptional activity, which led to a further understanding of why VC cells do not commit to differentiation.

Immunostaining to detect the functional status of RNAPII revealed two important facts. First, I provide evidence for the occurrence of MBT at early gastrulation (NF10) in *E. coqui*. The MZ cells of NF8 embryos are transcriptionally repressed and become active at NF10. Second, I show that VC cells remain transcriptionally silent, not only during gastrulation, but throughout development.

Table 10. Differential Smad2 and Smad4 expression during early embryogenesis between *X. laevis* and *E. coqui*. This comparison is focused on two of the early stages, namely blastula (NF8) and early gastrula (NF10). It illustrates how the temporal and spatial expressions of the Smad2 and Smad4 proteins are similar or different.

Table 10:

Protein	*X. laevis*	*E. coqui*
Smad2	1. Native Smad2 protein was present through all early embryonic stages. 2. Active PSmad2 expression began around NF8 and peaked at NF9.5. 3. PSmad2 expression remains high until late gastrulation and decreased at NF12. 4. At early gastrulation, Smad2 and PSmad2 proteins were detected in both MZ and VC cells, but not in animal cap cells. 5. Both Smad2 isoforms were phosphorylated in all the stages proteins were detected. (Faure et al., 2000) 6. During early gastrulation, PSmad2 signal was high in VC and moderate in MZ. There was nuclear accumulation of PSmad2 in both cell types, but no cell counts were provided. (Schohl & Fagotto, 2002)	1. Native EcSmad2 protein was detected through early embryogenesis. 2. No specific temporal expression pattern for active EcSmad2 was observed. 3. Activation of EcSmad2 was observed in later post-gastrulation stages. 4. Both MZ and VC at NF10 showed expression of native and active EcSmad2, but expression in VC is much lower. 5. Activation of both isoforms was not clearly evident. Usually one prominent band for active EcSmad2 was detected. 6. During early gastrulation, 100% of MZ cells and only 12% of VC cells showed nuclear accumulation of PSmad2.

Protein	*X. laevis*	*E. coqui*
Smad4	1. Two isoforms of *X. laevis* Smad4, XSmad4α and XSmad4β were reported to show different temporal expression patterns during early embryogenesis. 2. From MBT (NF8) until midgastrulation (NF11), expression of XSmad4β was much higher than that of XSmad4α. During this period, XSmad4β regulate Nodal-signaling. 3. After mid-gastrulation, expression of XSmad4α predominated and responded to Nodal-signal. 4. There have been no reports of spatial expression of Smad4 isoforms. (Howell et al., 1999) 5. No subcellular localization of Smad4 protein by immunostaining has been reported.	1. Both Smad4 isoforms were detected. 2. At NF10, EcSmad4β expression predominated. 3. Expression patterns of the two isoforms did not change after mid-gastrulation. Expression of EcSmad4α remained low during post-gastrulation. 4. Spatial expression of the isoforms differed radically between MZ and VC at NF10. EcSmad4β expression was higher than EcSmad4α in MZ tissue. Both EcSmad4β and EcSmad4α expression was negligible in VC cells. 5. At NF10, 100% of MZ cells showed nuclear Smad4 signal, whereas 50% of VC were positive for Smad4 nuclear signal.

I. *EcSmad2* RNA is maternally contributed and expressed in VC during early gastrulation.

Previous work from our lab provided evidence in favor of VC cells expressing Nodal ligand genes like *EcActivinB* and *EcDerriere* (Karadge, 2012) and the transcription factor *EcSox17* (Buchholz et al., 2007) during early gastrulation. These results suggest that the Nodal-signaling pathway is active in VC cells. In order to investigate this possibility, I examined the expression patterns of *EcSmad2*, which resides immediately downstream of the Nodal-receptors in the signaling cascade and which in *X. laevis*, carries the signal into the nucleus for endoderm/mesoderm specification.

I used degenerate PCR and 5'RACE to clone both the full length as well as the shorter splice variant of *EcSmad2* cDNA. Both the isoforms, *Smad2* and *Smad2Δexon3*, have been cloned and characterized in *X. laevis* (Graff et al., 1996; Faure et al., 2000) and in human (Takenoshita et al., 1998). The *EcSmad2* sequence revealed a high level of evolutionary conservation. When the nucleotide sequence of *EcSmad2* cDNA was aligned with that from human, mouse, *X. laevis* and *X. tropicalis*, there was 86% sequence conservation among the frog species, whereas the *E. coqui* sequence was 82% similar to those from human and mouse (Table 6). Alignment of the predicted peptide sequences indicated 99% conservation of the EcSmad2 amino acid sequence with those of *X. laevis* and *X. tropicalis* and 98% conservation with human, mouse and chicken Smad2 sequences (Table 7). Like in *X. laevis*, the splice variant, *EcSmad2Δexon3*, is 90 bp shorter than the full-length isoform. In the full-length isoform, exon3 codes for an insert in the MH1 domain preventing the protein from binding DNA (Faure et al., 2000). In *X. laevis*, deletion of the exon3 from the full-length ORF was sufficient to introduce

DNA binding activity to Smad2 (Dennler et al., 1998). Like *X. laevis*, the EcSmad2 protein also has MH1 (amino acid 11 to 171) and MH2 (amino acid 268 – 467) domains (Fig. 15). Based on the sequence conservation with *X. laevis* and *X. tropicalis*, amino acids 79 to 108 within the MH1 domain represents the position of the exon3 encoded region.

I used the ΔΔCt method of qPCR to detect the temporal and spatial expression of *EcSmad2*. The primers designed for the reaction span the exon3 region, so that the reaction results include both isoforms. *EcL8*, a ribosomal protein-coding gene was used as an endogenous control for all of my expression studies. Temporal expression data for *EcSmad2* RNA indicated a strong maternal contribution (Fig. 16). During NF8 and NF10, *EcSmad2* RNA amounts remained high, followed by a gradual decrease with development.

For spatial expression analysis, I compared MZ and VC for NF8 and NF10, and between embryonic tissues and NE in post-gastrulation stages. Like *EcVegT* and *EcVeg1* expression, I expected absence of *EcSmad2* RNA in VC. In contrast, the VC tissue had higher amounts of *EcSmad2* RNA than MZ at both NF8 and NF10. Moreover, for NF10 VC cells, the *EcSmad2* RNA level was more than that in NF8 VC cells (Fig. 17). There are two possible explanations for this higher expression of *EcSmad2* RNA in NF10 VC cells. First, more *EcSmad2* RNA may be made in VC cells during early gastrulation. This seems unlikely given the transcriptional repression that I found in VC cells. Second, the magnitude of *EcSmad2* RNA degradation in MZ may be higher compared to VC. This would result in a relatively higher *EcSmad2* expression in VC cells. At post gastrulation

stages, the expression profile became more like what I expected. Expression of *EcSmad2* was much lower in NE compared to the embryo at later stages.

In summary, *EcSmad2* RNA was maternally contributed, and the RNA level remained high until gastrulation. The spatial expression results contradicted my speculation that VC was devoid of molecular determinants. These results support the possibility of an active Nodal-signaling pathway in VC, and they provide evidence for differential expression patterns in MZ vs. VC cells. Based on the available results, I speculate that MZ cells, which become committed to form endoderm and mesoderm in the gastrulating embryo, have sufficient *EcSmad2* RNA to carry out the required level of Nodal-signaling. This would lead to the proper specification of endoderm and mesoderm. On the other hand, VC cells have more *EcSmad2* RNA than required, leading to improper Nodal-signal and resulting in faulty or no differentiation.

However, speculating final outcomes based on RNA amounts does not provide us with the real picture. Functional activity of a gene depends on variety of parameters like protein translation from RNA, proper processing of the protein products, and post-translational modifications leading to either activation or degradation of the protein. To analyze the behavior of the EcSmad2 proteins, I performed western blotting.

II. VC cells from *E. coqui* early gastrulae have Smad2.

Investigation of the endogenous Nodal-signaling pathway is important to understand its temporal and spatial activity and how it influences early embryogenesis. Faure et al. (2000) reported that in *X. laevis*, Smad2 and Smad2Δexon3 are highly phosphorylated towards the vegetal pole of *X. laevis* embryos during early gastrulation.

The expression of active Smad2/Smad2Δexon3 is moderate near marginal zones and undetectable in the animal cap (Faure et al., 2000). Anti-Smad2/3 (BD transduction) and anti-PSmad2 ($Ser^{465/467}$) antibodies (Cell Signaling) were used to detect the native and the active form of the protein respectively in *X. laevis* (Faure et al., 2000). They also detected the presence of shorter splice variant (Faure et al., 2000). In *X. laevis*, Smad2/Smad2Δexon3 activation coincides with the expression of VegT protein. VegT protein is exclusively and extensively present in the nuclei of the cells from vegetal hemisphere of blastulae, which become committed to form endoderm during gastrulation (Stennard et al., 1999). At the onset of gastrulation, VegT was not only detected in endoderm, but also in presumptive mesoderm (Stennard et al., 1999).

I used anti-Smad2/3 (BD transduction) and anti-PSmad2 (Cell Signaling) antibodies to detect the native and the active form of EcSmad2 on western blots.

A. EcSmad2Δexon3 is expressed, but not activated in VC cells.

Detection of EcSmad2Δexon3 in both temporal (Fig. 19) as well as spatial (Fig. 20) expression experiments is significant. The intrinsic DNA-binding activity is present in Smad2Δexon3, but not in Smad2 (Dennler et al., 1998; Takenoshita et al., 1998). Smad3 shares this property of Smad2Δexon3 (Yagi et al., 1999). *Goosecoid*, a homeobox gene, known as a marker of gastrulation and anterior mesoderm induction, is an early target of Smad2/3-FAST mediated transcription complex under Activin/Vg1 signaling in mouse and *X. laevis* (Blum et al., 1992; Cho et al., 1991; Labbé et al., 1998). When Smad2 was replaced with Smad3 in a FAST mediated complex targeting the mouse *goosecoid* promoter, transcription was repressed instead of activated (Labbé et al., 1998).

Due to the presence of similar DNA-binding property between Smad3 and Smad2Δexon3, this result could indicate an opposing behavior for the Δexon3, differing significantly from Smad2. Although these two functionally different isoforms are co-regulated in the same pathway, they could display significant differences in their responses to Nodal ligands during early embryonic development (Faure et al., 2000; Labbé et al., 1998).

My results indicate that not only the alternatively spliced mRNA, but also the protein product is maternally contributed. Following fertilization, the inactive EcSmad2Δexon3 is present in all stages. Western blots with anti-PSmad2 ($Ser^{465/467}$) antibody detected one strong band in most cases (Fig. 19 and 20). Based on size markers, phosphorylation of the shorter isoform was not detected in the stages examined. Although EcSmad2Δexon3 is expressed in the oocytes and in both MZ and VC regions of early embryos, it is not activated. These results suggest no functional role for EcSmad2Δexon3.

B. EcSmad2 is expressed and activated in VC cells at gastrulation.

Full length EcSmad2, on the other hand, was not only expressed but also was activated in oocytes as well as in MZ and in VC of early gastrulae (Fig. 20). During early gastrulation when specification of the germ layers was initiated, Nodal-signaling via active Smad2 played the pivotal role in *X. laevis*. The vegetal half of the embryo expressed VegT protein (Stennard et al., 1999), which in turn activated *Sox17* expression. VC cells were also characterized by extensive expression of Smad2, which finally committed these cells to become endoderm and mesoderm (Faure et al., 2000).

In *E. coqui*, the VC cells contain *Sox17* RNA but lack *EcVegT* RNA. Although my western blotting results indicated expression and activation of EcSmad2 in VC cells, it is not sufficient to answer the question why *E. coqui* VC cells do not commit to endoderm/mesoderm and become undifferentiated nutritional reserve. Presence of the active form of a signal transducer in a cell, however, does not necessarily mean that the signaling pathway is active. Lack of differentiation definitely argues against an active Nodal-signaling in VC cells.

