

Investigating the biochemical properties of two Nanog orthologues identified in the axolotl

Luke Alexander Simpson

The University of Nottingham

School of Life Sciences

Student No. 4267936

Supervisor: Prof. Andrew Johnson

Second Supervisor: Prof. Ramiro Alberio

Thesis submitted to the University of Nottingham for the degree of Master of Research, 30 April 2017

Academic Year: 2016 – 2017

Contents

Acknowledgements **5**

Abstract **6**

1. Introduction **7**
1.1 The role of Nanog and SMAD transcription factors in early development 7
1.2 Oct4 and Sox2 in early development 16
1.3 Cellular reprogramming 18
1.4 Epigenetic control of gene expression and changes during reprogramming 21
1.4.1 Histone phosphorylation 23
1.4.2 Histone acetylation 23
1.4.3 Histone methylation 25
1.4.4 Spatial territories, heterochromatin and euchromatin 26
1.4.5 TADS and sub-TADs 27
1.4.6 Chromatin loops 28
1.5 Aims and objectives **31**

2. Materials & Methods **34**
2.1 Chemicals and Reagents 34
2.2 Methodology 34
2.2.1 Cell Lines and oocytes 34
2.2.2 Cell culture 34
2.2.2.1 Mammalian cell culture 34
2.2.2.2 Cryopreservation of cell lines 35
2.2.3 Creation of plasmids 35
2.2.3.1 Creation of NanoBiT Vectors 35
2.2.3.2 Creation of Gal4 Vectors 38
2.2.3.3 Creation of PSIN Vectors 38
2.2.4 Plasmid transformation 38
2.2.5 Plasmid mini-preparation 38
2.2.6 Establishment of Gal4 system cell lines 39
2.2.7 Transient transfection of NanoBiT constructs 41
2.2.8 Validation of antibodies 43
2.2.8.1 *In vitro* translation of proteins 43
2.2.8.1.1 Rabbit reticulocyte lysate *in vitro* translation 43
2.2.8.1.2 Wheat germ lysate *in vitro* translation 44
2.2.8.2 Production of proteins in *Xenopus* oocytes 44
2.2.8.3 Production of expression cell lines 44
2.2.8.4 Antibodies 44
2.2.8.5 Dot Blots 44
2.2.8.6 Western blotting 45
2.2.8.6.1 Polyacrylamide gel electrophoresis 45
2.2.8.6.2 Semi-dry transfer of proteins 46
2.2.8.7 Immunoprecipitation 47
2.2.9 Axolotl and *Xenopus* oocyte extract treatments 47

2.2.10 ChIP-QPCR 48
2.2.10.1 *In vivo* cross-linking 48
2.2.10.2 Cell and nuclear lysis 49
2.2.10.3 Chromatin shearing 49
2.2.10.4 Chromatin immunoprecipitation 50
2.2.10.5 Quantitative PCR 50
2.2.10.5.1 Primer design 50
2.2.10.5.2 QPCR conditions 51
2.2.11 Eomesodermin transcriptional activation assay 52

3. Results 53
3.1 Testing EggNog and Nanog antibodies 53
3.2. Detection of EggNog and Nanog proteins in the AOE 56
3.3. Biochemical properties of axNanog and EggNog 58
3.4. Reprogramming with amphibian oocyte extracts 65

4. Discussion 76

References 83

Appendix 90

Acknowledgements

There are many people who have supported me during my mRes. I would like to take this opportunity to thank Professor Johnson and all the members of his team for their support and guidance throughout this project. In particular I would like to thank Professor Johnson for the opportunity to study as part of his team and the fascinating project he has given me. I would also like to thank Darren Crowley for teaching me a host of important skills and for the countless hours of discussions. Thank you to Jodie Chatfield, Teri Forey, Matt Loose and Helena Acosta for helping me at every turn with their expertise and providing a welcoming, enjoyable working atmosphere.

Abstract

Pluripotent cells give rise to the germ line and the soma. The expression of the Nanog orthologue axNanog is required to establish pluripotency during axolotl development and has a conserved role interacting with axSMAD2 and DPY30 to deposit H3K4me3 through COMPASS. Transcriptome analysis has revealed a second Nanog orthologue: EggNog. EggNog possesses a nearly identical intron- exon structure and yet has a profoundly different role to axNanog, acting to suppress primordial germ cell specification. We aimed to identify whether axNanog and EggNog exhibit different translational regulation or biochemical properties. We also aimed to further define AOE reprogramming of mammalian cells. AOE was probed using western blotting for the presence of axNanog and EggNog proteins. We explored the biochemical properties of axNanog and EggNog using a series of luciferase assays. RT-qPCR and ChIP-qPCR were used to investigate the changes to gene expression and chromatin structure of cells treated with AOE. Luciferase assays demonstrate that axNanog forms stable interactions with axSMAD2 while EggNog shows a reduced affinity to axSMAD2. Western blotting demonstrates that axNanog but not EggNog is present within the axolotl oocyte as a protein. Finally, we show that cells treated with AOE show increased expression of Nanog and Sox2. These data demonstrate that axNanog and EggNog differ in their post-transcriptional regulation and biochemical properties which likely contributes to their differing roles in the developing axolotl.

1. Introduction

1.1 The Role of Nanog and SMAD transcription factors in early development

The homeodomain-containing transcription factor Nanog is critical in early vertebrate development and has a conserved role as a master regulator of pluripotency (Dixon *et al*, 2010). In mice, Nanog transcripts are present within the inner cells of the late morula and throughout formation and expansion of the inner cell mass (ICM) (Mitsui *et al*, 2003; Chambers *et al*, 2003; Dietrich & Hiragi, 2007). Nanog-null morphants are able to establish an ICM; however, Nanog is indispensable to the establishment and subsequent expansion of pluripotent epiblast (Mitsui *et al*, 2003; Silva *et al*, 2009). Thus, *in vivo* Nanog is required for the maintenance and expansion of pluripotent cells.

Nanog was initially identified by Chambers and colleagues (2003) due to its role in mouse embryonic stem cells (mESCs) *in vitro*. Yamanka and colleagues performed digital differential display analysis identifying *Nanog* by comparing gene expression data from somatic tissues and mouse ES cells. Further investigation identified that *Nanog* maintains pluripotency in mESCs and native cells within the ICM independent of leukaemia inhibitory factor (LIF) (Mitsui *et al*, 2003). A summary of the main pathways involved in pluripotency can be seen in figure 1. *In vitro*, mESCs require a specific cell culture 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 leukaemia inhibitory factor (LIF), when it is 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 Klf4 and Tbx3, respectively. Klf4 drives expression of the transcription factor Sox2 while Tbx3 promotes Nanog expression (Saunders *et al*, 2013). BMP4 signalling induces the phosphorylation of SMAD proteins 1, 5 and 8, resulting in the upregulation of inhibitor of differentiation (Id) genes, which inhibit the expression of numerous genes associated with differentiation to promote

pluripotency. Expression of pluripotency genes and repression of differentiation associated genes maintains a positive auto-regulatory loop between Nanog, Sox2 and Oct4 to maintain pluripotency (Mitsui *et al*, 2003; Boyer *et al*, 2005; Niwa *et al*, 2009, Saunders *et al*, 2013).

While Nanog was initially identified and characterised in its role in mESC pluripotency, the pluripotency network of mESC has evolved adaptations which to date, have not been observed in other vertebrates including humans. However, cells isolated from the post-implantation mouse embryo just after implantation known as mouse epiblast stem cells (mEpiSCs) have a network governing pluripotency which more closely resembles Human ESCs (hESCs) than mESCs. In particular, BMP signalling mediated by SMADs 1 and 6 drives hESC and mEpiSC differentiation, which is in contrast to the role of BMP observed in mESCs (Saunders *et al*, 2013). Unlike mESC, both mEpiSCs and hESCs do not require LIF to maintain pluripotency in culture, instead they require insulin or insulin-like growth factor (IGF) and basic fibroblast growth factor (bFGF) signalling (Vallier, 2005; Bendall *et al*, 2007; Silva *et al*, 2008). In this context, bFGF activates the mitogen-activated protein kinase (MAPK) pathways while IGF activates the Ras and PI3K pathways. Both of these pathways lead to a sustained state of pluripotency.

Like mESC, both mEpiSCs and hESC require Activin/Nodal signalling to sustain pluripotency, propagated through the effector SMAD2/3 (Saunders *et al*, 2013). The mechanism by which Activin/Nodal signalling promotes pluripotency was clearly established when the TGF-β RI kinase inhibitor SB431542 was shown to prevent SMAD2/3 phosphorylation. The inactive SMAD2/3 directly resulted in the downregulation of Nanog and Oct4 expression (James *et al*, 2005). Later work directly identified SMAD2/3 consensus sequences within the human 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*, 2009; Greber *et al,* 2010). Activin/Nodal signalling is also associated with mesendodermal differentiation and inhibition of neuroectodermal specification through SMAD2/3 transcriptional regulation (Vallier *et al*, 2005; Brown *et al*, 2011).

In a landmark study Bertero *et al* (2015) demonstrated that Nanog, SMAD2/3 and DPY30 proteins physically interact in human pluripotent stem cells (hPSCs). Furthermore, the Nanog-SMAD2/3 complex recruits DPY30 a subunit of the complex of proteins associated with Set1 (COMPASS) in response to Nodal signalling. COMPASS is a conserved family of proteins that function to maintain specific patterns of gene expression throughout cellular development. COMPASS has known H3K4 methyltransferase activity. When it is complexed with Nanog, SMAD2 and DPY30 COMPASS facilitates trimethylation of H3K4 residues in histones bound to nodal-responsive genes (Bertero *et al*, 2015). Previously, Wang *et al* (2006) created a Nanog interactome, which shows Nanog has a complex network of interactions in mESCs. The interactome highlights Nanogs physical interaction with known co-repressors like the SWI/SNF, NuRD and Polycomb complexes. The Polycomb complexes themselves have been associated with the deposition of H3K27me3 (Schuettengruber *et al*, 2007). Thus it appears that Nanog also has a primary function in depositing both H3K4me3 and H3K27me3 marks associated with transcriptional activation and repression respectively. The presence of both activating and repressive marks creates a 'bivalent' epigenetic state, which has been hypothesised to allow rapid activation or silencing of specific genes during differentiation (Bernstein *et al*, 2006). Altogether it would appear that Nanog functions to maintain pluripotency and to deposit epigenetic marks which will orchestrate cell fate decisions upon differentiation. Both of these key roles is performed through interactions with SMAD2/3 in human ESC (Vallier *et al*, 2009a; Bertero *et al*, 2015).

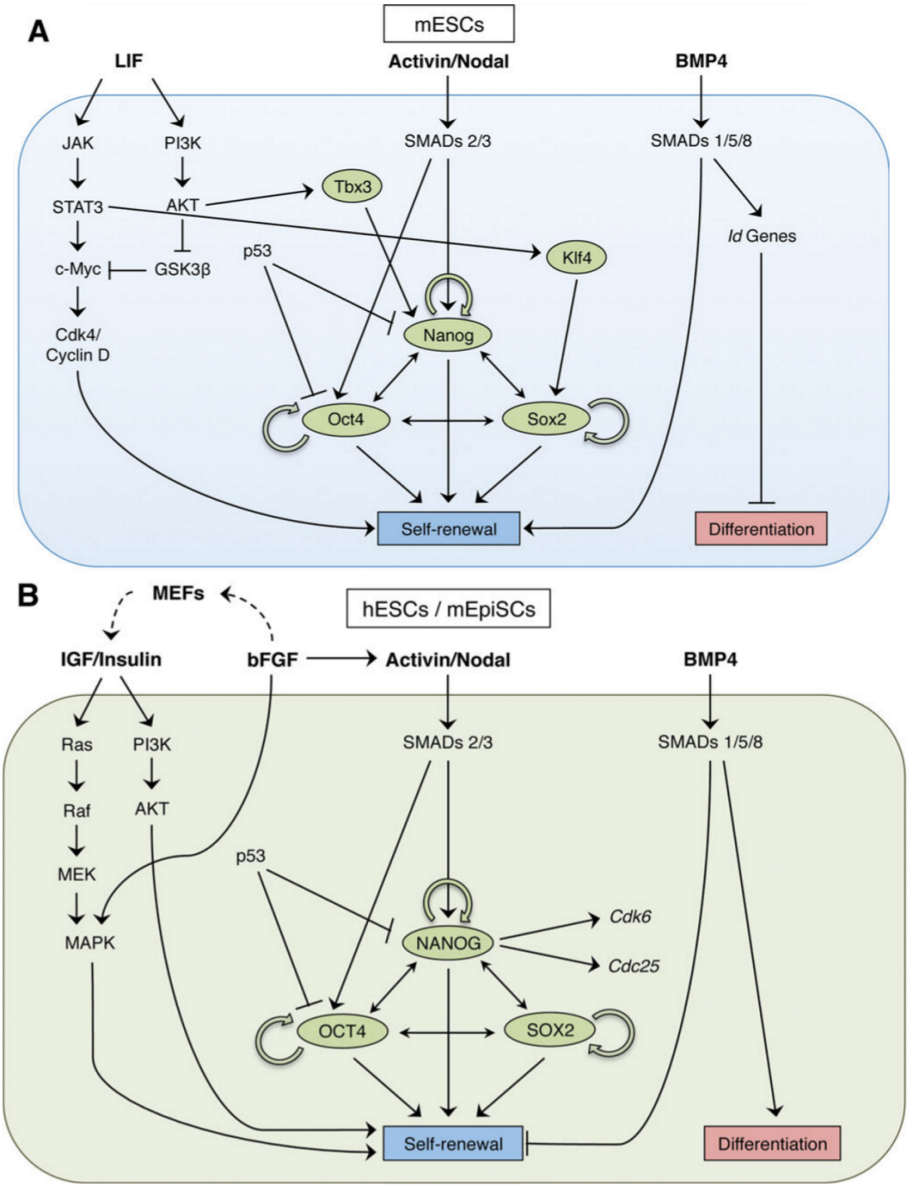


Figure 1. **Showing the main signalling mechanisms involved in pluripotency in mESC, hESC and mEpiSC pluripotency** .Main mouse and human pluripotency pathways. (A) MESCs require LIF and BMP4 for maintenance. (B) Human ESCs and mouse EpiSCs require IGF/insulin and bFGF for maintenance. Human ESC-derived fibroblast-like cells and MEFs are also stimulated by bFGF in culture to secrete IGF (dashed lines). In both cell types, Nanog, Oct4, and Sox2 form a positive auto-regulatory loop. Image taken from Saunders *et al*, (2013).

Key to the functionality of mammalian 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 with alanine resulted in an inability to dimerize. Curiously, the ability of Nanog to dimerise was demonstrated to be a requirement for self-renewal and pluripotency *in vitro* (Wang *et al*, 2008). Co-immunoprecipitations (Co-IP) also revealed that specific factors including Oct4, 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 key pluripotency factors' expression (Wang *et al*, 2008; Kim *et al*, 2008; Fidalgo *et al*, 2011). Intriguingly, the ability of the Nanog protein to dimerise does not appear to be conserved across vertebrates (Dixon *et al*, 2010).

The core mechanisms governing pluripotency are conserved across vertebrates, as evidenced by a high degree of conservation between urodele amphibians and mammals. A Nanog orthologue, axNanog has been identified by Johnson and colleagues in the urodele salamander *Ambystoma mexicanum* (axolotl). AxNanog, was discovered from RNA isolated from the animal hemisphere (caps) of blastula stage axolotl embryos. The greatest degree of homology between axNanog and mammalian Nanog genes is in the homeodomain (HD) and caspase cleavage site (CCS). Encoded by four exons, axNanog has an intron/exon structure which is also conserved in mammalian Nanog genes. Protein complementation assays demonstrated that unlike mammalian Nanog proteins, axNanog does not form a homodimer. Indeed, investigation into the biochemical properties of axNanog demonstrated that axNanog can drive the expression of mammalian Nanog and Rex-1 as a monomer. AxNanog was also demonstrated to enhance reprogramming when over-expressed in mESCs before fusion with neural stem cells (NSCs). The enhancement of reprogramming was comparable to mouse Nanog, demonstrating similar efficiencies (Dixon *et al*, 2010). Nanog orthologues have also been identified in Chick and Zebrafish. These Nanog orthologues also do not possess tryptophan repeat regions and cannot rescue self-renewal in Nanog null mESCs after LIF withdrawal. However, these same genes can substitute for mouse Nanog in somatic cell reprogramming (Theunissen, 2011). There is, however, evidence suggesting Zebrafish Nanog dimerises via a differing mechanism so it is unclear whether this may be the case for Chick Nanog (Schuff *et al*, 2012). AxNanog retains the ability to physically interact with Oct4 in protein complementation assays (PCAs) despite acting as a monomer; the induced monomeric form of mouse Nanog, however, does not retain this ability. Johnson and colleagues also demonstrated that dimerization was not essential in mouse or axolotl Nanog for DNA binding or transcriptional

activation. However, like mouse and human Nanogs, AxNanog dimerization was necessary and sufficient to rescue self-renewal of mESC in the absence of LIF. Thus, monomeric and dimeric forms of Nanog have some differing roles, as Nanog does not dimerize in lower vertebrates the functional importance of Nanog dimerization is still not fully understood. Moreover, it is unclear whether Nanog preferentially interacts with SMAD2/3 and DPY30 as a dimer in mammals.

Unpublished data by Johnson and colleagues (Figure 2A and B) has demonstrated that axNanog, like its mammalian counterparts, exhibits control of the epigenetic marks H3Kme3 and H3K27ac. This demonstrates that orchestrating cell fate decisions is a function of all vertebrate Nanogs. The team disrupted the Nodal signalling pathway with SB431542 inhibition, which resulted in the loss of H3Kme3 and H3K27ac epigenetic marks, as determined by Western blotting and immunofluorescence. It was then hypothesised that like mammalian Nanogs, axNanog likely orchestrates the deposition of H3Kme3 and H3K27ac through an association with DPY30 and COMPASS. Johnson and colleagues then used morpholinos to knockdown either axNanog or DPY30 and this also resulted in the loss of H3Kme3 and H3K27ac. These findings suggest that Nodal signalling and a complex between axNanog and DPY30 are required for the deposition of H3Kme3 and H3K27ac in axolotl embryos, and this observation mirrors results obtained by Bertero *et al* (2015) who inhibited Nodal signalling in hESC. Protein complementation assays (Figure 4A) have also demonstrated that axNanog can physically interact with SMAD2/3. Altogether, this suggests that axNanog may have a conserved role in regulating H3Kme3 and H3K27ac deposition through its interactions with SMAD2/3, DPY30 and COMPASS orthologues. In the absence of Nodal signal axolotl embryos arrest development at mid-blastula stage and fail to undergo lineage specification, mimicking the effects of Nanog knockdown. This suggests that axNanog plays a role in the establishment and maintenance of pluripotency as well as poising genes for lineage specification through the deposition of H3Kme3 and H3K27ac in the developing axolotl.

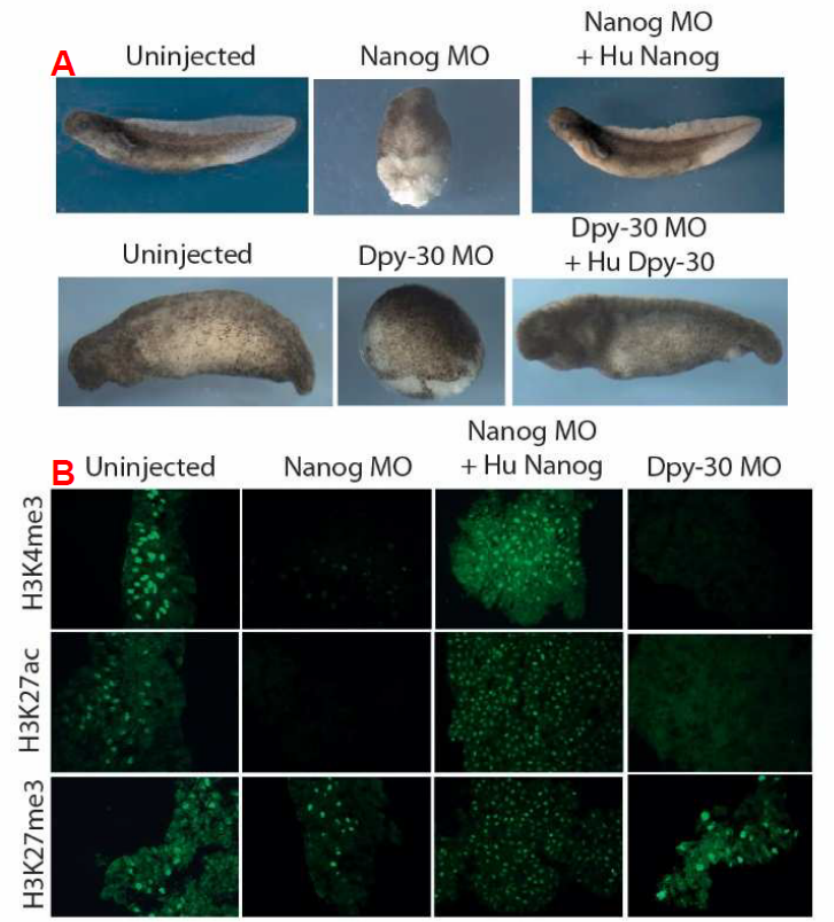


Figure 2. AxNanog and DPY30 knockdown and rescue data (A) The Nanog Interaction with COMPASS Initiates Embryogenesis as embryos fail to develop. This phenotype is observed in both Nanog and DPY30 knockdowns. Both phenotypes can be rescued by human RNA. (B) Depletion of Nanog or DPY30 inhibits the deposition of transcriptionally active epigenetic marks (H3K4me3; H3K27AC) in embryo chromatin. Repressive marks (H3K27me3) are unaffected.

Interestingly, Johnson and colleagues have also identified a second Nanog orthologue in the axolotl. The orthologue was identified from egg mRNA and has been named EggNog. Curiously, unpublished data demonstrates drastically different roles for EggNog and axNanog in developing embryos. An EggNog morpholino successfully knocked down EggNog activity resulting in a phenotype profoundly different from that of an axNanog knockdown, i.e. morphants, embryo. EggNog morphants (Figure 3A) exhibited some loss of the ventral posterior tissue, but developed relatively normally, with the exception of the mesodermal domains. The mesodermal domains usually produce the somites, the precursors of muscle and skeleton; however, EggNog morphants displayed structures resembling mesonephric ducts. These structures continued across the entirety of the anterior-posterior axis

and contained cells with enlarged nuclei and reduced cytoplasm resembling primordial germ cells (PGCs). QPCR analysis (Figure 3B) revealed that indeed the EggNog morphants highly expressed the PGC marker: DeadEnd. These data suggest that EggNog supresses PGC development in axolotl in order to promote somite development. AxNanog and EggNog have nearly identical intron/exon structures (Figure 3C); however, they exhibit different patterns of RNA expression. EggNog is expressed as maternally inherited RNA and its levels decline during cleavage stages to undetectable levels by the onset of gastrulation; in contrast, axNanog is expressed at low levels in the egg and its expression peaks just before stage 11, the mid stage of gastrulation (Figure 3D). However it is unclear whether axNanog and EggNog are present as proteins within the oocyte or whether their translation is initiated during development.

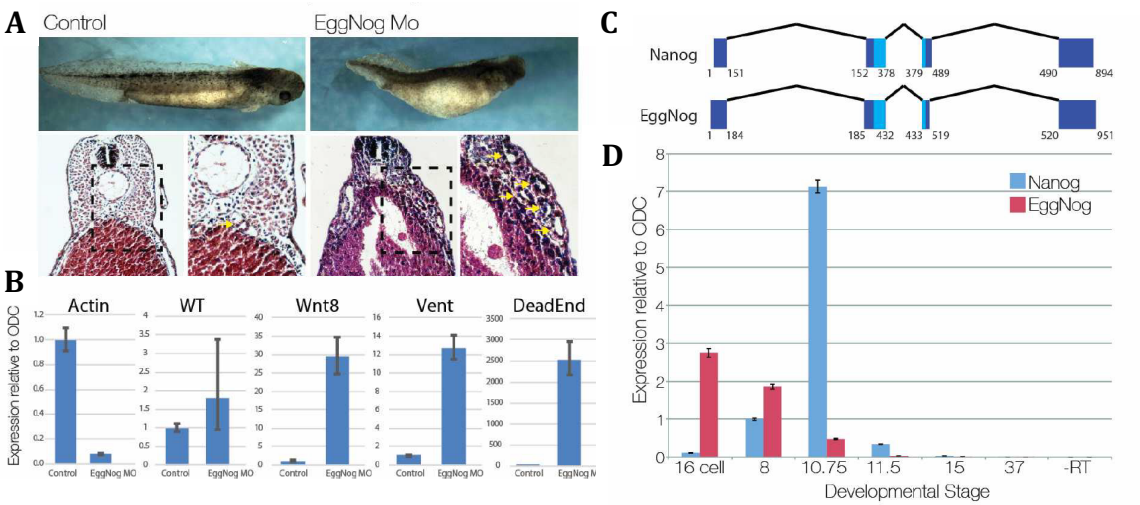


Figure 3. **Knockdown data, intron-exon structure and expression profiles of EggNog**. (A): EggNog morphants are hyper-ventralized. EggNog knockdown induces lateral mesoderm at the expense of somites (arrows show vessels containing PGC-like cells). (B): EggNog depletion induces upregulation of Wilms Tumor-1 (WT) at the expense of muscle (Actin), with upregulation of the Wnt8/VentX pathway and PGCs (DeadEnd) (C): EggNog and axNanog intron/exon structures. (D): RT-qPCR showing EggNog and axNanog expression profiles. EggNog RNA is maternally inherited and declines through early development. Nanog RNA is expressed zygotically during gastrulation.

