

**Conserved Principles of The
Pluripotency and Primordial Germ
Cell Gene Regulatory Networks:
From Basal Vertebrates to Mammals**

By Luke Alexander Simpson

BSc, MRes



**Thesis submitted to the University of Nottingham for the degree of Doctor
of Philosophy, September 2021**

School of Life Sciences

University of Nottingham

United Kingdom

Declaration

I declare that this thesis is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the acknowledgments or main text, and that has not previously been presented as part of any other degree.

Abstract

Pluripotency defines the unlimited potential of cells in the primitive ectoderm of vertebrate embryos, from which all adult somatic cells and germ cells are derived. Understanding how the programming of pluripotency evolved has been obscured by the study of early development in models from lower vertebrates in which pluripotency is not conserved. Using Axolotl to model the development of the tetrapod ancestor, I examined the role of two transcription factors in early development associated with pluripotency in human embryonic stem cells (hESC), NANOG and ELK1.

In chapter three I investigated how the core pluripotency factor *nanog* programs pluripotency in axolotl development. *Nanog* marks the pluripotent domain in mammals and NANOG knockout results in the loss of pluripotency *in vitro* and the developmental arrest prior to the establishment of the epiblast *in vivo*. Here I show that in axolotl animal caps (AC), NANOG synergizes with NODAL activity and the epigenetic modifying enzyme DPY30 to direct the deposition of H3K4me3 in chromatin prior to the waves of transcription required for lineage commitment and developmental progression. I show that the interaction of NANOG and nodal with DPY30 is required to direct development downstream of pluripotency and this is conserved in axolotls and humans. These data also demonstrate that the interaction of NANOG and NODAL signaling represents the basal state of vertebrate pluripotency.

In chapter 4, I explored the role of ELK1 in germ layer formation and PGC specification in early development. ELK1 is a prototypical ETS domain transcription factor, conserved across metazoans and known to govern cell-fate decisions, acting downstream of fibroblast growth factor (FGF) signalling.

ELK1 has also been identified as a factor essential to maintain the pluripotent state of hESC. Here I show that the lateral and intermediate mesoderm in axolotl embryos is diverted to somitic tissue in response to ELK1 knockdown, suggesting that the regulation of mesodermal differentiation by ELK1 may be conserved from amphibians to humans. Further, I uncover a novel role, showing that it is required for the formation of primordial germ cells (PGCs) in axolotl embryos. Together, these results establish ELK1 as a central player in the evolution of vertebrate mesoderm. Moreover, they align with the concept that as mechanisms of PGC specification evolve they can lead to fundamental changes in somatic development.

Acknowledgements

Throughout my studentship I have been supported by many people. I would like to thank them for their contributions.

Firstly, I want to thank Andrew Johnson. It saddens me that you will never get to read this, your mentorship during the course of my studentship will stay with me. Your unwavering passion, enthusiasm and creativity was infectious. Thank you for all the arguments, debates and conversations, your insights and questioning nature has made me a better scientific thinker. I am proud to have called you my supervisor and my friend.

I want to extend my deepest gratitude to Matt Loose and Ramiro Alberio for their co-supervision during the writing of this thesis, I think very highly of you both and am glad for the opportunity to learn from each of you.

Matt, you have provided a massive amount of support for many years here in Nottingham. You were the ego, to AJ's Id. You have always provided much needed push-back, you have been a grounding influence and a reminder that 'the details' do matter! Ramiro, your valuable input throughout my research has always been appreciated. I look forward to working with you both closely in the future.

Contents

Declaration.....	2
Abstract.....	3
Acknowledgements.....	5
Contents.....	6
List of Figures	9
List of Abbreviations	13
1 Introduction	16
1.1 Literature Review	16
1.2 Cell potency and pluripotency.....	16
1.3 Early mouse development.....	17
1.4 Early human development	19
1.1.3 The core pluripotency GRN	22
1.5 Conservation of the pluripotency GRN in mammals	36
1.6 Pluripotency in non-mammals.....	40
1.7 The germline-soma relationship.....	53
1.8 Aims and objectives.....	56
2 Materials and Methods.....	59
2.2 Axolotl embryos and explants	59
2.3 Morpholino and RNA microinjections	59
2.4 Western blotting.....	60

1.9	Immunofluorescence.....	61
2.5	First-Strand cDNA synthesis.....	62
2.6	Quantitative PCR.....	62
2.7	RNA-seq analysis.....	67
2.8	Comparison of Human, Pig, Xenopus and tissues	68
2.9	Gene-set enrichment using Xenopus cell type markers	68
2.10	ChIP qPCR	69
3	Nanog and TGF- β signalling establish developmental competency	70
3.1	Introduction.....	70
3.2	Nanog results.....	75
3.3	Discussion: Nanog activity is conserved	129
3.10.1	Recap of aims and objectives.....	129
3.10.2	Summary of key findings.....	129
3.10.3	Discussion of key findings	130
3.10.4	Study limitations	133
3.10.5	Future work.....	134
4	ELK1 modulates mesodermal development and germ line competence	136
4.1	Introduction.....	136
4.2	Results	140
4.3	Discussion: ELK1 has both conserved and novel roles in axolotl.....	169
4.1.1	Recap of aims and objectives	169
4.1.2	Summary of key findings	169

4.1.3	Discussion of key findings.....	170
4.1.4	Limitations of this study	174
4.1.5	Future work.....	175
5	General discussion	177
5.1	Axolotl as a model for early development.....	177
5.2	Axolotl and Xenopus.....	179
5.3	Germ plasm and the need for more model organisms.....	181
6	Appendices.....	182
7	Bibliography	183

List of Figures

Figure 1.1.1. Overview of human and mouse development. Human and mouse have a similar morphology prior to gastrulation.	21
Figure 1.2. Differentiation potential of axolotl embryo explants.....	50
Figure 3.1. Cross-species comparison of pluripotent tissues.	76
Figure 3.2. Overview of axolotl early development.....	78
Figure 3.3. Expression of axolotl genes associated with pluripotency in mammals.	79
Figure 3.4. NANOG KD in axolotl embryos	80
Figure 3.5. Confirmation of NANOG depletion.	81
Figure 3.6. NANOG is required to extinguish pluripotency.....	84
Figure 3.7. Differential gene expression following NANOG KD.	85
Figure 3.8. Amphibian cell-type marker expression in NANOG depleted embryos.	87
Figure 3.9. QPCR validation of key germ layer markers in NANOG KD embryos.	90
Figure 3.10. Schematic: NANOG KD prevents waves of gene expression.....	91
Figure 3.11. Germ-layer marker expression in NANOG KD ACs.	93
Figure 3.12. Mesendodermal induction in NANOG KD ACs.....	94
Figure 3.13. Schematic: NANOG acts as a rheostat of SMAD2 activity.	96
Figure 3.14. Epigenetic and transcriptional development in embryo explants.	97
Figure 3.15. Immunofluorescent imaging of NANOG KD ACs.	99

Figure 3.16. Expression of axolotl orthologues of TGF- β pathway related genes across developmental stages. The majority of TGF- β pathway related genes are maternally expressed and are highly expressed until mid-gastrula stages.	101
Figure 3.17. SB431542 treatment arrests development prior to gastrulation.....	102
Figure 3.18. Differential gene expression following SB431542 treatment.....	103
Figure 3.19. Amphibian cell type marker expression following SB431542 treatment.	104
Figure 3.20. QPCR validation of key germ layer markers in SB431542 treated embryos.	105
Figure 3.21. Germ-layer marker expression in SB431542 treated embryo explants.	107
Figure 3.22. Immunofluorescent staining's of SB431542 ACs.	108
Figure 3.23. Axolotl NANOG can physically interact with axolotl SMAD2.....	111
Figure 3.24. DPY30 KD arrests development post-gastrulation.....	112
Figure 3.25. Immunofluorescent staining's of DPY30 KD ACs	113
Figure 3.26. Differential gene expression in DPY30 KD embryos.	114
Figure 3.27. Amphibian cell-type marker expression in DPY30 depleted embryos.	115
Figure 3.28. QPCR validation of key germ layer markers in DPY30 KD embryos.	116
Figure 0.29. Germ-layer markers in DPY30 KD ACs.	117

Figure 3.30. Mesendoderm induction assays in DPY30 KD ACs	118
Figure 3.31. Western blots showing acquisition of epigenetic marks in axolotl embryos.....	119
Figure 3.32. Immunofluorescent staining's of SB431542 treated, Nanog and DPY30 depleted ACs.	120
Figure 3.33. Overlapping effects of SB431542 treatment, NANOG and DPY30 depletion.....	122
Figure 3.34. Clustering of experimental treatments and key gene expression.....	123
Figure 3.35. Correlation heatmap of experimental conditions.....	124
Figure 3.36. H3K4me3 ChIP in ACs following SB431542 treatment, NANOG or DPY30 depletion.....	125
Figure 3.37. H3K4me3 ChIP in mesoderm-induced ACs following SB431542 treatment, NANOG or DPY30 depletion.....	126
Figure 3.38. Schematic: Successive waves of gene expression drive development	126
Figure 3.39. Schematic: NANOG complexes with SMAD2 and DPY30 to deposit H3K4me3	128
Figure 4.1. Alignments of amino acid sequences of human, mouse, axolotl and xenopus ELK1 sequences.	141
Figure 4.2. Overview of FGF signalling in early axolotl development.	143
Figure 4.3. ELK1 depletion in axolotl embryos.....	144
Figure 4.4. Titration of ELK1 morpholino.....	145

Figure 4.5. Efficiencies of KD and characterisation of ELK1 KD phenotype.....	147
Figure 4.6. Germ-layer marker expression in ELK1 depleted embryos.	148
Figure 4.7. Elk1 KD does not affect the expression of pluripotency factors.	150
Figure 4.8.	Error! Bookmark not defined.
Figure 4.9. Differential gene expression in ELK1 depleted ACs.....	153
Figure 4.10. Key marker gene expression in hESCs 4 days following ELK1 depletion.....	155
Figure 4.11. Differentiation of late stage ACs following ELK1 depletion.	156
Figure 4.12. Mesoderm induction in ELK1 depleted ACs.....	158
Figure 4.13. Rescue of ELK1 depletion with mutant hELK1 mRNA.	161
Figure 4.14. Expression of PGC-associated genes in axolotl embryos.....	163
Figure 4.15. PGC related gene expression in PGC-induced ACs.	164
Figure 4.16. Expression of PGC genes in ELK1 depleted and hELK1 rescued ACs.	165
Figure 4.17. Gene expression of uninjected and PGC induced ACs with or without MED23 KD. (n=10 ×3).....	168

List of Tables

Table 1. List of primers used in this study.....	64
--	----

List of Abbreviations

ACAnimal Cap

cDNAcoding DNA

CCScaspase cleavage site

CDScoding sequence

DNADeoxyribonucleic acid

ESCEmbryonic Stem Cells

MESC.....Mouse Embryonic Stem Cells

HESCHuman Embryonic Stem Cells

EpiSCEpiblast-derived Stem cells

GFPGreen fluorescent protein

GRNGene Regulatory Network

H&E Haematoxylin and Eosin

H3K4me3 Histone 3 Lysine 4 trimethylation

H3K4me2 Histone 3 Lysine 4 dimethylation

H3K9me3Histone 3 Lysine 9 trimethylation

H3K27ac Histone 3 Lysine 27 acetylation

H3K27me3 Histone 3 Lysine 27 trimethylation

H3K36me3 Histone 3 Lysine 36 trimethylation

ICMInner Cell Mass

iPSCInduced Pluripotent Stem Cells

LB Lysogeny broth

LIF Leukemia inhibitory factor

MEFMouse embryonic fibroblasts

OSKMOct4 Sox2 Klf4 cMyc (Yamanaka factors for reprogramming)

PBS Phosphate-buffered saline

PCRPolymerase chain reaction

PGCPrimordial Germ Cell

qPCRQuantitative (Real time) PCR

RNARibonucleic acid

RNAiRNA interference

WRTryptophan repeat domain

Gene Nomenclature

This thesis frequently refers to the activities of genes, RNA and proteins in different organisms each with their own gene nomenclature. To aid the reader when referring to a gene/RNA/protein across multiple species the Human nomenclature has been used.

1.1 Literature Review

1.2 Cell potency and pluripotency

Every multicellular life-form, from whales to the algae on which they feed, develops from a single cell. Conrad Hal Waddington first introduced the concept of the epigenetic landscape, likening the zygote giving rise to genetically identical progenitors undergoing a process of irreversible lineage-commitment to marbles rolling down a hill (Waddington, 1957). Remarkably, almost 65 years later, this metaphor is still relevant. Cell potency refers to a progenitor cell's capacity to give rise to other specialised cell types. The term totipotency, for example, denotes a cell's ability to form both embryonic and extraembryonic cells needed to support mammalian offspring to term. By contrast, unipotency describes the ability to produce only a single cell type and multipotency several cell types. The capacity of the zygote for differentiation and self-organisation underpins the vast array of phenotypic diversity on the planet. Pluripotency describes the capacity to produce all the cells of the adult organism, and given that the majority of multicellular life forms do not require extraembryonic tissues, it is therefore, embryonic pluripotency which is conserved. Understanding to what extent this trait differs between organisms, particularly in regard to the genetic and epigenetic regulatory networks, can inform us on how diverse animal phylogenies evolved.

1.3 Early mouse development

While there is still much that is not understood about early development across metazoans, vertebrates stand out as perhaps the most well studied. Early mouse development, in particular, has been very well characterised (Overview in Fig. 1.1), given their use as a 'model' for human development. Following fertilisation of the oocyte by a spermatozoon, the nascent zygotic genome is reprogrammed to a state of totipotency by maternal factors in the egg (Schulz and Harrison, 2019). Within a day of fecundation (Embryonic day 1; E1), the maternal-to-zygotic transition (MBT, also referred to as zygotic genome activation or ZGA) occurs: the previously silent zygotic genome becomes transcriptionally active, and oocyte mRNA's are actively destroyed (Matsumoto et al., 1994, Bouniol et al., 1995, Aoki et al., 1997). Subsequently, the murine embryo undergoes a series of asynchronous cell divisions known as cleavage (Bischoff et al., 2008).

Cleavage itself is characterised by S phase DNA replication and mitosis without a G phase from within its proteinaceous envelope: the zona pellucida (Aiken et al., 2004). Totipotency is maintained after the first cleavage division (E1.5), demonstrable by an individual blastomere's ability to develop to term if the other is experimentally destroyed (Morris et al., 2012, Tarkowski, 1959). This potency is only transient as blastomeres isolated at the four or eight-cell stage (E2-2.5) cannot develop beyond implantation (Rossant, 1976, Tarkowski and Wroblewska, 1967). From the eight-cell stage (E2-2.5) onward, the embryo undergoes compaction whereby cells form a structure known as the morula. At E2-3.25, all early zygotic cells retain their ability to integrate into both somatic and extraembryonic tissues if combined with other developing

blastomeres to form a chimaera (Piotrowska-Nitsche et al., 2005, Kelly, 1977, Pinyopummin et al., 1994).

At the thirty-two cell stage (E3.5), cells have formed a blastula and initiate the process of cavitation, and for the first time, embryonic cells begin to acquire spatially derived identities. The positioning of blastomeres at this stage is critical, as inner and outer groups of cells, the inner cell mass (ICM) and trophoctoderm (TE) respectively, possess differing cell potencies (Posfai et al., 2017, Mihajlovic and Bruce, 2017). While the early TE lineage is irreversibly committed, fated to become part of the placenta, the ICM retains its potential to form TE until the sixty-four cell stage (Posfai et al., 2017). Heterogeneity within the ICM enables the formation of pluripotent-epiblast and extraembryonic primitive-endoderm (PrE) progenitors. As epiblast and PrE cells commit to their respective lineages, they also become spatially separated, with the PrE lining the blastocyst cavity at the embryonic pole by E4.5 (Chazaud et al., 2006, Meilhac et al., 2009, Plusa et al., 2008).

The nascent pluripotent epiblast can give rise to all embryonic cell types but is distinct from earlier ICM cells, as it is incapable of producing PrE or TE (Gardner and Rossant, 1979). After seven cleavage divisions, the blastula expands, breaking through the zona pellucida, forming the 'late blastocyst', denoting the final stage of pre-implantation development in the mouse (Mihajlovic and Bruce, 2017).

Immediately prior to the embedding of the murine embryo into the uterine lining around E4.75-5 (Wang and Dey, 2006), the blastocyst begins a series of morphogenetic reorganisations prior to gastrulation, including the formation of a transient structure called the epiblast rosette (around E5). Implantation itself is initiated through the mural trophoctoderm. During this time, apolar

epiblast and polar-trophectoderm derived extraembryonic ectoderm undergo lumenogenesis, forming a cup-shaped epithelium surrounding the pro-amniotic cavity. The completion of the egg cylinder occurs by E5.5 just prior to gastrulation and is marked by the extraembryonic endoderm forming the parietal and visceral endoderm within a cylindrical wall made up of mural-trophectoderm derived giant trophoblast cells (Bedzhov et al., 2014, Latos and Hemberger, 2016). Interestingly, the egg cylinder structure itself appears to be unique in animal development, only appearing within the rodent clade. In mice, gastrulation takes place from E6.5, denoted by the formation of the primitive streak in the posterior epiblast.

The completion of gastrulation is the knell for pluripotency, as lineage specification events drive epiblast cells to move through the primitive streak between the posterior epiblast and parietal endoderm, irreversibly committing to an endodermal, mesodermal, ectodermal or primordial germ cell (PGC) fate. Once cells are committed to these progenitor states, they undergo a process of further differentiation and morphogenesis to form the body plan of the future foetus (Shahbazi and Zernicka-Goetz, 2018).

1.4 Early human development

Early development of humans follows roughly the same trajectory as mice, particularly in the pre-implantation stages (Overview in Fig. 1.1). Human blastocysts are made up of the trophectoderm, epiblast and hypoblast (equivalent to the primitive endoderm in mice) and are unable to fully implant until around nine days post-fertilisation (E9). Following implantation, the morphological differences between mice and humans become increasingly pronounced (Shahbazi and Zernicka-Goetz, 2018). At E7, a subset of epiblast

cells are fated to become extraembryonic amnion; subsequently, through E9-11, these cells form the lining of the amniotic cavity. Concomitantly, the hypoblast forms the visceral yolk sac endoderm, lining the yolk sac cavity, while the epiblast forms a disk structure beneath the amnion. Around this time, the epiblast rosette undergoes morphological changes from a rosette to a flat disk shape, aptly called the epiblast disk.

Indeed, the disk shape structure of human embryos contrasts the conical shape of the equivalent stage mouse embryo (E5.5). Moreover, this disk-shaped morphology is conserved in non-rodent mammals and non-mammalian amniotes such as birds and reptiles (Alberio et al., 2021). By the end of E11, the trophoderm derived, cytotrophoblast, and syncytiotrophoblast enclose the embryo prior to gastrulation (Shahbazi and Zernicka-Goetz, 2018). Human gastrulation initiates at E14 when the primitive streak forms in the posterior epiblast, cells then migrate between the epiblast and visceral endoderm, eventually acquiring progenitor fates as in mice (Shahbazi and Zernicka-Goetz, 2018).

While the direct study of mammalian embryos provides insights into mammalian development, the study of embryonic stem cells (ESC) derived from the pluripotent cells of the embryo has also offered a great deal of understanding of pluripotency in mammals. These cells maintain their developmental competency and can be maintained and propagated indefinitely in culture. Given these properties, they can be easily subjected to experimental manipulation and have provided insights into the gene regulatory network (GRN) that governs the maintenance of the pluripotent state.

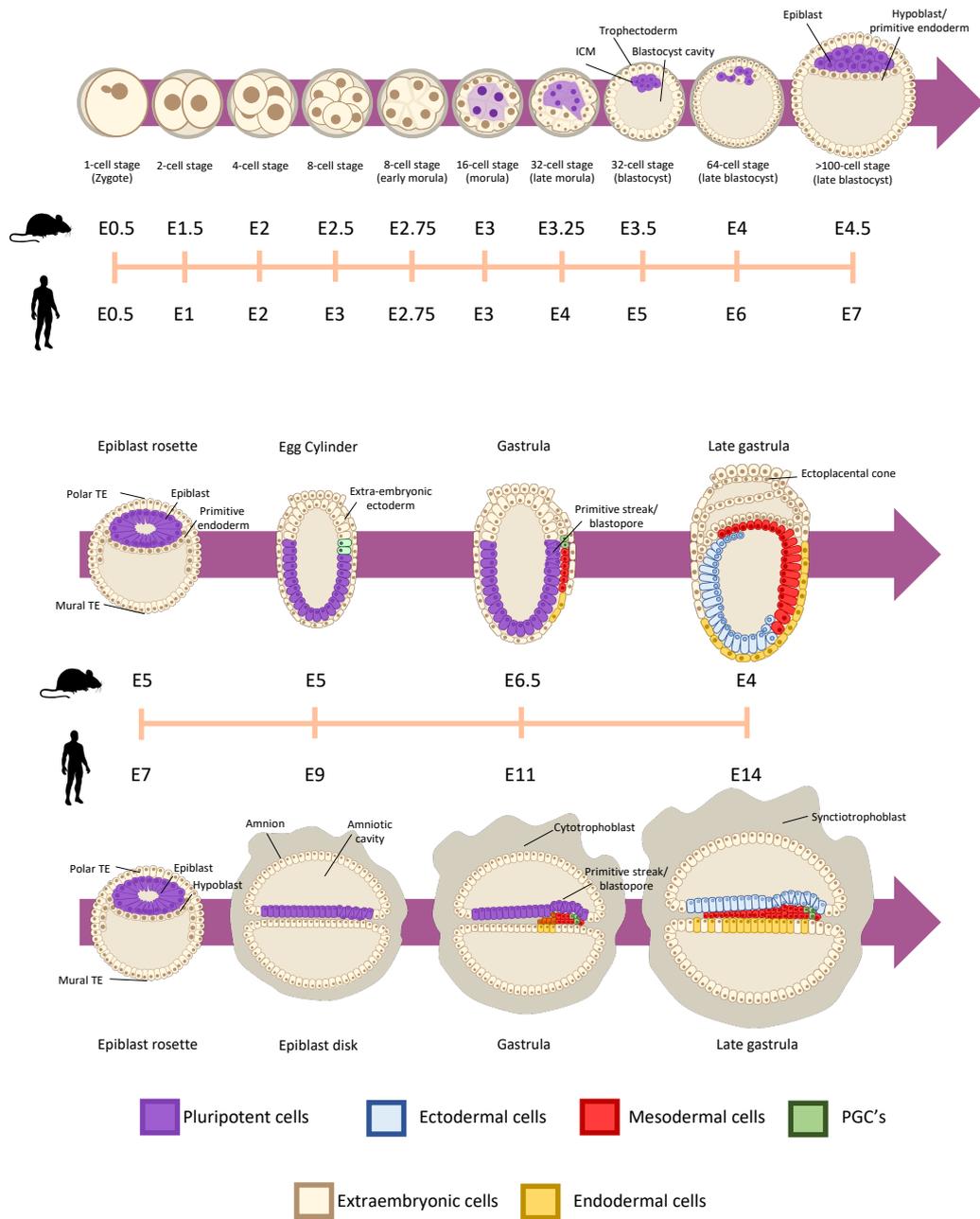


Figure 1.1.1. Overview of human and mouse development. Human and mouse have a similar morphology prior to gastrulation.

1.1.3 The core pluripotency GRN

1.1.3.1 Oct4

The transcription factor OCT4 is encoded by the *POU5F1* gene and belongs to the POU family of DNA-binding proteins. These proteins regulate target gene expression via the octamer motif ATGCAAAT on promoter and enhancer sequences (Scholer et al., 1989, Yeom et al., 1991). OCT4 expression is associated with maintenance of the pluripotent state in mESCs and hESCs and is absolutely required for both murine and human embryogenesis (Boiani and Scholer, 2005, Nichols et al., 1998, Scholer, 1991, Fogarty et al., 2017).

In mice, maternal *Pou5f1* mRNA and Oct4 protein are inherited by the fertilised zygote until the four-cell stage. Both maternal transcripts and protein are required for development until the four-cell stage, where zygotic transcription of *Pou5f1* begins. Oct4 protein can be readily detected in all blastomere nuclei up until the blastula stage, whereby Oct4 is only detected in the nuclei of the ICM. Transient upregulation of Oct4 in a subpopulation of the ICM may facilitate primitive endoderm formation, likely through modulating FGF (Palmieri et al., 1994, Pesce and Scholer, 2001, Jerabek et al., 2014, Zeineddine et al., 2014, Le Bin et al., 2014).

Pou5f1 mRNA transcripts continue to be highly expressed in the epiblast cells until at least E8.5 gastrula stages, overlapping with PGCs (Pijuan-Sala et al., 2019); more to this, *Pou5f1*/Oct4 expression can also be detected in the primitive streak. However, Oct4 protein expression gradually reduces during gastrula stages and by E8 is confined only to primordial germ cells (Pesce and Scholer, 2001, Jerabek et al., 2014, Grose and Spigland, 1976, Zeineddine et al.,

2014). *Pou5f1* null mice fail to form an ICM (Nichols et al., 1998); instead, embryos divert to trophectoderm and die around the time of implantation (Nichols et al., 1998, Chambers and Tomlinson, 2009). Later work suggests that maternal *Pou5f1* /Oct4 is required for the formation of the ICM while zygotic *Pou5f1* /Oct4 is required for the formation of epiblast and primitive endoderm (Le Bin et al., 2014). *In vitro*, *Pou5f1* /Oct4 is highly expressed in mESC cells, mouse embryonic carcinoma (EC) cells and embryonic germ cells. Induced differentiation of mESC subsequently results in the downregulation of *Pou5f1* /Oct4 (Cauffman et al., 2005, Trouillas et al., 2009b, Trouillas et al., 2009a). *Pou5f1* /Oct4 knockdown in mESC results in an up-regulation of genes associated with endoderm and trophoblast differentiation (Niwa et al., 2000, Niwa et al., 2005).

In humans, the *POU5F1* gene encodes 2 OCT4 isoforms, which are generated by alternative splicing (Takeda et al., 1992, Cauffman et al., 2006). The isoforms: OCT4-IA and OCT4-IB (which share a common C terminal but have 360 and 265 amino acids, respectively) possess differential activity. OCT4-IA is present in high levels in hESC, human EC cells and human embryonic germ cells (Reubinoff et al., 2000, Pera and Herszfeld, 1998, Goto et al., 1999) of the isoforms, therefore OCT4-IA is required to sustain hESC pluripotency. In contrast, OCT4-IB is unrelated to maintenance of the pluripotent state (Cauffman et al., 2006) as such, OCT4-IA in stem cell and developmental biology is usually referred to as simply OCT4.

As well as being highly expressed in hESCs, OCT4 expression is necessary to maintain pluripotency in hESCs where it acts as a repressor of the gene encoding for human chorionic gonadotropin (hCG), a placental marker. Depletion of *POU5F1* /OCT4 in hESC results in rapid differentiation and up-

regulation of genes associated with trophoblast and endoderm lineages (Hay et al., 2004, Zafarana et al., 2009).

As in mouse embryos, human *POU5F1* mRNA is maternally inherited, present until uncompact morula stage (Huntriss et al., 2017, Abdel-Rahman et al., 1995) (Hansis et al., 2000). However, unlike mice, *POU5F1*/*OCT4* expression varies between individual blastomeres and is localised in the cytoplasm as opposed to the nucleus, prior to compaction (Cauffman et al., 2005). During compaction, *OCT4* protein is highly expressed and translocates into the nuclei of all blastomeres of the morula. As the blastocyst develops, *POU5F1* transcripts and proteins are expressed exclusively in the ICM (Hansis et al., 2000). Human *POU5F1* mRNA transcripts continue to be highly expressed in the epiblast cells and PGCs until at least the late gastrula stages; lower-level *Oct4* expression can also be detected in the emerging germ layer progenitors. Deletion of *POU5F1* in human embryos results in the downregulation of markers of all three pre-implantation lineages: trophoctoderm, epiblast and hypoblast, as opposed to mice where *Pou5f1* deletion drives trophoctoderm, suggesting that the role of *POU5F1*/*OCT4* differs between mice and humans (Fogarty et al., 2017). Despite *POU5F1*/*OCT4* being expressed in the post-implantation epiblast in both mice and humans, to date, its role in this context *in vivo* is unexplored. *Pou5f1*/*Oct4* was also identified as a critical factor along with *Klf4*, *Sox2*, and *c-myc* (OKSM) capable of reprogramming somatic cells back to a pluripotent state (Takahashi and Yamanaka, 2006).

1.1.1.1 Sox2

SOX2 is a transcription factor and part of the SRY related HMG box (*SOX*) family of proteins, which all contain an HMG domain with at least 50%

sequence similarity to that of the *SRY* gene. Around 20 *SOX* genes have been identified in the mouse and human genomes; these genes are divided into eight subgroups based on sequence and functional similarity (Wegner, 2010, Weina and Utikal, 2014). *SOX2* itself is part of the *SOXB1* subgroup and, as previously mentioned, is part of the *OKSM* factors, which reprogram somatic cells into induced pluripotent stem cells (iPSCs) (Takahashi and Yamanaka, 2006). Crucially, *SOX2* forms part of the core pluripotency triumvirate in mESCs along with fellow core pluripotency factors *POU5F1/OCT4* and *NANOG* whereby they modulate each other's expression through promoter binding (Gong et al., 2015, Boyer et al., 2005, Loh et al., 2006). *SOX2* acts as an effector protein through its HMG domain which bind to specific DNA consensus sequences; this domain also contains a nuclear localisation and export signal.

The *SOX2* C-terminal transactivation domain facilitates gene activation or repression through promoter binding (Castillo and Sanchez-Cespedes, 2012). *Sox2* expression begins in the murine embryo at the 2-cell stage, and expression increases up to the blastocyst stage (Pan and Schultz, 2011). *Sox2* deletion is lethal, and embryos fail to establish an ICM (Avilion et al., 2003). After blastocyst formation, *Sox2* is restricted to the ICM and later the epiblast (Pijuan-Sala et al., 2019).

Following gastrulation, *Sox2* expression characterises primordial germ cells, gut endoderm, presumptive neuroectoderm, sensory placodes and pharyngeal arches (Wood and Episkopou, 1999, Yabuta et al., 2006, Pijuan-Sala et al., 2019). *In vitro*, *SOX2* is highly expressed in mESCs and hESCs. *Sox2* deletion in mESC results in differentiation to trophectoderm-like cells, consistent with its role in mouse *in vivo* (Boyer et al., 2005). *Sox2* is able to

reprogram somatic cells to pluripotency as part of OKSM or with only *Pou5f1*/*Oct4* and histone deacetylase inhibitor: valproic acid (Huangfu et al., 2008). Moreover, *SOX2* promotes neural development by antagonising transcription factors such as brachyury that drive the development of non-neural lineages (Wang et al., 2012, Thomson et al., 2011). Over-expression of *Sox2* in mouse ESCs induces non-specific lineage differentiation, neuronal differentiation or cell death (Mitsui et al., 2003, Zhao et al., 2004, Kopp et al., 2008).

In humans, zygotic *SOX2* transcripts are detectable from the eight-cell stage (E3) later than that observed in mice (Yan et al., 2013). Like *POU5F1*, *SOX2* mRNA expression increases up until implantation (E9) and reduces thereafter (Mole et al., 2021). Also, like *POU5F1* human *SOX2* can still be detected in the epiblast of late gastrula embryos after fellow core pluripotency factor *NANOG* has greatly reduced in expression. In contrast to mice, however, *SOX2* is not detected in PGCs (Perrett et al., 2008, Tyser et al., 2020). To date, the effects of *SOX2* depletion in human embryos has not been tested, so it remains unclear; however, depletion or over-expression of *SOX2* in human ESCs induced trophectoderm differentiation; therefore, *SOX2* may have a similar role in early human embryogenesis as mouse (Adachi et al., 2010).

1.1.1.2 NANOG

The homeodomain-containing transcription factor *NANOG* is critical in early mouse development acts as a master regulator of pluripotency *in vitro* in both mESCs and hESCs (Chambers et al., 2003, Mitsui et al., 2003, Zaehres et al., 2005, Darr et al., 2006) and is associated with cell fate determination, proliferation, and apoptosis (Kerr et al., 2008, Galan et al., 2010, Hoei-Hansen

et al., 2005, Yu and Cirillo, 2020, Grubelnik et al., 2020). In mouse embryos, *Nanog* mRNA transcripts can be detected from fertilisation (Grubelnik et al., 2020); however, there are some incongruities in the literature, one study reporting the detection of maternal *Nanog* proteins (Suzuki et al., 2015), another study shows *Nanog* protein expression does not occur until the eight-cell stage where only low amounts of *Nanog* protein can first be detected in a stochastic pattern (Komatsu and Fujimori, 2015). It is clear that *Nanog* protein can readily be detected by the sixteen cell stage (Komatsu and Fujimori, 2015, Suzuki et al., 2015, Dietrich and Hiiragi, 2007), from then, *Nanog* protein and mRNA is restricted to the inner cells of the late morula and continues to be expressed throughout formation and expansion of the inner cell mass (ICM) and elaboration of the epiblast (Chambers et al., 2003, Mitsui et al., 2003, Dietrich and Hiiragi, 2007, Biase et al., 2014, Komatsu and Fujimori, 2015).

Nanog expression in the epiblast during mouse implantation is also somewhat unclear. Acampora and colleagues show that immediately following implantation in mice, the *Nanog* protein is undetectable (E4.7-5) before gradually reactivating in the later epiblast and remaining active until the end of gastrulation where it's expression is confined to primordial germ cells (Acampora et al., 2013, Yamaguchi et al., 2005, Pijuan-Sala et al., 2019). These observations are consistent single-cell mRNA expression data which suggests *Nanog* transcripts are present mainly in the epiblast and primitive streak from E6.5. As PGCs emerge at E7, they too express *Nanog*. As gastrulation continues, *Nanog* transcripts are lost in the epiblast by E8 and in the streak at E8.25 before being confined to only PGCs (Pijuan-Sala et al., 2019). Early experiments demonstrated that *Nanog*-null mice are able to establish an ICM; however, *Nanog* is indispensable to the establishment of a viable epiblast

(Mitsui et al., 2003, Silva et al., 2009). Thus, *in vivo*, Nanog is required for the maintenance and expansion of pluripotent epiblast cells.

Nanog was initially identified by Chambers and colleagues due to its role in mouse embryonic stem cells (mESCs) *in vitro* (Chambers et al., 2003). Early culture of mESCs required specific cell culture conditions (Serum-LIF; Basic conditions) containing both leukaemia inhibitory factor (LIF) and bone morphogenic protein 4 (BMP4) to maintain a pluripotent state. Nanog was identified through its ability to maintain pluripotency whilst allowing the indefinite expansion (self-renewal) of mESCs in the absence of LIF when overexpressed in mESCs. In mice, LIF facilitates self-renewal through activation of the JAK/STAT3 and PI3K/AKT signalling pathways, resulting in the increased expression of pluripotency factors Klf4 and Tbx3, respectively. Klf4 promotes the expression of Sox2 while Tbx3 drives Nanog expression. BMP4 signalling induces the phosphorylation of Smad proteins 1, 5 and 8, resulting in inhibitor of differentiation (Id) gene upregulation, inhibiting the expression of numerous genes associated with differentiation.

A Nanog interactome created by Wang and colleagues show that Nanog has a complex network of interactions in mESCs (Wang et al., 2006). The interactome highlights its physical interaction with known co-repressors like the SWI/SNF, NuRD and Polycomb complexes. The Polycomb complexes themselves are associated with the deposition of H3K27me2 (Wang et al., 2006). Therefore, Nanog may directly repress differentiation factors through the recruitment of epigenetic modifiers. Also, key to the functionality of mouse Nanog proteins is a tryptophan-rich (WR) domain which drives the formation of homodimers. *In vitro*, Nanog homodimers are necessary to support self-renewal. Indeed, substitutions of the tryptophan residues for

alanine resulted in an inability to dimerise. Curiously, the ability for Nanog to dimerise was demonstrated to be a requirement for self-renewal and pluripotency *in vitro* (Wang et al., 2008). Co-IP also revealed that specific factors including Pou5f1, Sall4, Zfp198, Zfp281, Dax1 and Nac1 preferentially interacted with Nanog homodimers. Zfp281, in particular, appears to play a role in facilitating the binding of Nanog to its own promoter as well as regulating *Nanog* and other core pluripotency factors expression (Wang et al., 2008, Kim et al., 2008, Fidalgo et al., 2011).

The regulation of NANOG expression is complex and highly regulated through many different mechanisms, including the methylation status of the *NANOG* gene, miRNA and protein level regulation (Gong et al., 2015, Mato Prado et al., 2015, Hart et al., 2005). OCT4 and SOX2 are one of the most important and investigated protein regulators of NANOG; they form complex with KLF4 and bind to the *POU5F1* /SOX2 motif upstream of the transcription start site (TSS) of the *NANOG* promoter (Gong et al., 2015). Chromatin immunoprecipitation (ChIP) followed by paired-end sequencing (ChIP-Seq) in mESCs identified binding sites for Nanog, Oct4 and Sox2. Of the 1,083 and 3,006 binding sites identified for Oct4 and Nanog, respectively, 44.5% of Oct4 bound genes were shared with Nanog; furthermore, the majority of Oct4 bound genes also had binding sites for Sox2. This suggested that these genes may regulate each other's expression in addition to target genes (Loh et al., 2006). Although Oct4 is believed to be one of the major Nanog regulators, Nanog in the mouse ICM is not dependent on *Pou5f1* /Oct4 expression (Wu and Scholer, 2014). However, the promotion of pluripotency gene expression and repression of differentiation-associated genes maintains a positive auto-

regulatory loop between Oct4, Sox2 and Nanog to maintain pluripotency (Mitsui et al., 2003, Boyer et al., 2005, Niwa et al., 2009).

Interestingly, while knockout of the *Nanog* gene mESC is detrimental, it is ultimately dispensable for self-renewal (Chambers et al., 2007). *Nanog* null mESC are prone to differentiation but are able to be propagated and continue to express other pluripotency factors, including Oct4 and Sox2. Moreover, *Nanog* null mESC maintain their developmental competency able to contribute to all of the tissues of chimeric embryos, including the germ cells; it is worth noting that PGCs are unable to differentiate into mature germ cells once they reach the genital ridges (Chambers et al., 2007). It is unclear why *Nanog* null mESC are able to form a viable embryo, and a *Nanog*-null mouse is not. In another context, exogenous *Nanog* expression is not required to reprogram somatic cells to a pluripotent state (Takahashi and Yamanaka, 2006); however, an endogenous *Nanog* is required (Silva et al., 2009), given that mESC are already pluripotent *Nanog* may be dispensable after the earlier establishment of pluripotency in the ICM alternatively the *Nanog*⁺ cells already present within the chimeric blastocysts may be sufficient to support the formation of the epiblast alone.

The role of NANOG in human embryos is less clear, given the vastly reduced number of embryos available for experimental manipulation, as well as the fact that only recent advances have allowed the *in vitro* culture of embryos to peri-implantation (Shahbazi et al., 2016, Deglincerti et al., 2016, Grubel'nik et al., 2020). As with mice, there are conflicting reports on when Human *NANOG* mRNA expression begins; one study reports that *NANOG* expression was absent in two-cell, four-cell, and beginning of the eight-cell stage but was detected in some cells of the compacted morula (Hambiliki et al., 2012).

In contrast, another study showed that NANOG expression was detected in all blastomeres from the five, six and eight cell stages (Galan et al., 2010). A recent study has corroborated the findings of Hambiliki et al, finding NANOG is expressed stochastically from the eight-cell stage (Mole et al., 2021). There is consensus that NANOG is expressed sporadically in the cells of the ICM but not TE and later in all cells of the epiblast but not the DE or TE (Hambiliki et al., 2012, Galan et al., 2010, Mole et al., 2021). To date, however, no study has tested the effects of NANOG depletion or deletion in a human embryo; instead, the majority of our understanding of NANOG's role in humans comes from the study of hESCs.

Following the derivation of hESCs, it became clear that they were distinct to mESCs, particularly in morphology and culture conditions required to promote self-renewal. Similarly, mESCs have a greater developmental potential than hESCs, and indeed ESCs derived from other mammals. This is demonstrated by the fact that mouse and rat ESCs can generate germline-transmitting chimaeras when introduced into E2.5 morulae or E3.5 blastocysts (Tesar et al., 2007, Brons et al., 2007).

It was initially unclear whether the differences between hESCs and mESCs were due to species nuances or due to the methods in which the cells themselves were derived. Later work discovered that mouse epiblast stem cells (mEpiSCs) could be derived from later stage mouse embryos and demonstrated a closer semblance to hESCs. In particular, BMP signalling mediated by Smads 1 and 6 drives hESC and mEpiSC differentiation but can promote self-renewal mESCs (Saunders et al., 2013). In addition, both mEpiSCs and hESCs do not require LIF to maintain pluripotency; rather, they require insulin or insulin-like growth factor (IGF) and basic fibroblast growth

factor (bFGF) signalling (Vallier et al., 2005, Bendall et al., 2007, Silva et al., 2008). In this context, bFGF activates the mitogen-activated protein kinase (MAPK), which subsequently activates ACTIVIN/NODAL signalling pathways. IGF activates the Ras and PI3K pathways to sustain self-renewal.

ACTIVIN/NODAL signalling is propagated through the phosphorylation of effector proteins SMAD2/3 via the TGF- β RI kinase receptor. The TGF- β RI kinase inhibitor SB431542 was shown to prevent SMAD 2/3 phosphorylation which led to downregulation of NANOG and OCT4 (James et al., 2005). Later work identified SMAD2/3 consensus sequences within the *NANOG* promoter. The binding of SMAD2/3 to the *NANOG* promoter proved indispensable for transcriptional activation of *NANOG* in response to NODAL signalling (Vallier et al., 2009b).

ACTIVIN/NODAL signalling is also associated with mesendodermal differentiation and inhibition of neuroectodermal specification, also through its effector SMAD2/3 (Vallier et al., 2005, Brown et al., 2011). Interestingly, unlike mESC, NANOG is indispensable for the self-renewal of hESCs (Vallier et al., 2009a). Later work demonstrated that NANOG, SMAD2/3, and DPY30 proteins physically interact in hESCs (Bertero et al., 2015). Furthermore, the complex formed by NANOG, SMAD2/3 and DPY30 acts through the recruitment of the complex of proteins associated with SET1 (COMPASS). COMPASS is a conserved family of proteins that function to maintain specific patterns of gene expression throughout cellular development and has known H3K4 methyltransferase activity. Through COMPASS, DPY30 facilitates trimethylation of H3K4 residues from H3K4me₂, and in combination with NANOG and SMAD2/3, it regulates nodal-responsive genes through H3K4me₃ (Bertero et al., 2015). Thus, it appears that NANOG also has a

primary function in depositing a combination of activating H3K4me3 and repressive H3K27me3 marks, poising genes to rapidly respond to nodal signalling and maintain pluripotency in epiblast-like cells (Bertero et al., 2015).

1.1.1.3 Naïve and primed pluripotency

As mentioned previously, discrepancies between mESCs and hESCs (the latter resembling mEpiSCs) suggested either a critical difference between species or a difference in the cell populations from which the cell lines were derived. Improved technologies have better allowed the study of mouse and human development *in vivo*. Crucially, defined culture conditions and subsequent study has been able to uncover the transcriptional and mechanistic differences between mESCs and hESCs. This has led to mammalian pluripotency being subdivided into two primary states: 'Naïve' and 'primed' pluripotency (Ghimire et al., 2018). It is worth noting that these two states were defined in mice and have later been applied to humans and other large mammals (Discussed later).

The term Naïve pluripotency is applied to mESC which are dependent on Jak/STAT signalling and are usually cultured in Serum-LIF conditions; this state has been likened to the state of the pluripotent cells in the pre-implantation embryo. The term Naïve ground-state was later applied to mESC co-cultured with mitogen-activated protein kinase (MEK) signalling and glycogen synthase kinase-3 (GSK3) inhibitors (2i) with LIF (Silva et al., 2008, Ying et al., 2008, Plusa and Hadjantonakis, 2014). Work by Nichols and colleagues suggests that mESC cultured in 2i have the greatest semblance to the E4.5 epiblast, evidenced by transcriptional similarities and derivation of

mESC only being possible from E3.75-5.5 (Plusa and Hadjantonakis, 2014, Boroviak et al., 2014).

Primed pluripotency refers to the state of mEpiSCs and is comparable to the epiblast after implantation, prior to gastrulation (Plusa and Hadjantonakis, 2014, Boroviak et al., 2014). Given the similarities in culture conditions and developmental potential to mEpiSCs, the term primed pluripotency was later applied to hESCs. Given the existence of the mouse naïve state, it has been hypothesised that an equivalent state exists in humans. To date, it has not been possible to derive a naïve hESCs from an early embryo that is LIF and 2i dependent; whether this is because there is no equivalent state or simply due to the lack of early human embryos available for research remains unclear. Rhesus monkey primed ESCs were able to be induced to resemble a Naïve like phenotype using a media containing 2i +LIF, moreover these cells can be propagated without TGF β ; however, these cell lines still required the addition of bFGF for propagation, an addition which induces the differentiation of mESCs, undoubtedly pointing to differences between the species. However, researchers have attempted to recapitulate a 'naïve-like' state in hESCs *in vitro* using a variety of methods.

Takashima and colleagues were able to induce a rest naïve-like state in hESCs using transient transgene overexpression of NANOG and KLF2; the resulting cell-type is able to self-renew in 2i/Lif even in the presence of FGF and activin inhibitors (Takashima et al., 2014); similarly, other cocktails of small molecule inhibitors have been reported to induce naïve-like qualities (Dodsworth et al., 2015). Later work compared genome-wide expression profiles of individual cells of human blastula ICM (E5), pre-implantation epiblast (E6-7), early post-implantation epiblast (E9) and late post-implantation epiblast (E11) with that

of hESCs and 'reset' hESCs (Mole et al., 2021). Interestingly, the transcriptional profiles of the reset hESCs aligned more closely to E6-7 epiblast cells than conventional hESCs, and likewise, hESCs aligned closer to E11 epiblast cells than reset cells.

Given the mannered nature of ESC culture and that pluripotency *in vivo* is transient and in a constant state of flux, undoubtedly, some level of artefact may be present. This does, however, present reasonable evidence that a naïve and primed states may exist in pre-and post-implantation epiblast cells, respectively, in humans. However, it is worth noting that rodent embryos may be uniquely able to self-renew, possibly due to their ability to undergo diapause (Nichols and Smith, 2009). By contrast no equivalent phenomenon has ever been described in human and thus, earlier embryonic cells may never be able to be maintained stably in a state of self-renewal in contrast to more advanced cell states such as formative/primed.

1.1.1.4 FGF signalling

In mammals, FGF/ERK signalling is critical to gastrulation, particularly in mesendoderm formation (Dorey and Amaya, 2010). ERK/MAPK acts as one of the major transducers of FGF signalling and is activated by MEK (MAPK/Erk kinase) (Thisse and Thisse, 2005). FGF/ERK signalling appears to play a role in the exit from naïve to primed pluripotency evidenced by the requirement for ERK inhibition to stabilise the 'naïve state' *in vitro* and the induction of priming through supplementation of FGF (Greber et al., 2010, Guo and Wang, 2009). FGF/ERK also plays a role in the differentiation of primed cells into the three germ layers and germline.

In mouse and human Erk1 (MAPK3) and Erk3 (MAPK1 and MAPK2) (Boulton et al., 1991). Mice deficient in Erk1 are viable, fertile and of normal size (Pages et al., 1999), suggesting no significant role in mouse development; by contrast, ERK2 KD mice do not form mesodermal or endodermal tissues and die during gastrula stages (Hatano et al., 2003, Saba-El-Leil et al., 2003, Yao et al., 2003). When Fgf/Erk signalling is disturbed pharmacologically or genetically, mESCs are unable to differentiate into neuroectodermal, mesodermal or endodermal lineages (Kunath et al., 2007, Stavridis et al., 2007). In stark contrast, mEpiSCs and hESC maintain developmental competency through activation of FGF/ERK and TGFB/SMAD signalling (Brons et al., 2007, Tesar et al., 2007, Vallier et al., 2005).

1.5 Conservation of the pluripotency GRN in mammals

The study of mouse and human *in vivo*, as well as *in vitro* models, has been invaluable for the identification of some conserved features of the mammalian pGRN. The study of two species is, however, insufficient to elucidate which of the species differences represent derived or conserved regulatory adaptations, if any. One issue of importance is the well-established differences between the signalling dependencies of mice and humans. As previously mentioned, mESCs require the addition of LIF, a member of the interleukin 6 (IL6) family; broadly, these ligands are capable of activating Janus kinase/Signal transducer and activator of transcription 3 (Jak/Stat3) signalling pathway (Hirano et al., 2000).

While mESCs do not express the IL6 receptor, they are able to be maintained in culture by the addition of IL6 and IL6-receptor in culture (Nichols et al.,

1994). Interestingly, mouse embryos genetically deficient for LIF, LIF receptor, Gp130 (one of the common subunits of all IL6 receptors), or even Stat3 can survive until the primitive streak stage (Nichols et al., 2001, Takeda et al., 1997). These findings, taken together with hESCs being derived independently of LIF, cast doubt on whether Stat3 dependent signalling is even required in mammals. However, it was later discovered that LIF and IL6 are responsible for the activation of Stat3 at the four-cell stage in mice (Do et al., 2013). Furthermore, the elimination of stat3 from both oocytes and zygotes resulted in a failure to expand the epiblast. Recent studies trying to recapitulate a naïve state in human cells have enabled the creation of modified hESCs, which are also dependent on Stat3 signalling (Chen et al., 2015, Gafni et al., 2013, Takashima et al., 2014, Theunissen et al., 2014, Zimmerlin et al., 2016). While the role of Stat3 signalling in humans has been poorly explored *in vivo*, this work suggests that Stat3 signalling may be a conserved feature of early development in both rodents and primates.

The increasing number of sequenced genomes has allowed for interrogation of genomes and identification of putative orthologues for a variety of genes, those identified in the mouse and human pGRN included. Accordingly, this has facilitated a greater number of functional studies that shed light on the conserved features of the mammalian pGRN. The derivation of iPSCs from different mammalian taxonomic groups provides some evidence of a conserved mammalian pGRN. This is because iPSCs have been derived using identical constructs encoding constitutive expression of human OSKM transcription factors.

iPSCs and have been derived from a wide variety of mammals, including endangered species, given that somatic cells are far easier to obtain (Ben-Nun et al., 2011). Consequently, iPSCs have been derived from a wide range of taxonomic groups, including Carnivora (Dog, Snow leopard, Mink) (Shimada et al., 2010, Verma et al., 2012, Menzorov et al., 2015), Cetartiodactyla (Pig, Cow) (Ezashi et al., 2009, Han et al., 2011), Chiroptera (Bat)(Mo et al., 2014), Lagomorpha (Rabbit) (Osteil et al., 2013), Metatheria (Tasmanian devil) (Weeratunga et al., 2018), Perissodactyla (White Rhino, Horse)(Ben-Nun et al., 2011, Breton et al., 2013), Rodentia (Mouse, Rat, Naked mole Rat) (Takahashi and Yamanaka, 2006, Liao et al., 2009, Lee et al., 2017), and Primates (Human, Marmoset, Chimps, Ape and macaques, Orangutans) (Takahashi et al., 2007, Tomioka et al., 2010, Marchetto et al., 2013, Wunderlich et al., 2014, Ramaswamy et al., 2015). The pluripotent state of iPSCs lines was demonstrated by canonical means such as teratoma formation or Alkaline phosphatase activity. This suggests that at least broadly, core factors such as OKSM are able to induce pluripotency and are likely conserved; however, to what extent this is true for the wider pluripotency networks identified in mice and humans is unclear.

Intriguingly, the culture conditions used to sustain iPSCs varied; many employed the use of feeder cells, and therefore the requirement for individual cytokines across a wide range of mammals is as yet uncharacterised. The morphology and gene expression profiles also varied between animal models, some cells bearing a closer semblance to human and some to mouse iPSCs (Ezashi et al., 2009, Han et al., 2011, Weinberger et al., 2016). Another prime example is that Nanog is not highly expressed in Mink iPSCs (Menzorov et al., 2015) given it's pivotal role in mice and humans; this suggests that divergence

in even the core pGRN is possible in mammals; however, the mechanism by which this may have occurred is unknown. In regard to the larger pGRN, this has been explored in depth in a few animals. Particularly relevant are numerous studies looking at the early development of monkeys and pigs.