Based on Fig. 20B, one can raise the question whether the single band recognized by the anti-PSmad2 antibody was EcSmad2Δexon3, rather than EcSmad2. In that case, a selective phosphorylation of EcSmad2Δexon3 could be proposed, which would lead to inhibition of transcription, as reported by Labbé et al. (1998) in *X. laevis*. Although that could explain why VC cells do not commit to endoderm/mesoderm specification, the position of the size marker on the western blot did not support this idea.

Therefore, these results could indicate another possibility, namely a potential block for Nodal-signaling downstream of Smad2 activation. The next step in nodal-signaling is the association of active Smad2 with its co-Smad partner, Smad4. This protein complex translocates inside the nucleus to turn on endoderm/mesoderm-specifying genes in vertebrates (Lagna et al., 1996; Zhang et al., 1996). The outcome of my western analysis of EcSmad2 led me to investigate the expression of EcSmad4.

III. VC cells from gastrulating embryos express Smad4.

There are no previous records for use of anti-Smad4 antibody to detect endogenous levels of Smad4 expression in *X. laevis*. I used rabbit anti-Smad4 polyclonal

antibody (Thermo Scientific), which was raised against the human Smad4 protein and which was predicted to cross-react with the *X. laevis* homolog based on sequence similarity. In the absence of *EcSmad4* sequence and based on my previous experience with EcSmad2 conservation with other species, I decided to use this antibody on western blots to detect endogenous levels of EcSmad4 expression.

Unlike human, there are two isoforms of Smad4 proteins in *X. laevis* (LeSueus and Graff, 1999; Masuyama et al., 1999; Howell et al., 1999; Hill, 2001). The potential *X. laevis* ortholog of *hSmad4* is *XSmad4*α, whereas *XSmad4*β is a novel one that does not have a homolog in other species. It is also known as *Smad10* (Howell et al., 1999; LeSueus and Graff, 1999; Masuyama et al., 1999). Based on the reported ORF sequences, the predicted molecular weights for *XSmad4*α and *XSmad4*β are 59.8 kDa and 61.2 kDa, respectively (Masuyama et al., 1999; *XSmad4*α Genebank ID: AB022721.1 and *XSmad4*β Genebank ID: AB022722.1). My western blot results (Fig. 21A and B) show the presence of two isoforms of Smad4 in *E. coqui* around 60 kDa.

Both the Smad4 α and β isoforms are expressed at NF8, NF10, and TS5. As I did not clone *EcSmad4*, it is unclear which band corresponds to which isoform. Based on the protein size prediction in *X. laevis*, the higher molecular weight band is tentatively, EcSmad4β and the lower molecular weight band is EcSmad4α.

Spatial expression analysis was performed only on proteins from NF10 embryos, as this could be the time for pattern formation. Investigating EcSmad4 expression at NF10 could boost our understanding of nature of Nodal signaling at this time. From the spatial expression pattern of EcSmad4 (Fig. 21B), both isoforms were detected in whole embryo preparations, as well as from MZ and VC. In both whole and MZ protein

preparations, the expression of EcSmad4β was higher than that of EcSmad4α. In contrast, VC showed a lower expression of EcSmad4β, almost equal to that of EcSmad4α. Overall, expressions of both isoforms were much less in VC compared to MZ or the whole embryo (Table 9).

According to Howell et al. (1999), Smad4 isoforms show strikingly different temporal expression patterns in early *X. laevis* embryos (Table 9). Adult tissues were also characterized by their expression in different ratios, which suggests their different and specific roles. First, the subcellular distributions for these two isoforms are different. Due to the lack of NES, Smad4β is exclusively nuclear. Therefore, it is Smad4α, which predominantly resides in cytoplasm and mediates Nodal-signaling via binding with active Smad2/3. XSmad4β protein expression is high during pre-gastrulation and early gastrulation in *X. laevis* (Howell et al., 1999; Hill, 2001). At mid-gastrulation, the ratio of the two isoforms changes depending on the availability of different sets of Nodal ligands, and expression of Smad4α predominates (Howell et al., 1999). There are some reports, which suggest that until early gastrulation, the complex formation between active Smad2 and Smad4β can occur inside the nucleus (Hill, 2001; Howell, 1999). The complex formation inside the nucleus could also work for Smad4α, as it could be present on either side of the nuclear membrane (Hill, 2001). Also Liu et al. (1997) showed that it is not always necessary for R-Smads to bind Smad4 to accumulate inside the nucleus. These reports leave us with multiple possibilities on which isoform is needed at a specific stage of development and where it is localized.

My EcSmad4 temporal expression result (Fig. 21A) shows that the expression of EcSmad4 increased as the embryo developed, like what I found for EcSmad2. This

temporal expression pattern is not in agreement with that reported in *X. laevis* (Table 9). According to reports in *X. laevis* (Howell et al., 1999; Masuyama et al., 1999), after mid-gastrulation expression of Smad4α is higher than that of Smad4β. My result (Fig. 21A) for EcSmad4 isoforms showed gradual increase in the intensity of tentative EcSmad4β band with development. None of the post-gastrulation stages showed higher expression of EcSmad4α than EcSmad4β (Fig. 21A).

On the other hand, the EcSmad4 spatial expression result (Fig. 21B) is more informative. According to the established concept in *X. laevis* (Howell et al., 1999; Masuyama et al., 1999) Smad4β is required to mediate Nodal-signaling from MBT until mid-gastrulation. Western blot analysis of the *E. coqui* early gastrula indicated the predominance of EcSmad4β in MZ tissues, which will differentiate to mesendoderm. In contrast, VC cells had a very low level of EcSmad4β expression (Table 9). Based on my results, I suggest that the low level of EcSmad4β expression prevents VC cells from differentiation. Unlike MZ cells, a lack of enough EcSmad4β, whether inside the nucleus or outside in the cytoplasm, could prevent Nodal-signaling, leading to the development of undifferentiated NE.

IV. ActRIIA, an activin receptor type II, may be present during early embryogenesis.

Karadge (2012) provided evidence that different activin ligands are expressed in prospective NE of *E. coqui*. This observation raised the possibility that TGF-β receptors might also be expressed. Although a couple of TGFβ cDNAs were cloned in our lab (Karadge, 2012), no receptor genes had been looked at. Martello et al. (2007) used anti-

human Activin RIIA antibody (R&D Systems) against the *X. laevis* receptor protein on western blots. Although the antibody was predicted to show less than 2% cross-reactivity with human Activin RIIB or any other Activin receptor, it detected the endogenous levels of the *X. laevis* homolog (Martello et al., 2007).

I used this antibody on western blots of *E. coqui* embryonic protein preparations. The predicted molecular weight of human Activin RIIA is 58 kDa and that of the *X. laevis* homolog is 57.9 kDa (Xenbase: XB-GENE-865037), but Martello et al. (2007) did not provide size information on their western blots. I made several attempts to use this antibody on western blots with *E. coqui* embryos (Fig. 22). In all my experiments, a specific band was detected at ~50 kDa, instead of 58 kDa. The same size band was detected consistently in NF8, NF10 and TS4. This could mean that the *E. coqui* Activin RIIA homolog is a shorter one, but shorter by almost 100 amino acids is not very likely, especially when these signaling proteins are usually highly conserved. I tried the same antibody to immunostain MZ and VC cells dissociated from NF10 embryos. Although only a few populations of both MZ (20%) and VC (17%) cells showed label on the surface (Fig. 25), there is nothing conclusive about it. In the absence of *EcActRIIA* sequence data, it is not known whether the band obtained around 50 kDa is actually *E. coqui* ActRIIA. Therefore, I decided not to proceed further with this study, which would best be advanced by cloning of the gene.

V. During gastrulation, only 12% of VC cells show nuclear localization of active Smad2, but 50% are positive for Smad4.

My results indicated that the VC cells at gastrulation contain *EcSmad2* transcripts, native and active forms of the EcSmad2 proteins, and the Smad2 partner, EcSmad4. However, these results do not indicate whether these proteins are translocated into the nucleus or are just resident in the cytoplasm of the VC cells. It was important to know the cellular location of these proteins to conclude whether Nodal-signaling is active in VC cells. The most reliable method of characterizing the events of a specific signaling cascade is the direct detection of its components in their active state. Upon Nodal-stimulation, the active Smad2 in combination with Smad4 should localize in the nucleus (Cobb and Goldsmith, 2000; Heldin et al., 1997; Massagué, 1998). As an exception, both active Smad2 (Liu et al., 1997) as well as Smad4 (Bai et al., 2002) can translocate into the nucleus independently of each other.

Identification of activated Smad2 and its localization by means of immunostaining has been reported in *X. laevis* (Christen and Slack, 1999; Faure et al., 2000; Schohl and Fagotto, 2002). All previously reported immunolocalization data were obtained using whole mount procedures or immunofluorescence on cryosections (Fagotto and Gumbiner, 1994; Fagotto, 1999). These techniques are not ideal for studying large embryos filled with yolk. I performed my immunostaining on individual cells dissociated from the dissected MZ and VC from the *E. coqui* early gastrulae.

I used the same anti-PSmad2 antibody, which was used for my western blotting experiments. Dissociated NF10 MZ and VC cells were stained with the antibody and counter stained with DAPI to identify the nucleus. In all attempts, 100% of the MZ cells showed nuclear accumulation of PSmad2 as evident by the strong PSmad2 signal overlapping with DAPI. On the other hand only 12% of VC cells with single or multiple

nuclei were positive for such overlap (Table 8; Fig. 23). The remaining cells showed either scattered cytoplasmic distribution of the active protein or no labeling. VC cells are characterized by a varying number of nuclei, ranging from no nucleus to 3-4 nuclei in one cell (Karadge, 2012). Therefore, extra precaution was taken to count only the VC cells, in which at least one nucleus was detected. Although there is nuclear accumulation of PSmad2, 12% is low to draw a conclusion as to its effects.

I investigated the nuclear localization of Smad2 protein into the VC nuclei using an alternative approach. 8-16 cell embryos were injected with tracer level of 10 pg *EGFP-Smad2* capped mRNAs and grown until early NF10. Dissociated cells from dissected VC tissues were stained with anti-GFP antibody to detect the location of the tagged Smad2 proteins. Only 8% of VC cells showed nuclear accumulation of recombinant protein (Fig. 26). This result strengthens my finding with endogenous EcSmad2 nuclear localization (Fig. 23).

MZ and VC cells from NF10 were also stained with anti-Smad4 antibody, which detected EcSmad4 on western blots. For MZ cells, 100% were positive for nuclear EcSmad4 localization, but only 50% of VC cells were positive for a nuclear EcSmad4 signal (Table 8; Fig. 24). Moreover almost half of the VC cells, which lacked nuclear signal, also lacked any cytoplasmic staining. The positive nuclear signal was observed even for multinuclear cells, and nuclear staining was uniform.

The 12% and 50% of VC cells, which showed nuclear location of the PSmad2 and Smad4, respectively, provide many possible explanations. First, it may be that PSmad2 and Smad4 translocate inside the nucleus of only subpopulations of VC cells, and this is quantitatively not strong enough to carry on Nodal-signaling for endoderm/mesoderm

specification. It is possible that those cells, which reside inside the overlapping domain between 12% VC cells for active Smad2 and 50% VC cells for Smad4, could carry out Nodal-signaling. To confirm this hypothesis, I needed co-immunostaining with antibodies against both active Smad2 and Smad4. Co-immunostaining was not currently possible to do, since both of these antibodies were raised in rabbit.