Currently it is unknown whether AxNanog and EggNog share biochemical properties. It is therefore unclear whether the drastically different roles in development result from different biochemical properties or their different patterns of expression. As mentioned, it has been demonstrated that axNanog can physically interact with the axolotl SMAD2 protein. Thus, it is unknown whether EggNog also exhibits a binding affinity for SMAD2. SMAD proteins themselves physically interact with numerous

transcription factors, which enables them to regulate the transcription of hundreds of genes simultaneously (Massagué *et al*, 2005). Whether EggNog displays differential interaction propensities to axNanog, despite similarities in primary structure, would be key to understanding their vastly differing roles in the embryo.

Johnson and colleagues have also shown that transcriptional activation by SMAD2 is reduced by axNanog over-expression (Figure 4B). Together with protein complementation data, suggests that axNanog could act as a rheostat of SMAD2 transcriptional activity wherein the formation of a SMAD2-axNanog heterodimer reduces the amount of free SMAD2 available for transcriptional activation. The axNanog-SMAD2 heterodimer is likely formed due to conserved Nanog functions which both require SMAD2. However it is unclear whether EggNog, can mimic axNanog's ability to modulate SMAD2 transcriptional activity. Determining this could point toward the existence of differential biochemical properties.

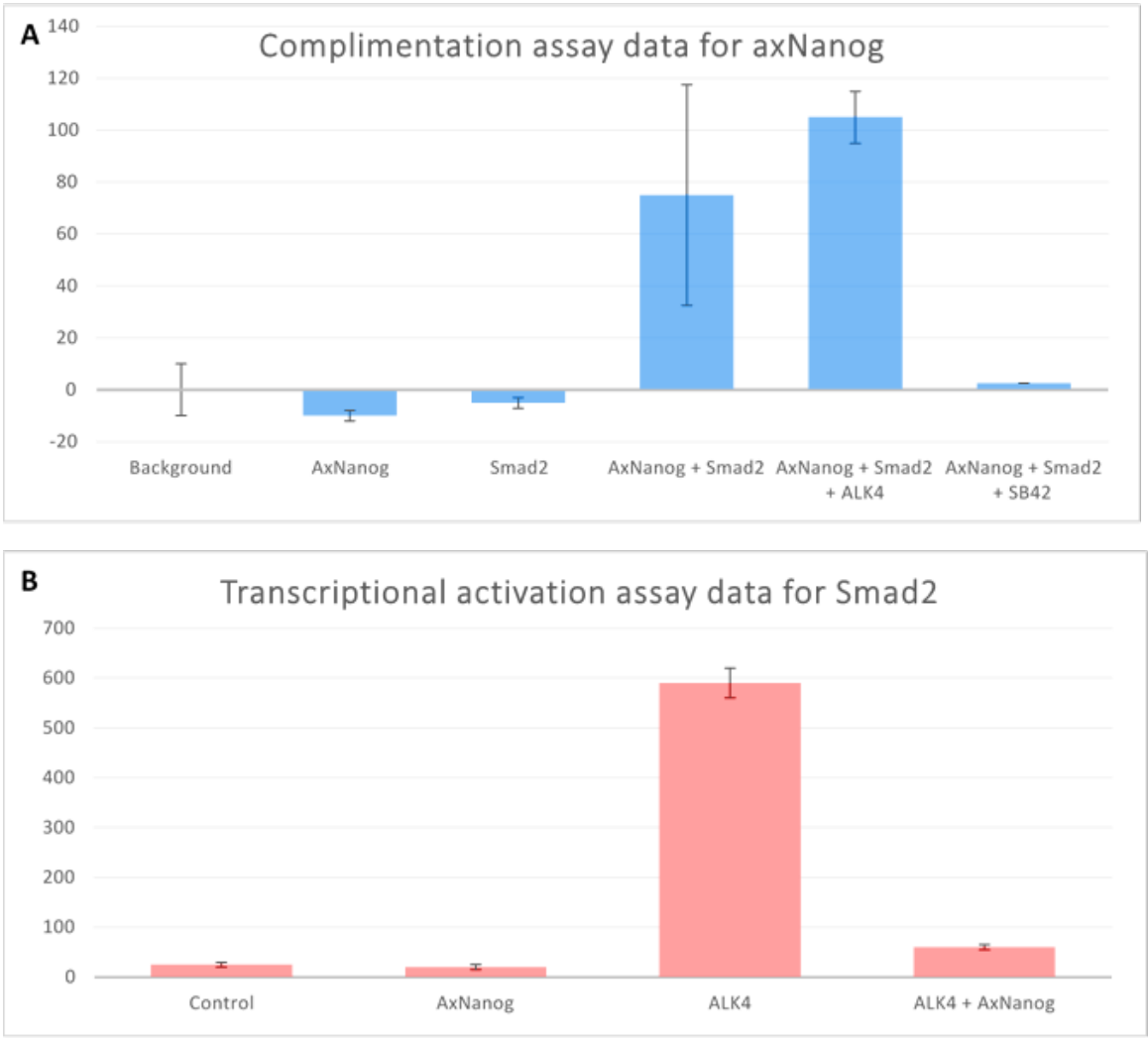


Figure 4. The biochemical properties of axNanog in two luciferase assays (A): Shows the binding affinity of axNanog to SMAD2 in the presence of constitutively active alkaline phosphatase 4 and SB42 inhibitor compared to control vectors where no binding partner is present. Binding affinity is measured in luminescence, increased luminescence is due to the presence of a functional luciferase enzyme forming upon PPI. (B): Shows transcriptional activation of gene encoding luciferase by SMAD2 with a GAL4 UAS fusion protein in the presence of ALK4 and axNanog overexpression compared to control in which no ALK4 is present.

1.2. Oct4 and Sox2 in early development

In vertebrates, Oct4 and Sox2 form part of the conserved pluripotency network along with Nanog, which establishes and maintains pluripotency both *in vivo* and *in vitro*. The Oct4 transcription factor is encoded by the pou5f1 gene and binds to the octamer motif ATGCAAAT in enhancer and promoter elements (Zeineddine *et al*, 2014). In mice, Oct4 is inherited by the zygote as mRNA and protein present in the

oocyte. Transcription of OCT4 begins at the 4-8 cell stage and is highly expressed in all blastomeres until morula stage when it is only found in the progenitors of the ICM (Palmieri *et al*, 1994; Pesce & Scholer, 2001; Jerabek *et al*, 2014). In Humans Oct4 is expressed in the oocyte and the developing zygote until the uncompacted morula (Abdel-Rahman *et al*, 1995; Verlinsky *et al*, 1998; Hansis *et al* 2000). Oct4 expression is required within the developing blastocyst of mice to repress Cdx2: a master regulator of trophectoderm development. Oct4 and Cdx2 mutually inhibit one another's expression, which promotes segregation of the ICM from the trophectoderm in the developing blastocyst. The Cdx2 expressing cells become the trophoblast while the Oct4 expressing cells later form the inner cell mass (Ralston & Rossant, 2005). Oct4 plays a similar role in human ES cells to downregulate the placental marker human chorionic gonadotropin (hCG) (Hay *et al*, 2004). Indeed embryo's lacking the Oct4 gene do not form ICM and die at implantation (Nichols *et al*, 1998). Oct4 is capable of binding to enhancers and promoting transcription of pluripotency associated genes including Sox2. Upon being translated Sox2 can form a heterodimer with Oct4, where it is then capable of binding to the enhancer of the NANOG gene to drive its transcription. This creates a feedforward circuit whereby the Oct4, Sox2 and Nanog bind to their respective enhancers and maintain their own expression, while simultaneously driving increased expression of other pluripotency associated genes through co-occupation of their respective enhancers. However, Oct4's role in development is dose-dependent. After implantation in mice, a subset of ICM cells increase Oct4 expression during the differentiation of primitive endoderm. This differentiation subsequently represses Oct4 transcription (Palmieri *et al*, 1994; Pesce & Scholer, 2001; Jerabek *et al*, 2014). Interestingly, increased Oct4 expression results in oct4-sox17 dimers perhaps indicating a rheostat activity of Sox2 in modulating Oct4-DNA interactions (Zeineddine *et al*, 2006; Stefanovic & Puceat, 2007; Stefanovic *et al*, 2009; Aksoy *et al*, 2013). During gastrulation Oct4 becomes further down-regulated in every cell type apart from primordial germ cells (Pesce & Scholer, 2001; Pan *et al*, 2001; Kehler *et al*, 2004). Oct4 homologs have been identified in monkeys, rats, pigs, bulls and goats (Zeineddine *et al*, 2014). It was demonstrated that Oct4's role to establish and maintain pluripotency is conserved across vertebrates as Oct4 homologues identified in axolotl were capable of rescuing

self-renewal in Oct4 -/- ES cell lines (Morrison & Brickman, 2006). Indeed the role of Oct4, Sox2 and Nanog is indispensable to establishing and maintaining pluripotency in vertebrate development and *in vitro*. Therefore they have also been implicated along with other key factors to establishing pluripotency during cellular reprogramming.

1.3 Cellular reprogramming

Reprogramming by somatic cell nuclear transfer (SCNT) has been used as a tool to investigate properties of the oocyte and cellular mechanisms relevant to early development. The successful transplantation of nuclei from donor blastula cells to enucleated host eggs was first achieved by Briggs and King in 1952 using *Rana pipens*. This conclusively demonstrated that the nuclei of the early blastula possessed the ability to direct development and the egg cytoplasm contained all necessary factors to initiate early development. Later work by John Gurdon and colleagues (1958) demonstrated the genomic equivalence of somatic and undifferentiated nuclei using *Xenopus laevis*, replacing the Egg nuclei with that of a somatic cell.. SCNT demonstrated that the matured *Xenopus* egg had all necessary factors to not only initiate development but also to remodel somatic nuclei to a state comparable to pluripotency. A method for producing mammalian blastocysts by SCNT later resulted in live offspring. This was the final proof of genomic equivalence, as mammalian development requires the formation of a trophectoderm and ICM, indicating that the somatic genome was being remodelled to a totipotent state (Wilmut *et al*, 1997; Wakayama *et al*, 1998). This work demonstrated that cytoplasmic factors within the oocyte, which normally govern nuclear reorganisation of the newly formed zygotic genome, are also able to reorganise somatic chromatin. Later work demonstrated that ESC lines could be produced from embryos derived from somatic cell nuclei- (Wakayama *et al*, 2001). Yet, despite some success, a large percentage of SCNT clones fail to reach full-term. Investigating why so many clones perish lead to advances in understanding several characteristic steps necessary for successful reprogramming. It was shown that successful SCNT-derived embryos underwent several stages of remodelling to become viable. These stages included the transcriptional silencing of donor nuclei, erasure of differentiated epigenetic

'cellular memory', the formation of the epigenetic 'embryo state', and this was required to achieve appropriate embryonic gene expression at all later stages. Unsuccessful SCNT clones often failed to complete these stages of remodelling. We now know the stages of cellular remodelling involve structural chromatin changes, covalent modification of histones, and in mammals, global DNA demethylation. Indeed most non-viable SCNT clones displayed aberrant epigenetic reprogramming (Reik *et al*, 2001; Wade & Kikyo, 2002; Dean *et al*, 2003).

While SCNT relies on non-heritable factors contained within the egg to remodel somatic nuclei, several other methods have been developed to remodel somatic nuclei to a pluripotent state. Most prominently, forced overexpression of defined groups of factors expressed in pluripotent stem cells, such as OCT4, SOX2, C-MYC and KLF4 (Takahashi and Yamanaka, 2006) or OCT4, SOX2, NANOG, and LIN28 (Thompson *et al*, 2007) can induce a pluripotent state in vitro. Defined factor reprogramming is useful to study the role of specific molecules involved in pluripotency; however in contrast to SCNT, these methods do not inform the mechanisms functioning in the oocyte.

Gurdon and colleagues showed that factors within amphibian oocytes could remodel the epigenome of mammalian nuclei. This was evidenced by demethylation of the *Oct4* promoter and increased *Oct4* expression in mammalian nuclei after injection into the germinal vesicle of *Xenopus* oocytes (Byrne *et al,* 2003; Simonsson & Gurdon, 2004).Furthermore XOE treatment of mammalian cells results in global reductions in H3K9me3 levels (Tamada *et al*, 2006). This conclusively demonstrated that factors within the amphibian oocyte can interact with the mammalian pluripotency network. Later work by Bian *et al* (2009) reported an ex-ovo system utilising axolotl oocyte extracts (AOE) that could induce targeted epigenetic modifications to the genome of somatic mammalian cells. In this system, permeabillized mouse embryonic fibroblasts (MEFs) were incubated in extracts made from mature oocytes of the axolotl supplemented with an energy generating system. Fluorescence activated cell sorting (FACS) of MEFs treated with AOE had global reductions in levels of repressive H3K9me3 much like cells treated in XOE.

Interestingly, HP1α levels were reduced to an intensity which more closely resembled mESC than control MEFs after treatment. 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. These data showed that the promoter of *Oct4* was significantly demethylated after just 5 hours incubation. Real-time qPCR later revealed that AOE treated cells showed increased Oct4 expression over control cells after 2 weeks in culture. Critically, just 5 hours of AOE treatment resulted constitutively increased Nanog expression. Altogether the data presented by Bian *et al* suggests that cells treated with AOE show an epigenetic profile which closely resembles that of pluripotent cells. While RT-qPCR verified directly some early transcription of pluripotency associated genes such as Nanog, the extent to which the AOE is able to remodel mammalian somatic cells toward pluripotency remains somewhat elusive. Unpublished ChIP qPCR data from our lab has compared the effects of reprogramming using AOE and XOE. These experiments showed AOE but not XOE induced activating marks H3K4me1 at the *Nanog* enhancer region and H3K27ac at both the *Nanog* promoter and enhancer. Critically, It appears that factors within axolotl oocytes but not Xenopus oocytes can remodel chromatin around the *Nanog* locus to facilitate *Nanog* expression. The axolotl Nanog orthologue axNanog has been demonstrated to be able to bind to mammalian promoters to drive transcription in a luciferase assay. Moreover, when constitutively expressed, axNanog was demonstrated to be able to bind to the native *Nanog* promoter (Dixon *et al*, 2010). Thus it would seem that axolotl pluripotency factors are highly compatible with the mammalian genome. This raises the question as to whether the axNanog protein binds to the mammalian Nanog promoter to drive Nanog expression during AOE reprogramming. *Xenopus* and axolotl represent two major groups of amphibians, Anurans and Urodeles respectfully. Urodele and Anuran amphibians diverged around 250 million years ago and while Urodeles appear to share a conserved vertebrate pluripotency network, Anurans have many unique adaptations to the network governing their early development. Perhaps most notably Anurans such as *Xenopus* lack Nanog or Oct4 orthologues. This suggests that the Anuran developmental network may have altered significantly since the

divergence of Anurans and Urodeles and perhaps explains why factors in Xenopus oocytes cannot reprogram mammalian cells to the same extent as those in axolotl oocytes (Johnson *et al*, 2003; Johnson *et al*, 2003b).

.

1.4 Epigenetic control of gene expression and changes during reprogramming

Changes in gene expression which occur during cellular reprogramming are a direct result of epigenetic changes to the genomic architecture, which alter a gene's transcriptional capacity. There are several 'layers' in which the genome is organised to facilitate control of gene expression. Perhaps the most basic level of regulation is the wrapping of DNA around histone and non-histone proteins into chromatin. This wrapping of DNA is pivotal to the transcriptional potential of any gene. Two copies of histones H2A, H2B, H3 and H4 form individual nucleosomes, while H1 histones act as linkers between nucleosomes. Histones are subject to differential post-translational modifications (PTMs) which alter the interaction of nucleosomal subunits with each other and with the DNA. There are two main hypotheses which aim to explain the mechanisms that regulate deposition of histone modifications. The first model suggests modifications to core histones attenuate the coulombic interactions between basic histone proteins and negative DNA. In essence, these modifications wrap or unwrap DNA, affecting its physical accessibility to non-histone proteins. The second, 'effector-mediated', mechanism posits that the histone modifications themselves directly interact with effector proteins, allowing or preventing downstream recruitment and binding of non-histone proteins to DNA (Taverna *et al*, 2007). PTMs themselves largely take place on the amino- and carboxy-terminal tails of individual histones. Histone PTMs include methylation, acetylation, phosphorylation, sumoylation, ADP-ribosylation, ubiquitination, deamination and proline isomerisation. Histone PTMs are produced rapidly, they are dynamic and in constant flux, catalysed by specific histone modifying enzymes (Kouzarides, 2007; Tessarz & Kouzarides, 2014). The histone-modifying enzymes themselves are subject to yet another layer of regulation, as there is evidence to suggest that association with co-factors may influence which residue is modified by a particular enzyme (Lee *et al*, 2005). Co-factors may also influence the degree of methylation imposed by a histone modifier (Steward *et al*, 2006). Some histone PTMs can appear within

minutes of a particular stimulus being presented to the cell surface, highlighting the rapid dynamics of histone-DNA interactions. Several modifications can occur on any given nucleosomal histone residue and modifications may interact with one another or with co-factors to create differential effects on an individual modification. The interactions of histone PTMs, or histone 'crosstalk', highlights the complexity of the histone code. Indeed, an individual histone mark at key loci, may not be indicative of the transcriptional state of a gene; rather, this must be considered in context. A multitude of data has elucidated some general properties of particular histone modifications and how they may influence the transcriptional potential of a gene. Histone PTMs play a vital role in the regulation of key genes in the acquisition and stabilisation of pluripotency during cellular reprogramming (Kouzarides, 2007; Tessarz & Kouzarides, 2014).

While there is a large body of research regarding these properties, this section will mainly cover methylation, acetylation and phosphorylation in the context of cellular reprogramming. Figure 5 shows H2A, H2B, H3 and H4 nucleosomal subunits and a multitude of well characterised histone modification sites.

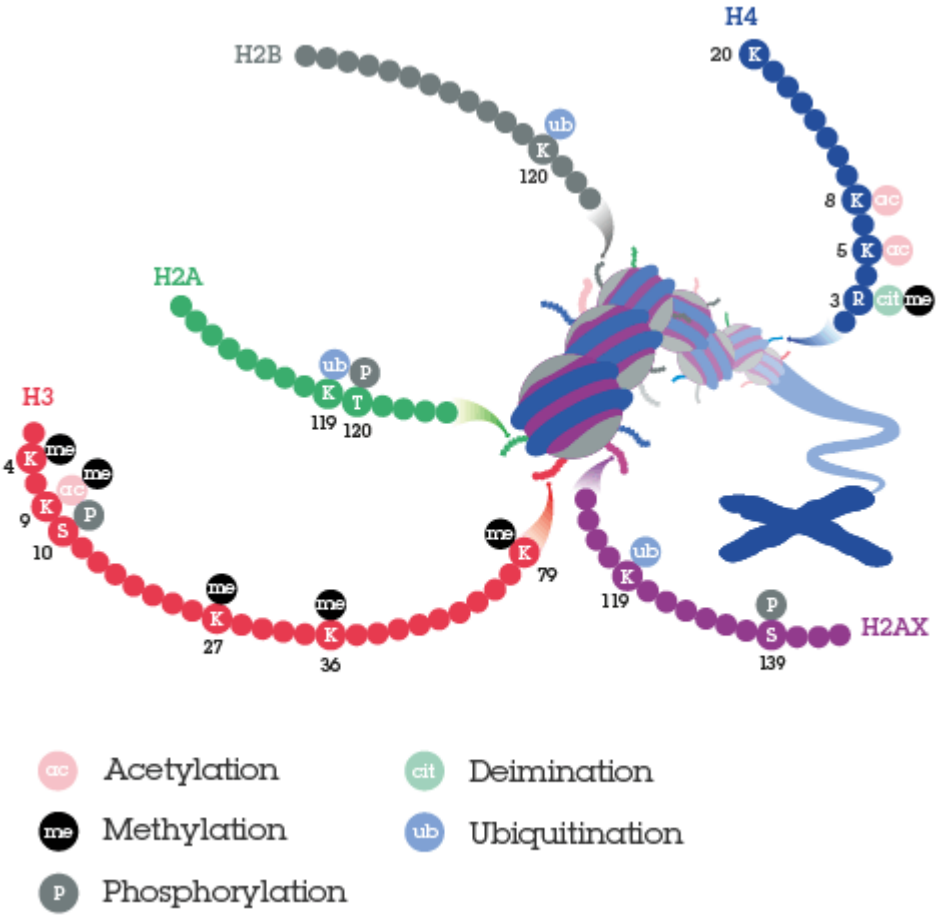


Figure 5. Common histone PTM sites. Showing Histones: H2A, H2AX, H2B and H4. Amino acid residues are numbered and labelled in standard single letter code. Key: ac, acetylation, me, methylation,

1.4.1 Histone Phosphorylation

Histone phosphorylation is a very abundant histone PTM and appears to function through its interactions with adjacent histone modifications, serving as a platform for effector proteins. This leads to a downstream cascade of events which can lead to differing effects on transcriptional activity. In essence histone phosphorylation can change the charge of an amino acid residue to facilitate further histone modifications either directly or indirectly (Wang *et al*, 2001; Nowak & Corces, 2004; Zippo *et al*, 2009).

1.4.2 Histone acetylation

Histone acetylation/deacetylation entails the addition or removal of an acetyl group on a lysine residue of a histone. Acetylation is carried out by histone acetyltransferases (HATs) which catalyse the addition of an acetyl group from acetyl

coenzyme A. Conversely, histone deacetylation is carried out by histone deacetylases (HDACs). Classically, it has been hypothesised that the acetylation of lysine residues reduces the electrostatic affinity between the histone and DNA, rendering the DNA more permissive to transcriptional activity (Roth *et al*, 2001). Furthermore, most active genes will feature some degree of histone acetylation within the gene coding region. In particular, enhancer and promoter regions generally display H3K9ac and H3K27ac marks. Recent research has highlighted that histone acetylation can often proceed or follow other histone PTMs to neighbouring histone tails leading to increased gene transcription (Lee *et al*, 2010). One study suggests that histone acetylation may occur as a result of interactions with co-factors and adjacent histones- to increase the transcriptional potential of a gene (Zippo *et al*, 2009). Zippo and colleagues also showed that acetylation of H4 tails recruited the RNA Pol II positive transcription elongation factor and that acetylation was dependent upon initial phosphorylation of H3 tails. Histone acetylation also plays a role in cellular reprogramming, as the increased acetylation of H3K9 residues was found in the nuclei of fibroblasts pre-treated with ooplasm before SCNT. The ooplasm pre-treatment greatly increased the efficiency of SCNT, and suggests that increased H3K9ac may facilitate stable reprogramming during SCNT (Bui *et al*, 2012). In the context of oocyte extract treatments, ChIP-qPCR data gathered by S Bereketoglu (unpublished) has shown that H3K27 residues around the Nanog enhancer regions become increasingly acetylated after only 3 hours of AOE treatment, compared to XOE treated cell or untreated cells. This demonstrates that the AOE is able to actively facilitate histone PTMs to facilitate transcription of the *Nanog* gene within only 3 hours of treatment.

In contrast to histone acetylation, histone deacetylation is almost universally associated with transcriptional repression. There are three distinct families of histone deacetylases, each of which has been linked in a multitude of signalling pathways present in numerous repressive chromatin complexes (Kouzarides, 2007).