Studies utilising cynomolgus monkeys have demonstrated that they share many gene expression patterns with humans; more to this, human and monkey embryo morphology are very similar (Alberio et al., 2021, Nakamura et al., 2017, Nakamura et al., 2016a, Liu et al., 2021). While these results are certainly interesting, they may not be entirely surprising given that they belong to the same order as humans and shared a common ancestor around 29 million years ago (MYA) (Kumar et al., 2017). Recent data gathered from studies of pigs suggests that they also have a greater semblance to humans than mice, both in terms of morphology, gene expression patterns and specification of PGCs (Kobayashi et al., 2017).

Given that mice shared a common ancestor with humans around 90 MYA and pigs and humans 96 MYA, these data may seem counterintuitive. However, when combined with observations of the development of other large mammals, it appears that much of the characteristics of early human development may be conserved across mammals, while rodents may have diverged (Johnson and Alberio, 2015, Alberio et al., 2021, Ramos-Ibeas et al., 2019). A recent study by Endo et al. (2020) used large-scale comparative genomics to assess the conservation of 127 pluripotency genes across 48 mammalian species and found that pluripotency genes and associated signalling networks are broadly conserved across mammals (Endo et al.,

2020). Together these data suggest that among mammals, there is indeed a great deal of conservation in the pGRN.

1.6 Pluripotency in non-mammals

Given that the pGRN appears to be conserved in mammals and that the property of pluripotency is a defining feature of embryology conserved across multicellular life, this begs the question: Is the pGRN also conserved in non-mammals? Surprisingly, the answer to this question is unclear. Given that many embryological features are shared in vertebrates, non-mammalian vertebrates present an interesting opportunity to study pluripotency separate from mammalian totipotency.

1.1.1.5 Non-mammalian amniotes

As amniotes, mammals shared a common ancestor with birds and reptiles around 312 MYA (Kumar et al., 2017). Possibly the most well studied non-mammalian amniotes are chickens. Like other birds, chicks develop a disc-shaped embryo which was likely present in the last common ancestor between birds and mammals (Pilato et al., 2013). In addition, early blastoderm contains cells able to contribute to both somatic and germinal tissue when injected into a recipient embryo (Petitte et al., 1990, Pain et al., 1996). Furthermore, early blastoderm cells can be cultured and maintained *in vitro* for long-term culture. These cells exhibit similar morphology to human (and, to some degree, mouse) ESC they also demonstrate high telomerase activity (Pain et al., 1996, Zhang et al., 2018).

While, in a strict sense, the chick embryo produces all three germ layers and germ cells, chick cells do not manifest true pluripotency *in vivo*. This is because

chicken epiblast does not produce the three primary layers and germ cells from the same pluripotent progenitor cell type. However, cESCs (and iPSCs) retain the ability to form PGCs in response to signalling (Shi et al., 2014, Zuo et al., 2019, Zhao et al., 2021). The chicken genome contains a pou factor with similar sequence similarity to *Pou5f1* (initially referred to as pouV) and Nanog (cNanog) as well as Sox2 (cSox2). PouV and cNanog are expressed in chicken ESCs (cESCs) and *in vivo* within the early and late-gastrula epiblast, as well as late-stage germ cells (Lavial et al., 2007). Functional analyses confirmed that cNanog can confer cytokine independence in mESC, and correspondingly, maintained expression of pluripotency genes. Further to this, cNanog can rescue Nanog null-mESC. However, pouV was only able to poorly maintain mESC AP activity and ES morphology for a limited period of time in an inducible *Pou5f1*/Oct4 KD cell line. (Lavial et al., 2007). Similar to observations with mESCs, however, KD of PouV in cESCs induced the expression of *Cdx2*, a marker of trophoderm in mammals. This suggests that the regulation of *Cdx2* precedes the advent of the trophoderm. PouV KD in cESCs also caused upregulation of endodermal markers *Gata4* and *Gata6*.

Later work identified that chick pouV was, in fact, *Pou5f3* (Frankenberg et al., 2014). Syntenic analysis suggests that *Pou5f1*/Oct4 and *Pou5f3* likely arose via a multigenic duplication event that occurred before the divergence cartilaginous and teleost fish (Frankenberg and Renfree, 2013). CSox2 is also expressed in the epiblast of pre-gastrula but not gastrula stage embryos (Streit and Stern, 1999); while this suggests it may have a role in pluripotency, to date, no functional studies have demonstrated this. CSox2 is also highly expressed in neural tissues, similar to mammalian Sox2 (mouse) (Streit et al., 1997).

Experimental KDs and overexpression experiments demonstrated that cSox2 likely maintains the progenitor identity of neural stem cells (Graham et al., 2003, Bylund et al., 2003).

Together these data suggests that of the core pluripotency transcription factors, PouV/Pou5f3 in chick is not analogous to mammalian *Pou5f1*/Oct4. However, cNanog does appear to have a conserved activity; this is particularly interesting given that mammalian Nanog's are regulated by *Pou5f1*/Oct4. While it is unclear as to whether cSox2 acts as a pluripotency factor, it appears to have a conserved function in neural progenitors.

Curiously, in regard to the wider pluripotency signalling network, cESCs rely on ESA medium supplemented with bFGF, h-IGF-1, avian-SCF, 1h-LIF and h-IL-11 for continuous propagation (Pain et al., 1996, Pain et al., 1999). This is particularly intriguing given that FGF is required for hESC culture while mESC require LIF; this suggests that chick pluripotency may share some features of naïve and primed mammalian pluripotency.

In vivo, Nakanoh and colleagues reported that chick orthologues of IL6, active Stat3 and Socs3 (part of the Jak/Stat signalling pathway) were significantly more abundant in pre-gastrula embryos than post-gastrula, correlating with the loss of multipotency in the epiblast. Indeed, Jak1 inhibition leads to reduced cNanog expression subsequent differentiation of cESCs to fibroblast-like cells and that this effect can be rescued by exogenous expression of active stat3 (Nakanoh et al., 2017). More to this, Jak1 inhibition prior to gastrulation in *ex ovo* blastoderm cultures disrupts primitive streak formation while inhibition after streak formation has no effect, suggesting Jak/Stat signalling may only be required prior to gastrulation in chick (Nakanoh and Agata, 2019). However, it is worth noting that it is somewhat implied that while

streak formation was perturbed, *ex ovo* cultures of chicken blastoderms did still form primitive streaks even in the presence of Jak1 inhibition (Nakanoh and Agata, 2019). Therefore, Jak1 signalling may not be required absolutely for gastrulation as observed in mice (Nichols et al., 2001).

Notably, the role of FGF in maintaining pluripotency in non-mammals is less clear. Bertocchini et al. (2004) demonstrated that pharmacologically blocking FGF/ERK signalling prevented Brachyury expression and disrupted gastrulation in chick. Furthermore, Erk inhibition resulted in a 5-fold increase in cNanog expression compared to untreated controls; Mek target Mkp3 was also disrupted, suggesting that Erk may be required for chick differentiation (Bertocchini et al., 2004). A later study showed that Fgf/Erk signal inhibition blocked neuronal specification in gastrulating chick embryos (Stavridis et al., 2007). Notably, as neuronal specification occurs later in gastrulation, this finding somewhat contradicts the conclusions drawn by Bertocchini as it suggests that the embryos do gastrulate following Fgf/Erk signal depletion, suggesting that the embryos can indeed exit pluripotency.

It has been suggested that chicken and turtle blastodermal cells can be cultured and that FGF 'prepares' avian and reptile cells for an exit from pluripotency (Nakanoh et al., 2013, Nakanoh et al., 2015, Nakanoh and Agata, 2019). However, this is particularly problematic due to the lack of evidence that cultured turtle cells are even pluripotent, and FGF's role is inferred by morphological features exhibited in culture. More to this, while Chicken blastodermal cells cultured in 2i & MEK inhibitor formed 'mESC/hESC-like' Nanog positive colonies, cells were only cultured for two days, and 2i conditions did not result in activation of cNanog or pouV above that of non-

cytokine media conditions, nor could cells be propagated in 2i (Nakanoh et al., 2013, Nakanoh and Agata, 2019). Therefore, it is unclear as to whether pluripotency in non-mammalian amniotes requires FGF or indeed JAK/STAT signalling; however, it does appear that FGF is required for differentiation.

Unlike birds, turtles (and likely other Testudines) manifest embryonic pluripotency in that all germ layers, including PGCs, are specified from pluripotent progenitors. Also, in contrast to birds (and some squamates), turtles have retained *POU5F1/OCT4* (Frankenberg et al., 2014). Indeed, ovarian stem-cell-like cells isolated from turtle express pluripotency factors OCT4 and NANOG as well as DAZL and VASA (Xu et al., 2018), in line with previous observations (Bachvarova et al., 2009b). While this may suggest that the role of PouV in birds may be divergent.

To date, there have been no studies investigating whether turtle *POU5F1/OCT4* is functionally equivalent to mammalian *POU5F1/OCT4* or indeed any studies investigating pluripotency in turtles. Overall, it would seem that, at least in part, there is evidence that some elements of the pGRN in mammals have been conserved from lower amniotes. However, to what extent this is the case remains somewhat unclear given that to date, data is only available from a handful of chick studies. Indeed, the vast majority of early developmental studies in non-mammals come from the study of amphibians and fish.

1.1.1.6 Xenopus

The extensive study of *Xenopus laevis* and *Xenopus tropicallis* has informed much of our understanding of early amphibian development, and indeed, many principles of early development characterised in *Xenopus* were later

applied to other vertebrate models. Indeed, *Xenopus* embryos were critical to understanding the mechanisms which govern PGC determination via preformation and, as such, do not specify germ layers and PGCs from a single cell progenitor type, indeed by MBT (stage 8), the germline is already transcriptionally distinct from the multipotent cells of the animal cap (Briggs et al., 2018). Correspondingly, *Xenopus* animal caps can be induced to form all three primary germ layers, but no one has reported an ability to produce PGCs (Boterenbrood and Nieuwkoop, 1973, Chatfield et al., 2014). The *Xenopus laevis* genome contains three members of the pou family, which show some similarity to *pou5f1*; *xlpou91*, *xlpou60* and *xlpou25*, curiously, however, *POU5F1/OCT4* appears to have been lost in the *Xenopus* genome as has *Nanog*, the *Xenopus* genome does contain *sox2* (Morrison and Brickman, 2006, Hinkley et al., 1992) (Frankenberg et al., 2014).

It has been suggested that the composite expression pattern of all three *Xenopus* pou orthologues during blastula and gastrula stages resembles the expression of murine *Pou5f1/Oct4*. The three homologues are expressed throughout blastula and gastrula stages in the multipotent animal cap, which produces ectodermal, mesodermal, and foregut endoderm (Briggs et al., 2018, Morrison and Brickman, 2006). However, targeted disruption of *xlpou60* or *xlpou25* had no discernible effect. While this was attributed to functional redundancy between the three homologues, a KD of *xlpou91* alone displayed a posterior truncation and anterior neural defects. Additionally, a KD of all three orthologues showed little difference to *xlpou91* KD.

Needless to say, the effects of pou factor depletion in frogs differs significantly from the *Pou5f1/Oct4*-null phenotype observed in mice, which may imply *POU5F1/OCT4* is less important in amphibian development.

Gastrula-stage pou morphant embryos showed reduced expression of genes associated with marginal zone cells (presumptive mesoderm and foregut) and increased expression of genes associated with more mature cell states. (Morrison and Brickman, 2006). Further, it was suggested that this effect was akin to premature differentiation and likened to *Pou5f1/Oct4* null mice (Nichols et al., 1998) Given that depletion of the three *Xlpou* molecules also induced expression of *xcad3*, a *cdx2* homologue, the authors proposed that the pou molecules have a role conserved between amphibians and mammals in that they suppress differentiation and maintain a multipotent population of cells. Over-expression of *Xlpou25*, *Xlpou60* or *Xlpou91* was shown to inhibit mesendoderm differentiation (Cao et al., 2004, Cao et al., 2006). However, *Xlpou91*, *Xlpou60* and *Xlpou25* were unable to facilitate the propagation of *Pou5f1/Oct4*-null mESC. This suggests that *Xenopus* pou factors are not functionally analogous to murine *Pou5f1/Oct4*.

Like chick, the *Xenopus* Sox2 (xSox2) orthologue has a role in neural development. XSox2 highly expressed dorsal animal cap (presumptive neuroectoderm) during early gastrula, stages when neural specification begins. More to this, it is absolutely required for neuroectoderm formation (Mizuseki et al., 1998, Kishi et al., 2000). XSox2 expression persists in neural tissues throughout embryonic development, including the central nervous system (CNS), neural crest, placodes and lateral line. While xSox2 does not appear to maintain the competency for gastrulation/differentiation, it does suggest that the role of Sox2 as a neural development factor is likely conserved in vertebrates.

As mentioned, no NANOG orthologue exists in the frog genome; however, it has been suggested that the closely related VentX family (Vent1.2 and Vent2.1)

are functionally equivalent to Nanog (Scerbo et al., 2012, Scerbo et al., 2014). This is evidenced by high Vent gene expression in *Xenopus* embryos prior to gastrulation. More to this, dual KD of Vent1.2/2.1 results in the formation of all three germ layers; however, embryos display a truncation along the anterior-posterior axis. Scerbo and colleagues reported that this phenotype can be rescued by mouse Nanog mRNA (Scerbo et al., 2012). However, a study the same year attempted to also rescue vent depletion with murine Nanog but was unable to do so (Schuff et al., 2012). The same study tested *Xenopus* Vent's ability to substitute mouse Nanog in mESC, but this was not possible (Schuff et al., 2012). Therefore, it seems unlikely that vent factors fulfil the role of Nanog in *Xenopus*; however, this has been largely accepted as fact, as it is commonly cited (Briggs et al., 2018).

LIF jak/stat signalling also plays a role in *Xenopus* development; LIF mRNA is expressed after MBT, while LIFR and IL6 are maternally inherited. Overexpression of LIF results in STAT3 phosphorylation, and embryos become centralised and microcephaly. This phenotype results mainly from stimulating BMP signalling, which in turn antagonises IGF. Microinjection of a dominant-negative LIFR perturbs embryonic kidney development (Jalvy et al., 2019). Another study employing dominant-negative Stat3 to perturb Stat signalling induced dorsalised *Xenopus* embryos (Nishinakamura et al., 1999). Together this suggests that LIF/STAT3 signalling likely functions to regulate dorsal-ventral axis formation but is not required for pre-gastrulation development as seen in mice (Nichols et al., 2001).

Interestingly, early work by Gurdon and colleagues (1958) demonstrated the genomic equivalence of somatic and undifferentiated nuclei using *Xenopus laevis*. This was evidenced by the transplantation of somatic nuclei into

enucleated host eggs (Gurdon et al., 1958). Importantly, this demonstrated that the *Xenopus* oocyte possessed all necessary factors to remodel somatic nuclei to a state which could support development to term. Later studies provided insights on the differences between mammalian and *Xenopus* developmental GRNs. Mammalian nuclei injected into *Xenopus* germinal vesicles also showed activation of native *Pou5f1/Oct4* (Byrne et al., 2003). Studies utilising *Xenopus* oocyte extracts to remodel permeabilised mouse embryonic fibroblasts (MEFs) found that *Xenopus* extract treatments were also able to diminish H3K9me3 levels, likely through heterochromatic decondensation (Bian et al., 2009, Tamada et al., 2006). Critically, however, *Xenopus* oocytes were not able to induce the expression of *Nanog*, suggesting that *Xenopus* and mammalian pGRNs are substantially different (Byrne et al., 2003).

Taken together, it would be reasonable to conclude through studies of *Xenopus* alone that pluripotency factors identified in mammals were not as critical to embryonic development in the amphibian-mammal common ancestor some 352 MYA (Kumar et al., 2017). Similarly, it could be concluded that the pGRN diverged around the time amniote common ancestor emerged; however, studies in urodele amphibians, namely axolotl (*Ambystoma mexicanum*), suggest that this may not be the case. While indeed an argument can be made that the frog embryo as a whole is pluripotent in that it produces both the soma and the germline. Crucially however, these come from different groups of cells with differing cellular potency, therefore frogs do not display cellular pluripotency which is conserved throughout metazoans.

1.1.1.7 Pluripotency in axolotl

The pluripotent properties of axolotl tissue were first identified using *ex vivo* explants of the animal cap cells taken from the top of the animal pole. Animal caps first acquire the competency to respond to mesoderm and endoderm inducing signals at late-cleavage stages of development, with competence peaking at the mid-late blastula stage, reducing toward gastrulation (Nieuwkoop, 1969). By dividing up blastula stage (stage 8-9) embryos into zones and observing their natural developmental progression, as well as the explanting and culturing these tissue zones in isolation, Nieuwkoop determined the relative contribution of each zone to the embryo (Shown in Fig. 1.2). In wild-type embryos, zone I develops into neuroectoderm, zone II develops predominantly into neuroectoderm and contributes to some mesoderm, zone III contributes to both mesoderm and endoderm derivatives and zone IV is comprised mainly of nutritive yolk that is absorbed by the embryo during development. When explanted and cultured in isolation, zones I and II form 'undifferentiated' ectoderm, Zone III forms ectoderm, mesoderm and endoderm, while zone IV develops only into 'undifferentiated' endoderm. Importantly, this demonstrated that the animal hemisphere contributed to all three germ layers. By combining the different zones of the embryo, Nieuwkoop showed that mesoderm and endoderm could be induced in the upper animal cap (zones I and II) in response to signals from the underlying endoderm (zone IV). This demonstrated unequivocally that while the position of animal cap cells may indicate a proclivity toward a particular route of differentiation, cell fate could be changed in response to signalling.

Later work demonstrated that AC's were able to be induced to form PGC's when explants of zone I and IV were combined (Boterenbrood and Nieuwkoop, 1973). Later work demonstrated that this was the result of FGF and BMP signalling, which specifies germline competency in the pluripotent AC (Chatfield et al., 2014). Crucially, this work demonstrates that axolotl ACs manifest a true pluripotent cell state which differs from other amphibians such as Frogs. While animals that specify PGC's conditionally like Frogs can be described as having a 'pluripotent embryo' in that the embryo does give rise to both the germ line and the soma, critically there is no cell state which is competent to produce all these lineages.

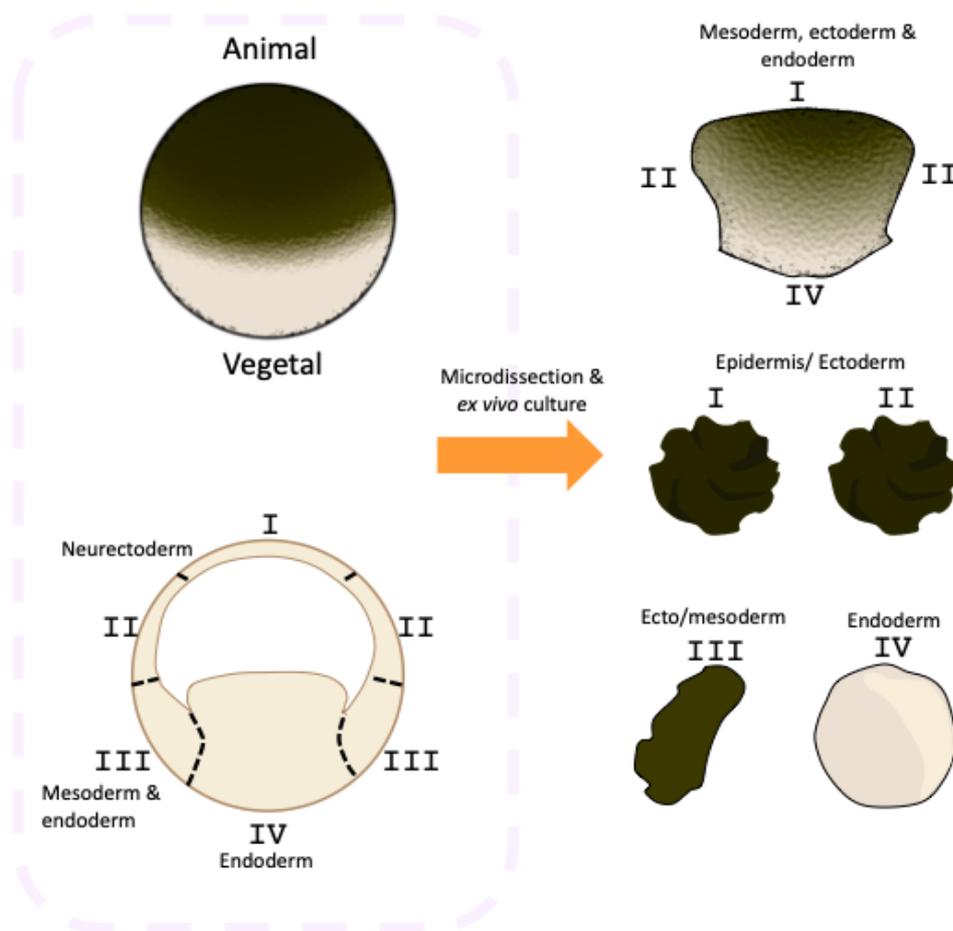


Figure 1.2. Differentiation potential of axolotl embryo explants.

Intriguingly, the axolotl genome has retained *NANOG*, *SOX2* and *POU5F1* as well as the *POU5F3* gene retained in anurans and chickens (Bachvarova et al., 2004, Dixon et al., 2010, Fei et al., 2014, Frankenberg et al., 2014). Axolotl *POU5F1* is expressed in the animal pole of blastula stage embryos and in the presumptive ectoderm and mesoderm of gastrula stage embryos (Bachvarova et al., 2004). Axolotl *POU5F1/OCT4* has been shown to rescue *Pou5f1/Oct4* deficient mESCs and *Xenopus* embryos depleted of all three homologues, suggesting that its role may have some overlapping function (Morrison and Brickman, 2006). Preliminary evidence (unpublished) demonstrates that *POU5F1/OCT4* KD in axolotl embryos arrests development at the late-blastula stage. *POU5F1/OCT4* KD also prevents mesodermal induction in explanted animal caps in response to activin, FGF or BMP. Thus, preliminary evidence indicates that Axolotl *POU5F1/OCT4* may also facilitate the acquisition of pluripotency and concomitantly the competence to produce mesoderm. Like *Sox2* in chickens and frogs, axolotl *SOX2* is also expressed in the developing neural tube. However, knockout of *SOX2* does not affect neural development, possibly owing to functional redundancy with *Sox3*, which is co-expressed in the same tissues. Axolotl *SOX2*, however, is required for spinal cord regeneration (Fei et al., 2014).

Axolotl *NANOG*, like *POU5F1*, is expressed in the pluripotent animal hemisphere on early blastula stage embryos (Dixon et al., 2010). Overexpression of Axolotl *NANOG* is able to promote LIF independence in mESC. Further to this, Luciferase assays demonstrated that axolotl *NANOG* was capable of driving transcription from the mammalian *NANOG* promoter (Dixon et al., 2010). Further experiments utilising over-expression of ChIP-qPCR showed that following overexpression of myc-tagged axolotl *NANOG*

in mESC, there is observable enrichment of axolotl *NANOG* at the *Nanog* and *Pou5f1* promoters.

Indeed, the differences between the regulatory networks in axolotl and *Xenopus* are exemplified in studies utilising oocyte extract remodelling of permeabilised mammalian cells. Despite both being amphibians, Axolotl oocyte extracts (AOE) and *Xenopus* oocyte extracts (XOE) possess differing abilities to remodel mammalian cells. Like XOE, AOE treatment induces replication-independent DNA demethylation and is also capable of decreasing the expression of nuclear lamins, which are expressed in somatic cells (Alberio et al., 2005). Unlike XOE, AOE treatments result in the acquisition of several epigenetic properties associated with pluripotency in mESC. Fluorescence-activated cell sorting (FACS) demonstrated global levels of repressive H3K9me3 and HP1 α marks were reduced to a level comparable to mESC. The study also showed global increases in H3K9ac, suggesting comprehensive increases in gene expression, a key characteristic of mESCs. Specific loci were also investigated to clarify whether the observed changes in gene expression were specific. The data showed that the promoter of *Pou5f1* was significantly demethylated after just 5 hours of incubation. Real-time qPCR later revealed that AOE treated cells showed increased Oct4 expression over control cells after two weeks in culture. Real-time QPCR showed increased *Nanog* expression after just 5 hours of incubation. Given that *Nanog* expression is generally considered an endpoint in reprogramming to pluripotency, this finding is particularly interesting. Given their embryonic origin, MEFs already display low levels of methylation at the *Nanog* promoter; however, they do not express Nanog. Therefore, AOE induced promoter demethylation may have allowed greater accessibility for transcription factors

within the oocyte extract. With this in mind, the inability of XOE to activate Nanog may suggest changes within the pluripotency gene regulatory network following the loss of the NANOG gene in Anurans. Replication independent demethylation in response to AOE treatment was also demonstrated by FACS using a 5-MeC antibody staining. FACS sorting showed that levels of 5-MeC in somatic nuclei were comparable to those of mESC after AOE treatment (Bian et al., 2009). Together this suggests that axolotl oocyte factors are able to somewhat reprogram mammalian cells to a state which has many features of pluripotency. Further to this, it implies that the core pGRN may be conserved in urodele amphibians and mammals but may have diverged to a greater extent in anurans.

1.7 The germline-soma relationship

The term mutation describes a change in the DNA sequence within an organism. Mutations can occur at any point in an organism's lifespan, where these mutations occur have profoundly different effects on evolution. In early metazoan development, PGC's are specified from the soma. PGC's give rise to gametes, themselves capable of giving rise to future generations of organisms. The somatic tissues are the precursors to all the other cells of the organism. Because of this segregation, mutations in an organism can occur either in the soma or in the germline. Somatic mutations generally can only affect the reproductive fitness of the individual in which they occur, whereas germline mutations can affect the reproductive fitness of the individual and of future generations of organisms. Germline mutations, therefore, can affect both somatic genes and germline genes of progeny.

Mutations can arise from a variety of factors, including high-energy sources such as radiation or chemicals in the environment; mutations may also occur as DNA replicates during mitosis. Broadly mutations can be split into two categories: point mutations and chromosomal aberrations. During DNA replication in the human genome, for example, around one in every 10,000,000,000 base pairs are mistakenly altered. While less common, bases may also be inserted or deleted. Chromosomal aberrations are far larger-scale mutations that generally occur during meiosis as a result of aberrant crossing over events, slippage during DNA recombination or due to the activities of transposable elements. These mutation events can result in the rearrangement, duplication or deletion of individual genes to entire chromosomes. Transposable elements, for example, can traverse the genome enabling large scale chromosomal organisation, which can have an array of effects. Insertion or removal of these elements at cis-regulatory loci can drastically alter transcriptional regulation through the formation of novel enhancers or insulators (Acemel et al., 2017, Maeso et al., 2017, Maeso and Tena, 2016). Therefore, germline mutations can be considered to be the single driving force behind evolution, creating ever greater phenotypic diversity. The process of meiosis itself appears to have been maintained in sexually reproducing populations because it increases phenotypic variation (Gould, 2002; van Valen, 1960). Concomitantly, because mutations are constantly acquired, they drive diversification, and thus any given species only exists within a window of time. Whether or not such germline mutations are carried forward is not dependent on whether it has an advantageous effect on reproductive fitness but rather whether it has an unresolvable detrimental effect. Indeed, DNA sequences can be carried forward with no apparent beneficial effects and even

moderately detrimental effects (Dawkins, 1976, Agren and Clark, 2018). However, if a mutation is detrimental to the point at which it compromises its own reproduction, then it will be eliminated.

At one level, these heritable mutagenic processes have driven the evolution of developmental GRNs. These finely controlled networks directly orchestrate the tissue composition and body plan of the developing organism (Davidson and Erwin, 2006, Erwin and Davidson, 2009). Moreover, changes within these systems can lead to both subtle and major phenotypic innovations (Britten and Davidson, 1971, Rebeiz and Tsiantis, 2017, Shubin et al., 2009). Nevertheless, the conservation of phylogenetic patterns, defined originally by conserved morphological traits and more recently by genome sequence, suggests underlying constraints. So, if germline mutations are considered the driving force behind diversification, constraint can be considered the guiding force that leads to stable phenotypic traits. Therefore, understanding which elements of development are conserved, and which are species or even order specific can provide powerful insights into the evolution of species.

Given their stochastic nature, most germline mutations are detrimental to reproductive fitness; they are more likely to be eliminated from a population over concurrent generations. However, it has been proposed that adverse somatic gene mutations may be better tolerated than germline gene mutations as the latter may be more likely to disrupt germline transmission and thereby removing themselves from the gene pool (Johnson and Alberio, 2015). Given that in the majority of animals, PGC's and soma are specified from pluripotent progenitors in response to signalling, then similarly, mutations that affect somatic development that perturb the formation of the germline would be equally defunct from an evolutionary perspective.

Incidentally, among metazoans, two forms of PGC specification have evolved, the aforementioned mechanism known as epigenesis whereby PGC's are specified via signalling from pluripotent precursors and preformation where maternally localised RNAs are inherited by cells during cleavage, thereby sequestering the germline away from the soma early in development. Indeed, this mechanism has been proposed to stabilise germline development and relieve constraints for mutations that could affect the early development of the soma (Johnson and Alberio, 2015, Johnson et al., 2003b, Evans et al., 2014). It has also been observed that animals which specify PGC's by preformation are more speciose and appear to evolve quicker when compared to closely related taxa, which have kept the basal mode of epigenesis (Crother, 2018, Johnson and Alberio, 2015, Evans et al., 2014). Curiously, birds, frogs and teleost fish all specify their germ cells via a mechanism of preformation, and this correlates with the loss of known germ cell marker and pluripotency factor *POU5F1* in the same taxa (Frankenberg et al., 2014).

1.8 Aims and objectives

While pluripotency is well understood in the context of ESC, ESC are not necessarily representative of an embryo, specifically ESC are held in a state of pluripotency by extracellular signalling in a process called self-renewal. This is useful for assaying factors, they can be knocked down/out or overexpressed and the effect of these perturbations on self-renewal can be studied. However, *in vivo* pluripotency is a transient state and therefore it can often be difficult to generalise *in vitro* findings to embryogenesis. Even less is known about pluripotency in non-mammalian vertebrates given that Zebrafish and *Xenopus* the two major models for non-mammalian vertebrate development

do not possess cellular pluripotency. Therefore, in this thesis I sought to know how specific pluripotency factors function in early development. I focused on two different transcription factors NANOG and ELK1.

In chapter 3, I investigated the role of NANOG in axolotl early development. NANOG has been the subject of a large body of research due to its function as a master regulator of pluripotency (discussed in detail in chapter 1.1.2.2 and chapter 3.1). Building on previous work carried out by our group, I attempted to address the following questions:

- Is NANOG required for axolotl development?
- Does depletion of NANOG effect early cell-fate decisions?
- Does NANOG regulate the expression of other pluripotency factors?
- Which genes are affected by NANOG depletion?
- Are there any similarities between the effects of NANOG depletion in axolotl and other animals?
- If there are any comparable effects, are the mechanisms by which these effects carried out also conserved?

The second factor ELK1, I investigated in chapter 4. While well characterised in a variety of different cellular contexts, ELK1 has not been extensively explored in the context of development. While it has been suggested that Elk1 plays a very pivotal role in the regulation of pluripotency downstream of FGF signalling, this was defined in hESC. Despite this role in humans *Elk1* knockout mice show no outward phenotype (discussed in chapter 4.1). Indeed, the only other *in vivo* knockdown of Elk1 was performed in *Xenopus* which as mentioned, do not possess cellular pluripotency. Therefore, I also

examined the role of ELK1 in axolotl development and address the following questions:

- Is ELK1 required for early development?
- Does ELK1 regulate the expression of pluripotency factors?
- Does ELK1 depletion effect early cell fate decisions?
- Are there any similarities between the effects of Elk1 depletion in other animals such as humans, mice or frogs?
- Does ELK1 act downstream of FGF?
- Which domains within the ELK1 protein are necessary for its function?

To address my research questions, I performed a series of knockdown and rescue experiments in axolotl embryos and aimed to characterise the functions of each gene using a variety of investigative tools including *ex vivo* inductive assays, high-resolution episcopic microscopy HREM, whole RNA sequencing and quantitative PCR (QPCR).

2.2 Axolotl embryos and explants

Embryos were collected following matings as described previously (Johnson et al., 2001). For microinjection, embryos were manually de-jellied and cultured in 1× modified Barth’s solution (MBS) with 4% Ficoll (Sigma) and injected at the one or 2-cell stage. Embryos were staged according to Bordzilovskaya and Dettlaff (1979), which are approximately equivalent to Nieuwkoop and Faber’s (1994) stages of *Xenopus*. From stage 7 onwards, embryos were maintained in 0.2× MBS, and dissected explants were maintained in 0.7× Marc’s modified ringers solution (MMR). Culture solutions were supplemented with antibiotics (50 µg/ml penicillin and streptomycin, and 50 µg/ml kanamycin), 100 µg/ml Ampicillin and 50 µg/ml fungizone). Embryos were staged as described previously (Bordzilovskaya et al., 1989).

2.3 Morpholino and RNA microinjections

Morpholino oligonucleotides (GeneTools, LLC, OR) were designed to block translation or disrupt target splice junctions. Intron/exon boundaries were predicted by homology. Sequences were obtained from our axolotl genomic resource (Evans et al., 2018) or via the axolotl genome website (Nowoshilow et al., 2018). The morpholino sequences used were as follows: Translation MO: *Nanog*, 5'- GGTC AATCCAAAAGCTCCTCCTAAG-3'; Splice MO: *Nanog* 5'- GGCAGGACTGAAACAAAACGAAGAC-3'; Translation MO: DPY30 5'-

ATGCTTTGTTCCGACTCCATTGTGA -3' and 5'-
TTATCGTAGCCCGTCACTCCAGCTC-3'. A nonspecific morpholino was
injected in each experiment at equivalent levels to the specific splice
morpholino combinations: MO: Control, 5'-
GGATTTC AAGGTTGTTTACCTGCCG-3'. Each morpholino experiment was
repeated at least three times, and the efficacy of the splice morpholinos was
tested by PCR in each experiment using primers detailed in supplementary
table S1. The Activin-nodal inhibitor SB431542 (Sigma) was solubilised in
dimethyl sulfoxide and used at a final concentration of 150 μ M. In vitro
transcription and microinjection mRNAs for microinjection were synthesised
using mMessage mMachine (Ambion) from plasmids encoding; *Xenopus*
eFGF, *Xenopus* Smad2C (XSmad2C), human *NANOG* and human DPY30.

High-resolution electron microscopy (HREM)

Samples were fixed overnight at four °C in 4% PFA in PBS before being
dehydrated overnight in increasing concentrations of methanol (50%, 60%,
70%, 80%, 90%,100% and 100%) before being mounted in JB4 medium with
acridine orange (SIGMA), and processed for HREM as described by Mohun
and Weninger (2012). Mounted samples were sectioned at 2 μ M. Section
images were converted to JPEG format using the image processor function on
photoshop. Embryo 3D reconstructions were created using the 3D volume
rendering function in Osirix MD.

2.4 Western blotting

For Western blot analysis, whole embryos were lysed in RIPA buffer and
homogenised using a Dounce homogeniser. Cell lysates were then centrifuged

at 13,000 rpm at four °C for 5 minutes. The supernatant was then collected, and protein concentration was assessed using Protein Assay reagent (BIORAD) according to the manufacturer's instructions. Approximately 25 µg of protein was used for each well for subsequent SDS-PAGE performed using standard methods. To test antibodies raised against axolotl *NANOG*, synthetic poly-A RNA encoding *NANOG*-HA were injected into mature *Xenopus* oocytes, from which lysates were prepared for Western blotting. In addition, an uninjected *Xenopus* oocyte lysate and a multi-antigen containing cell lysate were also used as HA-negative and positive controls, respectively. Extracts were produced using a process outlined by Hutchinson *et al* (1988). Axolotl *Nanog* antibodies were produced by DundeeCell and raised against antigenic peptides specific to Axolotl *NANOG*. All other antibodies are commercially available and listed in supplementary table S1, as well as the used concentrations.

1.9 Immunofluorescence

For paraffin-embedded sections, embryos were fixed overnight at four °C in 4% PFA in PBS before being dehydrated overnight in increasing concentrations of methanol (50%, 70%, 80%, 90% and 100% before being processed in xylene and mounted in paraffin wax. Slides were dewaxed with xylene for 30min and rehydrated with decreasing concentrations of ethanol (100%, 90% and 70%) and transferred to PBS for 15min before processing for immunofluorescence. For cryo-sections, fixed embryos were incubated in 30% sucrose/PBS overnight at four °C prior to mounting in OCT compound and sectioning. Sections were left to dry for at least two h before immunofluorescence. For both cryo-sections and wax sections, antigen

retrieval was performed by boiling the slides in 0.01M citrate buffer (pH 6.0) for 10min. Sections were permeabilised with 1% Triton X-100 in PBS for 15min. Slides were then immersed in blocking solution (PBS supplemented with 5% BSA for two h. After blocking, sections were incubated with primary antibody (see Extended Data Table S1) overnight at four °C in a humidified chamber. Slides were then washed three times with 0.1% Tween-20/PBS before being incubated with a fluorescent secondary antibody (Extended Data table S2) for 45min at room temperature. Slides were treated with mounted with Fluoroshield with DAPI (Sigma) and sealed with nail varnish. Slides were kept at -20°C until observed.

2.5 First-Strand cDNA synthesis

A reaction consisting of 1µl of Oligo(dT)₂₀ (50uM), 2ug of total RNA, and 1µl of 10mM dNTP mix was topped up to 13µl using sterile distilled water. The reaction was heated to 65°C for 5 minutes and then incubated at 4°C for 5 minutes. The tube was briefly centrifuged before the addition of 4µl 5X First-Strand Buffer, 1 µl 0.1M DTT, 1 µl RNaseOUT recombinant RNase inhibitor (40 units/ µl), and 1 µl SuperScript III RT (200 units/ µl). All reagents from Thermo Fisher Scientific. The reaction mixture was then heated to 25°C for 5 minutes, 50°C for 60 minutes, and 70°C for 15 minutes before being chilled to 4°C.

2.6 Quantitative PCR

QPCR was performed on cDNA libraries prepared as previously described (Swiers et al., 2010) and assayed with SYBR green or TaqMan probes and primers listed in supplementary table S2. Each qPCR reaction contained 5 µl

of SYBR-Green Jumpstart Taq ReadyMix (SIGMA), 1 μ l (10uM [final]) of each of the forward and reverse primers (SIGMA), 2 μ l of nuclease-free water, and 1 μ l of cDNA template. Each reaction was prepared in triplicate on an ABI FAST Systems 0.2ml 96-well PCR plate (STARLAB) and then sealed with an Optical Adhesive cover (Life Technologies).

The following qPCR conditions were utilised for each run on the QuantStudio 6 Flex (Life Technologies) qPCR instrument: The plate was heated to 105°C to activate the Jumpstart Taq before being held at 50°C for 2 minutes. The initial denaturation step was at 94°C for 10 minutes, followed by 40 cycles of 94 °C for 15 seconds (denaturation) and 60 °C for 1 minute (annealing and extension). The raw data was then extracted to be analysed by comparative CT (Cycle threshold). An endogenous control gene was chosen for each qPCR run to normalise the data in order to compare the relative fold change of target genes amongst cDNA samples.

The double delta CT value ($\Delta\Delta$ CT) was calculated using the following formula: Δ CT target- Δ CT reference= $\Delta\Delta$ CT. As all calculations are in logarithm base 2, the expression fold change of each target gene was calculated using $2^{-\Delta\Delta$ Ct.

Table 1. Primers used in this study

Gene	Forward/reverse	Sequence
<i>ACTA1</i>	Reverse	ATCCACATCTGCTGGAAGGT
<i>ACTA1</i>	Forward	CGCATGCAGAAGGAGATCAC
<i>BRACHYURY</i>	Forward	CATTGACCACATGTACCAA
<i>BRACHYURY</i>	Reverse	GATCAAGGGTCAATCGTGAGTTC
<i>BRACHYURY</i>	Reverse	TGTGTCCACTCCTCACCTTACTAT
<i>BRACHYURY</i>	Forward	GAAGAAGCAAGTGTCTGGAGAGAG
<i>C8B TSS</i>	Forward	CGATTCTCAAGCTCGGCAAC
<i>C8B TSS</i>	Reverse	CTGTACTCCAAAGGAAGACTGT
<i>CYTOK</i>	Reverse	GGAGCCGCGTCCATCTC
<i>CYTOK</i>	Forward	AACCACCAAGAGGAATTGCAA
<i>CYTOK TSS</i>	Forward	TATTATAGTGGTGCGGTGCC
<i>CYTOK TSS</i>	Reverse	AGGTGGAGATGCGTAGTTC
<i>DES</i>	Reverse	GTCTGGATTGGCATGGTGAC
<i>DES</i>	Forward	CGAGATCCGCAACTTGAAGG
<i>DND1</i>	Forward	CAACCTAATCCAGATCAATG
<i>DND1</i>	Reverse	CGTAGATGTCTTGAGGTA
<i>EEF1A1</i>	Reverse	CATGCAGCCAACGAACTATGTATT
<i>EEF1A1</i>	Forward	GTATGATGAGGTTCTGTGCATTG
<i>ENDODERMIN</i>	Reverse	GAAACCCGTAGGTGGACAGAGA
<i>ENDODERMIN</i>	Forward	CTGACCAGGAGGGAAAAGCTT
<i>FABP1</i>	Reverse	CCGTTCTGCTCCATCTCAGA
<i>FABP1</i>	Forward	AGTCCAGTCCCAGGAAATC
<i>FLK1</i>	Forward	GACTCAGAAAAGACACTG
<i>FLK1</i>	Reverse	GCAACTTGATCTTGTAATAAC
<i>FOXA2</i>	Forward	ACGACTGGAGCAGCTACTAC
<i>FOXA2</i>	Reverse	CGGTGTTACGTAGGACATG
<i>GATA4 TSS</i>	Forward	GCAGGATTGGCAGACACAAG
<i>GATA4 TSS</i>	Reverse	TCTGGGTCCGTCTCCTCA
<i>GRHL1</i>	Reverse	CTTTGTTTGGCCGTGTGCTG
<i>GRHL1 TSS</i>	Forward	CAACCCGAAAGTCCAGTTC
<i>GRHL1 TSS</i>	Reverse	TTCCAGGCTTCGTCTTCACT
<i>GRHL2</i>	Forward	CGGCAACAAAGGCATACACC
<i>HAND1</i>	Reverse	CTGTCCGCCCTTTCATCTTC
<i>HAND1</i>	Forward	CGGGCCGAGCTGAAGAA
<i>HNF4G TSS</i>	Forward	CCGAGTCCAAGTTTCTGCA
<i>HNF4G TSS</i>	Reverse	TGAGACTCAGCATCCAAAAGG
<i>HOXB3 TSS</i>	Reverse	GGTTTTGGTAGCTCGTGGAC
<i>HOXB3 TSS</i>	Forward	AGGACCAGAAGATCAAGGGC
<i>HOXB9 TSS</i>	Forward	CTTCTGCGCCGGTGATTTAC
<i>HOXB9 TSS</i>	Reverse	GGTGGGAAGAGAGGAGACG
<i>HOXC4 TSS</i>	Forward	GTGCCAATTCAGTAGTGCCG
<i>HOXC4 TSS</i>	Reverse	AAAGGAACGTAGACCCAGCC
<i>HOXC6</i>	Reverse	TGGCGATTTCGATCCTCCTG

<i>HOXC6</i>	Forward	ATGAATCCCACAGTGGCGT
<i>HOXD1 TSS</i>	Forward	CCTTTGTCATTGAGCAGCGA
<i>HOXD1TSS</i>	Reverse	CTCGGCTGAAGTCCTGGA
<i>MIX</i>	Reverse	GCTTCTGGGTGGATTTGATTTATAA
<i>MIX</i>	Forward	GTCCAGGATCCAGGTCTGGTT
<i>MYF5</i>	Reverse	GGCGCTGTCAAAGCTGTG
<i>MYF5</i>	Forward	GGGAGCCCCCTTTCCAA
<i>MYF5 TSS</i>	Reverse	ACTAGGTCCATGCTGTCACC
<i>MYF5 TSS</i>	Forward	CAGATCAGGACCTTGCCCTC
<i>NANOG</i>	Forward	ACTTTACCAAAAAGCGTGACACTAGA
<i>NANOG</i>	Reverse	ACAGAGCACCCAATTTTCCAA
<i>NANOS3</i>	Forward	CTGGTGGAGTACTGATAC
<i>NANOS3</i>	Reverse	GAGTCTAAGCGAATCTACA
<i>NCAM</i>	Forward	GCCCCTAAGTTGCAAGGCC
<i>NCAM</i>	Reverse	TCTCGTTTGTCTGTGGGGC
<i>NESTIN</i>	Forward	GTGGTTGAAGGAGAAAGGCG
<i>NESTIN</i>	Reverse	GTAGTCCTCGATCTCCACGG
<i>NEUROD4</i>	Reverse	GGTGGCTGACAAAAGAAGGG
<i>NEUROD4</i>	Forward	ACTATCACAAACCGACCAGCA
<i>NEUROG</i>	Reverse	CTCCTCGTCCGAGTAAGA
<i>NEUROG</i>	Forward	GTCGTTCAAACCGAGAG
<i>NKX2.5 TSS</i>	Forward	GAGGACATGACTGCTCTGG
<i>NKX2.5 TSS</i>	Reverse	GGTGACGGGGCTGTGGAA
<i>NODAL1</i>	Reverse	GGGTCGGGTGGTACAGCTT
<i>NODAL1</i>	Forward	CCCAGTGGATGAAACGTTACG
<i>NODAL1 TSS</i>	Forward	CGTTCACAGCCCACAAATA
<i>NODAL1 TSS</i>	Reverse	ACCCTCAAAGCGAAAATCCG
<i>NODAL2</i>	Reverse	CCCGCTCTGGAATGTACAATTT
<i>NODAL2</i>	Forward	CATACCGCTGTGATGGAAAGT
<i>ODC1</i>	Reverse	CCCGGACCCAGGTTACG
<i>ODC1</i>	Forward	ATGCCCGTCATGAGTAGTACCA
<i>PAX6</i>	Forward	GAAGTGGAGAAGGGAAGAGAAACTG
<i>PAX6</i>	Reverse	TGATGTAAATGAAACTGGTGTCTGTG
<i>PIWI</i>	Forward	AGCTTCGTTAAGAGCTTGGTTC
<i>PIWI</i>	Reverse	CGCCATCACGGTAGACAATG
<i>POU5F1/OCT4</i>	Reverse	ACATCCGCCTGCGTAAAGC
<i>POU5F1/OCT4</i>	Forward	GCGGACCTTGAACAGTTTIGC
<i>RUNX1</i>	Reverse	TTGGGAAGTGTCTGGTCAATTC
<i>RUNX1</i>	Forward	CGCCTCTCTGGTGCATCTG
<i>SOX11</i>	Forward	TGAGCCTGAACTTCTCCTCG
<i>SOX11</i>	Reverse	AGTCTGAGAAGTTGGCCTCC
<i>SOX17</i>	Forward	TTTTTGTGGAAACCTATGGGCCAC
<i>SOX17</i>	Reverse	CGATTTTTATTAGCCGACCACACA
<i>SOX2</i>	Reverse	GGCAGGTACATGCTAATCA
<i>SOX2</i>	Forward	GGTCAAGTCCGAATCGAG
<i>SOX21</i>	Reverse	GTGTGTGTGCGTGTCTTCAT

<i>SOX21</i>	Forward	AGAAGGGCCTTGCAAAATGG
<i>SOX7</i>	Forward	CGAGCTGCTAGAGATGGACA
<i>SOX7</i>	Reverse	GTTGTAGTATGCGGCTGTGG
<i>SOX8</i>	Reverse	TGCTTGCCTCCATGATGAAG
<i>SOX8</i>	Forward	CGGGAGGCCAACTCTACAA
<i>VASA</i>	Reverse	TGTTTGCCGTTCTTCTTTGGT
<i>VASA</i>	Forward	GATCGAATGCTTGATATGGGTTT
<i>VENTX</i>	Forward	TTCAAGCGGCAGAAGTACCT
<i>VENTX</i>	Reverse	CACGTCTGAGCTGCACTTC
<i>VIM</i>	Reverse	CGTCCTGAAGGCGGTTAATG
<i>VIM</i>	Forward	CGAAGTGGATGCCCTGAAAG
<i>VIM TSS</i>	Forward	GTGACTCAGTCCAGTAGCCT
<i>VIM TSS</i>	Reverse	TCCGGGGCGCTCTTAATAC
<i>WNT8</i>	Forward	GCAGAGTTCAGGGATATAGGCAAT
<i>WNT8</i>	Reverse	TCCTGCGCTGTCCATCTC

2.7 RNA-seq analysis

For axolotl sample only RNA-seq, each sample was mapped to the transcriptome assembly described in Evans et al. 2017 using RSEM with default parameters. This was used as it also contains several developmental genes identified by our group in previous publications. Two of the *NANOG* *KD* stage 22 biological replicates were removed based on their lack of correlation with all other samples. The resulting count files were input to edgeR to calculate differentially expressed genes ($FDR < 0.05$, $\logFC > 1$). Heatmaps were drawn using the \log_2 TPM values in R using heatmap.3 with default cluster settings.

The seventeen stages of wild-type axolotl early development was acquired by mapping the reads from the Jiang *et al* (2018) datasets to the same transcriptome assembly as before. Mean TPM values for each stage were calculated in R. These values were then used to determine whether each gene was 'early', 'late' or 'global'. Genes that were only expressed with a $TPM > 10$ before stage 12 were classed as 'early', genes that were only expressed with a $TPM > 10$ on or after stage 12 were classed as 'late'. The remaining genes were all considered global and were divided into 'global high and 'global low based on whether they had a $TPM > 10$ in every stage of the Jiang data or not. In order to construct the heatmap displaying every sample condition vs every sample condition, the heatmap row dendrogram was split at a height of 21, and each cluster was assigned the most common expression pattern of the genes within that cluster.

2.8 Comparison of Human, Pig, Xenopus and tissues

In total, single-cell transcriptomes from 144 pig peri-gastrulation stage epiblast cells from Ramos-Ibeas *et al.* (2019), 152 human peri-gastrulation stage epiblast cells retrieved from Petropoulos *et al.* (2016) were merged by calculating the geometric mean across all cells to get the simulated average count for each pluripotent tissue, and this was then compared with whole RNA-seq data from st10.5 *Xenopus* animal caps retrieved from Angerilli *et al.* 2018 and stage 10.5 axolotl animal caps (our study). Given the differences in sequencing methodology and sample preparation etc, I employed quantile normalisation to compare the relative abundance of expressed genes. For comparing whether a gene's expression profile is conserved, genes were grouped into two categories 'high' and 'low' a gene's expression was determined to be relatively high if its expression was higher than the median expression across all four organisms. Correspondingly, a gene's expression was deemed low if it was lower than the median expression value across all four species.

2.9 Gene-set enrichment using *Xenopus* cell type markers

The R package hypeR was used (Federico & Monti, 2020) to perform a custom gene-set enrichment analysis of differentially expressed genes using a list of amphibian cell type markers identified by Briggs *et al.* (2018). Analysis was carried out using an FDR threshold of 0.01. Cell types that were significantly enriched in the dataset were plotted using GOchord in R.

2.10 ChIP qPCR

50 AC explants per experimental condition were processed using the truCHIP Chromatin Shearing Tissue Kit with Formaldehyde (Covaris). Caps were fixed with 1% methanol-free formaldehyde-PBS (Covaris) for 15 minutes before quenching. Fixed caps were processed according to the manufacturer's protocols. 1ml of chromatin was sheared using the S220 ultra-sonicator (Covaris) in AFA fibre containing vesicles (Covaris). Chromatin shearing was performed under the following conditions: duty cycle 5%, intensity level 4, cycles/burst 200. The shearing program was run for 15 minutes resulting in chromatin fragments of 100–500 bp. 100 μ l of sheared chromatin was used per ChIP using the EZ MAGNA ChIP kit (Millipore) according to the manufacturer's instructions. Purified DNA was then used for subsequent qPCR analysis. QPCR primers are listed in supplementary table S2.