Second, the overall expression of EcSmad4 may not be sufficient to carry on Nodal-signaling. According to the reports in *X. laevis*, Smad4β is exclusive to the nucleus, whereas Smad4α shuttles between cytoplasm and nucleus (Howell et al., 1999; Masuyama et al., 1999). At pre-gastrulation and early gastrulation in *X. laevis*, it is Smad4β, not Smad4α (Howell et al., 1999; Masuyama et al., 1999) that predominates in expression. Western blots using anti-Smad4 antibody revealed that at early gastrulation, expression of both Smad4 isoforms in VC cells is weak compared to MZ cells (Fig. 21A & B). Although, the immunostaining signal for VC cells was bright when compared to that for MZ, we are looking at individual cells. Western blot, on the other hand, represent a broader view by looking at the whole tissue.

Third, Smad4 can act as a transcriptional activator that is independent of Smad2-mediated Nodal-signaling. That was evident in case of SMIF, a protein that translocates into the nucleus along with Smad4 without involving Smad2/3 (Bai et al., 2002). Continuous nucleocytoplasmic shuttling of Smad4 on its own has been suggested by Pierreus et al. (2000), Watanabe et al. (2000) and Shi and Massagué (2003), but this is true only for human Smad4 and *X. laevis* Smad4α, which share the NES motif in their MH1 domain. Hence, the nuclear signal for EcSmad4 in those 50% VC cells could be due to Smad4α, which is brought inside the nucleus along with some unidentified protein

rather than Smad2/3. In addition to the differential expression of Nodal-signaling, all these possibilities suggest the presence of another regulatory block, further downstream in the pathway, which ultimately determines the fate of VC cells.

VI. Global transcriptional repression may prevent VC cells from differentiation.

Although several lines of evidence in this work suggest a block of the Nodal-signaling pathway in the *E. coqui* VC, I also investigated the global transcriptional status of the VC cells. Staining of NF10 VC cells with antibodies against phosphorylated RNAPII CTD-Ser 2, phosphorylated CTD-Ser 5 and unphosphorylated RNAPII large subunit showed a lack of transcriptional elongation.

MZ cells, which form endoderm/mesoderm, showed positive staining for all three antibodies at NF10. This means that MZ cells underwent transcription initiation as well as transcription elongation. Unlike MZ cells, VC cells were positive for unphosphorylated RNAPII and phosphorylated CTD-Ser 5. Not only was RNAPII present in nuclei of VC cells, but also they underwent transcription initiation. VC cells lacked any signal for phosphorylated RNAPII CTD-Ser 2, which means no transcription elongation (Fig. 27). A small number of VC cells showed a faint signal for RNAPII CTD-Ser 2, which could indicate that there was a low level of transcription. Based on these results, I conclude that VC cells at NF10 stage are transcriptionally repressed globally. On the other hand, MZ cells shows proper transcription initiation and elongation during early gastrulation, which would allow them to commit to endoderm/mesoderm formation.

The global transcriptional repression existed for both MZ and VC cells at NF8 (Fig. 27). Only a few MZ and VC cells were positive for transcription initiation, but neither cell type showed transcription elongation. At the onset of gastrulation, MZ cells became transcriptionally active (Fig. 28) and remained active at every developmental stage examined. VC cells failed to undergo transcription elongation at gastrulation (Fig. 28) and remained transcriptionally silent through development. By the end of gastrulation, VC becomes nutritional endoderm (NE). I tested dissociated NE cells from post-gastrulation stages TS3, TS8, TS11 and TS14, a stage right before hatching (Fig. 29, 30, 31 and 32). No sign of transcription initiation or elongation was detected in them with the exception of a few NE cells from each stage, which showed a faint signal for transcription initiation or elongation. This may indicate that although VC cells of the pre- and early-gastrulating embryos and NE cells post-gastrulation, are transcriptionally repressed globally, there could be a few genes, which are actively transcribing. The *E. coqui* thyroid hormone receptor gene (*EcTRβ*) could represent one such gene, whose RNA level was upregulated by TS14 (Singamsetty and Elinson, 2010).

Similar global transcriptional repression events have been reported in *X. laevis* PGC formation (Venkatrama et al., 2010). Cells containing germ plasm are transcriptionally repressed while neighboring cells of endoderm lineage are transcriptionally active. Similar events are reported in *C. elegans* (Guven-Ozkan et al., 2008) and *D. melanogaster* (Hanyu-Nakamura et al., 2008). In mouse, PGCs are formed after gastrulation, and they do not require germ plasm. Seki et al. (2007) showed that the migrating mouse PGCs were negative for P-Ser5 or P-Ser2, also suggesting no transcription initiation or elongation.

VII. MBT in *E. coqui* may occur by NF10.

During early embryogenesis, MBT is a landmark event. Before MBT, early embryonic cells are transcriptionally quiescent, and survive on maternal supplies of proteins and RNAs (Newport and Kirshner, 1982; Kisielewska and Blow, 2012). Embryonic gene expression is initiated at MBT (Gerhart, 1980; Heasman, 2006; Shiokawa, 2012). In *X. laevis*, NF8 is designated as MBT.

In *E. coqui*, MBT has never been defined, but the global transcriptional status of MZ cells provides strong evidence for MBT in *E. coqui*. At NF8, MZ cells showed no sign of transcriptional initiation or elongation (Fig. 27). At NF10, MZ cells became transcriptionally active, as they were positive for both transcription initiation and elongation (Fig. 28). This event of global transcriptional activation in MZ cells marks the onset of zygotic gene expression during *E. coqui* early embryogenesis, thus establishing MBT by NF10.

VIII. A model for Nutritional Endoderm formation.

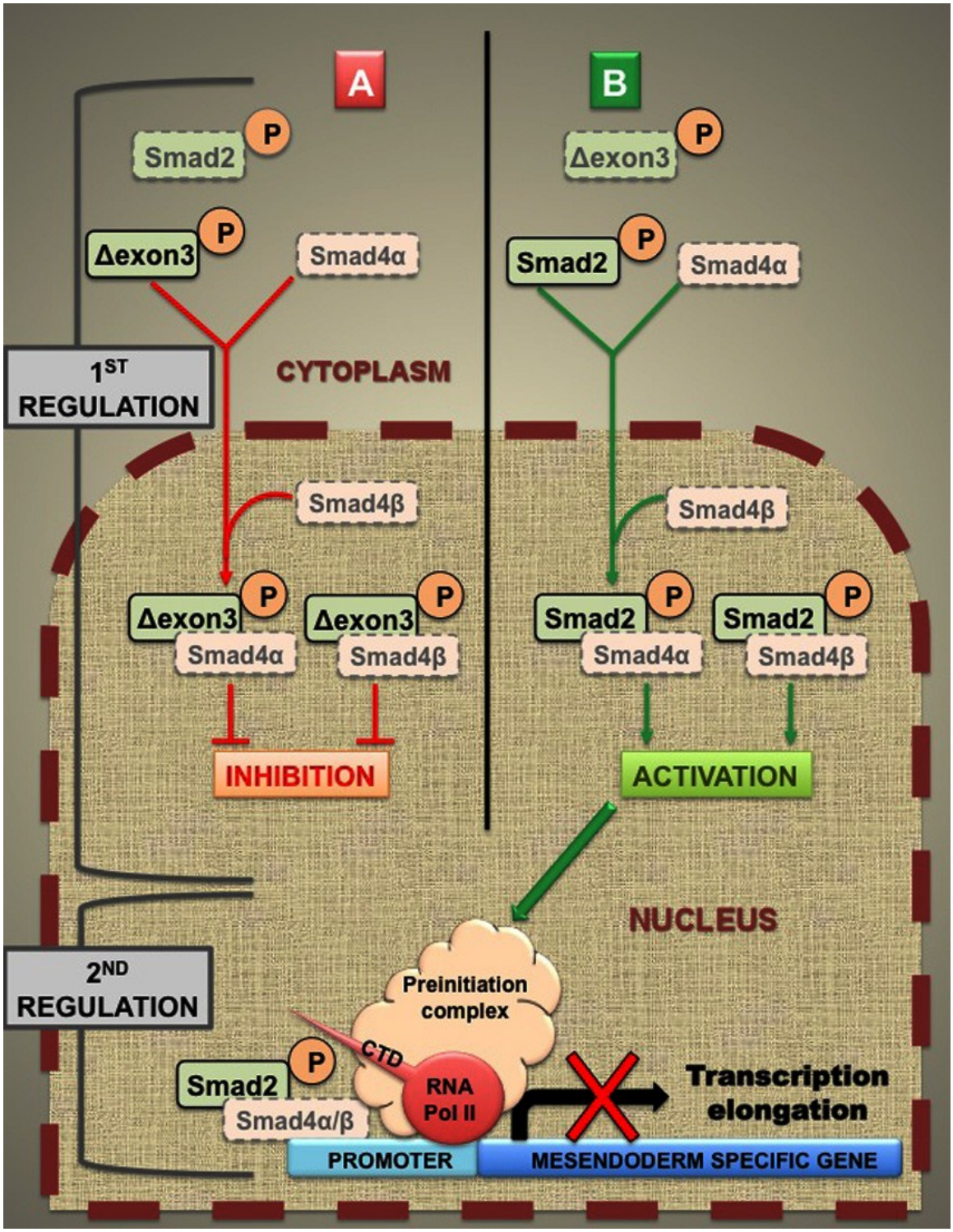
Taking all of these finding into consideration, I propose a model for the differences in the state of differentiation of MZ and VC cells (Fig. 33). My data suggests that in *E. coqui* the Nodal-signaling pathway (Fig. 1) could account for endoderm and mesoderm specification in MZ cells as occurs in *X. laevis*. VC cells, on the other hand, differ from the specification pathway, at two levels: first, by differential expression of Nodal-signaling components like Smad2 and Smad4, and second, by global transcriptional repression. My model (Fig. 33) represents the modifications of the

specification pathway and incorporates different possibilities for the generation of NE from VC cells.

This model proposes a two-phase regulation of NE development from VC cells at NF10. In the first phase, there are two possibilities. According to possibility 'A', only the alternatively spliced variant, EcSmad2Δexon3 is activated and undergoes nuclear translocation before association with Smad4β, or after associating with Smad4α in the cytoplasm. EcSmad2Δexon3 inside the nucleus could exert an inhibitory effect, as described in *X. laevis*, and block the Nodal-pathway. According to possibility 'B', it is the full length EcSmad2, which is activated and translocates inside the nucleus in association with poorly expressed Smad4α (in the cytoplasm) or Smad4β (inside the nucleus). This event is then subjected to a second phase of regulation. In the absence of P-Ser2 at CTD of RNAPII, a global transcriptional repression prevails in the vegetal half, blocking any effects of Nodal-signaling. This repression could also explain why *EcSox17*, whose RNA is present in VC cells (Buchholz et al., 2007), does not cause endodermal differentiation. Even if *EcSox17* RNA were translated, this transcription factor would not be able to activate downstream endodermal genes.

Fig. 33: Two-phase regulation for deciding the fate of VC cells. At NF10, both Smad4 isoforms are expressed weakly. Smad4α resides in cytoplasm and Smad4β localizes exclusively in the nucleus. **Phase 1 regulation:** Regulation of Nodal-signaling. There are two possibilities. (A) If the spliced variant, EcSmad2Δexon3, is the only form that is phosphorylated, its association with either of the Smad4 isoforms would lead to inhibition of Nodal-signaling. (B) If it is EcSmad2, which is phosphorylated, it could associate with Smad4 isoforms and may induce Nodal-signaling in a small subpopulation of VC cells. **Phase 2 regulation:** Global transcriptional repression. Repression occurs via absence of P-Ser2 at CTD of RNAPII, leading to no elongation. If an active PSmad2-Smad4 complex were present inside the nucleus of VC cells, it will not turn on the transcription machinery, as CTD of RNAPII is not properly phosphorylated. This leads to the development of NE instead of DE.