1.4.3 Histone and DNA Methylation

Histone methylation occurs on all basic amino acid residues including arginines, lysines and histidines. In the context of gene regulation, arginine and lysine methylation appear to play a more significant role. Histones with arginine methylation are generally associated with actively transcribed genes but they make up a small minority of targeted residues. The vast majority of transcription-associated histone methylation occurs at lysine residues, catalysed by lysine methyltransferases (Greer & Shi, 2012). Lysine methyltransferases have a high level of specificity, usually modifying only a single lysine residue. Lysine methylation is associated with both transcriptional activation and repression depending on the residue modified and whether the residue is mono-, di- or tri-methylated (Bannister and Kouzarides, 2005). There are three particularly well characterised histone methylation sites associated with transcriptional activation: H3K4, H3K36, and H3K79. H3K4me and H3K36me in particular have been implicated in transcriptional elongation in budding yeast. Data suggests H3K4me3 localizes to the 5' of active genes and is associated with active RNA Polymerase II. Furthermore, H3K4me1 and H3Kme3 are well characterised markers of active gene enhancers and promoters respectively. Data shows an enrichment of H3K36me3 at transcribed regions of gene bodies. Moreover, H3K36me3 accumulates at the 3' end of active genes and associates with the elongating form of RNA polymerase II (Kouzarides, 2007; Greer & Shi, 2012). Data gathered by S Bereketoglu (unpublished) shows a significant enrichment of H3K4me1 at the NANOG enhancer region in NIH3T3 cells after 6 hours of AOE treatment. In contrast, NIH3T3 cells treated with XOE for 6 hours showed no significant enrichment of H3K4me1. Together, these results suggest factors within AOE facilitate the monomethylation of the H3K4 residues around the NANOG enhancer while factors within XOE do not.

There are three main lysine methylation sites associated with transcriptional repression: H3K9, H3K27, and H4K20. H3K27me3, in particular, is associated with promotor repression. While H3K9me is associated generally with transcriptional repression in heterochromatin it is also has a strong role in the silencing of genes in the euchromatin. Transcriptional repression involves the recruitment of methylating

enzymes and HP1 to the promoter of repressed genes. Delivery of these components of methylation-based silencing is a process which also requires co-repressors (Kouzarides, 2007). However, this is not always the case. As highlighted above, the context of histone PTMs ultimately determines the transcriptional potential of a gene. Research carried out by Vakoc and colleagues (2005) demonstrated that the classic marker of repression H3K9me, along with the g isoform of HP1, can be enriched in the coding region of active genes. It is therefore postulated that H3K9me2 and H3K9me3 within the coding regions is an activating mark, whereas H3K9me in promoter regions is repressive (Vakoc *et al*, 2005). However, until a definitive mechanism has been identified this remains unclear. In the context of reprogramming, reducing the methylation of H3K9 residues may play a role in facilitating stable SCNT reprogramming. Bui *et al* (2012) found that pre-treating cells with ooplasm greatly increased SCNT reprogramming efficiency and globally decreased H3K9 methylation. Nevertheless the repression of lineage-specific genes is also required for effective reprogramming, a process facilitated by histone methylation. Specifically, the repressive histone mark H3K27me3, deposited by Polycomb complex 2, has been found to be essential for ESC-fusion reprogramming of B cells (Montgomery *et al*, 2005; Pasini *et al*, 2007).

While histone modification forms a key component of genomic regulation so too does the alteration of the genome directly. DNA methylation or demethylation is the addition or removal of methyl groups to nucleotide bases within DNA. DNA methylation is catalysed by DNA methyltransferases while DNA demethylation can occur via passive loss through cellular divisions or through several active processes (Guo *et al*, 2011; Hacket *et al*, 2013). Global patterns of DNA methylation present after fertilisation are extinguished passively during development (Smallwood & Kelsey, 2012). 5-methylcytosine immunostaining of staged embryos demonstrated that global losses in DNA methylation are observed across multiple cellular divisions (Morgan *et al*, 2005). The importance of this activity during development was highlighted by early knockout experiments of DNA methyltransferase: Dnmt1 which resulted in embryo death during mid-gestation (Li *et al*, 1992). Later Dnmt3a and Dnmt3b knockouts in mice showed embryos failed to reach full term and exhibit

stunted growth among other developmental defects (Okano *et al*, 1999). These observations and others made by the same team highlight a role of DNA methyltransferase activity in differential gene regulation within the developing embryo (Li *et al*, 1993; Beard *et al*, 1995; Gladych *et al*, 2015). Concurrently, it also appears to play a role in gene regulation during cellular reprogramming whereby the genome is reorganised from a somatic to pluripotent state (Gladych *et al*, 2015). The introduction of the OKSM factors used to create the first generation iPSC demonstrated demethylation of the Fbx15 and Nanog promoters, a key characteristic of ESCs (Takahashi and Yamanaka, 2006). DNA methylation profiling studies revealed even greater global similarities between ESCs and iPSCs. Many somatic unmethylated genes that are methylated in ESCs revert to a methylated state in iPSCs. In contrast many somatic hypermethylated genes are reprogrammed to an unmethylated state (Lister *et al*, 2011, Ziller *et al*, 2011). Other reprogramming techniques including reprogramming cells with amphibian oocyte extracts induced genome-wide demethylation in MEFs to levels resembling mESCs (Bian *et al*, 2009). Moreover it would also appear that agents which block DNA methyltransferase activity are able to increase the efficiency of iPSC production and even facilitate the transition of partially reprogrammed cells to a fully pluripotent state (Mikkelsen *et al*, 2008; Hou *et al*, 2013). These studies suggest specific DNA methyltransferase activity plays an important role in gene regulatory activity which is required for both cellular differentiation and reprogramming to a pluripotent state.

1.4.4 Spatial territories, Heterochromatin and euchromatin

While the organisation of DNA into chromatin forms the basic regulatory unit of the genome, the organisation of the genome into the 3D nuclear space also has a direct regulatory impact. The mechanisms which determine a genome's unique 3D structure in the nuclear space during interphase are largely unknown, even though this plays a fundamental role in gene regulation. DNA control elements and their target genes are not often in direct proximity along the linear genome sequence. Recent studies have elucidated that distal control elements act on target genes by a direct physical interaction and, thus, arrangement in 3D space shapes the specificity of these interactions.

Within the 3D nuclear space there are notable structures which have an impact on transcriptional regulation. These structures include the nuclear lamina, nuclear pore complex and nucleolus. The nuclear lamina complex is implicated in transcriptional repression through interactions with lamina associated domains (LADs) in chromosomes. In essence, the nuclear lamina physically move sections of chromosomes to the nuclear periphery which generally diminishes their transcriptional activity. The nuclear pore complex (NPC) is a separate substructure which mediates the transport of factors between the cytoplasm and the nucleus. The NPC has been linked to the increased transcriptional activity in the nuclear periphery even in domains termed heterochromatin exclusion zones. The nucleolus is another substructure that has been linked with transcriptionally inactive genes and may have a similar mechanism of chromatin interaction as the nuclear lamina, as they interact with chromatin within shared domains (Fraser *et al*, 2015).

During interphase, chromosomes adopt specific territories (figure 5A) which seem to reduce interchromosomal interactions. Fluorescence *in situ* hybridisation (FISH) studies have clearly demonstrated this propensity of chromosomes to form individual domains and more modern techniques such as 4C and Hi-C analysis have shown that intrachromosomal contacts hundreds of megabases apart are greatly favoured over interchromosomal contacts. While less frequent, chromosomes are more likely to interact with adjacent territories rather than the most distal points in their own territories. Gene rich chromosomes also tend to localize nearer to the centre of the nucleus, whilst gene poor chromosomes locate nearer the periphery. This localisation of gene rich chromatin to the centre of the nucleus has led to speculation that there is greater access to transcriptional machinery and other factors important for increased transcription at the centre of the nucleus. Nevertheless the chromosome territory in which a particular gene is confined, and adjacent territories, will further define a gene's transcriptional potential (Fraser *et al*, 2015).

1.4.5 TADS and sub-TADs

Studies utilising 5C and Hi-C methodologies have demonstrated dense chromatin domains within chromosome compartments which more frequently interact within themselves, rather than outside these domains. These dense chromatin domains are delineated by CTCF and cohesin, and are referred to as topologically associated domains or TADs (Figure 6B). Each TAD can have a self-contained transcriptional activity. Indeed, genes within a TAD may have similar transcriptional activity. Genes governing lineage specification which exhibit simultaneous expression during development tend to be within the same TAD. Indeed the spatial position of these particular TADs correlate with the activating or repressing histone marks seen at that loci during that stage of development. TADs can remain spatially distant from one another, even when adjacent along chromosomes. Specific TADS are conserved between cell types and even between species. Deletion of CTCF binding motifs which define TAD boundaries has a notable effect on the expression of nearby genes. CTCF deletion demonstrates the powerful influence of TAD boundaries on transcriptional regulation. At higher resolution, TADs can be further sub-divided based on the frequency of intra-domain interactions vs inter-domain interactions into structures called sub-TADs (Figure 6C). Sub-TADs are also delineated by cohesin and CTCF binding. The smaller the scale of organisation the more dynamic interactions become to facilitate complex gene regulation. In ChIP-seq data, CTCF and cohesin enrichment represent the boundaries of these transcriptionally associated loops or sub-TADs; however, 3D structure cannot be inferred from ChIP data alone. Sub-TADs can be both conserved between cell types or cell-type specific (Fraser *et al*, 2015).

1.4.6 Chromatin loops

Chromatin loops occur on a scale even smaller than sub-TADs and function to facilitate interactions between gene promoters and regulatory elements such as enhancers and insulators (figure 6D). The bridge between promoters and enhancers is formed by the mediator complex: a large multimeric complex which literally mediates the effect of regulatory elements to the RNA polymerase II bound at a gene promoter. The Mediator complex with RNA polymerase II forms the pre-initiation complex necessary to initiate transcription. The physical proximity of the

enhancer and promoter forms the chromatin loop, a structure stabilised by the cohesin protein and in some cases CTCF. (Malik and Roeder, 2010; Kim and Kaang, 2015). While it is apparent that chromatin loops must be necessary to facilitate the transcription of Nanog (and later Oct4) following AOE treatment, it is so far unclear as to what extent the chromatin structure becomes altered to promote the increased gene transcription. Indeed, as Bereketoglu's data (unpublished) suggests, histone alterations around the Nanog regulatory regions are induced within as little as 3 hours, and gene expression is altered after only 6 hours, raising the question of whether structural changes to chromatin also occur within a short time frame. It is also of interest whether XOE treatment effects chromatin structure around similar key loci.

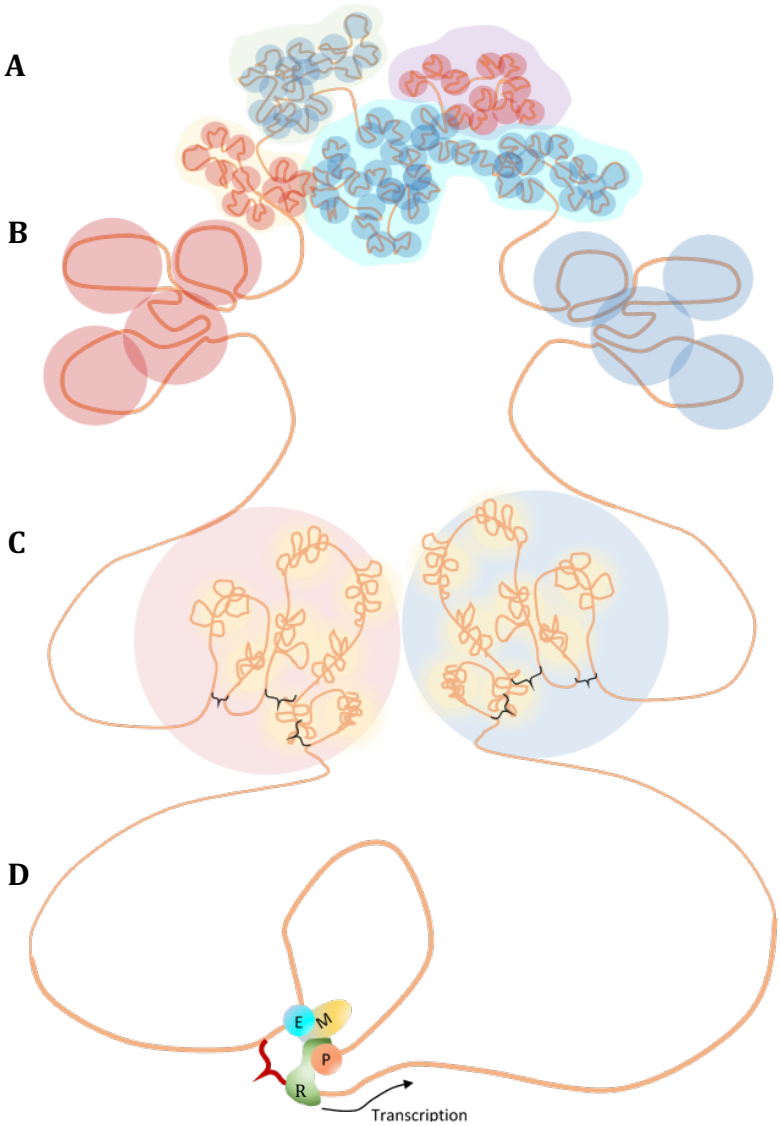


Figure 6. Levels of physical genome organisation. (A) At the lowest resolution chromatin forms chromosome territories located in differential nuclear space. (B) Chromatin can be defined by the frequency of interactions and CTCF enrichment into TADs. (C) Individual TADs can be further dived into sub-TADs. (D) Chromatin loops dynamically form to facilitate gene transcription. Key: Heterochromatin is represented by red circles, Euchromatin is seen in blue circles, DNA bound to nucleosomes in peach, black braces represent CTCF enrichment, while red braces represent cohesin enrichment. E, M, P and R represent -enhancer regions, mediator complex, promoter regions and RNA polymerase II respectively.

Enhancer elements activate transcription at short or long distances both upstream and downstream of a target gene. While it is rare, there is evidence that enhancers are even capable of driving transcription from within a gene and even from a different chromosome (Fraser *et al*, 2015). While the same enhancer and promoter sequences are present in every genome, enhancer-RNA polymerase II interactions are determined through the binding of specific transcription factors to the enhancer.

Once bound to an enhancer, transcription factors can enhance or supress the ability of RNA polymerase II to initiate transcription. A specific combination of enhancer bound TF's is required for the co-ordination of gene transcription. Due to the fact that Promoter-Enhancer interactions are governed by transcription factors these interactions are largely tissue specific (Fraser *et al*, 2015). The regulation of the genome at many levels can be probed to investigate a gene's likely transcriptional state. Chromatin immunoprecipitation remains the gold standard method for probing the potential transcriptional state of genes within a cell population (Liu *et al*, 2010).

1.5 Aims and objectives

One of the aims of this research was to further characterise the biochemical similarities of the axolotl Nanog orthologues: axNanog and EggNog. We aimed to investigate this in two different assays. The first set of experiments were NanoBiT complementation assays similar to those used by Dixon *et al* (2010). The NanoBiT complementation assay was used to compare the protein-protein interactions of axNanog and EggNog. The NanoBiT system attaches one of two luciferase subunits via a flexible linker molecule to proteins of interest. An active luciferase molecule can only be formed upon the fusion of the two subunits which possess a low affinity for one another. This assay can thus be utilised to identify even transient interactions and identify differences in binding affinity. Specifically, we sought to test whether axNanog and EggNog can physically interact with axolotl SMAD1 and SMAD2 proteins. As EggNog has a role in PGC specification, which does not involve Nodal signalling, there is no reason, a priori, to assume it synergizes with Nodal signalling through interaction with the SMAD2 effector protein. Previous experiments have suggested a physical interaction between axNanog and SMAD2, and results from morpholino depletion suggest that axNanog has a conserved role in which it physically interacts with SMAD2, DPY30 and COMPASS. However, the shared homology between EggNog and axNanog suggests that the two proteins exhibit similar biochemical properties which may include the ability to interact with SMAD2. Furthermore, SMADs exhibit a wide range of binding partners due to their regulatory functions in hundreds of genes, thus a differential binding affinity between closely

related proteins could be functionally significant. Currently there is no evidence to suggest either axNanog or EggNog interact with the effector of BMP signalling-SMAD1, so we sought to discern the potential for differences in binding affinity with SMAD1.

The second set of experiments aimed to create stable cell lines which utilised the GAL4 UAS transcriptional assay system. The Promega GAL4 UAS system is based on a protocol devised by Halder and colleagues (1995), which fuses the yeast GAL4 transcriptional activator protein onto a protein of interest. A second vector contains a regulatory sequence containing 5 GAL4 binding sites upstream of a gene encoding luciferase. Binding of the Gal4 fusion protein to an active transcription factor is sufficient to drive the transcription of an active luciferase enzyme. The activity of this enzyme can then be measured by an illuminometer. I have used this system to test if axNanog or EggNog overexpression affects the transcriptional activity of SMAD1 or SMAD2. A reduction in SMAD2 transcriptional activity has been demonstrated in our lab (unpublished data), suggesting that AxNanog acts as a rheostat of SMAD2 activity similar to its mouse counterpart. On this basis, I aimed to test if EggNog was also capable of damping down SMAD2 transcriptional stimulation. Determining whether axNanog and EggNog exhibit differential binding affinities, or effects on SMAD transcriptional activity, could help determine whether the dramatically different roles observed in response to their respective depletions from the embryo are due to their differential expression patterns or their possible divergent biochemical properties.

While the expression of axNanog and EggNog mRNA transcripts has been determined, it is unknown whether these transcription factors are present as translated proteins within the oocyte or whether their emergence occurs after translation from zygotically derived transcripts. To address this I performed immunoblotting using newly produced, affinity-purified, antibodies raised against AxNanog and EggNog. These experiments also served to validate the specificity of the new antibodies in immunoblotting, immunocytochemistry and IP experiments. The ultimate goal of antibody validation was use their use in ChIP-qPCR to elucidate

the binding of axolotl Nanog at key developmental loci of the mouse genome, after NIH3T3 cells have been treated with AOE or XOE. Published data has demonstrated that axNanog is capable of binding to a mammalian NANOG promoter when it is overexpressed in ES cells (Dixon *et al*, 2010). Unpublished ChIP qPCR experimental evidence has also shown an enrichment of axNanog molecules at the promoter and enhancer regions of the NANOG gene after AOE remodelling. Thus I used the newly acquired antibodies to determine whether axNanog molecules in the AOE were bound to these regulatory elements, to elucidate whether the binding of these molecules could drive expression of Nanog in AOE treated somatic cells.

ChIP qPCR experiments also aimed to elucidate whether treatments with AOE or XOE induced changes in sub-TAD structure. To accomplish this, we attempted to identify areas of mESC specific sites of CTCF enrichment around the NANOG and OCT4 loci, comparing with published ChIP-seq data sets available from the ENCODE database. Sites of CTCF enrichment specific to these sites represent active sub-TADs in a pluripotent state. We then aimed to identify sites where CTCF enrichment was present, specifically in somatic cells representing active TADs in somatic cells. From identification of pluripotency and somatic gene associated sites of CTCF enrichment, we then aimed to determine the extent of chromatin remodelling after 6 hours of AOE or XOE treatment, looking for differential responses in CTCF enrichment at these key sites.

2. Materials and Methods

2.1 Chemicals and Reagents

Most chemicals and reagents were purchased from Sigma-Aldrich. Reagents purchased elsewhere are stated as such. Solutions were made to standard protocols which are described in Appendix A.

2.2 Methodology

2.2.1 Cell Lines and Oocytes

For both luciferase assays an immortalised human embryonic kidney (HEK293T) cell line was used due to its longevity in culture and propensity for transient transfection. The line expresses the large T-antigen, which facilitates the production of lentiviral packaging from vectors containing the SV40 origin of replication, making it ideal for both assay types. NIH3T3 Fibroblasts were used for oocyte extract treatments along with CGR8 mouse embryonic stem cells (mESCs) for subsequent ChIP-qPCR. Oocytes and egg extracts were prepared from *Xenopus laevis* and *Ambystoma mexicanum* oocytes. Further details on extract preparation can be found in section 2.2.9.

2.2.2 Cell culture

2.2.2.1 Mammalian cell culture

All cell lines were cultured at 37°C in 5% CO_2 conditions. Lines were passaged at 80-90% confluency. This was achieved by washing once in PBS and incubation with 0.1% Trypsin-EDTA for 2-3 minutes at 37°C. Upon cell detachment serum-containing media (3 times the volume of Trypsin) was added before centrifugation at 1300rpm for 5 minutes. Pelleted cells were resuspended in fresh media, counted and plated at varying densities. HEK293T cells were cultured in high glucose Dulbecco's Modified Eagles Medium (DMEM) supplemented with 10% Foetal bovine serum (FBS), 1% Penicillin/Streptomycin, 2mM L-Glutamine and 2mM non-essential amino acids. NIH3T3 cells were cultured under the same conditions with addition of 0.1mM ß-Mecaptoethanol (BME). Wild-type mESCs were cultured in identical conditions to NIH3T3 cells with the addition of 100ng/ml leukaemia inhibitory factor (LIF-Chemicon) in 6-well plates treated with 0.1% gelatine.

2.2.2.2 Cryopreservation of cell lines

Cells were collected for cryopreservation in the same manner of passaging however, following centrifugation, pelleted cells were resuspended in a freeze medium containing 10% Dimethyl sulfoxide (DMSO), 60% FBS and 30% complete medium. Cells were aliquoted into cryopreservation tubes labelled and stored in a Mr Frosty cooling rack at -80°C for a minimum of 24 hours before being transferred into liquid nitrogen storage.

2.2.3 Creation of plasmids

2.2.3.1 Creation of NanoBiT Vectors

The NanoBiT vectors fuse luciferase subunits to C or N terminals of the gene of interest. Primers were designed for Axolotl SMAD1, SMAD2, Nanog and EggNog as well as mouse Nanog and SMAD2/3 and human VentX to amplify genes with appropriate restriction sites for cloning into NanoBiT C or N vectors which fuse either the large or small luciferase subunit (LgBiT/SmBiT) to the protein of interest when expressed. All primers used to amplify genes for a C terminus fused LgBiT/SmBiT incorporated an EcoRI restriction site on the 5’ end of the gene and a XhoI site on the 3’ of the gene and removed the stop codon from the gene inserts. All primers used to amplify genes for insertion into N vectors which fuse LgBiT/SmBiT incorporate and EcoRI site to the 5’ end of the gene and an XbaI site to the 3’ end, leaving the stop codon intact. All primer sequences and construct sequences can be found in Appendix B. Gene inserts for vectors were produced by amplification from cDNA or from other vectors (made in house) via polymerase chain reaction (PCR). For each PCR reaction a reaction mix was made up as seen in Table 1.

Table 1. Showing PCR reaction mix used

Reagent	**Volume (µl)**
RedTaq Taq polymerase ready mix	5
1µM Forward and reverse primers	1
50ng cDNA template	1
Nuclease-free H_2O	3
Final volume	**10**

A gradient PCR of 5 reaction mixes for each gene was then run covering +/- 2°C annealing temperatures around the predicted melting temperature (TM) for each gene primer pair. The annealing temperature producing the greatest degree of amplification was used for downstream applications. The general PCR program can be seen in table 2 below.

Table 2. Showing a standard PCR program used

Cycle	**Temperature (°C)**	**Duration (S)**	**Number of cycles**
Initial denaturation	96	120	1
Denaturation	96	30	35
Annealing	55-70	30	
Extension	72	30-90	
Final extension	72	300	1

A negative control PCR without template was also run alongside each primer set to ensure there was no contamination in primer solutions.

After PCR amplification, samples were run with a 1000bp reference DNA ladder on a 1% Agarose gel containing 5µg/ml ethidium bromide. The DNA bands were excised with a scalpel using a UV transluminator. DNA was purified using the QIAGEN Gel extraction kit according to the manufacturer's protocol.
Purified DNA was briefly vortexed before quantifying 1µl on a NanoDrop1000 spectrophotometer. Samples were digested with restriction enzymes if the

concentration of DNA was ≥100ng/µl. Genes and vectors with the EcoRI and XhoI restriction sites were digested for 1 hour at 37°C in a reaction mix of 2µg DNA, 2.5µl New England Biolabs (NEB) 10X Cutsmart buffer and 1µl each of EcoRI HF and XhoI restriction enzymes brought to a final volume of 25µl with nuclease-free water. The same method was used for genes and vectors with EcoRI and XbaI sites, but using XbaI enzyme instead of XhoI.

After restriction digestion, DNA was purified from a 1% agarose gel as previously described. The digested gene inserts and vectors were ligated together overnight at room temperature in a reaction mix detailed in table 3.

Table 3 Showing ligation reaction mixture

Reagent	**Volume (µl)**
Digested vector DNA (50ngµl^{-1})	1
Digested Gene insert (~70ngµl^{-1})	7
10X T4 DNA ligase buffer (Promega)	1
T4 DNA ligase (Promega)	1

Amplified gene inserts which had a concentration below 100ng/µl after gel extraction were ligated into a TA cloning vector PGemT Easy (Promega) using the same method, then transformed according to manufacturer's protocols. DNA from white insert-positive colonies was purified with the QIAGEN Mini-Prep kit and digested and purified as before to ensure a sufficient level of insert for cloning into the NanoBiT vectors.