2.11 Experimental numbers and repetitions

For each experiment the figure legend will indicate the number of embryos used for each experiment. "n" represents the number of independent biological samples (for example embryos) per experimental condition. "m" represents the number of separate matings. "t" represents the total number of biological samples used.

Nanog and TGF- β signalling

establish developmental competency

3.1 Introduction

Embryogenesis is directed by sequential waves of gene activation and inactivation that direct cell fate decisions, leading to loss of cellular potency and the differentiation of adult tissues. The cellular potency of the developing zygote is therefore determined by the tissue complexity of the animal. The mammalian zygote, for example, forms from a totipotent cell, which gives rise to both embryonic and extraembryonic tissues. This stands in contrast to the anuran zygote, which gives rise to distinct lineages with differing cellular potency; the multipotent cells of the animal hemisphere are capable of producing the soma, while unipotent, germplasm containing cells within the vegetal hemisphere produce the germline. Pluripotency is therefore defined as a cell state that is competent to produce both the soma and the germline in response to signalling, such as the cells of the mammalian epiblast (Gardner and Rossant, 1979, Tam and Zhou, 1996). While the extraembryonic tissues of mammals and the germplasm of frogs are examples of derived developmental traits, pluripotency, appears to be conserved in vertebrate development. Therefore, understanding pluripotency and its underlying gene regulatory network (pGRN) can enable a better understanding of vertebrate evolution.

The pGRN has been studied extensively *in vitro* using embryonic stem cells (ESC). In this case, cytokine signalling is able to maintain the balanced

expression of a core set of pGRN factors allowing indefinite propagation and retention of developmental potential. Consequently, sufficient disruption to key components of the pGRN results in the activation of differentiation programs (Smith et al., 1988, Boyer et al., 2005, Li and Belmonte, 2017, Li and Izpisua Belmonte, 2018). The study of the pGRN *in vitro* led to the discovery of *Nanog*, a core component of the transcriptional pGRN, and its expression marks the pluripotent domain in early mammalian embryos (Chambers et al., 2003, Mitsui et al., 2003).

In mice, *Nanog* is first expressed in internal cells of morula stage embryos (Chambers et al., 2003, Goolam et al., 2016), which it biases towards the development of the embryonic lineage (Torres-Padilla et al., 2007). Later, *Nanog* contributes to patterning the ICM towards epiblast or primitive endoderm specification (Chazaud et al., 2006, Frankenberg et al., 2011, Messerschmidt and Kemler, 2010). Given this prominent role in establishing the pluripotent domain, it is perhaps unsurprising that the ICM of mouse embryos harbouring a homozygous deletion of *Nanog* fail to establish pluripotency and are not viable (Mitsui et al., 2003, Silva et al., 2009). The *Nanog* null ICM aborts development prior to specification of the primed status of the mature epiblast, diverting development toward trophoblast development or engaging the apoptotic pathway.

Human (h)ESCs manifest a pGRN, which has been equated to that of mouse epiblast-derived stem cells (EpiSCs), as their self-renewal requires TGF- β signalling, stimulated by ACTIVIN *in vitro* and NODAL *in vivo* (James et al., 2005, Vallier et al., 2005). TGF- β signalling triggers phosphorylation of SMAD2/3 (Gaarenstroom and Hill, 2014, Massague, 2012), which is then

translocated to the nucleus where it forms a complex with NANOG (Vallier et al., 2009a). NANOG and SMAD2/3 also associate with the Trithorax complex, which contains the MLL/SET family of histone methyltransferases that deposit trimethylated Histone H3 lysine 4 (H3K4me3) on specific loci in the hESC genome (Bertero et al., 2015). H3K4me3 is a transcription promoting histone modification that overlaps with SMAD2/3 binding sites in the enhancer regions of pluripotency genes in hESC (Bertero et al., 2015, Ruthenburg et al., 2007).

The regulatory protein DPY30 is a fundamental component of the MLL/SET complex, which synergizes with NANOG and SMAD2/3 to facilitate H3K4me3 deposition in hESCs. DPY30 KD in hESCs severely impairs self-renewal, driving cells toward neural differentiation (Bertero et al., 2015). Mouse embryos harbouring a deletion of the gene encoding DPY30 fail to establish a 'primed' pluripotent epiblast (Bertero et al., 2015); concomitantly, this effect is mimicked by deletion of the *NODAL* gene (Camus et al., 2006, Conlon et al., 1991, Mesnard et al., 2006). Together these results suggest that the association of the NANOG-SMAD2/3 –Trithorax complex may program the primed state of pluripotency in the epiblast during embryogenesis in mammals, but because the removal of these factors results in the degradation of the embryo, the underlying mechanisms are not fully understood.

Phylogenetic analysis suggests that NANOG is an ancient gene that is found in most vertebrate lineages (Scerbo et al., 2014); furthermore, its reprogramming capacity is conserved within Nanog's unique homeodomain (Silva et al., 2009). It seems likely that as pluripotency itself is also conserved in the trunk of vertebrates that the study of non-mammalian vertebrates has the potential to discern whether the regulatory networks which govern

pluripotency share conserved mechanisms. Specifically, this could elucidate the mechanism by which *Nanog* programs the pluripotent domain. As non-mammalian model organisms develop externally and are in greater abundance, this would also be of benefit.

To date, a conserved role for *NANOG* in the development of non-mammalian vertebrates has remained elusive. In teleosts such as Zebrafish, for example, two *Nanog* genes exist but do not govern pluripotency (Camp et al., 2009, Lee et al., 2013, Xu et al., 2012). Maternal *Nanog* primarily governs the formation of the yolk syncytial layer (YSL), the source of Nodal signalling for endoderm specification (Xu et al., 2012). Since the YSL itself is a structure unique to teleosts, Xu et al. have suggested that *Nanog* was recruited into a novel regulatory circuitry within the teleost lineage (Xu et al., 2012). Others later suggested that zebrafish *Nanog* may also play a pivotal role in zygotic genome activation (ZGA), which again appears to be a feature that is not conserved in mammals (Camp et al., 2009, Lee et al., 2013, Xu et al., 2012, Mitsui et al., 2003, Chambers et al., 2003).

Curiously, *NANOG* is encoded in the *Ambystoma mexicanum* (axolotl) genome but not within the *Xenopus* genome (Dixon et al., 2010, Hellsten et al., 2010, Scerbo et al., 2012, Schuff et al., 2012). Thus, *NANOG* was deleted from the genome of anurans after their divergence from urodeles; indeed, this also corresponds with the emergence of germline (Johnson et al., 2001, Johnson et al., 2003b, Johnson et al., 2003a). It has, however, been postulated that *Vent1/2* fulfils the role of *NANOG* in anurans, and indeed this has gained widespread acceptance (Scerbo et al., 2014, Scerbo et al., 2012, Schuff et al., 2012, Briggs et al., 2018).

In order to clarify whether there is a conserved role for Nanog in vertebrate embryogenesis, I aimed to investigate the role of NANOG in the urodele amphibian, axolotl. Urodeles represent basal tetrapods which later gave rise to amniotes, urodeles like most vertebrates, retained the basic skeletal structure of the tetrapod ancestor (Callier et al., 2009, Niedzwiedzki et al., 2010). Furthermore, an ancestral urodele-like embryology appears to be conserved throughout the evolution of amniotes (Bachvarova et al., 2009a). This includes several fundamental embryological features, such as a surficial origin for mesoderm (Smith and Malacinski, 1983), a dorsally restricted blastopore that gives rise to the Notochord (Shook et al., 2002, Shook and Keller, 2008). Moreover, urodeles exhibit true pluripotency demonstrated by the animal cap's competence to form all three germ layers as well as germ cells in response to inductive signals (Swiers et al., 2010, Chatfield et al., 2014).

Here I show that in axolotl embryos NANOG and SMAD2/3 interact with the Trithorax complex to control the deposition of transcription activating marks throughout the chromatin of pluripotent cells in the AC. These modifications are required for germ layer commitment in response to early germ layer specification signalling. The lack of these marks arrests development prior to gastrulation and perturbs the sequential waves of gene activation that direct development. These results demonstrate that the mechanism by which Nanog programs pluripotency and thus embryogenesis is conserved between axolotl and human embryos and likely the trunk of vertebrates.

3.2 Nanog results

In order to understand which features of the mammalian pGRN are conserved from basal vertebrates, I sought to investigate the genome-wide RNA profiles of the pluripotent tissue of the axolotl. It has previously been demonstrated that the core pluripotency regulators *NANOG* and *POU5F1* are expressed in the AC from around stage 8.5 following zygotic genome activation to stage 12, which marks the completion of gastrulation (Dixon et al., 2010). Therefore, I performed whole RNA sequencing (RNA-seq) on explanted stage 10.5 (mid-gastrula) ACs (Figure 3.1). As expected, RNA-seq confirmed that the pluripotency genes: *NANOG* and *POU5F1* were highly expressed (Fig. 3.1b). I next integrated publicly available bulk and sc-RNASeq datasets to compare the gene expression in pluripotent tissues from mammals human (*Homo sapiens*) and pig (*Sus scrofa*) as well as a frog (*Xenopus tropicalis*) at an equivalent developmental stage (Fig. 3.1a-c) (Li et al., 2017, Angerilli et al., 2018, Zhu et al., 2021, Ramos-Ibeas et al., 2019).

Peri-gastrulation staged pig and human epiblast cells, frog and axolotl ACs showed similar expression profiles of core pluripotency factors *SOX2* and *LIN28A*. However, only axolotl, pig and human pluripotent cells showed similar expression profiles for the pluripotency genes *NANOG*, *POU5F1* and *PRDM14*. Some genes, such as *TFCP2L1* and *OTX2*, showed similar expression between human-axolotl and pig-axolotl, respectively. Interestingly, few genes were expressed in mammalian and *Xenopus* cell populations but not axolotl. Given that *NANOG* and *POU5F1/OCT4* are not conserved in the *Xenopus*

genome (Hellsten et al., 2010), differences observed here may reflect the divergence of the pGRN in frogs from a conserved vertebrate state.

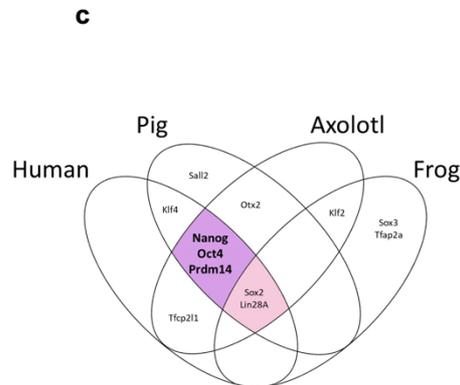
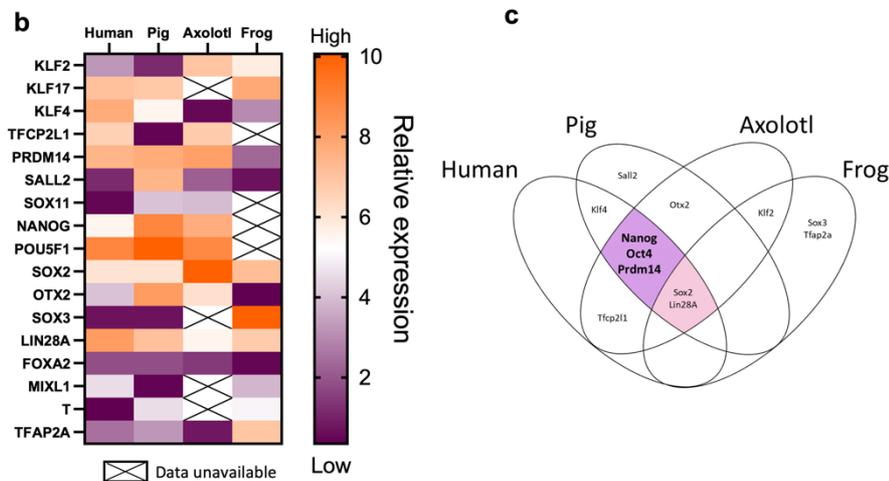
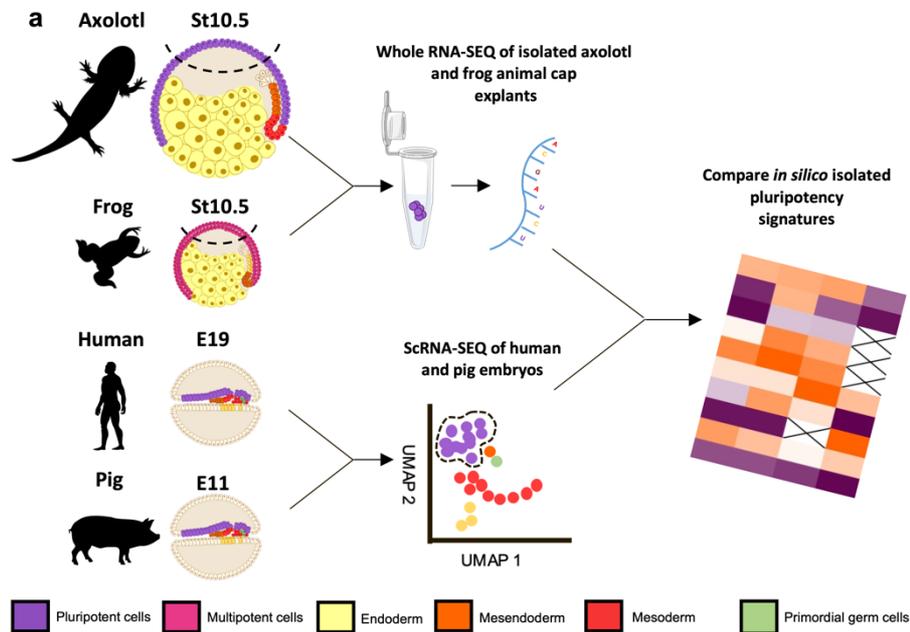


Figure 3.1. Cross-species comparison of pluripotent tissues.

a, Outline of transcriptomic data collection and cross-species comparison. **b**, Key pluripotency gene expression in the peri-gastrula primitive ectoderm of axolotl ($n=15$, $m=3$, $t=45$), frogs and mammals. Relative expression indicates the \log_2 transformed TPM after quantile normalisation (see methods). **c**, Venn diagram showing overlapping highly expressed (>1.5 times the median expression across all genes) pluripotency genes between the four organisms.

Mouse pluripotent cells are classified as either naïve or primed, defined by their developmental potential and their dependence on *nodal* signalling (Nichols and Smith, 2009). In axolotl, both *NODAL* and *NANOG* are activated between stages 8 and 9 in the AC following zygotic genome activation (ZGA)(Swiers et al., 2010, Dixon et al., 2010) (Fig. 3.2). Moreover, there is no stage at which *NANOG* is expressed in the absence of *NODAL*, suggesting that their co-expression represents basal pluripotency (Fig. 3.2).

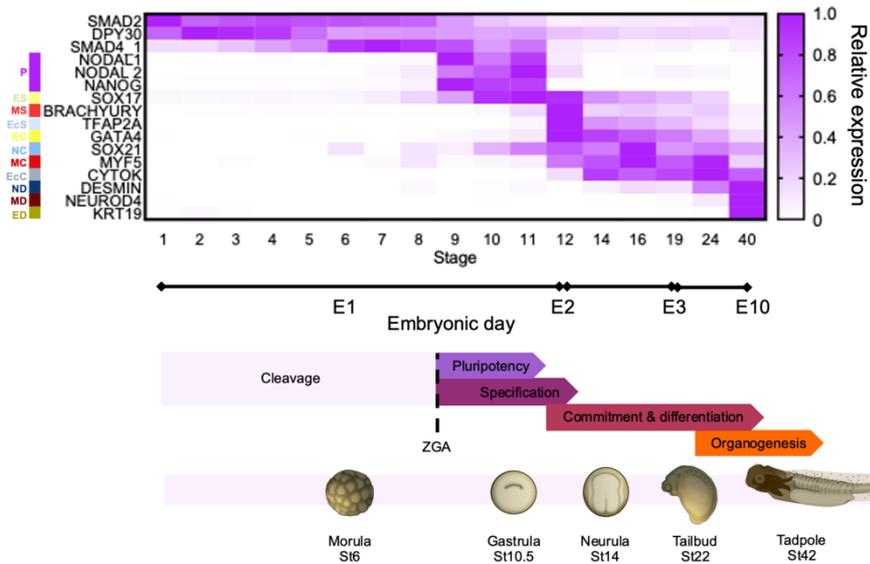


Figure 3.2. Overview of axolotl early development.

Heatmap showing gene expression data for representative marker genes across 17 developmental stages. Raw data was taken from whole embryo transcriptomes produced by Jiang, et al (2017). Gene expression was normalised to its own peak expression. Left hand side colour key indicates what each marker is associated with: Pluripotency (P), endoderm specification (general) (ES), mesoderm specification (general) (MS), ectoderm specification (general) (EcS), mesoderm commitment (MC), endoderm commitment (EC), ectoderm commitment (EcC), mesoderm differentiation (MD), endoderm differentiation (ED), Neural differentiation (ND). Also shown is a schematic indicating the embryonic day in which each developmental stage occurs (when embryos develop at 20°C), major milestones in cell differentiation and critical developmental time points: morula, gastrula, neurula tailbud and tadpole. Images adapted from the Axolotl Newsletter, Spring 1979 (Bordzilovskaya and Dettlaff, 1979).

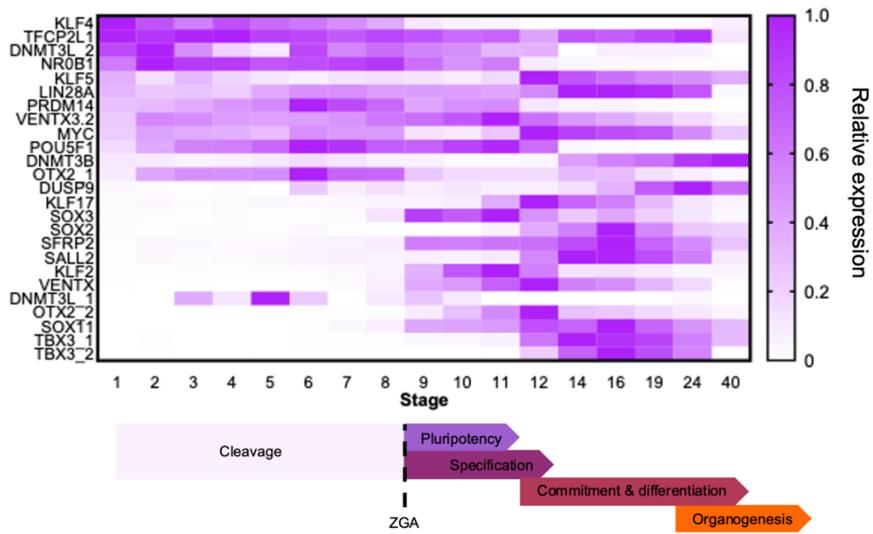


Figure 3.3. Expression of axolotl genes associated with pluripotency in mammals. Heatmap showing gene expression data for axolotl orthologues of mammalian pluripotency genes across 17 developmental stages. Raw data was taken from whole embryo transcriptomes produced by from Jiang, et al (2017). Gene expression was normalised to each gene’s peak expression.

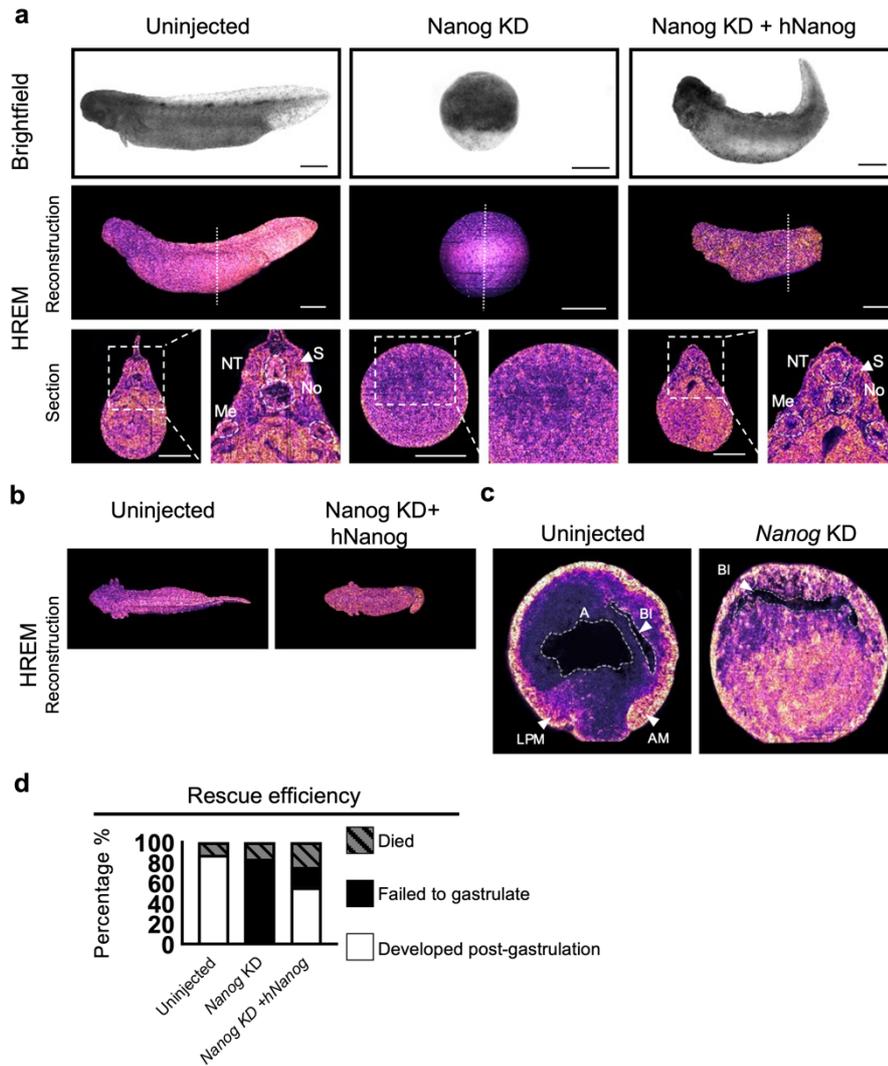


Figure 3.4. NANOG KD in axolotl embryos

a, NANOG KD arrests development prior to gastrulation and can be rescued by co-injection of hNanog RNA. Brightfield images and HREM reconstructions of uninjected, NANOG depleted and hNanog rescued embryos at equivalent stage 42 (Ventral view). HREM (n=2, m=1, t=6) Dotted line marks plane of section reconstruction (transverse). Dashed lines highlight: somites (S), neural tube (NT) notochord (No), meso/pronephric ducts (Me). Scale bar, 1mm. **b**, Dorsal view of HREM reconstructions of st42 uninjected and NANOG KD + hNanog rescued embryos seen in **a**. **c**, HREM reconstructed sagittal sections (n=2, m=1, t=6) of mid-gastrula embryos with and without NANOG KD at stage 10.5. Visible structures highlighted: Involuting axial mesoderm (AM), ingressing ventral mesoderm (VM), archenteron (A), blastocoel (B). Scale bar, 1mm. **d**, NANOG KD and rescue efficiencies (n=25, m=1, t=75).

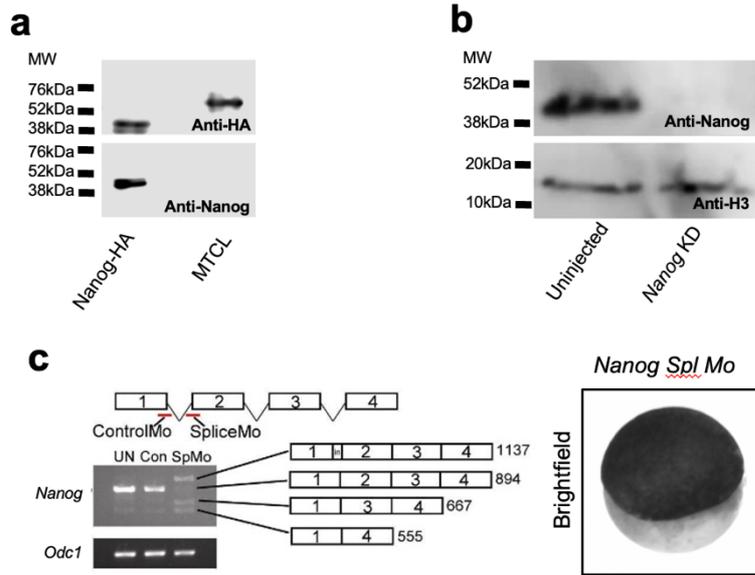


Figure 3.5. Confirmation of NANOG knockdown

a, Validation of antibodies raised against axolotl NANOG using western blotting. Lanes were loaded with lysates made from mature *Xenopus* oocytes following injection with Synthetic poly-A RNA encoding NANOG-HA as well as a multi-antigen cell lysate. **b**, Western blot confirming complete KD of NANOG following MO injection at stage 10.5. **c**, Validation of aberrant splicing in response to *Nanog* splice-morpholino and brightfield image of a stage 22 embryo following an injection of 80ng *Nanog* splice MO at the 1 cell stage. (n=15, m=1, t=30).

Given that broadly, the expression of several factors associated with mammalian pluripotency were also expressed in equivalent stage axolotl ACs, I decided to investigate the *in vivo* role of NANOG in further detail. Due to their extremely high efficiency in axolotl, I utilized a morpholino (MO) KD approach. Antisense morpholinos were designed to sterically block the formation of the translation initiation complex resulting in NANOG depletion. KD of NANOG resulted in complete developmental arrest at equivalent stage 9, prior to the onset of gastrulation (Fig. 3.4).

HREM revealed that NANOG morphants show no evidence of any involution or ingression characteristic of the early stages of gastrulation (Fig. 3.4c). To test the efficiency of KD, antibodies were raised against antigenic peptides specific to NANOG. Synthetic poly-A RNA encoding NANOG-HA, as well as

an uninjected *Xenopus* oocyte lysate and a multi-antigen containing cell lysate, was also used as HA-positive control.

Western blotting confirmed the specificity of the NANOG antibody (Fig 3.5a). Subsequent Western blotting confirmed complete KD of the NANOG protein after injection with 80ng NANOG morpholino (Fig. 3.5b). To verify the specificity of this effect, Splice MOs, which span the intron/exon boundaries of the gene, were designed to disrupt RNA splicing. This resulted in an identical phenotype (Fig. 3.5c). Disrupted splicing was detected via PCR (Fig. 3.5c). Superficially the effect of NANOG depletion is similar to that observed in Zebrafish, where *Nanog* depletion also results in developmental arrest before gastrulation (Xu et al., 2012, Gagnon et al., 2018). Unlike Zebrafish, however, NANOG morphants maintained this morphology indefinitely and remained alive long after sibling embryos had reached the Tailbud stages of development. It has been suggested that in Frogs (*Xenopus*), *Ventx1/2* is functionally equivalent to NANOG and has allowed for the subsequent deletion of *Nanog* from the *Xenopus* genome (Scerbo et al., 2012). *Ventx1/2* KD embryos develop to Tailbud stages, exhibiting compression along the anterior-posterior axis (Scerbo et al., 2012).

Given this stark contrast between *Ventx1/2* KD in *Xenopus* and NANOG KD, it seems unlikely that the two genes are functionally equivalent. The developmental progression of NANOG morphants can be rescued by injection of human NANOG (hNANOG) mRNA (Fig 3.4), suggesting a conserved activity between Urodele and mammalian NANOG orthologues. In mammals, *Nanog* forms part of the OCT4/SOX2/NANOG triumvirate of transcription factors that constitutes the core network controlling pluripotency (Boyer et al., 2005, Loh et al., 2006, Chambers and Tomlinson, 2009) and is critical for the

maintenance of the epiblast *in vivo* (Mitsui et al., 2003). Moreover, it has been proposed that in mammalian ESC Nanog regulates it's own expression and that of other key pluripotency factors: Oct4 and Sox2 (Loh et al., 2006). Given that previous work carried out by our group has suggested conservation of the gene regulatory network (GRN) between mammals and Urodeles (Dixon et al., 2010),

I next asked whether NANOG depletion affected *NANOG* expression and the expression of *POU5F1* and *SOX2* using quantitative PCR (qPCR) (Fig. 3.6). While NANOG depletion decreased the expression of *POU5F1* and *SOX2* at equivalent stage 11, *NANOG* expression increased. Counter-intuitively, unlike uninjected sibling embryo's, NANOG morphants failed to extinguish pluripotency factor gene expression after gastrula stages and expression continued long after siblings reached tadpole stages. This result suggests that NANOG protein may promote pluripotency gene expression before gastrulation. Another possibility is that the decreased pluripotency factor expression observed at stage 11 is due to morphants being incorrectly staged due to their unchanging morphology. Whichever the case, the data also suggests that NANOG is required to silence pluripotency gene expression post-gastrulation, including it's own expression. It is also worth noting that while morpholinos remain stable for around 5 days, we cannot rule out the reappearance of the NANOG protein later. However, it is clear that any reappearance is insufficient to restart a normal developmental trajectory.

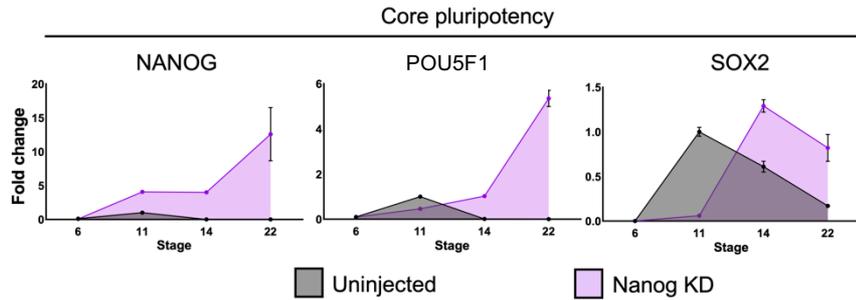


Figure 3.6. NANOG is required to extinguish pluripotency.

Line graphs of comparative expression of core pluripotency genes NANOG, *POU5F1* and *SOX2* across 4 developmental stages with and without NANOG KD, assayed using QPCR (n=10, m=3, t=60).

To further characterize the NANOG KD phenotype at different developmental stages, I measured the global gene expression of gastrula and late Neurula equivalent stage embryos (Fig. 3.7). The transcriptome data supported our previous observation that *POU5F1* and *SOX2* are downregulated at gastrula stages in morphants embryos (Fig. 3.7a). Further to this, I observed several transcription factors, associated with pluripotency in mammals, were down-regulated in NANOG morphants at equivalent stage 10.5, notably *POU5F1*, *SOX2*, *PRDM14*, *KLF2*, *ZIC3* and *REX1*. This suggests that several factors involved in the wider pluripotency GRN in mammals may also be regulated by NANOG, and this feature has been conserved from a Urodele-like ancestor. Transcriptome data also validated our observation that NANOG morphants fail to transcriptionally silence pluripotency gene expression after gastrula stages (Fig. 3.7.a).

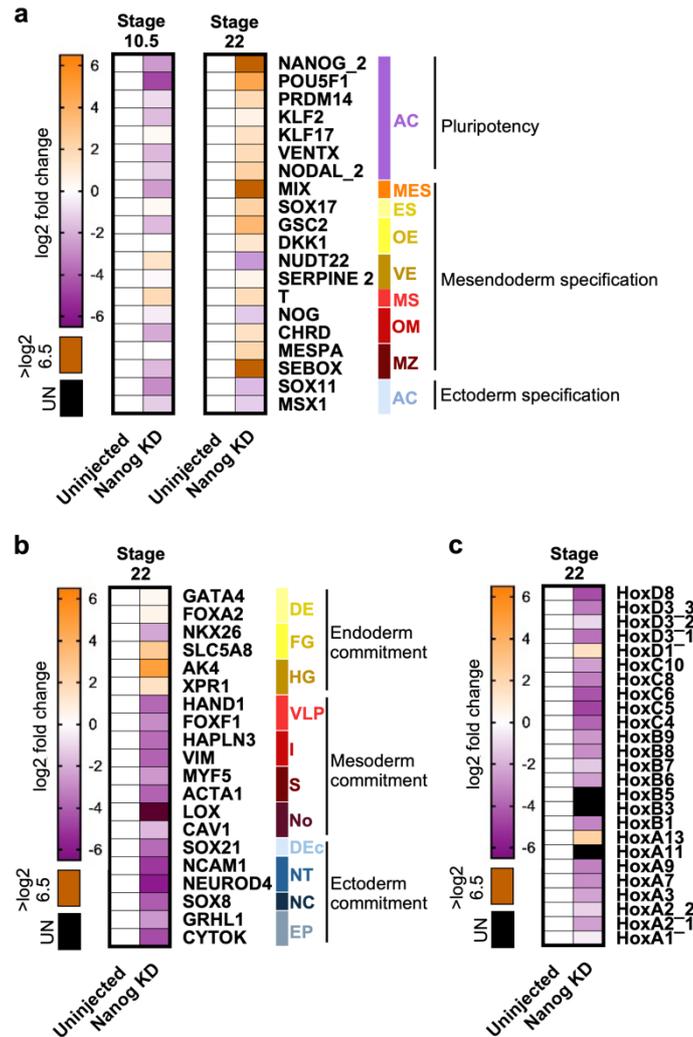


Figure 3.7. Differential gene expression following NANOG KD.

a, Heatmap showing differential gene expression of early marker genes at stages 10.5 and 22 in uninjected ($n=7$ and 6 , respectively, $m=1$, $t=13$) and NANOG depleted ($n=6$ and 3 , respectively, $m=1$, $t=9$) embryos. Black indicates no detectable expression. Cell types: animal cap (AC), mesendoderm specification (general) (MES), endoderm specification (general) (ES), organiser endoderm (OE), vegetal endoderm (VE), mesoderm specification (general) (MS), organiser mesoderm (OM), marginal zone (MZ). **b**, Differentially expressed germ-layer commitment marker genes at stage 22 in uninjected and NANOG KD embryos ($n=6$ and 3 , respectively). Black indicates no detectable expression. Cell types: definitive endoderm (general) (DE), foregut (FG), hindgut (HG), ventral-lateral plate (VLP), intermediate mesoderm (I), somite (S), notochord (No), definitive ectoderm (general) (Dec), neural tube (NT), neural crest (NC), epidermal progenitors (EP). **c**, Differential gene expression of Hox gene family members in response to NANOG KD at stage 22.

Transcriptomic data revealed that NANOG KD also affected germ layer specification genes. Indeed, genes involved in the specification of all three

germ layers show differential gene expression upon NANOG KD. Notably, *NODAL*, *MIX*, *GOOSECOID*, *CHORDIN*, *NOGGIN* and *SOX2* were downregulated while factors such as *SOX17* and *RUNX1* were upregulated at stage 10.5. Interestingly, many of these specification genes, like pluripotency genes, failed to be transcriptionally silenced and are highly upregulated at stage 22 (Fig 3.7a). This suggests NANOG is required for the global organization of lineage specification but is not required for global zygotic genome activation (ZGA) as proposed in Zebrafish (Lee et al., 2013, Leichsenring et al., 2013).

Ectodermal commitment markers were also downregulated, including *SOX21*, *NCAM1* and *CYTOK*, which delimit general definitive ectoderm, neural tube and epithelial tissues, respectively. By contrast, endodermal gene commitment markers were upregulated, including *GATA4*, *FOXA2* and *AK4*, suggesting that endodermal gene commitment is less affected by the absence of *Nanog*. This is reminiscent of *Nanog* deletion in hESC, which results in an endoderm-like cell identity that also expresses *GATA6*, *GATA4*, *HNF1B* and *IGF2* (Mitsui et al., 2003). Strikingly, transcriptomic analysis also revealed significant down-regulation of several members of the *HOX* genes indicative of the failure of *NANOG* morphants to establish a body axis (Fig 3.7c).

In order to elucidate the effects of *NANOG* depletion on cell-fate decisions, I employed a gene-set enrichment analysis (GSEA) on *NANOG* KD differentially expressed genes (DEGs) using a list of amphibian cell-type markers generated from staged single cell-transcriptome in sister taxa *Xenopus tropicalis* (Briggs et al., 2018) (Fig 3.8). Given that many of the markers identified in this paper have not been extensively validated in the axolotl, this analysis cannot be used in isolation to infer which cell types are affected by *NANOG* depletion. This approach, however, introduces less bias than manually looking at cell-type markers in a large dataset. GSEA showed enrichment of lateral plate, Notochord, cardiac mesoderm, neural tube, neural crest and epidermal progenitor markers among down-regulated genes at equivalent stage 22 in *NANOG* morphants (Fig 3.8b). Upregulated genes at the same stage showed enrichment of blastula and germ cell markers (Fig 3.8b). QPCR experiments on a subset of genes validated the accuracy of the transcriptome data (Fig. 3.9a-c).

As NANOG morphants displayed markers indicating the endodermal commitment, I next looked for late-stage differentiation markers to see if there is any evidence of more complex tissue formation (Fig. 3.9d). Intriguingly, NANOG morphants did not express any of the differentiation markers, including those of the endodermal lineage, suggesting that in the absence of Nanog, embryos can form an endoderm-like tissue but not more complex endodermal tissue types, an observation consistent with the morphants simple morphology. Together these findings suggest that NANOG is required for the organization of lineage specification to engage the forward momentum of the embryo. When taken in conjunction, these findings suggest in the absence of NANOG, embryos fail to silence early pluripotency/specification gene expression resulting in the ablation of mesodermal and ectodermal germ layer commitment, the embryos then default to a transcriptionally active, endoderm-like state (Fig 3.10).

Importantly, these observations are in stark contrast to both *Ventx1/2* KD in *Xenopus* and MZ *Nanog* KO in Zebrafish (Lee et al., 2013, Leichsenring et al., 2013, Xu et al., 2012, Gagnon et al., 2018, Scerbo et al., 2014, Scerbo et al., 2012, Schuff et al., 2012, Briggs et al., 2018).

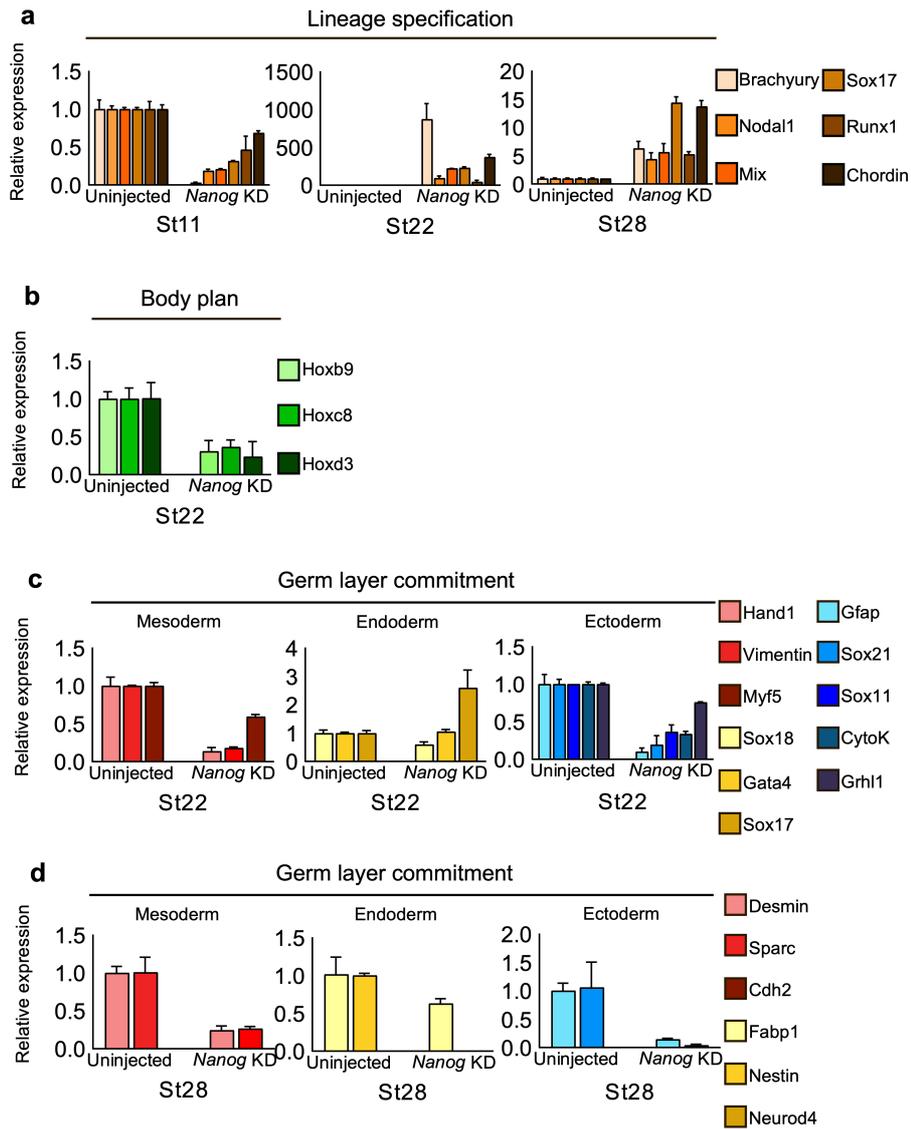


Figure 3.9. QPCR validation of key germ layer markers in NANOG KD embryos.

a, Expression of germ-layer specification markers in stage 10.5 embryos with and without NANOG depletion. (n=10, m=1, t=20). **b**, Hox-gene expression. (n=10, m=1, t=20) **c**, Germ-layer commitment gene expression at stage 22. (n=10, m=1, t=20) **d**, Late stage (stage 28) differentiation markers of uninjected and NANOG depleted embryos. (n=10, m=3, t=60).

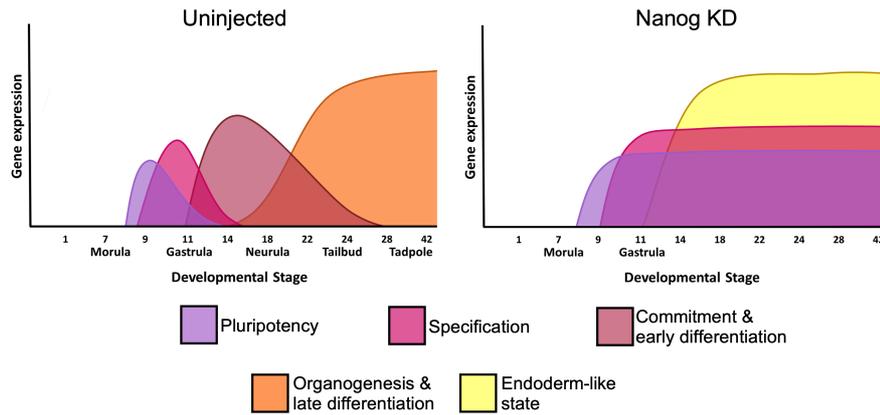


Figure 3.10. Schematic: NANOG KD prevents waves of gene expression.

Schematic diagram showing that NANOG depletion prevents the sequential waves of gene expression that direct embryogenesis. NANOG depleted embryos continue to express high levels of pluripotency and germ-layer specification associated genes and express a subset of endodermal commitment markers.

To test the developmental potential of *NANOG* depleted ACs directly, I explanted uninjected and depleted caps and investigated differentiation marker expression using QPCR (Fig. 3.11). ACs normally differentiate into epidermis, marked by upregulation of *Grhl1* and *Cytok*, and down-regulation of pluripotency markers. Morphant caps, however, showed reduced levels of epidermal markers and maintained high *NANOG* and *POU5F1* levels (Fig. 3.11b). Given that NANOG depleted embryos failed to produce mesoderm, I next tested whether mesoderm could be induced directly from morphant caps using ACTIVIN (Fig. 3.12). In ACs, 1pg of RNA encoding activin induces mesoderm, marked by elongation and *BRACHYURY* expression (Swiers et al., 2010). In contrast, 200fg of RNA is insufficient to induce mesoderm, while 1pg induces elongation and expression of *BRACHYURY*, suggesting the formation of mesoderm. It is worth noting, however, that at this level of ACTIVIN, you also get high expression of *SOX17*, indicating that perhaps a subset of cells are also forming endoderm (Fig. 3.12b). Remarkably, when combined with

NANOG depletion, caps expressing sub-mesodermal (200fg) doses of ACTIVIN expressed elevated levels of *NODAL1*, *MIX* and *SOX17* but not *BRACHYURY* (Fig. 3.12b). 1pg of ACTIVIN in morphants also induced extremely high levels of these markers compared to activin-only controls, also without induction of brachyury. Moreover, NANOG KD caps failed to elongate instead became spherical, usually indicative of endoderm (Fig. 3.12b) (Swiers et al., 2010). I next looked for definitive endodermal marker expression in the 1pg activin-injected morphant caps (Fig. 3.12c). Morphant caps showed high expression of *C8B*, a definitive endoderm marker but not foregut progenitor marker *NKX2-5* or hindgut progenitor marker *HNF4G*. Therefore, NANOG depleted caps are able to form definitive endoderm but not more specialized cell types in line with the whole embryo data. Moreover, these data suggest that ACs depleted of NANOG are hyper-sensitized to TGF- β signalling. I posit, therefore, that NANOG may act as a rheostat of SMAD2/3, as it does with SMAD1/5 in mESC (Mullin et al., 2008), which in ACs, modulates the response to nodal signalling levels (Fig. 3.13). This aligns with the inability of *Nanog* null mouse embryos to complete epiblast specification (Mitsui et al., 2003, Silva et al., 2009).

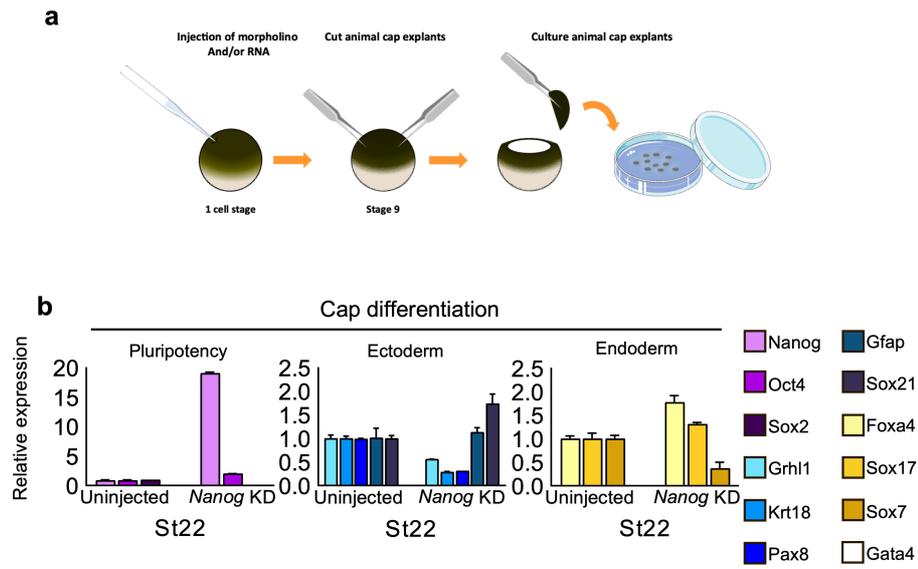


Figure 3.11. Germ-layer marker expression in NANOG KD ACs.

a, Schematic of morpholino injection regime and animal cap assay. **b**, Germ-layer differentiation markers of uninjected and NANOG depleted AC explants at stage 22. (n=15, m=1, t=30)

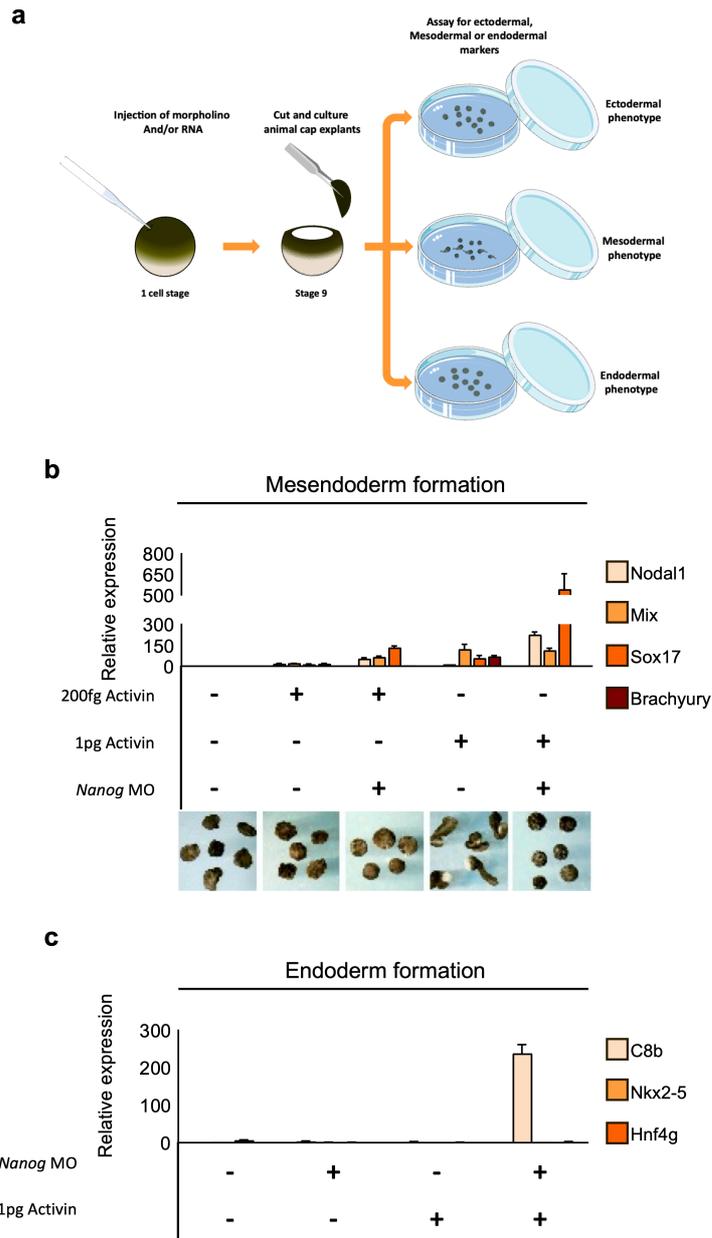


Figure 3.12. Mesendodermal induction in NANOG KD ACs.

a, Schematic of mesoderm induction animal cap assay. **b**, QPCR of mesodermal and endodermal markers in uninjected and NANOG depleted stage 22 caps following injection with different activin concentrations at the 1-cell stage (n=15, m=1, t=75). **c**, QPCR showing NANOG depleted caps express foregut/hindgut markers but not mature foregut/hindgut markers in response to activin. (n=15 (n=10, m=1, t=20), m=1, t=75).

It has previously been demonstrated that *NANOG* is expressed between stages 9 and 12 exclusively in the animal cap of the embryo (Dixon et al., 2010). In amphibians, the foregut derives from the dorsal marginal zone, while the

hindgut is produced from the vegetal hemisphere (Briggs et al., 2018). Previous data within our lab has established that vegetal hemisphere explants at late stages express *ENDODERMIN*, a marker of definitive endoderm (also shown in Fig. 3.30b&c). Given that NANOG morphants appeared to be able to form definitive endoderm in the absence of gastrulation, I tested whether the endodermal gene expression seen in morphant transcriptomes was due to the vegetal hemisphere forming endodermal tissue in a cell-autonomous manner (Fig. 3.14). Thus, I prepared and cultured animal and vegetal explants. It has previously been established that 1pg of RNA encoding activin is sufficient to induce animal cap elongation and *BRACHYURY* expression, both indicators for the induction of mesoderm (Green et al., 1992, Swiers et al., 2010). Using titration, I established that 100pg of activin RNA was sufficient to induce the formation of endoderm as determined by upregulated expression of *SOX17* and *MIX* and downregulated *BRACHYURY* expression. Vegetal pole explants exhibited higher expression of *SOX17* and *MIX* than the endoderm induced animal caps and, by stage 30, expressed high levels of the definitive endoderm marker: *ENDODERMIN*. The data from both time points suggests that the vegetal pole explants are able to form definitive endoderm in a cell-autonomous manner. Thus, it seems reasonable to conclude that vegetally derived definitive endoderm is NANOG independent and may be contributing to endodermal marker expression in the NANOG depleted whole embryos.