Fig. 33:


A
B
Smad2
P
Δexon3
P
Δexon3
P
Smad4α
Smad2
P
Smad4α
1ST REGULATION
CYTOPLASM
Smad4β
Smad4β
Δexon3
P
Smad4α
Δexon3
P
Smad4β
Smad2
P
Smad4α
Smad2
P
Smad4β
INHIBITION
ACTIVATION
NUCLEUS
2ND REGULATION
Preinitiation complex
CTD
Smad2
P
Smad4α/β
RNA Pol II
Transcription elongation
PROMOTER
MESENDODERM SPECIFIC GENE

IX. Future perspective.

My findings have indicated a probable mode of NE development. Smad2 is a well characterized TGF-β pathway intermediate performing many roles during vertebrate and invertebrate development. The role of Smad2 in germ layer specification has been well documented in *X. laevis*, but nothing was known for direct developing frogs. I addressed the question whether NE develops because of the absence of the molecular determinants in the VC cells. My work clearly indicated that the VC was not devoid of molecular determinants; rather during early embryogenesis, it expressed some of the Nodal-pathway specific proteins, especially Smad2. Characterization of EcSmad2 and EcSmad4 expressions in *E. coqui* early embryos showed us how differential expression in different regions of embryo may lead to the development of NE and DE. Finally, determination of the transcriptional status of early *E. coqui* embryonic cells furnished us with another layer of regulation. My studies revealed that MZ cells, the precursors of future definitive endoderm, became transcriptionally active at early gastrulation and remained active thereafter. On the other hand, the large, yolky VC cells stayed transcriptionally repressed throughout development and served only as nutritional reserves.

My results raised some interesting questions regarding *E. coqui* early embryogenesis. There is a possibility of a small population VC cells containing nuclear PSmad2 and Smad4. It would be interesting to do a dual immunostaining for Smad2 as well as Smad4 in the same cells. In that case, specific antibodies against Smad2 and Smad4 have to be raised in different animal species. Availability of such antibodies could provide us with more direct and detailed analysis. In future, this question could be addressed with the application of more sophisticated tools. Use of Bimolecular

Fluorescence Complementation (BiFC) (Hu et al., 2002) to investigate the association between PSmad2 and Smad4 *in vivo* in dissociated MZ and VC cells could furnish a clearer picture on the molecular interactions and nuclear translocation of the Smad2-Smad4 complex. This technique has been optimized for use in *X. laevis* embryos by the use a mutant version of VENUS (Saka et. al., 2007, 2008). It is a semi-quantitative approach, which would permit a real-time analysis of Nodal signaling by monitoring the different signaling events during *E. coqui* normal development.

The idea emerged from my transcriptional activity assays that VC cells did not show transcriptional initiation or elongation. In some instances, however, there were a few cells, which showed faint nuclear signals. VC cells are not dead, and they are definitely involved in trafficking of vesicles containing nutrients. This must require at least a minimal set of gene products to be active. It raises the question what genes in the VC could be transcribed and whether they are PSmad2/Smad4 targets. There are several approaches ways to address these questions. One such approach would be to perform chromatin immunoprecipitation (ChIP) with antibodies against Smad2 or Smad4. ChIP has been successful in identification of several important genes involved in the early embryogenesis of *X. laevis* (Blythe et al., 2009).

To understand the complex molecular mechanisms underlying the early embryonic development, whole transcriptome analysis would be of great interest. Analysis of VC RNA expression would provide us with insights into the molecular mechanisms of different signaling pathways, which determines the fate of NE. A successful whole transcriptome analysis depends on the quality of the RNA and efficient

production to cDNA libraries. Annotation of the resulting sequences could be a potential problem, because the absence of an *E. coqui* genome currently making gene identification difficult.

REFERENCES

Abdollah S., Macías-Silva M., Tsukazaki T., Hayashi H., Attisano L., and Wrana J.L. 1997. TβRI phosphorylation of Smad2 on Ser^{465} and Ser^{467} is required for Smad2-Smad4 complex formation and signaling. *J. Biol. Chem*. 272: 27678 – 27685.

Adamson E.D., Minchiotti G., and Salomon D.S. 2001. Cripto: a tumor growth factor and more. *J. Cell. Physiol*. 190: 267 – 278.

Alexander J. and Stainier D.Y.R. 1999. A molecular pathway leading to endoderm formation in zebrafish. *Curr. Biol*. 9: 1147 – 1157.

Altmann C.R., Chang C., Munoz-Sanjuan I., Bell E., Heke M., Rifkin D.B., and Brivanlou A.H. 2002. The latent-TGFbeta-binding-protein-1 (LTBP-1) is expressed in the organizer and regulates nodal and activin signaling. *Dev. Biol*. 248: 118 – 127.

Attisano L. and Wrana J.L. 2002. Signal transduction by the TGF-beta superfamily. *Science*. 296: 1646 – 1647.

Bai R.Y., Koester C., Ouyang T., Hahn S.A., Hammerschmidt M., Peschel C., and Duyster J. 2002. SMIF, a Smad4-interacting protein that functions as a co-activator in TGF-β signaling. *Nat Cell Biol*. 4: 181 – 190.

Baker JC and Harland RM. 1996. A novel mesoderm inducer, Madr2, functions in Activin signal transduction pathway. *Genes & Dev.* 10: 1880 – 1889.

Baroux C., Autran D., Gillmor C.S., Grimanelli D., and Grossniklaus U. 2008. The maternal to zygotic transition in animals and plants. *Cold Spring Harb. Symp. Quant. Biol*. 73: 89 – 100.

Beard K.H. 2007. Diet of invasive frog, *Eleutherodactylus coqui*, in Hawaii. *Copea*. 2007: 281 – 291.

Beck S., Le Good J.A., Guzman M., Ben Haim N., Roy K., Beerman F., and Constam D.B. 2002. Extraembryonic proteases regulate Nodal signaling during gastrulation. *Nat Cell Biol.* 4: 981 – 985.

Beckham Y.M., Nath K., and Elinson R.P. 2003. Localization of RNAs in oocytes of *Eleutherodactylus coqui*, a direct developing frog, differs from *Xenopus laevis*. *Evol. Dev.* 5: 562 – 571.

Birsoy B., Berg L., Williams P.H., Smith J.C., Wylie C.C., Christian J., and Heasman J. 2005. XPACE4 is a localized pro-protein convertases required for mesoderm induction and the cleavage of specific TGF-β proteins in *Xenopus* development. *Development*. 132: 591 – 602.

Blum M., Gaunt S. J., Cho K. W. Y, Steinbeisser H., Blumberg B., Bittner D. and De Robertis E. M. 1992. Gastrulation in the mouse: the role of the homeobox gene *goosecoid. Cell.* 69: 1097-1106.

Blythe S.A., Reid C.D., Kessler D. S. and Klein P.S. 2009. Chromatin Immunoprecipitation in Early *Xenopus laevis* Embryos. *Dev Dyn*. 238: 1422 – 1432.

Branford W.W. and Yost H.J. 2002. Left dependent inhibition of nodal and wnt-responsive organizer gene expression is essential for normal gastrulation. *Curr. Biol.* 12: 2136 – 2141.

Buchholz D., Singamsetty S., Karadge U., Williamson S., Langer C.E., and Elinson R.P. 2007. Nutritional endoderm in a direct developing frog: A potential parallel to the evolution of amniote egg. *Dev Dyn.* 236: 1259 – 1272.

Burdine R.D. and Schier A.F. 2000. Conserved and divergent mechanisms in left-right axis formation. *Genes & Dev*. 14: 763 – 776.

Callery M.E. and Elinson R.P. 1996. Developmental regulation of Urea-Cycle enzyme arginase in the direct developing frog *Eleutherodactylus coqui. J Exp Zool.* 275: 61 – 66.

Callery E.M. and Elinson R.P. 2000a. Operculum development and ontogenetic reorganization in a direct-developing frog. *Dev Genes Evol.* 210: 377 – 381.

Callery E.M. and Elinson R.P. 2000b. Thyroid hormone-dependent metamorphosis in a direct developing frog. *Proc. Natl. Acad. Sci. USA*. 97: 2615 – 2620.

Capdevila J., Vogan K.J., and Belmonte J.C.I. 2000. Mechanisms of left-right determination in vertebrates. *Cell*. 101: 9 – 21.

Chalmers A.D. and Slack J.M. 1998. Development of the gut in *Xenopus laevis*. *Dev Dyn*. 212: 509 – 521.

Chalmers A.D. and Slack J.M. 2000. The *Xenopus* tadpole gut: Fate maps and Morphogenetic Movements. *Development*. 127: 381 – 392.

Chang C. and Hemmati-Brivanlou A. 2000. A post-midblastula transition requirement for TGF-β signaling in early endodermal specification. *Mech. Dev*. 90: 227 – 235.

Chea H.K., Wright C.V., Swalla B.J., 2005. Nodal signaling and the evolution of deuterostome gastrulation. *Dev. Dyn*. 234, 269–278.

Chen X., Rubock M.J., and Whitman M., 1996. A transcriptional partner for MAD proteins in TGFβ signaling. *Nature*. 386: 691 – 696.

Chen Y.G., Hata A., Lo R.S., Wotton D., Shi Y., Pavletich N., and Massagué J. 1998. Determinants of specificity in TGF-β signal transduction. *Genes Dev*. 12: 2144 – 2152.

Chen Y. and Schier A.F. 2001. The zebrafish Nodal signal Squint functions as a morphogen. *Nature*. 411: 607 – 610.

Chen Y. and Schier A.F. 2002. Lefty proteins are long-range inhibitors of squint-mediated nordal signaling. *Curr. Biol*. 12: 2124 – 2128.

Cho K. W. Y., Blumberg B., Steinbeisser H. and De Robertis E. M. 1991. Molecular nature of Spemann's organizer: the role of the *Xenopus* homeobox gene *goosecoid. Cell.* 67: 1111-1120.

Christen B. and Slack J. 1999. Spatial response to fibroblast growth factor signaling in *Xenopus* embryos. *Development*. 126: 119 – 125.

Clements D., Friday R.V., and Woodland H.R. 1999. Mode of action of VegT in mesoderm and endoderm formation. *Development.* 128: 4903 – 4911.

Clements D., Rex M., and Woodland H.R. 2001. Initiation and early patterning of the endoderm. *Int. Rev. Cytol.* 203: 383 – 446.

Clements D. and Woodland H.R. 2000. Changes in embryonic cell fate produced by expression of an endodermal transcription factor, XSox17. *Mech. Dev*. 99: 65 – 70.

Clements D. and Woodland H.R. 2003. VegT induces endoderm by a self-limiting mechanism and by changing the competence of cells to respond to TGF-β signals. *Dev Biol.* 258: 454 – 4463.

Cobb M.H. and Goldsmith E.J. 2000. Dimerization in MAP-kinase signaling. *Trends. Biochem. Sci*. 25: 7 – 9.

Collazo A. 1996. Evolutionary correlations between early development and life history in plethodontid salamanders and teleost fishes. *Am Zool*. 36:116 – 131.

Collazo A., Bolker J.A., and Keller R. 1994. A phylogenetic perspective on teleost gastrulation. *Amer Nat.* 144: 133 – 152.

Corden, J.L. 1990. Tails of RNA polymerase II. *Trends Biol. Sci.* 15: 383–387.

Correia J.J., Chacko B.M., Lam S.S., and Lin K. 2001. Sedimentation studies reveal a direct role of phosphorylation in Smad3:Smad4 homo and hetero-trimerization. *Biochem*. 40: 1473 – 1482.

Dale L. and Slack J.M. 1987. Fate map for the 32-Cell stage of *Xenopus laevis*. *Development*. 99: 527 – 551.