The following constructs were generated for use in the NanoBiT protein complementation assay: axSMAD1 C-LgBiT, axSMAD1 N-LgBiT, axSMAD2 C-LgBiT, axSMAD2 N-LgBiT, EggNog C-LgBiT, axNanog C-LgBiT, mSMAD2 C-LgBiT, mSMAD2 N-LgBiT, EggNog C-smlBiT, axNanog C-smlBiT, hVentX C-smlBiT, mNANOG C-smlBiT, mNANOG C-LgBiT, mSMAD2 C-LgBiT.

2.2.3.2 Creation of Gal4 Vectors

Cloned DNA encoding SMAD1 and SMAD2 was again amplified from cDNA using primers with a 5' Promega SgF1 restriction site and 3' NotI site. GAL4 PFN26A (BIND*/ hluc*-neo FLEXI Vector (Promega) and gene inserts were purified and digested with SgF1 and NotI restriction enzymes in Buffer C without BSA (Promega). Vectors were ligated as previously described.

2.2.3.3 Creation of PSIN Vectors

Axolotl Nanog, EggNog and Human VentX cDNAs were amplified via PCR using primers with 5' EcoRI and 3' BamHI sites on each fragment. The purified inserts were digested along with pSIN expression vectors with EcoRI HF and BamHI HF in cutsmart buffer before purification. Ligation was performed as described above.

2.2.4 Plasmid transformation

After vectors had been ligated, 100µl of chemically competent E.coli were thawed on ice for 10 minutes before adding 5µl of the ligation. Transformations were mixed briefly and placed on ice for 30 minutes before heat shock for 1 minute at 42°C in a water bath. Transformations were incubated on ice for a further 5 minutes. Finally 700µl of SOC media was added to cells prior to plating on 100µgml^{-1} Ampicillin lysogeny broth (LB) agar plates. Plates were incubated overnight (16 hours) at 37°C.

2.2.5 Plasmid mini-preparations

Following overnight incubation, 6-10 colonies were picked and added to 2ml of LB containing 100µgml^{-1} Ampicillin and incubated in an orbital shaker overnight at 37°C at 220rpm. After incubation 1ml of the mini-culture was pelleted and supernatant was removed. Pellet was resuspended in 250µl resuspension buffer P1. Cells were lysed using 250µl buffer P2 and incubated at room temperature for 3 minutes before adding 300µl neutralisation buffer. Lysed mini-cultures were centrifuged at 17,000g for 5 minutes to remove cellular debris. 700µl of the supernatant was transferred into a separate tube. Phenol chloroform was added and the mixture vortexed briefly to allow DNA precipitation. The sample was spun at 17,000g for 10 minutes to allow separation of the phases. 600µl of the aqueous phase was transferred to a separate

tube and equal volume of isopropanol added before briefly vortexing and centrifugation for 10 minutes at 17,000g. The supernatant was discarded and the pellet washed with 100µl 100% ethanol, followed by 70% ethanol. The pellet was air dried and resuspended in de-ionised water. A restriction digest was performed with the same restriction enzymes and buffer used to insert the gene of interest into the vector. The digested DNA was run on a 1% agarose gel to confirm presence of the gene insert in the mini-culture. Upon confirmation of an insert DNA from the corresponding 1ml of mini-culture was purified according to the standard protocol of a QIAGEN mini-prep kit. Prepared plasmids were briefly vortexed and quantified on the NanoDrop 1000.

2.2.6 Establishment of Gal4 system cell lines

To create cell lines, cells were first transiently transfected with the vector of interest and DNA encoding a lentiviral packaging. The transfected cells then produced a lentivirus media containing the GAL4 UAS HYGRO vector to be delivered to host cells via the lentivirus. The following protocol was used to establish cell lines for the GAL4 system: Day one: cells were plated on sterile T25 cell culture vessels at a density of $1.25x10^6$ and allowed 6 hours for attachment and to enter growth phase. The transfection was performed using Lipofectamine 3000 transfection reagent (Thermofisher Scientific) as shown below in Table 4. All DNA was at a concentration of 1g/l.

Table 4 Showing Gal4-Lipofectamine reaction mixtures

Tube A		**Tube B**	
Reagent	Volume (µl)	Reagent	Volume (µl)
Serum-free DMEM	250	Serum-free DMEM	250
Lipofectamine 3000 reagent	15	Vector DNA	2.5
		PsPPAX2 packaging DNA	1.67
		pMD2-D envelope DNA	0.83
		p3000 reagent	10

Tubes A and B were mixed thoroughly and incubated at room temperature for 5 minutes. The tubes were mixed and incubated at room temperature for 15 minutes. The mixture was added to 3ml of fresh complete medium, placed on cells and left to transfect overnight. Day 2: the media was replaced with fresh medium and left to create virus media overnight. Day 3: the virus media was removed, filtered with a 45 micron filter and added to untransfected HEK 293T cells that had been seeded as above 6 hours prior. Cells were transduced overnight before media was replaced with fresh complete medium. Cells were passaged once in regular medium before being placed in medium containing 200µgml^{-1} Hygromycin B (Thermofisher). Cells were passaged at normal intervals at around 80-90% confluency and plated in medium without antibiotic. Fresh antibiotic medium was added to the plate after attachment. Antibiotic selection continued for 10 days.

The cells selected with antibiotic were passaged until a regular growth pattern and morphology was observed. After selection, cells were transduced with lentiviral media containing either axolotl SMAD1 or SMAD2 PFN26A (BIND) *hluc*-neo FLEXI vectors produced in the same fashion as the GAL4 UAS HYGRO lentiviral media. Cell lines were selected using medium containing G418 at a concentration of 200µg/ml.

Finally a third transduction was performed on lines with lentivirus-containing pSIN expression vectors containing either axolotl Nanog, EggNog or human VentX. Cell lines underwent final selection in medium containing 1µg/ml puromycin.
At each stage after a cell line was stably expressing antibiotic resistance, samples were cryogenically preserved during passaging. In total 6 cell lines were created each containing three vectors: GAL4 UAS HYGRO, either PFN26A (BIND) *hluc*-neo FLEXI SMAD1 or PFN26A (BIND) *hluc*-neo FLEXI SMAD2 and pSIN Nanog, pSIN EggNog or pSIN Human VentX. After selection cells were plated on a 96 well assay plate and allowed to grow to between 80-90% confluency before cells were washed in PBS and lysed in 20µl passive lysis buffer (Promega) for 15 minutes at room temperature. Plates were then placed in a Promega GLOMAX and wells were injected with 100µl of Luciferase assay reagent II (Promega) and fluorescence was read immediately after injection.

2.2.7 Transient transfection of NanoBiT constructs

NanoBiT constructs were transiently transfected using Lipofectamine 3000 transfection reagent. HEK293Ts were seeded in 100µl media in the inner 60 wells of a 96 well cell culture plate at a density of 500 cells per well. Upon reaching 50-60% confluence the following constructs were prepared as described in table 5 below.

Table 5. Showing NanoBiT-Lipofectamine reaction mixtures

Reagent	Volume (µl)	Tubes
Serum-free DMEM	5	A1-11 B1-11
Lipofectamine 3000 reagent	0.3	A1-11
P3000 reagent	0.2	B1-11
DNA (6.25ng/µl)		
LgBiT-PRKAR2A Positive control	4	B1
SmBiT-PRKACA positive control	4	B1
NanoBiT Negative Control	4	B2
LgBiT axSMAD1-N	4	B3-5
LgBiT axSMAD2-N	4	B2, B6-8
SmBiT axNanog-N	4	B3, B6, B10
SmBiT EggNog-N	4	B4, B7, B11
SmBiT VentX-N	4	B5 + B8
SmBiT mNanog-N	4	B9
LgBit mSMAD2-N	4	B9
LgBiT axNanog-N	4	B10
LgBiT EggNog-N	4	B11

Tubes were mixed and incubated at room temperature for 5 minutes. Corresponding tubes (A1 + B1, A2 + B2 etc) were then mixed and incubated at room temperature for 15 minutes. A total of 5 wells were transfected with each construct condition. This was repeated for constructs with LgBiT-N and SmBit-C combinations. A subsequent experiment was performed using double the volume of each DNA. Cells were incubated for 24 hours at 37°C, 5% CO_2. After 24, 48 and 72 hours media was replaced with 40µl serum-free medium and 10µl of Nano-Glo reagent was added after 1 in 20 dilution with the Promega LCS dilution buffer. Plates were mixed for one minute in the absence of light on a rotary shaker at 30rpm. Plates were then placed

in a GLOMAX illuminometer and read after a minute equilibration. For each condition 3 repeats were made and read separately.

2.2.8 Validation of antibodies

2.2.8.1 *In vitro* translation of proteins

2.2.8.1.1 Rabbit reticulocyte lysate *in vitro* translation

To validate affinity-purified antibodies to axolotl Nanog, EggNog proteins were first produced using the Promega rabbit reticulocyte lysate kit. The kit synthesises proteins from a synthetic poly-A RNA template. To compare test antibodies against a well-validated anti-HA antibody RNAs used to produce proteins of interest also encoded the production of a double human influenza hemagglutinin (HA) tag. A control luciferase RNA was also used to measure the efficiency of protein production by measuring luminescence in the GLOMAX illuminometer. The *in vitro* translation mixture was assembled as detailed in Table 6 below.

Table 6. Showing the *in vitro* translation reaction mixture

Reagent	**Volume (µl)**
Rabbit Reticulocyte Lysate	35
Amino Acid Mixture Minus Leucine, 1mM	0.5
Amino Acid Mixture Minus Methionine, 1mM	0.5
RNasin Ribonuclease Inhibitor (40u/µl)	1
RNA (1µg/µl)	2
Nuclease-Free Water	11

Three differing RNA concentrations were also trialled to optimise protein production, in amounts of 50ng, 100ng and 1µg. The reaction mixture was mixed thoroughly and incubated at 30°C for 90 minutes. To measure luciferase activity, 2.5µl of luciferase translation mixture was aliquoted into 6 wells of a 96 well plate. Fifty microliters of room temperature standard luciferase assay reagent (LAR) was

added and mixed. The plate was placed into the GLOMAX illuminometer and luminescence was measured for one minute after a 10 second equilibration.

2.2.8.1.2 Wheat germ lysate *in vitro* translation

Proteins were also produced using the Promega wheat germ lysate kit. The reaction mixture was identical to the rabbit reticulocyte mixture described previously, however 35µl of wheat germ lysate was used in its place.

2.2.8.2 Production of proteins in *Xenopus* oocytes

Axolotl Nanog, EggNog and Oct4 proteins with HA tags were produced by injecting poly-adenylated RNA into mature *Xenopus* oocytes. Extracts were produced using a process outlined by Hutchinson *et al* (1988) described in detail in section 2.2.9.

2.2.8.3 Production of expression cell lines

Cell lines expressing Nanog and EggNog were produced by transducing HEK 293T cells with lentivirus containing either Nanog or EggNog pSIN vectors. Transductions, lentiviral transfections and selections were performed according to the protocol described in section 2.2.6, selection was performed using 1µgml-1 puromycin.

2.2.8.4 Antibodies

In total there were 4 test rabbit antibodies produced by Dundee Cell against two different epitopes for axolotl Nanog and two for EggNog. Antibodies were affinity purified but not validated for use in Western blotting, immunofluorescence or ChIP experiments. Antibodies were resuspended following standard procedures previously established by our lab. Antibody concentrations were unknown at time of arrival. All other antibodies were validated antibodies supplied by Abcam.

2.2.8.5 Dot Blots

Dot blots were performed to test antibody specificity and sensitivity. By dotting 5µl of the antigen each antibody was generated against on a membrane activated in methanol and equilibrated in transfer buffer for 5 minutes. The membrane was left to dry and washed in 0.1% PBST before being blocked in PBST 5% milk for one hour.

The membranes were incubated for one hour at room temperature with 0.5µgml^{-1} of each primary test antibody before 3 washes for 5 minutes in 0.1% PBST. Membranes were then incubated in a mouse anti-rabbit secondary antibody with a conjugated horseradish peroxidase enzyme for one hour at room temperature in 0.1% PBST 5% milk. Finally membranes were washed 3 times in 0.1% PBST before being analyzed at high-sensitivity mode in the C-Digit chemilumincence imager for 12 minutes after adding complete Signal-fire ECL substrate (Cell Signalling).

2.2.8.6 Western blotting

2.2.8.6.1 Polyacrylamide gel electrophoresis

BIO-RAD mini gel casts were assembled according to manufacturer's instructions. All SDS-PAGE used 5ml, 8% resolving gel and standard 5% stacking gel detailed in Table 7 below.

Table 7. Reagents used to create 5ml of 8% resolving gel

Solution component	Volume (ml)
Deionised H_2O	2.6
30% acrylamide mix (BIORAD)	1.0
1.5M Tris pH8.8	1.3
10% SDS	0.05
10% Ammonium persulphate	0.05
Tetramethylethylenediamine (TEMED)	0.004

Components were mixed and poured into gel cassettes leaving a 1cm gap between the resolving gel meniscus and well comb of the gel mould. Approximately 200µl was poured on top of the resolving gel to flatten the meniscus and prevent dehydration of gels, before leaving to polymerise at room temperature for 30 minutes. Once gels were set, isopropanol was removed and a 2ml 5% stacking gel was made for each gel using reagents listed in Table 8.

Table 8. Recipe for 2ml 5% stacking gel

Solution component	Volume (ml)
Deionised H_2O	1.4
30% acrylamide mix (BIORAD)	0.33
1.0M Tris pH6.8	0.25
10% SDS	0.02
10% Ammonium persulphate	0.02
Tetramethylethylenediamine (TEMED)	0.002

Reagents were mixed and poured into gel moulds and the well combs were added ensuring not to form air bubbles. 30 minutes was allowed for polymerisation. All samples were boiled at 95°C in SDS gel loading buffer for 5 minutes to denature proteins. Gel cassettes were loaded into the electrophoresis tank after and wells were rinsed with deionised water. The tank was filled with Tris-glycine electrophoresis buffer and gels were loaded with 15µl of each sample alongside a rainbow marker protein ladder. Empty wells were loaded with 15µl 1X gel loading buffer. Gels were run at 16V until the electrophoretic front (marked by bromophenol blue) reached the resolving gel and then run at 100V until the electrophoretic front ran out of the gel. After completion, gel moulds were opened, stacking gel removed, the right hand corner of the gels was excised and the gels were gently placed in Tris-glycine electrophoresis buffer to prevent sticking or dehydration.

2.2.8.6.2 Semi-dry transfer of proteins

Gels were rinsed and placed in semi-dry transfer buffer for 5 minutes. PVDF membranes were cut to the gel size and activated for 30 seconds in 100% methanol before equilibration in transfer buffer for 5 minutes. Two 9mm BIORAD extra thick filter papers were soaked in transfer buffer for each gel and placed flat onto the semi-dry transfer anode plate. The PVDF membrane was placed on the filter paper followed by gel and another layer of filter paper. This was rolled to remove any air bubbles before the second cathode plate was placed on top and secured. Semi dry transfer was accomplished by running the transfer at 300 milliamps for 1 hour and

30 minutes until the rainbow ladder had visibly transferred onto the PVDF membrane.

After transfer gels were stained using BIORAD gel code blue and membranes were stained with Ponceau stain to check transfer efficiency. Membranes were incubated in 0.01% PBST 5% milk for blocking and then incubated overnight at 4°C with 1µgml^{-1} of each primary antibody or a control mouse Anti-HA primary antibody. Membranes were washed 3 times for 5 minutes in 0.1% PBST. Membranes probed with Rabbit primary antibodies were incubated in mouse anti-Rabbit secondary antibody with a conjugated horseradish peroxidase. Membranes probed with mouse anti-HA were incubated in goat anti-mouse secondary antibody also with horseradish peroxidase conjugate. All membranes were incubated in 1µgml^{-1} secondary antibodies in 0.1% PBST 5% milk for one hour at room temperature. After a final 3 washes in 0.1% PBST, membranes were placed in the Licor C-Digit Western blot imager after adding complete Signal fire ECL substrate (Cell Signalling).

2.2.8.7 Immunoprecipitation

Immunoprecipitations were performed according to the protocol in Gambus & Blow (2013). Axolotl oocyte extract was diluted five times with IP buffer and spun for 15 minutes at 18,500g at 4°C. 100µl aliquots were supplemented with 2 or 4µg affinity purified rabbit antibodies or IgG from rabbit serum (Millipore) as a control. After an hour-long incubation
on ice, immunoprecipitation samples were mixed with 30µl washed Protein A Sure beads (Biorad) and incubated with mixing at room temperature for 1 hour. Beads were washed twice as above, once with the buffer supplemented with 0.1% triton X-100 and twice with the buffer alone. Immunoprecipitated proteins were eluted off beads by boiling in 1x Laemmli loading buffer.

2.2.9 Axolotl and *Xenopus* oocyte extract treatments

Axolotl oocyte extracts were prepared according to protocols outlined by Hutchinson *et al* (1988). Oocyte extracts were prepared from the ovaries of either Axolotl or *Xenopus* females by digesting follicular cells for 2 hours at 25°C using 1gl^{-1}

Collagenase in OR2 media (Appendix A). Oocytes between stages 4-6 were selected by size before washing in ice-cold extraction buffer containing protease inhibitor cocktail. Oocytes were centrifuged at 4°C for 10 minutes at 10,000g. The cytoplasmic layer was collected, supplemented with 10% glycerol and snap frozen in liquid nitrogen.

For the reprogramming experiments and ChIP, NIH 3T3 cells were used for extract treatments. Cells were trypsinised at ~80% confluence and washed in DMEM. Cells were counted using a haemocytometer. Cells were pelleted at 10,000g for 5 minutes and media was removed. $1x10^6$ cells were resuspended in 1300µl fresh DMEM and cooled on ice for two minutes. Cells were then permeabilized for 1 minute 20 seconds in 1.5ml permeabilization buffer (PB) containing 173µgml^{-1} digitonin (Appendix A). The reaction was deactivated by addition of 10ml PB, and cells were then pelleted at 10,000g for 5 minutes. The extent of permeabilization was assessed using fluorescein isothiocyanate-Dextran (FITC-Dextran) and propidium iodide (PI) staining. Cells were stained with FITC-dextran and PI after permeabilization for 70 seconds, 90 seconds or 110 seconds.

Permeabilized cells were then incubated in 50µl of Axolotl or *Xenopus* extract with 100µl PB containing 150µgml^{-1} creatine phosphokinase, 60mM phosphocreatine and 1mM Adenosine triphosphate. Cells were incubated at 17°C for 6 hours in extracts along with a control group without extract and were incubated in 150µl PB buffer. After 6 hours cells were washed in PB and pelleted at 10,000g for 5 minutes. The supernatant was discarded, cells were snap frozen in liquid nitrogen and stored at -80°C for Chip-QPCR experiments.

2.2.10 ChIP-QPCR

2.2.10.1 *In vivo* cross-linking

All reagents used for preparation of chromatin were used from the (EZ Magna ChIP A/G Kit, Millipore) unless otherwise stated.

To start cross-linking the chromatin protein to DNA, each group of 2 x 10^6 cells was re-suspended gently with 1 ml of fixative (1% of formaldehyde in PBS) and incubated at room temperature for 10 minutes. To terminate cross-linking, 100µl of 10X Glycine was added to each tube, to quench un-reacted formaldehyde, and then incubated a further 5 minutes at room temperature. Cross-linked cells were pelleted by centrifuging for 5 minutes at 10,000Xg at room temperature. Supernatant was discarded and pelleted cells were washed twice in 1 ml of ice-cold PBS including 1µl of protease inhibitor cocktail II and pelleted by centrifuging for 5 minutes at 1200Xg at 4^{o}C.

2.2.10.2 Cell and nuclear lysis

Lysed cell extracts were prepared from 8x10^6 pelleted cells in 400µl of cell lysis buffer including 0.5µl of protease inhibitor cocktail II and incubated on ice for 15 minutes. The cell suspension was mixed every 5 minutes. The cell suspension was then centrifuged at 1200g at 4^{o}C for 5 minutes. The supernatant was removed and the pellet was resuspended in 800µl of Nuclear Lysis Buffer with 4µl protease inhibitor and transferred into an 800µl 12mmx12mm Covaris shearing vesicle with AFA fibre.

2.2.10.3 Chromatin shearing

Samples were sheared using a Covaris S220 ultra-sonicator in vesicles containing AFA fibres and run on a chromatin-shearing program on the S-Series SonoLAB program under the following conditions: duty cycle 5%, intensity level 4, cycles/burst 200. The shearing program was run between 5 and 15 minutes. Twenty microliters of sheared chromatin was mixed with 100µl ChIP elution buffer and 1µl proteinase K and incubated at 62°C for 2 hours in a rotary shaker. Samples were mixed with 5 volumes (100µl) bind reagent A and run through a spin filter collection tube for 30 seconds at 10,000g. The spin filter column was washed with 500µl Wash reagent B before being spun for 5 minutes at 10,000g and air dried for one minute before being eluted for 10 minutes with 20µl elute reagent C into a clean tube. DNA was extracted from the sonicated chromatin and run on a 2% agarose gel alongside a 100bp DNA marker to assess the optimum amount of chromatin shearing which results in DNA fragments

between 200-1000bp. All samples were then sheared for 15 minutes before being centrifuged at 10,000g for 10 minutes at 4°C to remove any cellular debris, and transferred into fresh tubes. 1% of the supernatant was saved as the input before immunoprecipitation (IP). All samples were stored at -80°C or immediately used for IP.

2.2.10.4 Chromatin Immunoprecipitation

For every 50µl of sheared chromatin, 450µl of ice-cold ChIP dilution buffer was added containing 2.25µl PIC for each ChIP reaction. For each antibody, amounts were optimised for use with $2x10^6$ cells worth of chromatin (~200µl). Three to six micrograms of either anti-TPA 104, anti-acetyl histone 3, anti-CTCF or normal rabbit IgG was used in ChIP. Antibody concentrations were varied to determine an optimum amount. 20µl of suspended protein A magnetic beads (BIORAD), the antibody of interest and control antibodies were added to diluted samples and incubated overnight at 4^oC with rotation at 15rpm on a rotating tube rack. Protein A bead complexes were collected with a Magna grip rack and were washed by re-suspending the beads in 0.5 ml of each of the following ice cold buffers for 5 minutes: low salt immune complex wash buffer, high salt immune complex wash buffer, LiCl immune complex wash buffer and TE buffer.

Following washing steps, input and immunoprecipitated samples were incubated in 100µl ChIP elution buffer (including SDS) and 1µl Proteinase K solution at 62°C for 4 hours with shaking, then at 95°C for 10 minutes. Samples were allowed to cool and DNA was purified as described for sonicated samples in 2.2.10.3 except eluted in 50µl nuclease-free water.

2.2.10.5 Quantitative PCR

2.2.10.5.1 Primer design

The initial ChIP aimed to investigate if the axNanog protein bound to specific loci in the mouse genome after AOE treatment. To test this we designed PCR primers probing for the Nanog promoter and enhancer sequences to investigate if these sequences were enriched in axNanog immunoprecipitations after AOE treatment compared to untreated cells. Cells treated with XOE were used for as a negative

control for the oocyte treatments. For each IP normal rabbit IgG (RIgG) was used as a negative control antibody. All analysis calculates enrichment after IP as a percentage of input prior to ChIP. GAPDH primers were used to identify any differences between comparative samples. Before immunoprecipitation, SYBR green primers which amplify the target sequences were tested using extracted mouse DNA to ensure suitability for qPCR analysis. Of all the primers tested some showed the formation of multiple products or unusual melt curves and thus for some qPCR analysis multiple primers were used for a particular region. For the initial axNanog IP 3µg of anti-axNanog and RIgG was used for IP's, this was later increased to 6µg.
We next investigated whether specific CTCF binding loci had greater enrichment before and after reprogramming with AOE or XOE and whether specific regions of CTCF enrichment resembled untreated cells or mESC after reprogramming. The next ChIP was performed with an anti-CTCF antibody, 5µg was used for the ChIP. The first set of primers used were a control set which amplified a region which was highly enriched in CTCF in both somatic NIH3T3 cells and mESC. Designs for CTCF primers were based on ChIP-seq data from the ENCODE database. ChIP-Seq performed on a variety of somatic cell types and mESC using an anti-CTCF antibody was downloaded from the ENCODE project open access database to the UCSC genome viewers. The each dataset was aligned to the mouse genome mm10 and visualised alongside one another. The UCSC genome viewer automatically labels peaks in which there is a significant amount of enrichment of a DNA sequence above background levels of IgG controls. We then looked for significantly sized peaks which differed between pluripotent and somatic tissue types around the NANOG and OCT4 loci. Datasets were gathered from anti-CTCF ChIP-seq in mouse cerebellum, liver and kidney cell lines and mES Bruce 4 lines.