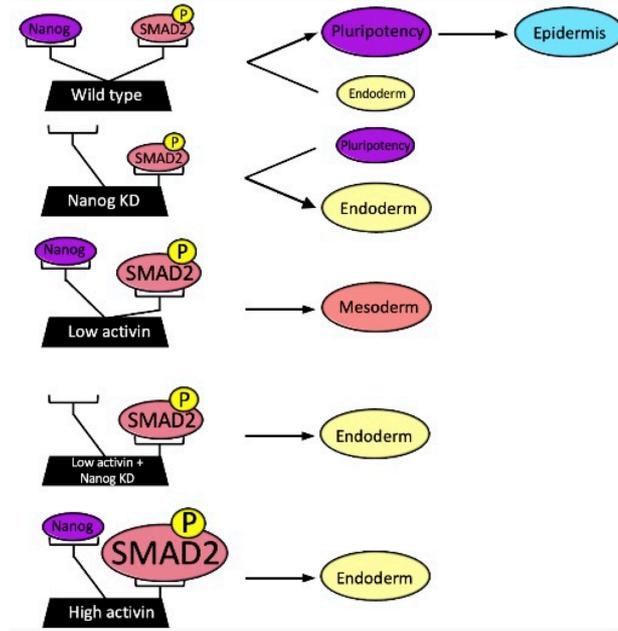


Figure 3.13. Schematic: NANOG acts as a rheostat of SMAD2 activity.

In wild type ACs, a balance of SMAD2 and NANOG levels must be maintained to promote pluripotency. Low amounts of activin promote the formation of mesoderm. High activin or knockdown of NANOG results in a shift toward endoderm suggesting that without Nanog ACs are hypersensitive to SMAD2 suggesting that NANOG may act as a rheostat of SMAD2.

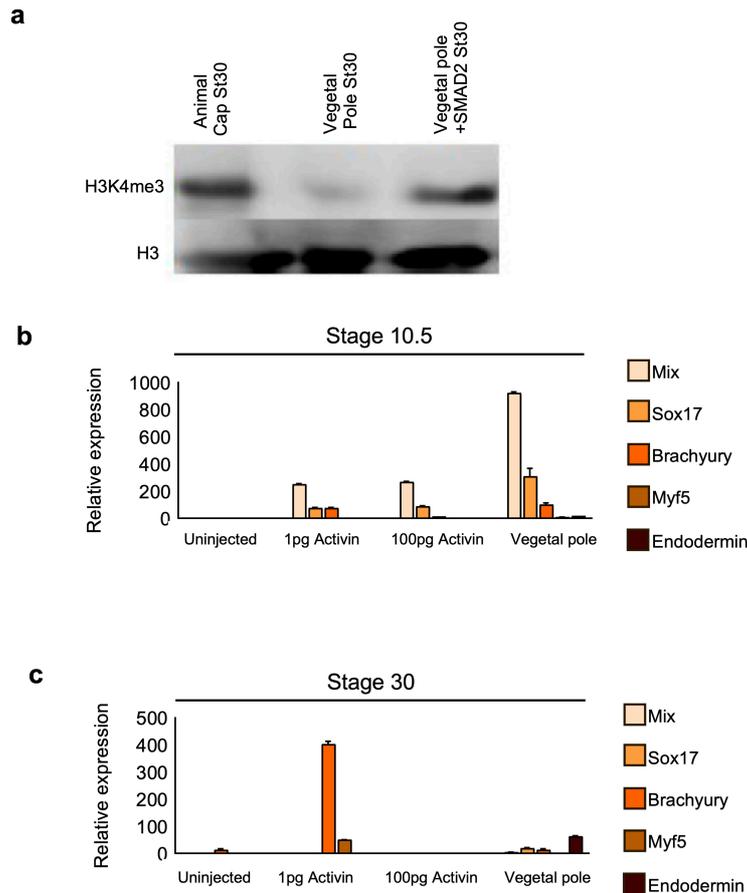


Figure 3.14. Epigenetic and transcriptional development in embryo explants.

a, Western blots showing the enrichment of H3K4me3 in explanted and cultured wild type animal and vegetal poles, as well as vegetal poles following injection of phospho-SMAD2 at the 1-cell stage. **b&c**, QPCR showing the expression of mesodermal and endodermal markers in vegetal pole explants, compared with animal caps with varying concentrations of activin at stages 10.5 and 30 respectively (n=15, m=1, t=120).

Having demonstrated the effects of NANOG depletion, I next sought to identify if the mechanism by which NANOG regulates pluripotency is conserved in vertebrates. Data in hESCs has demonstrated that Nanog depletion causes a loss of the transcription promoting H3K4me3 at gene promoters (Bertero *et al.*, 2015). Western blotting performed by Helena Acosta (HA) determined that histone modifications H3K4me3, H3K27me3 and H3K27ac are not present prior to ZGA but are instead gradually acquired between stages 8.5 and 10 (Fig. 3.31a). These observations are consistent with

the notion that the majority of enhancer chromatin marks appear to arise during the major ZGA phase. Therefore, before ZGA, chromatin is free of both the activating enhancer marks (Akkers et al., 2009, Vastenhouw et al., 2010, Hug et al., 2017).

Immunofluorescent staining (Raw data also provided by Helena Acosta) sought to determine whether NANOG has a conserved role in the deposition of H3K4me3 (Fig. 3.15a). Immunostaining demonstrated a complete loss of H3K4me3 at equivalent stage 14 in NANOG morphants when compared to uninjected. Moreover, this effect was rescued when MO was co-injected with hNanog mRNA. This suggests that Nanog is required for the deposition of H3K4me3 in axolotl, a role conserved in hESCs (Bertero *et al.*, 2015).

It is worth noting that the widespread loss of H3K4me3 shown in the immunofluorescence staining is far more extensive than the H3K4me3 loss reported in hESCs, which accounted for only 12.3% of the total H3K4me3 peaks. While the reason for this is unclear, the more pronounced loss of H3K4me3 in axolotl embryos than hESCs may be because, as a prerequisite, pluripotency and H3K4me3 is already established in hESC cell lines at the time of KD. Concomitantly, direct and downstream NANOG target genes are already actively being transcribed and have acquired the H3K4me3 modification. It may be the case that *in vivo* NANOG has a more prominent role in the establishment of H3K4me3 than the maintenance of this mark. Unexpectedly, immunostaining also revealed the loss of the active enhancer marker H3K27ac in response to NANOG KD (Fig. 3.15a). As this had not been previously reported, it was unclear as to whether this is a direct or indirect consequence of NANOG KD. H3K27me3 was unaffected by NANOG depletion as was Phospho-POLII (Fig 3.32b), H3K36me3 and H3 (Fig. 3.15b)

stainings confirmed that the loss of these chromatin modifications is not due to transcriptional silencing or loss of histones, respectively. Western blotting corroborated the global loss of H3K4me3 and H3K27ac (Fig. 3.31c).

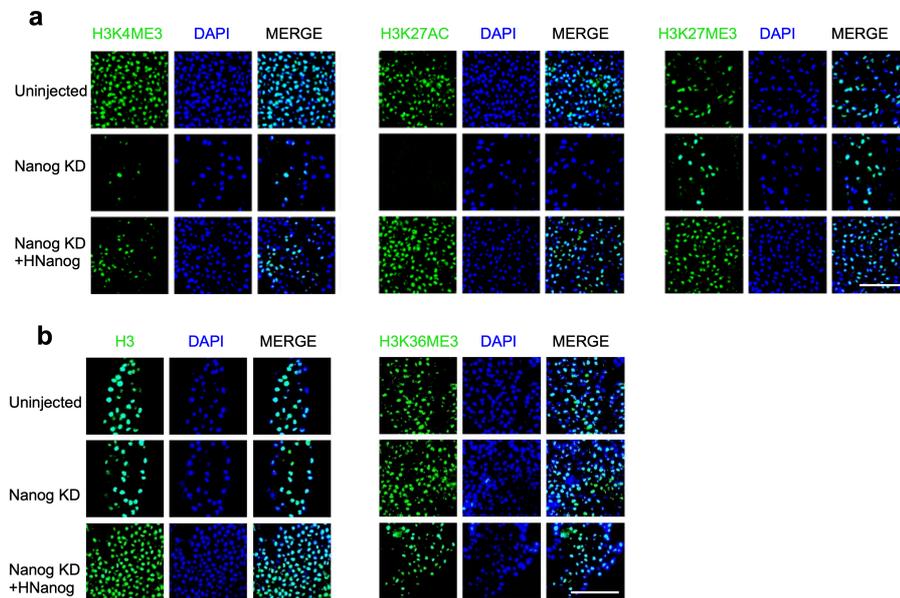


Figure 3.15. Immunofluorescent imaging of NANOG KD ACs.

Uninjected, NANOG depleted and hNanog rescued AC explants cultured to equivalent stage 14 and probed with fluorescent antibodies. **a**, Stained for H3K4me3, H3K27ac, H3K27me3 using and co-stained with DAPI. Scale bar, 60 μ m. **b**, Stained for H3, H3K36me3 and DAPI. Scale bar, 60 μ m

Interestingly, explanted animal and vegetal hemispheres from stage 9 embryos stained for H3K4me3 and H3K27me3 (Fig. 3.30a). Notably, the vegetal hemisphere showed reduced levels of epigenetic marks associated with both transcriptional activation and silencing. Having confirmed that the vegetal hemisphere explants form definitive endoderm in a cell-autonomous fashion, this may imply that endodermal commitment in the vegetal hemisphere requires less chromatin organization.

Nanog's epigenetic regulation in hESC depends on the formation of the NANOG-SMAD2/3 complex, which recruits the COMPASS methyltransferase complex (Bertero *et al.*, 2015), so I asked if this too functions

in pluripotent AC. Indeed, numerous members of the TGF- β signalling pathways were highly expressed prior to gastrulation (Fig. 3.16). It has been previously demonstrated that axolotl mesoderm induction is mediated by activin/nodal signalling, which acts upstream of SMAD2, and indeed this mesoderm GRN is conserved in mammals (Swiers et al., 2010). Moreover, the phenotype from nodal signalling inhibition using SB431542 (SB), an antagonist of the ALK4 and ALK7 type I activin/nodal receptors, resembles the axolotl *NANOG* morphant phenotype shown above, which arrests development prior to gastrulation. Brightfield images and HREM reconstructions confirmed the absence of morphological development following SB exposure (Fig 3.17).

The transcriptomic analysis confirmed previous observations that SB treatment results in downregulation of mesendodermal specification genes such as *NODAL1*, *MIX*, *SOX17* and *BRACHYURY* (Fig. 3.18a). RNA-seq also revealed that stage 22 SB treated embryos like *NANOG* morphants lacked mesodermal commitment markers, including *VIMENTIN*, *HAND1* and *MYF5* (Fig. 3.18b). Non-neural ectodermal commitment markers *GRHL1* and *KRT18* showed increased expression; however, in contrast, neural ectodermal markers such as *SOX21*, *NCAM1* and *SOX8* were downregulated. Like *NANOG* morphants, SB treatment resulted in downregulation of the HOX genes (Fig. 3.18c). Endodermal commitment marker *GATA4* was unaffected by SB treatment, as was a hindgut marker *AK4*.

As some endodermal markers and hindgut markers appeared unaffected by SB treatment, I investigated if the expression of mature endodermal markers in SB treated embryos was due to nodal independent expression from the vegetal pole. I treated vegetal pole explants with SB and analyzed endodermal

gene expression (Fig 3.21c&d). Indeed, *SOX17*, *GATA4* and *ENDODERMIN* expression was not greatly affected by SB, indicating that hindgut endoderm can develop independently of nodal signalling. Interestingly, the endoderm is the most ancient metazoan germ layer (Technau and Scholz, 2003, Hashimshony et al., 2015) and may have predated the evolution of Nodal signalling (Grande et al., 2014), suggesting that the vegetal pole endoderm GRN may reflect an ancient metazoan state. GSEA showed that among the significantly downregulated genes at stage 22 following SB treatment were markers of Notochord, Somite, Lateral plate, Dorsal lateral plate and intermediate mesodermal cell types as well as Chordal neural crest and anterior neural tube cell type markers (Fig. 3.19b). Significantly enriched in the upregulated genes at the same stage included markers of Ciliated epidermal progenitors Beta ionocytes, goblet cells and ionocytes. QPCR analysis confirmed the validity of the transcriptome data as well as demonstrating that no late differentiation markers of any of the three germ layers were present in equivalent stage 28 morphants (Fig. 3.20).

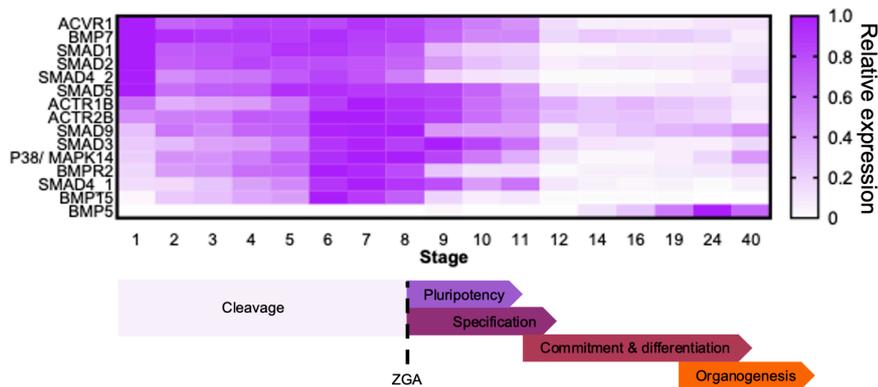


Figure 3.16. Expression of axolotl orthologues of TGF- β pathway related genes across developmental stages. The majority of TGF- β pathway related genes are maternally expressed and are highly expressed until mid-gastrula stages.

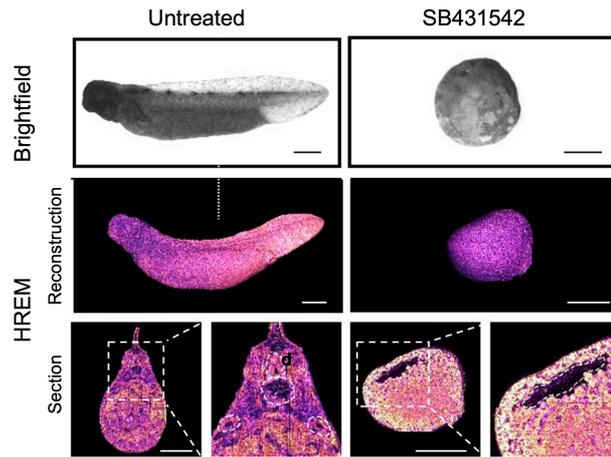


Figure 3.17. SB431542 treatment arrests development prior to gastrulation.

Brightfield and HREM images of uninjected and SB431542 treated embryos. (HREM $n=3$, $m=1$, $t=6$). Dotted line marks plane of section reconstruction (transverse). Dashed lines mark somites (S), neural tube (NT) Notochord (No), Meso/pronephric ducts (Me), Blastocoel (B). Scale bar, 1mm.

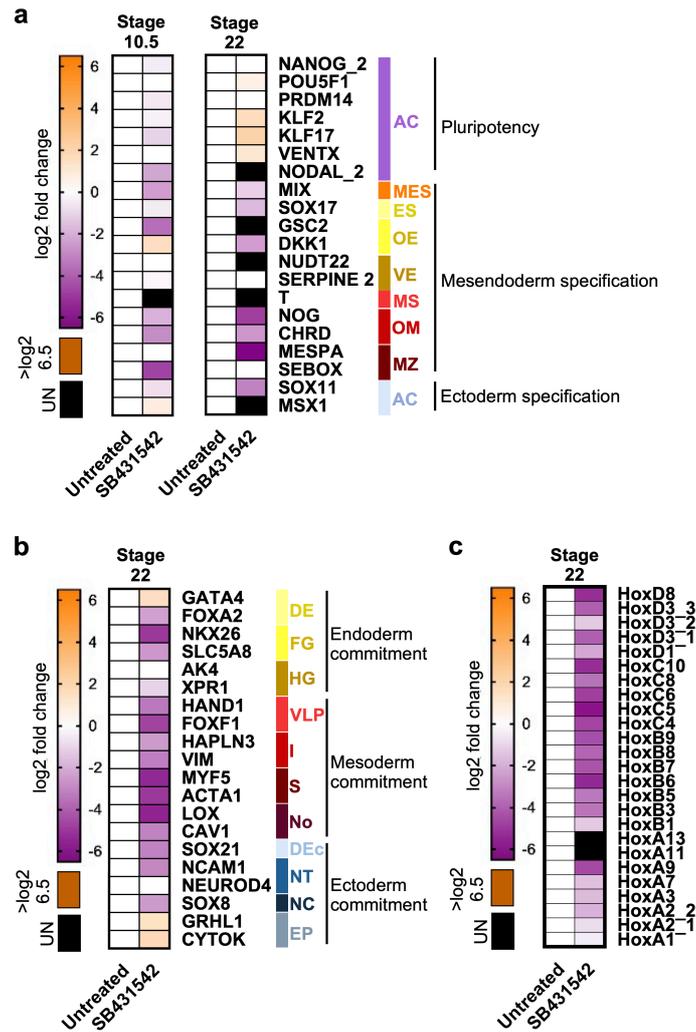


Figure 3.18. Differential gene expression following SB431542 treatment.

a, Heatmaps showing differential gene expression of early marker genes at stages 10.5 and 22 in uninjected (n=7 and 6, respectively, m=1, t=13) and SB431542 treated (n=6 and 6, respectively, m=1, t=12) embryos. Black indicates no detectable expression. Cell types: animal cap (AC), mesendoderm specification (general) (MES), endoderm specification (general) (ES), organiser endoderm (OE), vegetal endoderm (VE), mesoderm specification (general) (MS), organiser mesoderm (OM), marginal zone (MZ). **b**, Differentially expressed germ-layer commitment marker genes at stage 22 in uninjected and SB431542 treated embryos (n=6 and 3, respectively). Black indicates no detectable expression. Cell types: definitive endoderm (general) (DE), foregut (FG), hindgut (HG), ventral-lateral plate (VLP), intermediate mesoderm (I), somite (S), notochord (No), definitive ectoderm (general) (Dec), neural tube (NT), neural crest (NC), epidermal progenitors (EP). **c**, Differential gene expression of Hox gene family members in response SB431542 treatment at stage 22.

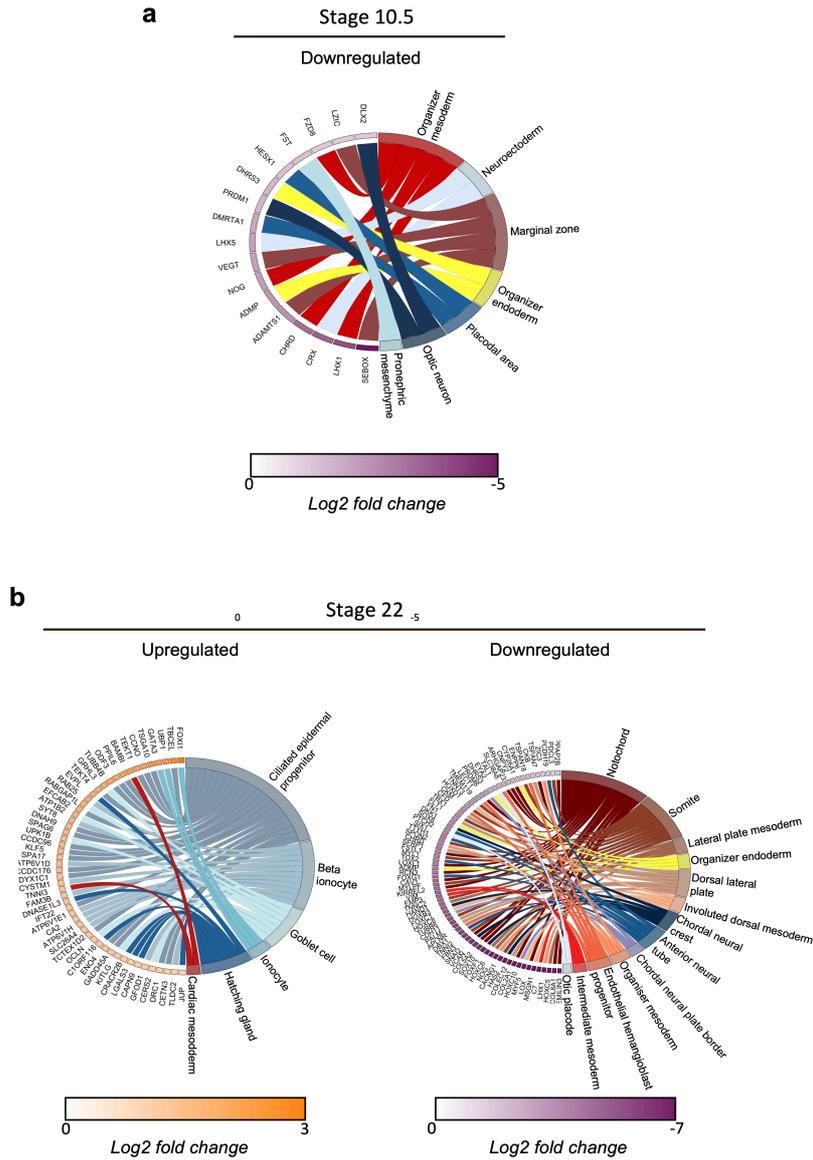


Figure 3.19. Amphibian cell type marker expression following SB431542 treatment. Chord diagrams showing statistically significant ($FDR < 0.01$) amphibian cell-type markers within SB431542 treated differentially expressed genes identified from gene set enrichment analysis (GSEA). **a**, Significant marker genes enriched in the up and down-regulated genes following SB431542 treatment at stage 10.5. **b**, Significant marker genes enriched in the up and down-regulated genes following SB431542 treatment at stage 22.

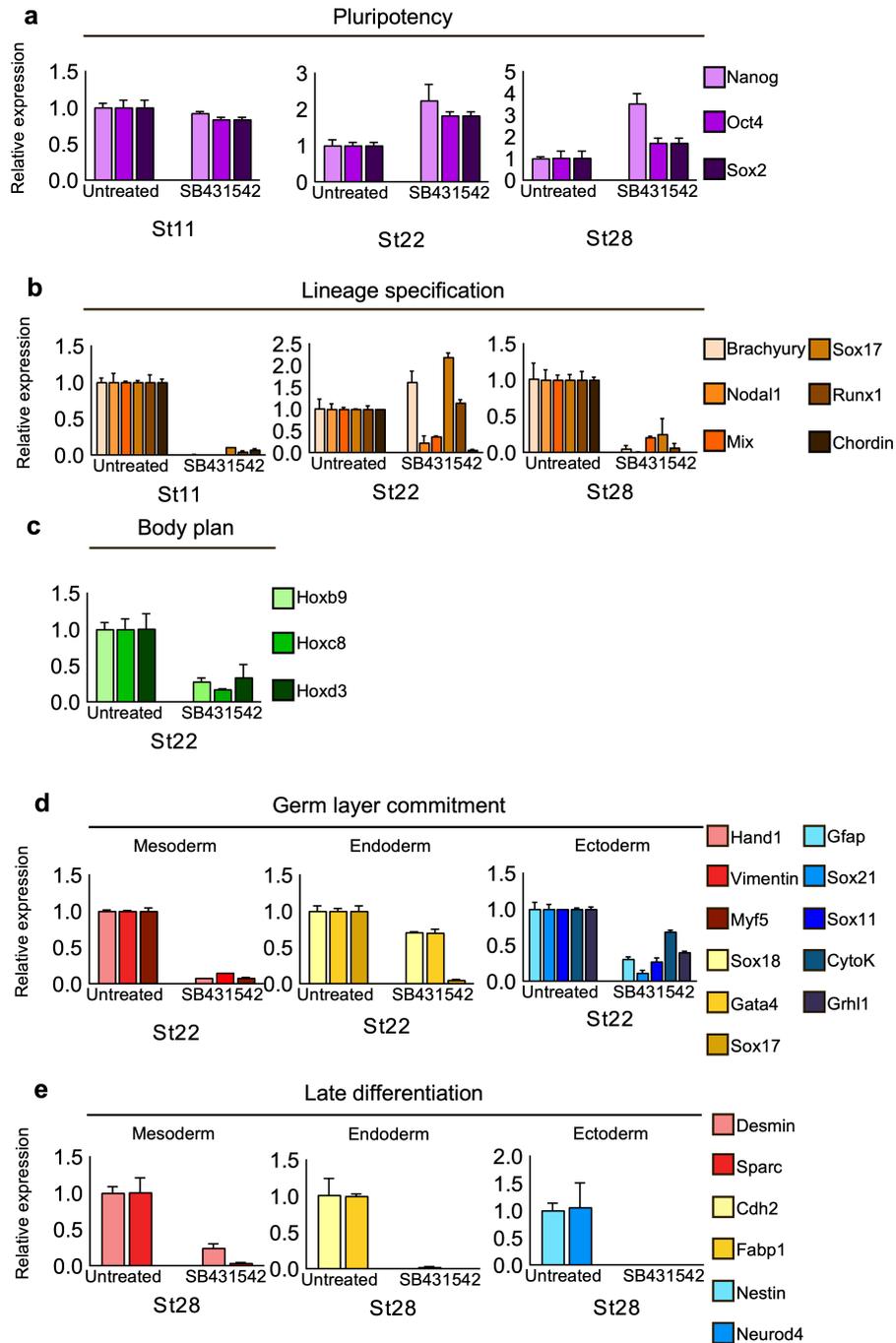


Figure 3.20. QPCR validation of key germ layer markers in SB431542 treated embryos.

a, Expression of core pluripotency factors in stage 10.5 embryos with and without SB431542 treatment (n=10, m=1, t=20). **b**, Germ-layer specification markers (n=10, m=1, t=20). **c**, Hox-gene expression. (n=10) **d**, Germ-layer commitment gene expression at stage 22 (n=10, m=1, t=20). **e**, Late stage (stage 28) differentiation markers of uninjected and SB431542 treated embryos. (n=10, m=3, t=60)

I next investigated the effect of NODAL signalling inhibition on the differentiation of animal cap explants (Fig. 3.21a&b). QPCR analysis

demonstrated that in the absence of NODAL signalling, animal caps show reduced epidermal markers indicating that in the absence of SMAD2 activity, the AC does not efficiently form any particular cell state.

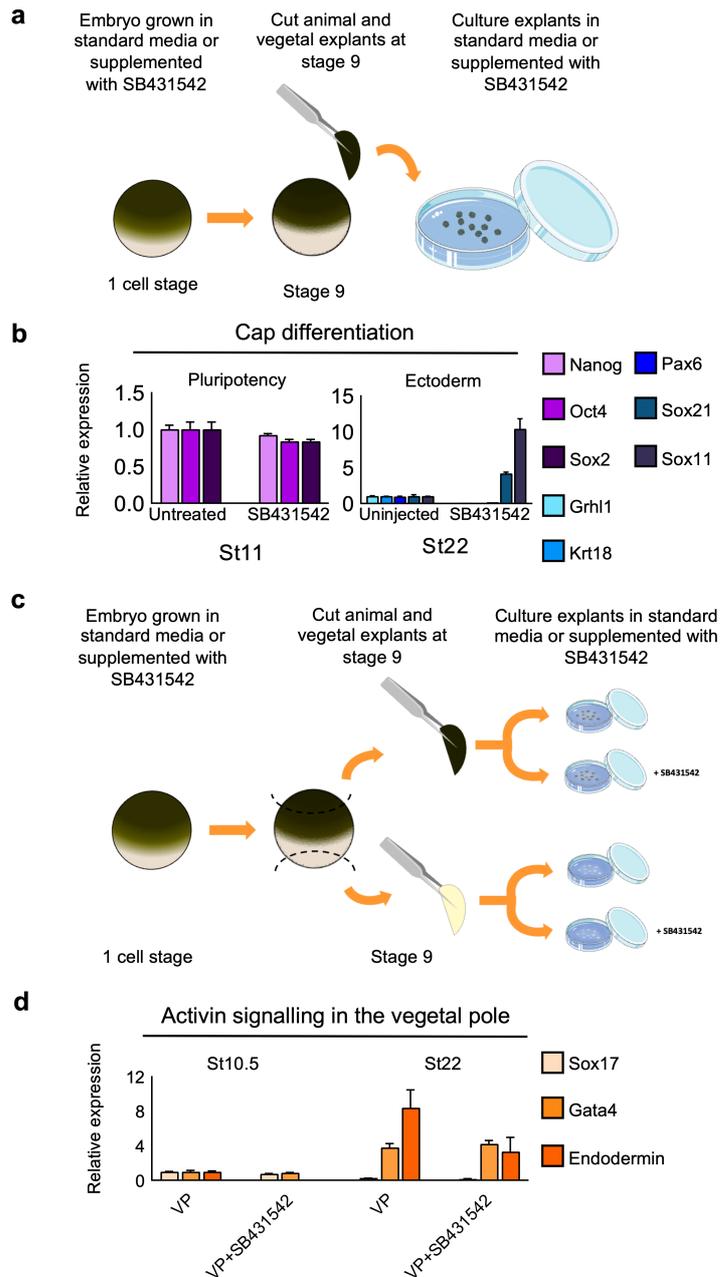


Figure 3.21. Germ-layer marker expression in SB431542 treated embryo explants.

a, Schematic of SB431542 treatment regime and explant culture. **b**, QPCR of pluripotency and ectodermal markers in uninjected and SB431542 treated stage 22 caps. **c**, Diagram of vegetal explant procedure and SB431542 treatment (n=15, m=1, t=30). **d**, QPCR of Vegetal explants from untreated or SB431542 treated embryos and assayed for endodermal markers (n=15, m=1, t=60).

Immunofluorescence revealed that SB treatment, like NANOG depletion, resulted in depletion of the H3K4me3 and H3K27ac epigenetic marks but not H3K27me3 (Fig. 3.22). Immunofluorescence also confirmed that co-injection

of SMAD2 rescued the deposition of these marks but was insufficient to fully rescue development (Not shown). POIII staining in the SB treated embryos demonstrates that the reduction of H3K4me3 and H3K27ac was not due to globally reduced transcriptional activity (Fig. 3.31b). Western blotting validated the results of the immunostaining (Fig. 3.31.c).

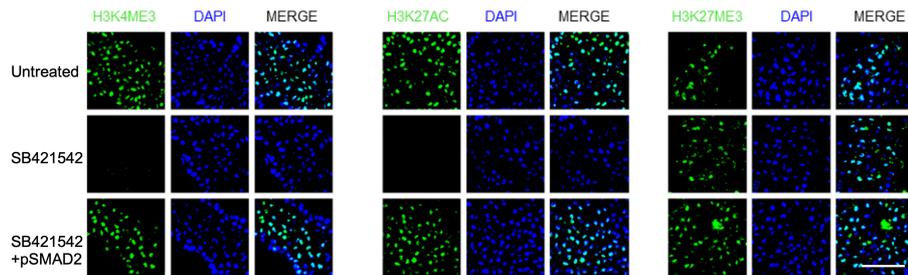


Figure 3.22. Immunofluorescent staining's of SB431542 ACs.

a, Uninjected, SB431542 treated and SMAD2 rescued AC explants cultured to equivalent stage 14 and probed for H3K4me3, H3K27ac, H3K27me3 using fluorescent antibodies and co-stained with DAPI. Scale bar, 60 μ m.

Data produced by James Dixon (JD) as well as my own data produced during my mRES suggest that when overexpressed in human embryonic kidney 293T cells axolotl NANOG and SMAD2 can interact, as determined by a luciferase complementation assay (Fig 3.23). Moreover, this interaction is disrupted by SB treatment. Together this suggested that the NANOG-SMAD2 complex identified in hESC may also have a role in axolotl development, also, given that both Nanog and nodal signal depletion led to the loss of H3K4me3. I next sought to identify the effects of removing H3K4me3.

It has been demonstrated that DPY30 is required for the deposition of trimethylated H3K4 residues but not H3K4 mono and demethylation (Ernst and Vakoc 2012). Moreover, data from ESC has shown that NANOG and SMAD2 recruit DPY30 for COMPASS mediated deposition of H3K4me3 (Bertero *et al.*, 2015); because of this, I next tested the effects of DPY30 KD in axolotl embryos.

Interestingly, DPY30 KD also resulted in developmental arrest; unexpectedly, however, this occurred post-gastrulation (Fig. 3.24). Like NANOG KD and SB treated embryos, the morphants halt development and remain alive long after siblings have reached tadpole stages. HREM reconstructions revealed morphants lacked a notochord, somites or other mesodermal structures, had a multi-layered AC and rudiments of a neural tube. Remarkably, this phenotype was rescued with 95% efficiency following co-injection of 200pg RNA encoding human DPY30 (100% n=30), demonstrating that DPY30 functionality is highly conserved in mammalian homologs.

As expected, DPY30 depletion also resulted in the loss of H3K4me3 and, like NANOG and NODAL signalling depletion, loss of H3K27ac as indicated by immunostaining and Western blotting (Fig. 3.25 and Fig. 3.31c, respectively). As DPY30 catalyzes the methylation of H3K4me2 to H3K4me3, this suggests that the loss of H3K27ac may be an indirect effect of NANOG/DPY30/NODAL signalling depletion. Phospho-POLII and H3K36me3 staining confirmed that DPY30 morphants remained transcriptionally active, while H3 staining confirmed no loss of histones (Fig 3.32a&b).

Transcriptomic analysis confirmed that much like NANOG KD, DPY30 morphants show reduced expression of pluripotency and lineage specification genes at equivalent stage 11 compared to uninjected controls (Fig. 3.26a). DPY30 morphants also fail to transcriptionally silence several pluripotency and specification associated genes. As with NANOG and NODAL signalling depletion, DPY30 KD embryos show severe reductions in the expression of mesodermal commitment markers and *HOX* gene expression (Fig. 3.26a&b). Interestingly, however, DPY30 KD embryos express neural tube markers such

as *NEUROD4* (Fig 3.26b) reminiscent of DPY30 KD in hESCs, which triggers neural differentiation (Bertero et al., 2015). GSEA showed that stage DPY30 morphant upregulated genes were enriched for markers of the blastula stage as well as organizer endoderm (Fig 3.27b). Downregulated genes were enriched for markers of several cell types: Lateral plate, Notochord, Somite and Dorsal lateral plate mesoderm. Interestingly, DPY30 morphant downregulated genes were also enriched for some neuroectodermal markers, including Chordal neural plate and anterior neural tube, suggesting that despite some neural tube formation, neural cell type formation is somewhat perturbed. QPCR confirmed and extended these observations, confirming that markers of differentiation were greatly reduced in morphants (Fig. 3.28). Explanted ACs depleted of DPY30 increased expression of pluripotency genes *NANOG* and *SOX2*, and reduced expression of epidermal markers (Fig 3.29). In a mesoderm induction assay, DPY30 depletion like *NANOG* KD reduced the expression of mesoderm marker *BRACHYURY*. However, caps did not display the same hypersensitivity to *ACTIVIN* (Fig 3.30a). Intriguingly, DPY30 depleted caps treated with FGF showed reduced expression of mesodermal marker *ACTA2* and increased expression of *NEUROD4* (Fig 3.30b). This suggests that DPY30 morphants are unable to properly form mesoderm but instead divert to a neural tissues. Taken together, these results suggest that DPY30 may regulate the formation of the same tissues in mammals and urodeles and the overlapping effects of *NANOG*, *SB* and *DPY30* suggest that the function of the *NANOG* -*DPY30*-*SMAD2* complex has been conserved through vertebrate evolution.

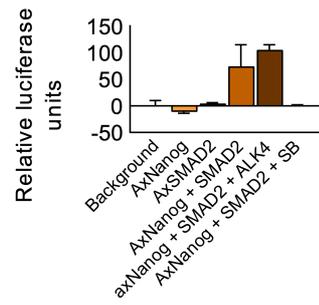


Figure 3.23. Axolotl NANOG can physically interact with axolotl SMAD2

a, Luciferase complementation assay. Axolotl NANOG can physically interact with axolotl SMAD2, interactions are increased in the presence of constitutively active ALK4, binding is disrupted with SB431542 treatment. (n=3, m=0, t=3)

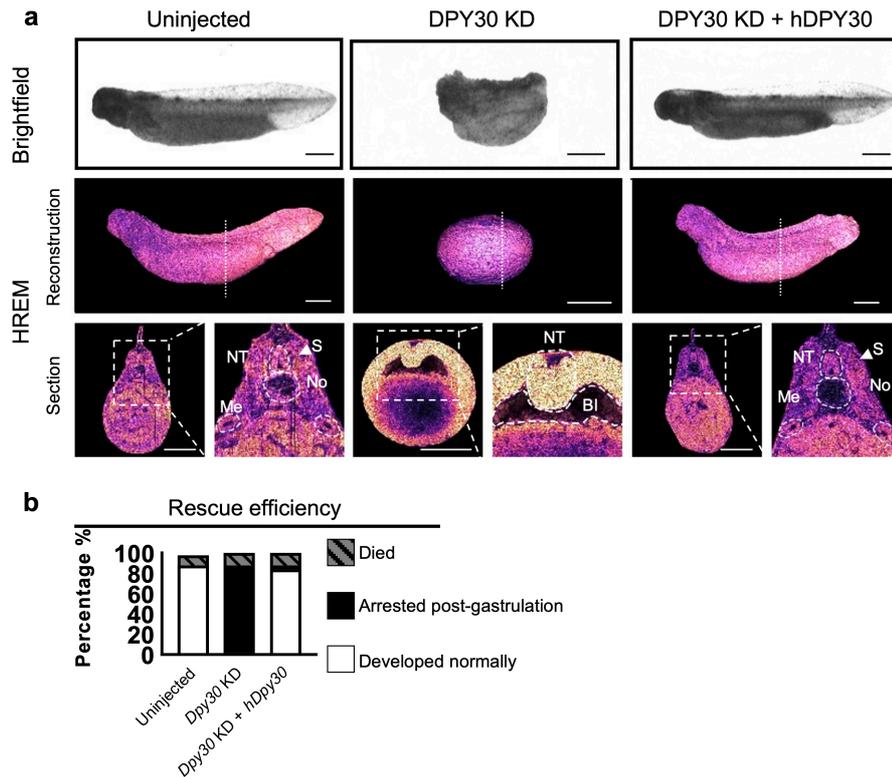


Figure 3.24. DPY30 KD arrests development post-gastrulation.

a, Brightfield and HREM images of uninjected and DPY30 depleted embryos at equivalent stage 42 (HREM $n=2$, $m=1$, $t=6$). Dotted line marks plane of section reconstruction (transverse). Dashed line delimits visible structures: Somites (S), Neural tube (NT) Notochord (No), Meso/pronephric ducts (Me), Blastocoel (B). Scale bar, 1mm. **b**, DPY30 KD and rescue efficiencies ($n= 30$, $m=1$, $t=90$).

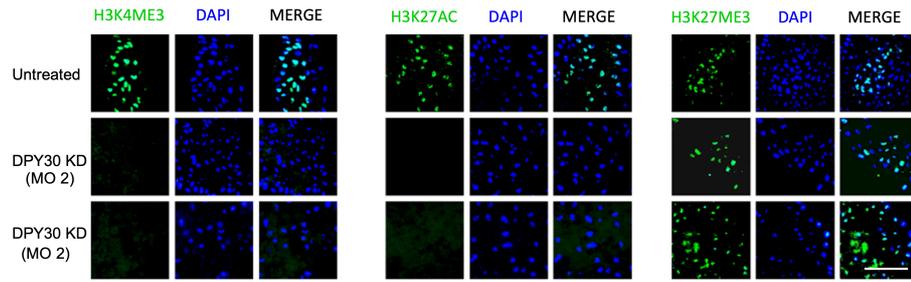


Figure 3.25. Immunofluorescent staining's of DPY30 KD ACs

a, Uninjected, DPY30 KD (morpholino 1) and DPY30 KD (morpholino 2) rescued AC explants cultured to equivalent stage 14 and probed for H3K4me3, H3K27ac, H3K27me3 using fluorescent antibodies and co-stained with DAPI. Scale bar, 60 μ m.

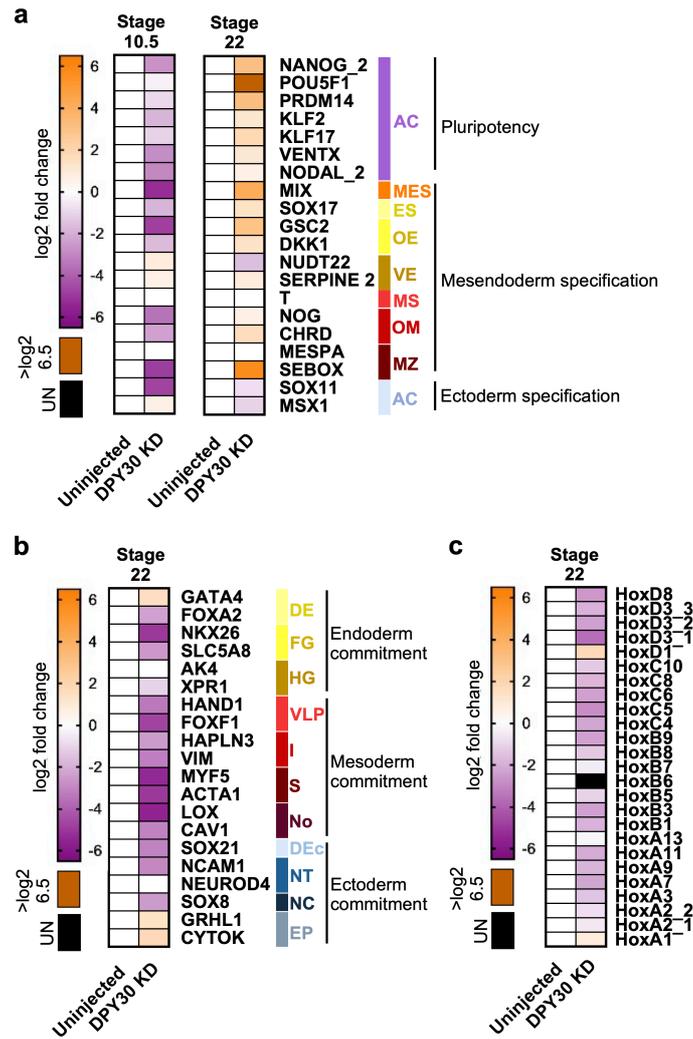


Figure 3.26. Differential gene expression in DPY30 KD embryos.

a, Heatmap showing differential gene expression of early marker genes at stages 10.5 and 22 in uninjected ($n=7$ and 6 , respectively, $m=1$, $t=13$) and DPY30 depleted ($n=6$ and 3 , respectively, $m=1$, $t=9$) embryos. Black indicates no detectable expression. Cell types: animal cap (AC), mesendoderm specification (general) (MES), endoderm specification (general) (ES), organiser endoderm (OE), vegetal endoderm (VE), mesoderm specification (general) (MS), organiser mesoderm (OM), marginal zone (MZ). **b**, Differentially expressed germ-layer commitment marker genes at stage 22 in uninjected and DPY30 KD embryos ($n=6$ and 3 , respectively). Black indicates no detectable expression. Cell types: definitive endoderm (general) (DE), foregut (FG), hindgut (HG), ventral-lateral plate (VLP), intermediate mesoderm (I), somite (S), notochord (No), definitive ectoderm (general) (Dec), neural tube (NT), neural crest (NC), epidermal progenitors (EP). **c**, Differential gene expression of Hox gene family members in response to DPY30 KD at stage 22.

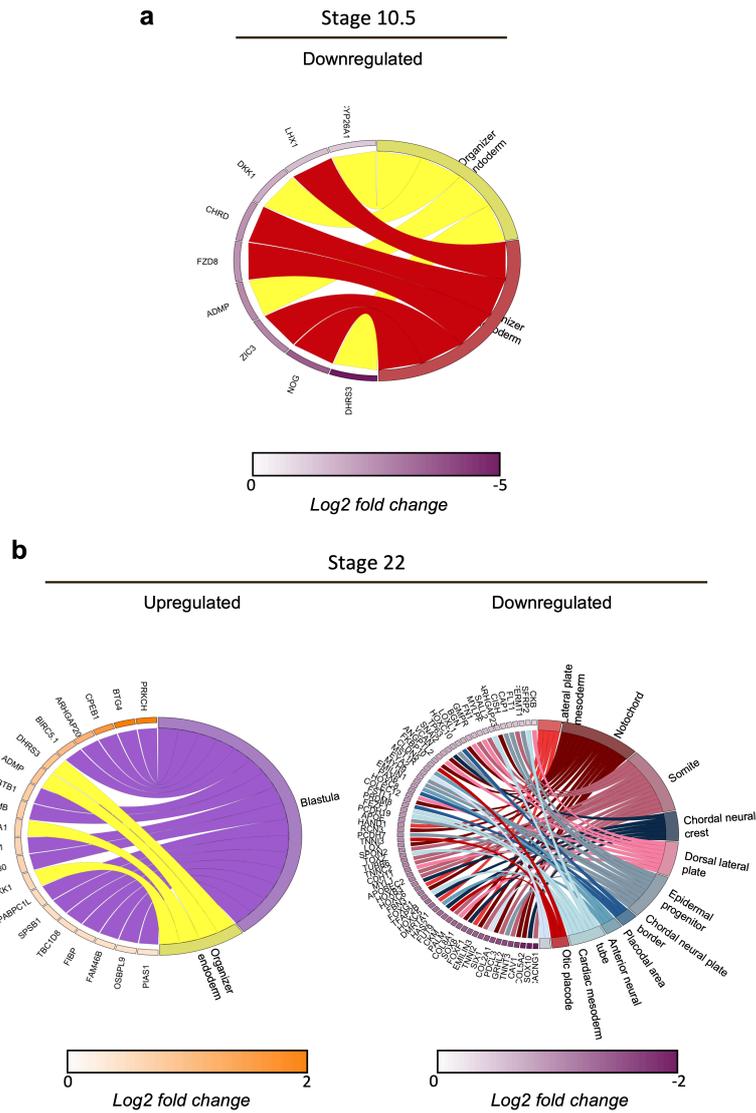


Figure 3.27. Amphibian cell-type marker expression in DPY30 depleted embryos. Chord diagrams showing statistically significant ($FDR < 0.01$) amphibian cell-type markers within DPY30 KD differentially expressed genes identified from gene set enrichment analysis (GSEA). **a**, Significant marker genes enriched in the up and down-regulated genes following DPY30 KD at stage 10.5. **b**, Significant marker genes enriched in the up and down-regulated genes following DPY30 depletion at stage 22.

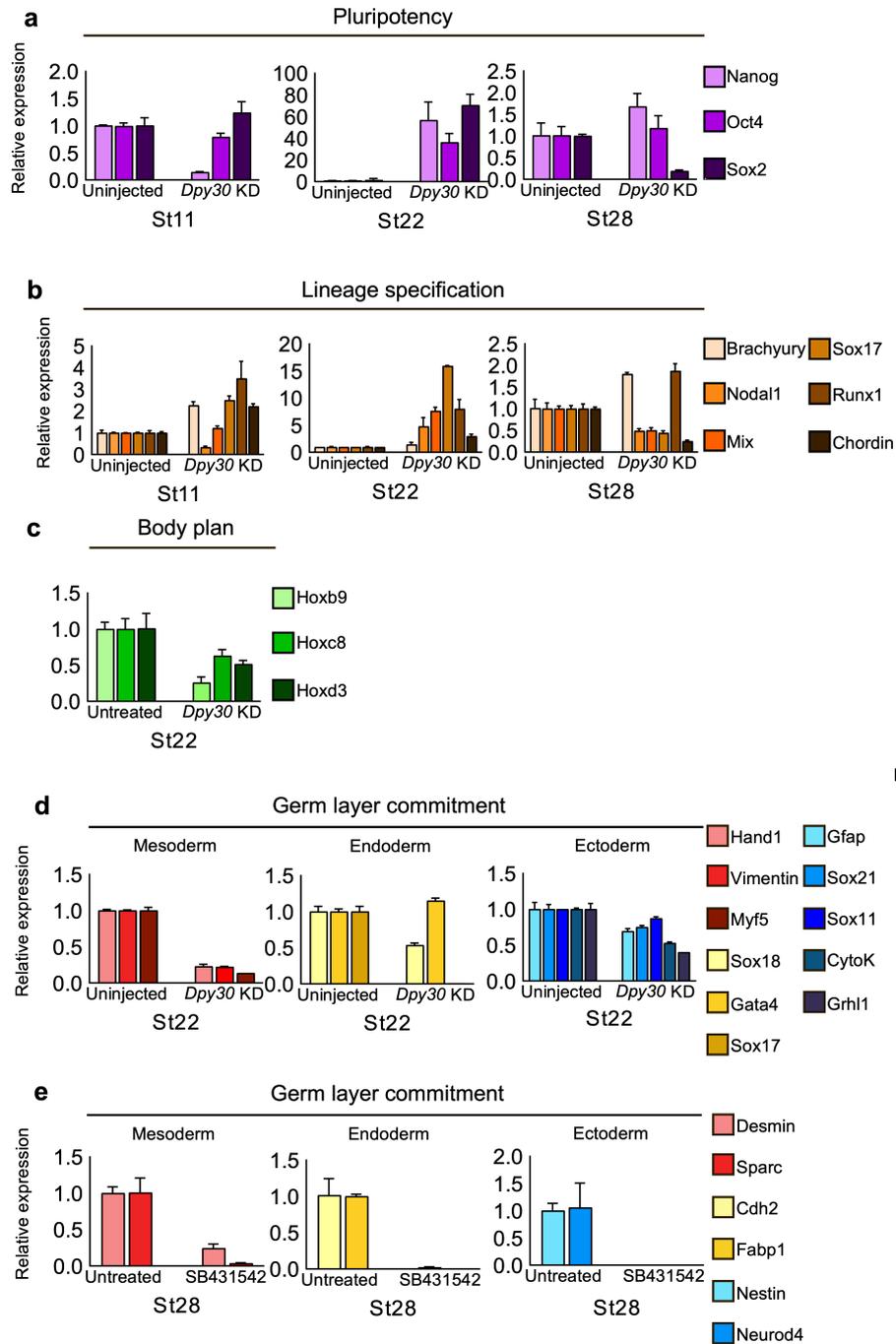


Figure 3.28. QPCR validation of key germ layer markers in DPY30 KD embryos.

a, Pluripotency factor expression in stage 10.5 embryos with and without DPY30 KD. (n=10, m=1, t=20) **b**, Germ-layer specification markers (n=10, m=1, t=20). **c**, Hox-gene expression. (n=10) **d**, Germ-layer commitment gene expression at stage 22. **e**, Late stage (stage 28) differentiation markers of uninjected and DPY30 depleted embryos. (n=10, m=3, t=60)

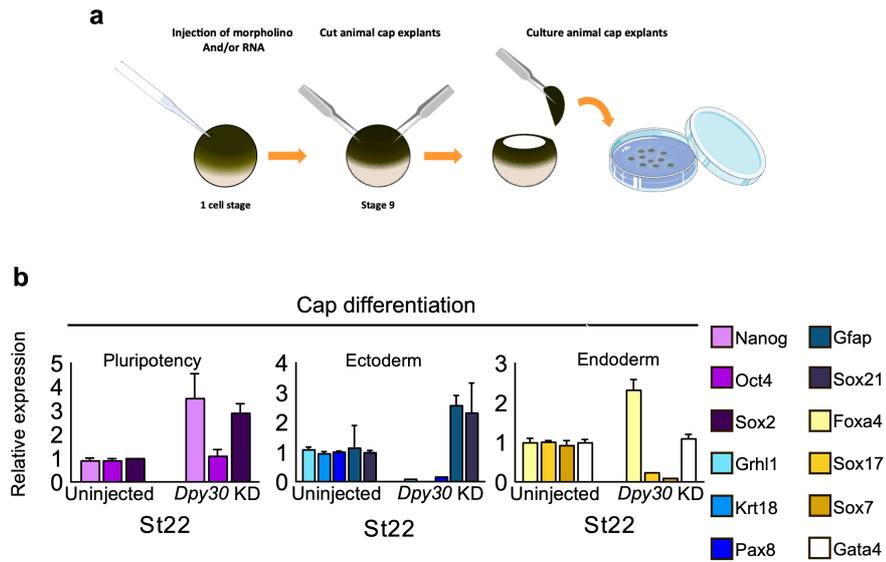


Figure 3.29. Germ-layer markers in DPY30 KD ACs.

a. Schematic of DPY30 KD regime and explant culture. **b.** Germ layer differentiation markers of uninjected and DPY30 depleted AC explants at stage 22. (n=15, m=3, t=90)

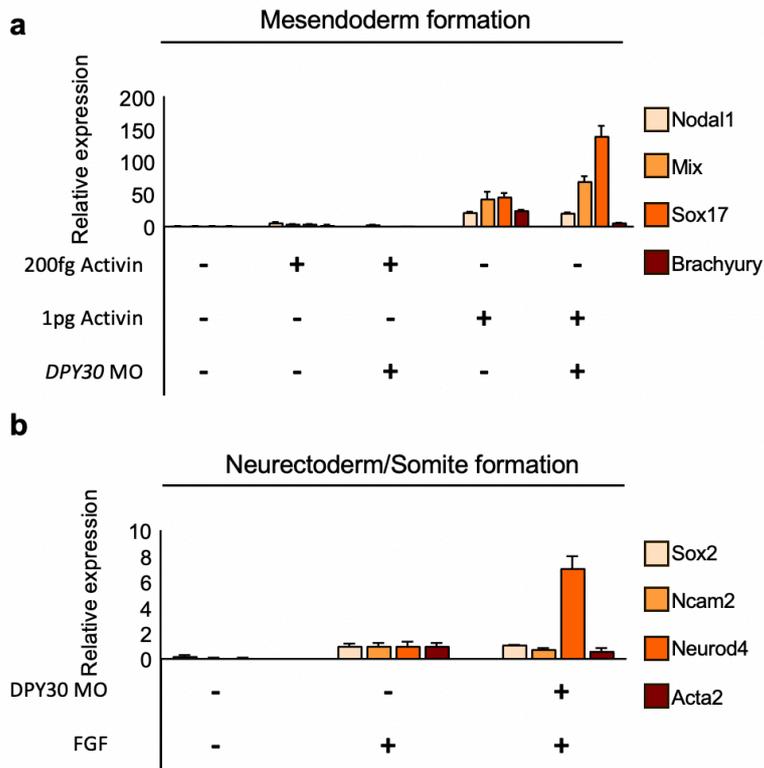


Figure 3.30. Mesendoderm induction assays in DPY30 KD ACs

a, QPCR showing expression of mesodermal and endodermal germ-layer markers in uninjected and DPY30 depleted stage 20 caps following treatment with different activin concentrations. (n=15). **b**, QPCR of germ-layer markers of uninjected and DPY30 depleted stage 20 caps following treatment with FGF. (n=15)

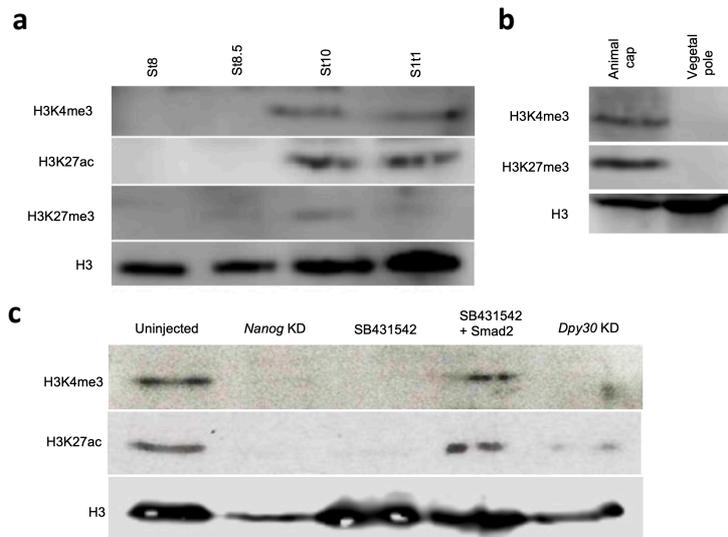


Figure 3.31. Western blots showing acquisition of epigenetic marks in axolotl embryos. NANOG KD results in the loss of markers of activating promoter marks. **a**, Western blots showing H3K4me3, H3K27ac and H3K27me3 appear after ZGA at stage 8.5. **b**, Western blots showing vegetal explants have lower levels of H3K4me3. **c**, Western blotting confirming the depletion of H3K4me3 and H3K27ac in response to NANOG KD, SB431542 and DPY30 KD.