Dennler S., Itoh S., Vivien D., ten Dijke P., Huet S., and Gauthier J.M. 1998. Direct binding of Smad3 and Smad4 to critical TGF beta-inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene. *EMBO J*. 17: 3091 – 3100.

Dent J.N. 1968. Survey of amphibian metamorphosis. In *Metamorphosis* (ed. Etkin W. and Gilbert L.I.), pp. 271 – 311; Appleton-Crofts, New York.

Derynck R. and Zhang Y.E. 2003. Smad-dependent and Smad-independent pathways in TGF-β family signaling. *Nature*. 425: 577 – 584.

Dickinson K., Leonard J., and Backer J.C. Genomic Profiling of Mixer and Sox17beta targets during *Xenopus* endoderm development. *Dev Dyn*. 235: 368 – 381.

D'Souza A., Lee A., Taverner N., Mason J., Carruthers S., Smith J.C., Amaya E., Papalopulu N., Zorn A.M. 2003. Molecular components of the endoderm specification pathway in *Xenopus tropicalis. Dev Dyn.* 226: 118 – 127.

Duellman W.E. and Trueb L. 1986. Biology of Amphibians. McGraw-Hill Book Co.

Ecochard V., Cayrol C., Rey S., Foulquier F., Caillol D., Lemaire, P., and Duprat A.M. 1998. A novel *Xenopus mix*-like gene milk involved in the control of endomesodermal fates. *Development*. 125: 2577 – 2585.

Eimon P.M. and Harland R.M. 2002. Effects of heterodimerization and proteolytic processing on Derriere and Nodal activity: implications for mesoderm induction in Xenopus. *Development*. 129: 3089 – 3103.

Elinson R.P. 1987. Change in developmental patterns: embryos of amphibians with large eggs. New York: Alan R. Liss, Inc.

Elinson R.P. 1987. Fertilization and aqueous development of the Puerto Rican terrestrial-breeding frog, *Eleutherodactylus coqui*. *J. Morph*. 193: 217 – 224.

Elinson R.P. 1990. Cortical and cytoplasmic phases in amphibian eggs. *Ann N Y Acad Sci*. 582: 1 – 9.

Elinson R.P. 2001. Direct development: An alternative way to make a frog. *Genesis*. 29: 91 – 95.

Elinson R.P., 2009. Nutritional endoderm: a way to breach the holoblastic–meroblastic barrier in tetrapods. *J. Exp. Zool*. (*Mol. Dev. Evol*.). 310B: 1 – 7.

Elinson R.P. 2013. Metamorphosis in a Frog That Does Not have a tadpole. *Curr Top Dev Biol*. 103: 259 – 276.

Elinson R.P. and Beckham Y. 2002. Development in frogs with large eggs and the origin of amniotes. *Zoology*. 105: 105 – 117.

Elinson R.P. and del Pino E.M. 2012. Developmental diversity of amphibians. *Wiley Interdiscip Rev Dev Biol*. 1: 345 – 369.

Elinson R.P. and Fang H. 1998. Secondary coverage of the yolk by the body wall in the direct developing frog, *Eleutherodactylus coqui:* an unusual process for amphibian embryos. *Dev Genes Evol.* 208: 457 – 466.

Elinson R.P., Sabo M.C., Fisher C., Yamaguchi T., Orii H., and Nath K. 2011. Germ plasm in *Eleutherodactylus coqui*, a direct developing frog with large eggs. *EvoDevo*. 2: 20.

Engleka M.J., Craig E.J. and Kessler D.S. 2001. VegT activation of *Sox17* at the midblastula transition alters the response to nodal signals in the vegetal endoderm domain. *Dev. Biol*. 237: 159 – 172.

Fagotto F. and Gumbiner B.M. 1994. β-catenin localization during *Xenopus* embryogenesis: accumulation at tissue and somite boundaries. *Development*. 120: 3667 – 3679.

Fagotto F. 1999. The Wnt pathway in *Xenopus* development. In *Signaling Through Cell Adhesion* (ed. J.-L. Guan). pp. 303 – 356: CRC Press.

Fang H. and Elinson R.P. 1996. Patterns of distal-less gene expression and inductive interactions in the head of the direct developing frog *Eleutherodactylus coqui*. *Dev Biol*. 179: 160 – 172.

Fang H. and Elinson R.P. 1999. Evolutionary alteration in the anterior patterning: *otx2* expression in the direct developing frog *Eleutherodactylus coqui*. *Dev Biol*. 205: 233 – 239.

Fang H., Marikawa Y. and Elinson R.P. 2000. Ectopic expression of *Xenopus* Noggin RNA induces complete secondary body axes in embryos of the direct developing frog *Eleutherodactylus coqui*. *Dev Genes Evol*. 210: 21 – 27.

Faure S., Lee M.A., Keller T., ten Dijke P., and Whitman M. 2000. Endogenous patterns of TGF-β superfamily signaling during early *Xenopus* development. *Development*. 127: 2917 – 2931.

Feldman B., Concha M.L., Saude L., Parsons M.J., Adams R.J., Winson S.W., and Stemple D.L. 2002. Lefty antagonism of squint is essential for normal gastrulation. *Curr. Biol*. 12: 2129 – 2135.

Frost D.R. 2007. *Amphibian Species of the World: an Online Reference. Version 5.0.* American Museum of Natural History, New York.

Fukuda K. and Kikuchi Y. 2005. Endoderm development in vertebrates: fate mapping, induction and regional specification. *Develop. Growth. Differ*. 47: 343 – 355.

Funaba M., Zimmerman C.M., and Mathews L.S. 2002. Modulation of Smad2-mediated signaling by extracellular signal-regulated kinase. *J. Biol. Chem*. 277: 41361 – 41368.

Gamer L.W. and Wright C.V.E. 1995. Autonomous endodermal determination in *Xenopus*: regulation of expression of the pancreatic gene *Xlhbox8*. *Dev Biol*. 171: 240 – 251.

Gerhart J.C. 1980. Mechanisms regulating pattern formation in the amphibian egg and early embryo. In *Biological Regulation and Development!* (ed. Goldberger R. F.), pp. 131-315. New York: Plenum Press.

Germain S., Howell M., Esslemont G.M., and Hill C.S. 2000. Homeodomain and winged helix transcription factors recruit activated Smads to distinct promoter elements via a common Smad interaction motif. *Genes Dev*. 14: 435 – 451.

Ghosh D. and Seydoux G. 2008. Inhibition of transcription by the Caenorhabditis elegans germline protein PIE-1: genetic evidence for distinct mechanisms targeting initiation and elongation. *Genetics*. 178: 235 – 243.

Gotoh T., Kishimoto T., and Sible J.C. 2012. Phosphorylation of Claspin is triggered by the nucleocytoplasmic ratio at the *Xenopus laevis* midblastula transition. *Dev Biol*. 353: 302 – 308.

Guven- Ozkan T., Nishi Y., Robertson S.M., and Lin R. 2008. Global transcriptional repression in C. elegans germline precursors by regulated sequestration of TAF-4. *Cell.* 135: 149 – 160.

Graff J.M., Bansal A, and Melton D.A. 1996. Xenopus Mad proteins transducer distinct subsets of signals for the TGF-β superfamily. *Cell.* 85: 479 – 487.

Green J. 2002. Morphogen gradients, positional information, and *Xenopus*: interplay of theory and experiment. *Dev.Dyn*. 225: 392 – 408.

Grishin N.V. 2001.Mh1 domain of Smad is a degraded homing endonuclease. *J. Mol. Biol.* 307: 31 – 37.

Gristman K., Zhang J., Cheng S., Heckscher E. Talbot W.S., and Schier A.F. 1999. The EGF-CFC protein one-eyed pinhead is essential for nodal signaling. *Cell*. 97: 121 – 132.

Gurdon J.B. and Bourillot P.Y. 2001. Morphogen gradient interpretation. *Nature*. 413: 797 – 803.

Hahn S.A., Schutte M., Hoque A.T.M.S., Moskaluk C.A., da Costa L.T., Rosenblum E., Weinstein C.L., Fischer A., Yeo C.J., Hruban R.H., and Kern S.E. DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science*. 271: 350 – 353.

Hamada H., Meno C., Watanabe D., and Saijoh Y. 2002. Establishment of vertebrate left-right asymmetry. *Nat. Rev. Genet*. 3: 103 – 113.

Hanken J. 1992. Life history and morphological evolution. *J. Evol. Biol*. 5: 549 – 557.

Hanyu-Nakamura K., Sonobe-Nojima H., tanigawa A., Lasko P., and Nakamura A. 2008. *Drosophila* PGC protein inhibits P-TEFb recruitment to chromatin in primordial germ cells. *Nature*. 451: 730 – 733.

Harland R. and Gerhart J.C. 1997. Formation and function of Spemann's Organizer. *Annu Rev Cell Dev Biol*. 13: 611 – 667.

Hata K., Lo R.S., Wotton D., Langa G., Massagué J. 1997. Mutations increasing autoinhibition inactivate tumor suppressors Smad2 and Smad4. *Nature.* 388: 82 – 87.

Heasman J. 2006. Patterning the early *Xenopus* embryo. *Development*. 133:1205 – 1217.

Hedges S.B., Duellman W.E., Heinicke M.P. 2008. New World direct-developing frogs (Anura: Terrarana): Molecular phylogeny, classification, biogeog- raphy, and conservation. Zootaxa 1737. 182 pp. (© 2008 Magnolia Press)

Heinicke, M.P., Duellman, W.E. & Hedges, S.B. (2007) Major Caribbean and Central American frog faunas originated by oceanic dispersal. *Proc. Nat. Aca. Sci. USA*. 104: 10092–10097.

Heldin C.H., Miyazono K. and ten Dijke P. 1997. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature*. 390: 465 – 471.

Hemmati-Brivanlou A., and Melton D.A. 1992. A truncated activin receptor inhibits mesoderm induction and formation of axial structures in *Xenopus* embryos. *Nature*. 359: 609 – 614.

Hennen S. 1973. Competence of tests of early amphibian gastrula tissue containing nuclei of one species (Rana palustris) and cytoplasm of another (Rana pipens). *J Embryol Exp Morphol*. 29: 529.

Henry G.L., Brivanlou I.H., Kessler D.S., Hemmati-Brivanlou A., and Melton D.A. 1996. TGFβ signaling and pre-patterning in *Xenopus laevis* endodermal development. *Development*. 122: 1007 – 1015.

Henry G.L. and Melton D.A. 1998. *Mixer*, a homeobox gene required for endoderm development. *Science.* 281: 91 – 96.

Heyer J., Escalante-Alcalde D., Lia M., Boettinger E., Edelmann W., Stewart C.L. and Kucherlapati R. 1999. Postgastrulation Smad2-deficient embryos show defects in embryo turning and anterior morphogenesis. *Proc. Natl. Acad. Sci*. 96: 12595 – 12600.

Hill C.S. 2001. TGF-β signaling pathways in early Xenopus development. *Current Opinion in Genetics & Development*. 11: 533 – 540.

Hirata T., Yamanaka Y., Ryu S.L., Shimizu T., Yabe T., Hibi M., and Hirano T. 2000. Novel mix-family homeobox genes in zebrafish and their differential regulation. *Biochem. Biophys. Res. Commun*. 271: 603 – 609.

Ho D.M., Yeo C.-Y., and Whitman M. 2010. The role and regulation of GDF11in Smad2 activation during tailbud formation in the Xenopus embryo. *Mech Dev*. 127: 485 – 495.

Horb M.E. and Thomsen G.H. 1997. A vegetally localized T-box transcription factor in *Xenopus* eggs specifies mesoderm and endoderm and is essential for embryonic mesoderm formation. *Development*. 124: 1689 – 1698.