2.2.10.5.2 QPCR conditions

Quantitative PCR (qPCR) reactions contained 10µl of SYBR®-Green Master Mix, 0.5µl of 10mM forward and reverse primer mixture, 6.5µl of nuclease-free water, and 3µl of sheared input or immunoprecipitated DNA sample. Samples were run in triplicate.

The following qPCR conditions were used: Initial denaturation at 94°C for 10 minutes, 42 cycles of 94°C for 20 seconds (denaturation) and 60°C for 1 minute (annealing and extension).

2.2.11 Eomesodermin transcriptional activation assay

Professor Andrew Johnson isolated Axolotl animal caps and performed all injections for this assay. Axolotl animal caps were injected with 160pg of a luciferase reporter construct with the promoter region of Eomesodermin. A number of animal caps were also incubated in 100uM SB42 inhibitor. Other animal caps were also injected with the reporter construct but received either; 2.5pg, 10pg, 40pg or 160pg of Nanog morpholino. A number of uninjected animal caps were also used as a negative control. All animal caps were lysed in 10ul of passive lysis buffer (Promega) for 30 minutes on ice. The amount of lysate used for the assay was varied from 20µl, 1µl or 0.5µl and brought to 20µl with passive lysis buffer before being pipetted into a well of an assay plate. Plates were then placed in a Promega GLOMAX and wells were injected with 100µl of Luciferase assay reagent II (Promega) and read immediately after injection.

3. Results

3.1 Testing Nanog and EggNog antibodies

A major aim of this project was to investigate potential differences between the two axolotl Nanog orthologues: axNanog and EggNog. Transcriptome data identified EggNog and axNanog from oocyte mRNA transcripts; it is unknown whether these transcription factors are present as translated proteins within the oocyte or if their translation occurs in the zygote. We used western blotting to test axolotl oocyte extracts for the presence of axNanog and EggNog proteins. Firstly, custom-designed rabbit antibodies were validated by immuno-dot blot assays. The four antibodies anti-Nanog1/2 and anti-EggNog1/2 were tested against serial dilutions of antigenic peptides used to raise each antibody (antigens N1, N2, E1, E2, respectively). The peptides were diluted to factors of 1, 5, 10, 100 and 1000 and spotted onto a membrane before incubation in one of the primary antibodies and then a secondary antibody, detailed in section 2.2.8.5. Dot blot results are shown in figure 7.

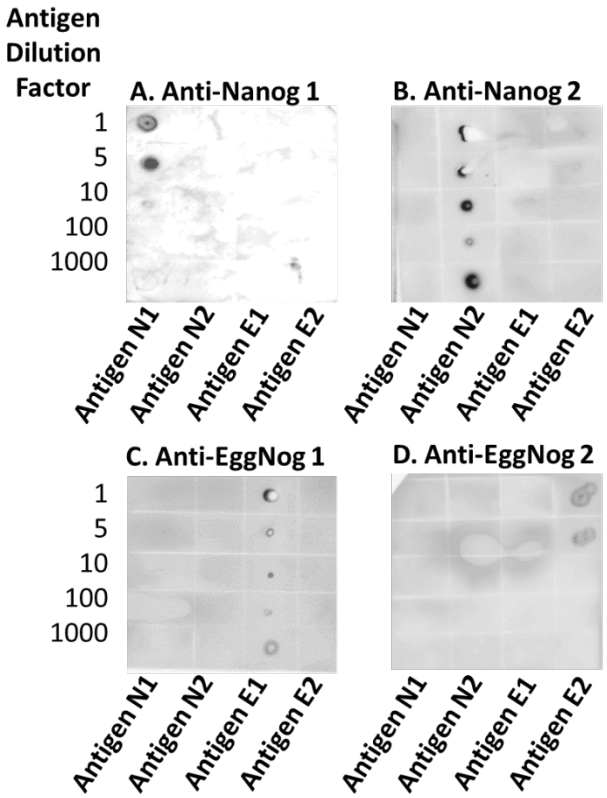


Figure 7. Antibodies recognise specific peptides in immuno-dot blots. Immuno-dot blot comparison of the specificity and affinity of the anti-axNanog and anti-EggNog antibodies. Antigenic peptides were serially diluted and spotted onto membranes before incubation in one of four primary antibodies (A) anti-Nanog 1 (B) anti-Nanog 2 (C) anti-EggNog 1 (D) anti-EggNog 2.

Immuno-dot blots showed that anti-Nanog 2 and anti-EggNog 1 specifically recognised peptides N2 and E1, respectively. Both antibodies produced strong signals at dilutions of 1:1000 and displayed no cross-reactivity with other antigens. Anti-Nanog 1 and anti-EggNog 2 recognised antigens N1 and E2 and also displayed no cross-reactivity with other antigens. However, they produced weaker signals and only detected antigens at concentrations 1:5 or higher. As anti-Nanog 2 and anti-EggNog 1 antibodies showed the highest affinity for their target antigen, they were further tested in western blotting.

We investigated whether antibodies could recognise denatured antigens following SDS-PAGE and Western blotting. Thus Western blotting was performed using the four antigenic peptides N1, N2, E1 and E2. Membranes were incubated in either anti-Nanog 2 or anti-EggNog 1 before incubation in an anti-rabbit IgG-HRP secondary antibody as detailed in section 2.2.8.6. The molecular weight of each peptide was predicted to be 14kDa. Figure 8 shows the western blot results.

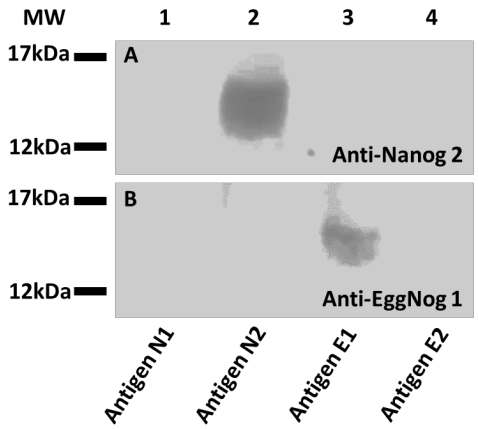


Figure 8. Both Nanog and EggNog antibodies recognise specific antigens in western blotting. Western blotting to detect antigenic peptides after SDS-PAGE and semi dry transfer**.** Lane 1, Antigen N1, lane 2, Antigen N2, lane 3, Antigen E1, lane 4, Antigen E2. Membrane A was probed with anti-Nanog 2 and membrane B probed with anti-EggNog 1.

The Western blot shows recognition of target antigens by anti-Nanog 2 and anti-EggNog 1 after western blotting. Furthermore, both antibodies produced strong signals at a working concentration of 0.5µg/ml. Again no cross-reactivity with the other peptides was observed. We next sought to test whether antibodies could

identify whole proteins after Western blotting. To accomplish this we used rabbit reticulocyte and wheat germ lysate *in vitro* translation kits programmed with synthetic poly-A RNA encoding either axNanog and EggNog with an in frame fusion to HA on its carboxy-terminal domain. To test the success of *in vitro* protein production, lysates from either kit were used to prepare a Western blot probed with primary mouse anti-HA and secondary anti-mouse-HRP antibodies (Figure 9). The predicted molecular weight of axNanog-HA and EggNog-HA were 38kDa and 41kDa, respectively.

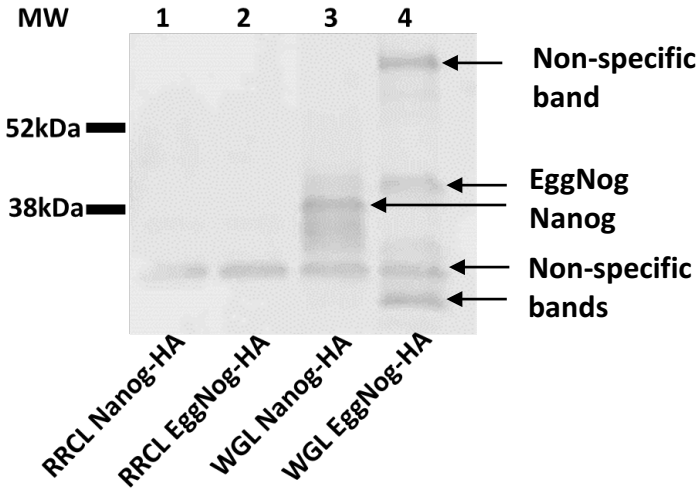


Figure 9. *In vitro* translation failed to produce sufficient amounts of target proteins. Western blotting to detect axNanog-HA and EggNog-HA after in vitro translation. Lane 1, rabbit reticulocyte lysate programmed with Nanog-HA RNA; Lane 2, rabbit reticulocyte lysate programmed with EggNog-HA RNA: Lane 3, wheat germ lysate programmed with Nanog-HA RNA; Lane 4, wheat germ lysate programmed with EggNog-HA RNA.

Non-specific bands are observed at 30kDa in lanes 1 to 4 at 28kDa and 70kDa. There are bands in lanes 3 and 4 at 38kDa and 41kDa, which correspond to the predicted molecular weights of axNanog-HA and EggNog-HA, respectively. In addition, there is also smearing in lane 3. All bands appear faint and were only visible after extensive exposure (5 minutes). This result suggests that the rabbit reticulocyte lysate kit failed to produce the target proteins and the wheat germ lysate kits produced them at very low levels. As both kits performed poorly, synthetic poly-A RNA encoding axNanog-HA, EggNog-HA and Oct4-HA was injected into mature *Xenopus* oocytes, from which lysates were prepared for Western blotting. In addition, an uninjected *Xenopus* oocyte lysate and a multiple antigen containing cell lysate were also used as HA-

negative and positive controls, respectively. The blots were incubated in anti-HA, anti-axNanog 2 or anti-EggNog antibodies (Figure 10).

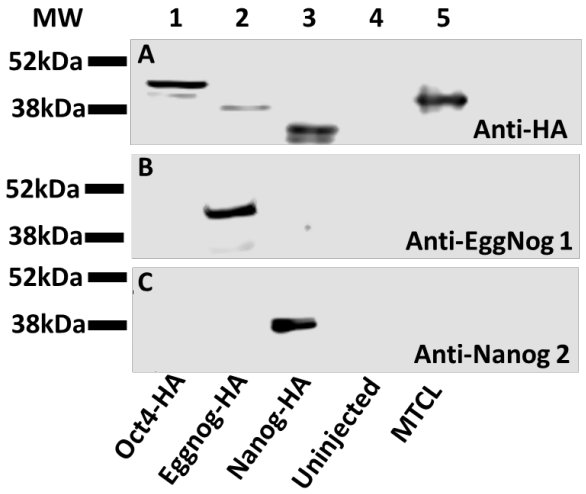


Figure 10. *Xenopus* oocytes produce target protein for custom axolotl antibodies. Western blot prepared with proteins produced in *Xenopus* oocytes**.** Lane 1, oocytes injected with Oct4-HA RNA; Lane 2, oocytes injected with EggNog-HA RNA; lane 3, oocytes injected with Nanog-HA RNA; Lane 4, lysate from uninjected oocytes; Lane 5, multiple tagged cell lysate. Membrane A, probed with anti-HA, membrane B, probed with anti-EggNog 1, membrane C, probed with anti-Nanog 2.

Figure 10A shows that *Xenopus* oocytes produced HA-tagged proteins: the bands in lanes 1-3 at approximately 49kDa, 41kDa and 38kDa correspond to predicted weights of Oct4-HA, EggNog-HA and axNanog-HA, respectively. In blots 10B and 10C, both custom antibodies reacted with proteins at the expected molecular weights for EggNog and axNanog. These antibodies show no cross-reactivity, indicating that they specifically recognise denatured target proteins.

3.2. Detection of axNanog and EggNog proteins in AOE

Transcriptome analysis suggests that the Nanog and EggNog genes are expressed as RNA transcripts in axolotl oocytes; however whether they are stored in the oocyte as proteins is unknown. Having confirmed the effectiveness of the custom antibodies we next aimed to investigate whether native axNanog and EggNog proteins could be detected within the AOE. Western blots were prepared using the same *Xenopus* oocyte extracts shown in Fig. 10, along with 5, 10 and 15µl of AOE. Proteins were isolated from oocyte lysates by precipitation with acetone and then re-dissolved in laemilli buffer before Western blotting (Figure 11).

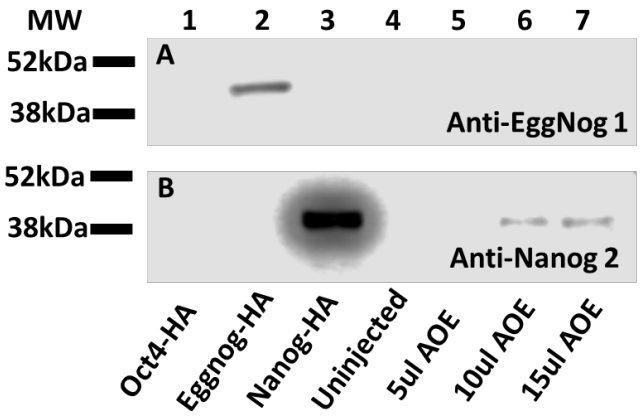


Figure 11. AxNanog is expressed as a protein in oocytes. Western blot to determine the expression of axNanog and EggNog in axolotl oocytes. Lane 1, lysate from oocytes injected with RNA encoding Oct4-HA; Lane 2, lysate from oocytes injected with RNA encoding EggNog-HA; Lane 3, lysate from oocytes injected with RNA encoding Nanog-HA; Lane 4, lysate from uninjected oocytes; lane 5, protein from 5ul AOE; lane 6, protein from 10ul AOE; lane 8, protein from 15ul AOE. Membrane A, Probed with anti-EggNog 1; membrane B, probed with anti-Nanog 2.

Western blots prepared with AOE did not detect EggNog in the oocyte as a translated protein as there are no specific bands at around 40kDA, the molecular weight of EggNog in figure 11A. In contrast, in Fig. 11B bands of approximately 37kDa are observed in lanes 6 and 7 corresponding to axNanog, confirming it is expressed as a protein in oocytes. This result suggests that axNanog is maternally inherited as a protein; however, the absence of maternal EggNog protein suggests that the morphants phenotype shown in figure xx results from depletion of protein translated post-fertilization, in the developing embryo.

To address whether EggNog detection was hindered by low abundance, I performed immunoprecipitations (IPs) using anti-EggNog 1 and anti-axNanog1 antibodies. IPs were carried out using either 2µg or 4µg of antibody as detailed in section 2.2.8.7. As there was a limited amount of AOE available, 5 lysed axolotl oocytes were used for each IP. Lysates prepared from *Xenopus* oocytes expressing Nanog-HA and EggNog-HA were used as positive controls to prepare Western blots, along with eluates from the beads following IPs (figure 12). Blots were incubated in either anti-Nanog2 or anti-EggNog1.

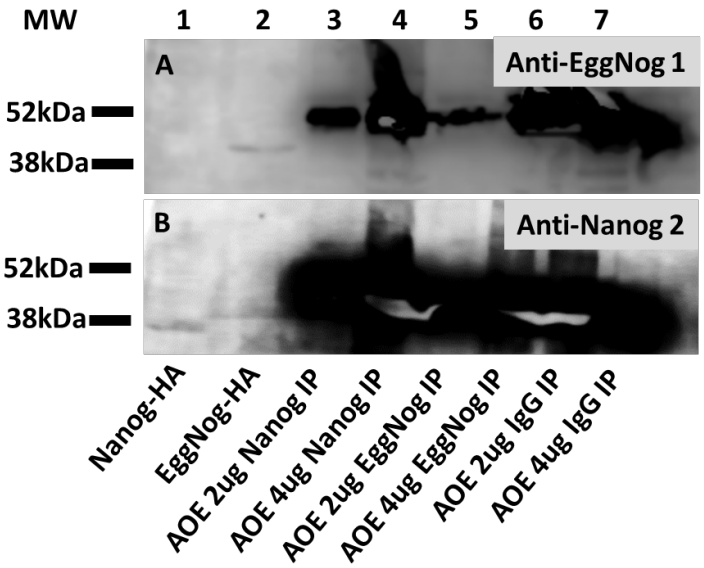


Figure 12. Immunoprecipitations using anti-Nanog2 and anti-EggNog1 antibodies. AOE immunoprecipitations performed with either 2µg or 4µg antibody after being ran on a western blot. Membrane A, anti-EggNog1 antibody, membrane B, anti-Nanog antibody.

There are bands of approximately 41kDa and 38kDa in molecular weight are in lanes 2 and 1 of figures 12 A and B, respectively. These correspond to EggNog-HA and axNanog-HA positive controls. There are no bands around the expected weights of axNanog and EggNog at low exposures in lanes 3-7 loaded with eluates from IPs in both figure 12 A and B. Thus there is no enrichment in either target protein. Both figures show lanes 3-7 only contain bright bands at 52kDa corresponding to the molecular weight of the heavy chain of IgG. While low amounts of target enrichment can sometimes be detected at higher exposure times, the bright signal from IgG saturates the image completely making any proteins present at ~40kDa indistinguishable.

3.3 Biochemical properties of axNanog and EggNog

To further characterise the biochemical properties of axNanog and EggNog I performed a NanoBiT protein complementation assay to compare their relative abilities to participate in protein-protein interactions (PPIs). In this assay, one of two luciferase subunits (LgBiT or SmBiT) is fused to a protein of interest. The holoenzyme is then reconstituted if/when the proteins interact, and a luminescent signal from a functional luciferase therefore serves as an indirect readout for protein complex formation. Experimental evidence suggests that mouse Nanog forms a complex with

SMAD2 in both human and mouse embryonic stem cells (Bertero *et al*, 2015). Initially I tested whether axNanog and EggNog were capable of interacting with either axSMAD1 or axSMAD2. We then compared axNanog/EggNog-SMAD1/2 interactions to mouse Nanog-SMAD2. A multitude of data generated in our lab (unpublished) indicates that human VentX and Nanog share a common ancestor. Unlike Nanog genes, Vent genes have not been found to interact with SMAD signalling, so we tested and human VentX-axSMAD1/2 interactions. Lastly, we tested whether EggNog like axNanog, was able to form a homodimer. Luminescence from protein interactions was compared to expression from a negative control construct. The negative control produces luminescence greater than untransfected cells, and represents the level of luminescence due to a small amount of catalytic activity due to expression of the Lrg-BiT luciferase subunit. Typically protein-protein interactions bring the two luciferase subunits into close proximity to form a functional enzyme that generates a bright, luminescent signal which is on average 10-1000 times brighter than that generated by the negative control. Thus, as a positive threshold for measuring a PPI we used a 10-fold greater luminescence value of the negative control. NanoBiT constructs utilise a weak HSV-TK promoter to reduce non-specific PPI, however whether sufficient expression of proteins occurred within 24 hours post-transfection was unknown. Levels of luminescence were therefore determined after addition of the NanoGlo reagent 24, 48 or 72 hours after transfections with NanoBiT constructs (figures 13A, 13B and 13C), as recommended by the manufacturer.

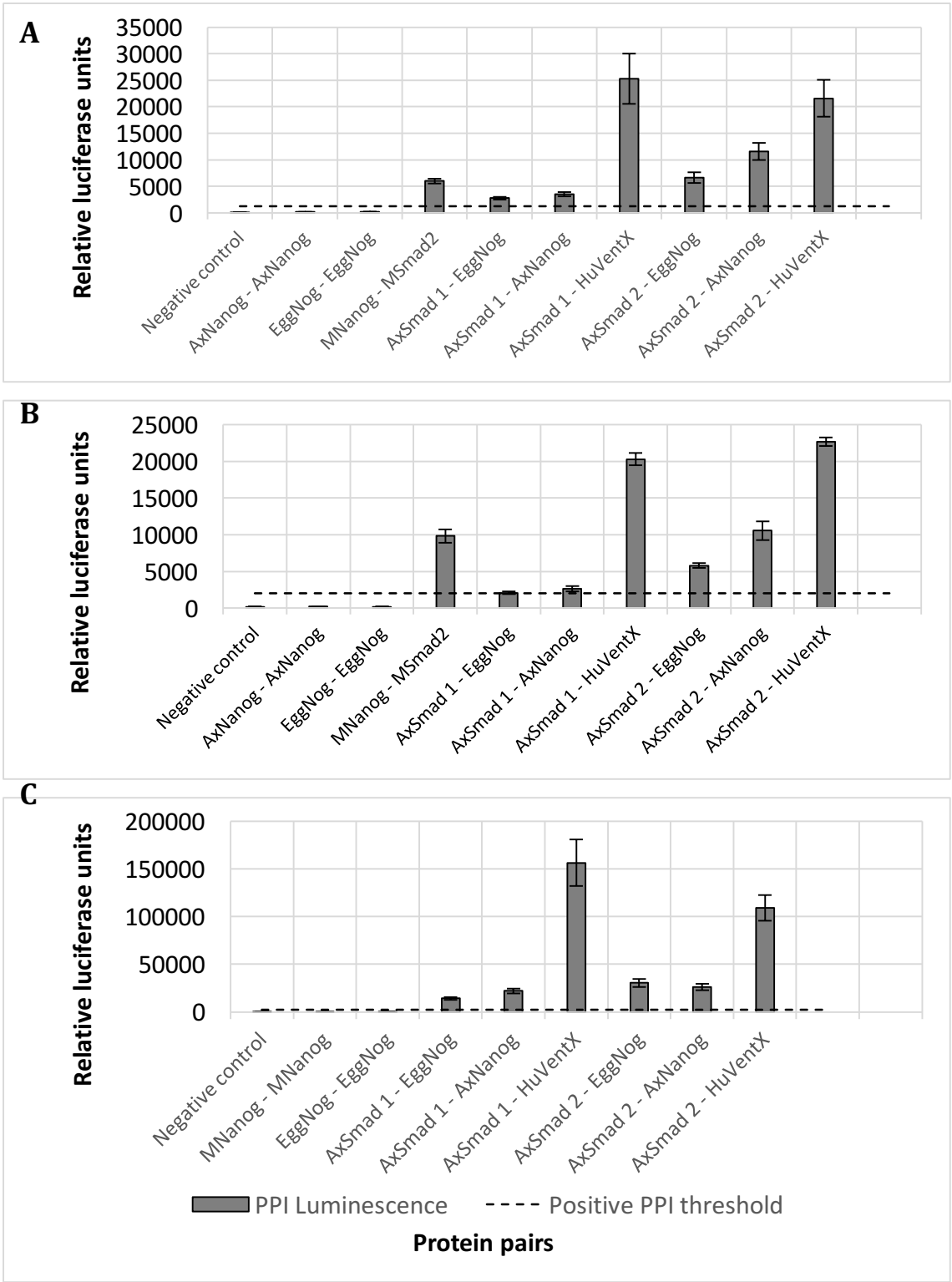


Figure 13. Average luminescence 24, 48 and 72 hours post-transfection. Average luminescence readings from protein-protein interactions across three time points (A) 24 hours, (B) 48 hours, (C) 72 hours post-transfection. Lanes show the luminescence from the following construct combinations; Lane 1, Promega negative control, Lane 2, AxNanog LgBiT-C: AxNanog SmBiT-C, Lane 3, EggNog LgBiT-C: EggNog SmBiT-C, Lane 4, MSMAD2 LgBiT- C: MNanog SmBiT –C, Lane 5, AxSMAD1 LgBiT-C: EggNog SmBiT-C, Lane 6, AxSMAD1 LgBiT-C: AxNanog SmBiT-C, Lane 7, AxSMAD1 LgBiT-C: HuVentX SmBiT-C, Lane 8, AxSMAD2 LgBiT-C: EggNog SmBiT-C, Lane 2, AxSMAD2 LgBiT-C: AxNanog SmBiT-C, Lane 10, AxSMAD2 LgBiT-C: HuVentX SmBiT-C. Dotted line represents ten times the average luminescence of the negative control, the positive PPI threshold.

The levels of luminescence 24 hours post-transfection indicate that lanes 4-10 exceeded the positive PPI threshold value of 1293 Relative luciferase units (RLU). This indicates that the following protein pairs physically interact: mouse Nanog and mouse SMAD2, axSMAD1 and EggNog, axSMAD1 and axNanog, axSMAD1 and human VentX, axSMAD2 and EggNog, axSMAD2 and axNanog, axSMAD2 and human VentX. Lanes 1-3 did not exceed the positive PPI threshold. The negative control construct (lane 1), axNanog-axNanog (lane 2) and EggNog-EggNog (lane 3) did not reconstitute luciferase activity. Measurements taken after 48 (13B) post-transfection closely mirror the results seen at 24 hours. At 72 hours signal increases substantially and does not follow the data trends observed at 24 and 48 hours, likely indicating non-specific interactions due to increasing protein expression. As readings at 24 hours represent the least amount of time in which PPIs are observed, these readings are most likely to represent specific interactions rather than artifacts caused by protein overexpression.