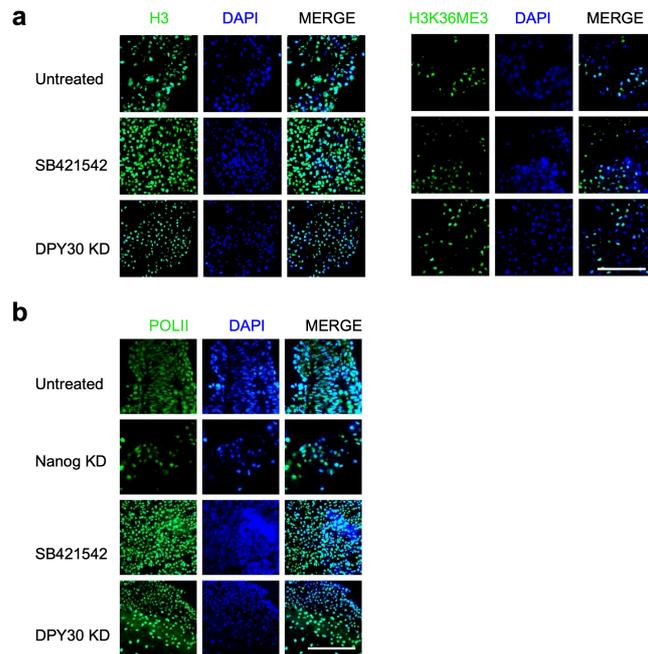


Figure 3.32. Immunofluorescent staining's of SB431542 treated, Nanog and DPY30 depleted ACs.

a, Uninjected, NANOG depleted and hNanog rescued AC explants cultured to equivalent stage 14 and probed for H3 and H3K36me3. Also shown is DAPI co-staining. Scale bar, 60μm.
b, Untreated, NANOG depleted, SB431542 treated and DPY30 depleted animal cap explants cultured to equivalent stage 14 and probed for phospho-POLII and co-stained with DAPI. Scale bar, 60μm.

To further elucidate the direct function of the NANOG -SMAD2-DPY30 complex in vertebrate embryogenesis, transcriptomic data from uninjected embryos, SB treated, NANOG KD and DPY30 KD was compared (Fig 3.33). With the help of Teri Forey, I identified 1508 'early activated' genes whose expression is usually reduced between stage 10.5 and 22 that are significantly upregulated in NANOG and DPY30 morphants only; this included pluripotency and early specification genes (Fig 3.33a). I also observed overlap in 470 genes normally upregulated between stages 10.5 and 22 that are instead down-regulated by each treatment. Notable genes included within this data set included numerous mesodermal commitment genes such as *MYF5*,

VIMENTIN and *HAND1*, as well as several members of the *HOX* gene family. GSEA analysis of these downregulated 470 genes showed significant enrichment for markers of mesodermal and a subset of neural tissues, consistent with the loss of these tissues after each treatment (Fig 3.33b). Remarkably, however, while mesendodermal specification genes are downregulated by nodal signalling inhibition, they are upregulated by NANOG or DPY30 KD at later stages, suggesting SMAD2/3 can act independently of interaction with *Nanog* in early specification events.

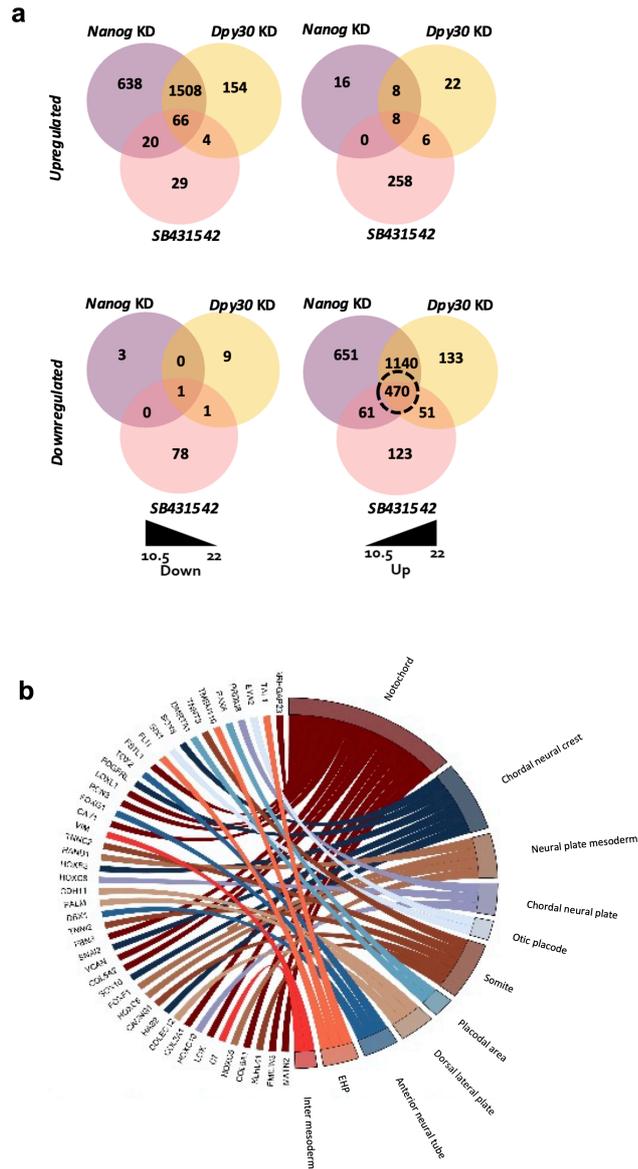


Figure 3.33. Overlapping effects of SB431542 treatment, NANOG and DPY30 depletion.
a, Venn diagram showing overlapping differentially expressed genes in NANOG, DPY30 and NODAL depleted embryos at stages 10.5 and 22. **b**, Chord diagram showing the results of gene set enrichment analysis, showing enrichment of amphibian cell-type specific markers in overlapping downregulated genes.

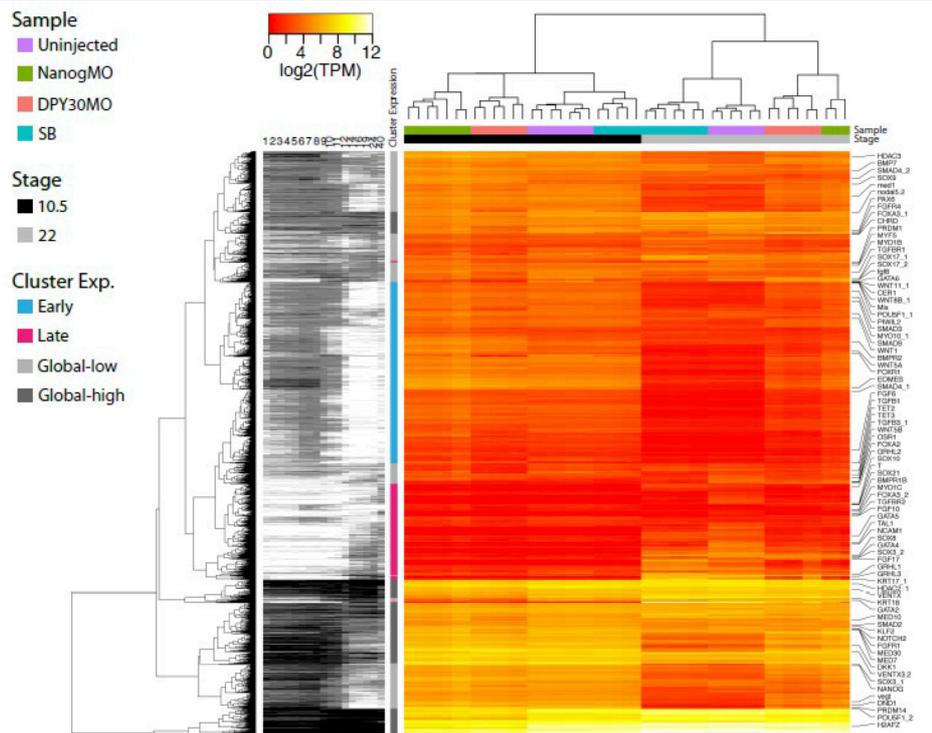


Figure 3.34. Clustering of experimental treatments and key gene expression.

Heatmap showing hierarchal clustering of ‘early’ or ‘late’ genes in uninjected, SB431542 treated, NANOG and DPY30 depleted embryo transcriptomes at stages 10.5 and 22. The ‘early’ genes were defined by genes which had an expression greater than 10TPM and had a significantly lower expression at stage 10.5 than at stage 22. Correspondingly, ‘late’ genes were defined by genes which had a TPM greater than 10 at stage 22 and had significantly lower expression at stage 10.5. Generally, early activated genes are more highly expressed even in late-stage embryos following Nanog/DPY30 KD while both early and late genes show downregulation in response to SB431542 treatment.

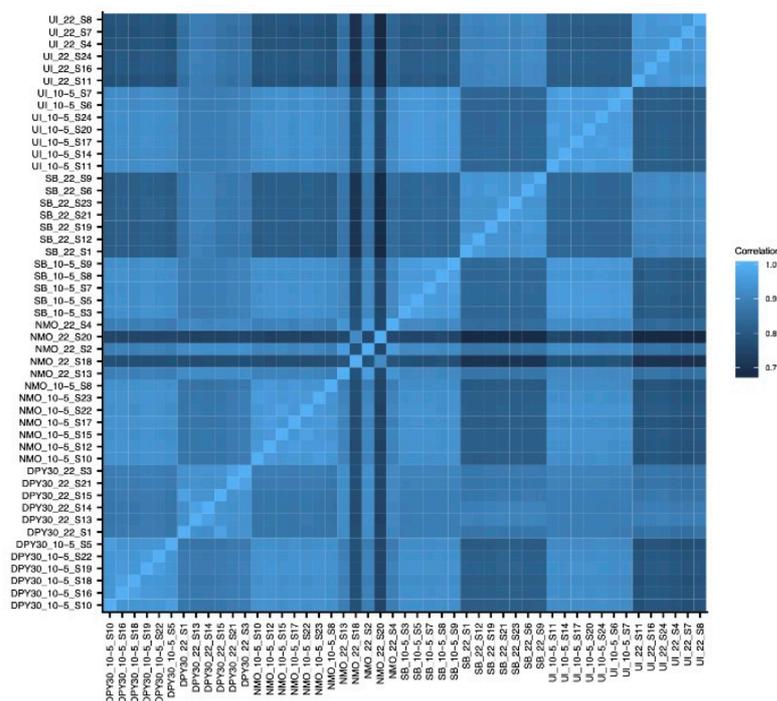


Figure 3.35. Correlation heatmap of experimental conditions.

Heatmap showing the Spearman correlation between the transcriptomes of uninjected, SB431542 treated, NANOG and DPY30 depleted embryos at stages 10.5 and 22.

Finally, I evaluated whether the transcriptional effects associated with diminished NANOG, DPY30 or NODAL signalling correlated with changes in H3K4me3 levels using ChIP-qPCR (Fig 3.36 & 3.37). ChIP revealed H3K4me3 levels on promoters of *EEF1A1* or *CYTOK* were unaffected in stage 10.5 caps under each experimental condition (Fig. 3.36) suggesting expression is independent of the NDS complex. However, H3K4me3 was reduced under all three conditions on the *NANOG* and *NODAL1* promoters. Given that these genes first express at ZGA, I posit that their activation is mediated by maternal factors not dependent on NDS but that the H3K4me3 mark may be required to integrate their expression into the pGRN that extinguishes their expression after gastrulation to initiate the subsequent waves of gene expression that drive development.

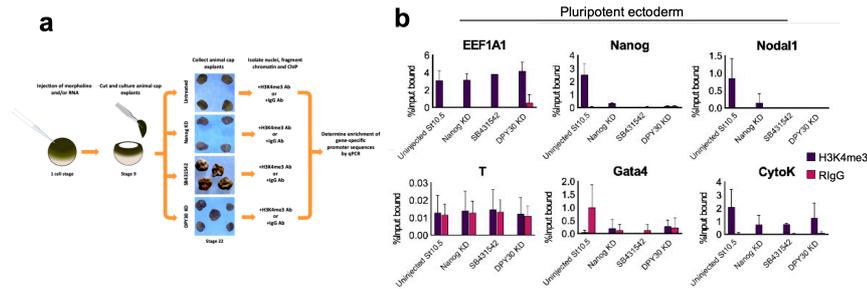


Figure 3.36. H3K4me3 ChIP in ACs following SB431542 treatment, NANOG or DPY30 depletion

a, Schematic of H3K4me3 ChIP with explant images shown. **b**, H3K4me3 ChIP of stage 10.5 uninjected, NANOG KD, DPY30 KD and SB431542 treated caps followed by qPCR using probes directed at gene promoter regions. (n=50, m=1, t=200)

I then tested the effects of each experimental regime on mesoderm induction by performing ChIP on stage 22 caps that had also been co-injected with 1pg of activin RNA along with either the NANOG or DPY30 MO. H3K4me3 levels were then compared with caps at the same stage that were untreated or exposed to SB. In addition, I included untreated caps at stage 10.5 to determine if the H3K4me3 mark is induced by ACTIVIN or laid down prior to mesoderm inducing signals (Fig. 3.37). ChIP revealed that ACTIVIN induced deposition of H3K4me3 on promoters of mesodermal commitment genes *MYF5* and *VIMENTIN*, and this was prevented by either NANOG or DPY30 KD. This mark was also absent in uninjected, and SB treated ACs correlating with the absence of expression. The *NODAL1* and *GATA4* promoters also showed reduced enrichment of H3K4me3 in response to depletion of NANOG and DPY30. Like uninjected caps, those treated with SB also showed no enrichment of H3K4me3. Whilst expression of these genes is not present in uninjected or SB treated caps, they were activated after either NANOG or DPY30 depletion, and their expression was not extinguished post-gastrulation.

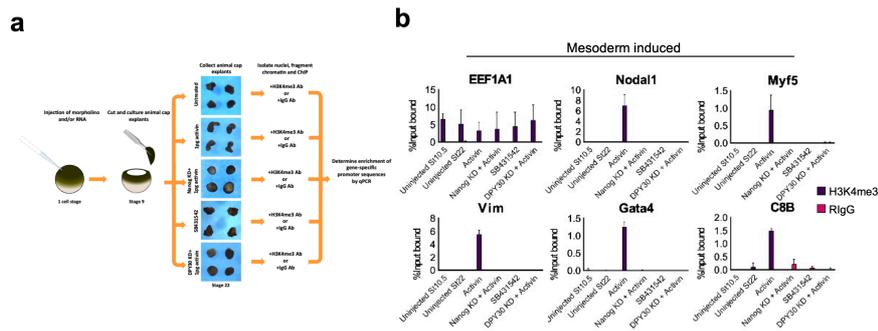


Figure 3.37. H3K4me3 ChIP in mesoderm-induced ACs following SB431542 treatment, NANOG or DPY30 depletion

a, Schematic of H3K4me3 ChIP with explant images shown. b, H3K4me3 ChIP-qPCR of equivalent stage 10.5 and 22 uninjected caps, stage 22 Activin treated caps with and without NANOG and DPY30 depletion as well as SB431542 treated caps. (n=50, m=1, t=300)

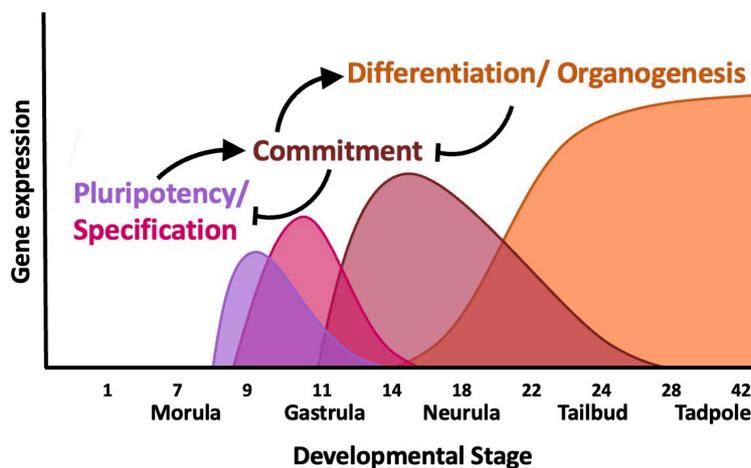


Figure 3.38. Schematic: Successive waves of gene expression drive development

a, Schematic of showing the successive waves of gene expression that drive embryogenesis.

Combined, these data reinforce the concept that NANOG and SMAD2/3 regulation of H3K4me3 is required for mesodermal commitment. Furthermore, this suggests SMAD2/3 regulates transcription of some early genes independent of NANOG; ultimately, though, both are required to regulate the sequential waves of embryonic gene expression in vertebrate development (Fig. 3.38).

When combined with the transcriptomic data and western blotting, this provides insight into the sequential events that take place, which are required for mesodermal commitment. We, therefore, posit the following model (Fig. 3.39) in which SMAD2 and DPY30 are maternally inherited molecules, highly expressed prior to ZGA in the stage 8 blastula. Around stages 8-11, zygotic *NODAL* and *NANOG* expression commences; the former facilitates phosphorylation of SMAD2 (SMAD2-P), enabling translocation into the nucleus and activation of Nodal target genes. SMAD2-P is also able to form a complex *de novo* with NANOG and DPY30/COMPASS, which are required to prime mesodermal commitment gene promoters directly or via genes acting upstream, through the deposition of H3K4me3. The activation of lineage commitment genes from stages 11 onward is required for the extinguishment of pluripotency and specification specific genes. Data from hESC has demonstrated that the loss of H3K4me3 due to DPY30 KD preceded changes in gene expression (Bertero *et al.*, 2015). Therefore, I posit that the NANOG - SMAD2-DPY30 complex appears to regulate the deposition of H3K4me3 and is required for germ layer commitment, particularly the formation of mesoderm. It, therefore, seems likely that the formation of all three germ layers is a prerequisite for the progression of development, organization of the body plan and organogenesis in vertebrates.

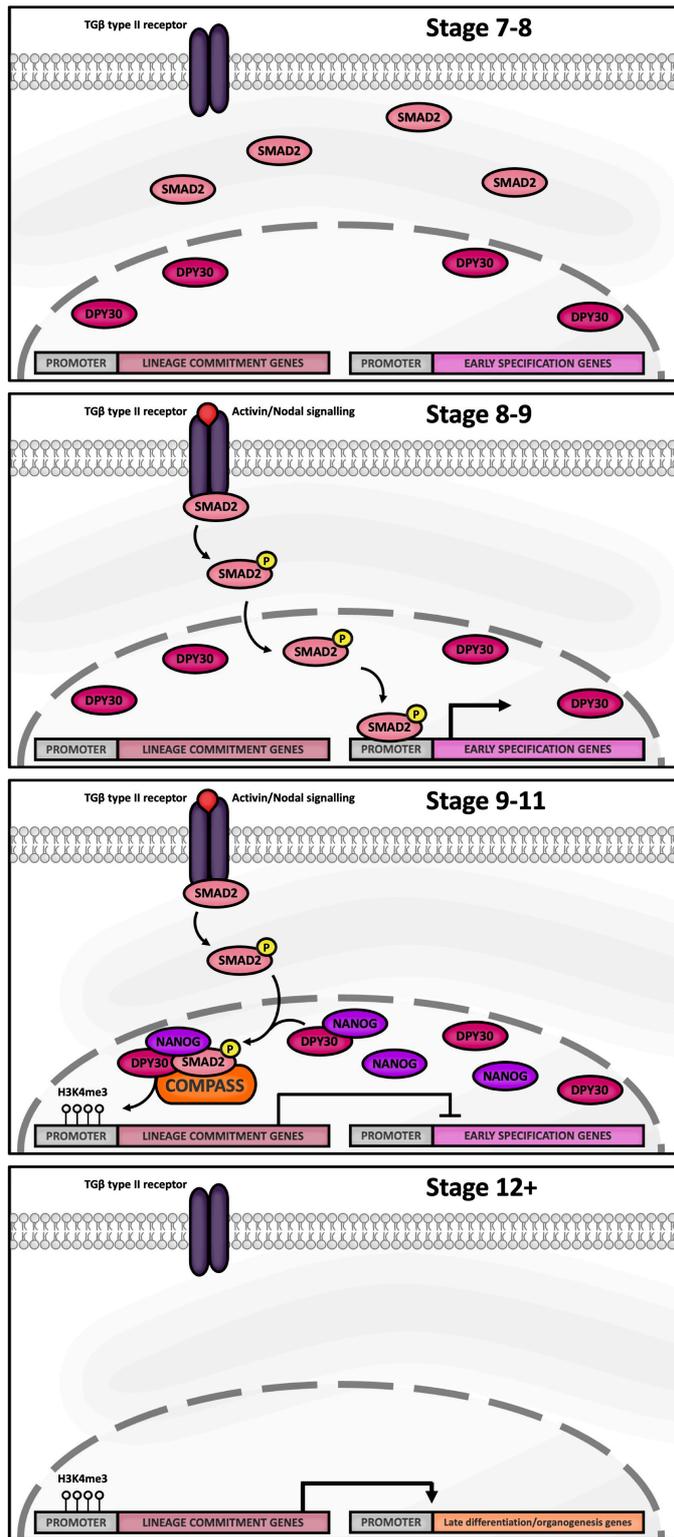


Figure 3.39. Schematic: NANOG complexes with SMAD2 and DPY30 to deposit H3K4me3. Schematic of the proposed temporal dynamics of NANOG and SMAD2 mediated deposition of H3K4me3 during axolotl development.

3.3 Discussion: NANOG activity is conserved

3.11.1 Recap of aims and objectives

In this chapter, I set out to address the following questions regarding NANOG's role in development:

- Is NANOG required for axolotl development?
- Does depletion of NANOG affect early cell-fate decisions?
- Does NANOG regulate the expression of other pluripotency factors?
- Which genes are affected by NANOG depletion?
- Are there any similarities between the effects of NANOG depletion in axolotl and other animals?

3.11.2 Summary of key findings

In this study, I addressed several research questions concerning the role of NANOG in early axolotl development. I have shown that NANOG is required for axolotl development beyond the late blastula stage. NANOG depletion does not affect the cell fate decisions in the vegetal pole but prevents the silencing of pluripotency gene expression and perturbs the AC from committing to all three germ layers. NANOG is required for the wide-spread deposition of H3k4me3 and H3K27ac in the AC following its activation post-ZGA, a function it likely carries out with SMAD2 and DPY30. Moreover, the mechanism by which NANOG carries out epigenetic modelling appears to be conserved between urodele amphibians and mammals.

3.11.3 Discussion of key findings

Urodeles (salamanders) represent the amphibious ancestor to reptiles, synapsids and mammals, which diverged from a common ancestor around 352 million years ago (Kumar et al., 2017). Urodeles retained the basic skeletal structure of the tetrapod ancestor, and importantly an ancestral urodele-like embryology was conserved during the evolution of amniotes (Bachvarova *et al.*, 2009a; Callier *et al.*, 2009; Niedzwiedzki *et al.*, 2010). In contrast to sister taxa anurans (frogs), the cells of the AC of urodeles manifest true pluripotency as they give rise to both the soma and the germline and therefore represent a useful model for the study of pGRN in vertebrates. In urodeles, as in amniotes, form PGC's through epigenesis within the posterior lateral mesoderm, a trait also conserved in chordate embryology (Bachvarova et al., 2009a, Bachvarova et al., 2009b, Chatfield et al., 2014, Johnson and Alberio, 2015). It has been previously demonstrated axolotl NANOG 's ability to functionally replace mouse Nanog *in vitro* (Dixon et al., 2010). Here I have presented evidence that the function of NANOG and the mechanism by which NANOG programs the pluripotent domain in the animal cap is conserved in vertebrates.

Interestingly, I found that the expression of many critical early specification genes are not NANOG dependent; however, NANOG, SMAD2 and DPY30 are all required to complete germ layer commitment. These data align with the inability of NANOG morphants to undergo gastrulation. Moreover, morphants fail to acquire the H3K4me3 mark after ZGA, which is evident at the loci of key lineage commitment genes. There is controversy surrounding a cause and effect relationship between H3K4me3 and gene transcription. Some suggest that the enrichment of H3K4me3 at TSS is a consequence of increased transcriptional activity rather than a cause (Howe et al., 2017). However, the

majority of studies that investigate this focus on the depleting SET1/COMPASS components after the establishment of H3K4me3 and then studying transcriptional rate. Here I have found that the KD of COMPASS component DPY30 prior to establishing H3K4me3 results in abrogated lineage commitment. Comparison of our transcriptome data shows 1610 of 1794 significantly downregulated genes at stage 22 in DPY30 morphants are also down-regulated in NANOG morphants at the same stage.

Unlike NANOG and SMAD2, DPY30 has not been shown to act as a transcription factor. Therefore, it seems likely that the downregulation of these genes and failure to commit to all three germ layers is a direct consequence of H3K4me3 ablation. Furthermore, this implies that H3K4me3 is a requisite for germ layer formation. Therefore, I propose that the epigenetic priming of regulatory elements by the NANOG -SMAD2-DPY30 complex confers competency for germ layer formation in response to early signalling events, which is the defining characteristic for pluripotency. Moreover, this study also allowed for greater insights into the temporal events leading to the establishment of pluripotency to be understood.

Transcriptomic and Western blotting data provide insight into the sequential events which drive the establishment of pluripotency required for embryogenesis. I posit the following model: SMAD2 and DPY30 are maternally inherited molecules, highly expressed prior to ZGA, which occurs at stage 8 (Jiang et al., 2017). At stage 8, zygotic *NODAL* and *NANOG* begin to be transcribed; the former leads to the phosphorylation of SMAD2, whereby it translocates the nucleus and activate target genes. SMAD2 can also form a complex *de novo* with NANOG and DPY30. Given that SB431542 treated embryos do not activate early lineage specification genes in contrast to the

NANOG and DPY30 morphants, SMAD2 signalling is first required for early signalling events independent of NANOG and DPY30. This hypothesis is supported by gene expression data which shows at this stage, *SMAD2* transcripts outnumber *NANOG* transcripts approximately 9 to 1. Thus, only a limited amount of phospho-SMAD2 is present in the NSD complex, and the surplus phospho-SMAD2 is free to activate NANOG -DPY30 independent gene targets, which likely include early germ layer specification genes. By stage 9, *Nanog* has reached its peak expression and is able to form the NSD complex and catalyse the deposition of H3K4me3 at key loci throughout the genome, evidenced by the emergence of H3K4me3 during stages 9 and 10. Interestingly, the absence of this mark results in an inability to form all three germ layers.

This finding has several implications for the evolution of the vertebrate embryo. In particular, there is a clear difference between the mechanisms which govern cellular potency between Urodeles and sister taxa Anurans, as *Nanog* was lost from the frog genome (Dixon et al., 2010). While it has been proposed that *XVentx1/2* may fulfil the role of *NANOG* (Scerbo et al., 2014, Scerbo et al., 2012), *XVentx1/2* morphants gastrulate normally and possess all three germ layers but exhibit compression along the anterior-posterior axis (Scerbo et al., 2012). Indeed, it was demonstrated that the injection of mNanog could rescue this phenotype. While mNanog may be able to functionally compensate the loss of *XVentx1/2* possibly due to shared ancestry, given the disparity between *XVentx1/2* and *NANOG* KD, it seems unlikely that *XVentx1/2* is of functional equivalence to *NANOG*. In the case of *Xenopus*, *Nanog* could not have been lost without significant compensatory effects. In particular, *Xenopus* embryos must have evolved a differential mechanism for

the deposition of H3K4me3, as it is NANOG independent. Furthermore, I have found no evidence of H3K4me3 prior to *Nanog* activation; in contrast, it has been demonstrated that a large proportion of H3K4me3, among other epigenetic marks, are maternally inherited in *Xenopus* (Hontelez et al., 2015, Bright et al., 2021). Interestingly, a significant portion of these genes show differential regulation following NANOG depletion in axolotl (Hontelez et al., 2015, Bright et al., 2021). This may suggest that Anurans had to first evolve mechanisms to deposit H3K4me3 at these loci before *Nanog* could be lost from the genome.

As urodeles represent the amphibious ancestor to mammals (Bachvarova *et al.*, 2009a; 2009b; Johnson *et al.*, 2001; Johnson *et al.*, 2003a; 2003b; Callier *et al.*, 2009; Niedzwiedzki *et al.*, 2010), then crucially the mammalian epiblast evolved from the AC. In mice, H3K4me3 is only established after ZGA at the two-cell stage (Liu *et al.*, 2016); thus, it seems likely that the widespread maternal deposition of H3k4me3 is not a conserved feature in vertebrates.

3.11.4 Study limitations

Unlike *Xenopus*, it is difficult to artificially fertilise a clutch of eggs without killing both the male and female. As a result, embryos can only be obtained following a natural mating, the success of which varies and can frequently be unsuccessful. During this study, often, long periods would go by (sometimes several months), yielding no successful matings. Given that a female may also only lay up to around 20 embryos per hour (internal records), obtaining a large number of embryos that can be injected at the one-cell stage can be difficult. Moreover, the number of fertilised embryos per successful mating varies from around 120-800 embryos. The result of these limitations of the axolotl meant

that getting sufficient numbers for experiments was difficult. In this study, we utilised ChIP-QPCR to examine the enrichment of H3K4me3 at specific loci. However, given the volume of embryos required for the ChIP (50 per condition), the experiment could not be repeated with three separate mating's. Therefore, results may not be as generalisable.

3.11.5 Future work

Here I looked at the genome-wide expression of the axolotl AC and compared it with the pluripotent/multipotent cell populations in humans, pigs and frogs. However, during this study, the first multi-stage sc-transcriptomes of early pig, frog and zebrafish development were published (Briggs et al., 2018, Wagner et al., 2018, Zhi et al., 2021). Sc-sequencing is a potent tool and, if applied to the axolotl, would allow for many insights into the regulation of cell fate decisions. Importantly, sc-sequencing datasets of axolotl early development could be compared at multiple levels to other vertebrate models, which could provide critical insights into the evolution of the vertebrate embryo. Moreover, Sc-sequencing could be applied to NANOG depleted embryos to study NANOG 's role in regulating the organiser, marginal zone and presumptive ectoderm. Also, given the increased availability of sc-sequencing technologies and reduced costs associated with sequencing, this avenue may be a goal of future work. In a similar vein, the decision to use qPCR assay for specific gene promoter sequences following ChIP in this study was made primarily out of the costs associated with ChIP-seq. However, qPCR required the testing of several primers and, as a result, meant only select promoter sequences could be interrogated. Future work could utilise lower sequencing costs and, in the process, map out loci in which H3K4me3 deposition is NANOG dependent; this could also be compared with

equivalent experiments already performed in *Xenopus* (Hontelez et al., 2015, Bright et al., 2021, Akkers et al., 2009) and provide insights into the mechanisms by which a gene as critical to development was lost in anurans.

ELK1 modulates mesodermal development and germ line competence

4.1 Introduction

More than 30 of the 35 described animal phyla in metazoans are triploblasts that produce their soma from three germ layers established during gastrulation: the outer ectoderm and the two inner layers, endoderm and mesoderm. It is well established that triploblasts evolved from ancient diploblasts, which derive their soma from endoderm and ectoderm only (Miller et al., 2005, Burton, 2008, Technau and Scholz, 2003). Despite this, few phyla exist, which are diploblasts and undoubtedly, the evolution of mesoderm has contributed to the expansion of triploblasts.

As mentioned in earlier chapters, pluripotency is a conserved property of embryonic cells early in triploblast development, whereby cells of the early zygote are competent to produce both three distinct germ-layers and the germline in response to inductive signals (Gardner and Rossant, 1979, Tam and Zhou, 1996). This suggests that pluripotency must have evolved prior to the evolution of mesoderm. Understanding the pGRN is key to understanding the evolution of bilateria, including vertebrates. To what extent the pGRN is conserved in triploblasts remains unclear. This is in part because many well-studied triploblasts have also developed species-specific innovations to

embryogenesis which have resulted in differing states of cellular potency. As mentioned earlier, early mammalian zygotes exist in a state of totipotency in order to give rise to the embryonic tissues and mammal-specific extra-embryonic tissues. In this case, pluripotency manifests after the segregation from the trophoblast. Therefore, the study of lower vertebrates that have not evolved these adaptations would be useful to study the pGRN. However, this is also problematic as the two main lower vertebrate models of early development, Frogs and Zebrafish, do not display true cellular pluripotency *in vivo*, as both models give rise to their soma and germline from separate groups of cells with differing cell potency.

The pGRN was first defined *in vitro* using embryonic stem cells (ESC) in mice and humans and has since been extensively studied. In this model, cytokine signalling can be used to manipulate and block the natural acquisition of cell identity, allowing indefinite propagation without loss of developmental potential. Alternatively, sufficient disruption to crucial components of the pGRN gives way to differentiation programs (Smith et al., 1988, Boyer et al., 2005, Loh et al., 2006, Li and Izpisua Belmonte, 2018, Li and Belmonte, 2017).

One such example is the transcription factor ETS-like protein 1 (ELK1), which has been shown to be essential in maintaining hESC pluripotency, acting as both an independent repressor of differentiation and as a substrate of ERK2 whereby, it acts as a transcription activator (Goke et al., 2013). *ELK1* knockdown (KD) resulted in upregulation of differentiation genes and the loss of pluripotency, resulting in differentiation toward a mesodermal, fibroblast-like cell type (Goke et al., 2013). This also suggested that ELK1, in conjunction with ERK2, act as the major effector of the MAPK pathway, which, as discussed in detail earlier, is critical to hESC pluripotency (Brons et al., 2007,

Tesar et al., 2007, Vallier et al., 2005). A more recent study implies that *Elk1* KD in hESC does not directly affect pluripotency but instead results in hypersensitivity to mesodermal differentiation (Prise and Sharrocks, 2019). ELK1 playing such a prominent role in the maintenance of pluripotency in hESC was also paradoxical specifically because a study in mice found that deletion of ELK1 through homologous recombination did not affect the ability of mESC to be propagated in culture.

Furthermore, *Elk1*^{-/-} mice developed with no abnormalities (Nordheim et al., 2004). Given that ELK1 orthologues have been identified throughout vertebrates and even in early extant metazoans (Saxton et al., 2016), it is unclear whether the role of ELK1 as defined in hESC, is ancient or derived. While the role of ELK1 has been explored in lower vertebrates, the only study to date depleted Elk1 in *Xenopus laevis* (Nentwich (Nentwich et al., 2009). Nentwich and colleagues reported that Elk1 KD prevented mesodermal tissues' formation, an effect that heavily contrasts the defined role in hESC. Interestingly, the role of MAPK signalling to maintain pluripotency may differ between mammals and amphibians (Chatfield et al., 2014, Nakanoh and Agata, 2019). However, as the sequence of ELK1 shows deep conservation across metazoan, this may suggest a conserved function (Saxton et al., 2016), and therefore the role of ELK1 as described in anurans may not represent a conserved function particularly, as *Xenopus laevis* like other anurans do not manifest cellular pluripotency.

ELK1 contains four conserved domains with well-defined, distinct functions in humans. The N-terminal ETS binding domain mediates direct DNA binding while the B box domain facilitates the physical interaction with

ELK1's binding partner: SRF (Hassler and Richmond, 2001, Janknecht and Nordheim, 1992, Shore and Sharrocks, 1994). The D domain acts as a docking site for mitogen-activated protein kinases (MAPKs) (Yang et al., 1998a, Ling et al., 1998, Jacobs et al., 1999) and the C domain functions as a MAPK inducible transcription activation domain (Hill et al., 1993, Marais et al., 1993, Zinck et al., 1993, Janknecht et al., 1994, Strahl et al., 1996, Cruzalegui et al., 1999). While the ELK1 protein is conserved throughout metazoans, ELK1 functional domains vary in their conservation (Saxton et al., 2016). While the function of these domains is well characterised in differing cellular contexts *in vitro*, it is unclear whether these domains are critical to ELK1's role in development.

Here I investigated ELK1's role in development in the basal vertebrate *Ambystoma mexicanum*. I demonstrate that ELK1's early role is not to act as a pluripotency factor directly but rather through transcriptional repression. This activity appears to be conserved from amphibians to large mammals. ELK1 morphants also show an expansion of the somitic mesoderm at the expense of intermediate and ventral mesodermal structures, including PGCs. I also demonstrate a previously unidentified role of ELK1 in safeguarding germline development which is also conserved between amphibians and large mammals. Using inductive assays, I show ELK1 is first required to activate canonical *Wnt* signalling to convey germline competency, an activity that requires the interaction with the downstream MAPK signalling effector ERK2. Later, ELK1 likely co-operates with the mediator subunit 23 (MED23) to drive mature PGC differentiation. Our results are consistent with the idea that the ancient role of ELK1 was likely to repress somitic mesodermal development and establish the germline.

4.2 Results

Previous work by our group has described the stepwise evolution of ELK1 orthologues in metazoans, including an axolotl Elk1 orthologue (Saxton et al., 2016). Axolotl Elk1 has high sequence conservation with other vertebrate ELK1s, including human ELK1 and *Xenopus* Elk1 (Fig. 4.1). Moreover, it contains several highly conserved domains, including the D-box, B-box, ETS binding and transactivation C domain (Saxton et al., 2016). Given ELK1's prominent role in the maintenance of pluripotency in hESC, I aimed to investigate the role of ELK1 in early axolotl development.

I first examined the gene expression profile of *Elk1* and other key members of the FGF/MAPK pathway during early development using an open-access transcriptome of axolotl early development (Jiang et al., 2017), which was mapped to a transcriptome published by our group (Evans et al., 2018). Gene expression data (Fig. 4.2) showed that *Elk1* mRNA transcripts were maternally inherited with expression peaking at stage 7 prior to Zygotic genome activation (ZGA); after this point, expression gradually declined until stage 14 following gastrulation. *Elk1* expression remained low from stage 14 onward throughout neurulation. Indeed, many other FGF/MAPK signalling pathway members are also expressed throughout gastrulation including *Ezh2*, *Mapk3(Erk1)*, *Mapk1(Erk2)*, *Fgfr1*, and *Suz12* suggesting they may function during this period of development. Some notable exceptions included *Fgf*'s 3,7,10 and 14, *Elk3* and *Elk4* suggesting which were only expressed post-gastrulation suggesting these genes may only function later in development. Interestingly, the expression profile of *Elk1* differs from that observed in

Xenopus, whereby maternal *Elk1* transcripts increase post-gastrulation (Nentwich et al., 2009).



Figure 4.1. Alignments of amino acid sequences of human, mouse, axolotl and xenopus ELK1 sequences.

Given that *Elk1* is expressed from ZGA and throughout gastrulation, I next explored whether ELK1, as demonstrated in hESC (Goke et al., 2013, Prise and Sharrocks, 2019), regulates pluripotency and differentiation in axolotl embryos. To do this, I employed an antisense-morpholino targeting the TSS of the *Elk1* gene (Figs 4.3-4.5). Embryos injected with 80ng of the ELK1 morpholino appeared phenotypically normal up until late neurula stages

where morphants appeared truncated (not shown). By tailbud stages (stage 28), ELK1 morphants showed distinct morphological differences when compared to uninjected sibling embryos (Fig. 4.3-4.4). Externally, ELK1 morphants displayed a severe truncation along the anterior, posterior axis, undefined pharyngeal arches, and no apparent tailbud. Morphants did form a neural tube and a head. HREM revealed that internally ELK1 KD embryos had formed a brain and brain ventricle as well as an undersized pharynx. Interestingly, intermediate mesodermal structures such as the meso/pronephric ducts were completely absent in KD embryos (Fig. 4.3). Furthermore, the notochord of morphants had failed to vacuolize. The somitic mesoderm, which is usually separated from the ventral-lateral plate mesoderm by the meso/pronephros, also extended across the ventral side of the embryos. This aberrant morphology was able to be rescued at high efficiency from 100ng of hELK1 RNA after co-injected with the MO at the 1-2 cell stage (Fig. 4.3 & 4.5a). This may be due to the highly conserved nature of the ELK1 sequence (Fig. 4.1) The severity of the ELK1 phenotype was reduced when only 60 or 40ng of MO was injected (Fig. 4.4& 4.5b); further to this, 160ng of ELK1 MO resulted in a similar phenotype as 80ng; however, most of the embryos died (Fig. 4.5c).

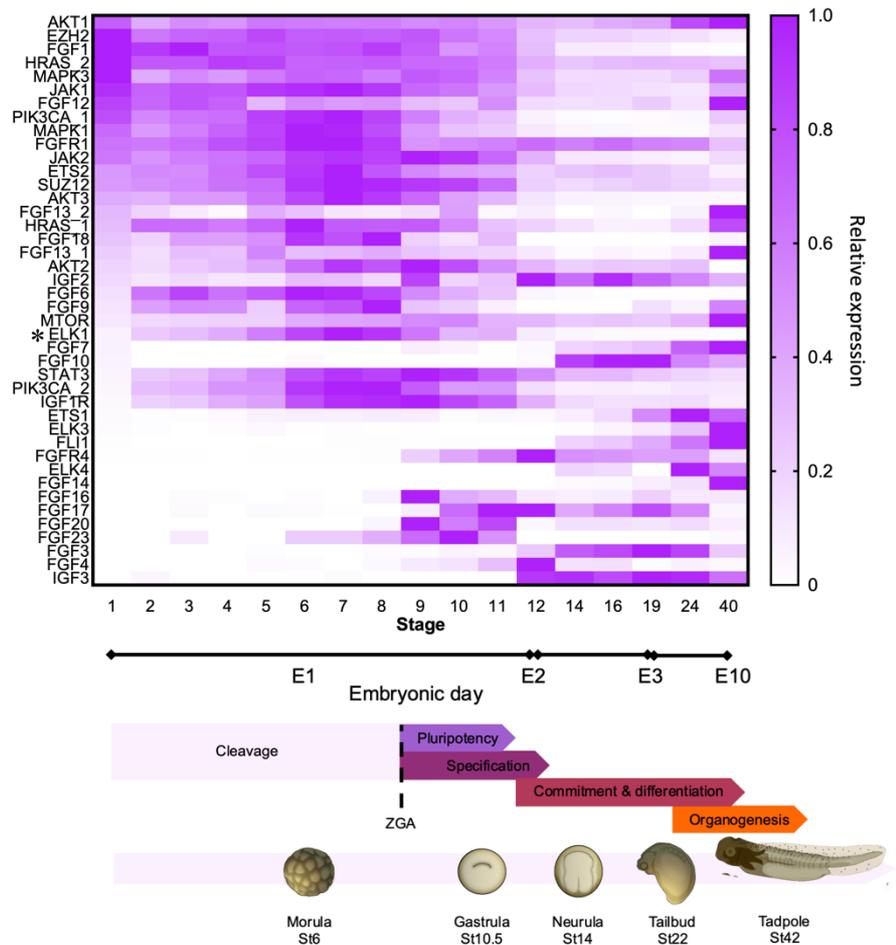


Figure 4.2. Overview of FGF signalling in early axolotl development.

Heatmap showing gene expression data for representative marker genes across 17 developmental stages. Raw data was taken from whole embryo transcriptomes produced by Jiang, et al (2017). Also shown is a schematic indicating the embryonic day in which each developmental stage occurs (When embryos are grown at 20°C), major milestones in cell differentiation and critical developmental time points: morula, gastrula, neurula tailbud and tadpole are labelled. Images adapted from the Axolotl Newsletter, Spring 1979 (Bordzilovskaya and Dettlaff, 1979).

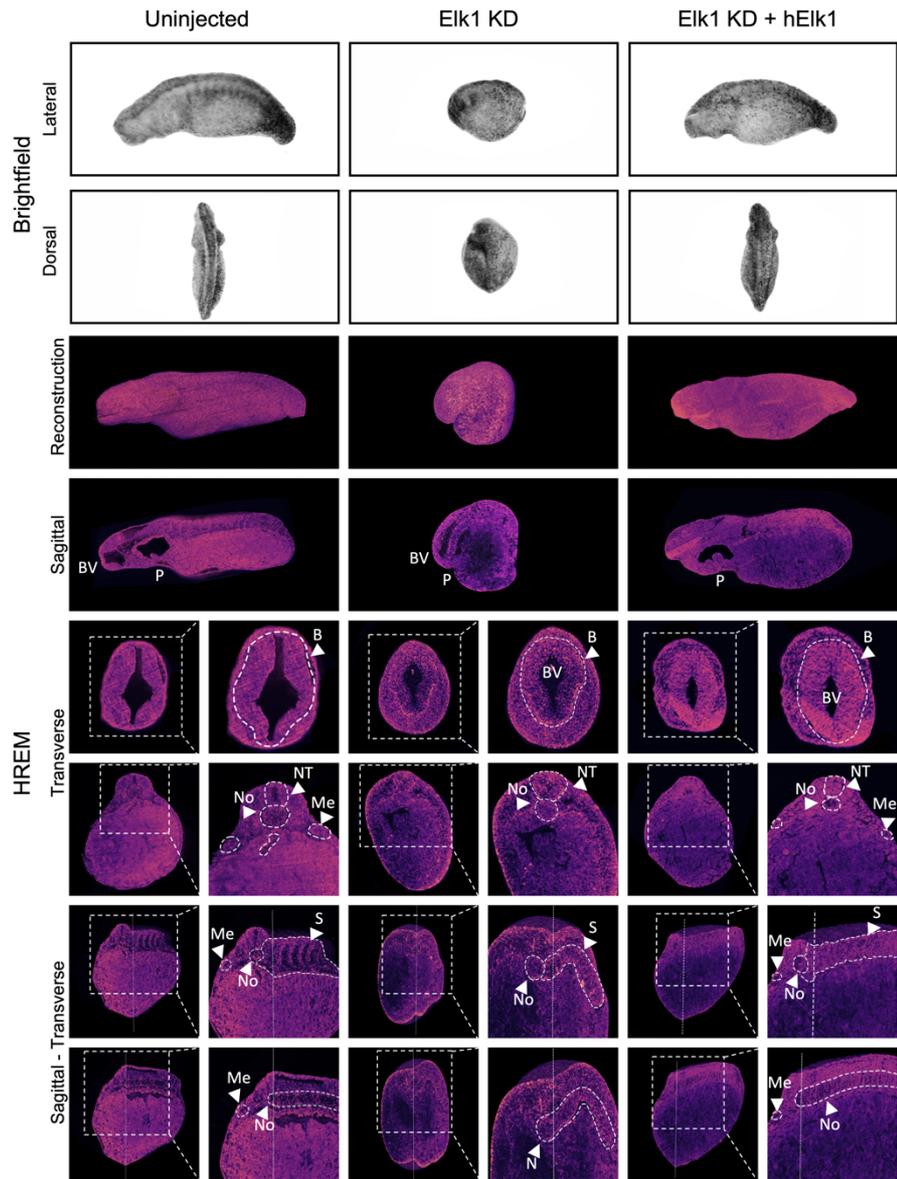


Figure 4.3. ELK1 depletion in axolotl embryos.

Brightfield images and HREM reconstructions of Uninjected, ELK1 depleted and hElk1 rescued embryos at equivalent stage 28. (HREM $n=2$, $m=1$, $t=6$). Brightfield images show the ventral view (top) and dorsal view of embryos. HREM images show a lateral view of 3D embryo reconstructions and section images: 3D reconstruction shows a lateral view. Transverse sections show the anterior (top) and posterior (bottom) views. Sagittal-transverse cross-sections are cropped at the mid-line (bottom) and medial-lateral (top) in the sagittal plane. Dashed lines highlight: brain (B), brain ventricle (BV), pharynx (P), somites (S), neural tube (NT) notochord (No), meso/pronephric ducts (Me). Scale bar, 1mm.