Horb M.E. and Slack J.M. 2001. Endoderm specification and differentiation in *Xenopus* embryos. *Dev Biol*. 236: 330–343.

Howard L., Rex M., Clements D., and Woodland H.R. 2007. Regulation of the *Xenopus* Xsox17α1 promoter by co-operating VegT and Sox17 sites. *Dev Biol.* 310: 402 – 415.

Howell M., Itoh F., Pierreux C.E., Valgiersdottir S., Itoh S., ten Dijke P. and Hill C.S. 1999. *Xenopus* Smad4β is the co-Smad component of the developmentally regulated transcription factor complexes responsible for induction of early mesodermal genes. *Dev Biol*. 214: 354 – 369.

Howell M., Inman G.J., and Hill C.S. 2002. A novel *Xenopus* Smad-interacting forkhead transcription factor (XFast-3) cooperates with XFast-1 in regulating gastrulation movement. *Development*. 129: 2823 – 2834.

Hu C.D., Chinenov Y., and Kerppola T.K. 2002. Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol. Cell.* 9: 789 – 798.

Hudson C., Clements D., Friday R.V., Scott D., and Woodland H.R. 1997. XSox17α and -β mediate endoderm formation in *Xenopus*. *Cell.* 91: 397 – 405.

Inman G.J., Nicolas F.J., and Hill C.S. 2002. Nucleocytoplasmic shuttling of Smad2, 3, and 4 permits sensing of TGFβ receptor activity. *Mol. Cell.* 10: 283 – 294.

Ishizuya-Oka A. and Ueda S. 1996. Apoptosis and cell proliferation in the *Xenopus* small intestine during metamorphosis. *Cell Tissue Res*. 286: 467 – 476.

Javelaud D. and Mauviel A. 2006. Interplay between Smad and map kinase signaling pathways. In *Smad signal transduction: Smads in proliferation, differentiation and disease* (ed. ten Dijke P. and Heldin C.H.), pp.317 – 334. Dordrecth: Springer.

Jennings D.H. and Hanken J. 1998. Mechanistic basis of life history evolution in anuran amphibians: Thyroid gland development in the direct-developing frog, *Eleutherodactylus coqui*. *Gen Comp Endo*. 111: 225 – 232.

Joglar R. 1998. Los Coquíes de Puerto Rico: Su Historia Natural y Conservación. Editorial de la Universidad de Puerto Rico. 232pp. San Huan, Peurto Rico.

Jones C.M., Kuehn M.R., Hogan B.L., Smith J.C., and Wright C.V. 1995. Nodal-related signals induce axial mesoderm and dorsalize mesoderm during gastrulation. *Development*. 121: 3651 – 3662.

Joseph E.M. and Melton D.A. 1998. Mutant Vg1 ligands disrupt endoderm and mesoderm formation in Xenopus embryos. *Development*. 125: 2677 – 2685.

Just A., Jorgensen H. and Fernandez J.A. 1981. The digestive capacity of the caecum-colon and the value of the nitrogen absorbed from the hind gut for protein synthesis in pigs. *Br. J. Nutr*. 46: 209 – 219.

Karadge U. 2012. Characterization of the nutritional endoderm in *Eleutherodactylus coqui*. Ph.D. Thesis. Duquesne University, 232 pp.

Kawabata M., Inoue H., Hanyu A., Imamura T. and Miyazono K. 1998. Smad proteins exist as monomers in vivo and undergo homo- and hetero-oligomerization upon activation by serine/threonine kinase receptors. *EMBO J.* 17: 4056 – 4065.

Kimelman D. 2006. Mesoderm induction: from caps to chips. *Nature*. 7: 360 – 372.

Kisielewska J. and Blow J.J. 2012. Dynamic interactions of high Cdt1 and geminin levels regulate S phase in early *Xenopus* embryos. *Development*. 139: 63 – 74.

King M.L., Messitt T.J., and Mowry K.L. 2005. Putting RNAs in the right place at the right time: RNA localization in frog oocyte. *Biol Cell*. 97: 19 – 23.

Kinoshita N., Minshull J., and Kirschner M.W. 1995. The identification of two novel ligands of the FGF receptor by a yeast screening method and their activity in *Xenopus* development. *Cell*. 83: 621 – 630.

Kikuchi Y., Agathon A., Alexander J., Thisse C., Waldron S., Yelon D., Thisse B., and Stainier D.Y. 2001. *casanova* encodes a novel Sox-related protein necessary and sufficient for endoderm formation in zebrafish. *Genes & Dev*. 15: 1493 – 1505.

Kofron M., Demel T., Xanthos J., Lohr J., Sun B., Sive H., Osada S., Wright C., Wylie C., and Heasman J. 1999. Mesoderm induction in Xenopus is a Zygotic Event regulated by Maternal VegT Via TGFbeta Growth Factors. *Development*. 126: 5759 – 5770.

Korzh V. 2009. Before maternal-zygotic transition ... There was morphogenetic function of nuclei. *Zebrafish*. 6: 295 – 302.

Kumar A., Novoselov V., Celeste A.J., Wolfman N.M., ten Dijke P. and Kuehn M.R. 2001. Nodal signaling uses activin and transforming growth factor-beta receptor-regulated Smads. *J. Biol. Chem*. 276: 656 – 661.

Kurokawa M., Mitani K., Irie K., Matsuyama T., Takahashi T., Chiba S., Yazaki Y., Matsumoto K., and Hirai H. 1998. The oncoprotein Evi-1 represses TGF-β signalling by inhibiting Smad3. *Nature*. 394: 92 – 96.

Kusanagi K., Kawabata M., Mishima H.K., and Miyazono K. 2001. α-helix 2 in the animo-terminal mad homology domain 1 is responsible for specific DNA binding of Smad3. *J. Biol. Chem*. 276: 28155 – 28163.

Kraus F., Campbell E.W., Allison A., and Pratt T.1999. *Eleutherodactylus* frog introduction to Hawaii. *Herp. Rev*. 30: 21 – 25.

Kretzschmar M., Doody J., and Massagué J. 1997. Opposing BMP and EGF signaling pathways converge on the TGFβ family mediator Smad1. *Nature*. 389: 618 – 622.

Kretzschmar M., Doody J., Timokhina I., and Massagué J. 1999. A mechanism of repression of TGF-β/Smad signaling by oncogenic Ras. *Genes Dev*. 13: 804 – 816.

Kurisaki A., Kose S., Yoneda Y., Heldin C.H., and Moustakas A. 2001. Transforming growth factor-β induces nuclear import of Smad3 in an importinβ1 and Ran-dependent manner. *Mol. Biol. Cell*. 12: 1079 – 1091.

Labbé E., Silvestri C., Hoodless P. A., Wrana J. L. and Attisano, L. 1998. Smad2 and Smad3 positively and negatively regulate TGF beta- dependent transcription through the forkhead DNA-binding protein FAST2. *Mol. Cell* 2, 109-120.

LaBonne C., and Whitman M. 1994. Mesoderm induction by activin requires FGF-mediated intracellular signals. *Development*. 120: 463 – 472.

Lagna G.A., Hata A., Hemmati-Brivanlou, and Massagué J. 1996. DPC4 and SMAD proteins in TGFβ signaling pathways. *Nature*. 383: 832 – 836.

Lee M.A., Heasman J., and Whitman M. 2001. Timing of endogenous Activin-like signals and regional specification of the *Xenopus* embryo. *Development.* 128: 2939 – 2952.

Lemaire P., Darras S., Caillol D., and Kodjabachian L. 1998. A role for the vegetally expressed *Xenopus* gene *Mix.1* in endoderm formation and in the restriction of mesoderm to the marginal zone. *Development*. 125: 2371 – 2380.

LeSueus J.A. and Graff J.G. 1999. Spemann organizer activity of Smad10. *Development*. 126: 137 – 146.

Liu F., Pouponnut C., and Massagué J. 1997. Dual role of the Smad4/DPC4 tumor suppressor in TGF-β-inducible transcriptional complex. *Genes Dev*. 11: 3157 – 3167.

Liu X., Luo M., Xie W., Wells J.M., Goodheart M.J., and Engelhardt J.F. 2010. Sox17 modulates Wnt3A/beta-catenin-mediated transcriptional activation of the *Lef-1* promoter. *Am J Physiol Lung Cell Mol Physiol*. 299: 694 – 710.

Loose M. and Patient R.K. 2004. A genetic regulatory network for *Xenopus* mesendodermal formation. *Dev Biol*. 271: 467 – 478.

Lustig K.D., Kroll K.L., Sun E.E., and Kirschner M.W. 1996. Expression cloning of a *Xenopus* T-related gene (*Xombi*) involved in mesodermal patterning and blastopore lip formation. *Development*. 122: 4001 – 4012.

Lynn W.G. 1948. The effects of thiourea and phenylthiourea upon the development of *Eleutherodactylus ricordii*. *Biol Bull*. 94: 1 – 15.

Martello G., Zacchigna L., Inui M., Montagner M., Adorno M., Mamidi A., Morsut L., Soligo S., Tran U., Dupont S., Cordenonsi M., Wessely O., and Piccolo S. 2007. MicroRNA control of Nodal signaling. *Nature*. 449, 183 – 187.

Macías-Silva M., Abdollah S., Hoodless P.A., Pirone R., Attisano L., and Wrana J.L. 1996. MADR2 is a substrate of the TGFbeta receptor and its phosphorylation is required for nuclear accumulation and signaling. *Cell*. 87: 1215 – 1224.

Massagué J. 1998. TGF-beta signal transduction. *Ann. Rev. Biochem*. 67: 753 – 791.

Massagué J. 2003. Integration of Smad and MAPK pathways: a link and a linker revisited. *Genes Dev*. 17: 2993 – 2997.

Massagué J., Seoane J., and Wotton D. 2005. Smad transcription factors. *Genes Dev*. 19: 2783 – 2810.

Mathews L.S. and Vale W.W. 1991. Expression cloning of an Activin receptor, a predicted transmembrane serine kinase. *Cell*. 65: 973 – 982.

Masuyama N., Hanafusa H., Kusakabe M., Shibuya H., and Nishida E. 1999. Identification of two Smad4 proteins in *Xenopus*. Their common and distinct properties. *J. Biol. Chem*. 274: 12163 – 12170.

Mead P.E., Brivanlou I.H., Kelly C.M., and Zon L.I. 1996. BMP-4-responsive regulation of dorsal-ventral patterning by the homeobox protein Mix.1. *Nature*. 382: 357 – 360.

Melton D.A. 1987. Translocation of a localized maternal mRNA to the vegetal pole of *Xenopus* oocyte. *Nature*. 328: 80 – 82.

Meno C., Takeuchi J., Sakuma R., Koshiba-Takeuchi K., Ohishi S., Saijoh Y., Miyazaki J.I., Peter ten Dijke P.T., Ogura T., and Hamada H. Diffusion of Nodal Signaling Activity in the Absence of the Feedback Inhibitor Lefty2. *Dev. Cell*. 1: 127 – 138.

Mercola M. and Levin M. 2001. Left-right asymmetry determination in vertebrates. *Annu. Rev. Cell Dev. Biol*. 17: 779 – 805.

Moustakas A., Souchelnytskyi A., and Heldin C.H. 2001. Smad regulation in TGF-β signal transduction. *J Cell Sci*. 114: 4359 – 4369.

Newport J. and Kirshner M. W. 1982. A major developmental transition in early Xenopus embryos: I. Characterization and Timing of cellular changes at the midblastula stage. *Cell*. 30: 675 – 686.

Niakan K.K., Ji H., Maehr R., Vokes S.A., Rodolfa K.T., Sherwood R.I., Yamaki M., Dimos J.T., Chen A.E., Melton D.A., McMahon A.P., and Eggan K. 2010. Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. *Genes Dev*. 24: 312 – 326.