At 24 hours, mouse Nanog-SMAD2/3 produced an average luminescence of 6040 RLU, over 4-fold greater than the threshold for a positive PPI indicating that the two proteins physically associate as expected. Interestingly, axNanog-axSMAD2 produced an average luminescence of 11573 RLU, 9-fold greater than the PPI threshold and twice that of the mice orthologues, suggesting that axNanog-axSMAD2 have a strong binding affinity and form a stable interaction. AxNanog-axSMAD1 also physically interact however luminescence was on average 3536 RLU, 3 times weaker than axNanog-axSMAD2, demonstrating that axNanog-axSMAD1 have a lower affinity for one another and interactions may be more transient. EggNog showed a similar interaction profile, also exhibiting a greater affinity for axSMAD2. The luminescence from EggNog-axSMAD2 was nearly half that generated by axNanog-axSMAD2. Therefore compared to axNanog, EggNog-axSMAD2 interactions are less stable. EggNog-SMAD1 interactions produced an average luminescence of around twice that of the PPI threshold indicating a transient interaction. Unexpectedly, human VentX-axSMAD1 showed the greatest luminescence of all tested conditions with an average luminescence nearly 20-fold above the threshold. VentX-axSMAD2 average

luminescence was also high, being 16-fold greater than the positive PPI threshold. This indicates that VentX has a high affinity for both axolotl SMAD orthologues in which it is able to form stable complexes. AxNanog-axNanog and EggNog-EggNog did not reconstitute luciferase activity and therefore do not form homodimers as expected, since both proteins lack a tryptophan repeat region (reference).

As axNanog appears to bind axSMAD2 with high affinity, and EggNog displayed a weaker binding affinity for axSMAD2, we tested whether axNanog or EggNog differentially affect axSMAD transcriptional activation activity using a GAL4 assay. The GAL4 vector system encodes axSMAD1 or axSMAD2 upstream of an in-frame yeast GAL4 binding domain (GBD) fusion protein. A second vector containing 5 GAL4 binding sites upstream of an in-frame firefly luciferase gene is also present within cell lines. Binding of the GBD to the GAL4 binding site is sufficient to drive transcription of an active luciferase enzyme. We tested SMAD1/2 transcriptional activation in the presence of axNanog, EggNog or human VentX overexpression vectors to determine if they augment or inhibit SMAD mediated transcriptional activation. Cell lines were lysed before the addition of LARII and luminescence reading (Figure 14).

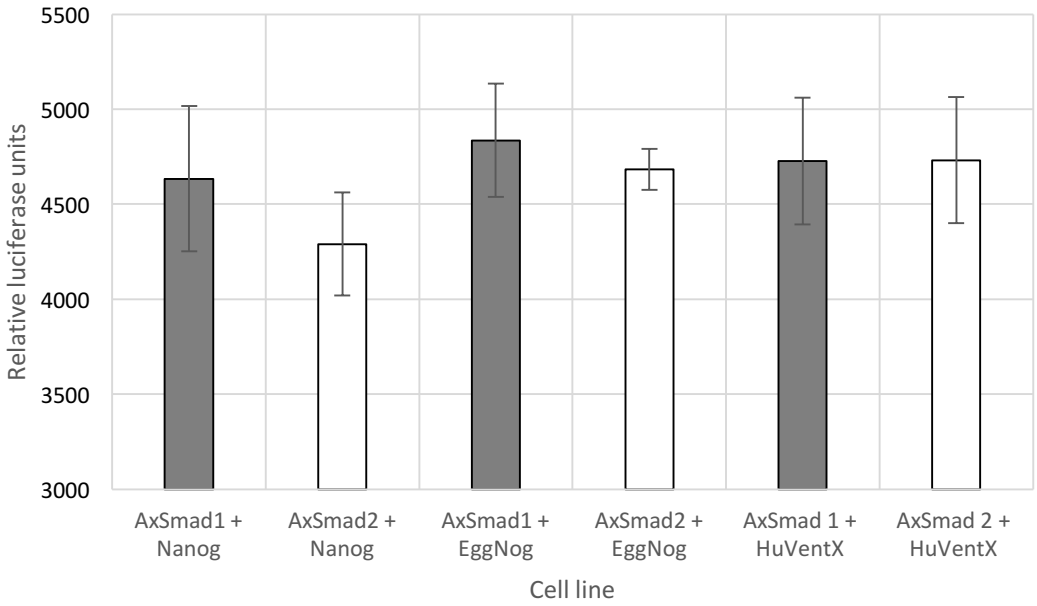


Figure 14. Gal4-SMAD transcriptional activation assay. Gal4-Luciferase transcription is driven by axSMAD1 or 2 fusion protein containing the Gal4 binding domain. Each cell line also constitutively expresses axNanog, EggNog or Human VentX.

Neither axNanog, EggNog nor VentX overexpression significantly altered the transcriptional activity of the axSMAD1/2 -GBD fusions. All conditions show very low luminescence. Poor transcriptional activation of SMAD-GBD fusions has been previously reported (Liu *et al*. 1996, Liu *et al*, 1997). SMAD1-GBD was a weak transcription activator and required strong BMP stimulation for this activation. This is because receptor SMADs require phosphorylation from extracellular signals to enter the nucleus and form transcriptional complexes. Furthermore, SMAD1-GBD proteins were found to have greater transcriptional activity when co-expressed with SMAD4 (Liu *et al*. 1996, 1997; Massagué *et al*, 2005). Thus SMAD proteins may require extracellular signalling and SMAD4 to activate transcription before the effects of axNanog and EggNog overexpression can be detected.

I next investigated the effects of axNanog/axSMAD2 interaction in an in vivo system, axolotl embryos, using a reporter construct for SMAD2 activity that contains a SMAD2 responsive fragment of the *Xenopus* Eomesodermin gene fused upstream of the firefly luciferase coding region . It has been proposed that Nanog acts as a rheostat of SMAD2 transcriptional activity (Mullin *et al*, 2008) so I tested if depletion of Nanog from axolotl embryos affected SMAD2 activity. Axolotl embryos were injected with the luciferase reporter construct. A number of embryos were also incubated in 100uM SB42 inhibitor (to block endogenous Nodal signalling, upstream of SMAD2) while other embryos were also injected with decreasing amounts of antisense morpholino against axNanog, to deplete its levels. Embryos were injected with 2.5pg, 10pg, 40pg or 160pg of morpholino, and uninjected embryos served as a negative control. . Animal caps were dissected from the embryos at early gastrula stage and they were lysed. Lysates were used in luciferase assays as described above (figure 15).

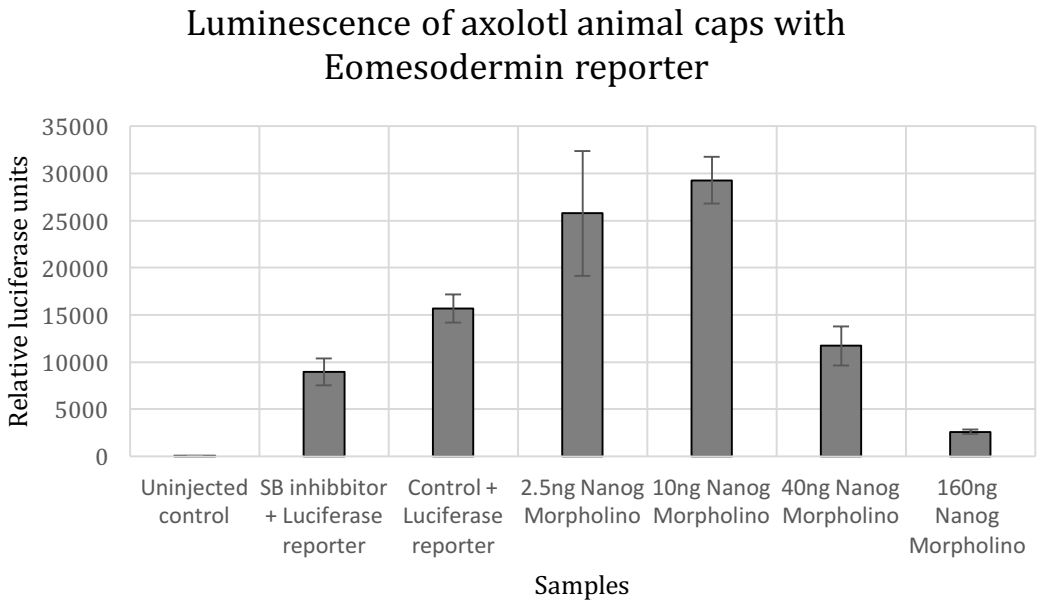


Figure 15. Eomesodermin transcriptional activation by SMAD2 in axolotl animal caps. Luminescence of lysed Axolotl animal caps; Lane 1, lysate from uninjected embryo; Lane 2, embryo incubated in SB42 inhibitor; Lane 3, Eomesodermin reporter alone, Lane 4, 2.5ng Nanog morpholino and Eomesodermin reporter; Lane 5, 10ng Nanog morpholino and Eomesodermin reporter; Lane 6, 40ng Nanog morpholino and Eomesodermin reporter; Lane 7, 160ng Nanog morpholino and Eomesodermin reporter.

Animal caps incubated in SB43 inhibitor generated luminescence on average 37% below that of control animal caps. This was expected as SB43 inhibits TGFβ receptor kinase preventing phosphorylation of axSMAD2, entry to the nucleus and formation of transcription activating complexes. Interestingly, addition of 2.5pg Nanog morpholino increased luminescence increased by an average of 23% above control levels, 10pg Nanog morpholino increased it by 38% and larger amounts had the opposite effect, drastically decreasing luminescence. Animal caps injected with 40pg and 160pg Nanog morpholino produced averages of 34% and 85% below that of control values, respectively. This suggests that injecting up to 10pg of Nanog morpholino proportionally increases SMAD2 transcriptional activation but larger doses ($\geq$40pg) of Nanog morpholino reduces SMAD2 activity. While unexpected, we postulated that higher amounts of the Nanog morpholino might damage the cells in the animal cap. Indeed, effects of the Nanog morpholino compromised embryo viability, as few embryos injected survived to gastrulation. Thus I propose the introduction of a second reporter construct, which constitutively expressed Renilla

luciferase, was introduced to address the potential of cell damage affecting the result. This would provide an internal control that serves as the baseline expression of host cells. Normalising activity of the experimental reporter to the internal control minimised variability caused by differences in sample amount, cell viability, cell lysis efficiency and assay efficiency.

3.4 Reprogramming with amphibian oocyte extracts

Gurdon and colleagues demonstrated that factors within amphibian oocytes could remodel the epigenome of mammalian nuclei. This was evidenced by demethylation of the *Oct4* promoter and increased *Oct4* expression in mammalian nuclei injected into the germinal vesicle of *Xenopus* oocytes (Byrne *et al,* 2003; Simonsson & Gurdon, 2004). Bian *et al* (2009) demonstrated an ex-ovo system utilising AOE induces targeted epigenetic modifications to the chromatin of MEFs. AOE treated cells displayed a reduction in H3K9me3, HP1 and DNA methylation and an increases in H3K9ac to levels comparable to mESCs. Molecules within the axolotl oocyte induced the expression of pluripotency factor Nanog after just 6 hours, an effect not observed in *Xenopus* oocyte reprogramming. Unpublished ChIP qPCR data from our lab has compared the effects of reprogramming using AOE and XOE (Bian *et al*, 2009). Critically these experiments showed AOE but not XOE induced activating marks H3K4me1 at the *Nanog* enhancer region and H3K27ac at both the *Nanog* promoter and enhancer. It appears that AOE and not XOE treatment remodels chromatin around the *Nanog* locus to facilitate *Nanog* expression. When constitutively expressed in a mouse cell line AxNanog has been demonstrated to bind to the native *Nanog* promoter. Moreover, axNanog can bind to the mouse Nanog promoter to drive transcription in a luciferase assay (Dixon *et al*, 2010). We aimed to test whether axNanog binds to Nanog regulatory elements after AOE treatment. We also aimed to see if sub-TAD boundaries around the Nanog and Oct4 loci was altered in response to AOE or XOE treatment. We aimed to test this using ChIP-qPCR.

Reprogramming with oocyte extracts is dependent on adequate permeabilization of the somatic cells with digitonin, which open pores that enable entry of molecules

from the extracts into the cytoplasm of host cells. However, over-permeabilisation can damage the cell nucleus, causing cell death. To test the extent of digitonin permeabilisation, NIH 3T3cells were incubated with two fluorescent molecules: propidium iodide (PI) and fluorescein isothiocyanate-Dextran (FITC-Dextran). PI is a red-fluorescent nuclear and chromosome counterstain that can bind DNA by intercalating with nucleotide bases. It is a small molecule only 668.4Da in size enabling it to freely enter the nucleus although it cannot enter the cytoplasm of unpermeabilised cells. Due to the bright signal produced upon DNA binding PI can also show nuclear fragmentation, a clear sign of cell damage typically detected in apoptotic or necrotic cells. Thus, PI was used as an indicator of cytoplasmic permeabilisation to small molecules and an indicator of nucleus integrity. The much larger FITC-Dextran molecule, 70kDa, can only enter a severely disrupted cytoplasm. Thus, we used FTIC-Dextran staining as an indicator or cytoplasmic permeabilisation (Moumaris *et al*, 2015). This experiment assessed the extent of cytoplasmic and nuclear integrity using FITC-Dextran and PI staining on cells permeabilised for 1 minute, 1 minute 30 seconds or 1 minute and 50 seconds in digitonin (figure 16).

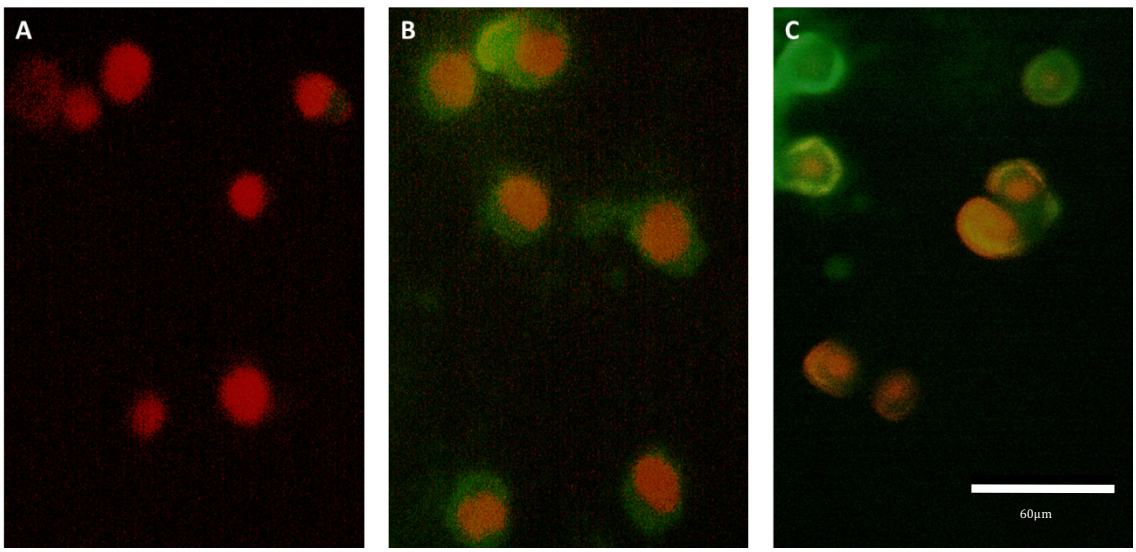
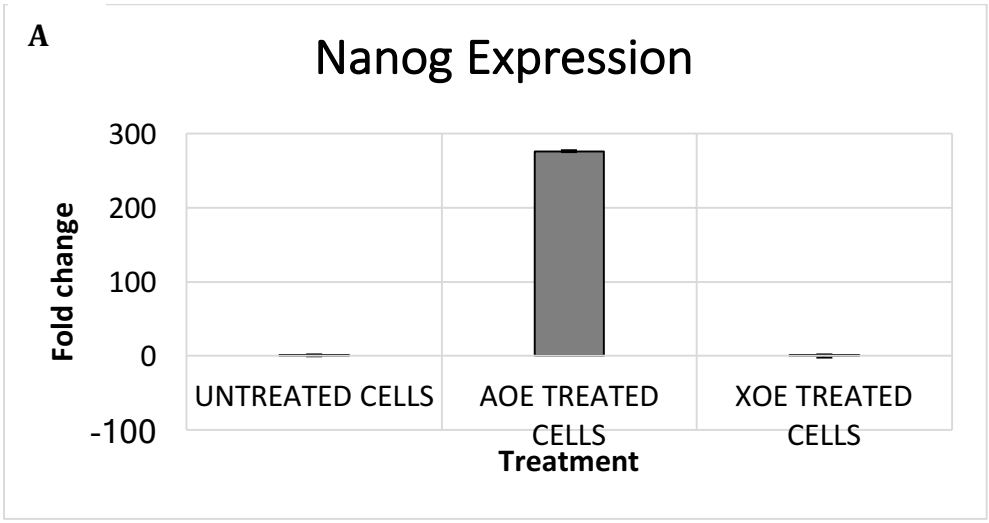


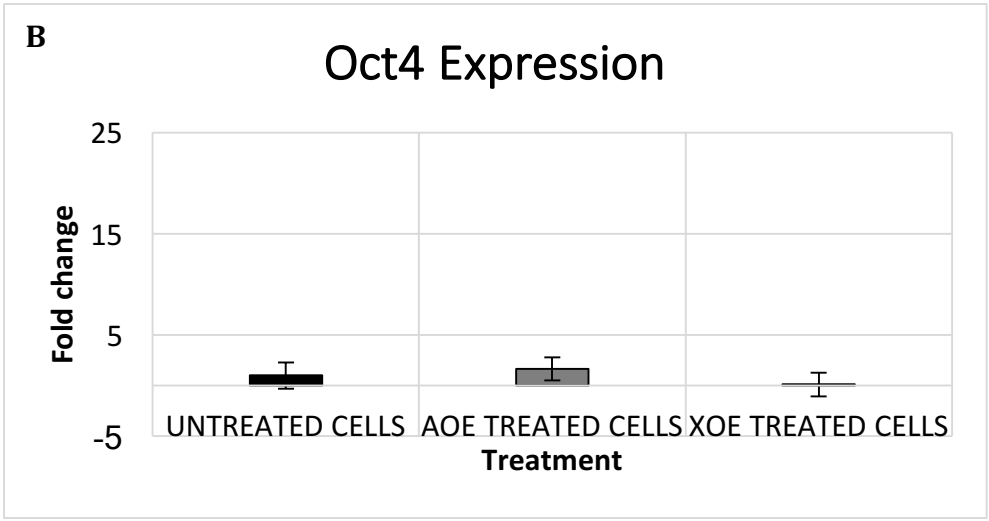
Figure 16. Rate of cell permeabilisation with digitonin. Merged images of cells after incubation with FITC-dextran (visible as green staining) and PI (visible as red staining) after permeabilisation with digitonin for (A) 1 minute (B) 1 minute 30 seconds or (C) 1 minute 50 seconds

Cells permeabilised for only 1 minute allow entry to the small PI molecule that upon entering the cytoplasm is able to traverse the nuclear envelope and bind the DNA within (16A). One minute 30 seconds of permeabilisation (16B) was optimum as cells

displayed green cytoplasmic staining and diffuse red nuclear staining, indicating the cells were being effectively permeabilised without causing excessive damage to the cell nucleus. Cells permeabilised for than 1 minute 50 seconds (16C) displayed uneven red nuclear staining resembling nuclear fragmentation possibly indicating pyknosis, likely due to increased nuclear damage with longer exposure to high digitonin concentrations. Therefore we confirmed that 1 minute 20 seconds was a sufficient period of incubation to effectively permeabilise cells without excessive damage to cell nuclei.

I then investigated whether cells were being remodelled by the oocyte extracts using real-time qPCR (RT-qPCR) to investigate gene expression of treated cells compared to untreated cells. RNA was extracted from untreated cells or cells treated with AOE or XOE and used to synthesize cDNA for qPCR. Bian *et al*, (2009) established that after 6 hours of AOE treatment Nanog transcription is increased significantly above that of untreated cells. They also found that Oct4 expression increases after two weeks in culture, so we compared the expression of Nanog and Oct4 via RT-qPCR (figure 17A and B). We also probed for expression of two other pluripotency markers, Sox2 and Klf4, a downstream target of Nanog, (figure 17C and D).







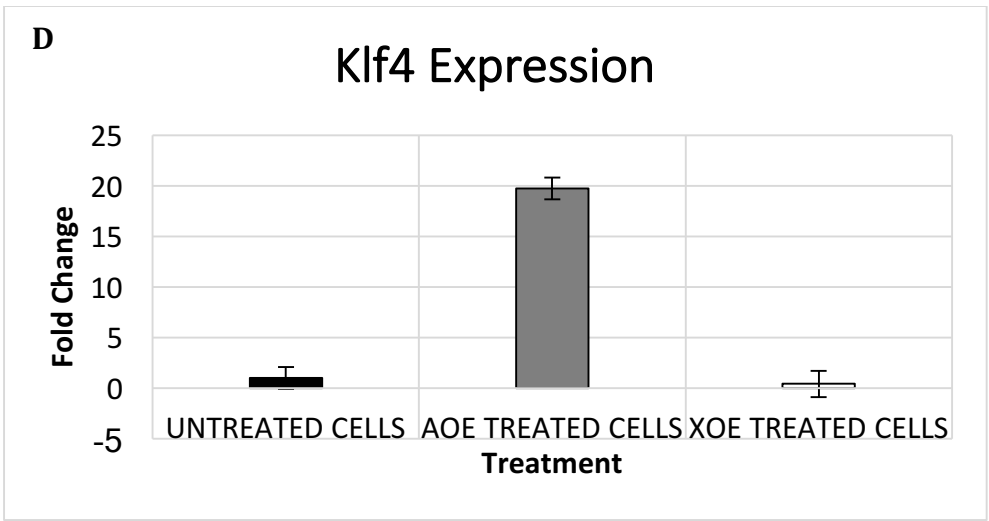


Figure 17. Pluripotency gene expression in oocyte treated cells. Showing fold changes of pluripotency gene expression in AOE (Lane 2) or XOE (Lane 3) treated cells compared to untreated controls (Lane 1). (A) Nanog expression (B) Oct4 expression (C) Sox2 expression (D) Klf4 expression. These should be discussed in numerical order.

The results showed that the nuclei were being reprogrammed to express Nanog, similar to results described by Bian *et al*, (2009). Cells treated with AOE for 6 hours showed a 219-fold increase in Nanog expression compared to the control untreated cells. On the contrary, XOE treated cells show no significant differences in Nanog expression compared to controls. Interestingly, AOE treated cells also displayed 116-fold increase in Sox2 and 20-fold increase in Klf4 expression against untreated cells. XOE treated cells showed no significant differences in the expression of Sox2 or Klf4. Neither AOE nor XOE treated cells showed a significant increase in Oct4 expression after only 6 hours of treatment a result also mirroring that of Bian *et al*. This not only demonstrates the effectiveness of the AOE reprogramming prior to ChIP, it also suggest that AOE is able to activate expression of genes that are components of the mammalian pluripotency network more efficiently than XOE.

Chromatin must be fragmented before ChIP to ensure the majority of the DNA is between 200-800bp, to ensure good resolution and that only DNA near a protein binding site is immunoprecipitated. Chromatin from untreated control NIH 3T3 cells, or 3T3s treated with either AOE or XOE was fragmented by sonication. All samples were sonicated at a 5% duty cycle, intensity level 4 cycles/burst setting of 200 on the S-Series Sonolab program. Samples were sonicated for 5, 10 or 15 minutes before crosslinks were reversed using proteinase K. DNA was extracted and run on an agarose gel to assess the efficiency. The results of sonication optimisation can be seen in figure 18.

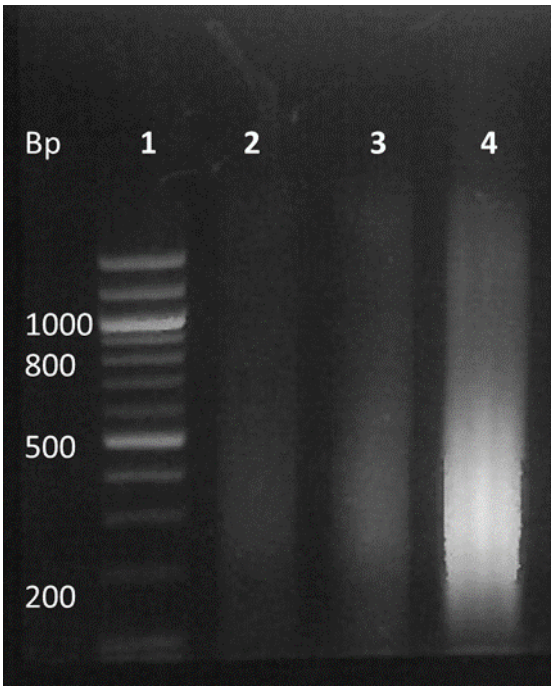


Figure 18. Sonication time optimisation. Extracted DNA from fragmented chromatin, after shearing, was run on a 2% agarose gel. Lane 1, 100bp Amarker; lane 2, DNA from 5 minutes of shearing; lane 3, DNA from 10 minutes of shearing; lane 4, DNA from 15 minutes of shearing.