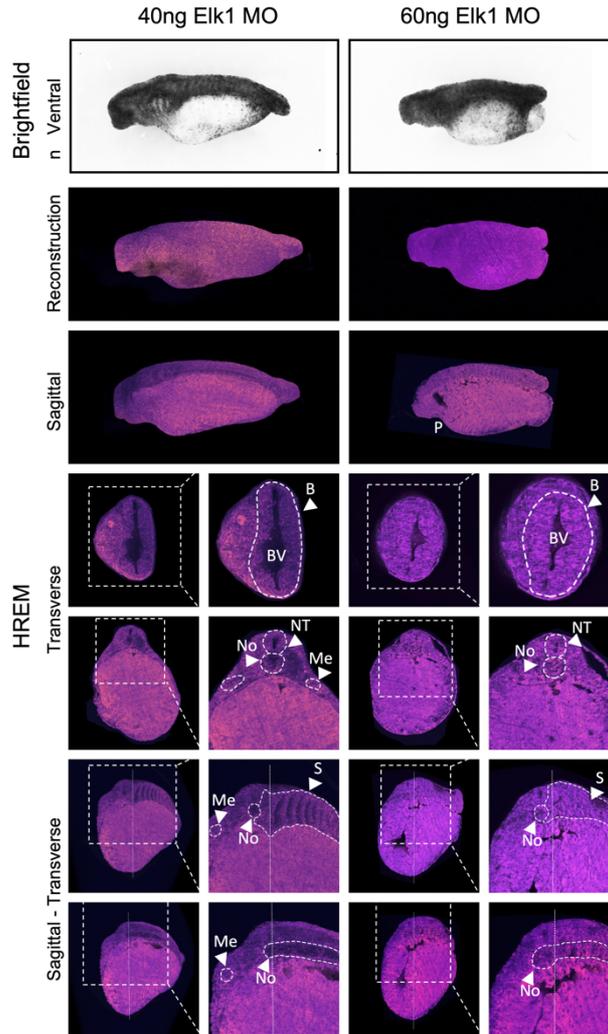


Figure 4.4. Titration of ELK1 morpholino.

Brightfield images and HREM reconstructions of embryos injected with 40ng and 60ng of ELK1 MO at equivalent stage 28. (n=1, m=1, t=2) Brightfield images show the ventral view (top) and dorsal view of embryos. HREM images show a lateral view of 3D embryo reconstructions and section images: 3D reconstruction shows a lateral view. Transverse sections show the anterior (top) and posterior (bottom) views. Sagittal-transverse cross-sections are cropped at the mid-line (bottom) and medial-lateral (top) in the sagittal plane. Dashed lines highlight: brain (B), brain ventricle (BV), somites (S), pharynx (P), neural tube (NT) notochord (No), meso/pronephric ducts (Me). Scale bar, 1mm.

Prior to the late neurula stages, ELK1 morphants appeared phenotypically normal. However, closer inspection showed gastrula stage morphants exhibited a wider, flatter, dorsal lip (Fig. 4.5c). Given that in axolotl, dorsal mesoderm involutes while lateral and ventral mesoderm ingresses, this may

suggest that reduced ELK1 levels results in an increased number of involuting cells moving through the blastopore during gastrulation (Shook et al., 2002, Shook and Keller, 2008, Kaneda and Motoki, 2012).

In axolotl, the first presumptive tissues to migrate inward during gastrulation are the precordal region followed by the presumptive notochord, somites, intermediate mesoderm, then presumptive ventral mesoderm, respectively (Shook et al., 2002, Shook and Keller, 2008, Kaneda and Motoki, 2012). This is due to the dorsal location of the blastopore, which extends lateral-ventrally throughout gastrulation. My observations suggested that ELK1 KD results in sections of the ventral mesoderm being re-specified to somites. As intermediate mesodermal structures are also absent, this may suggest ELK1 is required to repress dorsal mesodermal development to establish intermediate and ventral mesodermal tissues, which incidentally are specified later than the precordal region, notochord or somites. Alternatively, previous published work by our lab has shown that meso/pronephros and PGCs' formation depends on FGF signalling (Chatfield et al., 2014). Given that ELK1 has been shown to act downstream of FGF signalling *in vitro* and *in vivo* (Yang et al., 1998a, Yang et al., 1998b, Nentwich et al., 2009, Goke et al., 2013, Prise and Sharrocks, 2019), this may suggest that ELK1 acts as a downstream effector of FGF to establish meso/pronephros and PGCs. These hypotheses, however, are not mutually exclusive. Strikingly, these observations more closely resemble ELK1 depletion in hESCs, which results in a hypersensitivity to mesodermal differentiation (Goke et al., 2013, Prise and Sharrocks, 2019).

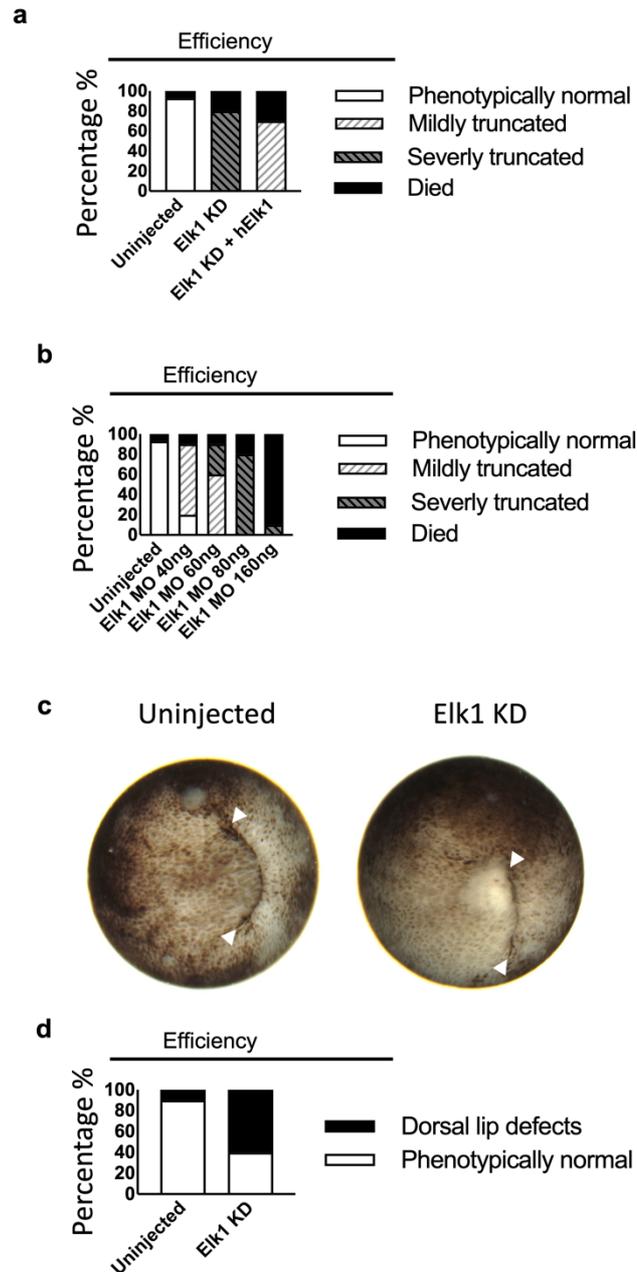


Figure 4.5. Efficiencies of KD and characterisation of ELK1 KD phenotype.

a, Numbers of embryos which presented with truncations following a single injection of 80ng ELK1 MO, co-injection of 100pg hELK1 and 80ng ELK1 MO or no injection ($n=25$, $m=1$, $t=75$).

b, Phenotypic characterisation of embryos following injection with different amounts of ELK1 MO. Embryos were either Uninjected or injected with 40, 60, 80 or 160ng of ELK1 MO ($n=25$, $m=1$, $t=100$).

c, Brightfield images of Uninjected and ELK1 KD embryos at stage 10.5. White arrows indicate the limits of the blastopore.

d, Number of embryos either Uninjected or ELK1 depleted which presented with dorsal lip defects ($n=18$, $m=1$, $t=36$).

To better characterise the ELK1 phenotype, I looked at markers of germ layer specification and commitment (Fig 4.6). Morphant embryos showed differential gene expression in several key specification genes around the gastrula stage (Fig 4.6a). Notably, ELK1 morphants exhibit upregulation of *NODAL1* and *NODAL2* by around 1.7 and 3.5-fold, respectively. *BRACHYURY* and *SOX17* were upregulated by 1.7 and 2.3-fold, respectively. In contrast *GATA2* expression was heavily depleted, present at around 0.2-fold lower than uninjected embryos. Together these results suggest that ELK1 may repress essential germ layer specification genes.

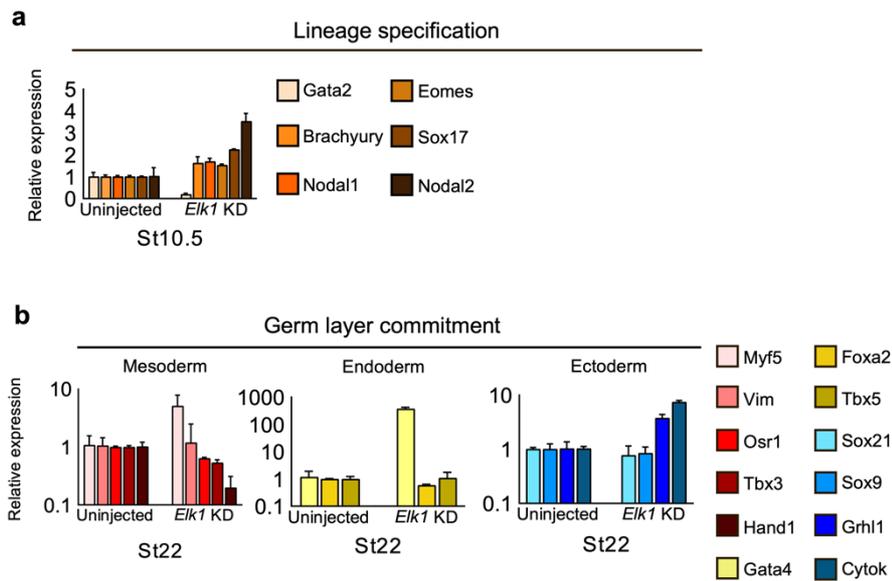


Figure 4.6. Germ-layer marker expression in ELK1 depleted embryos.

QPCRs showing the gene expression of markers of cell fate decisions in uninjected and ELK1 depleted embryos **a**, Expression of germ layer specification markers at stage 10.5 (n=10, m=3, t=60). **b**, Germ-layer commitment gene expression at stage 22. (n=10, m=3, t=60).

While ELK1 depletion enhanced transcription of key specification genes, the downstream consequences of these changes remained uncertain. Moreover, the morphology of ELK1 morphants was somewhat ambiguous, particularly as to whether the ventral lateral plate mesoderm is thickened or whether the somitic mesoderm had extended to the ventral side of the embryo. To this end,

I analysed the expression of key tissue commitment and differentiation markers at late neurula (Stage 22) stages via qPCR (Fig. 4.6b). Interestingly, qPCR analysis showed increased expression of somitic mesoderm marker *MYF5* around 5-fold, which may suggest increased somite formation as indicated by HREM. Intermediate mesoderm and meso/pronephros marker *OSR1* was around 0.6 fold lower than uninjected embryos. Intermediate marker *VIMENTIN* showed a moderate upregulation, although it is worth noting that *VIMENTIN* is also expressed in hindbrain/neural tissues which seem largely unaffected in the ELK1 morphants. Ventral lateral plate marker *HAND1* and notochord marker *TBX3* were downregulated to around 0.2 and 0.5-fold, respectively, which may suggest impairments in the formation of these tissues, both of which appeared abnormal in HREM imaging. Compared to uninjected sibling embryos, endoderm marker *GATA4* showed massive upregulation around 370-fold higher than controls. The hindgut marker *TBX5* was unaffected, while foregut marker *FOXA2* was downregulated to around 0.6-fold compared to Uninjected siblings. Moreover, endodermal structures do not show overrepresentation in ELK1 morphants; thus, it is unclear as to the effects of *GATA4* upregulation. Interestingly, Neural ectoderm and epidermis markers *SOX9* and *SOX21* showed only moderate downregulation, while epidermal markers *GRHL1* and *CYTOK* were upregulated around 3.7 and 7.5-fold respectively, suggesting epidermal formation is also repressed by ELK1.

Previous work by our group, as well as work presented in the last chapter, has demonstrated conservation in the GRN governing pluripotency and mesoderm specification between axolotl and mammals (Dixon et al., 2010, Swiers et al., 2010) as it has been suggested that ELK1 may synergise with

ERK2 to promote the expression of pluripotency factors in hESC and repress mesodermal differentiation (Goke et al., 2013, Prise and Sharrocks, 2019). Given that germ layer specification and commitment were affected by ELK1 depletion, I investigated whether this was due to premature loss of pluripotency. Therefore, I assayed the expression of pluripotency genes *NANOG*, *POU5F1* and *SOX2* following ELK1 KD during the gastrula stages of development using qPCR (Fig. 4.7). qPCR analysis showed no notable differences in expression of *POU5F1* and *SOX2* expression at early, mid or late gastrula stages (Stage 10, 11 and 12, respectively) in response to ELK1 KD. Together, this suggests that pluripotency gene expression is not dependent on ELK1, nor is the window of pluripotency shortened due to ELK1 depletion in contrast to the paradigm proposed by Goke et al. in hESC (Goke et al., 2013).

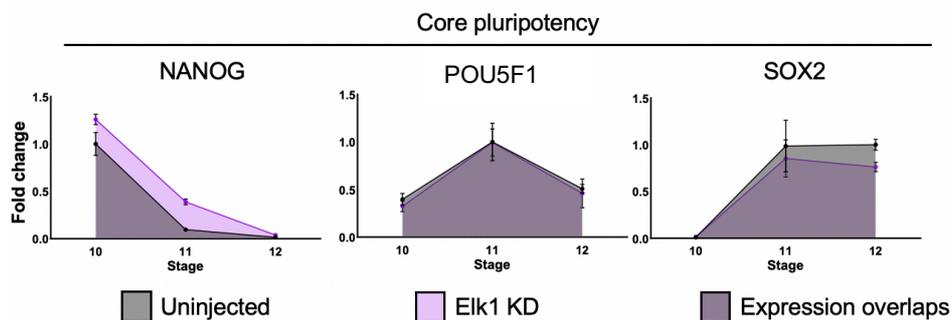


Figure 4.7. Elk1 KD does not affect the expression of pluripotency factors.

Expression of core pluripotency genes *Nanog*, *Pou5f1* and *Sox2* across 3 developmental stages with and without Elk1 KD (n=10, m=3, t=60).

ELK1 is known to have a role in repressing differentiation genes in hESCs. Therefore, I next performed whole RNA-seq on mid-gastrula (stage 10.5) explanted animal caps (Fig 4.8). Caps at this stage can be easily explanted and represent a pluripotent cell population. RNA seq confirmed that *ELK1* was expressed in the AC explants along with several other members of the MAPK/FGF family (Fig 4.6a). Moreover, several pluripotency factors,

including *NANOG*, *POU5F1*, *PRDM14*, *KLF2*, *OTX2*, *LIN28A* and *TFCP2L1*, were unaffected by ELK1 depletion (Fig. 4.8a), confirming our observations at the whole embryo level via QPCR. I next looked for differentially expressed genes using the DESEQ2 pipeline. Statistical analysis revealed that thirty-nine genes showed significantly (<0.05) different expression following, ELK1 depletion (Fig. 4.8.B). Importantly, however, only eight genes were significantly affected following p-value adjustment (Fig 4.9). Given this small number of DEGs, this suggests that ELK1 does play a significant role in regulating the pluripotent presumptive ectoderm at this stage. However, it is important to note that the presumptive ectoderm is not the only pluripotent tissue. The entire animal hemisphere, including the marginal zone (the presumptive mesoderm) of the axolotl embryo, expresses pluripotency genes *NANOG* and *POU5F1* (Dixon et al., 2010). However, while this can be dissected, it is difficult to do so without also taking underlying vegetal cells which are not pluripotent. Given that I have no data to suggest where *Elk1* is expressed, it is possible that ELK1 may only regulate the marginal zone/presumptive mesoderm. Indeed, *ELK1* is predominantly expressed in the marginal zone in *Xenopus* (Nentwich et al., 2009).

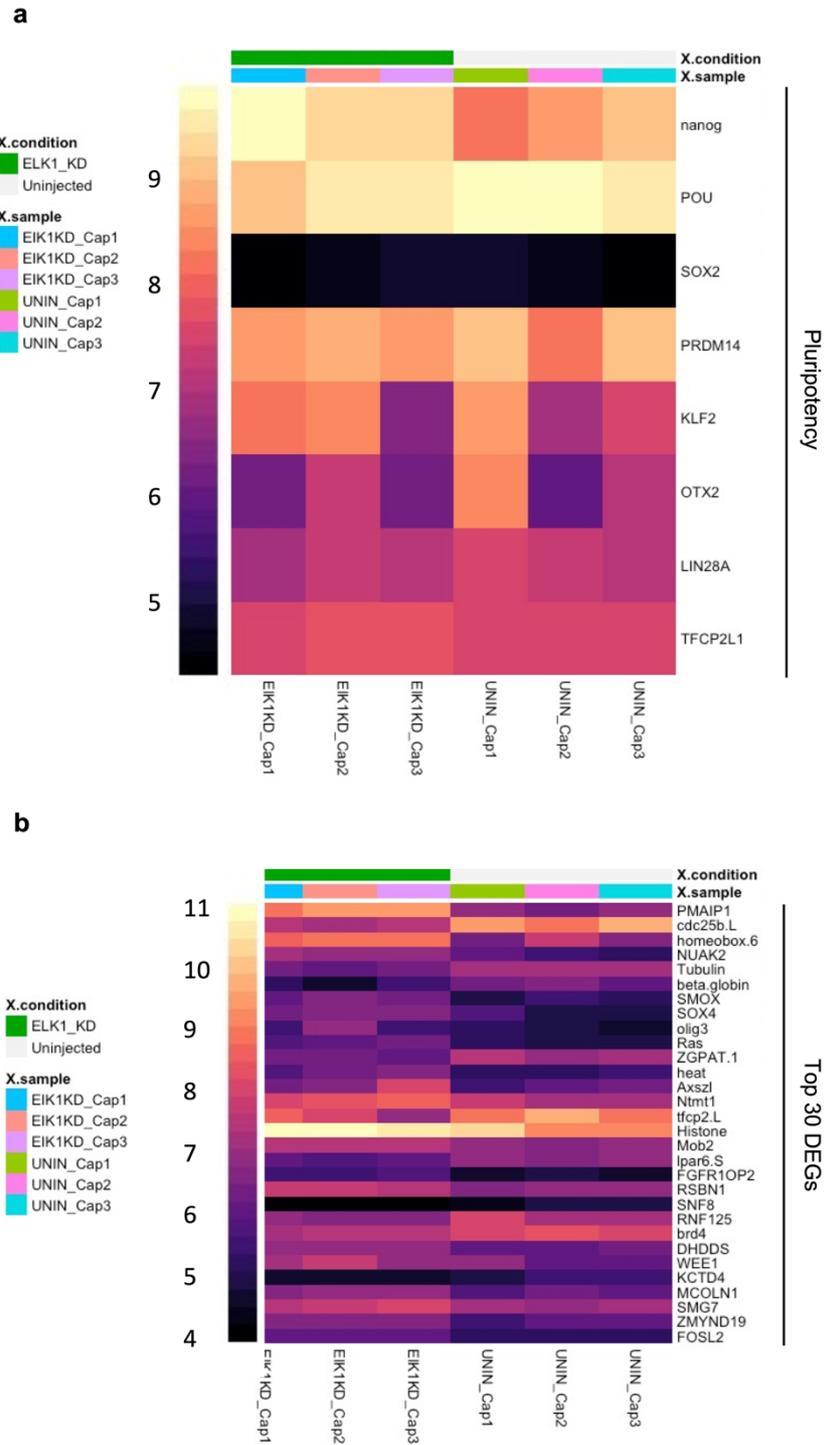


Figure 4.8. Expression of key genes in ELK1 depleted ACs.

a, Heatmap showing pluripotency factor expression in uninjected and ELK1 KD AC. **b**, Heatmap showing the top 30 DEG's expression in uninjected and ELK1 KD AC. (n=15, m=3, t=90).

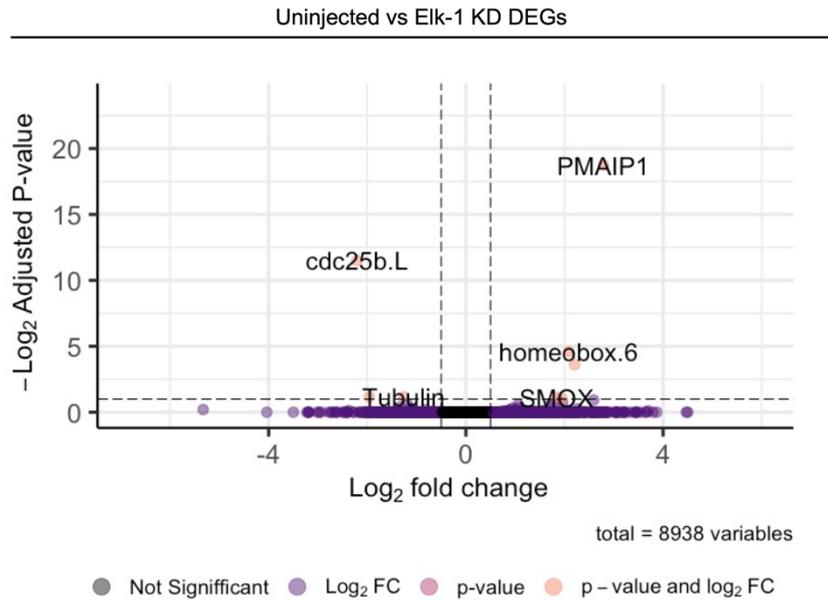


Figure 4.8. Differential gene expression in ELK1 depleted ACs.

Volcano plot showing significant DEGs following p-value correction in uninjected and ELK1 KD AC.

Given that it appeared that ELK1 doesn't regulate the pluripotent cells of the presumptive ectoderm (equivalent to the epiblast in humans) and that hESC are derived from the late epiblast, I decided to revisit the gene expression data produced by Goke et al. (Goke et al., 2013) following ELK1 depletion. It was suggested that ELK1-ERK2 bound loci were positively regulated while ELK1 only bound loci were repressed. Indeed, the publication included the processed ChIP-seq data, which showed that ELK1-ERK2 bound loci included pluripotency factors NANOG, POU5F1, KLF2 and KLF4. ChIP data also showed that among loci bound only by ELK1 were classic differentiation markers including TBXT (BRACHYURY), NODAL, HAND1, VIMENTIN, and members of the hox family. However, the processed microarray data following ELK1 KD was not made available, nor were individual genes named rather a fisher's exact test was performed to look at the correlation with 'pluripotency genes' and 'differentiation genes', and indeed, this seemed to

support the overall hypothesis (Goke et al., 2013). In order to investigate whether the regulation of specific target genes is conserved between axolotl and humans, I attempted to create a list of genes differentially expressed following ELK1 depletion in hESC (Fig 4.10).

Interestingly, many of the targets of ELK1-ERK2, which would be expected to decrease following ELK1 depletion, including NANOG, POU5F1 and KLF4, showed no significant differences in expression 4 days after ELK1 depletion. This also seemed to be the case with other ELK1 bound markers of differentiation such as NODAL or HAND1 which would be predicted to increase in expression, showed no difference following ELK1 KD. However, as predicted by the Goke model, VIMENTIN increased following ELK1 depletion (Goke et al., 2013). This does somewhat contradict the conclusions of the paper and may suggest a problem with the data. A more recent study suggested that ELK1 does not directly regulate pluripotency but instead represses differentiation, specifically mesodermal differentiation (Prise and Sharrocks, 2019). In their methodology, they also employ hESC but use directed differentiation. Given that I found upregulated members of the NODAL family and later increased somitic mesoderm markers, I investigated whether ELK1 depletion causes hypersensitivity to mesoderm induction using animal cap assays (Fig. 4.11-4.12).

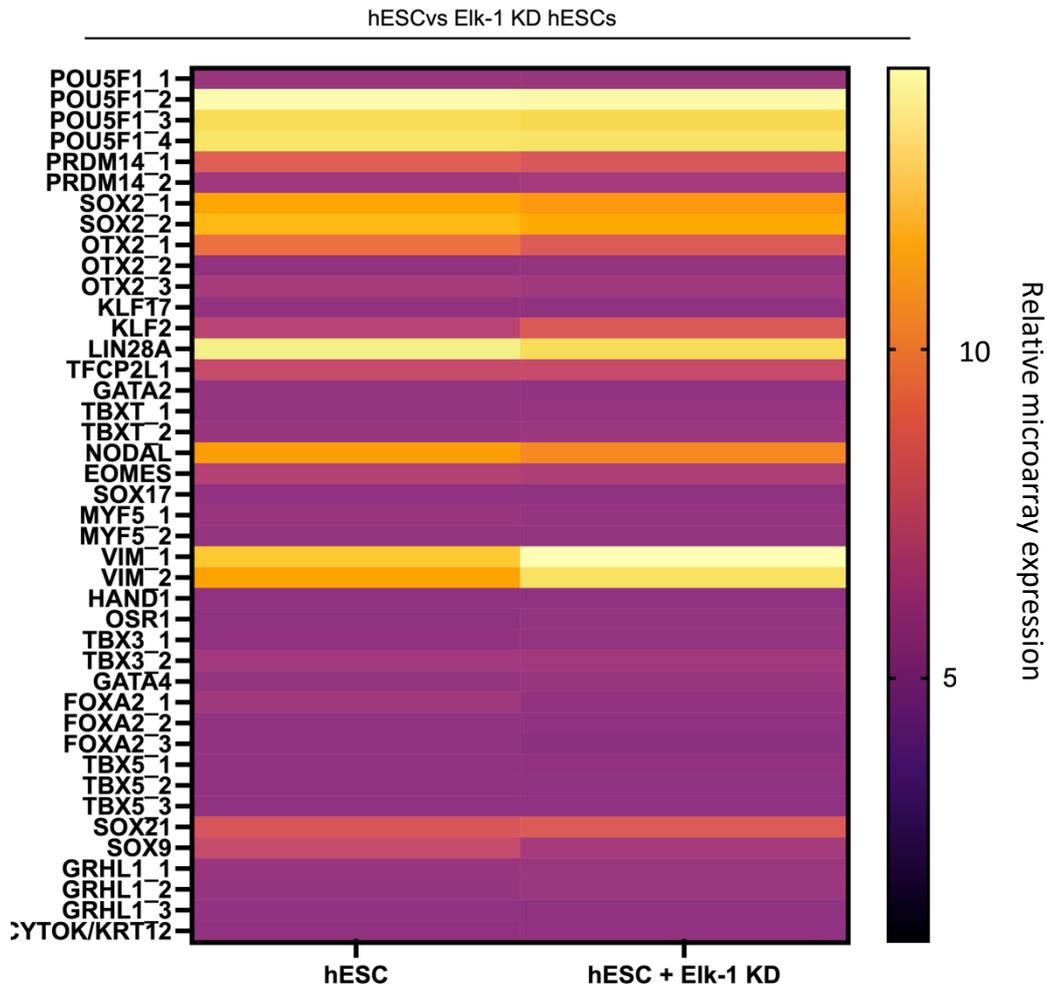


Figure 4.9. Key marker gene expression in hESCs 4 days following ELK1 depletion
Heatmap showing expression of key cell-type markers in uninjected and Elk1 KD hESC (Microarray data from Goke et al., 2013).

Dissected ACs naturally differentiate toward epidermis. To explore the effect of spurious gene activation driven by ELK1 KD on cap differentiation, I cultured the caps *ex vivo* to equivalent stage 30. I assayed them for germ layer commitment markers (Fig. 4.11). Remarkably, epidermal gene markers *CYTOK* and *GRHL1* were highly down-regulated, while somatic mesodermal marker *MYF5* was upregulated around 14-fold. Definitive endodermal marker *GATA4* was also upregulated 20-fold, and neural ectodermal markers *PAX6* and *NEUROG* were upregulated 4 and 3-fold, respectively, compared to uninjected caps of equivalent stage. In axolotl, only the ventral marginal

zone contains cells that produce mesoderm, endoderm and neural ectoderm; this suggests that alleviating ELK1 mediated gene repression is sufficient to alter the natural differentiation potential of the cells of the animal hemisphere. Therefore, it appears ELK1's repressive activity is conserved between axolotl and humans; however, in the axolotl, it appears that this affects the differentiation potential to multiple germ layers. Moreover, this implies ELK1's primary function during gastrulation may be of gene repression, which is required to facilitate temporal differentiation events key to embryogenesis.

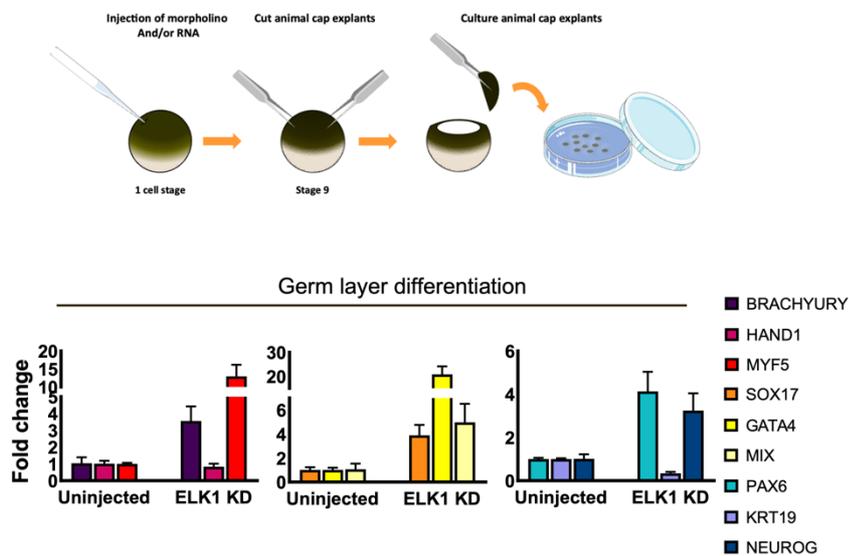


Figure 4.10. Differentiation of late stage ACs following ELK1 depletion.

Schematic of animal cap assay and qPCRs showing the gene expression of markers germ layer commitment in uninjected and ELK1 depleted AC explants cultured to stage 20. (n=15, m=3, t=90).

Given that our whole embryo data suggests the ELK1 morphants lack intermediate and ventral lateral plate mesodermal structures at the expense of somitic mesoderm, I next tested whether ELK1 KD exerted specific effects in response to two potent differentiation inducers via the FGF and ACTIVIN signalling pathways. To do this, I performed an inductive animal cap assay

whereby caps were induced to form mesoderm with and without ELK1 KD (Fig. 4.12). Our group has previously demonstrated that 200fg of ACTIVIN is sufficient to induce cap elongation – indicative of mesodermal induction (Swiers et al., 2010) and that 1pg of ACTIVIN or higher induces endoderm at the expense of mesoderm indicated by a rounder, paler appearance, reduced *BRACHYURY* expression and elevated *SOX17* expression. I found that ELK1 KD caps did produce higher *NODAL1* gene expression in response to 200fg of ACTIVIN and expressed nearly twice the level of *MYF5*, suggesting ELK1 depletion hyper-sensitises AC cells to activin. ELK1 KD caps also showed greater expression of *SOX17* in response to 4pg of activin than sibling caps without ELK1 KD. Unexpectedly, however, ELK1 KD caps injected with 4pg activin circumvented loss of *MYF5*, maintaining high expression suggestive of mesodermal and endodermal induction, by contrast, to control caps which showed evidence only of endodermal induction. Accordingly, ELK1 depletion also resulted in upregulation of *MYF5* in response to 60pg FGF and showed upregulation of neural differentiation marker *NEUROG*. These observations support a paradigm where ELK1 depletion hyper-sensitises cells to neural and mesendodermal differentiation in response to FGF and activin, respectively. It remains unclear whether the presence of both mesodermal and endodermal markers upon endodermal induction suggests a cellular heterogeneity within the ELK1 KD caps or whether an inability to repress lineage-specific gene expression drives the formation of erroneous cell types that co-express genes of mutually exclusive lineages.

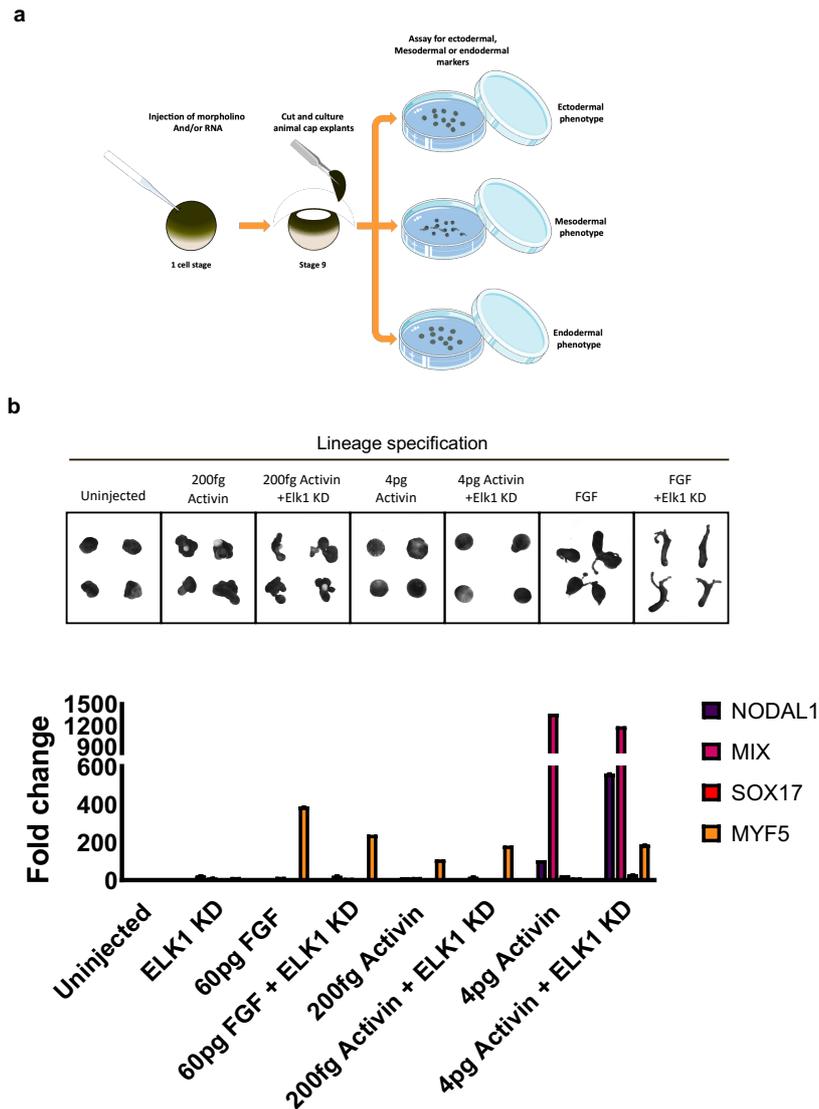


Figure 4.11. Mesoderm induction in ELK1 depleted ACs

a, Schematic of mesoderm induction animal cap assay. **b**, Brightfield images and qPCR of mesodermal, neuroectodermal and endodermal markers in uninjected and ELK1 depleted stage 22 caps following injection with 60pg FGF, 200fg of Activin or 4pg Activin at the 1-cell stage. (n=15, m=3, t=360).

ELK1, like most transcription factors, has several functional domains key to its functionality. While ELK1 orthologues are conserved throughout metazoans, its functional domains are conserved to differing extents (Saxton et al., 2016). While these domains have been well studied in various cellular contexts, little is known as to which domains are required for ELK1's function in early development. Because of this, I next decided to test the ability of

human ELK1 variants with mutations in specific functional domains to rescue ELK1 morphant development. The first variant tested was hELK1, which has a mutation within the D domain (Fig. 4.13). The D (or DEJL) domain is responsible for binding ELK1 to activated MAP kinases of the ERK, JNK, and p38 subtypes. The DEF domain within this region is specifically required for the recruitment of activated ERK (Zhang et al., 2008, Jacobs et al., 1999, Besnard et al., 2011). Critically, ELK1 phosphorylation by ERK increases ELK1 nuclear translocation (Lavaur et al., 2007). As such, I tested a hELK1 with a mutated D domain (AA307–428 $-/-$, Herein, D mutant) which is unable to recruit activated ERK. HELK1 D mutant-rescued morphants appeared to develop normally up to neurulation; however, as development continued, it became apparent that morphants lacked the ventral posterior region entirely. HREM confirmed that in addition to a reduction to the posterior ventral area, D mutant-rescued morphants showed reductions in the meso/pronephric ducts (Fig. 4.13). As axolotl PGCs have a ventral marginal zone origin and develop within the LPM before migrating dorso-ventrally to reside just lateral to the meso/pronephros, this suggests that the ELK1 likely acts downstream of FGF signalling through interactions with ERK; moreover, this interaction is critical to establishing the AGP region and PGCs (Johnson et al., 2001, Bachvarova et al., 2004, Bachvarova et al., 2009b, Bachvarova et al., 2009a, Chatfield et al., 2014). This observation is consistent with FGF's prominent role along with BMP signalling in establishing the PGC-AGP region (Chatfield et al., 2014). The C (or transactivation) domain contains the amino acids that are phosphorylated by MAP kinases. Among these residues, phosphorylation of Serine 383 is a crucial event to activate the ELK1-mediated transcription (Yang et al., 2002, Besnard et al., 2011). As such, I next tested a mutant in which the

crucial 383 serine residue had been changed to an alanine (S383A, Herein C mutant), rendering ELK1's ability to be phosphorylated inert. Outwardly, C mutant rescued morphants appeared relatively normal apart from a moderately reduced ventral posterior region, and their tails show an upward inflexion compared to uninjected controls. More to this there appeared be no difference to their internal morphology as analysed by HREM.

The ELK1 SM domain is involved in binding to the known ELK1 co-factor SRF. Thus, I also tested the ability of an SM mutant hELK1 (L158P/Y159A, Herein SM mutant) to rescue ELK1 KD embryos. This mutant cannot bind to SRF but has an intact ETS domain allowing DNA binding to the SRE. Intriguingly, the morphants closely resembled the C mutant rescued morphants, also presenting with a slightly reduced ventral posterior and upward tail inflexion. HREM showed that SM mutant rescues showed slightly enlarged meso/pronephric ducts in contrast to the D mutant rescued morphants.

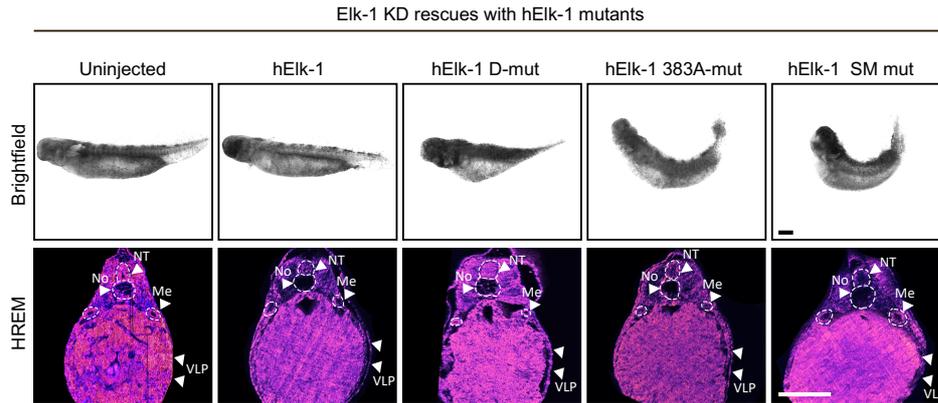


Figure 4.12. Rescue of ELK1 depletion with mutant hELK1 mRNA.

a, Brightfield images and HREM reconstructions of uninjected and hELK1 rescued embryos ($n=15$, $m=2$, $t=150$ and HREM $n=1$, $m=1$, $t=5$). Rescues were performed with WT hELK1, hELK1 D mut, hELK1 383A mut or hELK1 SM mut. Brightfield images show the ventral view (top). HREM images show a reconstructed transverse sections of the anterior of embryos. Dashed lines highlight: brain (B), brain ventricle (BV), somites (S), pharynx (P), neural tube (NT) notochord (No), meso/pronephric ducts (Me). Scale bar, 1mm.

Given that ELK1 morphants showed defects in intermediate mesoderm and VLP structures and that in axolotl PGCs develop within the posterior lateral plate mesoderm and migrate dorsomedial where they reside just ventral to the intermediate mesoderm (Johnson et al., 2001, Bachvarova et al., 2004, Bachvarova et al., 2009b, Bachvarova et al., 2009a, Chatfield et al., 2014). Moreover, the ACs of axolotl embryos can efficiently induce PGCs as well as some other cell types of the AGP, such as meso/pronephros, in response to injection of mRNA encoding FGF4 and BMP4 (Chatfield et al., 2014). As ELK1 appears to act downstream of FGF, I next investigated if ELK1 has a function in the formation of PGC's. I first sought to define our *ex vivo* PGC induction better by inducing PGC like cells *ex vivo* and collecting and assaying cap gene expression at five developmental time points (Fig. 4.15). My results show a clear progression of gene expression throughout the time course. Prior to gastrulation, ACs express the pluripotency markers *POU5F1* and *NANOG*

(Dixon et al., 2010). Interestingly, in *FGF* and *BMP* injected caps, *NANOG* and *POU5F1* expression is detectable in stage 15 and 19 caps, equivalent to late neurula stages; this may suggest that the expression of *NANOG* and *POU5F1* in early PGCs may be conserved from early tetrapods as the expression of these factors in early PGC's have been reported in both mammals and chick (Choi et al., 2021, Choi et al., 2018, Laval et al., 2007, Han et al., 2018, Yamaguchi et al., 2005, Murakami et al., 2016, Aeckerle et al., 2015, Kobayashi et al., 2017). Early programmed caps expressed *BRACHYURY*, consistent with the idea that PGC specification is brachyury dependent (Chatfield et al., 2014). *SOX17* expression also showed increased expression in stage 15 *FGF* and *BMP* caps, which resembles findings in mice and humans. At stage 15, early germ cell markers *NANOS* and *DEADEND* are also upregulated compared to uninjected control caps. Interestingly, stage 19 marks the peak of *NANOS* and *DEADEND* expression; this stage also marks upregulation of the LPM marker: *WNT8* low-level expression of the definitive PGC marker *Dazl* is also exhibited. Suggesting that by stage 19, PGC fate has commenced. Unpublished data gathered by myself and Darren Crowley has recently demonstrated that the transcription factor *VENTX* is required for PGC development; therefore, I also looked at *VENTX* expression. *VENTX* also shows increased expression compared to Uninjected caps at stage 19. By stage 26, Pluripotency gene expression is undetectable, and early PGC markers *NANOS* and *DEADEND* show a gradual decline in expression whilst mature PGC marker *DAZL* increases in expression. Stage 26 also marked the peak of *WNT8* and *VENTX*. Subsequently, stage 35 marked the decline of *WNT8* and *VENTX* and increased definitive PGC marker *DAZL*. By stage 42, early PGC markers, intermediate and LPM markers, are only lowly expressed while *DAZL*

expression peaked, suggesting completion of PGC fate. Together these data suggest three distinct phases in PGC development: Phase 1 where ‘pluripotent’ germ-line competent mesodermal precursors emerge post gastrulation. Phase 2: The emergence of early PGC and somatic mesodermal precursors and Phase 3, where definitive PGCs and somatic mesoderm are formed.

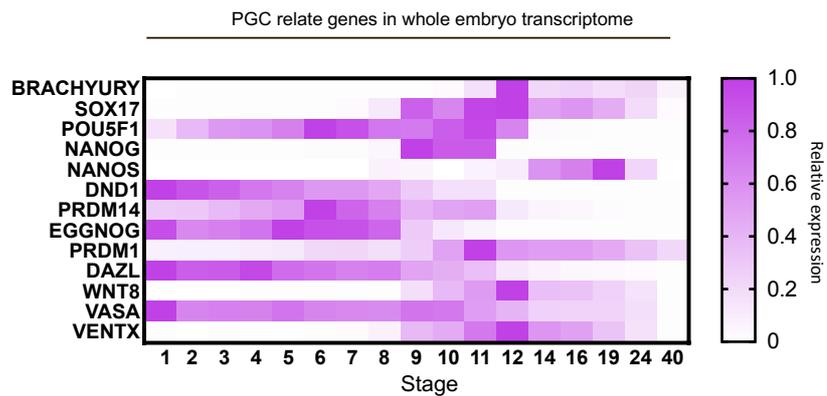


Figure 4.13. Expression of PGC-associated genes in axolotl embryos. Heatmap showing gene expression data for key genes involved in PGC development in embryos across different developmental time points. Raw data was taken from whole embryo transcriptomes produced by from Jiang, et al (2017).

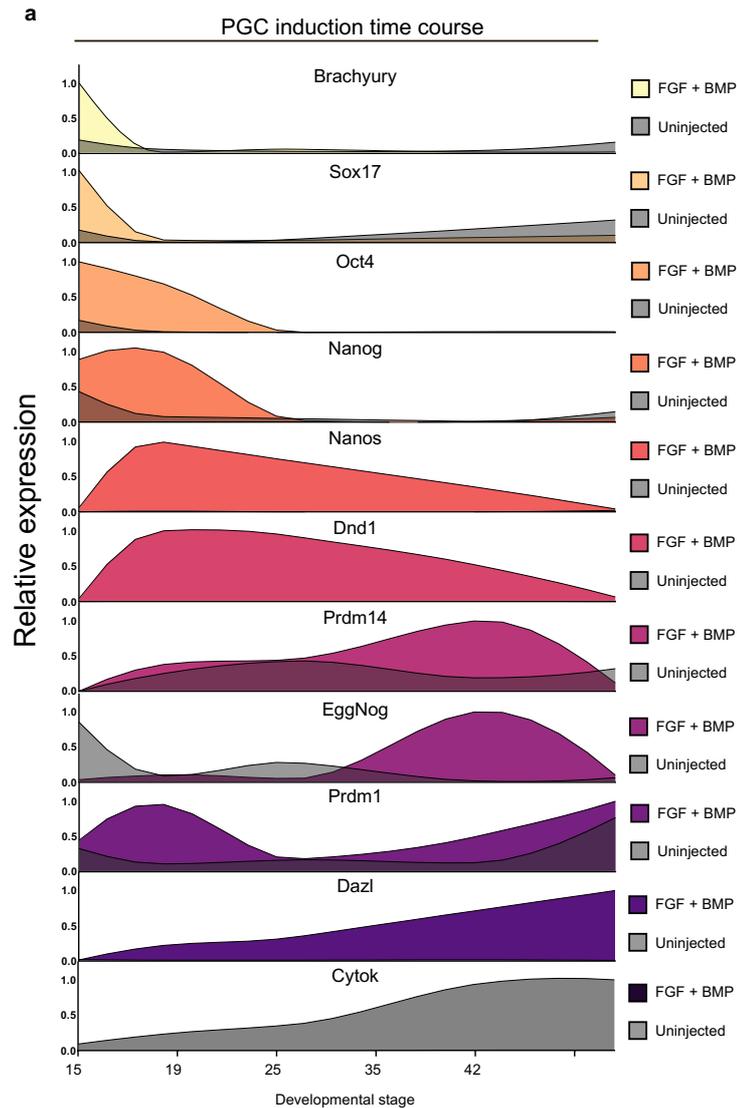


Figure 4.14. PGC related gene expression in PGC-induced ACs.

Expression of key genes involved in PGC development at different developmental stages in explanted animal caps, either Uninjected or injected with FGF & BMP. Gene expression was measured via qPCR. (n=15, m=3, t=225).

After defining our *ex vivo* PGC induction, I next sought to understand better ELK1's role in PGC development using our *ex vivo* PGC induction on ELK1 depleted ACs and analysing the expression of key genes using QPCR (Fig.4.16). As expected, co-injection of ELK1 morpholino severely reduced the expression of early PGC markers *DEADEND* and *NANOS* as well as intermediate and lateral plate markers *OSR1* and *WNT8* respectively at stage 20 in response to *FGF* and *BMP* than caps without ELK1 depletion. ELK1 KD

caps also showed downregulation of definitive PGC markers *DAZL*, *VASA* and *PIWI* and definitive meso/pronephros marker *PAX2* at stage 35. Moreover, the expression of all these markers can be rescued with co-injection of hELK1 mRNA alongside the ELK1 MO. HELK1 variants with D and C domain mutations were unable to rescue definitive PGC development at stage 35. Still, only the D mutant HELK1 could not rescue early PGC and LPM markers. By contrast, morphant caps rescued with C and SM mutant hELK1 were able to rescue early PGC/LPM mesodermal markers at stages 20. Interestingly, rescues with SM mutant hELK1 increased PGC/LPM markers above that of hELK1. Together with our whole embryo observations, this suggests that the interaction of ELK1 and ERK is critical to forming the PGC competent ‘pluripotent’ mesoderm. Still, the phosphorylation of the transactivation domain is required to produce definitive PGCs from early, while ELK1-SRF interactions are not necessary for either, given the enlargement of the meso/pronephros region and increased PGC marker expression, SRF-ELK1 interactions may favour non-PGC/LPM targets.

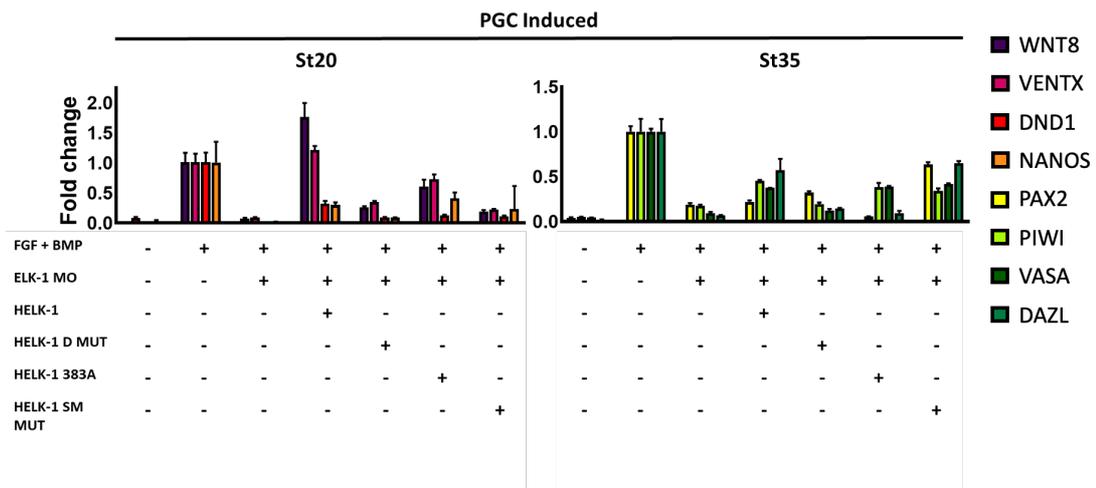


Figure 4.15. Expression of PGC genes in ELK1 depleted and hELK1 rescued ACs. Gene expression (qPCR) of uninjected and PGC induced animal caps with or without ELK1 KD and co-injection of hELK1 mutant variants. (n=15, m=3, t= 270).

WNT8 gene expression tracks with the development of the LPM from gastrulation through tailbud stages, the tissue which gives rise to PGCs mesoderm (Johnson et al., 2001, Bachvarova et al., 2004, Bachvarova et al., 2009b, Bachvarova et al., 2009a, Chatfield et al., 2014). As such, *WNT8* was highly upregulated at stages 19-26 in my *ex vivo* PGC induction time course.