Nieuwkoop P.D. and Faber J. 1994. Normal table of *Xenopus laevis*. Garland publishing., New York.

Ninomiya H., Zhang Q., Elinson R.P. 2001. Mesoderm Formation in *Eleutherodactylus coqui*: Body Patterning in a frog with a Large Egg. *Dev Biol*. 236: 109 – 123.

Nishimatsu S., Suzuki A., Shoda A., Murakami K., and Ueno N. 1992. Genes for bone morphogenetic proteins are differentially transcribed in early amphibian embryos. *Biochem. Biophys. Res. Commun*. 186: 1487 – 1495.

Nomura M. and Li E. 1998. Smad2 role in mesoderm formation, left-right patterning and craniofacial development. *Nature*. 393: 786 – 790.

Osada S.I. and Wright C.V. 1999. *Xenopus* nodal-related signaling is essential for mesendoderm patterning during early embryogenesis. *Development*. 126: 3229 – 3240.

Packard M.J., Jennings D.H. and Hanken J. 1998. Growth and uptake of mineral by embryos of the direct-developing frog *Eleutherodactylus coqui*. *Comp. Biochem. Physiol.* 113A (4): 343 – 349.

Palancade B., Bellier S., Almouzni G., and Bensaude O. 2001. Incomplete RNA polymerase II phosphorylation in *Xenopus laevis* early embryos. *J. Cell Sci*. 114: 2483 – 2489.

Patternson E.S., Addis R.C., Shamblott M.J., and Gearhart J.D. 2008. SOX17 directly activates Zfp202 transcription during in vitro endoderm differentiation. *Physiol Genomics*. 34: 277 – 284.

Patturajan M., Schulte R.J., Sefton B.M., Berezney R., Vincent M., Bensaude O., Warren S.L., and Corden J.L. 1998. Growth-related changes in phosphorylation of yeast RNA polymerase II. *J. Biol. Chem*. 273: 4689 – 4694.

Pearce J.J. and Evans M.J. 1999. *Mm1*, a mouse mix-like gene expressed in the primitive streak. *Mech. Dev*. 87: 189 – 192.

Pérez O., Benítez M.-S., Nath K., Heasman J., del Pino E.M., and Elinson R.P. 2007. Comparative analysis of *Xenopus VegT*, the meso-endodermal determinant, identifies an unusual conserved sequence. *Differentiation*. 75: 559 – 565.

Phatnani H.P. and Greenleaf A.L. 2006. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev*. 20: 2922 – 2936.

Piccolo S., Agius E., Leyns L., Bhattacharyya S., Grunz H., Bouwmeester, T., and DeRobertis E.M. 1999. The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. *Nature*. 397: 707 – 710.

Pierreus C.E., Nicolas F.J., and Hill C.S. 2000. Transforming growth factor β-independent shuttling of Smad4 between the cytoplasm and nucleus. *Mol. Biol. Cell*. 8: 9041 – 9054.

Price D.H. 2000. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol. Cell. Biol*. 20: 2629 – 2634.

Pogoda, H.M., Solnica-Krezel L., Driever W., and Meyer D. 2000. The zebrafish forkhead transcription factor FoxH1/Fast1 is a modulator of Nodal signaling required for organizer formation. *Curr. Biol*. 10: 1041 – 1049.

Randall R.A., Germain S. Inman G.J., Bates P.A., and Hill C.S. 2002. Different Smad2 partners bind a common hydrophonic pocket in Smad2 via a defined proline rich motif. *EMBO J*. 21: 145 – 156.

Randall R.A., Howell M., Page C.S., Daly A., Bates P.A., and Hill C.S. 2004. Recognition of phosphorylated Smad-containing complexs by a novel Smad interaction motif. *Mol. Cell. Biol.* 24: 1106 – 1121.

Rebagliati M.R., Weeks D.L., Harvey R.P., and Melton D.A. 1985. Identification and cloning of localized maternal mRNAs from *Xenopus* eggs. *Cell*. 42:769 – 777.

Reissman E., Jornvall H., Blokzijl A., Andersson O., Chang C., Minchiotti G., Persico M.G., Ibanez C.F., and Brivanlou A.H. 2001. The orphan receptor ALK7 and the Activin receptor ALK4 mediate signaling by Nodal proteins during vertebrate development. *Genes Dev.* 15: 2010 – 2022.

Rex M., Hilton E. and Old R. 2002. Multiple interactions between maternally-activated signalling pathways control *Xenopus* nodal-related genes. *Int. J. Dev. Biol*. 46: 217 – 226.

Robb L., Hartley L., Begley C.G., Brodnicki T.C., Copeland N.G., Gilbert D.J., Jenkins N.A., and Elefanty A.G. 2000. Cloning, expression analysis, and chromosomal localization of murine and human homologue of *Xenopus Mix* gene. *Dev Dyn*. 219: 497 – 504.

Roth G. and Wake D.B. 1989. Conservatism and innovation of the evolution of feeding in vertebrates. In *Complex organismal functions: integration and evolution in vertebrates* (ed. Wake D.B. and Roth G.) pg. 7 – 21. Wiley Intersciences, Chichester.

Rosa F.M. 1989. *Mix.1*, a homeobox mRNA inducible by mesoderm inducers, is expressed mostly in the presumptive endodermal cells of *Xenopus* embryos. *Cell*. 57: 965 – 974.

Ross S. and Hill C.S. 2008. How the Smads regulate transcription. *Int J Biochem Cell Biol.* 40: 383 – 408.

Saka Y., Hagemann A.I., Piepenburgh O., and Smith J.C. 2007. Nuclear accumulation of Smad complexes occur only after the midblastula transition in *Xenopus*. *Development*. 134: 4209 – 4218.

Saka Y., Hagemann A.I., and Smith J.C. 2008. Visualizing protein interaction by bimolecular fluorescence complementation in *Xenopus*. *Methods*. 45: 192 – 195.

Salthe S.N. and Mecham J.S. 1974. Reproductive and courtship patterns. In B. Lofts (ed). *Physiology of the Amphibia*, *Vol. 2*. pp. 309 – 521. Academic press, New York.

Schier A.F. 2003. Nodal Signaling in Vertebrate Development. *Annu. Rev. Cell Dev. Biol.* 19: 586 – 621.

Schier A.F. 2009. Nodal Morphogens. *Cold Spring Herb Perspect Biol doi:* 10.1101/cshperspect.a003459.

Schier A.F. and Shen M.M. 2000. Nodal Signaling in Vertebrate Development. *Nature.* 403: 385 – 389.

Schier A.F. and Talbot W.S. 2005. Molecular genetics of axis formation in zebrafish. *Annu Rev Genet.* 39: 561 – 613.

Schlosser G., Kintner C., and Northcutt R.G. 1999. Loss of ectodermal competence for lateral line placode formation in the direct-developing frog *Eleutherodactylus coqui*. *Dev Biol*. 213: 354 – 369.

Schohl A. and Fagotto F. 2002. β catenin, MAPK and Smad signaling during early *Xenopus* development. *Development*. 129: 37 – 52.

Seki Y., Yamaji M., Yabuta Y., Sano M., Shigeta M., matsui Y., Saga Y., Tachibana M., Shinkai Y., and Saitou M. 2007. Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. *Development*. 134: 2627 – 2638.

Seydoux G. and Dunn M.A. 1997. Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of *Caenorhabditis elegans* and *Drosophila melanogaster*. *Development*. 124: 2191 – 2201.

Seydoux G., Mello C.C., Pettitt J., Wood W.B., Priess J.R., and Fire A. 1996. Repression of gene expression in the embryonic germ lineage of *C. elegans*. *Nature*. 382: 713 – 716.

Sinner D., Kirilenko P., Rankin S., Wei E., Howard E., Kofron M., Heasman J., Woodland H.R., and Zorn A.M. 2006. Global analysis of the transcriptional network controlling *Xenopus* endoderm formation. *Development*. 133: 1955 – 1966.

Sinner D., Raskin S., Lee M., and Zorn A.M. 2004. Sox17 and Beta-catenin cooperate to regulate the transcription of endodermal genes. *Development*. 131: 1955 – 1966.

Signoret J. and Lefresne, J. (1971). Contribution a I'etude de la segmentation de I'oef d'axolotl. Definition de la transition blastuleenne. *Ann Embryol Morphogen*. 4: 113 – 123.

Shen M.M. 2007. Nodal Signaling: Developmental roles and regulation. *Development.* 134: 1023 – 1034.

Shen M.M. and Schier A.F. 2000. The EGF-CFC gene family in vertebrate development. *Trend Genet*. 16: 303 – 309.

Shi Y. 2001. Structural insights on Smad function in TGF-β signaling. *Bioessays*. 23: 223 – 232.

Shi Y. and Massagué J. 2003. Mechanisms of TGF-β signaling from cell membrane to the nucleus. *Cell.* 113: 685 – 700.

Shibuya E.K. and Masui Y. 1988. Stabilization and enhancement of primary cytostatic factor (CFS) by ATP and NaF in amphibian egg cytosols. *Dev Biol*. 129: 253 – 264.

Shimizu A., Kato M., Nakao A., Imamura T., ten Dijke P., Heldin C.H., Kawabata M., Shimada S., and Miyazono K. 1998. Identification of receptors and Smad proteins involved in activin signalling in a human epidermal keratinocyte cell line. *Genes Cells*. 3: 125 – 134.

Shimuta K., Nakajo N., Uto K., Hayano Y., Okazaki K., and Sagata N. 2002. Chk1 is activated transiently and targets Cdc25A for degradation at the *Xenopus* midblastula transition. *EMBO J*. 21: 3694 – 3703.

Shiokawa K. 2012. Maternally-preset program of apoptosis and caspases involved in execution of the apoptosis at midblastula transition (MBT) but not before in *Xenopus laevis* embryogenesis. *Adv Biosci Biotech*. 3: 751 – 769.

Shivdasani R.A. 2002. Molecular regulation of vertebrate early endoderm development. *Dev. Biol*. 249: 191 – 203.

Shook D.R. and Keller R. 2008. Epithelial type, ingression, blastopore architecture and the evolution of chordate mesoderm morphogenesis. *J Exp Zool* (*Mol Dev Evol*). 310B: 85 – 110.

Singamsetty S. and Elinson R.P. 2010. Novel regulation of yolk utilization by thyroid hormone in embryos of direct developing frog *Eleutherodactylus coqui*. *Evol. Dev*. 12: 437 – 448.

Sirotkin H.I., Gates M.A., Kelly P.D., Schier A.F., and Talbot W.S. 2000. Fast1 is required for the development of dorsal axial structures in zebrafish. *Curr. Biol*. 10: 1051 – 1054.

Skirkanich J., Luxardi G., Yang J., Kodjabachian L., and Klein P.S. 2011. An essential role for transcription before the MBT in Xenopus laevis. *Dev Biol*. 357: 478 – 491.

Smith W.C., McKendry R., Ribisi S., Jr., and Harland R.M. 1995. A nodal-related gene defines a physical and functional domain within the Spemann organizer. *Cell*. 82: 37 – 46.

Souchelnytskyi S., Tamaki K., Engstrom U., Wernstedt C., ten Dijke P., and Heldin C.H. 1997. Phosphorylation of Ser^{465} and Ser^{467} in the C terminus of Smad2 mediates interaction with Smad4 and is required for transforming growth factor-β signaling. *J. Biol. Chem*. 272: 28107 – 28115.

Stainier D.Y.R. 2002. A glimpse into the molecular entrails of endoderm formation. *Genes Dev.* 16: 893 – 907.

Stennard F. 1998. Xenopus differentiation: VegT gets specific. *Curr Biol*. 8: R928 – R930.