Lanes 2 and 3 represent chromatin sheared for 5 and 10 minutes respectively. DNA is fragmented to various lengths ranging outside the desired 200-800bp range. Lane 4 shows DNA sheared for 15 minutes is mostly between 200-800bp and is the optimum amount of time for chromatin shearing.

Initial ChIP experiments were intended to investigate if the axNanog protein was enriched at the mouse Nanog promoter and enhancer regions after 6 hours of AOE treatment. PCR primers were designed for the Nanog promoter and enhancer sequences to determine if they were enriched in immunoprecipitations performed with axNanog antibody after AOE treatment compared to controls. Initial axNanog ChIP was performed using 3µg of anti-axNanog and IgG for ChIP (figure 19).

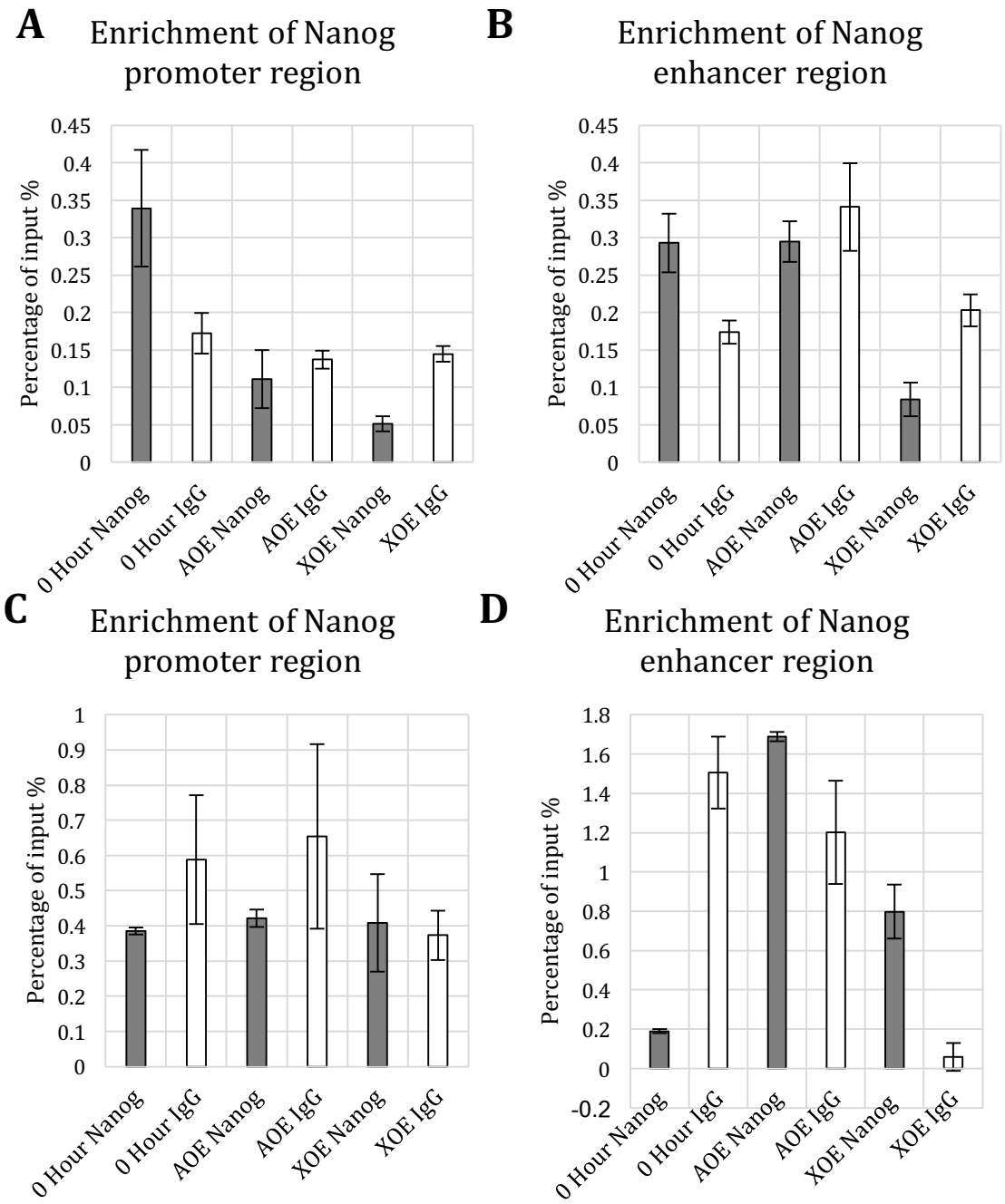


Figure 19. **Anti-Nanog 2 ChIP looking at enrichment of Nanog promoter and enhancer sequences**. QPCR data showing the enrichment of specific loci after ChIP with anti-Nanog 2 or IgG on untreated, AOE or XOE treated cells. (A) Using primers to amplify the Nanog promoter sequence. (B) Using primers to amplify the Nanog enhancer sequence. Antibody concentration was increased to 6ug anti-Nanog or IgG on untreated, AOE or XOE treated cells. (C) Using primers to amplify the Nanog promoter sequence. (D) Using primers to amplify the Nanog enhancer sequence. All sample enrichment is represented as the percentage of the input DNA.

A significant enrichment of Nanog enhancer or promoter sequences is not observed in samples from anti-axNanog2 ChIP over IgG controls. This is consistent in pulldowns using AOE, XOE or untreated cells, suggesting that there is no enrichment of sequences using the anti-Nanog 2 antibody above controls performed with non-specific rabbit IgG. We postulated that greater amounts of antibody might be required to maximise enrichment. Antibody concentrations were increased to 6µg to test whether this would result in greater enrichment (19C and D).

Performing anti-Nanog2 ChIP with an increased antibody titre also did not enrich Nanog promoter or enhancer sequences above the levels of IgG ChIP in any samples. Again this means there is no difference in sequence enrichment using specific anti-Nanog 2 antibody or non-specific rabbit IgG. While a greater range of antibody concentrations could have been tested, the scarcity of AOE available limited the amount of treated cells available for testing. Pull-downs with axNanog 2 appeared to show enrichment of Nanog promoter and enhancer sequences in untreated and XOE treated cells, so we postulated an underlying methodological problem with the ChIP. We then performed ChIP with a well-validated anti-CTCF antibody and positive control anti-acetyl H3 antibody. The aims were two-fold: 1) to use two ChIP-validated antibodies to ensure there were no underlying problems with the existing protocol and 2) to use the anti-CTCF antibody to investigate if sub-TADs around Nanog and Oct4 loci were altered in cells treated with AOE and XOE. Initially we sought to identify CTCF rich regions around the Nanog and Oct4 loci that differed between somatic and pluripotent tissues. CTCF ChIP-Seq datasets from a variety of somatic cell types and mESC was downloaded from the ENCODE project open access database and uploaded to UCSC genome viewer. Each dataset was aligned to the mouse genome (mm10) and visualised. UCSC automatically labels peaks with a significant amount of enrichment above IgG controls. We looked for peaks that differed between pluripotent and somatic tissue types around the Nanog and Oct4 loci. Figure 21 shows the comparison of ChIP-seq datasets and highlights regions of CTCF enrichment with genes that are associated with either a somatic or pluripotent identity.

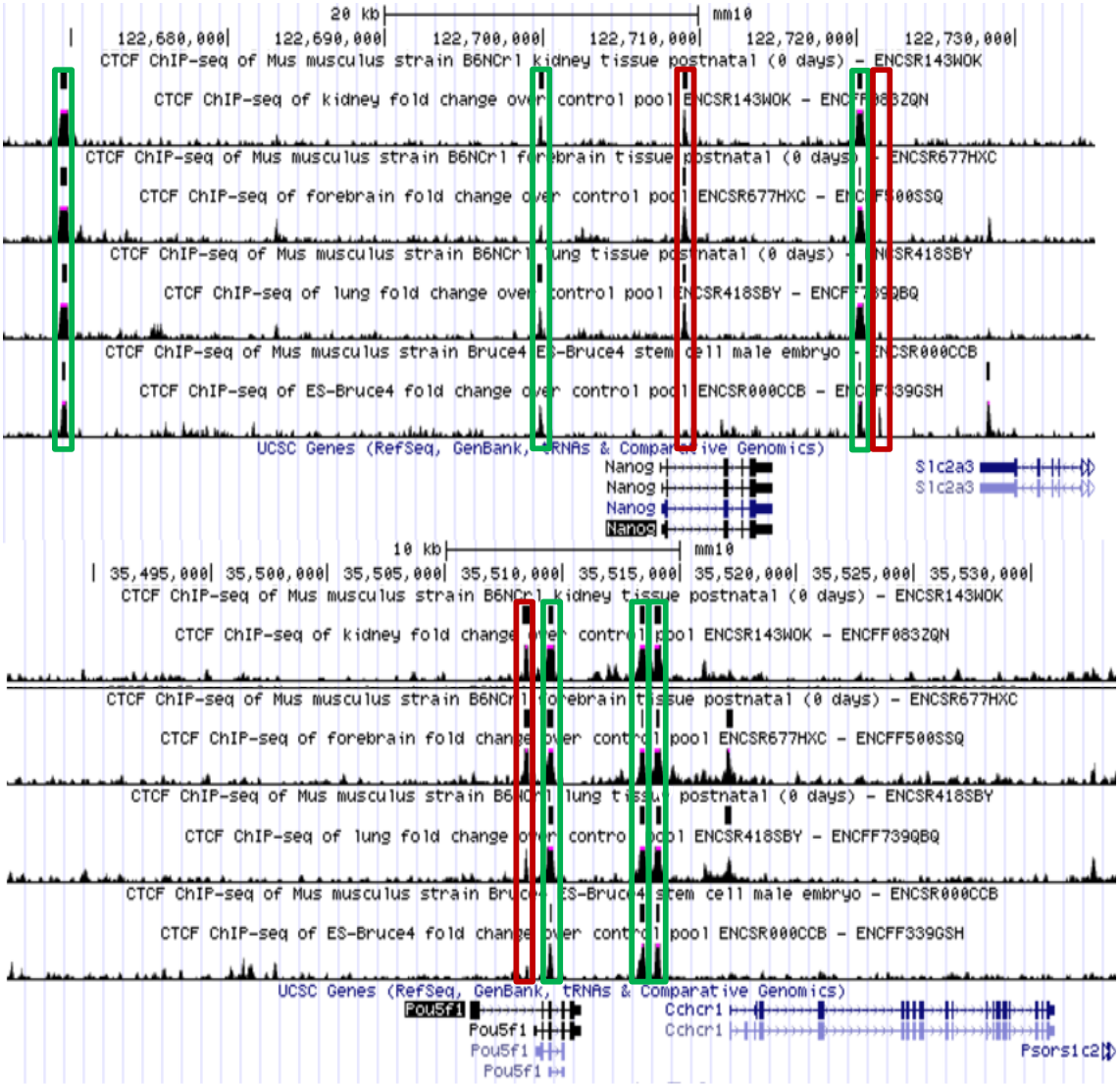


Figure 21. Comparison of CTCF ChIP-seq datasets in pluripotent and somatic tissues around the NANOG and OCT4 loci. CTCF ChIP-seq datasets obtained from kidney, forebrain, lung and ES cells from the ENCODE database aligned to the murine genome assembly mm10 using UCSC genome browser. Highlighted data show the Nanog and Oct4 loci. Red highlights show areas of CTCF enrichment that differ between pluripotent and differentiated tissues. Green highlights show CTCF enrichment that is not cell type specific.

Around these loci there are mESC-specific CTCF rich regions (red in figure 21). It has been shown that the Nanog locus becomes transcriptionally silent in somatic cells. Therefore the CTCF rich regions present in somatic cells are likely to be the boundaries of somatic cell associated, transcriptionally silent sub-TADs. Two CTCF-rich sites were identified, one in the Nanog locus and one in the Oct4 locus which were present in somatic tissues but not in pluripotent tissues. Additionally, one CTCF rich region next to the Nanog locus in pluripotent tissues, but not somatic tissues, was also identified. Several CTCF-rich sites were observed that did not vary between

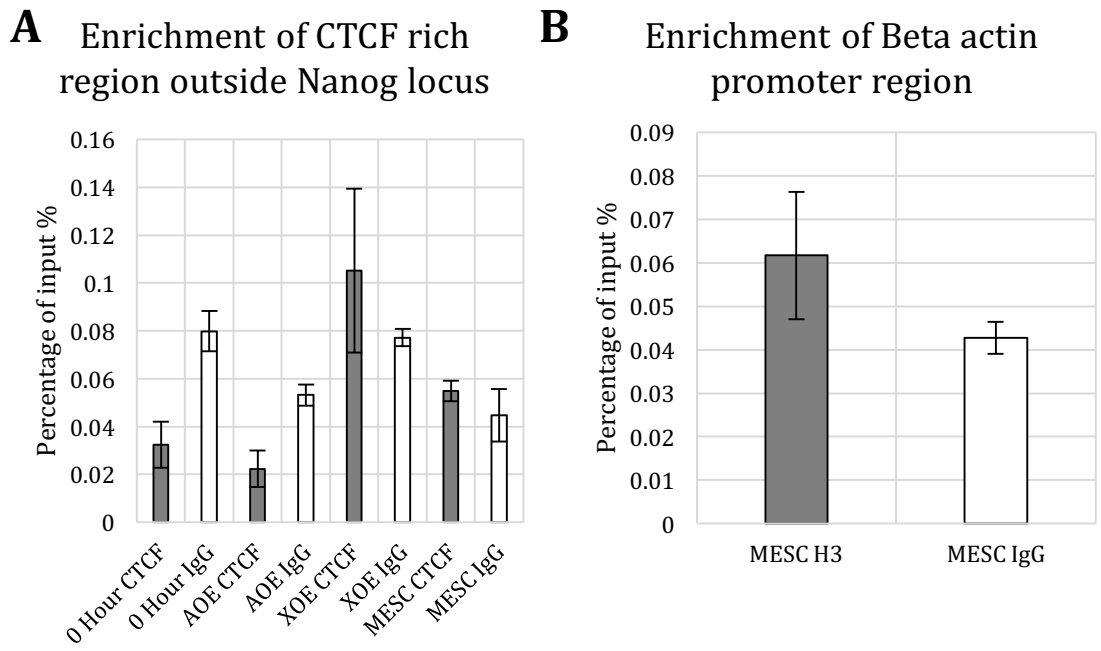


Figure 22. **Showing ChIP using anti-CTCF and anti-acetyl H3 antibodies** (A) qPCR data showing the enrichment of specific loci after ChIP with anti-CTCF or RIgG on mESC, untreated, AOE or XOE treated cells. QPCR performed using primers to amplify a control non cell type specific locus of CTCF enrichment near the Nanog locus. (B) Anti-Acetyl H3 ChIP performed on mESC. QPCR performed using primers to amplify the Beta actin promoter region. All sample enrichment is represented as the percentage of the input DNA.

somatic and pluripotent cell types, which are likely non-cell type specific sub-TADs. Primers were designed for both cell type variable and non-variable regions. The regions that are not cell type specific acted as positive controls for ChIP. I then conducted a control ChIP experiment using anti-CTCF and primers for region N1, a CTCF rich region present in all cell types (figure 21A). Alongside this we performed ChIP with an antibody which recognises the acetylated N terminus of histone 3 (anti-acetyl histone H3) and tested samples for enrichment of the mouse Beta actin promoter region, previously shown to be enriched in acetyl-histone H3 (figure 21B). A specific enrichment of CTCF bound sequences was not observed. Furthermore there was no enrichment in beta actin promoter sequences following immunoprecipitation with anti-acetyl H3. These control ChIP experiments seem to point to an underlying problem in the ChIP. There are several factors that can affect a ChIP experiment. The antibody used, the beads, abundance of target protein and amount of input material are critical factors for maximising enrichment of the target protein. As this experiment used well-validated antibodies in which the target protein is in high abundance, this is not likely to be the problem. Furthermore, both

Millipore kits used have been shown to perform successful ChIP with anti-acetyl H3 antibody with as few as 5×10^5 cells, so lack of chromatin input can be ruled out. Since IgG controls all show enrichment of target sequences it appears that there may be a high degree of background in the immunoprecipitation. Repeating the ChIP experiments is necessary as it requires further optimisation. Including a control antibody and performing a control IP using the magnetic protein A beads (BIORAD) alone could determine if the lack of difference between the test antibody and IgG pulldowns is due to poor quality beads.

4. Discussion

A major aim of this project was to investigate potential differences in the expression or biochemical activities of the axolotl Nanog orthologues, axNanog and EggNog, which might account for their very divergent roles in regulating axolotl development. Transcriptome data (unpublished) identified EggNog and axNanog from maternal RNA transcripts. Thus, initially I tested whether axNanog and EggNog were present as translated proteins in axolotl oocytes using custom antibodies. Both anti-Nanog 2 and anti-EggNog 1 exhibited no cross reactivity and they identified cognate HA-tagged target proteins produced in *Xenopus* oocytes at a working concentration of 0.5µg/ml, rivalling the performance of commercial anti-HA antibodies. Western blots detected the axNanog protein in the oocyte. In contrast, Western blotting for EggNog within AOE yielded a more ambiguous result as immunoreactivity was undetectable at the expected molecular weight, 40kDa, within low exposure times. Faint bands, however, were visible at this approximate position at higher exposures in samples including the uninjected lysate and the AOE, which therefore represent non-specific protein of a similar molecular weight. Nevertheless, this does not fully exclude the possibility that EggNog is present as a translated protein in oocytes, but it may be in low abundance. Because EggNog could not be detected at significant levels within the AOE it suggests that EggNog is translated after fertilisation. This suggests shows that axNanog and EggNog may be differentially regulated after transcription, which may explain in part why the molecules exhibit profoundly different roles in development.

I next tried to discern any biochemical differences between Nanog and EggNog by comparing their associations with other molecules using the NanoBiT complementation assay system. This was chosen because the LgBiT and SmBiT luciferase subunits, which are fused to a protein of interest, have very little affinity for one another. This allowed us to identify even transient interactions and identify differences in binding affinity only 24 hours after co-transfecting constructs. Data published by Bertero *et al*. (2015) demonstrated that in humans Nanog physically interacts with SMAD2/3 and DPY30 to recruit the COMPASS epigenetic remodelling

complex and deposit the H3K4me3 epigenetic mark on key developmentally regulated genes. They suggested that the role of Nanog-SMAD2/3-DPY30 is also conserved in mice as DPY30-deficient embryos displayed impaired gastrulation of the pluripotent epiblast and failed to specify the mesendoderm, diverting to a neuroectodermal fate. For this reason we used mNanog SmBiT-SMAD2/3 LgBiT constructs to directly compare with axNanog-axSMAD2. Axolotl orthologues produced a luminescence nearly double that of their mouse counterparts, suggesting axNanog may bind SMAD2 with higher affinity. These data suggest a physical interaction between axNanog and SMAD2, and unpublished results from morpholino depletion suggest that axNanog has a conserved role in which it physically interacts with axSMAD2, DPY30 and COMPASS, similar to mouse and human. As the axolotl has two Nanog orthologues with distinctly different roles, axNanog may have evolved an increased affinity for SMAD2 after diverging from EggNog. Similarly, EggNog's affinity for axSMAD2 appears significantly reduced, almost half that of axNanog. This suggests that its ability to bind with axSMAD2 has been partially lost. Interestingly, EggNog appears to repress PGC specification, which presumably does not involve Nodal signalling. While axNanog and EggNog differ in affinity for axSMAD2, both showed weak physical association with SMAD1. We did not anticipate that axNanog or EggNog as previous reports suggest that mouse Nanog does not interact with SMAD1 *in* vivo (Bertero *et al,* 2015). On the other hand, Suzuki *et al*, (2008) provide some evidence that SMAD1 and Nanog orthologues in mice physically interact under in vitro conditions, as demonstrated using co-IP performed in NIH3T3 cells transfected with Nanog and SMAD1 expression constructs. They also performed co-IP in mouse ES cells, which suggested the interaction might occur with the endogenous factors. ChIP-Seq experiments to map binding several sequence-specific transcription factors, including SMAD1 and Nanog, showed the two co-localised at the Nanog enhancer region. It was later proposed that SMAD1 interacts with Nanog to act as an activator of the Nanog enhancer in mouse ES cells, while Nanog interacts with SMAD2 to drive Nanog expression in mouse epiblast stem cells (Chen *et al*, 2008; Jang & Ng, 2008). Given that the Nanog protein has been implicated in functions where an interaction with both SMAD1 and 2 is necessary in mice, it seems unsurprising that axNanog is capable of physically interacting with

axSMAD1 in the context of this assay. SMAD proteins exhibit a great deal of homology and interact with a wide range of binding partners due to their regulatory functions in hundreds of genes. Thus the ability for SMAD proteins to interact with both axNanog and EggNog may simply be due to their shared homology.

Data produced in our lab (unpublished) suggests that vertebrate Nanog and Vent genes diverged from an ancient Vent gene family. We initially intended human VentX to act as a negative control as there is to date no evidence to suggest that human VentX physically interacts with either SMAD protein. However, human VentX showed the highest levels of luminescence among conditions tested when paired with axSMAD1 or axSMAD2, producing luminescence 20 and 16-fold that of the positive PPI threshold, respectively. This suggests that VentX is capable of physically binding with both axSMAD1 or 2 with high affinity, which was an unexpected finding. There is evidence that SMAD1 signalling regulates the expression of Vent2 in *Xenopus* but no evidence to date suggests a protein-protein interaction (Hata *et al*, 2000). There is evidence to suggest Nanog and SMAD1 interact to regulate Nanog expression in mice (Chen *et al*, 2008; Jang & Ng; 2008). Given that vertebrate Vent and Nanog may have emerged from a common ancestral Vent, it presents the possibility that both molecules interact with SMAD1 to bind to their own regulatory elements to drive their own expression. While the focus of this study was not human VentX, further investigation could be undertaken to determine if human VentX and Nanog proteins interact *in vivo*.

We also created axNanog SmBiT-axNanog LgBiT and EggNog SmBiT –EggNog LgBiT constructs. Previous experimental evidense demonstrated axNanog's inability to form a homodimer due to the absence of a tryptophan repeat region (Dixon *et al*, 2010). EggNog also lacks a tryptophan repeat. The inability of EggNog to form a homodimer was demonstrated by the inability of EggNog SmBiT-EggNog LgBiT fusion proteins to reconstitute luciferase activity. These protein pairs were therefore used as negative PPI controls in other experiments. This assay provides evidence to suggest that axNanog and EggNog do have similar biochemical properties but differ in their ability to form stable interactions with target proteins which may contribute

to their different roles *in vivo.* To determine the biological relevance of the interactions a goal for future research would be to conduct a co-IP. This could determine whether endogenous axNanog and EggNog interact with SMAD1 and SMAD2 during development.

We next tested whether axNanog or EggNog had an effect on SMAD protein transcriptional activation activity using a GAL4 assay transcription-activation assay. Overexpression of axNanog, EggNog or human VentX did not affect the transcriptional activity when axSMAD1 or 2 proteins were fused with the GAL4 binding domain (GBD). While the MH2 domains of receptor SMADs, such as SMAD1, have demonstrated transcriptional activation potential when tethered to a GBD in published work (Liu *et al*. 1996), a GBD fusion to the full SMAD1 was not a potent transcriptional activator on its own but required strong BMP stimulation for strong transcriptional activation. This is because receptor SMADs require phosphorylation from extracellular signals to enter the nucleus and form transcriptional complexes. Furthermore, SMAD1-GBD proteins were found to have greater transcriptional activity when co-expressed with SMAD4 (Liu *et al*. 1996, 1997; Massagué *et al*, 2005). The most likely explanation for the poor luminescence observed in every condition is that both SMAD1 and 2 require the addition of either BMP or activin, as well as co-expression of SMAD 4, to be effective transcription activators. The alternative explanation is that neither axNanog, EggNog nor human VentX affect the transcriptional activity of SMAD. It has been postulated that Nanog acts as a rheostat of SMAD2 transcriptional activity (Mullin *et al*, 2008) , however it has not been demonstrated empirically and so we performed a SMAD2 transcriptional activation reporter assay *in vivo*. The reporter assay used the promoter region of the SMAD2 target, Eomesodermin which was placed upstream of an in frame firefly luciferase gene similar to that described by Ryan *et al* (2000), providing a luminescent signal upon transcriptional activation by SMAD2. This assay is more biologically relevant than the GAL4 assay as reporter constructs were injected into axolotl animal caps where extracellular signalling and SMAD expression are naturally occurring. Interestingly, luminescence, and therefore transcriptional activity, increased above the levels of control caps upon addition of 2.5pg Nanog morpholino. This further

increased when 10pg was injected. This suggests that depletion of Nanog results in an increase in free SMAD2, which is then able to bind to Eomesodermin regulatory sequences to activate reporter gene transcription. However, the assay also showed that 40pg and 160pg has the opposite effect, drastically decreasing luminescence below control levels. While this was unexpected, we postulated that higher amounts of the Nanog morpholino may damage the developing cells in the animal cap, resulting in reduced expression of luciferase by the reporter construct.