The product of the *WNT8* gene signals through the canonical *WNT* pathway and is expressed in the posterior-lateral mesoderm in the embryos of all vertebrates examined (Nakamura et al., 2016). Accordingly, *WNT8* is also up-regulated following transitioning to media containing GS3K inhibitor:

CHIR99021, MEK inhibitor: PD0325901, TGF β inhibitor: SB203580 and JNK inhibitor: SP600125 (4i conditions), which produce PGC competent cells from hESC (Irie et al., 2015). Moreover, several *WNT* gene promoters are bound by ELK1 in hESC (Goke et al., 2013). Given that our data suggest that *WNT8* expression is dependent on ELK1/ERK, this indicated that there might be a conserved role of *WNT* signalling, which conveys competency for PGC/LPM development downstream of ELK1. Preliminary data gathered by ADJ (unpublished) used antisense morpholinos to KD *WNT8* protein expression in axolotl embryos. Interestingly, injection of the MO phenocopied the effects of ELK1 morphant rescue with hELK1 D mutant mRNA. This phenotype also resembles *Wnt8* knockdown in *Xenopus* (Nakamura et al., 2016b). Importantly, *Xenopus* PGCs are specified by germplasm, so *WNT8* is not involved in PGC specification. But in embryos, *WNT8* knockdown depleted the entire aorta-gonad meso/pronephros (AGP) region, including the PGCs, similar to the effects of *Wnt8* depletion in mice (Medvinsky and Dzierzak, 1996). This result is consistent with our previous observations. *WNT8* depletion also negatively impacts PGC marker expression much like ELK1

depletion, QPCR of ACs induced with *FGF* and *BMP* containing the *WNT8* morpholino demonstrated downregulation of both early PGC markers *DEADEND* and *NANOS* at stage 20. Interestingly, I found that expression of *VENTX* was also downregulated, suggesting *VENTX* may act downstream of *WNT* signalling in PGC development. After the same induction at stage 35, *WNT8* depletion downregulated PGC-specific markers *DAZL*, *VASA* and *PIWI*. They also inhibited the expression of *FLK1* and *PAX2*. Together, this suggests that *Wnt8* expression is required for induction of the mesoderm that gives rise to PGCs and the AGP and that its transcription is dependent on ELK1-ERK signalling downstream of FGF.

The human Mediator comprises 26 subunits forming three modules termed Head, Middle and Tail. *MED23* belongs to the Tail module of Mediator, whose primary function is to connect Mediator to sequence-specific transcription factors (Borggreffe (Borggreffe and Yue, 2011, Monte et al., 2018). *In vitro* and *in vivo* studies have demonstrated a specific interaction between *Med23* and *Elk1* which is activated by MAPK signalling (Stevens et al., 2002). Moreover, it has been proposed that *Med23* binding is required for *Elk1* activity (Wang et al., 2009). To test whether ELK1 requires MED23 for its developmental activity I used an antisense morpholino approach to KD MED23 protein expression in *Fgf* and *Bmp* induced caps (Fig. 4.17).

Animal caps induced with *FGF* and *BMP* with MED23 KD showed no changes to early PGC markers at stage 19, however early meso/pronephros marker *Osr1* is upregulated 12-fold compared to the induced control caps. At stage 42, early and late PGC markers are heavily downregulated while meso/pronephros marker *Pax2* is upregulated fourfold, suggesting that PGCs are respecified to meso/pronephros in the absence of MED23. Given the

similarities with ELK1 depleted *FGF* and *BMP* caps rescued with hELK1 C mutant mRNA, this may suggest MED23 acts as a cofactor in conjunction with Phosphorylated ELK1 to regulate the formation of definitive PGCs.

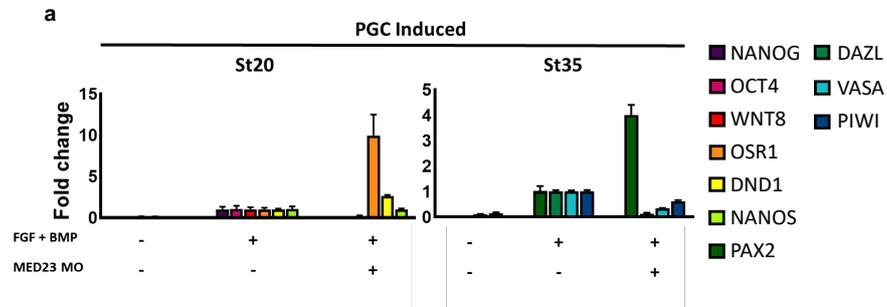


Figure 4.16. Gene expression of uninjected and PGC induced ACs with or without MED23 KD. (n=10, m=3, t=180).

4.3 Discussion: ELK1 has both conserved and novel roles in axolotl

4.1.1 Recap of aims and objectives

In this chapter I set out to address the following questions regarding ELK1's role in development:

- Is ELK1 required for early development?
- Does ELK1 regulate the expression of pluripotency factors?
- Does ELK1 depletion affect early cell fate decisions?
- Are there any similarities between the effects of ELK1 depletion in other animals such as humans, mice or frogs?
- Does ELK1 act downstream of FGF?
- Which domains within the ELK1 are necessary for it's function?

4.1.2 Summary of key findings

In this study I was able to address many of my key research questions around ELK1's role in early development. I demonstrated that ELK1 is required for the proper development of axolotl embryos. Depletion of ELK1 resulted in mild gastrulation defects, and disrupted ventral and intermediate mesoderm as well as expanded somitic mesoderm. ELK1 morphants also failed to develop following the end of neurulation and exhibited severe truncations along the anterior-posterior axis and notochord defects. ELK1 depletion did not appear to directly affect pluripotency but rather it resulted in a hypersensitivity to differentiation particularly in response to *FGF* and *ACTIVIN* a trait that may be conserved in mammals. ELK1 depletion also

ablated the production of PGC's in response to *FGF* and *BMP*. ELK1's role in PGC specification does appear to act downstream of FGF/MAPK signalling, however its role in mesodermal repression appears to be independent of FGF. It appears that ELK1's D-domain is required to initiate early PGC specification genes, possibly through activation of *WNT8*. Late PGC marker expression was also depleted when the phosphorylation of ELK1 was compromised. This effect was mirrored when the known ELK1 binding partner MED23 was depleted in PGC induction assays.

4.1.3 Discussion of key findings

Initially, I sought to investigate whether ELK1 regulates pluripotency as it has been proposed regulate pluripotency factors hESC (Goke et al., 2013). Indeed, the paradigm suggested by Goke and colleagues is that ELK1 has a dual role in regulating the pluripotent status of hESC. It is proposed that on a subset of loci ELK1 acts in conjunction with ERK2 to bind to the loci of target genes and positively regulate their transcription. The available ChIP-seq data shows that among the ELK1-ERK2 bound loci are factors essential to hESC self-renewal including NANOG, POU5F1/OCT4, NODAL and SMAD2. The second role, where ELK1 binds independent of ERK, is one of repression. ELK1 only loci included differentiation factors including T, SOX17, EOMES, and GATA2. These roles were evidenced by knockdown of either ELK1 or Erk2 resulting in the differentiation of hESC as confirmed broadly by marker immunostaining and microarray data (Goke et al., 2013). Despite these data in hESC, I found no evidence that ELK1 regulates pluripotency directly in axolotl. However, re-evaluating RNA-seq data from hESC following ELK1 depletion revealed several discrepancies in the model proposed by Goke and when taken in conjunction with the more recent paper on ELK1 activity (Prise and Sharrocks,

2019) suggests that ELK1's role in hESC is primarily repression of mesodermal differentiation. Interestingly, my data showed that ELK1 is also required for the proper formation of mesodermal tissues and depleting ELK1 resulted in the expansion of somitic mesoderm. Moreover, animal cap assays demonstrated that ELK1 KD hypersensitised ACs to mesodermal differentiation through both activin and FGF pathways, suggesting that ELK1's role as a mesodermal repressor may be conserved between amphibians and humans. Intriguingly, the expansion of somitic mesoderm was also accompanied by the depletion of ventral and intermediate mesodermal structures including the ventral lateral plate and meso/pronephros. Given that axial and paraxial mesoderm are the first cells to involute through the blastopore this may suggest a greater number of marginal zone cells differentiate to somitic mesoderm, thus there is a reduced number of cells to form intermediate and ventral mesoderm. This hypothesis is supported by evidence that ELK1 morphants show defects in the blastopore lip which resembles increased involution and reduced ingression, the former being the mode of migration of dorsal mesoderm, the latter being the mode of intermediate/ventral mesoderm (Shook et al., 2002, Shook and Keller, 2008).

Interestingly, the observed effects of ELK1 depletion in axolotl show some outward similarities with a phenotype presented by Nentwich and colleagues whereby they overexpressed a dominant negative form of *Xenopus Elk1*. The *Xelk1-EN^R* embryos also showed severe truncations across the anterior-posterior axis. However, *Elk1* depleted xenopus embryos unlike axolotl lacked a notochord and somitic mesoderm. Given that in ELK1 depleted axolotl both these structures are present or even over developed this may suggest differential activities of ELK1 in urodeles and anurans (Nentwich et al., 2009).

Both of these phenotypes contrast observations made in mouse embryos where *elk1* null mice showed no outward phenotype (Cesari et al., 2004). Concomitant with my observations that ELK1 depletion more closely resembles ELK1 depletion in hESC, the ELK1 phenotype in axolotl was able to be rescued by co-injection of hELK1 with the ELK1 MO at the 1 cell stage with a high efficiency suggesting the human gene can functionally compensate the loss of ELK1. This is likely due to the high degree of sequence similarity between the human and axolotl genes. This functional equivalence also allowed for the testing of hELK1 rescues with specific mutations in functional domains.

Previous research into the role of *FGF* in axolotl has shown that abrogation of *FGF* signalling using expression of a dominant-negative *FGF* receptor (XFD) resulted in a mild posterior truncation (Chatfield et al., 2014) which somewhat contrasted the effects observed in *Xenopus* which showed a severe truncation along the A-P axis (Amaya et al., 1991). Expression of XFD in axolotl did result in ablation of both PGCs, as well as the meso/pronephric ducts. Treatment of PI3K signalling inhibitor LY2941002 applied before the ZGA did result in a truncation of the A-P axis, more comparable, but still milder than the effects reported in *xenopus* (Carballada et al., 2001). Interestingly, U0126, a soluble inhibitor of MAPK activation appeared to moderately enlarge meso/pronephric ducts, while depleting the genital ridge and PGCs. This result again contrasted inhibition of MAPK in *Xenopus* embryos which causes gastrulation defects (Gotoh and Nishida, 1995, Sivak et al., 2005, Umbhauer et al., 1995). Given that ELK1 is suggested to act downstream of *FGF* signalling, I performed rescues of ELK1 morphants with hELK1 RNA encoding a mutation in D domain. The ELK1 D domain acts as a docking site for mitogen-

activated protein kinases (MAPKs) (Yang et al., 1998a, Ling et al., 1998, Jacobs et al., 1999), therefore this allowed me to study the effects of depleting ERK2 dependent ELK1 activity. Remarkably, hELK1 D mut was able to rescue both somitic mesodermal defects and anterior-posterior truncations but critically, still showed defects in intermediate mesoderm as well as loss of the posterior ventral-lateral region. This result suggested that the formation of meso/pronephros and LPM is dependent on ELK1 and ERK2 activity. Interestingly, the dependence of meso/pronephric development on ERK2 activity is supported by a separate study that has shown that the meso/pronephros in axolotl highly express phosphorylated ERK2 (Villiard et al., 2017), however the study did not investigate the VLP. The ability of the D mut hElk to rescue somitic mesodermal and A-P defects also suggests that mesodermal repression may be independent of *FGF*, which also appears to be conserved in hESC (Goke et al., 2013, Prise and Sharrocks, 2019).

Among multicellular lifeforms the germ line carries forward genetic material required to make the next generation of organism. Indeed, how the germ line segregates from soma is a fundamental question in developmental biology. As discussed in chapter one, there are two known modes of PGC specification, preformation and epigenesis however the majority of studies into the mechanisms of epigenesis have focused on mammals, despite epigenesis being the conserved mode of germ line formation in all animals. In particular the mechanisms governing germ line formation in lower vertebrates are poorly understood this is in part because the two main non-mammalian vertebrate model organisms: frogs and zebrafish, segregate their germline by preformation. During this study, I also was able to better characterise the process of PGC specification in axolotl.

During this study I also identified a novel role for ELK1 in which it regulates PGC specification. ELK1's role in PGC specification as with its role in the formation of intermediate mesoderm and VLP, likely requires MAPK1 / ERK2. This was evidenced by the hELK1 D mut being insufficient to rescue PGC specification. Given that in axolotl, PGCs develop from within the ventral-lateral mesoderm and following gastrulation this tissue migrates dorso-medially, the PGC's eventually reside just ventral to the intermediate mesoderm (Johnson et al., 2001, Bachvarova et al., 2004, Bachvarova et al., 2009b, Bachvarova et al., 2009a, Chatfield et al., 2014). This may suggest that ELK1-ERK2 regulates the ingressing ventral-lateral marginal zone as this area gives rise to both the VLM and intermediate mesoderm (Johnson et al., 2001, Bachvarova et al., 2004, Bachvarova et al., 2009b, Bachvarova et al., 2009a, Chatfield et al., 2014).

Moreover, ELK1 likely regulates early PGC specification in conjunction with ERK2 through modulation of *WNT8* expression. It also appears that later PGC differentiation is also dependent on ELK1 activity, while the mechanism by which this occurs is unclear it appears to be dependent on the phosphorylation status of ELK1 and may act in conjunction with the co-factor MED23. Given that the ELK1 sequence and indeed the D domain and C activation domain is conserved throughout vertebrates, this could suggest that ELK1's role in PGC specification is conserved in vertebrates (Saxton et al., 2016).

4.1.4 Limitations of this study

The findings of this study are subject to several limitations. While I was able to examine the expression profile of ELK1 across developmental stages I was unable to gather any spatial information on ELK1 transcripts. Techniques such

as in situ hybridisation would have been particularly useful as it could have informed in which tissues ELK1 functions and at what stage. Given time and embryo restrictions this was unfortunately not possible but would likely have been useful. As mentioned in the previous chapter a major challenge to this study has been with axolotl husbandry and embryo numbers. HREM only became available during the last few months of our axolotl colony. Furthermore, I was unable to have optimised the mounting conditions for embryos prior to performing experiments and embryo collections. While all experiments were able to be repeated 3 times, in some cases due to the relatively low number of embryos, intact samples from 3 separate mating's did not withstand mounting. As a result, not all imaging results may be generalisable.

While qPCR is a reliable method of interrogating gene expression, it only gives information on a gene-by-gene basis, a solution to this would be to perform RNA-seq which gives expression information genome-wide. During this study, RNA-sequencing costs were reduced significantly and while I had planned to perform RNA-seq on a range of samples due to the dismantling of our axolotl colony this was not possible.

4.1.5 Future work

As mentioned in the limitation section of this discussion future work could implement in situ hybridisations to probe for ELK1 and tissue marker expression at different stages in development, as well as performing whole embryo bulk RNA-seq or indeed single-cell sequencing at the same stages. This could provide useful information as to ELK1's expression domains and better characterise the effects of ELK1 depletion on specific tissues.

Given the novel role identified in this study for ELK1 in germline, it would be interesting to investigate whether this role like ELK1's role in mesodermal repression, is also conserved in mammals. Experiments could focus on performing timed knock-downs of hELK1 in hESC's at different stages of a GC induction using an inducible dead cas9 system. Indeed, this line of inquiry was originally planned for this study and creation of a dead cas9 hESC line had already been achieved.

5.1 Axolotl as a model for early development

Here I have explored the role of several genes in early axolotl development which had previously been implicated in the regulation of pluripotency in mammalian ESCs. I have shown key examples of where mechanisms involved in the regulation of pluripotency and early cell fate decisions are conserved between amphibians and mammals, building upon previous research which suggests that the core GRN governing early development are conserved between urodeles and mammals (Dixon et al., 2010, Swiers et al., 2010, Johnson et al., 2003b, Johnson et al., 2001, Johnson and Alberio, 2015, Chatfield et al., 2014, Bian et al., 2009, Johnson et al., 2003a). This suggests that these features existed in the tetrapod ancestor, which later gave rise to amniotes. In addition urodeles like most vertebrates, retained the basic skeletal structure of the tetrapod ancestor (Callier et al., 2009, Niedzwiedzki et al., 2010). Furthermore, an ancestral urodele-like embryology appears to be conserved throughout the evolution of amniotes (Bachvarova et al., 2009a).

Practically there a variety of features of axolotl embryos which have advantages and drawbacks as a model system for research which have influenced this study. As Amphibians develop externally this allows them to be easily injected, staged and their development can be tracked without any invasive procedures which would be required in mammals. Further, a single mating can produce up to 800 viable embryos (Internal records). However, it

is worth noting that the embryo availability varies to great extents in axolotl in some cases being as low as 120 embryos from a mating. During the course of this study problems with axolotl husbandry often resulted in periods of months without a successful mating as discussed in earlier chapters. Indeed, little is known about the mating habits of axolotl in general. If a natural mating is successful a female may also only lay up to around 20 embryos per hour (internal records) and thus, obtaining a large number of embryos which can be injected at the one cell stage can also be difficult. Despite such difficulties, the axolotl as a model offers some advantages over mammalian systems where embryos develop internally and are far fewer in number.

The axolotl embryo is uniquely large, measuring around 2mm which is advantageous in that it allows specific tissues to be explanted or transplanted, however this also means that axolotl embryos are very delicate especially when compared to other non-mammalian model systems such as zebrafish. The delicate nature of axolotl embryos also meant a high amount of sample loss particularly when mounting for imaging such as HREM. Given that the axolotl is less popular than other models of early development there also exists many challenges regarding experimental resources, such as functioning antibodies.

Over the course of this study however, there were large advances in the field of axolotl genomics with the publication of transcriptomes (Jiang et al., 2017, Evans et al., 2018) and the axolotl genome (Nowoshilow et al., 2018) which has and will continue to be an incredibly useful resource to study axolotl development. While there are some practical limitations of the axolotl as a model system in my view, the biological significance of the axolotl outweighs such limitations. More to this, the data presented in this study and others

demonstrates that axolotl is a useful model organism for studying the evolution of developmental processes between amphibians and mammals.

5.2 Axolotl and Xenopus

Interestingly, this has further highlighted differences between urodele and anuran amphibians. Based on their outward appearance particularly at the early embryonic stages, and the phylogenetic relationship between urodeles and anurans it would be reasonable to predict a high biological similarity. However, there are several key morphological and genetic differences which suggest this may not be the case (Dixon et al., 2010, Swiers et al., 2010, Johnson et al., 2003b, Johnson et al., 2001, Johnson and Alberio, 2015, Chatfield et al., 2014, Bian et al., 2009, Johnson et al., 2003a, Frankenberg et al., 2014, Shook et al., 2002, Shook and Keller, 2008, Kaneda and Motoki, 2012). The case of *NANOG* is a good example of this, *NANOG* has been lost in anurans and indeed it has been suggested that *Vent1/2* may have functionally replaced *NANOG* in frogs, however given my observations in this study, and given the study of *VENTX* activity in axolotl (Appendix 1) this does not appear to be the case. Interestingly, a subset of genes which are regulated by *NANOG* in axolotl likely through the deposition of H3K4me3 have maternally deposited H3K4me3 in *Xenopus* (Hontelez et al., 2015, Bright et al., 2021, Akkers et al., 2009). In anurans, it seems likely that the evolution of germlasm which repositioned PGCs to the vegetal hemisphere, was instrumental to the divergence from an ancestral pGRN. Crucially the frog animal cap no longer needed to produce both the soma and germline from a single population of pluripotent cells, therefore, rendering the pGRN functionally redundant, allowing divergence toward a multipotency GRN (mGRN). The extent to

which the mGRN has diverged from the ancestral pGRN is unclear. Further to this, our hypothesis is also supported by the research of NANOG'S function in germlasm containing Zebrafish whereby Nanog has evolved a separate function unique to teleost fish (Camp et al., 2009; Xu et al., 2012; Lee et al., 2013).

We have recently reported that NANOG likely evolved from VENT genes in diploblasts and that NANOG activity arose with the advent of mesoderm in triploblasts (Appendix 1). This may suggest that the increased cell diversity seen between vertebrates and their invertebrate last common ancestor may have been predicated on the evolution of an expanded pGRN. Given that the PGC's of axolotl like mammals arise from a pluripotent population in response to inductive signalling events, I also posit that the requirement to produce PGC's may also act as a constraint for pGRN and mesodermal GRN (mGRN) evolution, as the two networks are linked. This is further supported by previous findings that there is a conserved mGRN between urodeles and mammals (Swiers *et al*, 2010). Together these results suggest that the core mechanisms which govern pluripotency are conserved in the trunk of vertebrates but have likely diverged in many animals which specify their germ cells conditionally (Johnson et al., 2003b, Johnson et al., 2003a, Johnson and Alberio, 2015, Evans et al., 2014). It is worth noting that these animals are heavily represented in models of early development.

5.3 Germ plasm and the need for more model organisms

It is also worth considering that often, why an animal is used as a model organism often is not due to the universality of their phenotypic traits. Indeed, available resources for experimentation, ease of access, generational time, ease of husbandry, number of offspring, perceived intelligence and size are sometimes more pressing considerations. This is perhaps particularly true for the study of development as organisms with short generation times which produce high numbers of embryos are heavily represented. It is also true that many of these organisms specify their germline conditionally, a factor that may also lead to increased diversification (Johnson and Alberio, 2015, Evans et al., 2014). Observed differences in the developmental mechanisms between closely related species, some highlighted in this thesis highlight the need for an increased diversity of model organisms. Moreover, with technological advances, particularly in the areas of genomics and transcriptomics, opportunities to study development in a greater number of organisms are increasing. Consequently, understanding which features of development are conserved and which are derived could have large-scale implications for the understanding the evolution of the vertebrate embryo.

Appendices

Appendix 1. Publication ready for review (In attached file).

Bibliography

- ABDEL-RAHMAN, B., FIDDLER, M., RAPPOLEE, D. & PERGAMENT, E. 1995. Expression of transcription regulating genes in human preimplantation embryos. *Hum Reprod*, 10, 2787-92.
- ACAMPORA, D., DI GIOVANNANTONIO, L. G. & SIMEONE, A. 2013. Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development*, 140, 43-55.
- ACEMEL, R. D., MAESO, I. & GOMEZ-SKARMETA, J. L. 2017. Topologically associated domains: a successful scaffold for the evolution of gene regulation in animals. *Wiley Interdiscip Rev Dev Biol*, 6.
- ADACHI, K., SUEMORI, H., YASUDA, S. Y., NAKATSUJI, N. & KAWASE, E. 2010. Role of SOX2 in maintaining pluripotency of human embryonic stem cells. *Genes Cells*, 15, 455-70.
- AECKERLE, N., DRUMMER, C., DEBOWSKI, K., VIEBAHN, C. & BEHR, R. 2015. Primordial germ cell development in the marmoset monkey as revealed by pluripotency factor expression: suggestion of a novel model of embryonic germ cell translocation. *Mol Hum Reprod*, 21, 66-80.
- AGREN, J. A. & CLARK, A. G. 2018. Selfish genetic elements. *PLoS Genet*, 14, e1007700.
- AIKEN, C. E., SWOBODA, P. P., SKEPPER, J. N. & JOHNSON, M. H. 2004. The direct measurement of embryogenic volume and nucleo-cytoplasmic ratio during mouse pre-implantation development. *Reproduction*, 128, 527-35.
- AKKERS, R. C., VAN HEERINGEN, S. J., JACOBI, U. G., JANSSEN-MEGENS, E. M., FRANCOIJS, K. J., STUNNENBERG, H. G. & VEENSTRA, G. J. 2009. A hierarchy of H3K4me3 and H3K27me3 acquisition in spatial gene regulation in *Xenopus* embryos. *Dev Cell*, 17, 425-34.
- ALBERIO, R., JOHNSON, A. D., STICK, R. & CAMPBELL, K. H. 2005. Differential nuclear remodeling of mammalian somatic cells by *Xenopus laevis* oocyte and egg cytoplasm. *Exp Cell Res*, 307, 131-41.
- ALBERIO, R., KOBAYASHI, T. & SURANI, M. A. 2021. Conserved features of non-primate bilaminar disc embryos and the germline. *Stem Cell Reports*, 16, 1078-1092.
- AMAYA, E., MUSCI, T. J. & KIRSCHNER, M. W. 1991. Expression of a dominant negative mutant of the FGF receptor disrupts mesoderm formation in *Xenopus* embryos. *Cell*, 66, 257-70.
- ANGERILLI, A., SMIALOWSKI, P. & RUPP, R. A. 2018. The *Xenopus* animal cap transcriptome: building a mucociliary epithelium. *Nucleic Acids Res*, 46, 8772-8787.
- AOKI, F., WORRAD, D. M. & SCHULTZ, R. M. 1997. Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. *Dev Biol*, 181, 296-307.

- AVILION, A. A., NICOLIS, S. K., PEVNY, L. H., PEREZ, L., VIVIAN, N. & LOVELL-BADGE, R. 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev*, 17, 126-40.
- BACHVAROVA, R. F., CROTHER, B. I. & JOHNSON, A. D. 2009a. Evolution of germ cell development in tetrapods: comparison of urodeles and amniotes. *Evol Dev*, 11, 603-9.
- BACHVAROVA, R. F., CROTHER, B. I., MANOVA, K., CHATFIELD, J., SHOEMAKER, C. M., CREWS, D. P. & JOHNSON, A. D. 2009b. Expression of Dazl and Vasa in turtle embryos and ovaries: evidence for inductive specification of germ cells. *Evol Dev*, 11, 525-34.
- BACHVAROVA, R. F., MASI, T., DRUM, M., PARKER, N., MASON, K., PATIENT, R. & JOHNSON, A. D. 2004. Gene expression in the axolotl germ line: Axdazl, Axvh, AxOct4, and Axkit. *Dev Dyn*, 231, 871-80.
- BEDZHOV, I., GRAHAM, S. J., LEUNG, C. Y. & ZERNICKA-GOETZ, M. 2014. Developmental plasticity, cell fate specification and morphogenesis in the early mouse embryo. *Philos Trans R Soc Lond B Biol Sci*, 369.
- BEN-NUN, I. F., MONTAGUE, S. C., HOUCK, M. L., TRAN, H. T., GARITAONANDIA, I., LEONARDO, T. R., WANG, Y. C., CHARTER, S. J., LAURENT, L. C., RYDER, O. A. & LORING, J. F. 2011. Induced pluripotent stem cells from highly endangered species. *Nat Methods*, 8, 829-31.
- BENDALL, S. C., STEWART, M. H., MENENDEZ, P., GEORGE, D., VIJAYARAGAVAN, K., WERBOWETSKI-OGILVIE, T., RAMOS-MEJIA, V., ROULEAU, A., YANG, J., BOSSE, M., LAJOIE, G. & BHATIA, M. 2007. IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro. *Nature*, 448, 1015-21.
- BERTERO, A., MADRIGAL, P., GALLI, A., HUBNER, N. C., MORENO, I., BURKS, D., BROWN, S., PEDERSEN, R. A., GAFFNEY, D., MENDJAN, S., PAUKLIN, S. & VALLIER, L. 2015. Activin/nodal signaling and NANOG orchestrate human embryonic stem cell fate decisions by controlling the H3K4me3 chromatin mark. *Genes Dev*, 29, 702-17.
- BERTOCCHINI, F., SKROMNE, I., WOLPERT, L. & STERN, C. D. 2004. Determination of embryonic polarity in a regulative system: evidence for endogenous inhibitors acting sequentially during primitive streak formation in the chick embryo. *Development*, 131, 3381-90.
- BESNARD, A., GALAN-RODRIGUEZ, B., VANHOUTTE, P. & CABOCHE, J. 2011. Elk-1 a transcription factor with multiple facets in the brain. *Front Neurosci*, 5, 35.
- BIAN, Y., ALBERIO, R., ALLEGRUCCI, C., CAMPBELL, K. H. & JOHNSON, A. D. 2009. Epigenetic marks in somatic chromatin are remodelled to resemble pluripotent nuclei by amphibian oocyte extracts. *Epigenetics*, 4, 194-202.
- BIASE, F. H., CAO, X. & ZHONG, S. 2014. Cell fate inclination within 2-cell and 4-cell mouse embryos revealed by single-cell RNA sequencing. *Genome Res*, 24, 1787-96.
- BISCHOFF, M., PARFITT, D. E. & ZERNICKA-GOETZ, M. 2008. Formation of the embryonic-abembryonic axis of the mouse blastocyst: relationships between orientation of early cleavage divisions and pattern of symmetric/asymmetric divisions. *Development*, 135, 953-62.

- BOIANI, M. & SCHOLER, H. R. 2005. Regulatory networks in embryo-derived pluripotent stem cells. *Nat Rev Mol Cell Biol*, 6, 872-84.
- BORGGREFE, T. & YUE, X. 2011. Interactions between subunits of the Mediator complex with gene-specific transcription factors. *Semin Cell Dev Biol*, 22, 759-68.
- BOROVIK, T., LOOS, R., BERTONE, P., SMITH, A. & NICHOLS, J. 2014. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat Cell Biol*, 16, 516-28.
- BOTERENBROOD, E. C. & NIEUWKOOP, P. D. 1973. The formation of the mesoderm in urodelean amphibians : V. Its regional induction by the endoderm. *Wilhelm Roux Arch Entwickl Mech Org*, 173, 319-332.
- BOULTON, T. G., NYE, S. H., ROBBINS, D. J., IP, N. Y., RADZIEJEWSKA, E., MORGENBESSER, S. D., DEPINHO, R. A., PANAYOTATOS, N., COBB, M. H. & YANCOPOULOS, G. D. 1991. ERKs: a family of protein-serine/threonine kinases that are activated and tyrosine phosphorylated in response to insulin and NGF. *Cell*, 65, 663-75.
- BOUNIOL, C., NGUYEN, E. & DEBEY, P. 1995. Endogenous transcription occurs at the 1-cell stage in the mouse embryo. *Exp Cell Res*, 218, 57-62.
- BOYER, L. A., LEE, T. I., COLE, M. F., JOHNSTONE, S. E., LEVINE, S. S., ZUCKER, J. P., GUENTHER, M. G., KUMAR, R. M., MURRAY, H. L., JENNER, R. G., GIFFORD, D. K., MELTON, D. A., JAENISCH, R. & YOUNG, R. A. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, 122, 947-56.
- BRETON, A., SHARMA, R., DIAZ, A. C., PARHAM, A. G., GRAHAM, A., NEIL, C., WHITELAW, C. B., MILNE, E. & DONADEU, F. X. 2013. Derivation and characterization of induced pluripotent stem cells from equine fibroblasts. *Stem Cells Dev*, 22, 611-21.
- BRIGGS, J. A., WEINREB, C., WAGNER, D. E., MEGASON, S., PESHKIN, L., KIRSCHNER, M. W. & KLEIN, A. M. 2018. The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science*, 360.
- BRIGHT, A. R., VAN GENESEN, S., LI, Q., GRASSO, A., FROLICH, S., VAN DER SANDE, M., VAN HEERINGEN, S. J. & VEENSTRA, G. J. C. 2021. Combinatorial transcription factor activities on open chromatin induce embryonic heterogeneity in vertebrates. *EMBO J*, 40, e104913.
- BRITTEN, R. J. & DAVIDSON, E. H. 1971. Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. *Q Rev Biol*, 46, 111-38.
- BRONS, I. G., SMITHERS, L. E., TROTTER, M. W., RUGG-GUNN, P., SUN, B., CHUVA DE SOUSA LOPES, S. M., HOWLETT, S. K., CLARKSON, A., AHRLUND-RICHTER, L., PEDERSEN, R. A. & VALLIER, L. 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature*, 448, 191-5.
- BROWN, S., TEO, A., PAUKLIN, S., HANNAN, N., CHO, C. H., LIM, B., VARDY, L., DUNN, N. R., TROTTER, M., PEDERSEN, R. & VALLIER, L. 2011. Activin/Nodal signaling controls divergent transcriptional networks in human embryonic stem cells and in endoderm progenitors. *Stem Cells*, 29, 1176-85.
- BURTON, P. M. 2008. Insights from diploblasts; the evolution of mesoderm and muscle. *J Exp Zool B Mol Dev Evol*, 310, 5-14.

- BYLUND, M., ANDERSSON, E., NOVITCH, B. G. & MUHR, J. 2003. Vertebrate neurogenesis is counteracted by Sox1-3 activity. *Nat Neurosci*, 6, 1162-8.
- BYRNE, J. A., SIMONSSON, S., WESTERN, P. S. & GURDON, J. B. 2003. Nuclei of adult mammalian somatic cells are directly reprogrammed to Oct4 stem cell gene expression by amphibian oocytes. *Curr Biol*, 13, 1206-13.
- CALLIER, V., CLACK, J. A. & AHLBERG, P. E. 2009. Contrasting developmental trajectories in the earliest known tetrapod forelimbs. *Science*, 324, 364-7.
- CAMP, E., SANCHEZ-SANCHEZ, A. V., GARCIA-ESPANA, A., DESALLE, R., ODQVIST, L., ENRIQUE O'CONNOR, J. & MULLOR, J. L. 2009. Nanog regulates proliferation during early fish development. *Stem Cells*, 27, 2081-91.
- CAMUS, A., PEREA-GOMEZ, A., MOREAU, A. & COLLIGNON, J. 2006. Absence of Nodal signaling promotes precocious neural differentiation in the mouse embryo. *Dev Biol*, 295, 743-55.
- CAO, Y., KNOCHER, S., DONOW, C., MIETHE, J., KAUFMANN, E. & KNOCHER, W. 2004. The POU factor Oct-25 regulates the Xvent-2B gene and counteracts terminal differentiation in *Xenopus* embryos. *J Biol Chem*, 279, 43735-43.
- CAO, Y., SIEGEL, D. & KNOCHER, W. 2006. *Xenopus* POU factors of subclass V inhibit activin/nodal signaling during gastrulation. *Mech Dev*, 123, 614-25.
- CARBALLADA, R., YASUO, H. & LEMAIRE, P. 2001. Phosphatidylinositol-3 kinase acts in parallel to the ERK MAP kinase in the FGF pathway during *Xenopus* mesoderm induction. *Development*, 128, 35-44.
- CASTILLO, S. D. & SANCHEZ-CEPEDES, M. 2012. The SOX family of genes in cancer development: biological relevance and opportunities for therapy. *Expert Opin Ther Targets*, 16, 903-19.
- CAUFFMAN, G., LIEBAERS, I., VAN STEIRTEGHEM, A. & VAN DE VELDE, H. 2006. POU5F1 isoforms show different expression patterns in human embryonic stem cells and preimplantation embryos. *Stem Cells*, 24, 2685-91.
- CAUFFMAN, G., VAN DE VELDE, H., LIEBAERS, I. & VAN STEIRTEGHEM, A. 2005. Oct4 mRNA and protein expression during human preimplantation development. *Mol Hum Reprod*, 11, 173-81.
- CESARI, F., BRECHT, S., VINTERSTEN, K., VUONG, L. G., HOFMANN, M., KLINGEL, K., SCHNORR, J. J., ARSENIAN, S., SCHILD, H., HERDEGEN, T., WIEBEL, F. F. & NORDHEIM, A. 2004. Mice deficient for the ets transcription factor elk-1 show normal immune responses and mildly impaired neuronal gene activation. *Mol Cell Biol*, 24, 294-305.
- CHAMBERS, I., COLBY, D., ROBERTSON, M., NICHOLS, J., LEE, S., TWEEDIE, S. & SMITH, A. 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell*, 113, 643-55.
- CHAMBERS, I., SILVA, J., COLBY, D., NICHOLS, J., NIJMEIJER, B., ROBERTSON, M., VRANA, J., JONES, K., GROTEWOLD, L. & SMITH, A. 2007. Nanog safeguards pluripotency and mediates germline development. *Nature*, 450, 1230-4.
- CHAMBERS, I. & TOMLINSON, S. R. 2009. The transcriptional foundation of pluripotency. *Development*, 136, 2311-22.
- CHATFIELD, J., O'REILLY, M. A., BACHVAROVA, R. F., FERJENTSCHIK, Z., REDWOOD, C., WALMSLEY, M., PATIENT, R., LOOSE, M. & JOHNSON, A. D. 2014. Stochastic specification of primordial germ cells from mesoderm precursors in axolotl embryos. *Development*, 141, 2429-40.

- CHAZAUD, C., YAMANAKA, Y., PAWSON, T. & ROSSANT, J. 2006. Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Dev Cell*, 10, 615-24.
- CHEN, H., AKSOY, I., GONNOT, F., OSTEIL, P., AUBRY, M., HAMELA, C., ROGNARD, C., HOCHARD, A., VOISIN, S., FONTAINE, E., MURE, M., AFANASSIEFF, M., CLEROUX, E., GUIBERT, S., CHEN, J., VALLOT, C., ACLOQUE, H., GENTHON, C., DONNADIEU, C., DE VOS, J., SANLAVILLE, D., GUERIN, J. F., WEBER, M., STANTON, L. W., ROUGEULLE, C., PAIN, B., BOURILLOT, P. Y. & SAVATIER, P. 2015. Reinforcement of STAT3 activity reprogrammes human embryonic stem cells to naive-like pluripotency. *Nat Commun*, 6, 7095.
- CHOI, H. J., JIN, S. D., RENGARAJ, D., KIM, J. H., PAIN, B. & HAN, J. Y. 2021. Differential transcriptional regulation of the NANOG gene in chicken primordial germ cells and embryonic stem cells. *J Anim Sci Biotechnol*, 12, 40.
- CHOI, H. J., KIM, I., LEE, H. J., PARK, Y. H., SUH, J. Y. & HAN, J. Y. 2018. Chicken NANOG self-associates via a novel folding-upon-binding mechanism. *FASEB J*, 32, 2563-2573.
- CONLON, F. L., BARTH, K. S. & ROBERTSON, E. J. 1991. A novel retrovirally induced embryonic lethal mutation in the mouse: assessment of the developmental fate of embryonic stem cells homozygous for the 413.d proviral integration. *Development*, 111, 969-81.
- CROTHER, B. I. A. M., C.M. 2018. Linking a biological mechanism to evolvability. *J. Phylogenetics Evol. Biol.*, 6, 192.
- CRUZALEGUI, F. H., CANO, E. & TREISMAN, R. 1999. ERK activation induces phosphorylation of Elk-1 at multiple S/T-P motifs to high stoichiometry. *Oncogene*, 18, 7948-57.
- DARR, H., MAYSHAR, Y. & BENVENISTY, N. 2006. Overexpression of NANOG in human ES cells enables feeder-free growth while inducing primitive ectoderm features. *Development*, 133, 1193-201.
- DAVIDSON, E. H. & ERWIN, D. H. 2006. Gene regulatory networks and the evolution of animal body plans. *Science*, 311, 796-800.
- DAWKINS, R. 1976. *The selfish gene*, New York, Oxford University Press.
- DEGLINCERTI, A., CROFT, G. F., PIETILA, L. N., ZERNICKA-GOETZ, M., SIGGIA, E. D. & BRIVANLOU, A. H. 2016. Self-organization of the in vitro attached human embryo. *Nature*, 533, 251-4.
- DIETRICH, J. E. & HIRAGI, T. 2007. Stochastic patterning in the mouse pre-implantation embryo. *Development*, 134, 4219-31.
- DIXON, J. E., ALLEGRUCCI, C., REDWOOD, C., KUMP, K., BIAN, Y., CHATFIELD, J., CHEN, Y. H., SOTTILE, V., VOSS, S. R., ALBERIO, R. & JOHNSON, A. D. 2010. Axolotl Nanog activity in mouse embryonic stem cells demonstrates that ground state pluripotency is conserved from urodele amphibians to mammals. *Development*, 137, 2973-80.
- DO, D. V., UEDA, J., MESSERSCHMIDT, D. M., LORTHONGPANICH, C., ZHOU, Y., FENG, B., GUO, G., LIN, P. J., HOSSAIN, M. Z., ZHANG, W., MOH, A., WU, Q., ROBSON, P., NG, H. H., POELLINGER, L., KNOWLES, B. B., SOLTER, D. & FU, X. Y. 2013. A genetic and developmental pathway from STAT3 to the OCT4-NANOG circuit is essential for maintenance of ICM lineages in vivo. *Genes Dev*, 27, 1378-90.

- DODSWORTH, B. T., FLYNN, R. & COWLEY, S. A. 2015. The Current State of Naive Human Pluripotency. *Stem Cells*, 33, 3181-6.
- DOREY, K. & AMAYA, E. 2010. FGF signalling: diverse roles during early vertebrate embryogenesis. *Development*, 137, 3731-42.
- ENDO, Y., KAMEI, K. I. & INOUE-MURAYAMA, M. 2020. Genetic Signatures of Evolution of the Pluripotency Gene Regulating Network across Mammals. *Genome Biol Evol*, 12, 1806-1818.
- ERWIN, D. H. & DAVIDSON, E. H. 2009. The evolution of hierarchical gene regulatory networks. *Nat Rev Genet*, 10, 141-8.
- EVANS, T., JOHNSON, A. D. & LOOSE, M. 2018. Virtual Genome Walking across the 32 Gb *Ambystoma mexicanum* genome; assembling gene models and intronic sequence. *Sci Rep*, 8, 618.
- EVANS, T., WADE, C. M., CHAPMAN, F. A., JOHNSON, A. D. & LOOSE, M. 2014. Acquisition of germ plasm accelerates vertebrate evolution. *Science*, 344, 200-3.
- EZASHI, T., TELUGU, B. P., ALEXENKO, A. P., SACHDEV, S., SINHA, S. & ROBERTS, R. M. 2009. Derivation of induced pluripotent stem cells from pig somatic cells. *Proc Natl Acad Sci U S A*, 106, 10993-8.
- FEI, J. F., SCHUEZ, M., TAZAKI, A., TANIGUCHI, Y., ROENSCH, K. & TANAKA, E. M. 2014. CRISPR-mediated genomic deletion of Sox2 in the axolotl shows a requirement in spinal cord neural stem cell amplification during tail regeneration. *Stem Cell Reports*, 3, 444-59.
- FIDALGO, M., SHEKAR, P. C., ANG, Y. S., FUJIWARA, Y., ORKIN, S. H. & WANG, J. 2011. Zfp281 functions as a transcriptional repressor for pluripotency of mouse embryonic stem cells. *Stem Cells*, 29, 1705-16.
- FOGARTY, N. M. E., MCCARTHY, A., SNIJDERS, K. E., POWELL, B. E., KUBIKOVA, N., BLAKELEY, P., LEA, R., ELDER, K., WAMAITHA, S. E., KIM, D., MACIULYTE, V., KLEINJUNG, J., KIM, J. S., WELLS, D., VALLIER, L., BERTERO, A., TURNER, J. M. A. & NIAKAN, K. K. 2017. Genome editing reveals a role for OCT4 in human embryogenesis. *Nature*, 550, 67-73.
- FRANKENBERG, S., GERBE, F., BESSONNARD, S., BELVILLE, C., POUCHIN, P., BARDOT, O. & CHAZAUD, C. 2011. Primitive endoderm differentiates via a three-step mechanism involving Nanog and RTK signaling. *Dev Cell*, 21, 1005-13.
- FRANKENBERG, S. & RENFREE, M. B. 2013. On the origin of POU5F1. *BMC Biol*, 11, 56.
- FRANKENBERG, S. R., FRANK, D., HARLAND, R., JOHNSON, A. D., NICHOLS, J., NIWA, H., SCHOLER, H. R., TANAKA, E., WYLIE, C. & BRICKMAN, J. M. 2014. The POU-er of gene nomenclature. *Development*, 141, 2921-3.
- GAARENSTROOM, T. & HILL, C. S. 2014. TGF-beta signaling to chromatin: how Smads regulate transcription during self-renewal and differentiation. *Semin Cell Dev Biol*, 32, 107-18.
- GAFNI, O., WEINBERGER, L., MANSOUR, A. A., MANOR, Y. S., CHOMSKY, E., BEN-YOSEF, D., KALMA, Y., VIUKOV, S., MAZA, I., ZVIRAN, A., RAIS, Y., SHIPONY, Z., MUKAMEL, Z., KRUPALNIK, V., ZERBIB, M., GEULA, S., CASPI, I., SCHNEIR, D., SHWARTZ, T., GILAD, S., AMANN-ZALCENSTEIN, D., BENJAMIN, S., AMIT, I., TANAY, A.,

- MASSARWA, R., NOVERSHTERN, N. & HANNA, J. H. 2013. Derivation of novel human ground state naive pluripotent stem cells. *Nature*, 504, 282-6.
- GAGNON, J. A., OBBAD, K. & SCHIER, A. F. 2018. The primary role of zebrafish nanog is in extra-embryonic tissue. *Development*, 145.
- GALAN, A., MONTANER, D., POO, M. E., VALBUENA, D., RUIZ, V., AGUILAR, C., DOPAZO, J. & SIMON, C. 2010. Functional genomics of 5- to 8-cell stage human embryos by blastomere single-cell cDNA analysis. *PLoS One*, 5, e13615.
- GARDNER, R. L. & ROSSANT, J. 1979. Investigation of the fate of 4-5 day post-coitum mouse inner cell mass cells by blastocyst injection. *J Embryol Exp Morphol*, 52, 141-52.
- GHIMIRE, S., VAN DER JEUGHT, M., NEUPANE, J., ROOST, M. S., ANCKAERT, J., POPOVIC, M., VAN NIEUWERBURGH, F., MESTDAGH, P., VANDESOMPELE, J., DEFORCE, D., MENTEN, B., CHUVA DE SOUSA LOPES, S., DE SUTTER, P. & HEINDRYCKX, B. 2018. Comparative analysis of naive, primed and ground state pluripotency in mouse embryonic stem cells originating from the same genetic background. *Sci Rep*, 8, 5884.
- GOKE, J., CHAN, Y. S., YAN, J., VINGRON, M. & NG, H. H. 2013. Genome-wide kinase-chromatin interactions reveal the regulatory network of ERK signaling in human embryonic stem cells. *Mol Cell*, 50, 844-55.
- GONG, S., LI, Q., JETER, C. R., FAN, Q., TANG, D. G. & LIU, B. 2015. Regulation of NANOG in cancer cells. *Mol Carcinog*, 54, 679-87.
- GOOLAM, M., SCIALDONE, A., GRAHAM, S. J. L., MACAULAY, I. C., JEDRUSIK, A., HUPALOWSKA, A., VOET, T., MARIONI, J. C. & ZERNICKA-GOETZ, M. 2016. Heterogeneity in Oct4 and Sox2 Targets Biases Cell Fate in 4-Cell Mouse Embryos. *Cell*, 165, 61-74.
- GOTO, T., ADJAYE, J., RODECK, C. H. & MONK, M. 1999. Identification of genes expressed in human primordial germ cells at the time of entry of the female germ line into meiosis. *Mol Hum Reprod*, 5, 851-60.
- GOTOH, Y. & NISHIDA, E. 1995. Activation mechanism and function of the MAP kinase cascade. *Mol Reprod Dev*, 42, 486-92.
- GRAHAM, V., KHUDYAKOV, J., ELLIS, P. & PEVNY, L. 2003. SOX2 functions to maintain neural progenitor identity. *Neuron*, 39, 749-65.
- GRANDE, C., MARTIN-DURAN, J. M., KENNY, N. J., TRUCHADO-GARCIA, M. & HEJNOL, A. 2014. Evolution, divergence and loss of the Nodal signalling pathway: new data and a synthesis across the Bilateria. *Int J Dev Biol*, 58, 521-32.
- GREBER, B., WU, G., BERNEMANN, C., JOO, J. Y., HAN, D. W., KO, K., TAPIA, N., SABOUR, D., STERNECKERT, J., TESAR, P. & SCHOLER, H. R. 2010. Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. *Cell Stem Cell*, 6, 215-26.
- GREEN, J. B., NEW, H. V. & SMITH, J. C. 1992. Responses of embryonic *Xenopus* cells to activin and FGF are separated by multiple dose thresholds and correspond to distinct axes of the mesoderm. *Cell*, 71, 731-9.
- GROSE, C. & SPIGLAND, I. 1976. Guillain-Barre syndrome following administration of live measles vaccine. *Am J Med*, 60, 441-3.

- GRUBELNIK, G., BOSTJANCIC, E., PAVLIC, A., KOS, M. & ZIDAR, N. 2020. NANOG expression in human development and cancerogenesis. *Exp Biol Med (Maywood)*, 245, 456-464.
- GUO, X. & WANG, X. F. 2009. Signaling cross-talk between TGF-beta/BMP and other pathways. *Cell Res*, 19, 71-88.
- GURDON, J. B., ELSDALE, T. R. & FISCHBERG, M. 1958. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature*, 182, 64-5.
- HAMBILIKI, F., STROM, S., ZHANG, P. & STAVREUS-EVERS, A. 2012. Co-localization of NANOG and OCT4 in human pre-implantation embryos and in human embryonic stem cells. *J Assist Reprod Genet*, 29, 1021-8.
- HAN, J. Y., LEE, H. G., PARK, Y. H., HWANG, Y. S., KIM, S. K., RENGARAJ, D., CHO, B. W. & LIM, J. M. 2018. Acquisition of pluripotency in the chick embryo occurs during intrauterine embryonic development via a unique transcriptional network. *J Anim Sci Biotechnol*, 9, 31.
- HAN, X., HAN, J., DING, F., CAO, S., LIM, S. S., DAI, Y., ZHANG, R., ZHANG, Y., LIM, B. & LI, N. 2011. Generation of induced pluripotent stem cells from bovine embryonic fibroblast cells. *Cell Res*, 21, 1509-12.
- HANSIS, C., GRIFO, J. A. & KREY, L. C. 2000. Oct4 expression in inner cell mass and trophectoderm of human blastocysts. *Mol Hum Reprod*, 6, 999-1004.
- HART, A. H., HARTLEY, L., PARKER, K., IBRAHIM, M., LOOIJENGA, L. H., PAUCHNIK, M., CHOW, C. W. & ROBB, L. 2005. The pluripotency homeobox gene NANOG is expressed in human germ cell tumors. *Cancer*, 104, 2092-8.
- HASHIMSHONY, T., FEDER, M., LEVIN, M., HALL, B. K. & YANAI, I. 2015. Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. *Nature*, 519, 219-22.
- HASSLER, M. & RICHMOND, T. J. 2001. The B-box dominates SAP-1-SRF interactions in the structure of the ternary complex. *EMBO J*, 20, 3018-28.
- HATANO, N., MORI, Y., OH-HORA, M., KOSUGI, A., FUJIKAWA, T., NAKAI, N., NIWA, H., MIYAZAKI, J., HAMAOKA, T. & OGATA, M. 2003. Essential role for ERK2 mitogen-activated protein kinase in placental development. *Genes Cells*, 8, 847-56.
- HAY, D. C., SUTHERLAND, L., CLARK, J. & BURDON, T. 2004. Oct4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. *Stem Cells*, 22, 225-35.
- HELLSTEN, U., HARLAND, R. M., GILCHRIST, M. J., HENDRIX, D., JURKA, J., KAPITONOV, V., OVCHARENKO, I., PUTNAM, N. H., SHU, S., TAHER, L., BLITZ, I. L., BLUMBERG, B., DICHMANN, D. S., DUBCHAK, I., AMAYA, E., DETTER, J. C., FLETCHER, R., GERHARD, D. S., GOODSTEIN, D., GRAVES, T., GRIGORIEV, I. V., GRIMWOOD, J., KAWASHIMA, T., LINDQUIST, E., LUCAS, S. M., MEAD, P. E., MITROS, T., OGINO, H., OHTA, Y., POLIAKOV, A. V., POLLET, N., ROBERT, J., SALAMOV, A., SATER, A. K., SCHMUTZ, J., TERRY, A., VIZE, P. D., WARREN, W. C., WELLS, D., WILLS, A., WILSON, R. K., ZIMMERMAN, L. B., ZORN, A. M., GRAINGER, R., GRAMMER, T., KHOKHA, M. K., RICHARDSON, P. M. & ROKHSAR, D. S. 2010. The genome of the Western clawed frog *Xenopus tropicalis*. *Science*, 328, 633-6.