Stennard F., Carnac G., and Gurdon J.B. 1996. The *Xenopus* T-box gene, *Antipodean*, encodes a vegetally localized maternal mRNA and can trigger mesoderm formation. *Development*. 122: 4179 – 4188.

Strome S. and Lehmann R. 2007. Germ versus soma decisions, lessons from flies and worms. *Science*. 316: 392 – 393.

Sun B.I., Bush S.M., Collin-Racie L.A., LaVallie E.R., DiBlasio-Smith, E.A., Wolfman N.M., McRoy J.M., and Sive H.L. 1999. Derriere: A TGFβ family member required for posterior development in *Xenopus*. *Development*. 126: 1467 – 1482.

Sutherland H. and Bickmore W.A. 2009. Transcription factories: gene expression in unions? *Nat Rev Gen*. 10: 457 – 466.

Tada M., Casey E.S., Fairclough L., and Smith J.C. 1998. Bix1, a direct target of *Xenopus* T-box genes, causes formation of ventral mesoderm and endoderm. *Development*. 125: 3997 – 4006.

Tada K., Inoue H., Ebisawa T., Makuuchi M., Kawabata M., Imamura T., and Miyazono K. 1999. Region between α-helices 3 and 4 of the mad homology 2 domain of Smad4: functional roles in oligomer formation and transcriptional activation. *Genes Cell*. 4: 731 – 741.

Tadros W. and Lipshitz H.D. 2009. The maternal-to-zygotic transition: a play in two acts. *Development*. 136: 3033 – 3042.

Takashi S., Yokota C., Takano K., Tanegashima K., Onuma Y., Goto J.I., and Asashima M. 2000. Two novel nodal-related genes initiate early inductive events in *Xenopus* Nieuwkoop center. *Development*. 127: 5319 – 5329.

Takenoshita S., Mogi A., Nagashima M., Yang K., Yagi K., Hanyu A., Nagamachi Y., Miyazono K., and Hagiwara K. 1998. Characterization of the MADR2/Smad2 gene, a human Mad homolog responsible for the transforming growth factor-beta and activin signal transduction pathway. *Genomics*. 48: 1 – 11.

Takeuchi M, Takahashi M., Okabe M., and Aizawa M. 2009. Germ layer patterning in bichir and lamprey; an insight into its evolution in vertebrates. *Dev Biol*. 332: 90 – 102.

ten Dijke P., yamashita H., Ichijo H., Franzen P., Laiho M., Miyazono K., and Heldin C.H. 1994. Characterization of Type I Receptors for Transforming Growth factor Beta and Activin. *Science*. 264: 101 – 104.

Thomsen G.H. and Melton D.A. 1993. Processed Vg1 Protein is an axial mesoderm inducer in *Xenopus*. *Cell.* 74: 433 – 441.

Timinszky G., Bortfeld M., and Ladurner A.G. 2008. Repression of RNA polymrase II transcription by a *Drosophila* oligopeptide. *PLoS ONE* 3(6): e2506. doi:10.1371/journal.pone.0002506

Tomioka M., Miya T., and Nishida H. 2002. Repression of zygotic gene expression in the putative germ line cells in ascidian embryos. *Zool. Sci*. 19: 49 – 55.

Townsend D.S. and Stewart M.M. 1985. Direct development in *Eleutherodactylus coqui (Anura: Leptodactylidae):* A staging table. *Copeia.* 1985: 423 – 436.

Tremblay K.D., Hoodless P.A., Bikoff E.K., Robertson E.J. 2000. Formation of the definitive endoderm in mouse is a Smad2 dependent process. *Development.* 127: 3079 – 3090.

Tsuchida K., Mathews L.S., and Vale W.W. 1993. Cloning and Characterization of a Transmembrane Serine Kinase that Acts as an Activin Type I Receptor. *Proc. Natl. Acad. Sci. USA*. 90: 11242 – 11246.

Tsukazaki T., Chiang T.A., Davison A.F. Attisano L., and Wrana J.L. 1998. SARA, a FYVE domain protein that recruits Smad2 to the TGFβ receptor. *Cell*. 95: 779 – 791.

Valett B.B. and Jameson D.L. 1961. The embryology of Eleutherodactylus augusti latrans. *Copea*. 1961: 103 – 109.

Van-Dorren M., Williamson A.L., and Lehmann R. 1998. Regulation of zygotic gene transcription in Drosophila primordial germ cells. *Curr. Biol*. 8: 243 – 246.

Velo-Antón G., Burrowes P.A., Joglar R.L., Martínez-Solano I., Beard K.H., and Parra-Olea G. 2007. Phytogenetic study of *Eleutherodactylus coqui* (Anura: Leptodactylidae) reveals deep genetic fragmentation in Puerto Rico and pinpoints origins of Hawaiian populations. *Mol Phylogen Evol*. 45: 716 – 728.

Venkatarama T., Lai F., Luo X., Zhou Y., Newman K., and King ML. (2010). Repression of zygotic gene expression in the *Xenopus* germline. *Development*. 137, 651 – 659.

Vize P.D. 1996. DNA sequences mediating the transcriptional response of the *Mix.2* homeobox gene to mesoderm induction. *Dev. Biol*. 177: 226 – 231.

Wake M.H. 1980. The reproductive biology of *Nectophrynoides malcomi* (Amphibia: Bufonidae), with comments on the evolution of reproductive modes in the genus Nectophrynoides. *Copia*. 1980: 193 – 209.

Wake M.H. 1993. Evolution of oviductal gestation in amphibians. *J. Exp. Zool*. 266: 394 – 413.

Watanabe M., Masuyama N., Fukuda M., and Nishida E. 2000. Regulation of intracellular dynamics of Smad4 by its leucine-rich nuclear export signal. *EMBO Rep*. 1: 176 – 182.

Watanabe M. and Whitman M. 1999. FAST-1 is a key maternal effector of mesoderm inducers in early *Xenopus* embryo. *Development.* 126: 5621 – 5634.

Weber H., Symes C.E., Walmsley M.E., Rodaway A.R.F., and Patient R.K. 2000. A role for GATA5 in *Xenopus* endoderm specification. *Development.* 127: 4345 – 4360.

Weeks D.L. and Melton D.A. 1987. A maternal mRNA localized to the vegetal hemisphere in Xenopus eggs codes for a growth factor related to TGF-Beta. *Cell.* 51: 861 – 867.

Weinstein M., Yang X., Li C., Xu X., Gotay J., and Deng C.X. 1998. Failure of egg cylinder elongation and mesoderm induction in mouse embryos lacking the tumor suppressor Smad2. *Proc. Natl. Acad. Sci. USA*. 95: 9378 – 9383.

Willis S.A., Zimmerman C.M., Li L.I. and Mathews L.S. 1996. Formation and Activation by Phosphorylation of Activin Receptor Complexes. *Mol Endocrinol*. 10: 367 – 379.

White J.A. and Heasman J. 2008. Maternal Control of Pattern Formation in Xenopus laevis. *J Exp Zool* (*Mol Dev Evol*). 310B: 73 – 84.

Whitman M. 1998. Smad2 and early developmental signaling by the TGF-β superfamily. *Genes Dev.* 12: 2445 – 2462.

Whitman M. 2001. Nodal signaling in early vertebraye embryos: themes and variations. *Dev. Cell*. 1: 607 – 617.

Wu M.Y. and Hill C.S. 2009. TGF-β superfamily signaling in embryonic development and homeostasis. *Dev Cell*. 16: 329 – 343.

Wu R. –Y., Zhang Y., Feng X. –H. and Derynck R. 1997. Heteromeric and homomeric interactions correlate with signaling activity and functional cooperativity of Smad3 and Smad4/DPC4. *Mol. Cell. Biol*. 17: 2521 – 2528.

Xanthos J.B., Kofron M., Wylie C., and Heasman J. 2001. Maternal VegT is the initiator of a molecular network specifying endoderm in *X. laevis. Development*. 128: 167 – 180.

Xiao Z., Liu X., Henis Y.I., and Lodish H.F. 2000. A distinct nuclear localization signal in the N-terminus of Smad3 determines its ligand-induced nuclear translocation. *Proc. Natl. Acad. Sci. USA*. 97: 7853 – 7858.

Xiao Z., Watson N., Rodriguez C., and Lodish H.F. 2001. Nucleocytoplasmic shuttling of Smad1 conferred by its nuclear localization and nuclear export signals. *J. Biol. Chem*. 276: 39404 – 39410.

Xu J. and Attisano L. 2000. Mutation in the tumor suppressors Smad2 and Smad4 activate transforming growth factor-β signaling by targeting Smads to the ubiquitin-proteasome pathway. *Proc. Natl. Acad. Sci. USA*. 97: 4820 – 4825.

Xu L., Chen Y.G., and Massagué J. 2000. The nuclear import function of Smad2 is masked by SARA and unmasked by TGFβ-dependent phosphorylation. *Nat. Cell. Biol*. 2: 559 – 562.

Xu L., Kang Y., Col S., and Massagué J. 2002. Smad2 nucleocytoplasmic shuttling by nucleoporins CAN/Nup214 and Nup153 feeds TGF-β signaling complexes in the cytoplasm and nucleus. *Mol. Cell.* 10: 271 – 282.

Yabe S.I., Tanegashima K., Haramoto Y., Takahashi S., Fujii T., Kozuma S., Taketani Y., and Asashima M. 2003. FRL-1, a member of the EGF-CFC family, is essential for neural differentiation in *Xenopus* early development. *Development*. 130: 2071 – 2081.

Yagi K., Goto D., Hamamoto T., Takenoshita S., Kato M., and Miyazono K. 1999. Alternatively spliced variant of Smad2 lacking exon 3. Comparison with wild type Smad2 and Smad3. *J. Biol. Chem*. 274: 703 – 709.

Yasuo H. and Lemaire P. 1999. A two-step model for the fate determination of presumptive endodermal blastomeres in *Xenopus* embryos. *Curr Biol*. 9: 8869 – 8879.

Yeo C. and Whitman M. 2001. Nodal signals to Smads through Cripto-dependent and Cripto- independent mechanisms. *Mol Cell.* 7: 949 – 957.

Zhang J., Houston D.W., King M.L., Payne C., Wylie C., and Haesman J. 1998. The role of maternal VegT in establishing the primary germ layers in *Xenopus* embryos. *Cell*. 94: 515 – 524.

Zhang J. and King M.L. 1996. *Xenopus* VegT RNA is localized to the vegetal cortex during oogenesis and encodes a novel T-box transcription factor involved in mesodermal patterning. *Development*. 122: 4119 – 4129.

Zhang J., Talbot W.S., and Schier A.F. 1998. Positional cloning identifies zebrafish one-eyed pinhead as a permissive EGF-related ligand required during gastrulation. *Cell.* 92: 241 – 251.

Zhang Y., Musci T., and Derynck R. 1997. The tumor-supressor Smad4/DPC4 as a central mediator of Smad function. *Curr. Biol*. 7: 270 – 276.

Zhang C., Basta T., Fawcett S.R., and Klymkowsky M.W. 2005a. SOX7 is an immediate-early target of VegT and regulates Nodal-related gene expression in *Xenopus. Dev Biol.* 278: 526 – 541.

Zhang C., Basta T., and Klymkowsky M.W. 2005b. SOX7 and SOX18 are essential for cardiogenesis in *Xenopus. Dev Dyn.* 234: 878 – 891.

Zhang C. and Klymkowsky M.W. 2007. The Sox axis, Nodal signaling, and germ layer specification. *Differentiation.* 75: 536 – 545.

Zorn A.M., Barish G.D., Williams B.O., Lavender P., Klymkowsky M.W., and Varmus H.E. 1999. Regulation of Wnt signaling by Sox proteins: XSox17 alpha/beta and XSox3 physically interact with beta-catenin. *Mol Cell*. 4: 487 – 498.