To summarise it would appear that evidence obtained in our lab (unpublished) demonstrates that axNanog and EggNog have drastically different roles in the developing axolotl. EggNog appears to be a suppressor of PGC development and evidence gathered in this study suggests that unlike axNanog, maternally inherited EggNog RNA is not translated at significant levels. It also forms more transient interactions binding with lower affinity to axSMAD2 than axNanog. It is unclear at this point how EggNog maintains its control over PGC specification and whether this involves interactions with SMAD signalling. However, it appears that axNanog and EggNog exhibit different post-transcriptional regulation and biochemical properties which may reflect some of their different functios during development. My results also suggest that axNanog forms stable interactions with SMAD2, the effector of nodal signalling, as part of a conserved role for Nanog in complexing with DPY30 and COMPASS to deposit the H3K4me3 epigenetic mark in pluripotent cells.

The secondary aim of this project was to further investigate axolotl oocyte reprogramming of mammalian cells. Bian *et al*. (2009) established that after 6 hours of incubation in AOE, MEF cells activated transcription of the pluripotency associated gene Nanog. They also published that Oct4 expression would increase after two weeks in culture. I extended these findings here, showing that along with *Nanog*, *Oct4, Sox2* and *Klf4* expression is activated by six hours of incubation in AOE. Under these conditions NIH3T3 cells showed a 219-fold increase in *Nanog* expression when compared to untreated cells. Cells treated with XOE for the same time showed no no detectable differences in *Nanog* expression. We also found cells treated with AOE displayed increases in KLF4 and Sox2 expression, again in contrast to XOE. Finally,

neither AOE nor XOE treated cells showed significant increase in the expression of Oct4 after 6 hours of treatment, supporting the findings of Bian *et al*. Anuran amphibians have lost both Oct4 and Nanog genes since diverging from Urodele amphibians around 250 million years ago. Given that these genes form the 'core' of the conserved pluripotency network in vertebrates, it is perhaps unsurprising that XOE exhibits a less profound effect than AOE on mouse cells. This may also suggest that since diverging from Urodeles, Anuran regulation of Sox2 has changed significantly. It may also be possible, if not likely, that *Xenopus* may have also lost other factors conserved within the vertebrate pluripotency network.

Previous ChIP qPCR data gathered by our lab have suggested that AOE treatment induces multiple epigenetic alterations around the Nanog promoter and enhancer, which are characteristic of transcriptionally active genes. Dixon *et al*, (2010) also demonstrated that axNanog can drive luciferase transcription by interacting with the mammalian *Nanog* promoter. Furthermore they showed that axNanog could bind to native mammalian targets including the *Nanog* promoter when overexpressed. Given this, we hypothesised that axNanog could be driving Nanog transcription through interactions with the mouse *Nanog* promoter in AOE treated NIH3T3's. We used Chip-qPCR with anti-Nanog 2 to investigate if the axNanog protein was enriched at the mouse Nanog promoter and enhancer regions after 6 hours of AOE treatment. We investigated whether these sequences were enriched following axNanog ChIP in AOE treated cells compared to *Xenopus* treated cells and negative controls. While it is not uncommon that antibody concentrations require optimisation during ChIP, both 3µg and 6µg anti-axNanog ChIP showed no significant enrichment above IgG ChIP. As we had previously demonstrated that the AOE treated cells were being effectively permeabilized and reprogrammed, this initially suggested that axNanog is not enriched at these loci.

Because evidence suggests that cells treated with AOE display epigenetic modifications around the Nanog promoter and enhancer, and express Nanog, we sought to investigate whether sub-TADs were altered in cells treated with AOE and XOE. To accomplish this, we attempted to identify whether CTCF enrichment was

altered near the *Nanog* and *Oct4* loci. As it has been shown that *Nanog* and *Oct4* loci become transcriptionally silenced in somatic cells, areas of CTCF-rich regions present in the somatic cells but not mESC are likely to be the boundaries of somatic cell associated, transcriptionally silent, sub-TADs. Indeed our analysis of CTCF ChIP-Seq data confirmed that CTCF enrichment does change around the *Nanog* and *Oct4* loci during the transition from a pluripotent to a somatic cell state. We identified 2 CTCF rich sites that likely appear during the transition from a pluripotent to a somatic cell type, one site within the Nanog locus and one site within the Oct4 locus (seen in figure 21). Interestingly, these sites were in the middle of each gene and while 3D structure cannot be directly inferred from ChIP-Seq data, these sites may represent 'looping-out' a process whereby the looping of DNA physically prevents transcription, resulting in the silencing of a gene. While we successfully identified DNA encoding CTCF test sites which may be altered during AOE reprogramming, anti-CTCF and anti-acetyl histone 3 ChIPs revealed an underlying problem in the ChIP methodology. Several factors may influence the success of ChIP, for this final experiment. The antibodies used are both well validated and have been used successfully with as few as 5×10^6, thus the antibody and chromatin input can be ruled out as possible problems. However, the magnetic protein A beads (BIORAD) may not be suitable for ChIP as most reports have used them only for protein IPs. Therefore the amount of background binding exhibited by these beads may make them unsuitable for highly sensitive ChIP-qPCR.

In summary we have provided evidence which suggests that axNanog and EggNog may be differentially regulated post translation. Furthermore we also elucidated differences in biochemical properties using PPIs. This may in part explain why these Nanog homologs exhibit profoundly different functions in the developing axolotl. Finally we have further defined the expressional differences of cells treated with AOE. Future experiments will primarily focus on improving ChIP methodology to further define AOE reprogramming of mammalian cells.

References

1. Abdel-Rahman, B., Fiddler, M., Rappolee, D. and Pergament, E. (1995). Expression of transcription regulating genes in human preimplantation embryos. *Human Reproduction*, 10(10), pp.2787-2792.

2. Aksoy, I., Jauch, R., Chen, J., Dyla, M., Divakar, U., Bogu, G., Teo, R., Leng Ng, C., Herath, W., Lili, S., Hutchins, A., Robson, P., Kolatkar, P. and Stanton, L. (2013). Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm. *The EMBO Journal*, 32(7), pp.938-953.

3. Alberio, R., Johnson, A., Stick, R. and Campbell, K. (2005). Differential nuclear remodeling of mammalian somatic cells by *Xenopus* laevis oocyte and egg cytoplasm. *Experimental Cell Research*, 307(1), pp.131-141.

4. Beard, C., Li, E. and Jaenisch, R., 1995. Loss of methylation activates Xist in somatic but not in embryonic cells. *Genes & development*, *9*(19), pp.2325-2334.

5. Bendall, S., Stewart, M., Menendez, P., George, D., Vijayaragavan, K., Werbowetski-Ogilvie, T., Ramos-Mejia, V., Rouleau, A., Yang, J., Bossé, M., Lajoie, G. and Bhatia, M. (2007). IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro. *Nature*, 448(7157), pp.1015-1021.

6. Bian, Y., Alberio, R., Allegrucci, C., Campbell, K.H. and Johnson, A.D., 2009. Epigenetic marks in somatic chromatin are remodelled to resemble pluripotent nuclei by amphibian oocyte extracts. *Epigenetics*, *4*(3), pp.194-202.

7. Boyer, L., Lee, T., Cole, M., Johnstone, S., Levine, S., Zucker, J., Guenther, M., Kumar, R., Murray, H., Jenner, R., Gifford, D., Melton, D., Jaenisch, R. and Young, R. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, 122(6), pp.947-956.

8. Byrne, J., Simonsson, S., Western, P. and Gurdon, J. (2003). Nuclei of adult mammalian somatic cells are directly reprogrammed to Oct-4 stem cell gene expression by amphibian oocytes. *Current Biology*, 13(14), pp.1206-1213.

9. Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S. and Smith, A., 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell*, *113*(5), pp.643-655.

10. Dean, W., Santos, F., Stojkovic, M., Zakhartchenko, V., Walter, J., Wolf, E. and Reik, W. (2001). Conservation of methylation reprogramming in mammalian development: Aberrant reprogramming in cloned embryos. *Proceedings of the National Academy of Sciences*, 98(24), pp.13734-13738.

11. Dietrich, J. and Hiiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. *Development*, 134(23), pp.4219-4231.

12. Dixon, J.E., Allegrucci, C., Redwood, C., Kump, K., Bian, Y., Chatfield, J., Chen, Y.H., Sottile, V., Voss, S.R., Alberio, R. and 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*(18), pp.2973-2980.

13. Evans, M. and Kaufman, M. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 292(5819), pp.154-156.

14. Fidalgo, M., Shekar, P., Ang, Y., Fujiwara, Y., Orkin, S. and Wang, J. (2011). Zfp281 functions as a transcriptional repressor for pluripotency of mouse embryonic stem cells. *Stem Cells*, 29(11), pp.1705-1716.

15. Gładych, M., Andrzejewska, A., Oleksiewicz, U. and Estécio, M.R., 2015. Epigenetic mechanisms of induced pluripotency. *Contemporary Oncology*, *19*(1A), p.A30.

16. Guo, J.U., Su, Y., Zhong, C., Ming, G.L. and Song, H., 2011. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell*, *145*(3), pp.423-434.

17. Hackett, J.A., Sengupta, R., Zylicz, J.J., Murakami, K., Lee, C., Down, T.A. and Surani, M.A., 2013. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science*, *339*(6118), pp.448-452.

18. Halder, S., Beauchamp, R. and Datta, P. (2005). A specific inhibitor of TGF-β receptor kinase, SB-431542, as a potent antitumor agent for human cancers. *Neoplasia*, 7(5), pp.509-521.

19. Hansis, C. (2000). Oct-4 expression in inner cell mass and trophectoderm of human blastocysts. *Molecular Human Reproduction*, 6(11), pp.999-1004.

20. Hata, A., Seoane, J., Lagna, G., Montalvo, E., Hemmati-Brivanlou, A. and Massagué, J. (2000). OAZ uses distinct DNA- and protein-binding zinc fingers in separate BMP-SMAD and Olf signaling pathways. *Cell*, 100(2), pp.229-240.

21. Hay, D., Sutherland, L., Clark, J. and Burdon, T. (2004). Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. *Stem Cells*, 22(2), pp.225-235.

22. Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., Zhao, T., Ye, J., Yang, W., Liu, K. and Ge, J., 2013. Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science*, *341*(6146), pp.651-654.

23. James, D. (2005). TGF/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development*, 132(6), pp.1273-1282.

24. Jerabek, S., Merino, F., Schöler, H. and Cojocaru, V. (2014). Oct-4: Dynamic DNA binding pioneers stem cell pluripotency. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1839(3), pp.138-154.

25. Jiang, J. and Ng, H. (2008). TGFβ and SMADs talk to Nanog in human embryonic stem cells. *Cell Stem Cell*, 3(2), pp.127-128.

26. Johnson, A., Crother, B., White, M., Patient, R., Bachvarova, R., Drum, M. and Masi, T. (2003). Regulative germ cell specification in Axolotl embryos: A primitive trait conserved in the mammalian lineage. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358(1436), pp.1371-1379.

27. Johnson, A., Drum, M., Bachvarova, R., Masi, T., White, M. and Crother, B. (2003). Evolution of predetermined germ cells in vertebrate embryos: implications for macroevolution. *Evolution and Development*, 5(4), pp.414-431.

28. Kehler, J., Tolkunova, E., Koschorz, B., Pesce, M., Gentile, L., Boiani, M., Lomelí, H., Nagy, A., McLaughlin, K., Schöler, H. and Tomilin, A. (2004). Oct4 is required for primordial germ cell survival. *EMBO reports*, 5(11), pp.1078-1083.

29. Kim, J., Chu, J., Shen, X., Wang, J. and Orkin, S. (2008). An Extended Transcriptional Network for Pluripotency of Embryonic Stem Cells. *Cell*, 132(6), pp.1049-1061.

30. Kouzarides, T., 2007. Chromatin modifications and their function. *Cell*, *128*(4), pp.693-705.

31. Lavial, F., Acloque, H., Bertocchini, F., MacLeod, D., Boast, S., Bachelard, E., Montillet, G., Thenot, S., Sang, H., Stern, C., Samarut, J. and Pain, B. (2007). The Oct4 homologue PouV and Nanog regulate pluripotency in chicken embryonic stem cells. *Development*, 134(19), pp.3549-3563.

32. Li, E., Beard, C. and Jaenisch, R., 1993. Role for DNA methylation in genomic imprinting. *Nature*, *366*(6453), pp.362-365.

33. Li, E., Bestor, T.H. and Jaenisch, R., 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, *69*(6), pp.915-926.

34. Lister, R., Pelizzola, M., Kida, Y.S., Hawkins, R.D., Nery, J.R., Hon, G., Antosiewicz-Bourget, J., O'malley, R., Castanon, R., Klugman, S. and Downes, M., 2011. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature*, *471*(7336), p.68.

35. Martin, G. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences*, 78(12), pp.7634-7638.

36. Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S. and Meissner, A., 2008. Dissecting direct reprogramming through integrative genomic analysis. *Nature*, *454*(7200), p.49.

37. Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M. and Yamanaka, S. (2003). The Homeoprotein Nanog Is Required for Maintenance of Pluripotency in Mouse Epiblast and ES Cells. *Cell*, 113(5), pp.631-642.

38. Morgan, H.D., Santos, F., Green, K., Dean, W. and Reik, W., 2005. Epigenetic reprogramming in mammals. *Human molecular genetics*, *14*(suppl_1), pp.R47-R58.

39. Mullin, N., Yates, A., Rowe, A., Nijmeijer, B., Colby, D., Barlow, P., Walkinshaw, M. and Chambers, I. (2008). The pluripotency rheostat Nanog functions as a dimer. *Biochemical Journal*, 411(2), pp.227-231.

40. Niwa, H., Ogawa, K., Shimosato, D. and Adachi, K., 2009. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature*, *460*(7251), pp.118-122.

41. Nowak, S. and Corces, V. (2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends in Genetics*, 20(4), pp.214-220.

42. Okano, M., Bell, D.W., Haber, D.A. and Li, E., 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, *99*(3), pp.247-257.

43. Palmieri, S., Peter, W., Hess, H. and Schöler, H. (1994). Oct-4 Transcription Factor Is Differentially Expressed in the Mouse Embryo during Establishment of the First Two Extraembryonic Cell Lineages Involved in Implantation. *Developmental Biology*, 166(1), pp.259-267.

44. Pan, G., Chang, Z., Schöler, H. and Pei, D. (2002). Stem cell pluripotency and transcription factor Oct4. *Cell Research*, 12(5-6), pp.321-329.

45. Pesce, M. and Schöler, H. (2001). Oct-4: Gatekeeper in the Beginnings of Mammalian Development. *Stem Cells*, 19(4), pp.271-278.

46. Reik, W. (2001). Epigenetic Reprogramming in Mammalian Development. *Science*, 293(5532), pp.1089-1093.

47. Ryan, K., Garrett, N., Bourillot, P., Stennard, F. and Gurdon, J. (2000). The *Xenopus* Eomesodermin promoter and its concentration-dependent response to activin. *Mechanisms of Development*, 94(1-2), pp.133-146.

48. Saunders, A., Faiola, F. and Wang, J., 2013. Concise review: pursuing self-renewal and pluripotency with the stem cell factor Nanog. *Stem Cells*, *31*(7), pp.1227-1236.

49. Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. and Cavalli, G. (2007). Genome Regulation by Polycomb and Trithorax Proteins. *Cell*, 128(4), pp.735-745.

50. Schuff, M., Siegel, D., Philipp, M., Bundschu, K., Heymann, N., Donow, C. and Knöchel, W. (2012). Characterization ofDanio rerio Nanogand Functional Comparison to*Xenopus* Vents. *Stem Cells and Development*, 21(8), pp.1225-1238.

51. Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T. and Smith, A. (2008). Promotion of Reprogramming to Ground State Pluripotency by Signal Inhibition. *PLoS Biology*, 6(10), p.e253.

52. Smallwood, S.A. and Kelsey, G., 2012. De novo DNA methylation: a germ cell perspective. *Trends in Genetics*, *28*(1), pp.33-42.

53. Smallwood, S.A., Tomizawa, S.I., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., Sato, S., Hata, K., Andrews, S.R. and Kelsey, G., 2011. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nature genetics*, *43*(8), pp.811-814.

54. Stefanovic, S. and Pucéat, M. (2007). Oct-3/4: Not Just a Gatekeeper of Pluripotency for Embryonic Stem Cell, a Cell Fate Instructor through a Gene Dosage Effect. *Cell Cycle*, 6(1), pp.8-10.

55. Stefanovic, S., Abboud, N., Désilets, S., Nury, D., Cowan, C. and Pucéat, M. (2009). Interplay of Oct4 with Sox2 and Sox17: a molecular switch from stem cell pluripotency to specifying a cardiac fate. *The Journal of Cell Biology*, 186(5), pp.665-673.

56. Suzuki, A., Raya, A., Kawakami, Y., Morita, M., Matsui, T., Nakashima, K., Gage, F., Rodriguez-Esteban, C. and Izpisua Belmonte, J. (2006). Nanog binds to SMAD1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. *Proceedings of the National Academy of Sciences*, 103(27), pp.10294-10299.

57. Tamada, H., Van Thuan, N., Reed, P., Nelson, D., Katoku-Kikyo, N., Wudel, J., Wakayama, T. and Kikyo, N. (2007). Chromatin Decondensation and Nuclear Reprogramming by Nucleoplasmin. *Molecular and Cellular Biology*, 27(18), pp.6580-6580.

58. Tessarz, P. and Kouzarides, T., 2014. Histone core modifications regulating nucleosome structure and dynamics. *Nature reviews Molecular cell biology*, *15*(11), pp.703-708.

59. Vakoc, C.R., Mandat, S.A., Olenchock, B.A. and Blobel, G.A., 2005. Histone H3 lysine 9 methylation and HP1γ are associated with transcription elongation through mammalian chromatin. *Molecular cell*, *19*(3), pp.381-391.

60. Vallier, L. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *Journal of Cell Science*, 118(19), pp.4495-4509

61. Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L., Trotter, M., Cho, C., Martinez, A., Rugg-Gunn, P., Brons, G. and Pedersen, R. (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development*, 136(8), pp.1339-1349.

62. Vallier, L., Reynolds, D. and Pedersen, R. (2004). Nodal inhibits differentiation of human embryonic stem cells along the neuroectodermal default pathway. *Developmental Biology*, 275(2), pp.403-421.

63. Verlinsky, Y., Morozov, G., Verlinsky, O., Koukharenko, V., Rechitsky, S., Goltsman, E., Ivakhnenko, V., Gindilis, V., Strom, C. and Kuliev, A. (1998). Isolation of cDNA libraries from individual human preimplantation embryos. *Molecular Human Reproduction*, 4(6), pp.571-575.

64. Wade, P. and Kikyo, N. (2002). Chromatin remodeling in nuclear cloning. *European Journal of Biochemistry*, 269(9), pp.2284-2287.

65. Wang, J., Levasseur, D. and Orkin, S. (2008). Requirement of Nanog dimerization for stem cell self-renewal and pluripotency. *Proceedings of the National Academy of Sciences*, 105(17), pp.6326-6331.

66. Wang, Z., Teh, C., Chan, C., Chu, C., Rossbach, M., Kunarso, G., Allapitchay, T., Wong, K. and Stanton, L. (2008). The Transcription Factor Zfp281 Controls Embryonic Stem Cell Pluripotency by Direct Activation and Repression of Target Genes. *Stem Cells*, 26(11), pp.2791-2799.

67. Xu, R., Sampsell-Barron, T., Gu, F., Root, S., Peck, R., Pan, G., Yu, J., Antosiewicz-Bourget, J., Tian, S., Stewart, R. and Thomson, J. (2008). NANOG Is a Direct Target of TGFβ/Activin-Mediated SMAD Signaling in Human ESCs. *Cell Stem Cell*, 3(2), pp.196-206.

68. Zeineddine, D., Papadimou, E., Chebli, K., Gineste, M., Liu, J., Grey, C., Thurig, S., Behfar, A., Wallace, V., Skerjanc, I. and Pucéat, M. (2006). Oct-3/4 Dose Dependently Regulates Specification of Embryonic Stem Cells toward a Cardiac Lineage and Early Heart Development. *Developmental Cell*, 11(4), pp.535-546.

69. Ziller, M.J., Müller, F., Liao, J., Zhang, Y., Gu, H., Bock, C., Boyle, P., Epstein, C.B., Bernstein, B.E., Lengauer, T. and Gnirke, A., 2011. Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS genetics*, *7*(12), p.e1002389.

Appendices

Appendix 1 Buffer recipes

Buffers for agarose gel electrophoresis

50x TAE:

For 1 L

• Dissolve 242.2 g Tris base in around 600 mL of ddH2O

• Slowly add 57.1 mL glacial acetic acid

• Add 100 mL 0.5 M EDTA, pH 8

• Bring up the volume to 1 L with ddH2O

6x DNA loading buffer:

For 100 mL

• Weigh 60 g glycerol into a 100-mL graduated cylinder

• Add 12 mL 0.5 M EDTA, pH 8

• Add 10 mg bromophenol blue

• Bring up the volume to 100 mL with ddH2O

10x DNA loading buffer:

For 100 mL

• Measure 20 mL 50x TAE into a 100-mL graduated cylinder

• Add 40 g sucrose

• Add 10 mg bromophenol blue

• Bring up the volume to 100 mL with ddH2O

Buffers for SDS-PAGE

1.5 M Tris, pH 8.8 (stock buffer for separating gels):

For 1 L

• Dissolve 181.65 g Tris base in around 800 mL of ddH2O

• Adjust the pH to 8.8 with concentrated HCl

• Bring up the volume to 1 L with ddH2O

1.5 M Tris, pH 6.8 (stock buffer for stacking gels):

For 1 L

• Dissolve 181.65 g Tris base in around 700 mL of ddH2O

• Adjust the pH to 6.8 with concentrated HCl

• Bring up the volume to 1 L with ddH2O

10x Tris-glycine running buffer:

For 4 L

• 121.1 g Tris base

• 576 g glycine

• 200 mL 20% SDS

• Bring up the volume to 4 L with ddH2O

2x sample loading buffer (reducing):

For 10 mL

• 0.5 mL 1 M Tris, pH 7

• 2.5 mL 20% SDS

• 2 mL glycerol

• 0.2 mg bromophenol blue

• 500 µL β-mercaptoethanol

• Bring up the volume to 100 mL with ddH2O

Buffers for western blotting

10x Transfer buffer:

For 4 L

• 121.1 g Tris base

• 576 g glycine

• Bring up the volume to 4 L with ddH2O

1x Transfer buffer:

For 1 L

• 700 mL cold ddH2O

• 100 mL 10x Transfer buffer

• 200 mL methanol

E. coli growth media

LB:

For 1 L

• 10 g tryptone

• 5 g yeast extract

• 10 g NaCl

• Bring up the volume to around 900 mL with ddH2O, then adjust the pH to 7.4

• Bring up the volume to 1 L with ddH2O

• Sterilize by autoclaving for 30 min

To make LB-agar, add 15 g of agar prior to autoclaving

Low-salt LB:

For 1 L

• 10 g tryptone

• 5 g yeast extract

• 5 g NaCl

• Bring up the volume to around 900 mL with ddH2O, then adjust the pH to 7.4

• Bring up the volume to 1 L with ddH2O

• Sterilize by autoclaving for 30 min

To make low-salt LB-agar, add 15 g of agar prior to autoclaving

SOC:

For 1 L

• 20 g tryptone

• 5 g yeast extract

• 0.5 g NaCl

• 0.186 g KCl

• 0.952 g MgCl2

• Bring up the volume to around 900 mL with ddH2O, then adjust the pH to 7.4 with 10 NaOH

• Bring up the volume to 1 L with ddH2O

• Sterilize by autoclaving for 30 min

• Add 20 mL of sterile 1 M glucose immediately before use