- HILL, C. S., MARAIS, R., JOHN, S., WYNNE, J., DALTON, S. & TREISMAN, R. 1993. Functional analysis of a growth factor-responsive transcription factor complex. *Cell*, 73, 395-406.
- HINKLEY, C. S., MARTIN, J. F., LEIBHAM, D. & PERRY, M. 1992. Sequential expression of multiple POU proteins during amphibian early development. *Mol Cell Biol*, 12, 638-49.
- HIRANO, T., ISHIHARA, K. & HIBI, M. 2000. Roles of STAT3 in mediating the cell growth, differentiation and survival signals relayed through the IL-6 family of cytokine receptors. *Oncogene*, 19, 2548-56.
- HOEI-HANSEN, C. E., ALMSTRUP, K., NIELSEN, J. E., BRASK SONNE, S., GRAEM, N., SKAKKEBAEK, N. E., LEFFERS, H. & RAJPERT-DE MEYTS, E. 2005. Stem cell pluripotency factor NANOG is expressed in human fetal gonocytes, testicular carcinoma in situ and germ cell tumours. *Histopathology*, 47, 48-56.
- HONTELEZ, S., VAN KRUIJSBERGEN, I., GEORGIOU, G., VAN HEERINGEN, S. J., BOGDANOVIC, O., LISTER, R. & VEENSTRA, G. J. C. 2015. Embryonic transcription is controlled by maternally defined chromatin state. *Nat Commun*, 6, 10148.
- HOWE, F. S., FISCHL, H., MURRAY, S. C. & MELLOR, J. 2017. Is H3K4me3 instructive for transcription activation? *Bioessays*, 39, 1-12.
- HUANGFU, D., OSAFUNE, K., MAEHR, R., GUO, W., EIJKELENBOOM, A., CHEN, S., MUHLESTEIN, W. & MELTON, D. A. 2008. Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nat Biotechnol*, 26, 1269-75.
- HUG, C. B., GRIMALDI, A. G., KRUSE, K. & VAQUERIZAS, J. M. 2017. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell*, 169, 216-228 e19.
- HUNTRISS, J., LU, J., HEMMINGS, K., BAYNE, R., ANDERSON, R., RUTHERFORD, A., BALEN, A., ELDER, K. & PICTON, H. M. 2017. Isolation and expression of the human gametocyte-specific factor 1 gene (GTSF1) in fetal ovary, oocytes, and preimplantation embryos. *J Assist Reprod Genet*, 34, 23-31.
- IRIE, N., WEINBERGER, L., TANG, W. W., KOBAYASHI, T., VIUKOV, S., MANOR, Y. S., DIETMANN, S., HANNA, J. H. & SURANI, M. A. 2015. SOX17 is a critical specifier of human primordial germ cell fate. *Cell*, 160, 253-68.
- JACOBS, D., GLOSSIP, D., XING, H., MUSLIN, A. J. & KORNFELD, K. 1999. Multiple docking sites on substrate proteins form a modular system that mediates recognition by ERK MAP kinase. *Genes Dev*, 13, 163-75.
- JALVY, S., VESCHAMBRE, P., FEDOU, S., REZVANI, H. R., THEZE, N. & THIEBAUD, P. 2019. Leukemia inhibitory factor signaling in *Xenopus* embryo: Insights from gain of function analysis and dominant negative mutant of the receptor. *Dev Biol*, 447, 200-213.
- JAMES, D., LEVINE, A. J., BESSER, D. & HEMMATI-BRIVANLOU, A. 2005. TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development*, 132, 1273-82.
- JANKNECHT, R. & NORDHEIM, A. 1992. Elk-1 protein domains required for direct and SRF-assisted DNA-binding. *Nucleic Acids Res*, 20, 3317-24.
- JANKNECHT, R., ZINCK, R., ERNST, W. H. & NORDHEIM, A. 1994. Functional dissection of the transcription factor Elk-1. *Oncogene*, 9, 1273-8.

- JERABEK, S., MERINO, F., SCHOLER, H. R. & COJOCARU, V. 2014. OCT4: dynamic DNA binding pioneers stem cell pluripotency. *Biochim Biophys Acta*, 1839, 138-54.
- JIANG, P., NELSON, J. D., LENG, N., COLLINS, M., SWANSON, S., DEWEY, C. N., THOMSON, J. A. & STEWART, R. 2017. Analysis of embryonic development in the unsequenced axolotl: Waves of transcriptomic upheaval and stability. *Dev Biol*, 426, 143-154.
- JOHNSON, A. D. & ALBERIO, R. 2015. Primordial germ cells: the first cell lineage or the last cells standing? *Development*, 142, 2730-9.
- JOHNSON, A. D., BACHVAROVA, R. F., DRUM, M. & MASI, T. 2001. Expression of axolotl DAZL RNA, a marker of germ plasm: widespread maternal RNA and onset of expression in germ cells approaching the gonad. *Dev Biol*, 234, 402-15.
- JOHNSON, A. D., CROTHER, B., WHITE, M. E., PATIENT, R., BACHVAROVA, R. F., DRUM, M. & MASI, T. 2003a. Regulative germ cell specification in axolotl embryos: a primitive trait conserved in the mammalian lineage. *Philos Trans R Soc Lond B Biol Sci*, 358, 1371-9.
- JOHNSON, A. D., DRUM, M., BACHVAROVA, R. F., MASI, T., WHITE, M. E. & CROTHER, B. I. 2003b. Evolution of predetermined germ cells in vertebrate embryos: implications for macroevolution. *Evol Dev*, 5, 414-31.
- KANEDA, T. & MOTOKI, J. Y. 2012. Gastrulation and pre-gastrulation morphogenesis, inductions, and gene expression: similarities and dissimilarities between urodelean and anuran embryos. *Dev Biol*, 369, 1-18.
- KELLY, S. J. 1977. Studies of the developmental potential of 4- and 8-cell stage mouse blastomeres. *J Exp Zool*, 200, 365-76.
- KERR, C. L., HILL, C. M., BLUMENTHAL, P. D. & GEARHART, J. D. 2008. Expression of pluripotent stem cell markers in the human fetal ovary. *Hum Reprod*, 23, 589-99.
- KIM, J., CHU, J., SHEN, X., WANG, J. & ORKIN, S. H. 2008. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell*, 132, 1049-61.
- KISHI, M., MIZUSEKI, K., SASAI, N., YAMAZAKI, H., SHIOTA, K., NAKANISHI, S. & SASAI, Y. 2000. Requirement of Sox2-mediated signaling for differentiation of early *Xenopus* neuroectoderm. *Development*, 127, 791-800.
- KOBAYASHI, T., ZHANG, H., TANG, W. W. C., IRIE, N., WITHEY, S., KLISCH, D., SYBIRNA, A., DIETMANN, S., CONTRERAS, D. A., WEBB, R., ALLEGRUCCI, C., ALBERIO, R. & SURANI, M. A. 2017. Principles of early human development and germ cell program from conserved model systems. *Nature*, 546, 416-420.
- KOMATSU, K. & FUJIMORI, T. 2015. Multiple phases in regulation of Nanog expression during pre-implantation development. *Dev Growth Differ*, 57, 648-56.
- KOPP, J. L., ORMSBEE, B. D., DESLER, M. & RIZZINO, A. 2008. Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells*, 26, 903-11.
- KUMAR, S., STECHER, G., SULESKI, M. & HEDGES, S. B. 2017. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol Biol Evol*, 34, 1812-1819.

- KUNATH, T., SABA-EL-LEIL, M. K., ALMOUSAILLEAKH, M., WRAY, J., MELOCHE, S. & SMITH, A. 2007. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development*, 134, 2895-902.
- LATOS, P. A. & HEMBERGER, M. 2016. From the stem of the placental tree: trophoblast stem cells and their progeny. *Development*, 143, 3650-3660.
- LAVAU, J., BERNARD, F., TRIFILIEFF, P., PASCOLI, V., KAPPES, V., PAGES, C., VANHOUTTE, P. & CABOCHE, J. 2007. A TAT-DEF-Elk-1 peptide regulates the cytonuclear trafficking of Elk-1 and controls cytoskeleton dynamics. *J Neurosci*, 27, 14448-58.
- LAVIAL, F., ACLOQUE, H., BERTOCCHINI, F., MACLEOD, D. J., BOAST, S., BACHELARD, E., MONTILLET, G., THENOT, S., SANG, H. M., STERN, C. D., SAMARUT, J. & PAIN, B. 2007. The Oct4 homologue PouV and Nanog regulate pluripotency in chicken embryonic stem cells. *Development*, 134, 3549-63.
- LE BIN, G. C., MUNOZ-DESCALZO, S., KUROWSKI, A., LEITCH, H., LOU, X., MANSFIELD, W., ETIENNE-DUMEAU, C., GRABOLE, N., MULAS, C., NIWA, H., HADJANTONAKIS, A. K. & NICHOLS, J. 2014. Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst. *Development*, 141, 1001-10.
- LEE, M. T., BONNEAU, A. R., TAKACS, C. M., BAZZINI, A. A., DIVITO, K. R., FLEMING, E. S. & GIRALDEZ, A. J. 2013. Nanog, *Pou5f1* and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. *Nature*, 503, 360-4.
- LEE, S. G., MIKHALCHENKO, A. E., YIM, S. H., LOBANOV, A. V., PARK, J. K., CHOI, K. H., BRONSON, R. T., LEE, C. K., PARK, T. J. & GLADYSHEV, V. N. 2017. Naked Mole Rat Induced Pluripotent Stem Cells and Their Contribution to Interspecific Chimera. *Stem Cell Reports*, 9, 1706-1720.
- LEICHSENTRING, M., MAES, J., MOSSNER, R., DRIEVER, W. & ONICHTCHOUK, D. 2013. *Pou5f1* transcription factor controls zygotic gene activation in vertebrates. *Science*, 341, 1005-9.
- LI, L., DONG, J., YAN, L., YONG, J., LIU, X., HU, Y., FAN, X., WU, X., GUO, H., WANG, X., ZHU, X., LI, R., YAN, J., WEI, Y., ZHAO, Y., WANG, W., REN, Y., YUAN, P., YAN, Z., HU, B., GUO, F., WEN, L., TANG, F. & QIAO, J. 2017. Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. *Cell Stem Cell*, 20, 858-873 e4.
- LI, M. & BELMONTE, J. C. 2017. Ground rules of the pluripotency gene regulatory network. *Nat Rev Genet*, 18, 180-191.
- LI, M. & IZPISUA BELMONTE, J. C. 2018. Deconstructing the pluripotency gene regulatory network. *Nat Cell Biol*, 20, 382-392.
- LIAO, J., CUI, C., CHEN, S., REN, J., CHEN, J., GAO, Y., LI, H., JIA, N., CHENG, L., XIAO, H. & XIAO, L. 2009. Generation of induced pluripotent stem cell lines from adult rat cells. *Cell Stem Cell*, 4, 11-5.
- LING, Y., WEST, A. G., ROBERTS, E. C., LAKEY, J. H. & SHARROCKS, A. D. 1998. Interaction of transcription factors with serum response factor. Identification of the Elk-1 binding surface. *J Biol Chem*, 273, 10506-14.
- LIU, T., LI, J., YU, L., SUN, H. X., LI, J., DONG, G., HU, Y., LI, Y., SHEN, Y., WU, J. & GU, Y. 2021. Cross-species single-cell transcriptomic analysis

- reveals pre-gastrulation developmental differences among pigs, monkeys, and humans. *Cell Discov*, 7, 8.
- LOH, Y. H., WU, Q., CHEW, J. L., VEGA, V. B., ZHANG, W., CHEN, X., BOURQUE, G., GEORGE, J., LEONG, B., LIU, J., WONG, K. Y., SUNG, K. W., LEE, C. W., ZHAO, X. D., CHIU, K. P., LIPOVICH, L., KUZNETSOV, V. A., ROBSON, P., STANTON, L. W., WEI, C. L., RUAN, Y., LIM, B. & NG, H. H. 2006. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet*, 38, 431-40.
- MAESO, I., ACEMEL, R. D. & GOMEZ-SKARMETA, J. L. 2017. Cis-regulatory landscapes in development and evolution. *Curr Opin Genet Dev*, 43, 17-22.
- MAESO, I. & TENA, J. J. 2016. Favorable genomic environments for cis-regulatory evolution: A novel theoretical framework. *Semin Cell Dev Biol*, 57, 2-10.
- MARAIS, R., WYNNE, J. & TREISMAN, R. 1993. The SRF accessory protein Elk-1 contains a growth factor-regulated transcriptional activation domain. *Cell*, 73, 381-93.
- MARCHETTO, M. C. N., NARVAIZA, I., DENLI, A. M., BENNER, C., LAZZARINI, T. A., NATHANSON, J. L., PAQUOLA, A. C. M., DESAI, K. N., HERAI, R. H., WEITZMAN, M. D., YEO, G. W., MUOTRI, A. R. & GAGE, F. H. 2013. Differential L1 regulation in pluripotent stem cells of humans and apes. *Nature*, 503, 525-529.
- MASSAGUE, J. 2012. TGFbeta signalling in context. *Nat Rev Mol Cell Biol*, 13, 616-30.
- MATO PRADO, M., FRAMPTON, A. E., STEBBING, J. & KRELL, J. 2015. Gene of the month: NANOG. *J Clin Pathol*, 68, 763-5.
- MATSUMOTO, K., ANZAI, M., NAKAGATA, N., TAKAHASHI, A., TAKAHASHI, Y. & MIYATA, K. 1994. Onset of paternal gene activation in early mouse embryos fertilized with transgenic mouse sperm. *Mol Reprod Dev*, 39, 136-40.
- MEDVINSKY, A. & DZIERZAK, E. 1996. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell*, 86, 897-906.
- MEILHAC, S. M., ADAMS, R. J., MORRIS, S. A., DANCKAERT, A., LE GARREC, J. F. & ZERNICKA-GOETZ, M. 2009. Active cell movements coupled to positional induction are involved in lineage segregation in the mouse blastocyst. *Dev Biol*, 331, 210-21.
- MENZOROV, A. G., MATVEEVA, N. M., MARKAKIS, M. N., FISHMAN, V. S., CHRISTENSEN, K., KHABAROVA, A. A., PRISTYAZHNYUK, I. E., KIZILOVA, E. A., CIRERA, S., ANISTOROAELI, R. & SEROV, O. L. 2015. Comparison of American mink embryonic stem and induced pluripotent stem cell transcriptomes. *BMC Genomics*, 16 Suppl 13, S6.
- MESNARD, D., GUZMAN-AYALA, M. & CONSTAM, D. B. 2006. Nodal specifies embryonic visceral endoderm and sustains pluripotent cells in the epiblast before overt axial patterning. *Development*, 133, 2497-505.
- MESSERSCHMIDT, D. M. & KEMLER, R. 2010. Nanog is required for primitive endoderm formation through a non-cell autonomous mechanism. *Dev Biol*, 344, 129-37.
- MIHAJLOVIC, A. I. & BRUCE, A. W. 2017. The first cell-fate decision of mouse preimplantation embryo development: integrating cell position and polarity. *Open Biol*, 7.

- MILLER, D. J., BALL, E. E. & TECHNAU, U. 2005. Cnidarians and ancestral genetic complexity in the animal kingdom. *Trends Genet*, 21, 536-9.
- MITSUI, K., TOKUZAWA, Y., ITOH, H., SEGAWA, K., MURAKAMI, M., TAKAHASHI, K., MARUYAMA, M., MAEDA, M. & YAMANAKA, S. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, 113, 631-42.
- MIZUSEKI, K., KISHI, M., MATSUI, M., NAKANISHI, S. & SASAI, Y. 1998. Xenopus Zic-related-1 and Sox-2, two factors induced by chordin, have distinct activities in the initiation of neural induction. *Development*, 125, 579-87.
- MO, X., LI, N. & WU, S. 2014. Generation and characterization of bat-induced pluripotent stem cells. *Theriogenology*, 82, 283-93.
- MOLE, M. A., COORENS, T. H. H., SHAHBAZI, M. N., WEBERLING, A., WEATHERBEE, B. A. T., GANTNER, C. W., SANCHO-SERRA, C., RICHARDSON, L., DRINKWATER, A., SYED, N., ENGLE, S., SNELL, P., CHRISTIE, L., ELDER, K., CAMPBELL, A., FISHEL, S., BEHJATI, S., VENTO-TORMO, R. & ZERNICKA-GOETZ, M. 2021. A single cell characterisation of human embryogenesis identifies pluripotency transitions and putative anterior hypoblast centre. *Nat Commun*, 12, 3679.
- MONTE, D., CLANTIN, B., DEWITTE, F., LENS, Z., RUCKTOO, P., PARDON, E., STEYAERT, J., VERGER, A. & VILLERET, V. 2018. Crystal structure of human Mediator subunit MED23. *Nat Commun*, 9, 3389.
- MORRIS, S. A., GUO, Y. & ZERNICKA-GOETZ, M. 2012. Developmental plasticity is bound by pluripotency and the Fgf and Wnt signaling pathways. *Cell Rep*, 2, 756-65.
- MORRISON, G. M. & BRICKMAN, J. M. 2006. Conserved roles for Oct4 homologues in maintaining multipotency during early vertebrate development. *Development*, 133, 2011-2022.
- MULLIN, N. P., YATES, A., ROWE, A. J., NIJMEIJER, B., COLBY, D., BARLOW, P. N., WALKINSHAW, M. D. & CHAMBERS, I. 2008. The pluripotency rheostat Nanog functions as a dimer. *Biochem J*, 411, 227-31.
- MURAKAMI, K., GUNESDOGAN, U., ZYLICZ, J. J., TANG, W. W. C., SENGUPTA, R., KOBAYASHI, T., KIM, S., BUTLER, R., DIETMANN, S. & SURANI, M. A. 2016. NANOG alone induces germ cells in primed epiblast in vitro by activation of enhancers. *Nature*, 529, 403-407.
- NAKAMURA, T., OKAMOTO, I., SASAKI, K., YABUTA, Y., IWATANI, C., TSUCHIYA, H., SEITA, Y., NAKAMURA, S., YAMAMOTO, T. & SAITOU, M. 2016a. A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature*, 537, 57-62.
- NAKAMURA, T., YABUTA, Y., OKAMOTO, I., SASAKI, K., IWATANI, C., TSUCHIYA, H. & SAITOU, M. 2017. Single-cell transcriptome of early embryos and cultured embryonic stem cells of cynomolgus monkeys. *Sci Data*, 4, 170067.
- NAKAMURA, Y., DE PAIVA ALVES, E., VEENSTRA, G. J. & HOPPLER, S. 2016b. Tissue- and stage-specific Wnt target gene expression is controlled subsequent to beta-catenin recruitment to cis-regulatory modules. *Development*, 143, 1914-25.
- NAKANO, S. & AGATA, K. 2019. Evolutionary view of pluripotency seen from early development of non-mammalian amniotes. *Dev Biol*, 452, 95-103.

- NAKANO, S., FUSE, N., TADOKORO, R., TAKAHASHI, Y. & AGATA, K. 2017. Jak1/Stat3 signaling acts as a positive regulator of pluripotency in chicken pre-gastrula embryos. *Dev Biol*, 421, 43-51.
- NAKANO, S., FUSE, N., TAKAHASHI, Y. & AGATA, K. 2015. Verification of chicken Nanog as an epiblast marker and identification of chicken PouV as Pou5f3 by newly raised antibodies. *Dev Growth Differ*, 57, 251-63.
- NAKANO, S., OKAZAKI, K. & AGATA, K. 2013. Inhibition of MEK and GSK3 supports ES cell-like domed colony formation from avian and reptile embryos. *Zoolog Sci*, 30, 543-52.
- NENTWICH, O., DINGWELL, K. S., NORDHEIM, A. & SMITH, J. C. 2009. Downstream of FGF during mesoderm formation in *Xenopus*: the roles of Elk-1 and Egr-1. *Dev Biol*, 336, 313-26.
- NICHOLS, J., CHAMBERS, I. & SMITH, A. 1994. Derivation of germline competent embryonic stem cells with a combination of interleukin-6 and soluble interleukin-6 receptor. *Exp Cell Res*, 215, 237-9.
- NICHOLS, J., CHAMBERS, I., TAGA, T. & SMITH, A. 2001. Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines. *Development*, 128, 2333-9.
- NICHOLS, J. & SMITH, A. 2009. Naive and primed pluripotent states. *Cell Stem Cell*, 4, 487-92.
- NICHOLS, J., ZEVNIK, B., ANASTASSIADIS, K., NIWA, H., KLEWENEBENIUS, D., CHAMBERS, I., SCHOLER, H. & SMITH, A. 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell*, 95, 379-91.
- NIEDZWIEDZKI, G., SZREK, P., NARKIEWICZ, K., NARKIEWICZ, M. & AHLBERG, P. E. 2010. Tetrapod trackways from the early Middle Devonian period of Poland. *Nature*, 463, 43-8.
- NIEUWKOOP, P. D. 1969. The Formation of the Mesoderm in Urodelean Amphibians I. Induction by the Endoderm. *Wilhelm Roux Archiv Fur Entwicklungsmechanik Der Organismen*, 162, 341-373.
- NISHINAKAMURA, R., MATSUMOTO, Y., MATSUDA, T., ARIIZUMI, T., HEIKE, T., ASASHIMA, M. & YOKOTA, T. 1999. Activation of Stat3 by cytokine receptor gp130 ventralizes *Xenopus* embryos independent of BMP-4. *Dev Biol*, 216, 481-90.
- NIWA, H., MIYAZAKI, J. & SMITH, A. G. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet*, 24, 372-6.
- NIWA, H., OGAWA, K., SHIMOSATO, D. & ADACHI, K. 2009. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature*, 460, 118-22.
- NIWA, H., TOYOOKA, Y., SHIMOSATO, D., STRUMPF, D., TAKAHASHI, K., YAGI, R. & ROSSANT, J. 2005. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell*, 123, 917-29.
- NOWOSHILOW, S., SCHLOISSNIG, S., FEI, J. F., DAHL, A., PANG, A. W. C., PIPPEL, M., WINKLER, S., HASTIE, A. R., YOUNG, G., ROSCITO, J. G., FALCON, F., KNAPP, D., POWELL, S., CRUZ, A., CAO, H., HABERMANN, B., HILLER, M., TANAKA, E. M. & MYERS, E. W. 2018. The axolotl genome and the evolution of key tissue formation regulators. *Nature*, 554, 50-55.

- OSTEIL, P., TAPPONNIER, Y., MARKOSSIAN, S., GODET, M., SCHMALTZ-PANNEAU, B., JOUNEAU, L., CABAU, C., JOLY, T., BLACHERE, T., GOCZA, E., BERNAT, A., YERLE, M., ACLOQUE, H., HIDOT, S., BOSZE, Z., DURANTHON, V., SAVATIER, P. & AFANASSIEFF, M. 2013. Induced pluripotent stem cells derived from rabbits exhibit some characteristics of naïve pluripotency. *Biol Open*, 2, 613-28.
- PAGES, G., GUERIN, S., GRALL, D., BONINO, F., SMITH, A., ANJUERE, F., AUBERGER, P. & POUYSSEGUR, J. 1999. Defective thymocyte maturation in p44 MAP kinase (Erk 1) knockout mice. *Science*, 286, 1374-7.
- PAIN, B., CHENEVIER, P. & SAMARUT, J. 1999. Chicken embryonic stem cells and transgenic strategies. *Cells Tissues Organs*, 165, 212-9.
- PAIN, B., CLARK, M. E., SHEN, M., NAKAZAWA, H., SAKURAI, M., SAMARUT, J. & ETCHES, R. J. 1996. Long-term in vitro culture and characterisation of avian embryonic stem cells with multiple morphogenetic potentialities. *Development*, 122, 2339-48.
- PALMIERI, S. L., PETER, W., HESS, H. & SCHOLER, H. R. 1994. Oct4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation. *Dev Biol*, 166, 259-67.
- PAN, H. & SCHULTZ, R. M. 2011. Sox2 modulates reprogramming of gene expression in two-cell mouse embryos. *Biol Reprod*, 85, 409-16.
- PERA, M. F. & HERSZFELD, D. 1998. Differentiation of human pluripotent teratocarcinoma stem cells induced by bone morphogenetic protein-2. *Reprod Fertil Dev*, 10, 551-5.
- PERRETT, R. M., TURNPENNY, L., ECKERT, J. J., O'SHEA, M., SONNE, S. B., CAMERON, I. T., WILSON, D. I., RAJPERT-DE MEYTS, E. & HANLEY, N. A. 2008. The early human germ cell lineage does not express SOX2 during in vivo development or upon in vitro culture. *Biol Reprod*, 78, 852-8.
- PESCE, M. & SCHOLER, H. R. 2001. Oct4: gatekeeper in the beginnings of mammalian development. *Stem Cells*, 19, 271-8.
- PETITTE, J. N., CLARK, M. E., LIU, G., VERRINDER GIBBINS, A. M. & ETCHES, R. J. 1990. Production of somatic and germline chimeras in the chicken by transfer of early blastodermal cells. *Development*, 108, 185-9.
- PIJUAN-SALA, B., GRIFFITHS, J. A., GUIBENTIF, C., HISCOCK, T. W., JAWAID, W., CALERO-NIETO, F. J., MULAS, C., IBARRA-SORIA, X., TYSER, R. C. V., HO, D. L. L., REIK, W., SRINIVAS, S., SIMONS, B. D., NICHOLS, J., MARIONI, J. C. & GOTTGENS, B. 2019. A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature*, 566, 490-495.
- PILATO, G., D'URSO, V., VIGLIANISI, F., SAMMARTANO, F., SABELLA, G. & LISI, O. 2013. The problem of the origin of primordial germ cells (PGCs) in vertebrates: historical review and a possible solution. *Int J Dev Biol*, 57, 809-19.
- PINYOPUMMIN, A., TAKAHASHI, Y., HISHINUMA, M. & KANAGAWA, H. 1994. Development of single blastomeres from 4-cell stage embryos after aggregation with parthenogenones in mice. *Jpn J Vet Res*, 42, 119-26.
- PIOTROWSKA-NITSCHKE, K., PEREA-GOMEZ, A., HARAGUCHI, S. & ZERNICKA-GOETZ, M. 2005. Four-cell stage mouse blastomeres have different developmental properties. *Development*, 132, 479-90.

- PLUSA, B. & HADJANTONAKIS, A. K. 2014. Embryonic stem cell identity grounded in the embryo. *Nat Cell Biol*, 16, 502-4.
- PLUSA, B., PILISZEK, A., FRANKENBERG, S., ARTUS, J. & HADJANTONAKIS, A. K. 2008. Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development*, 135, 3081-91.
- POSFAL, E., PETROPOULOS, S., DE BARROS, F. R. O., SCHELL, J. P., JURISICA, I., SANDBERG, R., LANNER, F. & ROSSANT, J. 2017. Position- and Hippo signaling-dependent plasticity during lineage segregation in the early mouse embryo. *Elife*, 6.
- PRISE, I. & SHARROCKS, A. D. 2019. ELK1 has a dual activating and repressive role in human embryonic stem cells. *Wellcome Open Res*, 4, 41.
- RAMASWAMY, K., YIK, W. Y., WANG, X. M., OLIPHANT, E. N., LU, W., SHIBATA, D., RYDER, O. A. & HACIA, J. G. 2015. Derivation of induced pluripotent stem cells from orangutan skin fibroblasts. *BMC Res Notes*, 8, 577.
- RAMOS-IBEAS, P., SANG, F., ZHU, Q., TANG, W. W. C., WITHEY, S., KLISCH, D., WOOD, L., LOOSE, M., SURANI, M. A. & ALBERIO, R. 2019. Pluripotency and X chromosome dynamics revealed in pig pre-gastrulating embryos by single cell analysis. *Nat Commun*, 10, 500.
- REBEIZ, M. & TSIANTIS, M. 2017. Enhancer evolution and the origins of morphological novelty. *Curr Opin Genet Dev*, 45, 115-123.
- REUBINOFF, B. E., PERA, M. F., FONG, C. Y., TROUNSON, A. & BONGSO, A. 2000. Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nat Biotechnol*, 18, 399-404.
- ROSSANT, J. 1976. Postimplantation development of blastomeres isolated from 4- and 8-cell mouse eggs. *J Embryol Exp Morphol*, 36, 283-90.
- RUTHENBURG, A. J., ALLIS, C. D. & WYSOCKA, J. 2007. Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. *Mol Cell*, 25, 15-30.
- SABA-EL-LEIL, M. K., VELLA, F. D., VERNAY, B., VOISIN, L., CHEN, L., LABRECQUE, N., ANG, S. L. & MELOCHE, S. 2003. An essential function of the mitogen-activated protein kinase Erk2 in mouse trophoblast development. *EMBO Rep*, 4, 964-8.
- SAUNDERS, A., FAIOLA, F. & WANG, J. 2013. Concise review: pursuing self-renewal and pluripotency with the stem cell factor Nanog. *Stem Cells*, 31, 1227-36.
- SAXTON, J., FERJENTSCHIK, Z., DUCKER, C., JOHNSON, A. D. & SHAW, P. E. 2016. Stepwise evolution of Elk-1 in early deuterostomes. *FEBS J*, 283, 1025-38.
- SCERBO, P., GIRARDOT, F., VIVIEN, C., MARKOV, G. V., LUXARDI, G., DEMENEIX, B., KODJABACHIAN, L. & COEN, L. 2012. Ventx factors function as Nanog-like guardians of developmental potential in *Xenopus*. *PLoS One*, 7, e36855.
- SCERBO, P., MARKOV, G. V., VIVIEN, C., KODJABACHIAN, L., DEMENEIX, B., COEN, L. & GIRARDOT, F. 2014. On the origin and evolutionary history of NANOG. *PLoS One*, 9, e85104.
- SCHOLER, H. R. 1991. Octamania: the POU factors in murine development. *Trends Genet*, 7, 323-9.

- SCHOLER, H. R., BALLING, R., HATZOPOULOS, A. K., SUZUKI, N. & GRUSS, P. 1989. Octamer binding proteins confer transcriptional activity in early mouse embryogenesis. *EMBO J*, 8, 2551-7.
- SCHUFF, M., SIEGEL, D., PHILIPP, M., BUNDSCHU, K., HEYMAN, N., DONOW, C. & KNOCHEL, W. 2012. Characterization of *Danio rerio* Nanog and functional comparison to *Xenopus laevis*. *Stem Cells Dev*, 21, 1225-38.
- SCHULZ, K. N. & HARRISON, M. M. 2019. Mechanisms regulating zygotic genome activation. *Nat Rev Genet*, 20, 221-234.
- SHAHBAZI, M. N., JEDRUSIK, A., VUORISTO, S., RECHER, G., HUPALOWSKA, A., BOLTON, V., FOGARTY, N. N. M., CAMPBELL, A., DEVITO, L., ILIC, D., KHALAF, Y., NIAKAN, K. K., FISHEL, S. & ZERNICKA-GOETZ, M. 2016. Self-organization of the human embryo in the absence of maternal tissues. *Nat Cell Biol*, 18, 700-708.
- SHAHBAZI, M. N. & ZERNICKA-GOETZ, M. 2018. Deconstructing and reconstructing the mouse and human early embryo. *Nat Cell Biol*, 20, 878-887.
- SHI, Q. Q., SUN, M., ZHANG, Z. T., ZHANG, Y. N., ELSAYED, A. K., ZHANG, L., HUANG, X. M. & LI, B. C. 2014. A screen of suitable inducers for germline differentiation of chicken embryonic stem cells. *Anim Reprod Sci*, 147, 74-85.
- SHIMADA, H., NAKADA, A., HASHIMOTO, Y., SHIGENO, K., SHIONOYA, Y. & NAKAMURA, T. 2010. Generation of canine induced pluripotent stem cells by retroviral transduction and chemical inhibitors. *Mol Reprod Dev*, 77, 2.
- SHOOK, D. R. & KELLER, R. 2008. Epithelial type, ingression, blastopore architecture and the evolution of chordate mesoderm morphogenesis. *J Exp Zool B Mol Dev Evol*, 310, 85-110.
- SHOOK, D. R., MAJER, C. & KELLER, R. 2002. Urodeles remove mesoderm from the superficial layer by subduction through a bilateral primitive streak. *Dev Biol*, 248, 220-39.
- SHORE, P. & SHARROCKS, A. D. 1994. The transcription factors Elk-1 and serum response factor interact by direct protein-protein contacts mediated by a short region of Elk-1. *Mol Cell Biol*, 14, 3283-91.
- SHUBIN, N., TABIN, C. & CARROLL, S. 2009. Deep homology and the origins of evolutionary novelty. *Nature*, 457, 818-23.
- SILVA, J., BARRANDON, O., NICHOLS, J., KAWAGUCHI, J., THEUNISSEN, T. W. & SMITH, A. 2008. Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol*, 6, e253.
- SILVA, J., NICHOLS, J., THEUNISSEN, T. W., GUO, G., VAN OOSTEN, A. L., BARRANDON, O., WRAY, J., YAMANAKA, S., CHAMBERS, I. & SMITH, A. 2009. Nanog is the gateway to the pluripotent ground state. *Cell*, 138, 722-37.
- SIVAK, J. M., PETERSEN, L. F. & AMAYA, E. 2005. FGF signal interpretation is directed by Sprouty and Spred proteins during mesoderm formation. *Dev Cell*, 8, 689-701.
- SMITH, A. G., HEATH, J. K., DONALDSON, D. D., WONG, G. G., MOREAU, J., STAHL, M. & ROGERS, D. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature*, 336, 688-90.

- SMITH, J. C. & MALACINSKI, G. M. 1983. The origin of the mesoderm in an anuran, *Xenopus laevis*, and a urodele, *Ambystoma mexicanum*. *Dev Biol*, 98, 250-4.
- STAVRIDIS, M. P., LUNN, J. S., COLLINS, B. J. & STOREY, K. G. 2007. A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. *Development*, 134, 2889-94.
- STEVENS, J. L., CANTIN, G. T., WANG, G., SHEVCHENKO, A., SHEVCHENKO, A. & BERK, A. J. 2002. Transcription control by E1A and MAP kinase pathway via Sur2 mediator subunit. *Science*, 296, 755-8.
- STRAHL, T., GILLE, H. & SHAW, P. E. 1996. Selective response of ternary complex factor Sap1a to different mitogen-activated protein kinase subgroups. *Proc Natl Acad Sci U S A*, 93, 11563-8.
- STREIT, A., SOCKANATHAN, S., PEREZ, L., REX, M., SCOTTING, P. J., SHARPE, P. T., LOVELL-BADGE, R. & STERN, C. D. 1997. Preventing the loss of competence for neural induction: HGF/SF, L5 and Sox-2. *Development*, 124, 1191-202.
- STREIT, A. & STERN, C. D. 1999. Establishment and maintenance of the border of the neural plate in the chick: involvement of FGF and BMP activity. *Mech Dev*, 82, 51-66.
- SUZUKI, S., NOZAWA, Y., TSUKAMOTO, S., KANEKO, T., MANABE, I., IMAI, H. & MINAMI, N. 2015. CHD1 acts via the Hmgpi pathway to regulate mouse early embryogenesis. *Development*, 142, 2375-84.
- SWIERS, G., CHEN, Y. H., JOHNSON, A. D. & LOOSE, M. 2010. A conserved mechanism for vertebrate mesoderm specification in urodele amphibians and mammals. *Dev Biol*, 343, 138-52.
- TAKAHASHI, K., TANABE, K., OHNUKI, M., NARITA, M., ICHISAKA, T., TOMODA, K. & YAMANAKA, S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131, 861-72.
- TAKAHASHI, K. & YAMANAKA, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126, 663-76.
- TAKASHIMA, Y., GUO, G., LOOS, R., NICHOLS, J., FICZ, G., KRUEGER, F., OXLEY, D., SANTOS, F., CLARKE, J., MANSFIELD, W., REIK, W., BERTONE, P. & SMITH, A. 2014. Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell*, 158, 1254-1269.
- TAKEDA, J., SEINO, S. & BELL, G. I. 1992. Human Oct3 gene family: cDNA sequences, alternative splicing, gene organization, chromosomal location, and expression at low levels in adult tissues. *Nucleic Acids Res*, 20, 4613-20.
- TAKEDA, K., NOGUCHI, K., SHI, W., TANAKA, T., MATSUMOTO, M., YOSHIDA, N., KISHIMOTO, T. & AKIRA, S. 1997. Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. *Proc Natl Acad Sci U S A*, 94, 3801-4.
- TAM, P. P. & ZHOU, S. X. 1996. The allocation of epiblast cells to ectodermal and germ-line lineages is influenced by the position of the cells in the gastrulating mouse embryo. *Dev Biol*, 178, 124-32.
- TAMADA, H., VAN THUAN, N., REED, P., NELSON, D., KATOKU-KIKYO, N., WUDEL, J., WAKAYAMA, T. & KIKYO, N. 2006. Chromatin decondensation and nuclear reprogramming by nucleoplasmin. *Mol Cell Biol*, 26, 1259-71.

- TARKOWSKI, A. K. 1959. Experiments on the development of isolated blastomers of mouse eggs. *Nature*, 184, 1286-7.
- TARKOWSKI, A. K. & WROBLEWSKA, J. 1967. Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage. *J Embryol Exp Morphol*, 18, 155-80.
- TECHNAU, U. & SCHOLZ, C. B. 2003. Origin and evolution of endoderm and mesoderm. *Int J Dev Biol*, 47, 531-9.
- TESAR, P. J., CHENOWETH, J. G., BROOK, F. A., DAVIES, T. J., EVANS, E. P., MACK, D. L., GARDNER, R. L. & MCKAY, R. D. 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature*, 448, 196-9.
- THEUNISSEN, T. W., POWELL, B. E., WANG, H., MITALIPOVA, M., FADDAH, D. A., REDDY, J., FAN, Z. P., MAETZEL, D., GANZ, K., SHI, L., LUNJANGWA, T., IMSOONTHORNRUKSA, S., STELZER, Y., RANGARAJAN, S., D'ALESSIO, A., ZHANG, J., GAO, Q., DAWLATY, M. M., YOUNG, R. A., GRAY, N. S. & JAENISCH, R. 2014. Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell*, 15, 471-487.
- THISSE, B. & THISSE, C. 2005. Functions and regulations of fibroblast growth factor signaling during embryonic development. *Dev Biol*, 287, 390-402.
- THOMSON, M., LIU, S. J., ZOU, L. N., SMITH, Z., MEISSNER, A. & RAMANATHAN, S. 2011. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell*, 145, 875-89.
- TOMIOKA, I., MAEDA, T., SHIMADA, H., KAWAI, K., OKADA, Y., IGARASHI, H., OIWA, R., IWASAKI, T., AOKI, M., KIMURA, T., SHIOZAWA, S., SHINOHARA, H., SUEMIZU, H., SASAKI, E. & OKANO, H. 2010. Generating induced pluripotent stem cells from common marmoset (*Callithrix jacchus*) fetal liver cells using defined factors, including Lin28. *Genes Cells*, 15, 959-69.
- TORRES-PADILLA, M. E., PARFITT, D. E., KOUZARIDES, T. & ZERNICKA-GOETZ, M. 2007. Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature*, 445, 214-8.
- TROUILLAS, M., SAUCOURT, C., GUILLOTIN, B., GAUTHEREAU, X., DING, L., BUCHHOLZ, F., DOSS, M. X., SACHINIDIS, A., HESCHELER, J., HUMMEL, O., HUEBNER, N., KOLDE, R., VILO, J., SCHULZ, H. & BOEUF, H. 2009a. Three LIF-dependent signatures and gene clusters with atypical expression profiles, identified by transcriptome studies in mouse ES cells and early derivatives. *BMC Genomics*, 10, 73.
- TROUILLAS, M., SAUCOURT, C., GUILLOTIN, B., GAUTHEREAU, X., TAUPIN, J. L., MOREAU, J. F. & BOEUF, H. 2009b. The LIF cytokine: towards adulthood. *Eur Cytokine Netw*, 20, 51-62.
- TYSER, R. C., MAHAMMADOV, E., NAKANOH, S., VALLIER, L., SCIALDONE, A. & SRINIVAS, S. 2020. A spatially resolved single cell atlas of human gastrulation. *bioRxiv*.
- UMBHAUER, M., MARSHALL, C. J., MASON, C. S., OLD, R. W. & SMITH, J. C. 1995. Mesoderm induction in *Xenopus* caused by activation of MAP kinase. *Nature*, 376, 58-62.
- VALLIER, L., ALEXANDER, M. & PEDERSEN, R. A. 2005. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci*, 118, 4495-509.

- VALLIER, L., MENDJAN, S., BROWN, S., CHNG, Z., TEO, A., SMITHERS, L. E., TROTTER, M. W., CHO, C. H., MARTINEZ, A., RUGG-GUNN, P., BRONS, G. & PEDERSEN, R. A. 2009a. Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development*, 136, 1339-49.
- VALLIER, L., TOUBOUL, T., CHNG, Z., BRIMPARI, M., HANNAN, N., MILLAN, E., SMITHERS, L. E., TROTTER, M., RUGG-GUNN, P., WEBER, A. & PEDERSEN, R. A. 2009b. Early cell fate decisions of human embryonic stem cells and mouse epiblast stem cells are controlled by the same signalling pathways. *PLoS One*, 4, e6082.
- VASTENHOEW, N. L., ZHANG, Y., WOODS, I. G., IMAM, F., REGEV, A., LIU, X. S., RINN, J. & SCHIER, A. F. 2010. Chromatin signature of embryonic pluripotency is established during genome activation. *Nature*, 464, 922-6.
- VERMA, R., HOLLAND, M. K., TEMPLE-SMITH, P. & VERMA, P. J. 2012. Inducing pluripotency in somatic cells from the snow leopard (*Panthera uncia*), an endangered felid. *Theriogenology*, 77, 220-8, 228 e1-2.
- VILLIARD, E., DENIS, J. F., HASHEMI, F. S., IGELMANN, S., FERBEYRE, G. & ROY, S. 2017. Senescence gives insights into the morphogenetic evolution of amniotes. *Biol Open*, 6, 891-896.
- WAGNER, D. E., WEINREB, C., COLLINS, Z. M., BRIGGS, J. A., MEGASON, S. G. & KLEIN, A. M. 2018. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science*, 360, 981-987.
- WANG, H. & DEY, S. K. 2006. Roadmap to embryo implantation: clues from mouse models. *Nat Rev Genet*, 7, 185-99.
- WANG, J., LEVASSEUR, D. N. & ORKIN, S. H. 2008. Requirement of Nanog dimerization for stem cell self-renewal and pluripotency. *Proc Natl Acad Sci USA*, 105, 6326-31.
- WANG, J., RAO, S., CHU, J., SHEN, X., LEVASSEUR, D. N., THEUNISSEN, T. W. & ORKIN, S. H. 2006. A protein interaction network for pluripotency of embryonic stem cells. *Nature*, 444, 364-8.
- WANG, W., HUANG, L., HUANG, Y., YIN, J. W., BERK, A. J., FRIEDMAN, J. M. & WANG, G. 2009. Mediator MED23 links insulin signaling to the adipogenesis transcription cascade. *Dev Cell*, 16, 764-71.
- WANG, Z., ORON, E., NELSON, B., RAZIS, S. & IVANOVA, N. 2012. Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell*, 10, 440-54.
- WEERATUNGA, P., SHAHSAVARI, A., OVCHINNIKOV, D. A., WOLVETANG, E. J. & WHITWORTH, D. J. 2018. Induced Pluripotent Stem Cells from a Marsupial, the Tasmanian Devil (*Sarcophilus harrisi*): Insight into the Evolution of Mammalian Pluripotency. *Stem Cells Dev*, 27, 112-122.
- WEGNER, M. 2010. All purpose Sox: The many roles of Sox proteins in gene expression. *Int J Biochem Cell Biol*, 42, 381-90.
- WEINA, K. & UTIKAL, J. 2014. SOX2 and cancer: current research and its implications in the clinic. *Clin Transl Med*, 3, 19.
- WEINBERGER, L., AYYASH, M., NOVERSHTERN, N. & HANNA, J. H. 2016. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat Rev Mol Cell Biol*, 17, 155-69.
- WOOD, H. B. & EPISKOPOU, V. 1999. Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages. *Mech Dev*, 86, 197-201.

- WU, G. & SCHOLER, H. R. 2014. Role of Oct4 in the early embryo development. *Cell Regen*, 3, 7.
- WUNDERLICH, S., KIRCHER, M., VIETH, B., HAASE, A., MERKERT, S., BEIER, J., GOHRING, G., GLAGE, S., SCHAMBACH, A., CURNOW, E. C., PAABO, S., MARTIN, U. & ENARD, W. 2014. Primate iPS cells as tools for evolutionary analyses. *Stem Cell Res*, 12, 622-9.
- XU, C., FAN, Z. P., MULLER, P., FOGLEY, R., DIBIASE, A., TROMPOUKI, E., UNTERNAEHRER, J., XIONG, F., TORREGROZA, I., EVANS, T., MEGASON, S. G., DALEY, G. Q., SCHIER, A. F., YOUNG, R. A. & ZON, L. I. 2012. Nanog-like regulates endoderm formation through the Mxtx2-Nodal pathway. *Dev Cell*, 22, 625-38.
- XU, H., ZHU, X., LI, W., TANG, Z., ZHAO, Y. & WU, X. 2018. Isolation and in vitro culture of ovarian stem cells in Chinese soft-shell turtle (*Pelodiscus sinensis*). *J Cell Biochem*, 119, 7667-7677.
- YABUTA, Y., KURIMOTO, K., OHINATA, Y., SEKI, Y. & SAITOU, M. 2006. Gene expression dynamics during germline specification in mice identified by quantitative single-cell gene expression profiling. *Biol Reprod*, 75, 705-16.
- YAMAGUCHI, S., KIMURA, H., TADA, M., NAKATSUJI, N. & TADA, T. 2005. Nanog expression in mouse germ cell development. *Gene Expr Patterns*, 5, 639-46.
- YAN, L., YANG, M., GUO, H., YANG, L., WU, J., LI, R., LIU, P., LIAN, Y., ZHENG, X., YAN, J., HUANG, J., LI, M., WU, X., WEN, L., LAO, K., LI, R., QIAO, J. & TANG, F. 2013. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol*, 20, 1131-9.
- YANG, S. H., BUMPASS, D. C., PERKINS, N. D. & SHARROCKS, A. D. 2002. The ETS domain transcription factor Elk-1 contains a novel class of repression domain. *Mol Cell Biol*, 22, 5036-46.
- YANG, S. H., WHITMARSH, A. J., DAVIS, R. J. & SHARROCKS, A. D. 1998a. Differential targeting of MAP kinases to the ETS-domain transcription factor Elk-1. *EMBO J*, 17, 1740-9.
- YANG, S. H., YATES, P. R., WHITMARSH, A. J., DAVIS, R. J. & SHARROCKS, A. D. 1998b. The Elk-1 ETS-domain transcription factor contains a mitogen-activated protein kinase targeting motif. *Mol Cell Biol*, 18, 710-20.
- YAO, Y., LI, W., WU, J., GERMANN, U. A., SU, M. S., KUIDA, K. & BOUCHER, D. M. 2003. Extracellular signal-regulated kinase 2 is necessary for mesoderm differentiation. *Proc Natl Acad Sci U S A*, 100, 12759-64.
- YEOM, Y. I., HA, H. S., BALLING, R., SCHOLER, H. R. & ARTZT, K. 1991. Structure, expression and chromosomal location of the Oct4 gene. *Mech Dev*, 35, 171-9.
- YING, Q. L., WRAY, J., NICHOLS, J., BATLLE-MORERA, L., DOBLE, B., WOODGETT, J., COHEN, P. & SMITH, A. 2008. The ground state of embryonic stem cell self-renewal. *Nature*, 453, 519-23.
- YU, S. S. & CIRILLO, N. 2020. The molecular markers of cancer stem cells in head and neck tumors. *J Cell Physiol*, 235, 65-73.
- ZAEHRES, H., LENSCH, M. W., DAHERON, L., STEWART, S. A., ITSKOVITZ-ELDOR, J. & DALEY, G. Q. 2005. High-efficiency RNA interference in human embryonic stem cells. *Stem Cells*, 23, 299-305.

- ZAFARANA, G., AVERY, S. R., AVERY, K., MOORE, H. D. & ANDREWS, P. W. 2009. Specific knockdown of OCT4 in human embryonic stem cells by inducible short hairpin RNA interference. *Stem Cells*, 27, 776-82.
- ZEINEDDINE, D., HAMMOUD, A. A., MORTADA, M. & BOEUF, H. 2014. The Oct4 protein: more than a magic stemness marker. *Am J Stem Cells*, 3, 74-82.
- ZHANG, H. M., LI, L., PAPADOPOULOU, N., HODGSON, G., EVANS, E., GALBRAITH, M., DEAR, M., VOUGIER, S., SAXTON, J. & SHAW, P. E. 2008. Mitogen-induced recruitment of ERK and MSK to SRE promoter complexes by ternary complex factor Elk-1. *Nucleic Acids Res*, 36, 2594-607.
- ZHANG, L., WU, Y., LI, X., WEI, S., XING, Y., LIAN, Z. & HAN, H. 2018. An Alternative Method for Long-Term Culture of Chicken Embryonic Stem Cell In Vitro. *Stem Cells Int*, 2018, 2157451.
- ZHAO, R., ZUO, Q., YUAN, X., JIN, K., JIN, J., DING, Y., ZHANG, C., LI, T., JIANG, J., LI, J., ZHANG, M., SHI, X., SUN, H., ZHANG, Y., XU, Q., CHANG, G., ZHAO, Z., LI, B., WU, X., ZHANG, Y., SONG, J., CHEN, G. & LI, B. 2021. Production of viable chicken by allogeneic transplantation of primordial germ cells induced from somatic cells. *Nat Commun*, 12, 2989.
- ZHAO, S., NICHOLS, J., SMITH, A. G. & LI, M. 2004. SoxB transcription factors specify neuroectodermal lineage choice in ES cells. *Mol Cell Neurosci*, 27, 332-42.
- ZHI, M., ZHANG, J., TANG, Q., YU, D., GAO, S., GAO, D., LIU, P., GUO, J., HAI, T., GAO, J., CAO, S., ZHAO, Z., LI, C., WENG, X., HE, M., CHEN, T., WANG, Y., LONG, K., JIAO, D., LI, G., ZHANG, J., LIU, Y., LIN, Y., PANG, D., ZHU, Q., CHEN, N., HUANG, J., CHEN, X., YAO, Y., YANG, J., XIE, Z., HUANG, X., LIU, M., ZHANG, R., LI, Q., MIAO, Y., TIAN, J., HUANG, X., OUYANG, H., LIU, B., XIE, W., ZHOU, Q., WEI, H., LIU, Z., ZHENG, C., LI, M. & HAN, J. 2021. Generation and characterization of stable pig pregastrulation epiblast stem cell lines. *Cell Res*.
- ZHU, Q., SANG, F., WITHEY, S., TANG, W., DIETMANN, S., KLISCH, D., RAMOS-IBEAS, P., ZHANG, H., REQUENA, C. E., HAJKOVA, P., LOOSE, M., SURANI, M. A. & ALBERIO, R. 2021. Specification and epigenomic resetting of the pig germline exhibit conservation with the human lineage. *Cell Rep*, 34, 108735.
- ZIMMERLIN, L., PARK, T. S., HUO, J. S., VERMA, K., PATHER, S. R., TALBOT, C. C., JR., AGARWAL, J., STEPPAN, D., ZHANG, Y. W., CONSIDINE, M., GUO, H., ZHONG, X., GUTIERREZ, C., COPE, L., CANTO-SOLER, M. V., FRIEDMAN, A. D., BAYLIN, S. B. & ZAMBIDIS, E. T. 2016. Tankyrase inhibition promotes a stable human naive pluripotent state with improved functionality. *Development*, 143, 4368-4380.
- ZINCK, R., HIPSKIND, R. A., PINGOUD, V. & NORDHEIM, A. 1993. c-fos transcriptional activation and repression correlate temporally with the phosphorylation status of TCF. *EMBO J*, 12, 2377-87.
- ZUO, Q., JIN, J., JIN, K., SUN, C., SONG, J., ZHANG, Y., CHEN, G. & LI, B. 2019. Distinct roles of retinoic acid and BMP4 pathways in the formation of chicken primordial germ cells and spermatogonial stem cells. *Food Funct*, 10, 7152-7163.